Polynomial chaos expansions for uncertain dynamical systems. Applications in earthquake engineering

Report

Author(s):
Mai, Chu V.

Publication date:
2018-08

Permanent link:
https://doi.org/10.3929/ethz-b-000287120

Originally published in:
IBK Bericht 502

Institut für Baustatik und Konstruktion, ETH Zürich

Polynomial chaos expansions for uncertain dynamical systems. Applications in earthquake engineering

Chu Van MAI

IBK Bericht Nr. 502, August 2018

KEYWORDS: Polynomial chaos expansions, uncertain dynamical systems, earthquake engineering, surrogate modelling, time-warping, nonlinear autoregressive with exogenous input models, fragility curves

Chu Van MAI: Polynomial chaos expansions for uncertain dynamical systems. Applications in earthquake engineering

Bericht IBK Nr. 502, März 2018

Sie finden das Verzeichnis der IBK-Publikationen auf unserer Homepage unter:
The catalogue of IBK publications is available on our homepage at:
http://www.ibk.ethz.ch/publikationen/ibk-reports/alle-berichte.html

Die meisten Berichte von Nr. 270 bis Nr. 333 sind auch noch in gedruckter Form unter Angabe der ISBN-Nr. erhältlich bei:
Most reports from No. 270 to No. 333 can still be purchased in printed form by indicating the ISBN number from:

AVA Verlagsauslieferung AG
Centralweg 16
CH-8910 Affoltern am Albis

Tel. ++41 44 762 42 00
Fax ++41 44 762 42 10
e-mail: avainfo@ava.ch

Berichte ab Nr. 334 sind nur noch in elektronischer Form verfügbar. Sie finden die entsprechenden Dateien in der research-collection der ETH Bibliothek unter https://www.research-collection.ethz.ch oder über die Links auf unserer Homepage.
Reports from No. 334 onwards are only available in electronic form. The respective files can be found in the research-collection of the ETH Library at https://www.research-collection.ethz.ch or through the links on our homepage.

POLYNOMIAL CHAOS EXPANSIONS FOR UNCERTAIN DYNAMICAL SYSTEMS

APPLICATIONS IN EARTHQUAKE ENGINEERING

Chu Van MAI

Institute of Structural Engineering
ETH Zurich

August 2018

Foreword

Computer simulations are nowadays a standard tool for modelling complex engineering systems such as bridges, nuclear power plants or wind turbines. However, simulation results allow the analyst to make decisions only if they provide an accurate description of the real world. Although the physics of the system behaviour is well understood in most cases, meaningful predictions of its future performance are obtained only if its characteristics and properties, as well as its loading and more generally its environmental conditions are perfectly well known. This is rarely the case in practice, unfortunately.

In this context, uncertainty quantification aims at properly modelling the lack of knowledge and/or the variability of parameters describing the system using probability theory. Then uncertainties on the input parameters can be propagated through the computational model of the system, so that certain statistics of the predicted performance indicators can be estimated: mean, standard deviation, probability of exceeding some admissible threshold, etc. Propagating uncertainties through a computational model requires repeated runs of the latter for a sample set of input parameters. Monte Carlo simulation usually require a large number of such simulations, which is not tractable in the case of costly, high-fidelity models. In this respect surrogate models such as polynomial chaos expansions have been developed in the last decades to make such analyses feasible.

This research report was written as a PhD thesis (ETH Dissertation Nr. 23822) by Chu Van MAI and presents innovative methods to build surrogate models of realistic dynamical systems found in civil, mechanical and chemical engineering. The new techniques have been applied in the context of earthquake engineering, where the estimation of the so-called fragility curves of structures submitted to seismic ground motion is of interest. I would like to thank Chu Van MAI for his high commitment all along this research project, and the quality of his contributions to the development of performance-based engineering.

Zurich, August 2018
Bruno Sudret

Abstract

In modern science and engineering, the consideration of uncertainties has become indispensable. The framework of uncertainty quantification, which aims at identifying, quantifying and hierarchizing the sources of uncertainties and studying their effects on the outputs of computational models, has been developed in the last decades. To this end, polynomial chaos expansions (PCEs) represent a powerful and versatile tool which is commonly used in various research fields. The objective of this PhD work is to develop PCE methods that can be applied to dynamical systems with uncertain parameters and/or uncertain excitation.

Chapter 2 presents the general mathematical foundation of generalized PCEs and all the aspects associated with their practical computation. An original analytical formulation of derivatives of PCEs, which allows a straightforward computation of sensitivity measures, is introduced.

In Chapter 3, a literature review on PCE methods for uncertain dynamical systems is thoroughly presented. It opens discussions on why pure vanilla PCEs fail to represent the uncertain behaviour of dynamical systems and how to overcome this issue. Successful existing methods are examined, which reveals their common strategy. Nonetheless, most of those methods are intrusive by construction, meaning that they are developed to solve specific uncertain evolution equations. The findings constitute the guidance upon which two non-intrusive, general-purpose methods are proposed in the remaining of the manuscript.

Chapter 4 introduces a PC-based *stochastic time-warping* method to solve problems of random oscillations. The idea is to capture the dynamics characterized by the vibration frequency with the stochastic time-warping process before applying PCEs to represent the effects of uncertainties on the random amplitudes.

In Chapter 5, a more general method is investigated to solve problems of mechanical systems subject to stochastic excitations. The dynamics is handled with a stochastic nonlinear autoregressive with exogenous input (NARX) model, whose stochastic parameters are modelled with PCEs. The use of a sparsity-promoting regression technique is considered for selecting appropriate NARX terms and polynomial chaos functions.

Finally, Chapter 6 features applications of PC-based surrogate models in the context of earthquake engineering. Predictions of the transient structural responses obtained with the proposed surrogates are used to compute seismic fragility curves. Original non-parametric methods for computing these curves are introduced, which allows one to assess the accuracy of the commonly used parametric methods based on the lognormal format.

The manuscript focuses on applications of PCEs in structural dynamics. However, the developed methods can be easily extended and used in various contexts as some numerical case studies from chemical engineering will illustrate. More importantly, the strategy utilized in the manuscript appears to be a promising research path which differs significantly from existing approaches and shall attract more attention from the uncertainty quantification community.

Keywords: Polynomial chaos expansions, uncertain dynamical systems, earthquake engineering, surrogate modelling, time-warping, nonlinear autoregressive with exogenous input models, fragility curves.

Résumé

De nos jours, il est indispensable de considérer les effets des incertitudes sur les prédictions des modèles numériques. La quantification des incertitudes, qui a pour but d'identifier, de quantifier et de hiérarchiser les sources d'incertitudes ainsi que d'étudier leurs effets sur les sorties des modèles numériques, a été constamment développée durant les deux dernières décennies. A ce titre, la décomposition en polynômes de chaos (PCs) représente un outil puissant et générique utilisé dans de nombreux domaines. L'objectif de la présente thèse est de développer des méthodes basées sur les PCs pour les systèmes dynamiques ayant des paramètres incertains et/ou soumis à des excitations incertaines.

Le chapitre 2 présente les fondations mathématiques des PCs ainsi que tous les aspects associés à leur implémentation d'un point de vue pratique. Une formulation analytique originale des dérivées de PCs, qui permet une estimation directe des indices de sensibilité, est proposée.

Dans le chapitre 3, nous présentons une étude bibliographique des méthodes basées sur les PCs pour résoudre les problèmes de systèmes dynamiques incertains. Nous évoquons les raisons pour lesquelles les PCs classiques n'arrivent pas à représenter le comportement incertain des systèmes dynamiques et analysons les méthodes les plus performantes proposées récemment. Cependant, ces dernières sont intrusives dans la plupart des cas, c'est-à-dire qu'elles sont développées pour un système spécifique. L'étude bibliographique nous permet d'orienter notre recherche vers deux méthodes non-intrusives, qui sont détaillées dans la suite du rapport.

Dans le chapitre 4, nous introduisons une méthode basée sur les PCs et la technique de *time-warping* stochastique pour résoudre les problèmes d'oscillations aléatoires. L'idée est de capturer la dynamique caractérisée par la fréquence de vibration par le processus de *time-warping* avant d'appliquer les PCs pour représenter les effets des incertitudes sur les amplitudes incertaines.

Dans le chapitre 5, une méthode plus générale est étudiée afin de résoudre les problèmes des systèmes mécaniques soumis à des excitations stochastiques. La dynamique est ainsi capturée par un modèle auto-régressif non-linéaire stochastique dont les paramètres incertains sont modélisés par PCs. L'utilisation d'une technique favorisant des structures creuses permet de sélectionner les fonctions auto-régressives et les polynômes appropriés de façon optimale.

Le chapitre 6 présente les applications de modèles de substitution basés sur PCs dans le contexte du génie parasismique. Les prédictions des réponses structurales transitoires obtenues avec les méta-modèles sont utilisées pour calculer les courbes de fragilité. Deux méthodes non-paramétriques sont proposées pour ce calcul. Elles permettent d'évaluer la pertinence des méthodes classiques basées sur l'hypothèse de distribution lognormale.

Le manuscrit aborde principalement des applications en dynamique des structures. Il est néanmoins possible d'appliquer les méthodes développées dans des contextes variés comme le montrent les exemples issus de l'ingénierie chimique. La stratégie utilisée semble prometteuse. Elle est différente des approches classiques et mérite plus d'attention de la communauté scientifique travaillant sur le thème de la quantification des incertitudes.

Mot-clés: Polynômes de chaos, systèmes dynamiques incertains, génie parasismique, méta-modèles, time-warping, modèles auto-régressifs non-linéaires, courbes de fragilité.

Table of contents

List of figures

List of tables

1

Introduction

1.1 Introduction to uncertainty quantification

1.1.1 Computer simulation

In modern science and engineering, computer simulations represent an indispensable tool for research and design processes. The computer models, which are based on rigorous mathematical equations and numerical methods, allow one to reproduce and predict the behaviour of complex physical systems. The use of computer models is unanimous in all contexts with applications including the simulation and forecast of large scale natural phenomena (*e.g.* typhoons, earthquakes, biological and chemical systems), human-systems (*e.g.* in economic and social science) and engineering systems (*e.g.* machines, buildings, aircrafts).

It is worth noting that the processor power is predicted to double every two years according to the so-called Moore's law which still holds true (Moore, 1965; Mack, 2011). For instance, during the 20-year period from 1986 to 2007 the general purpose computing capacity grew at an annual rate of 58% (Hilbert and López, 2011). In particular, the development of super-computers with several thousands of processors has allowed the simulation of extremely large and complex systems, *e.g.* the behaviour of the entire Tokyo city during an earthquake (Ichimura et al., 2014).

Despite the drastic growth of the computing power, one challenge will always remain for practitioners: the more complex the considered system is, the harder it becomes to represent the system accurately with a computer model. First of all, computer models are always based on assumptions simplifying or approximating the real world phenomena, thus resulting in inevitable modelling errors. Second, the parameters used as input of the simulations are hardly known perfectly because of their inherent natural variability or lack of knowledge. In other words, effects of uncertainty on analyses with computer simulations are *certain*. This is a fact that modellers must always be aware of. In this context, uncertainty quantification, which aims at dealing with uncertainties, has become a must in all branches of science and engineering.

1.1.2 Uncertainty quantification

In engineering, uncertainty quantification (UQ) is conducted by a multi-step analysis, which can be graphically summarized by the sketch depicted in Figure 1.1.

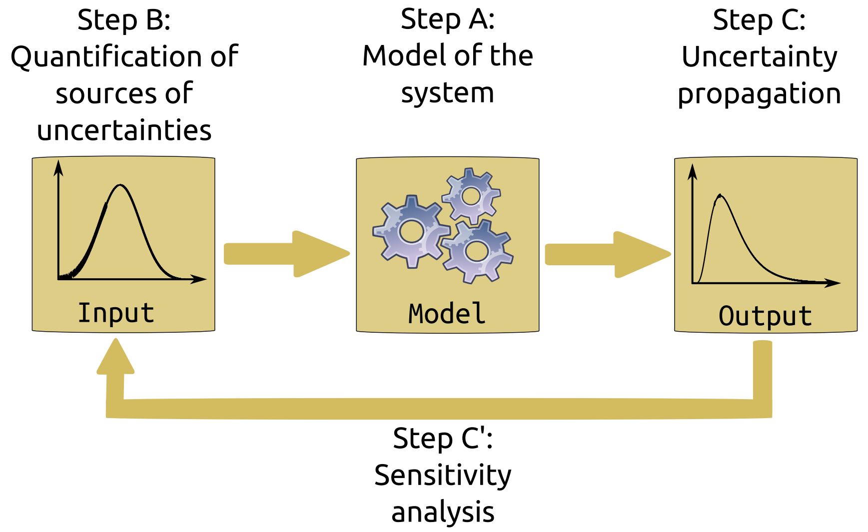


Figure 1.1 – General sketch for uncertainty quantification in engineering (adapted from Sudret (2007))

Step A consists in defining the model of the considered physical system and the associated criteria used to assess the performance of the system. Herein, all the ingredients necessary for a classical *deterministic* analysis are defined, namely the input and output of the computer model and the associated algorithms. The model is commonly defined by the function $\mathcal{M}$: $\boldsymbol{x} \mapsto \boldsymbol{y} = \mathcal{M}(\boldsymbol{x})$ where $\boldsymbol{x}$ gathers the input parameters and $\boldsymbol{y}$ is the model response. In a general case, $\boldsymbol{x}$ includes parameters describing and governing the system, which is problem-dependent whereas $\boldsymbol{y}$ includes the quantities of interest.

Step B consists in quantifying the sources of uncertainties. More precisely, it relies on the identification of input parameters that are not perfectly known, *i.e.* prone to uncertainties and the modelling of these parameters in a probabilistic context by means of statistical methods using data from experiments, legacy data or expert judgement. Step B results in the random vector of input parameters $\boldsymbol{X}$ modelled by a joint probability density function (PDF) $f_{\boldsymbol{X}}$.

Step C consists in propagating the input parameters uncertainties through the model to the output quantities of interest. The objective is to study the probabilistic content of the response. Various methods for propagating uncertainties can be used to obtain different levels of information on the output (Sudret, 2007).

- Second moment methods are used to compute the mean and variance of the response. To this end, one can use Monte Carlo simulation, perturbation, quadrature and weighted integral methods.

- Structural reliability methods (Ditlevsen and Madsen, 1996; Lemaire, 2009) are used to compute the probability of not fulfilling a prescribed safety criterion, which is obtained from the tail of the PDF of the model output. These methods include Monte Carlo simulation, importance sampling, subset simulation, first- and second-order reliability methods.

- Spectral analysis allows one to represent the complete probabilistic content (*i.e.* the entire PDF) of the response by a functional tool, which represents the response as a function of the random input parameters. The spectral stochastic finite element method introduced by Ghanem and Spanos (1991) is such a tool.

Step C' aims at conducting sensitivity analysis on the considered model. To simulate complex engineering problems, a large number of input parameters is commonly required, as specified in Step A. Quantifying the probabilistic model (Step B) for each of these parameters might require important efforts and in reality appears unnecessary. This is due to the fact that usually only a small subset of the parameters play a significant role whereas the variabilities of the remaining parameters do not induce important effects. Therefore it makes sense to rank the parameters in accordance with their relative importances so as to identify those with dominant and negligible effects. Sensitivity analysis techniques such as perturbation-, FORM- and variance-based methods (Sudret, 2007), are used for this purpose.

At this point, it is worth emphasizing that one can distinguish two types of uncertainties, namely aleatory and epistemic (Der Kiureghian and Ditlevsen, 2009). Uncertainties are characterized as aleatory if it is impossible to reduce them. They are related to the inherent randomness of the considered phenomenon. As an example, modelling uncertain parameters with probabilistic distributions using empirical data (Step B) consists in representing the intrinsic aleatory variability of the parameters. In contrast, when modellers foresee a possibility to reduce uncertainties by gathering more knowledge, information and data, they are considered epistemic. For instance, the parameters defining the probabilistic models which are commonly determined using empirical data are prone to epistemic uncertainties which reduce with more observations. From this simple example, one sees that most problems in engineering involve both types of uncertainties. The current manuscript investigate the problem of uncertainty propagation (Step C), in which uncertainties from input parameters are classified as aleatory. The joint effects of aleatory and epistemic uncertainties are not in the scope of the work. However, an investigation on epistemic uncertainties due to limited data (*i.e.* statistical uncertainty) affecting the estimated output quantities of interest (*i.e.* fragility curves) will be conducted in Chapter 6.

1.2 Uncertainty quantification in earthquake engineering

The general UQ framework specified in the previous section can be applied to all engineering problems, *e.g.* design of cars and aircrafts. In this section, the framework will be explained step-by-step in detail in the context of earthquake engineering, which is specifically of interest in civil and nuclear engineering.

1.2.1 Defining model

In Step A, the computer model of the physical system (*i.e.* structure, mechanical component) is defined. It includes the system of equations governing the structural behaviour and the numerical methods used as solvers. The simplest models are single- and multi-degree-of-freedom oscillators, which have been commonly used to describe in a simplified manner the displacements and restoring forces of the structures (Chopra, 1995). More complicated models are obtained with the finite element (FE) method, in which the entire large, complex structure is numerically divided into smaller, simpler parts called finite elements. Each of the finite elements is modelled by simple equations, which are then assembled into a large system of equations describing the original problem. In order to define a FE model, the following input parameters are required:

- Parameters determining the geometry of the system including dimensions of elements, cross-sections, etc.
- Parameters determining the material constitutive law, *e.g.* the Young's modulus, the yield stress and strain, parameters governing the shape of hysteretic loops, etc.
- Parameters determining the loading applied to the system which might be time-independent (*e.g.* dead, live, intermittent loads) or time-dependent (*e.g.* earthquake and wind excitations).
- Parameters determining the boundary conditions, *e.g.* the contact between components, the imposed displacement, the degree of freedom at the boundary nodes.
- "Parameters" (more precisely user-defined options) determining the algorithm used to solve the FE code. For instance, these parameters concern the consideration of the $P - \Delta$ effects, the use of force- or displacement-based elements, the refinement of the FE mesh.

The output quantities of the model include the responses of the components and structures to the earthquake excitation and other types of loadings. The responses commonly searched for are time histories of nodal displacements, reaction forces, stress and strain fields, crack width at a section, etc.

1.2.2 Quantifying sources of uncertainty

After the computer model of the system has been defined in Step A, the quantification of sources of uncertainty (Step B) is now of interest. Earthquake engineering is particularly exposed to uncertainties from different sources as listed in the following:

- Geometry of the system: the mechanical components and structures are built according to a given design. Due to human- or equipment-related imperfections during the manufacturing process, the geometry might not correspond exactly the original design. In some cases, the discrepancies between the design and the real model represent a harmful flaw for the components and structures, see *e.g.* Le Gratiet et al. (2016). The probabilistic models for the geometrical parameters can be obtained by statistical analysis on the manufactured elements or the legacy database.

- Material constitutive law: in reality, the parameters governing the material behaviour differ from the codified properties. Note that the recommended design values are conventionally chosen as prescribed quantiles of the properties distribution functions. The statistical models of material properties can be found in the probabilistic codes (Joint Committee on Structural Safety, 2001b) or statistical studies (Mirza and MacGregor, 1979; Bartlett and Macgregor, 1996). Ghanem and Spanos (1991); Zhang and Ellingwood (1995) modelled the uncertain material properties as random fields by series expansions involving orthogonal functions in the context of stochastic FE method. Recently, Kwon and Elnashai (2006); Liel et al. (2009) considered uncertainties in structural component strength, stiffness, deformation capacity and cyclic deterioration in seismic risk assessment.

- Loading: the loads (either dead, live or intermittent) applied to the structures are always different from the codified values used in the design process. Probabilistic models for these loads can be found in the codes, for instance see Joint Committee on Structural Safety (2001a). In earthquake engineering, special attention has been given to the variability of seismic excitations which represents a paramount source of uncertainty. The randomness of earthquake excitations concerns the magnitude of the event and the associated return period, the direction of the shaking, the spatial variability of the ground motions, the temporal and spectral contents of the motions, etc. The variability of seismic motions has been a long-term topic in earthquake engineering. Probabilistic models have been proposed to model the seismic excitations. In the 1960s, random process models were used to represent earthquake accelerograms (Bogdanoff et al., 1961; Rosenblueth, 1964; Iyengar and Iyengar, 1969). Sophisticated probabilistic models for seismic motions have been continuously introduced and improved, see *e.g.* Conte and Peng (1997); Boore (2003); Rezaeian and Der Kiureghian (2010a); Konakli and Der Kiureghian (2012), among others.

1.2.3 Propagating uncertainty

The third step consists in propagating uncertainties from the uncertain parameters identified and modelled by probabilistic laws so as to study the responses statistical content.

1.2.3.1 Second-moment analysis

The mean and standard deviation of the responses are determined by second moment methods. The most commonly used include Monte Carlo simulation and equivalent linearization.

The Monte Carlo method is based on the law of large numbers, of which the general idea can be summarized as follows. By means of random sampling techniques (Rubinstein, 1981; Owen, 1992), a large sample of the uncertain parameters $\boldsymbol{x}$ is generated in accordance with the joint PDF $f_{\boldsymbol{X}}$. Using the deterministic numerical model, an equally large population of resulting responses $\boldsymbol{y}$ is obtained. This allows the Monte Carlo estimates of the statistical moments of the response quantities. The Monte Carlo estimates have been commonly used as references to validate any newly proposed technique, see *e.g.* Liu et al. (1986); Ghanem and Spanos (1991). Use of MCS for second-moment analysis in structural dynamics and earthquake engineering is universal, see *e.g.* Shinozuka (1972); Papadrakakis and Papadopoulos (1996); Hwang et al. (1998); Duenas-Osorio et al. (2007).

Equivalent linearization (EQL) is a method developed specifically for estimating the first two moments of the responses of a weakly or slightly non-linear system to random excitation (Caughey, 1963). The method aims at replacing a non-linear system with an equivalent linear one such that an approximation error is minimized (Proppe et al., 2003). Multiple linearization criteria might be of interest, *e.g.* the equality between the first two moments of the linearized and non-linearized responses (Proppe et al., 2003), criterion based on the dissipative energy (Elishakoff and Zhang, 1992), tail-equivalent probability (Fujimura and Der Kiureghian, 2007; Der Kiureghian and Fujimura, 2009), etc.

1.2.3.2 Reliability analysis

Reliability analysis in the context of earthquake engineering aims at computing the probability that the considered component or structure fails to fulfil a prescribed safety criterion, which is commonly specified by a limit value of structural responses, namely displacement, deformation or stress. To clearly define the failure of the system, a limit state function, which depends on the input parameters $\boldsymbol{x}$, is specified. In the space of input parameters, the limit state function $\boldsymbol{x} \mapsto g(\boldsymbol{x})$ conventionally defines two domains, namely the safe domain $\mathcal{D}_s = \{\boldsymbol{x} : g(\boldsymbol{x}) > 0\}$ and the failure domain $\mathcal{D}_f = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq 0\}$. Given a set of parameters $\boldsymbol{x}^*$, the system is considered in the failure mode (resp. safe mode) if $\boldsymbol{x}^* \in \mathcal{D}_f$ (resp. $\boldsymbol{x}^* \in \mathcal{D}_s$). The common reliability methods include Monte Carlo simulation (MCS) and its variants (subset simulation, importance sampling), first- and second-order reliability methods.

MCS is the universal tool to conduct reliability analyses in earthquake engineering. Rosenblueth (1964) performed a probabilistic earthquake-resistant design in which probabilistic distributions of parameters governing the earthquake excitations and the random characteristics of the structures were taken into account. MCS was used to compute the probability of survival (*i.e.* non-failure) of the structures, based on which the adequacy of the proposed designs could be assessed. With the increasing computing power, use of MCS has become more and more popular for complex structures, see *e.g.* Karamlou and Bocchini (2015); Mai et al. (2016a). However, the low convergence rate of MCS hinders its application to the computation of small failure probabilities, which is of particular interest in nuclear engineering when low-probability events might lead to considerable damages and consequences. Therefore, variance-reduction techniques, *e.g.* importance sampling (Harbitz, 1983; Melchers, 1989) and subset simulation (Au and Beck, 2001, 2003) have been proposed to accelerate the convergence, thus improving the efficiency of MCS.

The First- and Second-Order Reliability Method (FORM and SORM) were introduced to compute the approximate probability of failure at a low cost compared to MCS. FORM (resp. SORM) relies on first transforming the random variables $\boldsymbol{X}$ into a set of standard normal variables $\boldsymbol{U}$, then approximating the failure surface in the space of $\boldsymbol{U}$ in the vicinity of the most likely failure point with a tangent hyperplane (resp. parabola) (Hurtado and Barbat, 1998). In principle, FORM and SORM are designed for time-invariant reliability problems, which is due to the fact that the time variable is not present in the procedure. However these methods have also been used for stochastic dynamic problems in earthquake engineering, provided that the excitations are discretized and represented in terms of a finite number of random variables (Der Kiureghian and Fujimura, 2009). Li and Der Kiureghian (1995) first used FORM for assessing the safety of non-linear structures subject to random excitations. Subsequent publications in this topic include the works by Zhang and Der Kiureghian (1997); Der Kiureghian (2000); Franchin (2004); Barbato and Conte (2006).

1.2.3.3 Spectral methods

MCS represents a pointwise characterization of the random response, *i.e.* the value of the response given a set of input parameters $\boldsymbol{x}$ is computed locally. Thus a sample with a sufficiently fine resolution in the parameters domain is required to capture the statistical properties of the response. The pointwise characterization can bee seen as a discretized form of a global continuous function relating the response $\boldsymbol{y}$ to the random input parameters $\boldsymbol{x}$. In contrary, spectral UQ methods aim at reconstructing a functional dependence of the response on $\boldsymbol{x}$. The functional form is typically written as follows:

$$\boldsymbol{y} \overset{\text{def}}{=} \mathcal{M}(\boldsymbol{x}) = \sum_{i \in \mathbb{N}} \boldsymbol{y}_i \, \phi_i(\boldsymbol{x}), \tag{1.1}$$

in which $\phi_i(\boldsymbol{x})$'s constitute a suitable basis of the space of second-order random variables $\boldsymbol{x}$ and $\boldsymbol{y}_i$'s are the coordinates of $\boldsymbol{y}$ on this basis. The spectral representation in Eq. (1.1) is exploited to obtain the statistical properties of $\boldsymbol{y}$, either analytically or numerically by random sampling.

Ghanem and Spanos (1991) pioneered the use of spectral representations in the context of the stochastic finite element method. The original work investigated the stochastic response of a structure with spatial random field of Young's modulus of the material. Use of Hermite polynomial chaos functions was introduced to represent the random response. Since then, this method has been generalized, popularized and applied in stochastic structural mechanics and several other fields (Stefanou, 2009). In earthquake engineering, recent applications of polynomial chaos expansions include the work by Ghanem et al. (2005); Sudret and Mai (2013a); Saha et al. (2013); Spiridonakos and Chatzi (2015a,b).

1.2.3.4 Sensitivity analysis

In the context of earthquake engineering, sensitivity analysis is conducted to identify the uncertain parameters with significant (resp. negligible) effects on structural responses and on the system's reliability.

Wall and Bucher (1987) investigated the influences of parameters uncertainties (namely soil parameters in the Kanai-Tajimi earthquake model (Kanai, 1957) and oscillator's parameters) on the expected exceedance rate of a linear oscillator. The results indicated that ground and system uncertainties cause large variations of the mean exceedance rate, which must not be neglected in reliability analysis. Jensen and Iwan (1992) considered the sensitivity of the random responses of a five-degree-of-freedom primary-secondary system with respect to the stochastic stiffness. By sensitivity analysis, Der Kiureghian (1996) found out that for the considered primary subsystem, the damping uncertainty is the most important whereas the mass and stiffness are the most relevant for the secondary subsystem. Kahan et al. (1996) analyzed the sensitivity of the structural responses with respect to small spatial variability of seismic motions. Based on the sensitivity analysis, Huh and Haldar (2001) could consider two parameters to be set at their mean values, namely the plastic section modulus of the beams and columns, which allowed them to reduce the complexity of the reliability analysis. Porter et al. (2002) examined the effects of various factors namely spectral acceleration, ground-motion details, mass, damping, structural force-deformation behaviour and building-component fragility on the repair cost of a building in future earthquakes.

1.3 Probabilistic seismic risk assessment

In earthquake engineering, it is of utmost importance to study the statistics of structural responses (*e.g.* time-dependent displacements, stresses), to quantify the probability of failure and conduct a sensitivity analysis on uncertain parameters. However, it has been shown that the mentioned information is not sufficient for properly assessing the performance of structures, in particular those of major importance such as nuclear power plants, high-rise buildings or other structures with important social impacts. There has been a need for a probabilistic framework to assess the seismic risk of structures in socio-economic terms, *e.g.* number of fatalities and injuries, downtime and consequent financial costs.

To this end, the Pacific Earthquake Engineering Center (PEER) has been developing a probabilistic performance-based earthquake engineering (PBEE) framework (Cornell and Krawinkler, 2000; Günay and Mosalam, 2013). In this multi-element framework, stochastic dynamics play an important role. The modeller/engineer conducts non-linear transient analyses of the structures, in some cases with stochastic structural properties, subject to random seismic excitations selected from a database of recorded motions or generated synthetically with probabilistic models. The seismic intensity measures and resulting random responses are retrieved to build a probabilistic seismic demand model, which consequently specifies the probability conditioned on an intensity measure that the structural response exceeds a critical value. The outcome is commonly represented graphically as seismic fragility curves.

Due to the high computational cost of the non-linear transient analyses, usually only a limited number of simulations is performed. In addition, use of simplifying assumptions for analytical and numerical conveniences is a widely accepted practice (*e.g.* set the shape of the fragility curves to a lognormal cumulative distribution function). However, the current development of advanced computational approaches, *e.g.* the spectral stochastic finite element method, might allow one to validate such assumptions and practice. This will be done with the use of spectral methods for surrogating the original structural model, thus obtaining reliable estimates at acceptable computational expenses.

1.4 Surrogate modelling

Use of Monte Carlo simulation (MCS) for propagating uncertainties (*e.g.* for second moment, reliability and sensitivity analyses) is universal in engineering. It is due to the fact that MCS is theoretically simple, computationally straightforward, generally applicable and well-suited for parallel processing (Schuëller, 2001). The MCS procedure does not depend on the complexity of the considered problem, *e.g.* the number of random parameters, the complexity of non-linear behaviours. However, the low convergence rate of MCS estimators represents a major issue that needs to be solved. For instance, to evaluate a probability of failure of magnitude 10^{-k} with relative accuracy of 5%, the number of model simulations required by MCS is $4 \times 10^{k+2}$. In earthquake engineering, each model simulation (*i.e.* non-linear transient analysis of a structure under earthquake shakings) is computationally expensive. Despite current considerable computing power, conducting a proper MCS on daily design leads to prohibitive costs for engineers. There are several techniques proposed to substitute MCS, *e.g.* equivalent linearization and perturbation method for second moment analysis, FORM and SORM for reliability analysis. In particular, MCS represents a collocation-type method which characterizes the random response point-wise. It is appealing to consider a method that allows a global representation of the response and can be used for several purposes (*i.e.* second moment, reliability and sensitivity analyses) at low computational expenses.

In this context, surrogate models, also known as response surface models, metamodels or emulators, appear as a promising alternative to MCS (Forrester et al., 2008). In a nutshell, surrogate modelling aims at constructing an approximate model that mimics the behaviour of the considered computer simulation while being computationally cheap to evaluate. The process of surrogate modelling consists of three stages (Simpson et al., 2001; Forrester et al., 2008):

- Stage 1: One first prepares the data, a.k.a. the experimental design, used for fitting the surrogate model. The data are collected from numerical simulations or real measurements. They consist of a set of realizations of random parameters which might be obtained by different sampling schemes and the associated response values. Next, one chooses a surrogate modelling approach to investigate. Forrester et al.

(2008) presented a wide variety of options, namely polynomial functions, radial basis function, Kriging, support vector regression, artificial neural networks. However, we will not go through all the mentioned approaches in this manuscript. We focus only on the polynomial chaos expansion metamodelling technique, which was introduced in the context of the spectral stochastic finite element method (Ghanem and Spanos, 1991).

- Stage 2: Using the prepared data, one determines the parameters defining the surrogate model. This is the training process, in which numerical algorithms are used to identify parameters that lead to a sufficiently accurate surrogate model in the sense that it satisfies specified criteria, *e.g.* minimizing an error estimator.

- Stage 3: The accuracy of the identified model is validated on an independent validation set which is different from the data used for training. The surrogate model is judged in terms of predicting the random responses. Only after the validation can the surrogate model be used for other purposes, *e.g.* second moment analysis, sensitivity analysis, design optimization. Note that one may also rely on cross-validation techniques in cases when it is not possible to have an independent validation set.

In the context of stochastic mechanics, Wong (1985) used the response surface method to replace a long-running FE code in the reliability analysis of soil slopes with uncertain soil density, elastic moduli, cohesion and angle of friction. Faravelli (1989) represented the structural response (stress) in terms of polynomial functions of spatial averages of design variables (*i.e.* the Young's modulus and the thickness of the component). The response surface model was then used to assess the reliability of a pressure vessel. Ghanem and Spanos (1991) used Hermite polynomial chaos basis functions to represent the stochastic response of a structure with spatial random field of Young's modulus of the material. Since then, spectral methods have proved their effectiveness when being used for surrogate modelling in various fields of science and engineering, see *e.g.* Le Maître and Knio (2010).

In the field of earthquake engineering and probabilistic risk assessment, surrogate modelling has been introduced recently. Huh and Haldar (2001) approximated the limit state functions with polynomial response surfaces then conducted FORM for time-dependent reliability of structures subject to earthquake motions. Iervolino et al. (2004) used second order polynomial expansion to represent the parameters of lognormal fragility curves (*i.e.* median and standard deviation) for ground supported liquid storage tanks with uncertain liquid filling height. From a similar perspective, Towashiraporn et al. (2008) presented the first and second order moment of the structural responses as polynomial response surfaces of uncertain structural and earthquake properties, namely the steel yield strength, the elastic modulus, the damping ratio and the spectral acceleration. The response surface was then used to compute fragility curves. Liel et al. (2009) also described the median collapse capacity of a structure as a function of the model random variables. Sudret and Mai (2013a) used PCEs to replace a linear steel frame subject to random seismic excitation, which was then combined with MCS to compute seismic fragility curves of the structure. Saha et al. (2013) utilized PCEs for statistical analyses of responses of base-isolated liquid storage tanks with uncertain structural parameters subject to random sinusoidal excitation. More recently, Gidaris et al. (2015) approximated the medians of structural responses (*e.g.* peak inter-story drift and peak floor acceleration) as functions of the uncertain structural and ground motions parameters by Kriging. The metamodel was then used to compute seismic fragility curves.

In this work, we are particularly interested in applying polynomial chaos expansions (PCEs) to represent the time-dependent stochastic responses of structures under earthquake excitations. Note that the objective is different from surrogating the first and second order moments of the responses (Towashiraporn et al., 2008; Liel et al., 2009; Gidaris et al., 2015). We aim at building PCEs that are capable of predicting the response time-histories to random values of uncertain parameters. This is a challenging task because plain vanilla PCEs are well known to fail when it comes to approximating stochastic dynamical systems (Wan and Karniadakis, 2006; Le Maître et al., 2010). A simple and straightforward application of PCEs is leading to inaccurate models. It is therefore of utmost importance to figure out the reasons of the failure and find a solution to overcome this challenging issue.

1.5 Goal and outline of the manuscript

Following this introduction, the goal of the PhD thesis is to develop polynomial chaos expansions for stochastic dynamics[1] and apply this technique to earthquake engineering. The manuscript aims at:

- presenting a systematic and thorough review of existing methods based on PCEs in the context of stochastic dynamics;
- investigating novel methods that can help to improve the accuracy and effectiveness of PCEs in stochastic dynamics;
- applying the investigated methods to solve stochastic dynamics problems, in particular realistic earthquake engineering applications;
- validating assumptions that are commonly used for computing seismic fragility curves of structures in seismic risk assessment.

To this end, the work conducted during the PhD is organized in six chapters as follows:

Chapter 2 presents briefly the general mathematical background of PCEs. In addition, it describes the aspects associated with the use of PCEs in practice, namely the handling of dependent random variables, the truncation of polynomial functions, the computation of relevant coefficients using an advanced sparse adaptive scheme, the estimation of errors and the post-processing of PCEs for obtaining statistical content of the responses. In particular, this chapter features a section which introduces a novel method to compute sensitivity indices using PCEs.

---

[1] For the sake of simplicity, the term “stochastic dynamics” is occasionally used in this manuscript to refer to the considered problems of dynamical systems subject to uncertainties.

Chapter 3 consists of a literature review of PCE methods that have been used so far to handle stochastic dynamics problems. Advantages and limitations of each method are pointed out and connections between them are clarified. In particular, the methods are classified in an unprecedented manner which allows to emphasize their characteristics, thus explaining the effectiveness (resp. ineffectiveness) of certain methods. Based on the established big picture of the framework, two novel methods are proposed in Chapters 4 and 5.

Chapter 4 presents a novel approach for surrogating the time-dependent responses of oscillatory systems with random parameters. The approach is based on the concept of stochastic time-warping, which introduces an auxiliary time scale specific to each realization of the uncertain response. The auxiliary time scale is determined so as to maximize the similarity between the random response and a deterministic reference trajectory. PCEs are then used to represent the transformed responses on the new time scale as well as the stochastic time-warping parameters.

Chapter 5 investigates the use of nonlinear autoregressive representations of time series (so-called NARX models) together with PCEs to surrogate the response time-history of systems with random parameters and external excitations. The NARX model which captures the dynamical behaviour of the system is identified by a regressors selection technique which has been commonly used to compute PCEs. Most importantly, the stochastic NARX coefficients are represented with PCEs.

Chapter 6 introduces applications of the developed PC-based surrogate models in probabilistic seismic risk assessment. A large number of simulations is obtained with surrogate models, which allows the use of recently introduced non-parametric methods to compute seismic fragility curves. The assumptions commonly used for computing these curves and the effect of epistemic uncertainties are then discussed.

Chapter 7 finally summarizes the findings and contributions of the work. It also opens discussions on the limitations of the proposed methods as well as new research paths that can follow up.

2

Polynomial chaos expansions

Monte Carlo simulation (MCS) is the classical tool for uncertainty quantification, which is based on repeated simulations of the computational model to estimate the statistics of the output, *e.g.* the first and second order moments (mean, variance), the distribution or the probability of exceeding a prescribed threshold. Use of MCS represents several advantages. It is robust in the sense that no hypothesis or condition on the input, output or computational model is required. In addition, it is applicable to all problems of arbitrary complexity given that the computational model is available. The error of the estimator converges independently of the dimensionality of the problem, therefore MCS is advantageous in case a large number of random variables is involved. However, the convergence rate of the MCS estimator is proportional to $1/\sqrt{N}$ where N is the number of numerical simulations, which is deemed a low rate. This hinders the applicability of MCS in the cases when running a large number of numerical simulations is prohibitive in terms of computational resources. To overcome this issue, several approaches were proposed in order to accelerate the statistical convergence of the MCS-based estimators. For instance, importance sampling and subset simulation are widely used in reliability analysis. However, the proposed approaches are not yet sufficient in practical applications. They are less robust than the traditional MCS, and have to be tailored to the problems under investigation, *e.g.* the selection of an importance sampling density has to be carefully addressed.

In this context, *spectral methods* have proven to be an effective alternative to MCS. The spectral methods are based on the concept of constructing the functional dependence of the output quantity on the input random variables. The computed Fourier-like representation is then used for estimating the statistics analytically (*i.e.* by using only the spectral expansion) or numerically (*i.e.* by sampling the input variables and computing the output with the closed-form spectral expression). Whether using the spectral expansions in an analytical or numerical manner, the required computational cost is reduced significantly compared to applying MCS directly to the original model.

The spectral approach was introduced by Ghanem and Spanos (1991) in the framework of the stochastic finite element method. The random spatial variation of the Young modulus was described as a Gaussian stochastic field by means of the Karhunen-Loève expansion. The resulting stochastic stiffness matrix of the system was obtained with the random Young modulus field. The response of the structure was represented by a series of Hermite polynomial chaos functions. Both expansions were introduced in the deterministic system of equations. The resulting set of *deterministic* equations allows one to determine the coefficients of the polynomial expansion of the response. Later, Xiu and Karniadakis (2002) generalized the spectral approach to the Askey scheme of polynomials, thus offering the possibility to represent the responses with different orthogonal polynomials, for instance Legendre, Laguerre, Jacobi. The approach is termed generalized polynomial chaos expansions (PCEs).

The spectral approach described above is intrusive in the sense that one needs to interfere with the finite element code or in a general case with the original system of equations describing the considered system. In most practical cases, this is not feasible because the numerical models are of “black box” type . Usually, only a set of model simulations is available for the analysis. This situation requires non-intrusive methods, see *e.g.* Choi et al. (2004); Xiu and Hesthaven (2005); Berveiller et al. (2006), which are in the focus of this manuscript. Note that the non-intrusive methods aim at determining the expansion coefficients using techniques such as stochastic collocation, projection and regression *without* modifying the original deterministic code, thus being applicable to daily-life engineering problems.

This chapter is organized as follows: first the non-intrusive generalized polynomial chaos expansion is presented. Then a sparse adaptive scheme for computing PCEs, which is used throughout this manuscript, is described. Next, one reviews the use of PCEs for computing the output statistics. Finally, a contribution to the framework is presented, namely the analytical estimation of derivative-based global sensitivity measures with PCEs which was published in Sudret and Mai (2015).

2.1 Generalized polynomial chaos expansions

2.1.1 Spectral representation

Let us consider the computational model $Y = \mathcal{M}(\boldsymbol{X})$. Herein $\boldsymbol{X} = (X_1, \dots, X_M)$ is a M-dimensional input vector of random variables with given joint probability density function (PDF) $f_{\boldsymbol{X}}$ defined over an underlying probability space $(\varOmega, \mathcal{F}, \mathbb{P})$ and $\mathcal{M} : \boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} \subset \mathbb{R}^M \mapsto \mathbb{R}$ is the computational model of interest, where $\mathcal{D}_{\boldsymbol{X}}$ is the support of the distribution of $\boldsymbol{X}$. Without loss of generality, one considers the case of a scalar valued output and assumes that the input random variables are independent, *i.e.* the joint PDF is the product of the marginal PDFs:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{X_1}(x_1) \dots f_{X_M}(x_M). \tag{2.1}$$

Assuming that the scalar output Y is a second order random variable, *i.e.* $\mathbb{E}\left[Y^2\right] < +\infty$, the computational model $\mathcal{M}$ belongs to the Hilbert space $\mathcal{H}$ of square-integrable functions with respect to the inner product:

$$< u \, , \, v >_{\mathcal{H}} = \int\limits_{\mathcal{D}_{\boldsymbol{X}}} u(\boldsymbol{x}) v(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \tag{2.2}$$

Denote by $\mathcal{H}_i$ the Hilbert space of square-integrable functions with respect to the marginal probability measure $\mathbb{P}_{X_i}(\mathrm{d}x_i) = f_{X_i}(x_i)\mathrm{d}x_i$. $\mathcal{H}_i$ is equipped with an inner product:

$$< u \, , \, v >_{\mathcal{H}_i} = \int\limits_{\mathcal{D}_{X_i}} u(x_i) v(x_i) f_{X_i}(x_i) \mathrm{d}x_i, \tag{2.3}$$

where $\mathcal{D}_{X_i}$ is the support of the distribution of X_i. Let us denote by $\{\phi^i_k, k \in \mathbb{N}\}$ an orthonormal basis of $\mathcal{H}_i$ satisfying:

$$< \phi^i_k \, , \, \phi^i_l >_{\mathcal{H}_i} = \delta_{kl}, \tag{2.4}$$

with δ_{kl} being the Kronecker symbol, which is equal to 1 if $k = l$ and equal to 0 otherwise.

Soize and Ghanem (2004) proved that the Hilbert space $\mathcal{H}$ is isomorphic to the tensor product $\otimes_{i=1}^M \mathcal{H}_i$ of vector spaces $\mathcal{H}_i$. Thus a basis of $\mathcal{H}$ may be obtained by the tensor product of the univariate bases $\{\phi^i_k, k \in \mathbb{N}\}$, $i = 1, \dots, M$. Consequently, the random variable $Y = \mathcal{M}(\boldsymbol{X})$ that results from the propagation of the uncertainties modelled by $\boldsymbol{X}$ through the computational model $\mathcal{M}$ may be cast as:

$$Y = \sum_{\alpha_1 \in \mathbb{N}} \dots \sum_{\alpha_M \in \mathbb{N}} y_{\alpha_1 \dots \alpha_M} \phi^1_{\alpha_1}(X_1) \dots \phi^M_{\alpha_M}(X_M). \tag{2.5}$$

with $y_{\alpha_1 \dots \alpha_M}$ being deterministic coefficients. For the sake of simplicity, the notation of multi-indices $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_M\}$ is used. The spectral representation of Y is rewritten as follows:

$$Y = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \boldsymbol{\phi}_{\boldsymbol{\alpha}}(\boldsymbol{X}), \tag{2.6}$$

where $\boldsymbol{\phi}_{\boldsymbol{\alpha}}(\boldsymbol{X}) = \prod\limits_{i=1}^{M} \phi^i_{\alpha_i}(X_i)$ are the multivariate basis functions and $y_{\boldsymbol{\alpha}}$ are the associated deterministic coefficients.

2.1.2 Polynomial chaos expansions

Let us define a space $\mathcal{H}_i$ spanned by univariate orthonormal polynomials $\{\psi^i_k, k \in \mathbb{N}\}$, $i = 1, \ldots, M$, which is associated with the probabilistic model of X_i. For instance, when X_i is a uniform over $[-1, 1]$ (resp. standard normal) random variable, the corresponding polynomial basis comprises orthonormal Legendre (resp. Hermite) polynomials (Abramowitz and Stegun, 1970). The spectral representation in Eq. (2.6) becomes the so-called generalized polynomial chaos expansion (PCE) (Xiu and Karniadakis, 2002; Soize and Ghanem, 2004):

$$Y = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X}), \tag{2.7}$$

in which $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X}) = \prod\limits_{i=1}^{M} \psi^i_{\alpha_i}(X_i)$ are multivariate *orthonormal* polynomials obtained by the tensor product of univariate polynomials and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M)$ are the multi-indices with $\alpha_i, i = 1, \ldots, M$, denoting the degree of the univariate polynomial in X_i.

In practice, the input random variables are usually not standardized, therefore it is necessary to transform the them into standard variables. We define the isoprobabilistic transform $\boldsymbol{Z} = \mathcal{T}^{-1}(\boldsymbol{X})$ which is a mapping from the original random space of X_i's onto a standard space of M basic independent random variables Z_i's. As an example Z_i may be a standard normal random variable or a uniform variable over $[-1, 1]$. Finally, Y can be represented by PCE as follows:

$$Y = \mathcal{M}(\boldsymbol{X}) = \mathcal{M} \circ \mathcal{T}(\boldsymbol{Z}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{Z}). \tag{2.8}$$

2.1.3 Case with dependent random variables

The above polynomial chaos expansion was derived with the assumption that the input random variables are independent. In the case when the input vector comprises dependent variables, an isoprobabilistic transform must be used (Sudret, 2015). Assuming that the dependence between the components of $\boldsymbol{X}$ is modelled by a Gaussian copula (Nelsen, 2006), the isoprobabilistic transform is built as follows. Each input random variable X_i of marginal cumulative distribution function (CDF) $F_{X_i}(X_i)$ is first transformed into a standard normal variable Z_i:

$$Z_i = \Phi^{-1}\left(F_{X_i}(X_i)\right), \tag{2.9}$$

where Φ is the standard normal CDF. The resulting Gaussian random vector $\boldsymbol{Z} = (Z_1, \dots, Z_M)$ has zero-mean and correlation matrix $\boldsymbol{R}$ with pseudo-correlation coefficients $R_{i,j} = \rho(Z_i, Z_j) = \operatorname{Cov}[Z_i, Z_j] / \left(\sigma_{Z_i} \sigma_{Z_j}\right)$ in which σ_{Z_i} and σ_{Z_j} are standard deviations of Z_i and Z_j respectively. Next, $\boldsymbol{Z}$ is transformed into a standard normal random vector $\boldsymbol{U}$ with identity covariance matrix by means of the following linear transform:

$$\boldsymbol{U} = \boldsymbol{L}^{-1} \cdot \boldsymbol{D}^{-1} \cdot \boldsymbol{Z}, \tag{2.10}$$

in which $\boldsymbol{D} = \operatorname{Diag}(\sigma_{Z_i}, \dots, \sigma_{Z_M})$ is the diagonal matrix with the i^{th} element on the diagonal being equal to σ_{Z_i} and $\boldsymbol{L}$ is the lower-triangular matrix obtained with the Cholesky decomposition of the correlation matrix $\boldsymbol{R}$. Finally, the PCE of the random variable Y reads:

$$Y = \mathcal{M}(\boldsymbol{X}) = \mathcal{M} \circ \mathcal{T}(\boldsymbol{U}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{U}). \tag{2.11}$$

For more general forms of dependence (*e.g.* other copulas), the Rosenblatt transform (Rosenblatt, 1952) may be used.

2.1.4 Truncation schemes

In practice, use of an infinite series expansion in Eq. (2.7) is not tractable. A truncated form of PCE with a finite number of terms must be used instead:

$$Y = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X}) + \epsilon, \tag{2.12}$$

in which $\mathcal{A}$ is the countable truncation set of the multi-indices $\boldsymbol{\alpha}$ and ϵ is the truncation-induced error. The most popular truncation scheme consists in considering only the polynomials with total degree not exceeding a prescribed value p, *i.e.* :

$$\mathcal{A} \equiv \mathcal{A}^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : ||\boldsymbol{\alpha}||_1 = \sum_{i=1}^{M} \alpha_i \leqslant p\}. \tag{2.13}$$

This standard scheme is, however, not effective in cases of high dimensionality in the random space and high degree polynomials. To overcome this issue, Blatman (2009) introduced two schemes that allow to reduce significantly the cardinality $D = \operatorname{card}(\mathcal{A}^{M,p}) = (M+p)!/(M!\,p!)$ of the truncation set.

In the first scheme, one assumes that the effects of univariate polynomials, and interaction between low-degree functions are more important than high-degree interaction effects. Thus, in the multi-indices space, the relevant polynomials lie in the subset defined by the following hyperbolic truncation scheme:

$$\mathcal{A} \equiv \mathcal{A}_q^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : ||\boldsymbol{\alpha}||_q = \left(\sum_{i=1}^{M} \alpha_i^q\right)^{1/q} \leqslant p\,\}, \tag{2.14}$$

where $0 < q \leq 1$ is the parameter governing the hyperbol and p is the prescribed maximum degree of the multivariate polynomials. Given a value of p, the cardinality of $\mathcal{A}_q^{M,p}$ increases as q increases but remains smaller than the size of $\mathcal{A}^{M,p}$ (which corresponds to $q = 1$).

In the second scheme, one assumes that low-rank[1] polynomials are more important than high-rank terms, *i.e.* two-dimensional interaction polynomials are more relevant than three-dimensional functions and so on. This assumption is based on the so-called hierarchy principle (Yuan et al., 2007), according to which the model can be approximated by low-rank terms. The low-rank truncation scheme reads (Blatman and Sudret, 2011):

$$\mathcal{A} \equiv \mathcal{A}_r^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : ||\boldsymbol{\alpha}||_0 = \sum_{i=1}^{M} \mathbb{1}_{\alpha_i > 0} \leq r,\ ||\boldsymbol{\alpha}||_1 = \sum_{i=1}^{M} \alpha_i \leqslant p\}, \tag{2.15}$$

in which $||\boldsymbol{\alpha}||_0$ is the rank of the multivariate polynomial $\boldsymbol{\psi}_{\boldsymbol{\alpha}}$, defined as the total number of non-zero component indices $\alpha_i, i = 1, \ldots, M$. The prescribed rank r is chosen as a small integer value, *e.g.* $r = 2, 3, 4$.

2.1.5 Computation of the coefficients

In the context of the spectral stochastic finite element method, the Galerkin scheme has been commonly used. The Galerkin scheme consists in converting the original system of equations to the weak formulation, resulting in a linear system of equations that allows to determine the coordinates of the response in the random space. This is an intrusive approach due to the fact that the original system of equations, *e.g.* the finite element code, is modified.

Non-intrusive methods were proposed as an alternative to the Galerkin scheme, in which the expansion coefficients are estimated using only an experimental design (ED), *i.e.* a series of samples of input parameters and associated output values obtained by running the original computational model. It is worth emphasizing that the latter is neither modified nor adapted. In the literature, the non-intrusive approach includes stochastic collocation, projection and regression methods.

The stochastic collocation method is based on the so-called Lagrange interpolation scheme. The computational model is approximated by means of Lagrange polynomials, which satisfy the solutions at the collocation nodes. The coefficients of the expansion are the output values at those specified points. This method must be used with structured collocation nodes, *i.e.* the points obtained with sparse grids (Nobile et al., 2006). The convergence of the approximation is crucially affected by the selected ED.

In the projection method (Le Maître et al., 2001; Matthies and Keese, 2005), using the orthogonality of the polynomials, the expansion coefficients are estimated by the inner products between the expansion and the corresponding polynomial functions. Numerical integration techniques are applied to obtain the approximate values of the estimates. The accuracy of this method relies on the choice of the integration weights and nodes, which

---

[1] The "rank" of a polynomial function $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\cdot)$ is the l_0-norm of the multi-index $||\boldsymbol{\alpha}||_0$.

can be random, quasi-random or determined by quadrature rules.

The regression method consists in estimating the set of coefficients $\hat{\boldsymbol{y}}_{\boldsymbol{\alpha}} = \{y_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{A}\}$ that minimizes the mean square error:

$$\mathbb{E}\left[\epsilon^2\right] \overset{\text{def}}{=} \mathbb{E}\left[\left(Y - \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\right)^2\right], \tag{2.16}$$

which means:

$$\hat{\boldsymbol{y}}_{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{y}_{\boldsymbol{\alpha}} \in \mathbb{R}^{\text{card}\mathcal{A}}} \mathbb{E}\left[\left(\mathcal{M}(\boldsymbol{X}) - \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\right)^2\right]. \tag{2.17}$$

In practice, the coefficients are obtained by minimizing an empirical mean over a sample set:

$$\hat{\boldsymbol{y}}_{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{y}_{\boldsymbol{\alpha}} \in \mathbb{R}^{\text{card}\mathcal{A}}} \frac{1}{N} \sum_{i=1}^{N} \left(\mathcal{M}(\boldsymbol{x}^{(i)}) - \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{x}^{(i)})\right)^2, \tag{2.18}$$

where $\mathcal{X} = \left\{\boldsymbol{x}^{(i)},\, i = 1, \ldots, N\right\}$ is obtained with random sampling of the input vector. The computational model $\mathcal{M}$ is run for each point in $\mathcal{X}$, resulting in the sample set of output quantity values $\mathcal{Y} = \left\{y^{(i)} = \mathcal{M}(\boldsymbol{x}^{(i)}),\, i = 1, \ldots, N\right\}$. By evaluating the polynomial basis onto each sample point in $\mathcal{X}$, one obtains the information matrix, which is defined as follows:

$$\boldsymbol{A} \overset{\text{def}}{=} \left\{A_{ij} = \boldsymbol{\psi}_j(\boldsymbol{x}^{(i)}),\, i = 1, \ldots, N,\, j = 1, \ldots, \text{card}\mathcal{A}\right\}, \tag{2.19}$$

i.e. the i^{th} row of $\boldsymbol{A}$ is the evaluation of the polynomial basis functions at the point $\boldsymbol{x}^{(i)}$. This is basically the problem of estimating the parameters of a linear regression model, for which the least squares solution reads:

$$\hat{\boldsymbol{y}}_{\boldsymbol{\alpha}} = \left(\boldsymbol{A}^{\mathsf{T}} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^{\mathsf{T}} \mathcal{Y}. \tag{2.20}$$

2.1.6 Error estimators

Using an experimental design (ED) containing the input samples $\mathcal{X}$ and the corresponding response values $\mathcal{Y}$, the coefficients of the polynomial chaos representation can be estimated, leading to the following approximative model:

$$\widehat{\mathcal{M}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} \hat{y}_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}). \tag{2.21}$$

The expected squared error of the approximation, also called the generalization error reads:

$$Err = \mathbb{E}\left[\left(\mathcal{M}(\boldsymbol{X}) - \widehat{\mathcal{M}}(\boldsymbol{X})\right)^2\right]. \tag{2.22}$$

In practice, it is not possible to calculate this quantity exactly. Instead, an empirical estimate over the ED can be obtained:

$$Err_{emp} = \frac{1}{N} \sum_{i=1}^{N} \left(\mathcal{M}(\boldsymbol{x}^{(i)}) - \widehat{\mathcal{M}}(\boldsymbol{x}^{(i)}) \right)^2 . \tag{2.23}$$

The relative empirical error is defined by:

$$\epsilon_{emp} = \frac{Err_{emp}}{\mathrm{Var}\left[\mathcal{Y}\right]}, \tag{2.24}$$

in which

$$\mathrm{Var}\left[\mathcal{Y}\right] = \frac{1}{N-1} \sum_{i=1}^{N} \left(y^{(i)} - \bar{\mathcal{Y}} \right)^2 , \; \bar{\mathcal{Y}} = \frac{1}{N} \sum_{i=1}^{N} y^{(i)}. \tag{2.25}$$

The generalization error Err is underestimated by the empirical error Err_{emp} (Blatman, 2009). This is due to the so-called overfitting phenomenon where the latter systematically decreases with the increasing degree of the polynomial basis while its generalization counterpart Err may indeed increase. It is of utmost importance to use an error indicator that is less prone to overfitting. Herein the leave-one-out (LOO) cross validation error is utilized.

Cross validation consists in partitioning the ED into two complementary subsets, training a model using one subset, then validating its prediction on the other one. In this context, the term LOO means that the validation set comprises only one sample. Setting one point $\boldsymbol{x}^{(i)}$ apart from $\mathcal{X}$, one can build a PCE model $\mathcal{M}^{\mathrm{PC}\backslash i}(\cdot)$ from the remaining points $\mathcal{X}\backslash\boldsymbol{x}^{(i)} = \left\{ \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(i-1)}, \boldsymbol{x}^{(i)}, \ldots, \boldsymbol{x}^{(n)} \right\}$. The predicted residual error at the point $\boldsymbol{x}^{(i)}$ reads:

$$\Delta^{(i)} \stackrel{\mathrm{def}}{=} \mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}^{\mathrm{PC}\backslash i}(\boldsymbol{x}^{(i)}). \tag{2.26}$$

The LOO error is defined as the empirical mean of the residual errors $\Delta^{(i)}$, *i.e.* :

$$\widehat{\mathrm{Err}}_{LOO} = \frac{1}{n} \sum_{i=1}^{n} \Delta^{(i)2}. \tag{2.27}$$

At first glance, computing the LOO error appears computationally demanding since it requires to train and validate N PCE models. However, due to the linear regression formula of PCE, one can compute the LOO error $\widehat{\mathrm{Err}}_{LOO}$ by means of algebraic derivations from a *single* PCE model $\mathcal{M}^{\mathrm{PC}}(\cdot)$ built with the full ED $\mathcal{X}$ as follows (Blatman, 2009):

$$\widehat{\mathrm{Err}}_{LOO} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}^{\mathrm{PC}}(\boldsymbol{x}^{(i)})}{1 - h_i} \right)^2 , \tag{2.28}$$

in which h_i is the i^{th} diagonal term of the matrix $\boldsymbol{A} \left(\boldsymbol{A}^{\mathsf{T}} \boldsymbol{A} \right)^{-1} \boldsymbol{A}^{\mathsf{T}}$. The relative LOO error is obtained by:

$$\widehat{\epsilon}_{LOO} = \frac{\widehat{Err}_{LOO}}{\mathrm{Var}\left[\mathcal{Y}\right]}. \tag{2.29}$$

2.2 Sparse adaptive polynomial chaos expansions

In practice, not all polynomial functions specified in the truncation basis are important and contribute significantly to the output. In addition, the ordinary least squares method requires a large number of model evaluations, typically two to three times the cardinality of the polynomial basis. In case the problems have high dimensionality and high degree polynomials must be used, *i.e.* P is large, the required size of the ED soon becomes excessive. However, the number of available simulations is usually limited. Therefore, it is of utmost importance to use a technique that can detect the relevant polynomials even with a limited number of simulations. Least angle regression (LARS) (Efron et al., 2004) is an iterative method for regressor selection, which was proven to be particularly powerful when the number of regressors is much larger than the size of the ED, and only a few regressors are relevant.

In the following, one will shortly describe the adaptive sparse PCE scheme proposed by Blatman and Sudret (2011) which is a non-intrusive least-square minimization technique based on the LARS algorithm. LARS allows one to achieve a sparse representation, *i.e.* the number of active polynomial functions is small compared to the size of the candidate set. This is done in an adaptive manner, *i.e.* the candidate functions become active one after another in the descending order of their importance. The relevance of a candidate polynomial is measured by means of its correlation with the residual of the expansion obtained in the previous iteration. The optimal PCE is chosen to minimize the LOO error estimator. The sparse adaptive PCE scheme can be summarized as follows (Blatman and Sudret, 2011):

1. Initialize the set of candidate regressors to the full basis and the set of selected regressors to $\varnothing$.
2. Initialize all coefficients equal to 0. Set the residual equal to the output vector.
3. For each iteration until $m = \min(P, N-1)$ regressors have been analyzed, perform the following steps:
    - Determine the candidate polynomial that is most correlated to the current residual and add it to the selected basis,
    - Update simultaneously all the active coefficients associated with the selected basis until another regressor is as correlated with the residual as they are. This is done using $\hat{\boldsymbol{y}}_{\boldsymbol{\alpha}}^{(k+1)} = \hat{\boldsymbol{y}}_{\boldsymbol{\alpha}}^{(k)} + \gamma^{(k)}\, \tilde{\boldsymbol{\omega}}^{(k)}$ with $\tilde{\boldsymbol{\omega}}^{(k)}$ and $\gamma^{(k)}$ are descent direction and step respectively. The residual is updated correspondingly.
4. Choose the best PCE from the iteration in which the LOO error indicator is minimized.

2.3 Post processing of polynomial chaos expansions

Let us discuss the use of PCEs for propagating the uncertainties from the input parameters to the output quantity of interest. The analytical formula of PCEs allows one to post-process the representation in order to obtain the output statistical moments without additional cost. The output probability density function can also be obtained at a computational expense significantly smaller than that of the traditional MCS approach. This is due to the fact that with PCE one can predict the output on a large set of input parameters by simply evaluating polynomial functions. In particular, the polynomial coefficients provide sensitivity measures that reveal directly the importance of the uncertain parameters in the model.

2.3.1 Statistical moments

The analysts are commonly interested in the statistical moments of the considered quantity, for instance its expected value and variance, which can partially describe the uncertainty characteristics.

PCEs bear a particular property which is the orthonormality of the polynomial functions. Indeed, the following equations hold:

$$\mathbb{E}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = 0 \;\forall \boldsymbol{\alpha} \neq \boldsymbol{0}, \tag{2.30}$$

and

$$\mathbb{E}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\,\boldsymbol{\psi}_{\boldsymbol{\beta}}(\boldsymbol{X})\right] = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}} \;\forall \boldsymbol{\alpha},\, \boldsymbol{\beta}. \tag{2.31}$$

Using this property, the estimates of the output first and second order moments, *i.e.* mean and variance, can be computed analytically as follows:

$$\mu_Y \overset{\text{def}}{=} \mathbb{E}\left[\sum_{\boldsymbol{\alpha}\in\mathbb{N}^M} y_{\boldsymbol{\alpha}}\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = y_{\boldsymbol{0}}, \tag{2.32}$$

$$D \overset{\text{def}}{=} \operatorname{Var}\left[\sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{A}\\ \boldsymbol{\alpha}\neq\boldsymbol{0}}} y_{\boldsymbol{\alpha}}^2 \operatorname{Var}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X})\right] = \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{A}\\ \boldsymbol{\alpha}\neq\boldsymbol{0}}} y_{\boldsymbol{\alpha}}^2. \tag{2.33}$$

One notes that the first and second order moments of the output are obtained using only the coefficients of PCE, which does not require additional computational cost. The higher order moments can also be estimated straightforwardly with the expansion coefficients by means of analytical derivations. The reader is referred to Sudret (2007) for the closed form expressions of those quantities.

2.3.2 Probability density function

Knowing the statistical moments may be insufficient to fully describe the uncertainty in the output. As an example, the multimodal distribution of a quantity can only be observed on the entire probability density function (PDF).

Due to its nature, PCEs can be used inexpensively for a large number of evaluations, providing a large sample set of the output. This allows to compute the PDF of the output numerically by means of the so-called kernel density estimation technique.

For a single random variable Y, the kernel density estimate of the PDF reads (Wand and Jones, 1995):

$$\hat{f}_Y(y) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{y - y_i}{h}\right), \tag{2.34}$$

where $\{y_1, \ldots, y_N\}$ is a sample set of Y, h is the *bandwidth* parameter and $K(\cdot)$ is the *kernel* function which integrates to one. Classical kernel functions are the Epanechnikov, uniform, normal and triangular functions. The choice of the kernel is known not to affect strongly the quality of the estimate provided the sample set is sufficiently large (Wand and Jones, 1995). In case a standard normal PDF is adopted for the kernel, *i.e.* $K(y) = \varphi(y) \equiv \exp\left(-y^2/2\right)/\sqrt{2\pi}$, the kernel density estimate rewrites:

$$\hat{f}_Y(y) = \frac{1}{Nh} \sum_{i=1}^{N} \varphi\left(\frac{y - y_i}{h}\right). \tag{2.35}$$

In contrast, the choice of the bandwidth h is crucial for the kernel density estimate. An inappropriate value of h can lead to an oversmoothed or undersmoothed estimated PDF. Different techniques for computing the bandwidth are thoroughly reviewed by Duong (2004).

2.3.3 Variance-based sensitivity analysis

Sensitivity analysis (SA) examines the sensitivity of the model output with respect to the input parameters, *i.e.* how the output variability is affected by the variability of the input factors (Saltelli et al., 2000, 2004, 2008). Two major classes of techniques for SA in the literature are variance-based and derivative-based methods. *Variance-based* SA relies upon decomposing the output variance into contributions of marginal effects of input parameters and interactions between them. The use of PCEs to derive analytically variance-based sensitivity measures is reviewed hereafter. For the sake of clarity, the author's original contribution is presented in the next section in which analytical derivations are introduced for computing derivative-based global sensitivity measures with PCEs.

Using the orthonormality property of the polynomial basis, Sudret (2006, 2008) proposed an original post-processing of PCEs for conducting sensitivity analysis. For any subset of variables $\boldsymbol{X}_{\mathbf{u}}$ with $\mathbf{u} = \{i_1, \ldots, i_s\} \subset \{1, \ldots, M\}$, one defines the set of multivariate polynomials $\boldsymbol{\psi}_{\boldsymbol{\alpha}}$ depending only on $\boldsymbol{X}_{\mathbf{u}}$, which is specified by a set of multi-indices:

$$\mathcal{A}_{\mathbf{u}} = \{\boldsymbol{\alpha} \in \mathcal{A} : \alpha_k \neq 0 \text{ if and only if } k \in \mathbf{u}\}\,. \tag{2.36}$$

The Sobol' decomposition of the truncated PCE may be written as follows:

$$Y = y_{\mathbf{0}} + \sum_{\substack{\mathbf{u} \subset \{1,\dots,M\} \\ \mathbf{u} \neq \emptyset}} \mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}}), \tag{2.37}$$

where:

$$\mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}}) \stackrel{\text{def}}{=} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{u}}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{X}). \tag{2.38}$$

The *Sobol' sensitivity index* $S_{\mathbf{u}}$ for the subset of variables $\boldsymbol{X}_{\mathbf{u}}$ is defined as follows (Sobol', 2001):

$$S_{\mathbf{u}} \stackrel{\text{def}}{=} \frac{D_{\mathbf{u}}}{D} = \frac{\mathrm{Var}\left[\mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}})\right]}{D}, \tag{2.39}$$

where the partial variance $D_{\mathbf{u}}$ reads:

$$D_{\mathbf{u}} = \mathrm{Var}\left[\mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}})\right] = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\mathbf{u}}} y_{\boldsymbol{\alpha}}^2. \tag{2.40}$$

The *total sensitivity index* for subset $\boldsymbol{X}_{\mathbf{u}}$ is given by (Sobol', 2001):

$$S_{\mathbf{u}}^T \stackrel{\text{def}}{=} \frac{D_{\mathbf{u}}^T}{D} = \frac{\sum_{\mathbf{v} \supset \mathbf{u}} D_{\mathbf{v}}}{D} = \frac{\sum_{\mathbf{v} \supset \mathbf{u}} \mathrm{Var}\left[\mathcal{M}_{\mathbf{v}}(\boldsymbol{X}_{\mathbf{v}})\right]}{D}, \tag{2.41}$$

where the sum in the numerator is extended over all sets $\mathbf{v} = \{j_1, \dots, j_t\}$ which contain $\mathbf{u}$. It represents the total amount of uncertainty apportioned to the subset of variables $\boldsymbol{X}_{\mathbf{u}}$, including its sole effects and interactions with other variables.

As a consequence, having a PCE at hand, the Sobol' indices at any order may be computed by a mere combination of the squares of the coefficients. For instance, for a single variable X_i, $i = 1, \dots, M$, the first order PC-based Sobol' index reads:

$$S_i = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_i} y_{\boldsymbol{\alpha}}^2 / D, \quad \mathcal{A}_i = \{\boldsymbol{\alpha} \in \mathcal{A} : \alpha_i > 0,\, \alpha_{j \neq i} = 0\}\,, \tag{2.42}$$

whereas the total PC-based Sobol' index is:

$$S_i^T = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_i^T} y_{\boldsymbol{\alpha}}^2 / D, \quad \mathcal{A}_i^T = \{\boldsymbol{\alpha} \in \mathcal{A} : \alpha_i > 0\}\,. \tag{2.43}$$

2.4 Polynomial chaos derivative-based global sensitivity measures

2.4.1 Derivative-based sensitivity measures

Derivative-based sensitivity analysis originates from the *Morris method* (Morris, 1991) which was introduced to detect input parameters having substantial impacts on the output and those with negligible influences using parameters elementary effects.

Consider a computational model $Y = \mathcal{M}(\boldsymbol{X})$ with Y being the scalar output and $\boldsymbol{X}$ being the vector of uncertain input variables. In the Morris method (Morris, 1991), when investigating the variable X_i one first samples an experimental design (ED) $\mathcal{X} = \left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(N)}\right\}$ and then varies this sample in the i^{th} direction. The elementary effect (EE) is defined as:

$$EE_i^{(j)} = \frac{\mathcal{M}(\boldsymbol{x}_r^{(j)}) - \mathcal{M}(\boldsymbol{x}^{(j)})}{\Delta}, \tag{2.44}$$

in which $\boldsymbol{x}^{(j)} = \left\{x_1^{(j)}, \dots, x_i^{(j)}, \dots, x_M^{(j)}\right\}$ is the j^{th} sample point and $\boldsymbol{x}_r^{(j)} = \left\{x_1^{(j)}, \dots, x_i^{(j)} + \Delta, \dots, x_M^{(j)}\right\}$ is the perturbed sample. The Morris importance measure (Morris factor) is defined as the average of the EE_i's:

$$\mu_i = \frac{1}{N} \sum_{j=1}^{N} EE_i^{(j)}. \tag{2.45}$$

By definition, the empirical variance σ_i^2 of the EEs is given by:

$$\sigma_i^2 = \frac{1}{N-1} \sum_{j=1}^{N} (EE_i^{(j)} - \mu_i)^2. \tag{2.46}$$

The resulting mean μ_i and standard deviation σ_i are usually plotted in a two-dimensional graph, see Figure 2.1. This graph is interpreted as follows: when μ_i and σ_i are jointly small (lower left corner of the plot), the parameter X_i is considered as unimportant. When μ_i is large and σ_i is small (lower right corner) then X_i is considered as an influential parameter, which is in a quasi-linear relationship with the output (a zero standard deviation indicates a fully linear dependence). Finally, when σ_i is large (upper part of the plot), the parameter X_i is deemed important and the output depends on this very parameter in a nonlinear way and/or it interacts with other parameters.

Kucherenko et al. (2009) generalized the quantities in Eqs.(2.45)-(2.46) as follows:

$$\mu_i \overset{\text{def}}{=} \mathbb{E}\left[\frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{X})\right] = \int_{\mathcal{H}^M} \frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \tag{2.47}$$

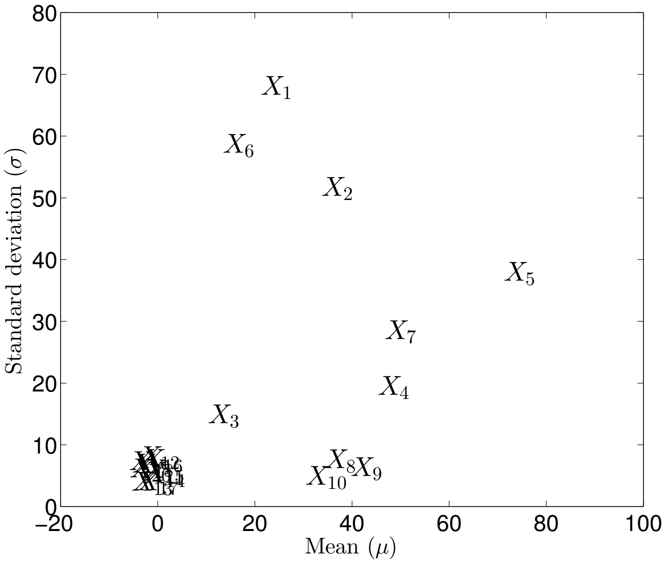


Figure 2.1 – Morris method: representation of the mean and standard deviation of the elementary effects (output plot of the UQLab software (Marelli et al., 2015))

$$\sigma_i^2 = \int_{\mathcal{H}^M} \left[\frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{x}) \right]^2 \mathrm{d}\boldsymbol{x} - \mu_i^2, \tag{2.48}$$

provided that $\dfrac{\partial \mathcal{M}}{\partial x_i}$ is square-integrable on the unit hypercube $\mathcal{H}^M$ (assuming the input parameters have been rescaled to this hypercube). The interpretation of these measures is similar as for the original Morris method.

The elementary effects may take positive or negative values, therefore by calculating their average (or expected value) the effects may counteract. This might lead to a misinterpretation of the importance of X_i. To avoid this, Campolongo et al. (2007) modified the Morris factor as follows:

$$\mu_i^* = \mathbb{E}\left[\left| \frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{X}) \right| \right]. \tag{2.49}$$

Recently, Sobol' and Kucherenko (2009) introduced a new sensitivity measure (SM) which is the mean-squared derivative of the model with respect to X_i:

$$\nu_i = \mathbb{E}\left[\left(\frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{X}) \right)^2 \right]. \tag{2.50}$$

Sobol' and Kucherenko (2009) and Lamboni et al. (2013) established a link between the measure ν_i in Eq. (2.50) and the total Sobol' indices. In case of a uniform variable $X_i \sim [a_i,\, b_i]$ this relationship reads:

$$S_i^T \leq S_i^{DGSM} = \frac{(b_i - a_i)^2}{\pi^2} \frac{\nu_i}{D}, \tag{2.51}$$

where S_i^{DGSM} is the upper-bound to the total sensitivity index S_i^T and D is the model output variance. The above result can be extended to other types of distributions. When $X_i \sim \mathcal{N}(a_i, b_i)$ is a Gaussian random variable with mean a_i and variance b_i^2, one gets:

$$S_i^T \leq S_i^{DGSM} = {b_i}^2 \frac{\nu_i}{D}. \tag{2.52}$$

In the general case, Lamboni et al. (2013) define the upper bound of the total Sobol' index of X_i as:

$$S_i^{DGSM} = 4{C_i}^2 \frac{\nu_i}{D}, \tag{2.53}$$

in which $C_i = \sup\limits_{x \in \mathbb{R}} \dfrac{\min\left[F_{X_i}(x), 1 - F_{X_i}(x)\right]}{f_{X_i}(x)}$ is the Cheeger constant, F_{X_i} is the cumulative distribution function of X_i and f_{X_i} is the probability density function of X_i.

In the following, one will demonstrate the use of PCEs to compute the DGSMs analytically, first in cases of Hermite and Legendre polynomials, then in a general case. This work has originally been presented in Sudret and Mai (2013c) and Sudret and Mai (2015).

2.4.2 Case of Hermite polynomial chaos expansion

In this section, we consider a computational model $Y = \mathcal{M}(\boldsymbol{X})$ where Y is the scalar output and $\boldsymbol{X} = (X_i, \ldots, X_M)$ is the input vector composed of M *independent* Gaussian variables $X_i \sim \mathcal{N}(\mu_i, \sigma_i)$. One first uses an isoprobabilistic transform which reads:

$$\boldsymbol{X} = \mathcal{T}(\boldsymbol{Z}): \qquad X_i = \mu_i + \sigma_i Z_i, \tag{2.54}$$

to convert X_i into a standard normal random variable $Z_i \sim \mathcal{N}(0,1)$. The truncated PCE of Y reads:

$$Y = \mathcal{M}_{\mathcal{A}}(\boldsymbol{X}) = \mathcal{M}_{\mathcal{A}}\left(\mathcal{T}(\boldsymbol{Z})\right) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{Z}), \tag{2.55}$$

in which $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_M\}$ is a multi-index, $\mathcal{A}$ is the truncated set of multi-indices $\boldsymbol{\alpha}$, $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z}) = \prod\limits_{i=1}^{M} \tilde{He}_{\alpha_i}(z_i)$ is the multivariate polynomial basis obtained by the tensor product of univariate orthonormal Hermite polynomials $\tilde{He}_{\alpha_i}(z_i)$ (see A.1) and $y_{\boldsymbol{\alpha}}$ is the deterministic coefficient associated with $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z})$.

Since $\mathcal{T}$ is a one-to-one mapping with $\dfrac{\partial z_i}{\partial x_i} = \dfrac{1}{\sigma_i}$, the derivative-based sensitivity index reads:

$$\nu_i = \mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}}}{\partial x_i}(\boldsymbol{X})\right)^2\right] = \mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}} \circ \mathcal{T}}{\partial z_i} \frac{\partial z_i}{\partial x_i}\right)^2\right] = \frac{1}{{\sigma_i}^2} \mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}} \circ \mathcal{T}}{\partial z_i}(\boldsymbol{Z})\right)^2\right]. \tag{2.56}$$

The DGSM of X_i, in other words the corresponding upper bound to the total Sobol' index S_i^T, is computed according to Eq. (2.52):

$$S_i^{DGSM} = \sigma_i^2 \frac{\nu_i}{D} = \frac{1}{D}\mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}} \circ \mathcal{T}}{\partial z_i}(\boldsymbol{Z})\right)^2\right] = \frac{1}{D}\mathbb{E}\left[\left(\frac{\partial}{\partial z_i}\sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{Z})\right)^2\right], \tag{2.57}$$

in which $D \stackrel{\text{def}}{=} \sum_{\boldsymbol{\alpha}\in\mathcal{A},\, \boldsymbol{\alpha}\neq\mathbf{0}} y_{\boldsymbol{\alpha}}^2$ is the PCE-based estimate of the output variance. This requires computing the partial derivatives of $\mathcal{M}_{\mathcal{A}}(\boldsymbol{z}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z})$. The first order derivative $\tilde{He}_n^{'}(z) = \frac{\mathrm{d}\tilde{He}_n}{\mathrm{d}z}(z)$ is proved to be (see A.1):

$$\tilde{He}_n^{'}(z) = \sqrt{n}\,\tilde{He}_{n-1}(z). \tag{2.58}$$

Consequently, the derivative of the multivariate orthonormal Hermite polynomial $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z}) = \prod_{i=1}^{M} \tilde{He}_{\alpha_i}(z_i)$ with respect to z_i is obtained with:

$$\frac{\partial\boldsymbol{\psi}_{\boldsymbol{\alpha}}}{\partial z_i}(\boldsymbol{z}) = \prod_{\substack{j=1\\ j\neq i}}^{M} \tilde{He}_{\alpha_j}(z_j)\,\sqrt{\alpha_i}\tilde{He}_{\alpha_i-1}(z_i) \tag{2.59}$$

provided that $\alpha_i > 0$ and $\frac{\partial\boldsymbol{\psi}_{\boldsymbol{\alpha}}}{\partial z_i}(\boldsymbol{z}) = 0$ otherwise. Then the derivative of a Hermite PCE with respect to z_i is given by the following expression:

$$\frac{\partial\mathcal{M}_{\mathcal{A}}}{\partial z_i}(\boldsymbol{z}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}^{(i)}} \sqrt{\alpha_i}\,y_{\boldsymbol{\alpha}}\boldsymbol{\psi}_{\boldsymbol{\alpha}_i^{'}}, \tag{2.60}$$

in which $\mathcal{A}^{(i)} = \{\boldsymbol{\alpha}\in\mathcal{A}, \alpha_i > 0\}$ is the set of multi-indices $\boldsymbol{\alpha}$ with a non-zero i^{th} component α_i and $\boldsymbol{\alpha}_i^{'} = \{\alpha_1, \ldots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \ldots, \alpha_M\}$ is the index vector derived from $\boldsymbol{\alpha}$ by subtracting 1 from α_i. The expectation of the squared derivative in Eq. (2.57) is reformulated as:

$$\mathbb{E}\left[\left(\frac{\partial\mathcal{M}_{\mathcal{A}}}{\partial z_i}(\boldsymbol{Z})\right)^2\right] = \mathbb{E}\left[\sum_{\boldsymbol{\alpha}\in\mathcal{A}^{(i)}}\sum_{\boldsymbol{\beta}\in\mathcal{A}^{(i)}} \sqrt{\alpha_i\,\beta_i}\,y_{\boldsymbol{\alpha}}y_{\boldsymbol{\beta}}\,\boldsymbol{\psi}_{\boldsymbol{\alpha}_i{'}}\boldsymbol{\psi}_{\boldsymbol{\beta}_i^{'}}\right]. \tag{2.61}$$

Due to the linearity of the expectation operator, the above equation requires computing $\mathbb{E}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}_i{'}}\boldsymbol{\psi}_{\boldsymbol{\beta}_i^{'}}\right]$. Note that the orthonormality of the polynomial basis leads to $\mathbb{E}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}_i{'}}\boldsymbol{\psi}_{\boldsymbol{\beta}_i^{'}}\right] = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ where $\delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ is the Kronecker symbol. Thus one obtains:

$$\mathbb{E}\left[\left(\frac{\partial\mathcal{M}_{\mathcal{A}}}{\partial z_i}(\boldsymbol{Z})\right)^2\right] = \sum_{\boldsymbol{\alpha}\in\mathcal{A}^{(i)}} \alpha_i\,y_{\boldsymbol{\alpha}}^2. \tag{2.62}$$

The result is that, the DGSM of a Hermite PCE is given the following *analytical expres-*

sion:

$$\hat{S}_i^{DGSM} = \frac{1}{D} \sum_{\boldsymbol{\alpha} \in \mathcal{A}^{(i)}} \alpha_i \, y_{\boldsymbol{\alpha}}^2 = \frac{\sum\limits_{\boldsymbol{\alpha} \in \mathcal{A}^{(i)}} \alpha_i \, y_{\boldsymbol{\alpha}}^2}{\sum\limits_{\boldsymbol{\alpha} \in \mathcal{A}, \, \boldsymbol{\alpha} \neq \mathbf{0}} y_{\boldsymbol{\alpha}}^2}. \tag{2.63}$$

Note that the total Sobol' indices S_i^T can be obtained directly from the PCE by $\hat{S}_i^T = \sum\limits_{\boldsymbol{\alpha} \in \mathcal{A}^{(i)}} y_{\boldsymbol{\alpha}}^2 / \sum\limits_{\boldsymbol{\alpha} \in \mathcal{A}, \, \boldsymbol{\alpha} \neq \mathbf{0}} y_{\boldsymbol{\alpha}}^2$ as shown in Eq. (2.43). With integer indices $\alpha_i > 0$, it is clear that the inequality $S_i^T \leq S_i^{DGSM}$ always holds.

2.4.3 Case of Legendre polynomial chaos expansion

Let us consider now the case when the input vector $\boldsymbol{X}$ contains M *independent* uniform random variables $X_i \sim \mathcal{U}[a_i, \, b_i]$. An isoprobabilistic transform is used to convert the input factors into standard uniform variables $\boldsymbol{Z} = (Z_i, \, \dots \, , Z_M)$:

$$\boldsymbol{X} = \mathcal{T}(\boldsymbol{Z}) : \qquad X_i = \frac{b_i + a_i}{2} + \frac{b_i - a_i}{2} Z_i, \tag{2.64}$$

where $Z_i \sim \mathcal{U}[-1, 1]$ is uniformly distributed over $[-1, 1]$. The Legendre PCE has the form of the expansion in Eq. (2.55), except that $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z}) = \prod\limits_{i=1}^{M} \tilde{Le}_{\alpha_i}(z_i)$ is now the tensor product of univariate orthonormal Legendre polynomials $\tilde{Le}_{\alpha_i}(z_i)$ (see A.2). The transform $\mathcal{T}$ is a one-to-one linear mapping with $\dfrac{\partial z_i}{\partial x_i} = \dfrac{2}{b_i - a_i}$, thus the derivative-based sensitivity index is given by:

$$\nu_i = \mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}}}{\partial x_i}(\boldsymbol{X})\right)^2\right] = \frac{4}{(b_i - a_i)^2} \, \mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}} \circ \mathcal{T}}{\partial z_i}(\boldsymbol{Z})\right)^2\right]. \tag{2.65}$$

Similarly to Eq. (2.57), the upper bound DGSM to the total Sobol' index S_i^T is computed from Eq. (2.51) as:

$$\begin{aligned} S_i^{DGSM} &= \frac{(b_i - a_i)^2}{\pi^2} \frac{\nu_i}{D} = \frac{4}{\pi^2 D} \, \mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}} \circ \mathcal{T}}{\partial z_i}(\boldsymbol{Z})\right)^2\right] \\ &= \frac{4}{\pi^2 D} \, \mathbb{E}\left[\left(\frac{\partial}{\partial z_i} \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \, \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{Z})\right)^2\right]. \end{aligned} \tag{2.66}$$

It is therefore necessary to compute the derivative of univariate and multivariate Legendre polynomials. Denoting the derivative of an univariate Legendre polynomial by $\tilde{Le}_i^{'}(z) \overset{\text{def}}{=} \dfrac{\mathrm{d}\tilde{Le}_i(z)}{\mathrm{d}z}$, it is proved that (see A.2):

$$\left\{\tilde{Le}_1^{'}(z), \, \dots \, , \tilde{Le}_n^{'}(z)\right\}^{\mathsf{T}} = \boldsymbol{C}^{Le} \cdot \{\tilde{Le}_0(z), \, \dots \, , \tilde{Le}_{n-1}(z)\}^{\mathsf{T}}, \tag{2.67}$$

in which $\boldsymbol{C}^{Le}$ is a constant matrix whose i^{th} row contains the coordinates of the derivative of $\tilde{Le}_i(z)$ onto a basis consisting of lower-degree polynomials $\left\{\tilde{Le}_j(z),\, j=0,\, \dots\, , i-1\right\}$. In other words, one has $\tilde{Le}_i^{'}(z) = \sum\limits_{j=1}^{i} C_{ij}^{Le} \tilde{Le}_{j-1}(z)$. Consequently, the partial derivative of the multivariate orthonormal Legendre polynomials $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z}) = \prod\limits_{i=1}^{M} \tilde{Le}_{\alpha_i}(z_i)$ with respect to z_i is given by:

$$\frac{\partial \boldsymbol{\psi}_{\boldsymbol{\alpha}}}{\partial z_i}(\boldsymbol{z}) = \prod_{\substack{j=1 \\ j\neq i}}^{M} \tilde{Le}_{\alpha_j}(z_j) \left(\sum_{l=1}^{\alpha_i} C_{\alpha_i l}^{Le} \tilde{Le}_{l-1}(z_i) \right). \tag{2.68}$$

For a given multi-index $\boldsymbol{\alpha} = \{\alpha_1,\, \dots\, , \alpha_M\}$, let us define by $\boldsymbol{\alpha}_i^r$ the index vector obtained by substituting the i^{th} component of $\boldsymbol{\alpha}$ with r:

$$\boldsymbol{\alpha}_i^r = \{\alpha_1,\, \dots\, , \alpha_{i-1}, r, \alpha_{i+1},\, \dots\, , \alpha_M\}. \tag{2.69}$$

Using this notation, Eq. (2.68) rewrites as follows:

$$\frac{\partial \boldsymbol{\psi}_{\boldsymbol{\alpha}}}{\partial z_i}(\boldsymbol{z}) = \sum_{l=1}^{\alpha_i} C_{\alpha_i l}^{Le} \prod_{\substack{j=1 \\ j\neq i}}^{M} \tilde{Le}_{\alpha_j}(z_j) \tilde{Le}_{l-1}(z_i). = \sum_{l=1}^{\alpha_i} C_{\alpha_i l}^{Le} \boldsymbol{\psi}_{\boldsymbol{\alpha}_i^{l-1}}. \tag{2.70}$$

Denote by $\mathcal{A}^{(i)}$ the set of multi-indices $\boldsymbol{\alpha}$ having a non-zero index α_i, *i.e.* $\mathcal{A}^{(i)} = \{\boldsymbol{\alpha} \in \mathcal{A}, \alpha_i > 0\}$. The partial derivative of a Legendre PCE with respect to z_i then reads:

$$\frac{\partial \mathcal{M}_{\mathcal{A}}}{\partial z_i}(\boldsymbol{z}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}^{(i)}} y_{\boldsymbol{\alpha}} \frac{\partial \boldsymbol{\psi}_{\boldsymbol{\alpha}}}{\partial z_i}(\boldsymbol{z}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}^{(i)}} \sum_{l=1}^{\alpha_i} y_{\boldsymbol{\alpha}}\, C_{\alpha_i l}^{\mathcal{L}}\, \boldsymbol{\psi}_{\boldsymbol{\alpha}_i^{l-1}}(\boldsymbol{z}). \tag{2.71}$$

Denote by $\mathcal{B}^{(i)}$ the set of multi-indices $\boldsymbol{\beta}$ representing the ensemble of multivariate polynomials generated by differentiating the linear combination of polynomials $\left\{\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{z}),\, \boldsymbol{\alpha}\in\mathcal{A}^{(i)}\right\}$. $\mathcal{B}^{(i)}$ is obtained with:

$$\mathcal{B}^{(i)} = \left\{\boldsymbol{\beta} = \boldsymbol{\alpha} + (k-\alpha_i)\cdot \boldsymbol{e}_i,\, \boldsymbol{\alpha}\in\mathcal{A}^{(i)},\, k=0,\, \dots\, , \alpha_i - 1\right\}, \tag{2.72}$$

where:

$$\boldsymbol{e}_i = (0, \dots, 0, \overbrace{\ 1\ }^{i^{th}\text{pos.}}, 0 \dots, 0). \tag{2.73}$$

The derivative of Legendre PCE rewrites:

$$\frac{\partial \mathcal{M}_{\mathcal{A}}}{\partial z_i}(\boldsymbol{z}) = \sum_{\boldsymbol{\beta}\in\mathcal{B}^{(i)}} b_{\boldsymbol{\beta}}\, \boldsymbol{\psi}_{\boldsymbol{\beta}}(\boldsymbol{z}), \tag{2.74}$$

in which the coefficient $b_{\boldsymbol{\beta}}$ is obtained from Eq.(2.71) as follows:

$$b_{\boldsymbol{\beta}} = \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{A}^{(i)} \\ \boldsymbol{\alpha}=\boldsymbol{\beta}+l\cdot\boldsymbol{e}}} y_{\boldsymbol{\alpha}}\, C_{\alpha_i\,(\beta_i+1)}^{\mathcal{L}}\ \forall\, l\in\mathbb{N}. \tag{2.75}$$

Since the polynomials $\boldsymbol{\psi}_{\boldsymbol{\beta}}$ are also orthonormal, one obtains:

$$\mathbb{E}\left[\left(\frac{\partial \mathcal{M}_{\mathcal{A}}}{\partial z_i}(\boldsymbol{Z})\right)^2\right] = \sum_{\boldsymbol{\beta}\in\mathcal{B}^{(i)}} b_{\boldsymbol{\beta}}^2. \tag{2.76}$$

Finally, the DGSMs of Legendre PCEs read:

$$\hat{S}_i^{DGSM} = \frac{4}{\pi^2}\frac{\sum\limits_{\boldsymbol{\beta}\in\mathcal{B}^{(i)}} b_{\boldsymbol{\beta}}^2}{\sum\limits_{\boldsymbol{\alpha}\in\mathcal{A},\,\boldsymbol{\alpha}\neq\boldsymbol{0}} y_{\boldsymbol{\alpha}}^2}. \tag{2.77}$$

2.4.4 General case

Consider now the general case where the input vector $\boldsymbol{X}$ contains M independent random variables with different prescribed PDFs, *i.e.* Gaussian, uniform or others. Such a problem can be addressed using generalized polynomial chaos expansions (Xiu and Karniadakis, 2002). As the derivatives of Hermite and Legendre polynomials in Eq. (2.58) and Eq. (2.67) hold component-wise, the derivative of a generalized PCE with respect to one parameter results in a generalized PCE of lower degree, as identically given by Eq. (2.60) and (2.71). This representation can be obtained given the proper matrix yielding the derivative of the univariate polynomials in the same univariate orthonormal basis, see A.1 for Hermite polynomials and A.2 for Legendre polynomials. The derivation for Laguerre polynomials is also given in A.3 for the sake of completeness.

2.4.5 Numerical applications

2.4.5.1 Morris function

We first consider the Morris function that is widely used in the literature for sensitivity analysis (Morris, 1991; Lamboni et al., 2013). This function reads:

$$y = \beta_o + \sum_{i=1}^{20}\beta_i\,\omega_i + \sum_{i<j}^{20}\beta_{ij}\,\omega_i\,\omega_j + \sum_{i<j<l}^{20}\beta_{ijl}\,\omega_i\,\omega_j\,\omega_l + \beta_{1234}\,\omega_1\,\omega_2\,\omega_3\,\omega_4, \tag{2.78}$$

in which:

- $\omega_i = 2\,(X_i - 1/2)$ except for $i = 3, 5, 7$ where $\omega_i = 2\left(1.2\dfrac{X_i}{X_i+1} - \dfrac{1}{2}\right)$,
- the input vector $\boldsymbol{X} = \{X_1, \ldots, X_{20}\}$ contains 20 uniform random variables $\{X_i \sim \mathcal{U}[0,1],\ i = 1, \ldots, 20\}$,
- $\beta_i = 20$ for $i = 1, 2, \ldots, 10$,
- $\beta_{ij} = -15$ for $i, j = 1, 2, \ldots, 6,\ i < j$,

- $\beta_{ijl} = -10$ for $i, j, l = 1, 2, \ldots, 5$, $i < j < l$,
- $\beta_{1234} = 5$,
- the remaining first and second order coefficients are defined by $\beta_i = (-1)^i$, $\beta_0 = 0$ and $\beta_{ij} = (-1)^{i+j}$,
- and the remaining third order coefficients are set to 0.

First, a PCE is built using the LARS-based sparse adaptive scheme using a Latin Hypercube experimental design of size $N_{PCE} = 500$. Then the PCE is post-processed to obtain the total Sobol' indices and the upper-bound derivative-based sensitivity measures (DGSMs) using Eq. (2.43) and Eq. (2.77), respectively. The procedure is replicated 100 times in order to provide the 95% confidence interval of the resulting sensitivity indices.

As a reference, the total Sobol' indices are computed by Monte Carlo simulation (MCS) using the `sensitivity` package in `R` (Pujol et al., 2013). One samples two experimental designs of size $N = 5,000$ denoted respectively by A and B then computes the corresponding output vectors Y_A and Y_B. To estimate the total sensitivity index S_i^T with respect to the random variable X_i, one replaces the entire i^{th} column in sample A (which contains the samples of X_i) by the i^{th} column in sample B to obtain a new experimental design denoted by C_i. Then the output Y_{C_i} is computed from the input C_i. The variance-based S_i^T is obtained by means of Y_A, Y_B and Y_{C_i} using the `sobol2007` function (Pujol et al., 2013; Saltelli et al., 2010). The total number of model evaluations required by the MCS approach is $5,000 \times (2 + 20) = 110,000$. After 100 replications we also obtain the 95% confidence interval on the sensitivity indices.

The DGSMs are also computed by MCS for comparison, using a finite difference scheme to evaluate the gradient for each realization. 500 realizations are used, leading to a total number of $500 \times (20 + 1) = 10,500$ model evaluations. Again the approach is replicated 100 times to obtain the confidence intervals.

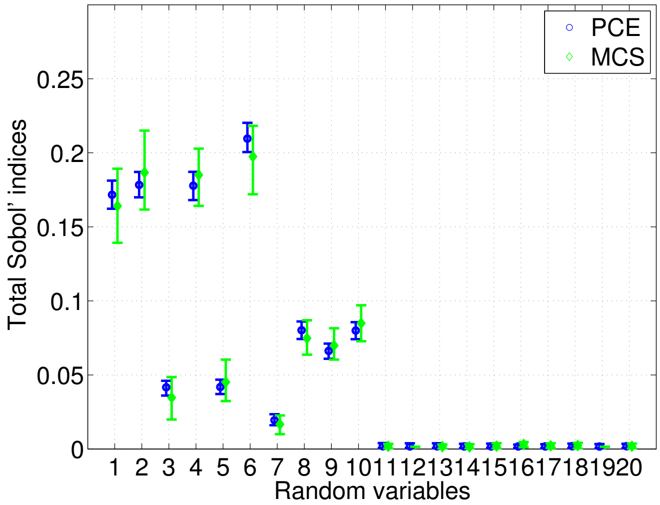


Figure 2.2 – Morris function: PCE-based *vs.* MCS-based total Sobol' indices

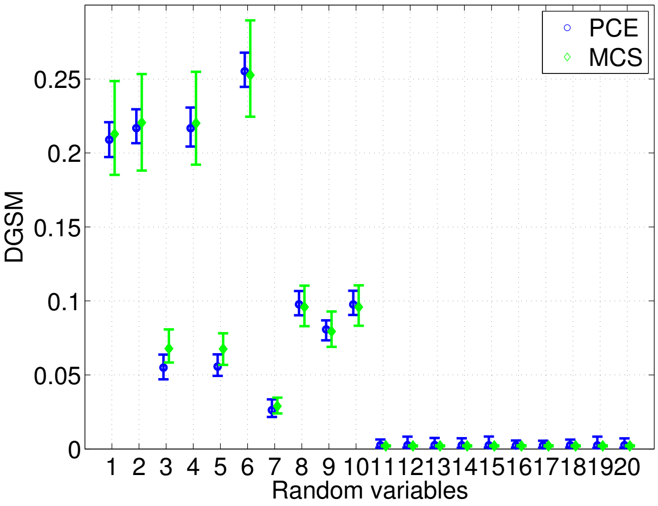


Figure 2.3 – Morris function: PCE-based *vs.* MCS-based derivative-based global sensitivity measures (DGSM)

Figure 2.2 shows the total Sobol' sensitivity indices computed by MCS and by PCE as well as their 95% confidence intervals. The median results (circles and diamonds) are close to each other, and show that input parameters $X_{11}, \ldots, X_{20}$ are unimportant factors, while X_1, X_2, X_4 and X_6 are substantial ones. It is observed that the confidence intervals are much smaller for the PCE-based indices than for the MCS-based indices, at a cost which is two order of magnitude smaller though (500 runs instead of 110,000).

Figure 2.3 shows the DGSMs computed by MCS and by PCE as well as their 95% confidence intervals. Again the results obtained by the two approaches compare very well to each other and it is observed that the confidence intervals are smaller when using PC expansions. By comparing Figures 2.2 and 2.3, one can verify that the obtained total Sobol' indices are always smaller than the DGSMs, as expected. Moreover, the less significant the parameter is, the closer the DGSM gets to the total Sobol' index.

2.4.5.2 Oakley & O'Hagan function

The second numerical application is the function proposed by Oakley and O'Hagan (2004) which reads:

$$f(\boldsymbol{X}) = \boldsymbol{a}_1^T \boldsymbol{X} + \boldsymbol{a}_2^T \cos(\boldsymbol{X}) + \boldsymbol{a}_3^T \sin(\boldsymbol{X}) + \boldsymbol{X}^T \boldsymbol{M} \boldsymbol{X}, \tag{2.79}$$

in which the input vector $\boldsymbol{X} = \{X_1, \ldots, X_{15}\}$ consists of 15 independent standard normal random variables $\{X_i \sim \mathcal{N}(0,1),\ i = 1, \ldots, 15\}$. The 15×1 vectors $\boldsymbol{a}_j$, $j = 1, 2, 3$ and the 15×15 matrix $\boldsymbol{M}$ are provided at `www.sheffield.ac.uk/st1jeo`.

The PCE-based approach is run with a Latin Hypercube experimental design of size $N_{PCE} = 600$. The size of a single sample set for the MCS approach is $N = 10,000$, resulting in $10,000 \times (2 + 15) = 170,000$ model runs. The procedure is similar as in Section 2.4.5.1.

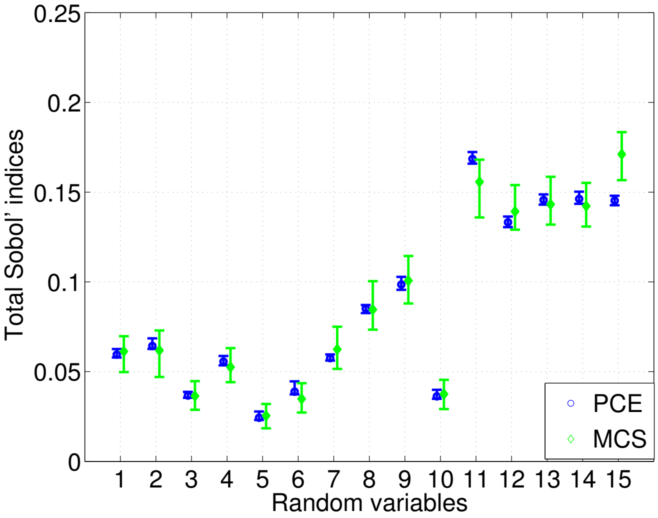


Figure 2.4 – Oakley & O'Hagan function: PCE-based *vs.* MCS-based total Sobol' indices

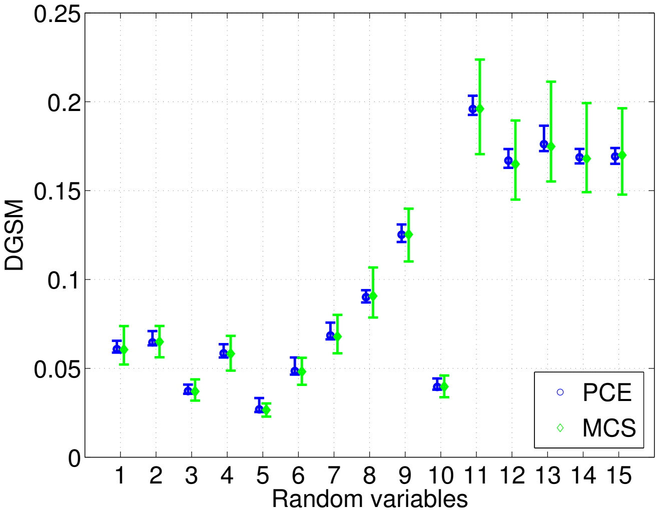
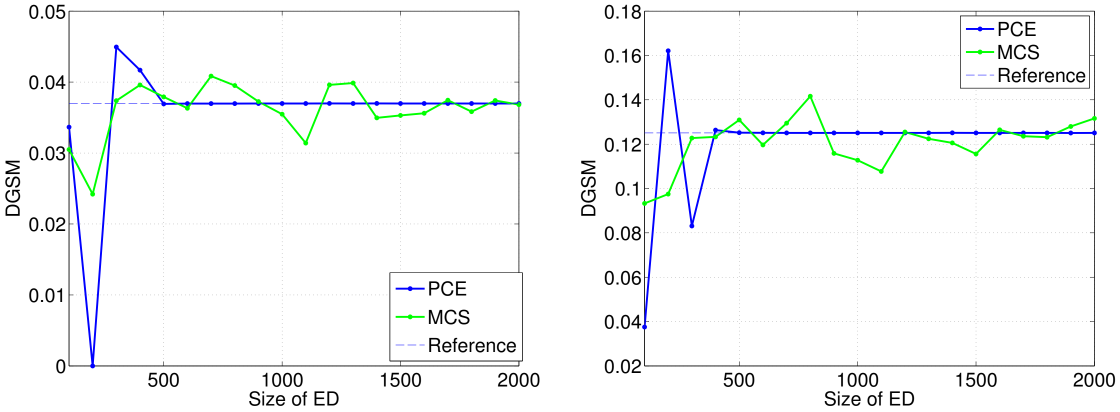


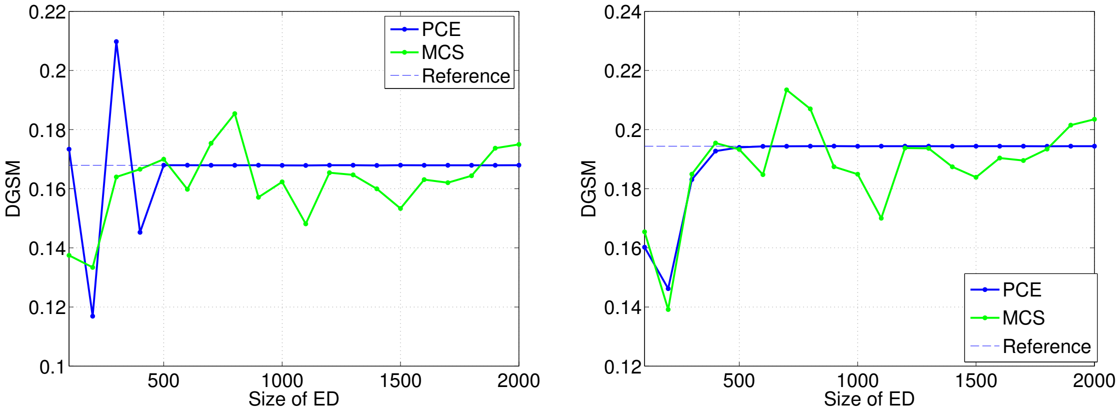
Figure 2.5 – Oakley & O'Hagan function: PCE-based *vs.* MCS-based derivative-based global sensitivity measures (DGSM)

Figures 2.4 and 2.5 show the total Sobol' indices and the DGSMs computed both from a PCE and by Monte Carlo simulation. The conclusions are similar to the ones already drawn from the first example: the median values of the PCE-based DGSMs are almost identical to the MCS-based estimators while the confidence intervals are much smaller. Here the computational cost is 600 runs for PCE against $600 \times (15 + 1) = 9,600$ for the MCS. From the values one can conclude that $X_{11}, \ldots, X_{15}$ are important parameters, whereas the other have medium to little importance.

In order to better assess the accuracy of polynomial chaos expansions as a tool for computing the DGSMs, we carry out a parametric study on the number of samples used in the analysis. Precisely, a Latin hypercube sample of size N is used as the experimental design for establishing the PCE, and as the set of points where the gradient is computed for the MCS-based approach (thus $2 \times N$ points are used for computing a single $DGSM$).



(a) X_3 (b) X_9



(c) X_{15} (d) X_{11}

Figure 2.6 – Oakley & O'Hagan function: convergence of the PCE-based (resp. MCS-based) DGSMs as a function of the number of runs (NB: abscissa is the size of the experimental design for PCE, whereas the actual number of runs for MCS is twice larger, for *each* DGSM).

The convergence plots are shown in Figure 2.6 for variables X_3, X_9, X_{15} and X_{11} which range from unimportant to most important. In each case the reference solution is attained using 500 runs or less using PCEs whereas the convergence is not attained even for $2 \times 2,000$ runs using MCS. Again it is emphasized that a *single experimental design* (*e.g.* of size 500) is used for computing all 15 DGSMs, whereas MCS requires $(15 + 1)$ times this number as a whole, which makes PCEs even more appealing in large dimensions.

2.4.6 Discussion

The scheme for computing PCE-based DGSMs faces certain limitations. First, only the cases with linear probabilistic transforms were considered. In case a *nonlinear* transform is required, for instance to convert a random variable of arbitrary distribution to a standard normal variable, the derivation of analytical formulations will not be straightforward as shown herein. The generalization requires further investigations. In addition, the DGSMs are always the upper bounds of the total Sobol' indices, which can also be computed using PCEs with no additional cost. The total Sobol' indices can yet indicate the importance of the input parameters. Therefore, obtaining the DGSMs only helps consolidate the conclusions on the relevance of each parameter.

However, a promising perspective for this approach should be underlined. More precisely, it is important to prove that the derivatives of PCEs can be obtained analytically and in particular can be represented as polynomial chaos expansions. This opens the path for different researches. For instance, it can be used in the context of the derivative-based optimization under uncertainties using PCEs, which deserves more attention in the future. In addition, it may be useful in the gradient-enhanced l_1-minimization scheme (Peng et al., 2016), in which the derivative information is computed to accelerate the identification of PC coefficients. Furthermore, it will be helpful in computing the gradient matrix in the iterative rotation approach aiming at enhancing the sparsity of PCEs (Yang et al., 2016). In general, our finding on the analytical formulations to represent derivatives of PCEs might serve researches involving the computation of gradients of the output quantity.

2.5 Summary

This chapter presented the theory of generalized polynomial chaos expansions. The aspects related to the use of PCEs in practice were outlined, *i.e.* the truncation schemes, the non-intrusive regression method for estimating the coefficients and the error estimators. The least angle regression-based scheme for computing sparse adaptive PCEs, which plays a significant role in the remaining part of this manuscript, was reviewed. Then, use of PCEs for estimating the output statistics was recalled. Finally, a contribution to the framework in which PCEs are used for computing analytically the derivative-based global sensitivity measures was introduced.

3

Literature review on polynomial chaos expansions for stochastic dynamics

The previous chapter presented the theory of generalized polynomial chaos expansions (PCEs) which allow the representation of random response quantities as functions of uncertain parameters of the considered systems. Note that the specified PC formulations are established only for the case of static systems, *i.e.* the random response at a specific location is time-independent. However, this is not the case in practice where the considered systems are commonly dynamical, *i.e.* the responses vary not only in the stochastic and spatial domain but also in time. Considering the stochastic dynamical problems has been a long-term challenge, in particular use of PCEs has been shown not effective by several works, see *e.g.* Wan and Karniadakis (2006); Beran et al. (2006b); Le Maître et al. (2010). The current chapter aims at summarizing and analysing the contributions that allow one to solve this problem.

The chapter is organized as follows: first, the theory of PCEs is applied to uncertain dynamical systems, *i.e.* to problems involving time-dependent random response quantities, which are encountered in various fields such as structural dynamics and fluid dynamics. The non-intrusive PCE-based approach commonly used to handle such problems is presented and applied to an illustrative numerical example. This allows one to underline the limitations of the classical approach and the reasons hindering its effectiveness are analyzed. Next, we present a review on various methods proposed in the literature which gives the big picture of the status quo. Finally, the approaches to be investigated by the author are clarified.

3.1 Problem set-up

3.1.1 Time-frozen polynomial chaos expansions

In this work, we focus on stochastic dynamical systems where the random response is a time-dependent quantity $y(t) = \mathcal{M}(\boldsymbol{\xi}, t)$. The latter is commonly the solution of a system of first-order ordinary differential equations (ODEs):

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{f}(\boldsymbol{y}, \boldsymbol{\xi}, t), \tag{3.1}$$

where $\boldsymbol{y}(t, \boldsymbol{\xi})$ is the time-dependent response of the system with the initial condition $\boldsymbol{y}(t = 0) = \boldsymbol{y}_0$ and the vector $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_M\}$ comprises M uncertain parameters of the system modelled by M independent second-order random variables $\Xi_i, i = 1, \ldots, M$ defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\boldsymbol{\xi}$ may include the parameters governing the system behaviour, *e.g.* masses, stiffness, damping ratio, reaction parameters. The initial condition can also be uncertain, in which case it becomes a random variable belonging to $\boldsymbol{\xi}$. Without loss of generality, we consider one component of the output quantity, *e.g.* $y(t, \boldsymbol{\xi})$ with the initial condition $y(t = 0) = y_0$. At each time instant, $y(t, \boldsymbol{\xi})$ is modelled by a second-order random variable. In this context, the polynomial chaos representation of the response is cast as:

$$y(t) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}}(t) \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) + \epsilon(t), \tag{3.2}$$

in which the notation $y_{\boldsymbol{\alpha}}(t)$ indicates time-dependent PCE coefficients and $\epsilon(t)$ is the residual at time t. The representation of a time-dependent quantity by means of PCEs as in Eq. (3.2), in which the effects of t and $\boldsymbol{\xi}$ are separated, is commonly used in the literature, see *e.g.* Wan and Karniadakis (2005); Pettit and Beran (2006); Le Maître et al. (2010); Gerritsma et al. (2010).

The time-dependent coefficients can be determined by means of intrusive or non-intrusive methods. Using the intrusive method, the truncated form of the PC representation is introduced in the original system of equations. Then one uses the properties of the orthonormality of the PC basis in order to extract a modified system of equations, which will be solved to obtain the time-dependent coefficients. The intrusive method is illustrated in the following by means of a numerical example.

Let us consider an undamped linear single-degree-of-freedom (SDOF) oscillator, whose equation of motion is an ODE as follows (Le Maître et al., 2010):

$$\ddot{y}(t, \xi) + k(\xi)\, y(t, \xi) = 0. \tag{3.3}$$

Herein $y(t, \xi)$ is the time-dependent displacement of the oscillator and $k(\xi)$ is the uncertain stiffness which is defined by $k = k_0 + k_1 \xi$ with $\xi \sim \mathcal{U}[-1, 1]$ being a uniform random variable. The truncated PC representation of the response reads:

$$y(t, \xi) = \sum_{i=1}^{P} y_i(t)\, \psi_i(\xi), \tag{3.4}$$

with $P = \mathrm{card}(\mathcal{A})$, where $\mathcal{A}$ is the set of multi-indices in the expansion in Eq. (3.2). Substituting Eq. (3.4) into Eq. (3.3), one obtains the approximate equation:

$$\sum_{i=1}^{P} \ddot{y}_i(t) \psi_i(\xi) + k(\xi) \sum_{i=1}^{P} y_i(t)\, \psi_i(\xi) = 0. \tag{3.5}$$

Taking the Galerkin projection of the above equation successively on each orthonormal polynomial $\psi_l(\xi), l = 1, \ldots, P$, one obtains the following deterministic system of P coupled ODEs:

$$\ddot{y}_l(t) + \sum_{i=1}^{P} y_l(t) \; < k(\xi)\, \psi_i(\xi)\,,\, \psi_l(\xi) >= 0, \qquad l = 1, \ldots, P, \tag{3.6}$$

in which $< \cdot\,,\, \cdot >$ is the inner product defined on the space of second-order random variables on $\mathbb{P}$. The P time-dependent PC coefficients are obtained by solving the system of equations (3.6). The intrusive method has been widely used in the literature. However, it is not usable when the original system of equations is not available, which is a common situation in practice when handling legacy computer codes. In addition, the intrusive solution is strongly problem-dependent, *i.e.* when considering a different problem, the Galerkin projection scheme must be adapted accordingly. Herein, it is observed that computing the deterministic coefficients $< k(\xi)\, \psi_i(\xi)\,,\, \psi_l(\xi) >, i, l = 1, \ldots, P$, in Eq. (3.6) is straightforward due to the form of the original ODE representing a simple linear oscillator. When another system with a different mechanism is of interest, the use of the Galerkin projection scheme might become troublesome. Let us make a minor modification to Eq. (3.3) so as to consider a Duffing oscillator with a cubic stiffness term:

$$\ddot{y}(t, \xi) + k(\xi)\, y^3(t, \xi) = 0. \tag{3.7}$$

The Galerkin projection scheme results in the following system of equations:

$$\ddot{y}_l(t) + < k(\xi) \left(\sum_{i=1}^{P} y_i(t)\, \psi_i(\xi) \right)^3 ,\, \psi_l(\xi) >= 0, \qquad l = 1, \ldots, P. \tag{3.8}$$

This final system of ODEs is still manageable, however is considerably more complicated than that defined by Eq. (3.6). The use of the intrusive method for problems involving complex dynamical behavior, *e.g.* an hysteretic restoring force, will certainly introduce issues in deriving a suitable system of equations. In general, the coupled equations obtained with the Galerkin projection scheme is significantly larger and more complicated than the original equations, but needs to be solved only once.

Despite the fact that the intrusive method has been considerably used in the literature, it is of utmost importance to introduce non-intrusive methods that do not rely explicitly on the system of equations describing the problems. Instead of manipulating the equations, non-intrusive methods only make use of a set of samples obtained by running the numerical model. The separated representation of the PCE in Eq. (3.2) makes it relatively simple to apply non-intrusive methods. At a given time instant t, using a sample set of the random parameters $\mathcal{X} = \left\{ \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)} \right\}$ and the corresponding responses $\mathcal{Y}(t) = \left\{ y(t, \boldsymbol{\xi}^{(1)}), \ldots, y(t, \boldsymbol{\xi}^{(N)}) \right\}$, the coefficients $\{ y_{\boldsymbol{\alpha}}(t), \boldsymbol{\alpha} \in \mathcal{A} \}$ and the accuracy of the PCEs may be estimated with different techniques presented in Chapter 2, namely stochastic collocation, projection, regression methods. In such an approach, the metamodel of the response would be computed *independently* at each time instant, hence the name *time-frozen* PCEs.

Once the PC coefficients are determined, Eq. (3.2) can be used to compute the evolution of the response statistics. The multivariate polynomial chaos functions are orthonormal, *i.e.* :

$$\mathbb{E}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\,\boldsymbol{\psi}_{\boldsymbol{\beta}}(\boldsymbol{\xi})\right] \overset{\text{def}}{=} \int\limits_{\mathcal{D}_{\boldsymbol{\Xi}}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\,\boldsymbol{\psi}_{\boldsymbol{\beta}}(\boldsymbol{\xi})\,f_{\boldsymbol{\Xi}}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}} \quad \forall\boldsymbol{\alpha},\,\boldsymbol{\beta}\in\mathbb{N}^M, \tag{3.9}$$

in which $\delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ is the Kronecker symbol that is equal to 1 if $\boldsymbol{\alpha}=\boldsymbol{\beta}$ and equal to 0 otherwise. In particular, each multivariate polynomial is orthonormal to $\boldsymbol{\psi}_{\mathbf{0}}(\boldsymbol{\xi}) = 1$, which leads to $\mathbb{E}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\right] = 0\ \forall\boldsymbol{\alpha}\neq\mathbf{0}$ and $\mathrm{Var}\left[\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\right] = \mathbb{E}\left[\left(\boldsymbol{\psi}^2_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\right)\right] = 1\ \forall\boldsymbol{\alpha}\neq\mathbf{0}$. Thus, the time-dependent mean and standard deviation of the response can be estimated with no additional cost by post-processing the truncated PC coefficients in Eq. (3.2) as follows:

$$\hat{\mu}_{Y(t)} \overset{\text{def}}{=} \mathbb{E}\left[\sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}(t)\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\right] = y_{\mathbf{0}}(t)\ , \tag{3.10}$$

$$\hat{\sigma}^2_{Y(t)} \overset{\text{def}}{=} \mathrm{Var}\left[\sum_{\boldsymbol{\alpha}\in\mathcal{A}} y_{\boldsymbol{\alpha}}(t)\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\right] = \sum_{\substack{\boldsymbol{\alpha}\in\mathcal{A}\\ \boldsymbol{\alpha}\neq\mathbf{0}}} y^2_{\boldsymbol{\alpha}}(t). \tag{3.11}$$

3.1.2 Failure of time-frozen polynomial chaos expansions

To illustrate the use of time-frozen PCEs, let us consider a simple numerical example which is the free vibration of a linear undamped oscillator (Eq. (3.3)). Note that Le Maître et al. (2010) also investigated this example to illustrate the limitation of the intrusive approach. Herein $y(t,\xi)$ is the time-dependent displacement of the oscillator and $k(\xi)$ is the uncertain stiffness of the system which is defined by $k = k_0 + k_1\,\xi$ with $\xi \sim \mathcal{U}[-1,1]$ being a uniform random variable, $k_0 = (2\,\pi)^2$ and $k_1 = 0.2\,k_0$. The initial conditions are deterministic with $y(t=0)=1$ and $\dot{y}(t=0)=0$. The exact solution of the problem is $y(t,\xi) = \cos(\sqrt{k(\xi)}\,t)$ (Le Maître et al., 2010).

Figure 3.1 depicts the displacement of the oscillator for three different values of ξ at the early instants $t < 10$ s. The effect of uncertainty in this parameter is pronounced, when the resulting solution trajectories differ significantly. It is seen that the differences between trajectories tend to increase as time progresses.

The non-intrusive time-frozen sparse PCE[1] approach is now applied to model the displacement $y(t,\xi)$ of the oscillator at every 0.5 s until 50 s, *i.e.* at $t = 0, 0.5, 1, \ldots, 50$ s. First $N = 500$ numerical simulations are conducted. At each considered time instant t, the set of parameter samples $\mathcal{X} = \{\xi_1, \ldots, \xi_N\}$ and the corresponding displacements $\mathcal{Y}(t) = \{y(t,\xi_1), \ldots, y(t,\xi_N)\}$ are used as the experimental design for the metamodel. In order to judge the effect of the polynomial degree on the accuracy of the method, PCEs of degree up to 10 and 30 are computed.

---

[1] The non-intrusive time-frozen sparse PCE approach refers to the instantaneous computation of PCEs with the LARS-based sparse adaptive scheme presented in Section 2.2. For the sake of simplicity, it will be denoted by time-frozen PCE throughout the manuscript.

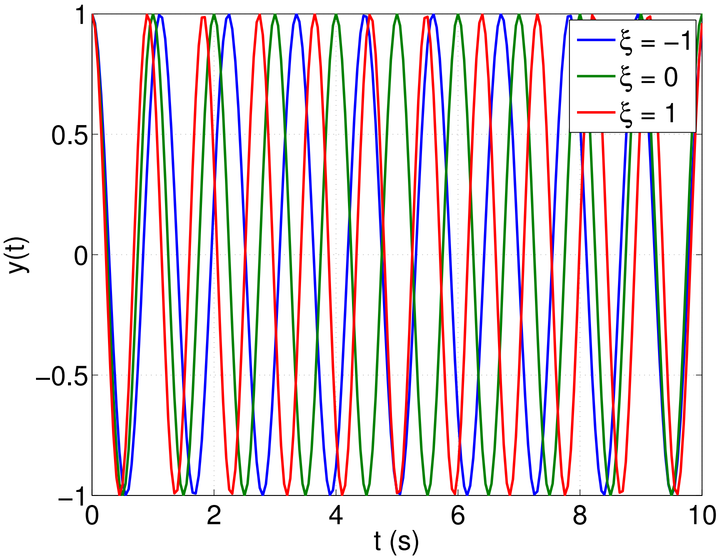


Figure 3.1 – Linear oscillator – Three solutions to different values of ξ .

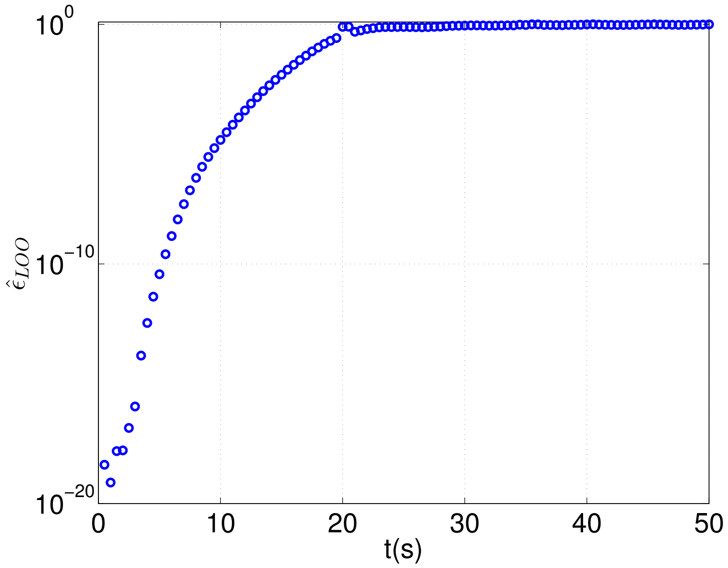
Figure 3.2 depicts the relative LOO error of time-frozen PCEs as a function of time t. When the maximal polynomial degree is 10, the LOO error exhibits a fast growth which starts right after the initial instant $t = 0$. At $t = 20$ s, the LOO error is about 1. Using higher maximal polynomial degree ($p = 30$), PCE remains highly accurate until $t = 20$ s, at which point the accuracy starts degenerating at a rate which is slightly slower than for PCEs with $p = 10$. Note that at certain instants, the resulting PCEs of maximal degree $p = 30$ are highly inaccurate whereas PCEs at neighbouring instants are of acceptable quality. This might be a sign of the instability that PCEs exhibit, which is due to the limited number of samples available. The conclusion is that by increasing the polynomial degree, the accuracy of PCEs can be improved. However, this only allows to delay the onset of the accuracy degeneration. At a certain point, high-degree PCEs will start loosing their accuracy, though at a slightly smaller rate compared to low-degree PCEs.

Let us now use the resulting PCEs to predict the time-dependent statistics of the response. Given the exact solution $y(t,\xi) = \cos(\sqrt{k(\xi)}\, t)$, the mean value of the response at instant t can be obtained as:

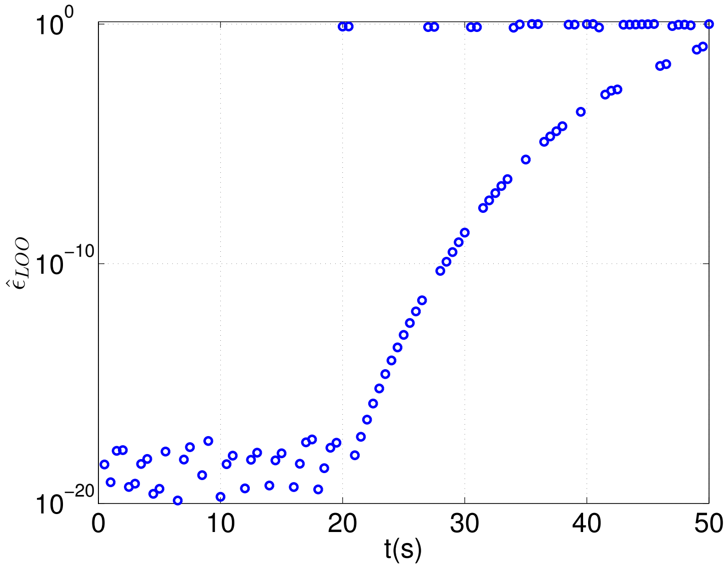
$$\mathbb{E}\left[y(t,\xi)\right] = \int_{-1}^{1} y(t,\xi)\, \frac{1}{2}\, \mathrm{d}\xi \tag{3.12}$$

The standard deviation of the response can be obtained approximately using the Taylor expansion for the moments of function of random variables (Benaroya et al., 2005). However the approximation might be prone to large inaccuracy due to the non-linear relationship between ξ and $y(t,\xi)$. Therefore, the empirical standard deviation obtained from 10^4 Monte Carlo samples is used herein as reference.

Figure 3.3 shows that PCEs of degree 10 are capable of predicting accurately the expected values of the responses at considered time instants. Regarding the standard



(a) Maximal polynomial degree $p = 10$

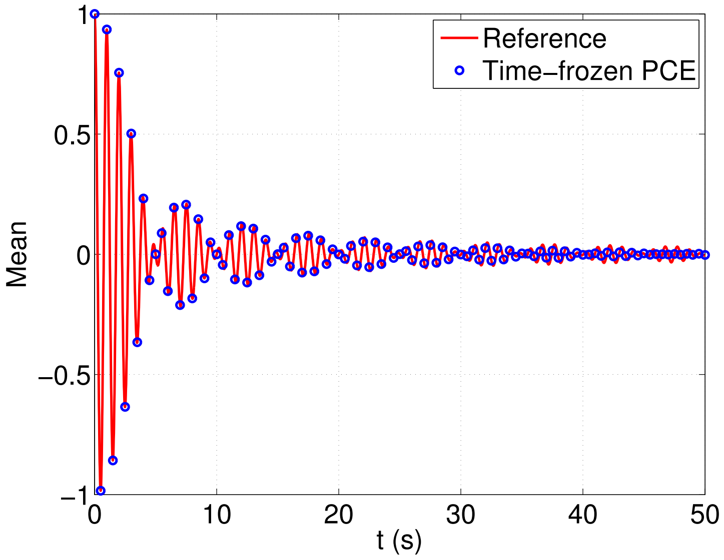


(b) Maximal polynomial degree $p = 30$

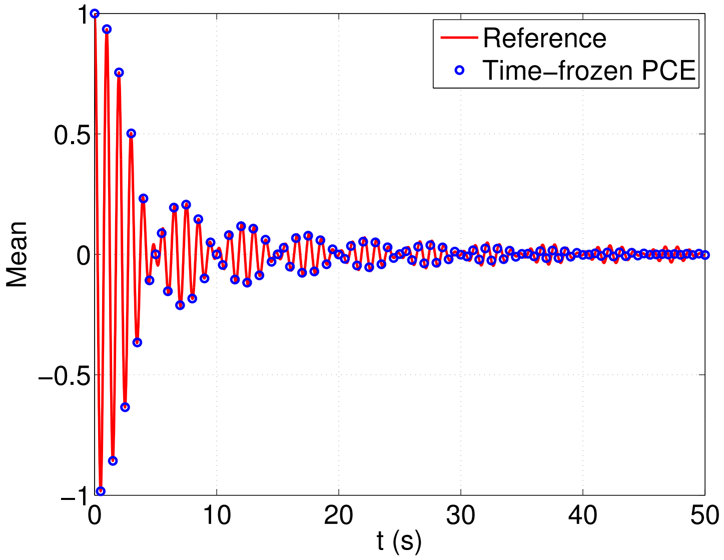
Figure 3.2 – Linear oscillator – Evolution of the LOO error in time t .

deviation (Figure 3.4), degree-10 PCEs fail to capture the trend of the trajectory from $t = 20$ s onwards. Using polynomials of degree 30, the standard deviation can be predicted relatively well except for the later instants ($t > 35$ s).

The observed results suggest that non-intrusive time-frozen PCEs are not effective even for the considered simple example concerning a linear oscillator with only one uncertain parameter and without external excitation or damping. Le Maître et al. (2010) showed that the intrusive method faces the same problem, when its accuracy degenerates in time. Therefore it is extremely difficult to apply PCEs to more complicated time-dependent problems, which has been reported in multiple publications, see *e.g.* Wan and Karniadakis (2005); Gerritsma et al. (2010).

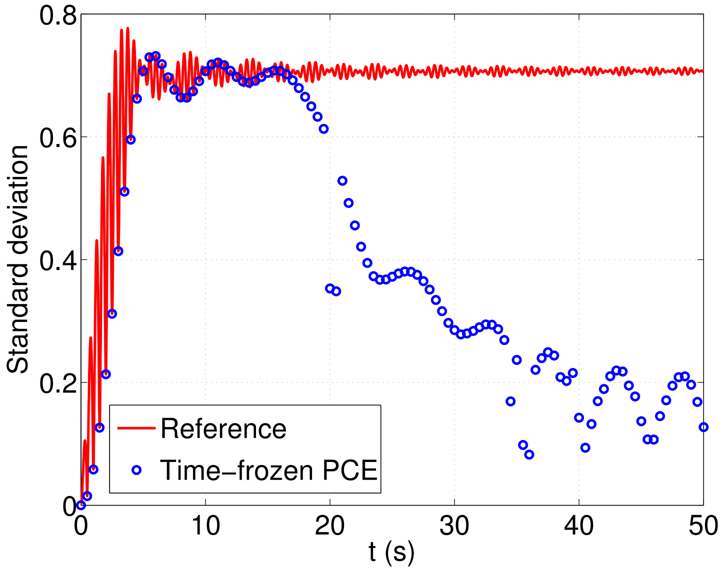


(a) Maximal polynomial degree $p = 10$

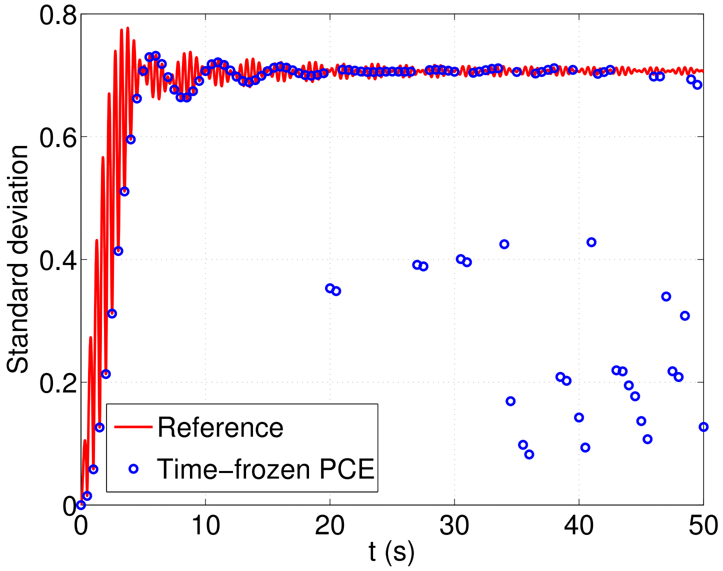


(b) Maximal polynomial degree $p = 30$

Figure 3.3 – Linear oscillator – Evolution of the response mean in time t .



(a) Maximal polynomial degree $p = 10$



(b) Maximal polynomial degree $p = 30$

Figure 3.4 – Linear oscillator – Evolution of the response standard deviation in time t .

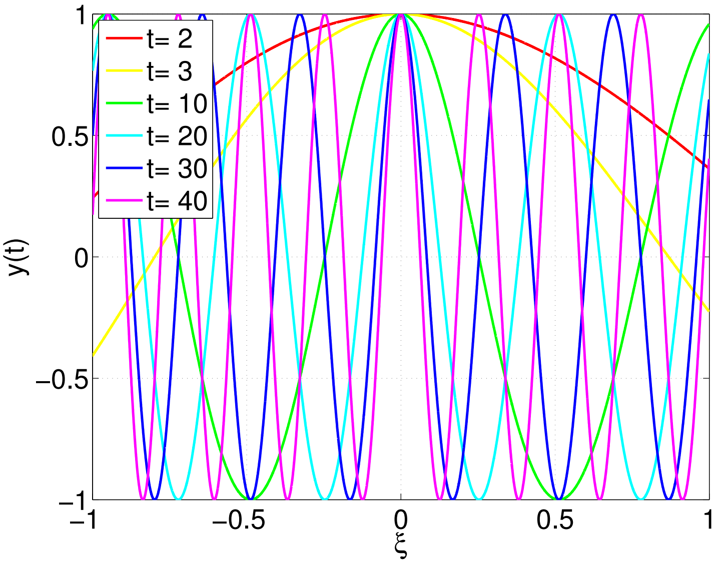
3.1.3 Why time-frozen polynomial chaos expansions fail

The current section aims at clarifying the reasons that cause time-frozen PCEs failure in representing time-dependent random responses.

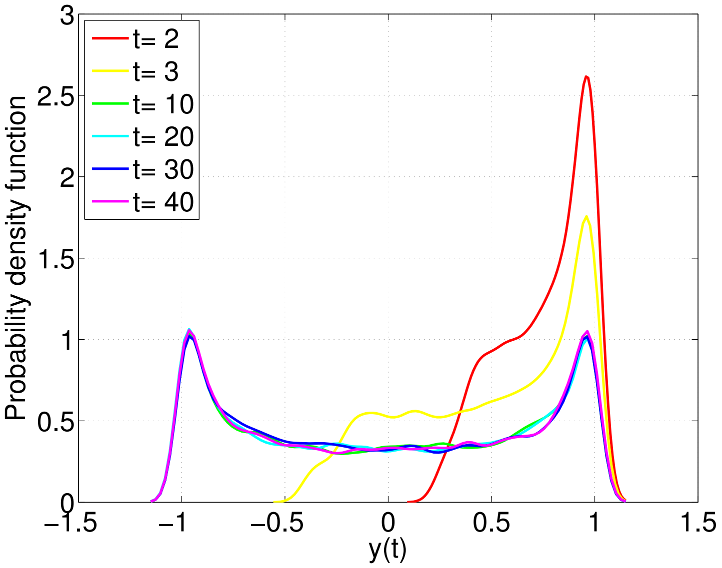
In the considered linear oscillator system, the relationship between the input parameter ξ and the solution becomes increasingly non-linear as time progresses. The increasingly non-linear dependence of the solution on the input parameters is also reported in different publications, see *e.g.* Wan and Karniadakis (2005); Pettit and Beran (2006). Figure 3.5 depicts the response $y(t, \xi)$ plotted as a function of the uncertain parameter ξ at different time instants $t = 2, 3, 10, 20, 30, 40$ s. At $t = 2, 3$ s, the response is a parabolic-like function of ξ. At $t = 10$ s, the function becomes sinusoidal-like in the considered range of ξ. The frequency of the sinusoidal-like function is increasing with time. This is consistent with the exact solution of the displacement $y(t, \xi) = \cos(\sqrt{k(\xi)}\, t)$, in which the response at the instant t is a cosine function of $\sqrt{k_0 + k_1\, \xi}$ with frequency $t/(2\,\pi)$.

The increasing non-linearity of the solution with respect to the parameter leads to the following by-products. First, the probability density function (PDF) of $y(t, \xi)$ becomes bi-modal at late instants, whereas they are uni-modal at the early stage. This multi-modality of the response can be seen as an indicator that approximation by means of time-frozen PCEs is becoming harder in time. Representation of multi-modal quantities using PCEs was investigated by Nouy (2010) and Soize (2015). Second, the expansions containing polynomials suitable for representing the response at the early instants become insufficient at later instants, as reported by Ghosh and Ghanem (2008); Gerritsma et al. (2010). In order to maintain the accuracy, the expansion needs to be updated with higher-degree polynomials or different types of basis functions that can represent well the non-linearity, see *e.g.* Wan and Karniadakis (2006); Pettit and Beran (2006); Paffrath and Wever (2007); Ghosh and Ghanem (2008).

At this point, we should clarify why the response becomes an increasingly non-linear function of the uncertain parameter as time progresses. The question itself contains a hint for the answer. Indeed, this phenomenon happens only when the temporal variable t intervenes in the systems. On the contrary, in static problems where the systems remain in the equilibrium state, the input-response function does not evolve, *i.e.* it remains the same at all instants. Therefore the evolution of the input-response function is uniquely due to the dynamics of the considered systems. For instance, in the linear oscillator under investigation, the response is a periodic function of time t with uncertain frequency. For the same amplitude of the response, its dynamics are governed solely by its frequency. The distinct values of the uncertain parameter result in out-of-phase trajectories. Consequently, the input-response function exhibits an increasing level of non-linearity. In more complicated problems, the dynamics is governed by more complex mechanism involving the solution history (or memory), as shown in the next section of literature review. Therefore, use of PCEs for capturing the increasingly non-linear relationship consists in capturing the *consequence* of the dynamics. The question is whether it is suitable to catch the dynamics by using larger and more complicated bases. PCEs is a tool for propagating uncertainties and is not a tool designed to represent the dynamics. It is therefore more effective to use



(a) Time dependent relationship between $y(t, \xi)$ and ξ



(b) Time dependent probability density function of $y(t, \xi)$

Figure 3.5 – Linear oscillator – Evolution of the response in time t .

PCEs only for the uncertainties-related part of the problem while capturing the dynamics with a different tool designed for this specific purpose.

Let us now discuss another aspect of the problem relating to the fundamental formula of PCEs for time-dependent problems, which consists in separating the effects of time t and the uncertain parameters. For the considered linear oscillator, the PCE reads:

$$y(t, \xi) = \sum_{i=1}^{P} \psi_i(\xi)\, y_i(t). \tag{3.13}$$

The representation is rewritten this way in order to emphasize the fact that from a different point of view $y_i(t)$ (resp. $\psi_i(\xi)$) can be considered basis functions (resp. coefficients). This represents the time-dependent stochastic response $y(t, \xi)$ as a linear combination of *deterministic* modes $y_i(t), i = 1, \ldots, P$ with corresponding stochastic coefficients $\psi_i(\xi), i = 1, \ldots, P$. This follows the principle of superposition that is unfortunately not valid for non-linear dynamics. In other words, for problems involving non-linearity, the classical representation of PCEs for time-dependent problems is merely a crude approximation of the dynamics. It has also been observed that even for systems with linear behaviour (*e.g.* the oscillator considered herein), this representation of PCEs is not effective (Le Maître et al., 2010). Therefore, a different representation in which the effects of time t and $\boldsymbol{\xi}$ are not separated might be of interest.

3.2 Current polynomial chaos expansion approaches for stochastic dynamics

This section includes an extensive literature review of the state-of-the-art methods for PCEs in the context of time-dependent problems, *e.g.* in structural and fluid dynamics and chemical systems. The different methods are classified into two major groups in accordance with the philosophy of how to handle the problems. For each method, the main features, differences and similarities with respect to the other methods, its effectiveness and limitations will be discussed.

There are different ways of classifying the methods presented hereinafter, for instance intrusive *vs.* non-intrusive PCEs, global *vs.* local PCEs, etc. Herein, it is proposed to classify the methods into two groups according to the guiding principles, which are explained in the next paragraph.

Uncertain dynamical system involves two aspects, namely uncertainty and dynamics. As explained in the previous section, the dynamics of the system results in the increasingly non-linear dependence of the response on the stochastic input parameters as time progresses. There are so far two groups of methods for handling this problem. The first one refers to methods that use PCEs to capture directly the increasing non-linearity, which is the consequence of the dynamics. The second group includes methods that directly capture the dynamics of the systems with a specific technique while representing the effect of uncertainties with PCEs, thus avoiding modelling the strong non-linearity with the uncertainty propagation tool. For the sake of clarity, the evolution of PCE methods for uncertain dynamics in the two groups is depicted in Figure 3.6.

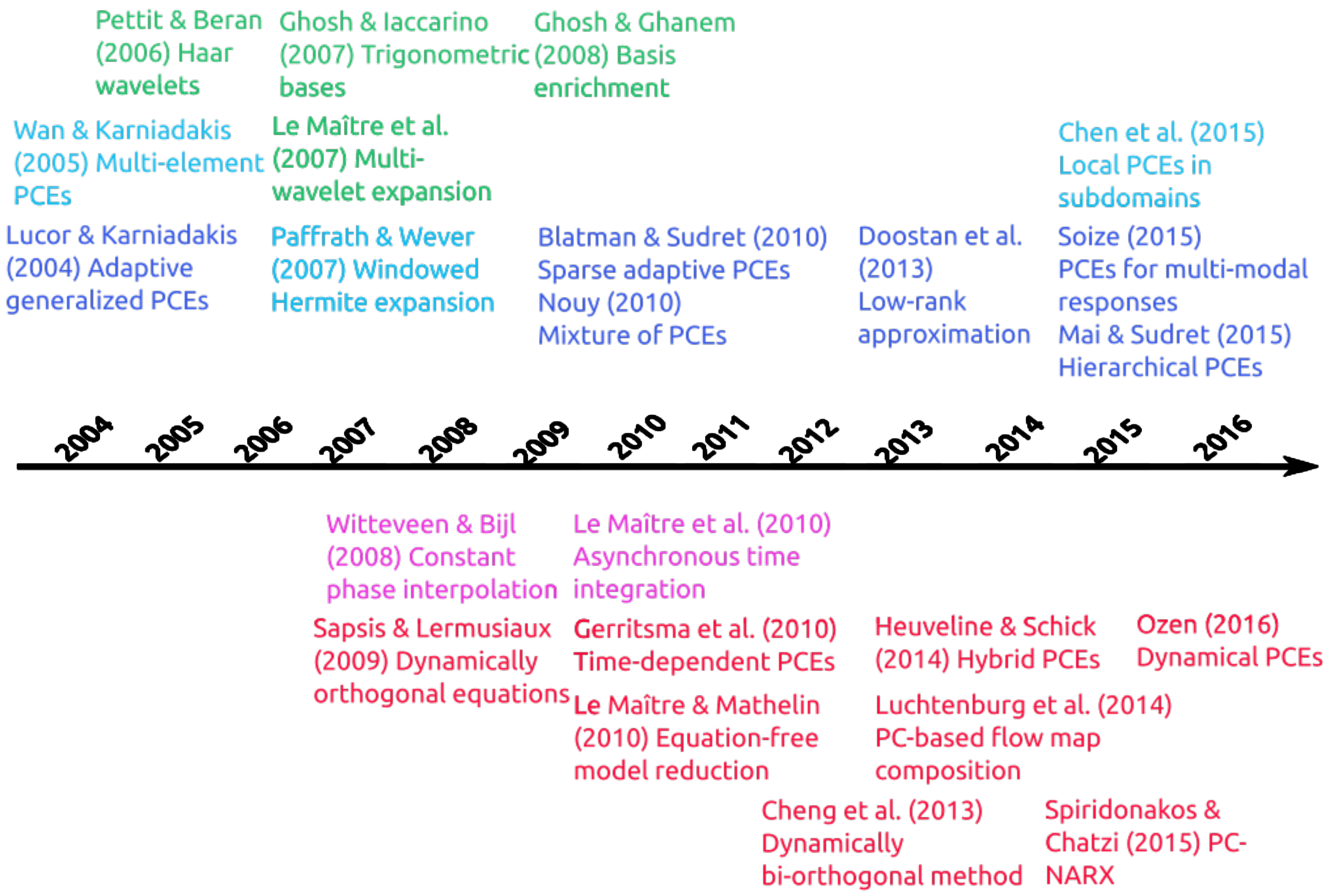


Figure 3.6 – PCE approaches for uncertain dynamics.

3.2.1 Approach 1: capture the consequence of the dynamics

The so-called consequence of the dynamics is the increasingly non-linear dependence of the response on the random input parameters. This has been currently handled by means of global high-degree PCEs, local low-degree PCEs or different types of basis functions.

3.2.1.1 Use of high-degree global functions

To represent a highly non-linear function, the use of high-degree PCEs is the first solution that one commonly investigates. For instance, Li and Ghanem (1998) and Lucor and Karniadakis (2004) suggested that strong non-linear dynamics can only be captured accurately using high polynomial degree.

The most relevant issue for high-degree PCEs is the curse of dimensionality, which means the cardinality of the PC basis increases exponentially with the number of random variables and the total degree. Meanwhile, usually a limited number of model simulations is available. In the literature, different methods have been designed to break the curse of dimensionality for high-degree and high-dimensional PCEs that can be used in the context of time-dependent problems.

Lucor and Karniadakis (2004) proposed an adaptive generalized PCE approach which reads:

$$y(t,\boldsymbol{\xi}) = \bar{y}(t) + \sum_{i=1}^{K} y_i(t)\,\xi_i + \sum_{i=K+1}^{M} y_i(t)\,\xi_i + \sum_{j=M+1}^{P} y_j(t)\,\boldsymbol{\psi}_j(\xi_i|_{i=1}^{K}), \tag{3.14}$$

where $\bar{y}(t)$ is the mean trajectory, the first two summations include the linear terms and the third summation comprises the non-linear terms of the expansion which are polynomials of random variables $\{\xi_i, i = 1, \dots, K\}$. The intrusive Galerkin projection method is used to compute the time-dependent PC coefficients $y_i(t), i = 1, \dots, P$. The linear terms are sorted and reordered according to the L_2-norm of the time-series $y_i(t), i = 1, \dots, M$ over the considered duration. Only the non-linear terms with contribution from K input variables with the largest L_2-norm of $y_i(t)$ are kept in the next iteration. During each iteration, all the PC coefficients are recomputed and the linear terms are reordered based on the resulting L_2-norm. The process is repeated until convergence is reached, *i.e.* the ordering of the K largest linear contributors is stable. The idea is to adaptively select the non-linear interaction polynomials, which actually allows one to reduce the cardinality of the PC basis. The method was applied to a Duffing oscillator with one uncertain parameter subject to a sinusoidal excitation with random amplitude (Lucor and Karniadakis, 2004).

To break the curse of dimensionality, Blatman and Sudret (2010) introduced truncation schemes that allow one to reduce significantly the size of the candidate basis. Furthermore, the authors used the least angle regression technique to effectively select the relevant PC functions. The resulting sparse adaptive PCE scheme was presented earlier in chapter 2 of the current manuscript and can be used to enhance the accuracy at late instants. From a similar perspective, different methods have been proposed to exploit the sparsity of PCEs, *e.g.* l_1-minimization technique (Jakeman et al., 2015; Hampton and Doostan, 2015).

Instead of enhancing the sparsity of the PC expansions, Doostan et al. (2013) introduced the use of low rank approximation (LRA) which relies on the canonical decomposition of the response. The method is based on solving a sequence of one-dimensional least squares regression problems. When considering one dimension, all variables in the other directions are frozen. This is alternated over all directions. The alternating approach allows the computation of high-degree and high-dimensional PCEs, which appears promising and might be applied to time-dependent problems.

As observed in the previous section, the probability density function of the response is evolving and becomes multimodal at late instants of the response. Nouy (2010) proposed the use of mixtures of PCEs to represent the multimodal distribution. Soize (2015) also focused on the representation of a multimodal response quantity. It can be expected that the proposed methods will help improving the accuracy of PCEs when applied to time-dependent problems.

3.2.1.2 Use of low-degree local functions

From a different perspective, Wan and Karniadakis (2005) realized that increasing the polynomial degree might be ineffective for long term responses. This is due to the use of global polynomial functions defined over the entire random space. When a highly non-linear relationship is of interest, a large number of global basis functions is required. However, over each sub-domain of the random space, the non-linearity is significantly lower and the degree of perturbation is scaled down; thus low-degree local basis functions can be used effectively. Based on this observation, Wan and Karniadakis (2005) proposed an adaptive multi-element PC approach. The space of random inputs is decomposed into non-overlapping sub-elements. In each sub-element, new local random variables are defined in accordance with their distribution functions. Finally, low-degree PCEs are applied element-wise. In order to maintain the accuracy, an adaptive decomposition of the random space is required. At the next time instant, an element is divided into two equal sub-elements or kept intact depending on the local decay rate of the relative error of the PC approximation on that element. An indicator is proposed to select the most sensitive random dimension to be decomposed. This method was applied to the Kraichnan-Orsag problem with three random variables representing the initial conditions of the responses.

Sharing a similar point of view, Paffrath and Wever (2007) used windowed Hermite expansion which consists in constructing conditional polynomials in small sub-elements in the space of random inputs. This allows one to zoom into the critical region in the random space, in this case the neighbourhood of the point with the highest probability of failure. The method is used for computing the time-dependent failure probability in systems of ODEs such as the predator-prey model (three random variables) and the Belousov-Zhabotinsky reaction (one random variable).

Recently, Chen et al. (2015) proposed to use local PCEs in non-overlapping subdomains which are obtained by dividing the spatial space and imposing artificial interface boundary conditions. The local solutions are obtained by using PCEs of local random variables and auxiliary variables governing the artificial boundary conditions.

As a summary, the methods presented in this section rely on decomposing the problems into several sub-problems either in the random or in the spatial space, which can be solved independently. The sub-problems are less complex than the original problem with respect to the dependence of the solution on the input parameters, thus local PCEs can be used effectively.

3.2.1.3 Use of different basis functions

In order to capture the non-smooth behaviour of the response as a function of random parameters, Ghosh and Ghanem (2008) introduced the basis enrichment of PCEs. By integrating a priori knowledge about the system and its non-smooth behaviour, the PC basis is updated with appropriate enrichment basis functions, namely absolute, step and inverse function etc. Although the orthogonality is not maintained for the enriched basis, this method allows one to capture the highly non-linear behaviour (including discontinuity) in the considered problems, namely the free vibration of a simply supported rectangular plate (two random variables) and a 2 DOF linear system with uncertain stiffness of the linear spring.

Another approach consists in substituting the full original polynomial bases. Ghosh and Iaccarino (2007) showed that trigonometric (sinus and cosine) functions perform better than PCEs for a simple problem of CO surface oxidation. The trigonometric expansion predicts the statistics of the response more accurately than PCEs. Pettit and Beran (2006) used Haar wavelets of random variables which are based on a piecewise-constant mother wavelet to represent the responses of limit cycle oscillations. Le Maître et al. (2007) proposed to substitute the PC representation by a multi-resolution scheme based on multi-wavelet expansion consisting of the piece-wise continuous polynomial functions (multi-wavelets) obtained by translations and dilations of the normalized Hermite polynomials. The multi-resolution scheme allows the local refinement of the expansion in appropriate areas of the random space.

The presented methods share the common feature that the steep, non-smooth dependence is captured by functions different from polynomials. However, the use of different basis functions, for instance compact polynomial multi-wavelets, has a much higher computational cost than PCEs (Le Maître et al., 2007).

3.2.2 Approach 2: capture the dynamics

The second approach relies on capturing the dynamics by specific tools so that low-degree PCEs can be used effectively to represent the uncertainties. Special techniques can be used for systems with particular characteristics, *e.g.* periodicity. In the following, we present different techniques that have been investigated to capture the dynamics.

3.2.2.1 Methods for periodic random responses

We first consider the problems involving periodic random responses, which are relevant in various fields of engineering. Examples include the steady-state response of systems subject to sinusoidal excitation (Lucor and Karniadakis, 2004), the limit cycle oscillations of aircraft wing in a transonic regime (Bunton and Denegri, 2000; Denegri, 2000) and the response of self-oscillation systems (Sarrouy et al., 2013). Uncertainty quantification of periodic responses has attracted a large attention, see *e.g.* Beran et al. (2006a); Le Meitour et al. (2010); Schick et al. (2014).

Many of the methods presented in Section 3.2.1 were illustrated with dynamical systems involving periodicity, however they are based on the ideas of using high-degree PCEs, or local basis functions or functions different than polynomials. For instance, Le Meitour et al. (2010) used the multi-element approach for the limit cycle oscillation of an air-foil, whereas Beran et al. (2006a) used Haar wavelets. Hereafter, two methods specifically designed for limit cycle oscillations are discussed.

Witteveen and Bijl (2008) proposed to use a non-intrusive constant phase interpolation which consists in interpolating the responses onto their phase space. The phase of the vibrations is extracted by measuring the local extrema in the sampled response trajectories. In the phase space, the trajectories become actually in-phase, thus use of PCEs is more effective. A further step to transform the response from the phase space back to the time space is required.

Le Maître et al. (2010) proposed the intrusive asynchronous time integration scheme, in which the transformed time variable $\tau(t, \boldsymbol{\xi})$ is introduced. $\tau(t, \boldsymbol{\xi})$ is adjusted in such a way that the spectrum of the solution in the transformed time scale remains narrow-banded. In other words, the transformed trajectories vary in a small neighbourhood of a reference trajectory. Finally, one needs to solve a new set of ODEs involving the transformed time $\tau(t, \boldsymbol{\xi})$ and the transformed trajectory $y(\tau, \boldsymbol{\xi})$.

The above two methods follow the same principle that the representation of the solution in the normal time space is too complex for using PCEs, thus a projection of the solution on a suitable space is sought so that it is more convenient for the PC representation. In particular, in the selected space (phase or transformed time), despite the uncertainties from the input parameters, the response trajectories exhibit a low level of variability by remaining in-phase. Indeed, the uncertain frequencies of the solution, which in this case characterize the dynamics of the oscillation, are captured by the intrusive time transform operator or the phase interpolation scheme.

3.2.2.2 Methods for general cases

In a general case, the responses are not periodic functions, thus a general approach must be used.

Gerritsma et al. (2010) observed that the PDF of the solution evolves in time, which makes the polynomials associated with initial distributions no longer optimal for representing long-term responses. Therefore, an adaptation of the set of uncertain variables is needed. Time-dependent PCEs are constructed by adding new random variables on-the-fly. At instant t, PCEs read:

$$y(t,\boldsymbol{\xi}) = \sum_{i=1}^{P} y_i(t)\,\boldsymbol{\psi}_i(\boldsymbol{\xi}). \tag{3.15}$$

When the non-linear coefficients become too large with respect to the linear coefficients, a new random variable equal to the solution at that time instant is added. At a later instant $t' = t + \Delta t$, time-dependent PCEs read:

$$y(t',\boldsymbol{\xi}) = \sum_{j=1}^{P'} y_j(t')\,\boldsymbol{\psi}_j(\boldsymbol{\xi},\xi^*) = \sum_{j=1}^{P'} \left[y_j(t')\,\psi_j^{(M+1)}(\xi^*)\right] \prod_{k=1}^{M} \psi_j^{(k)}(\xi_k), \tag{3.16}$$

in which $\xi^* \equiv y(t,\boldsymbol{\xi})$ is the newly added random variable and $\psi_j^{(k)}(\cdot)$ is the univariate polynomial in the direction k of the multivariate polynomial $\boldsymbol{\psi}_j(\cdot)$. Time-dependent PCEs represent the response at the considered instant as polynomial chaos functions of the responses at early instants and original random variables. This is a major difference compared to the classical representation of PCEs, where the dynamics of PC coefficients are not evident, *i.e.* no connection between the PC coefficient $y_j(t')$ and the past values of the response is specified. The classical representation led to the idea that PC coefficients can be determined instant-wise in a non-intrusive manner. One can argue that in the intrusive method, the PC coefficients are described by ordinary differential equations, thus they exhibit the dynamics. However, it is reminded that even the intrusive PCEs fail in relatively simple problems, which might be due to the separated form of the representation. In time-dependent PCEs, the PC-coefficients are functions of the responses past values (ξ^*), which are themselves functions of $\boldsymbol{\xi}$. The clear distinction between the effects of the temporal variable t and the uncertain parameters $\boldsymbol{\xi}$ disappears, when $\boldsymbol{\xi}$ governs indirectly the PC-coefficients $y_j(t')\,\psi_j^{(M+1)}(\xi^*)$ in Eq. (3.16). Time-dependent PCE were used to solve the Kraichnan-Orszag problem with three random parameters. Recently, Heuveline and Schick (2014) combined time-dependent PCEs with the multi-element approach. The time-dependent PCEs are applied in each sub-domain in the random space. The hybrid approach allows to take advantages of both methods.

From a similar perspective, Luchtenburg et al. (2014) proposed the PC-based flow map composition, in which a short term flow map is the solution of the original system of ODEs in small time window. The long-term response can be obtained by the composition of intermediate flow maps:

$$\varphi_0^{N\,\Delta t} = \varphi_{(M-1)\,\Delta t}^{M\,\Delta t} \circ \varphi_{(M-2)\,\Delta t}^{(M-1)\,\Delta t} \circ \cdots \circ \varphi_{\Delta t}^{2\,\Delta t} \circ \varphi_{\Delta t}^{0}. \tag{3.17}$$

Each of the short term flow maps can be approximated by means of PCEs as follows:

$$\varphi_{t_0}^{t_0+\Delta t}(y,\boldsymbol{\xi}) = \sum_{i=1}^{P}\sum_{j=1}^{Q} \varphi_{t_0,ij}^{t_0+\Delta t}\,\psi_i(y)\boldsymbol{\psi}_j(\boldsymbol{\xi}). \tag{3.18}$$

The PC coefficients $\varphi_{t_0,ij}^{t_0+\Delta t}$ are computed by means of the Gaussian quadrature method. The representation of the PC-flow map composition is similar to time-dependent PCEs in the sense that current responses are used to predict their future values. The method was applied successfully to predict the positions of particles in a double gyre flow with uncertain initial conditions (two random variables) (Luchtenburg et al., 2014).

The idea that the projection coefficients follow dynamical behaviours was also investigated by Le Maître and Mathelin (2010). The equation-free model reduction method consists in, first projecting the response on a reduced basis using the Karhunen-Loève expansion and second, seeking a mapping for the projection coefficients which is of the form:

$$y_i(t+T) = \mathcal{M}_T(y_i(t)). \tag{3.19}$$

Le Maître and Mathelin (2010) proposed a polynomial representation for the mapping function $\mathcal{M}_T(\cdot)$, of which the coefficients are determined with Gaussian quadrature.

Recently, in the context of structural dynamics, Spiridonakos and Chatzi (2015a,b); Spiridonakos et al. (2016) introduced the PC-nonlinear autoregressive with exogenous input (NARX) model which reads:

$$y(t,\boldsymbol{\xi}) = \sum_{i=1}^{Q}\sum_{j=1}^{P} \vartheta_{i,j}\, \boldsymbol{\psi}_j(\boldsymbol{\xi})\, g_i(\boldsymbol{z}(t)), \tag{3.20}$$

where $g_i(\boldsymbol{z}(t)), i = 1, \ldots, Q$ are non-linear autoregressive functions of past values of the responses and the excitation gathered in the vector $\boldsymbol{z}(t)$. In the proposed approach, the NARX model is used to represent the dynamical behaviour of the system, whereas PCEs tackle the uncertainty part. A two-phase scheme is employed. First, a *stochastic* NARX model is identified to represent the dynamical system. It is characterized by a set of specified NARX model terms and associated *random coefficients*. Second, the latter are represented as PCEs of the random input parameters which govern the uncertainties in the considered system. In the two phases, both the NARX terms and the polynomial functions are selected with the heuristic genetic algorithm. The approach proved its effectiveness in several case studies in structural dynamics with limited number of random variables (Spiridonakos and Chatzi, 2015a,b). It is worth mentioning that early combinations of system identification tools with polynomial chaos expansions can be found in the literature. Ghanem et al. (2005) regressed the restoring force of an oscillator on the Chebychev polynomials of state variables of the system, then used PCEs to represent the polynomial coefficients. Wagner and Ferris (2007) used PC-ARIMA models with a-priori known deterministic coefficients for characterizing terrain topology. Linear ARX-PCE models were also used by Kopsaftopoulos and Fassois (2013), Samara et al. (2013) and Sakellariou and Fassois (2016).

The PC-NARX model can be considered as an extension of time-dependent PCEs or PC-flow map composition. It takes into account the effect of the excitation in the system and past values from several previous steps are included.

Observing that a fixed-in-time basis in KL expansion (or in other reduced basis model) may not describe effectively the solution with strong time-dependent form, Sapsis and

Lermusiaux (2009) proposed dynamically orthogonal field equations, which consists in deriving intrusively a system of equations describing the evolution of the spatial and stochastic basis. The bases are therefore constructed on-the-fly instead of being computed a priori using the covariance matrix and solving eigen-problems. The dynamically orthogonal condition imposes that the evolution of the spatial basis is orthogonal to the space spanned by the basis themselves. Later, Cheng et al. (2013) proposed a dynamically bi-orthogonal method, which is considered the reformulation of the dynamically orthogonal field equations by Choi et al. (2014). These approaches, which are intrusive in essence, are not considered in the present work.

A common feature of the methods presented in this section is that the PC coefficients follow dynamical evolutions, which can be represented by a suitable mapping from the past values of the response. Their formulas differ fundamentally from the classical separated representation, in which the response is projected onto a set of deterministic time-dependent bases $y_i(t)$, which have no connection with the past values of the responses, thus do not involve the dynamics.

3.2.3 Approaches to be investigated

At the beginning, the author considered that high-degree polynomials might help to solve time-dependent problems. Therefore, Mai and Sudret (2015c) developed the hierarchical PCE approach, in which the non-linear interaction terms are selectively added to the expansion depending on the selected linear terms. This approach shares similarities to the adaptive PCEs proposed by Lucor and Karniadakis (2004), when the interaction polynomial functions are selected adaptively under a specified constraint. However, it was soon realized that increasing the polynomial degree in time is not a sustainable approach due to the fact that it only deals with the consequences of the dynamics. It is nothing more than using an uncertainty quantification tool to mimic the dynamics. The root of the problem, which is associated with the dynamical behaviour of the systems, is thus not efficiently tackled.

From that point on, all efforts are focused on the dynamics of the problems. It was observed that one can take advantages of particular characteristics of the problems. For instance, when the responses of interest are periodic, the ideas introduced by Witteveen and Bijl (2008) and Le Maître et al. (2010) presented in Section 3.2.2.1 appear relevant. By means of the time transform or the interpolation on the phase space, the complexity of the derived problem is reduced significantly. This allows an effective use of PCEs. Inspired by those ideas, a non-intrusive approach is proposed in Chapter 4, in which stochastic time transform is used to obtain in-phase trajectories and then, PCEs are applied in the transformed time scale.

In a general case, for instance in structural dynamics, it is of utmost importance to introduce a non-intrusive generalized method. The combination of PCEs and NARX model (Spiridonakos and Chatzi, 2015a,b) seems promising. In this approach, the dynamical behavior is handled effectively by the NARX tool. In Chapter 5, the computation of PC-NARX model based on least angle regression is investigated with applications on structures with random properties subject to stochastic excitations.

3.3 Summary

In the current chapter, classical time-frozen non-intrusive PCEs for time-dependent problems are first investigated. By means of a simple numerical example, the limitations of the classical approach are revealed. We attempted a thorough review of state-of-the-art methods that were introduced in the last decade to solve such problems. Inspired by the ideas of Witteveen and Bijl (2008), Le Maître et al. (2010), Spiridonakos and Chatzi (2015a), two approaches will be investigated in the next two chapters of this manuscript. For problems with periodic or quasi-periodic responses, the time transform idea is developed in an original completely non-intrusive way. For more general cases, the use of the least angle regression technique for computing PC-NARX model is considered.

4

Stochastic time-warping polynomial chaos expansions

The use of polynomial chaos expansions (PCEs) in problems involving stochastic dynamics has so far attracted great attention. In mechanics and fluid dynamics, this problem has a major importance, see for instance the stochastic flow simulations (Wan and Karniadakis, 2006) or random responses of nonlinear aeroelastic systems (Pettit and Beran, 2006). However, the application of PCEs to such problems is facing important challenges, in particular PCEs loose accuracy with time. PCEs are currently not capable of representing the time-dependent responses at late instants and it is of utmost importance to find a solution to this issue.

In this context, an interesting problem emerges in nonlinear oscillating systems possessing a limit cycle[1] which may depend on the uncertain parameters. Limit cycle oscillations (LCO) represent a class of time-dependent problems that plays an important role in several fields, see *e.g.* aerospace engineering (Bunton and Denegri, 2000) and mechanical engineering (Sarrouy et al., 2013) among others. Use of PCEs to represent LCO systems has attracted a large attention and actually almost all novel ideas with PCEs are applied first to LCO systems or systems involving periodicity. For instance, Wan and Karniadakis (2006) used multi-element PCEs whereas Beran et al. (2006a) proposed different methods namely use of Haar wavelets as local bases or use of B-spline functions. These approaches aim at resolving the highly nonlinear behaviour of LCO responses in the stochastic domain. There are also techniques that are designed specifically for LCO. Le Maître et al. (2010) proposed an intrusive time transform of the trajectories which aims at representing the transformed time-histories in a small neighborhood of a reference trajectory, *i.e.* to reduce their variability by making them in-phase. A transformed time line τ is introduced, of which the varying clock speed $\dot{\tau} = \dfrac{\mathrm{d}\tau}{\mathrm{d}t}$ is adjusted in an intrusive setup at each time step. This is achieved by minimizing the Euclidean distance between the distinct trajectories and the reference counterpart. From a similar perspective, Witteveen and Bijl (2008) interpolated the oscillatory responses on the phase space to obtain in-phase oscillations. Inspired by the two mentioned approaches, a non-intrusive time transform, which consists in finding a suitable *stochastic warping* of the time line to increase the *similarity* between different trajectories in the transformed (warped) time scale, is introduced in the present chapter. The proposed approach focuses on increasing the frequency and phase similarity of the considered trajectories in problems involving periodicity.

It is worth noting that in the engineering literature, the time-warping technique has

---

[1] Limit cycle is a closed isolated trajectory in the phase-space of self-oscillated oscillators. The nearby trajectories can either spiral in toward or away from the limit cycle.

been of interest for decades. For instance, in the context of voice recognition, Sakoe and Chiba (1978) first proposed time-warping to eliminate the timing differences and obtain maximum coincidences between two speech patterns. Wang and Gasser (1997) proposed a novel cost function to determine the time-warping function. Later, Ramsay and Li (1998) used the technique under the name "curve registration" for biological data. The essential idea consists in the registration (or alignment) of salient curve features by means of a suitable *smooth monotone transformation* of the temporal variable t. The actual analyses are then carried out on the aligned curves. Note that the same idea can also be conducted in the spatial domain. For instance, registering the outcomes over surfaces or volumes are particularly important in medical imaging (Bookstein, 1997).

In this chapter, we are adding one dimension to the time-warping technique by incorporating the effects of uncertainties in the transformation function. This results in a stochastic time-transform framework. Indeed, due to the inherent randomness of the stochastic problem, a time transformation function with deterministic parameters is not suitable. Therefore, stochastic transform parameters must be used and will be cast as functions of the original random parameters.

The chapter is organised as follows: in Section 4.1, the theory of the stochastic time-warping framework is introduced. A method to determine the stochastic parameters of the transformation function is proposed. Use of principal component analysis is presented for the purpose of reducing the associated computational cost. Section 4.2 illustrates the proposed framework with numerical engineering applications. Discussions on this approach are given in the next section, followed by a summary of the conducted work. The theoretical foundation of the approach was originally introduced by Mai and Sudret (2015a,b) and has been recently submitted for publication (Mai and Sudret, 2016a).

4.1 Stochastic time-warping polynomial chaos expansions for random oscillations

4.1.1 Stochastic time-warping

Consider a dynamical system (*e.g.* a structural dynamic or chemical system) whose behaviour is modelled by a system of first-order differential equations:

$$\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{f}(\boldsymbol{y}, \boldsymbol{\xi}, t), \tag{4.1}$$

where the initial condition is $\boldsymbol{y}(t = 0) = \boldsymbol{y}_0$ and the random vector $\boldsymbol{\xi}$ comprises independent second-order random variables defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\boldsymbol{\xi}$ may include the parameters governing the system behaviour, *e.g.* masses, stiffness, damping ratio, reaction parameters, frequency and amplitude of excitation. The initial condition can also be uncertain, in which case it becomes a random variable belonging to $\boldsymbol{\xi}$. The time-dependent response of the system is denoted by $\boldsymbol{y}(t, \boldsymbol{\xi})$. Without loss of generality, we consider one component of the output quantity, *e.g.* $y(t, \boldsymbol{\xi})$ with the initial condition $y(t = 0) = y_0$. At each time instant, $y(t, \boldsymbol{\xi})$ is assumed to be a second-order random variable. As in the works by Wan and Karniadakis (2005, 2006); Witteveen and Bijl (2008); Le Maître et al. (2010), herein we focus on the class of problems when $y(t, \boldsymbol{\xi})$ is an oscillatory response with random frequencies and amplitudes.

The time dependent response $y(t, \boldsymbol{\xi})$ is represented by means of time frozen PCEs as follows:

$$y(t, \boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}}(t) \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) + \epsilon(t, \boldsymbol{\xi}). \tag{4.2}$$

A virtual time variable $\tau(t, \boldsymbol{\xi})$, which is obtained by a stochastic time-warping, is introduced as follows:

$$\tau(t, \boldsymbol{\xi}) = \sum_{i=0}^{N_\tau} c_i(\boldsymbol{\xi})\, f_i(t) = F(t, \boldsymbol{\xi}), \tag{4.3}$$

where $\{f_i(t), i \in \mathbb{N}\}$ are functions of the physical time t and $\{c_i(\boldsymbol{\xi}), i \in \mathbb{N}\}$ are coefficients which depend on the input random variables $\boldsymbol{\xi}$. The coefficients $\{c_i(\boldsymbol{\xi}), i \in \mathbb{N}\}$ can be represented by means of PCEs as follows:

$$c_i(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} c_{i\boldsymbol{\alpha}}\, \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{4.4}$$

where $\boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ and $c_{i\boldsymbol{\alpha}}$ are respectively the orthonormal polynomial functions and the coefficients of the expansion. The only constraint on the time-warping is that τ is a strictly monotonically increasing function of t. Then the inverse transform may be cast as:

$$t(\tau, \boldsymbol{\xi}) = F^{-1}(\tau, \boldsymbol{\xi}). \tag{4.5}$$

In other words, for each realization $\boldsymbol{\xi}_0$, *i.e.* each trajectory of the system response, we assume a one-to-one mapping between t and τ. The response trajectory may then be represented in the transformed (warped) time scale as follows:

$$y(\tau, \boldsymbol{\xi}) = \sum_{\beta \in \mathcal{B}} y_{\beta}(\tau) \boldsymbol{\psi}_{\beta}(\boldsymbol{\xi}) + \epsilon(\tau, \boldsymbol{\xi}), \tag{4.6}$$

in which $\mathcal{B}$ is the truncation set of the multi-indices $\boldsymbol{\beta}$. The inverse time transform allows one to obtain the PCEs of the response in the physical time scale as follows:

$$y(t, \boldsymbol{\xi}) = y(F^{-1}(\tau, \boldsymbol{\xi}), \boldsymbol{\xi}). \tag{4.7}$$

The objective is to find a suitable time-warping defined by Eq. (4.3) and (4.4) so that the cardinality of $\mathcal{B}$ remains small (*i.e.* low-degree PCEs can be used) to achieve an acceptable error $\epsilon(\tau)$ even at late instants. This can be obtained if the trajectories $y(\tau(t, \boldsymbol{\xi}))$ become in-phase, as suggested by Le Maître et al. (2010) and Witteveen and Bijl (2008). First, a deterministic reference trajectory $y_r(t)$ is introduced. The stochastic time-warping (Eq. (4.3)) is determined by maximizing the similarity between $y(\tau(t, \boldsymbol{\xi}))$ and the reference counterpart $y_r(t)$ for all values of $\boldsymbol{\xi}$, which makes the responses become in-phase. This allows the effective computation of Eq. (4.6). Having at hand the time-warping (Eq. (4.3)) and the PCEs of the response in the virtual time line τ (Eq. (4.6)), one can finally obtain the PCEs in the physical time line t by conducting the inverse time-warping. The proposed non-intrusive time-warping approach is explained in detail in the following. For the sake of clarity, it is graphically summarized in Figure 4.1.

- One first chooses a reference trajectory $y_r(t)$ which is for instance obtained by considering the mean values of the input vector $\boldsymbol{\xi}$, *i.e.* $y_r(t) = y(t, \mathbb{E}[\boldsymbol{\xi}])$. In general, $y_r(t)$ may be any realization of the response quantity $y(t)$ obtained with a specific sample $\boldsymbol{\xi}_0$. For the numerical case studies considered in the current chapter, the choice of $y_r(t)$ did not affect the final results.
- Let us start now the time-warping, which consists in transforming the time line with the purpose of increasing the similarity between different realizations of the output $y(t, \boldsymbol{\xi})$. Assume that one is given a set of trajectories $y_i(t) \equiv y(t, \boldsymbol{\xi}_i)$, $i = 1, \ldots, n$ for n realizations of $\boldsymbol{\xi}$ corresponding to an experimental design in the input space $\mathcal{D}_{\boldsymbol{\Xi}}$. Then for each i, the following steps are performed:
  - Define a linear time-warping $\tau = k_i\, t + \phi_i$. In general, the functions $f_i(t)$ in Eq. (4.3) might be polynomials of t. However, when investigating the problem of vibration with random frequencies, a linear transform usually suffices. This is due to the periodicity of the considered response trajectories. In the intrusive time transform approach (Le Maître et al., 2010), although a linear warping function is not specified for the considered examples, the resulting transformed time τ represents a linear relationship when plotted against t. Wang and Gasser (1997) also used a linear warping function. In particular, given the complexity of the problems under investigation, use of a linear function facilitates the inverse transform in the next phase, which is highly convenient. This linear warping represents two actions, namely scaling and shifting, respectively driven by the parameters k_i and ϕ_i. The time line is stretched (resp. compressed) when $k_i > 1$ (resp. $0 < k_i < 1$) and is shifted to the left (resp. to the right) when $\phi_i < 0$ (resp. $\phi_i > 0$). In fact, the scaling factor k_i (resp. shifting factor ϕ_i) allows to maximize the similarity in frequency (resp. phase) between the considered trajectories.
  - Determine the parameters $(k_i,\, \phi_i)$ governing the time-warping as the solution of an optimization problem which aims at maximizing the similarity between the response trajectory $y_i(k_i\, t + \phi_i)$ and the reference counterpart $y_r(t)$. The details of the optimization problem, in which a measure of similarity is introduced, will be described in the next section.
  - Represent $y_i(t)$ on the transformed time line τ. For this purpose, one chooses a grid line of τ with the desired time interval. In fact, the finer the grid is, the smaller is the error introduced by the *interpolation*. The trajectory $y_i(t)$ is projected onto $\tau_i = k_i\, t + \phi_i$ to obtain $y_i(\tau_i)$. In order to assure that all transformed time lines τ_i start at 0, when $t \leq t_0$, one uses the following transform $\tau_i = \dfrac{k_i\, t_0 + \phi_i}{t_0}\, t$. The small value t_0 is chosen so that $k_i\, t_0 + \phi_i > 0\, \forall i$.

For instance, $t_0 = 0.2$ s is used for the numerical applications in this chapter. Finally the projected trajectory is linearly *interpolated* on the selected time line τ yielding $y_i(\tau)$.

- One builds PCEs of $k(\boldsymbol{\xi})$, $\phi(\boldsymbol{\xi})$ and $y(\tau, \boldsymbol{\xi})$ using the realizations $\{k_i, \phi_i, y_i(\tau), i = 1, \dots, n\}$ as the experimental design (or training set):

$$k(\boldsymbol{\xi}) = \sum_{\boldsymbol{\gamma} \in \mathcal{G}} k_{\boldsymbol{\gamma}} \, \boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\xi}) + \epsilon_k, \tag{4.8}$$

$$\phi(\boldsymbol{\xi}) = \sum_{\boldsymbol{\theta} \in \mathcal{T}} \phi_{\boldsymbol{\theta}} \, \boldsymbol{\psi}_{\boldsymbol{\theta}}(\boldsymbol{\xi}) + \epsilon_\phi, \tag{4.9}$$

$$y(\tau, \boldsymbol{\xi}) = \sum_{\boldsymbol{\beta} \in \mathcal{B}} y_{\boldsymbol{\beta}}(\tau) \boldsymbol{\psi}_{\boldsymbol{\beta}}(\boldsymbol{\xi}) + \epsilon_y(\tau, \boldsymbol{\xi}). \tag{4.10}$$

In the above equations, $\boldsymbol{\gamma}$, $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are multi-indices belonging to the truncation set $\mathcal{G}$, $\mathcal{T}$ and $\mathcal{B}$ of the expansions. $k_{\boldsymbol{\gamma}}$, $\phi_{\boldsymbol{\theta}}$ and $y_{\boldsymbol{\beta}}(\tau)$ are coefficients computed by means of sparse adaptive PCEs (Blatman and Sudret, 2011). $k(\boldsymbol{\xi})$ and $\phi(\boldsymbol{\xi})$ are scalar quantities, therefore the computation of their PCE models is straightforward. However, for the vector-valued response $y(\tau, \boldsymbol{\xi})$, it might be computationally expensive when τ is of important length. This computational cost can be reduced significantly by coupling PCEs with the principal component analysis (Blatman and Sudret, 2013). The combination of PCA and PCEs will be described in detail in Section 4.1.3.

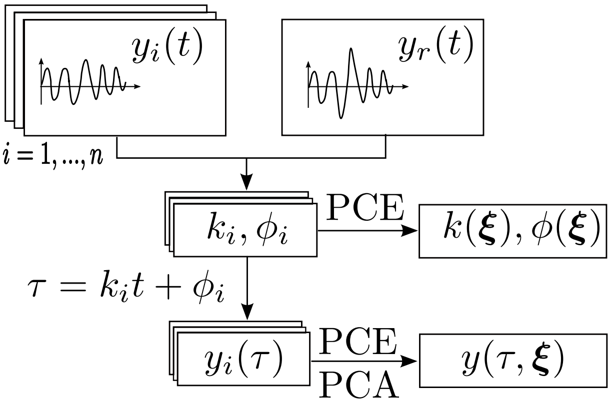


Figure 4.1 – Stochastic time-warping approach: computation of PCEs

4.1.2 Determination of time-warping parameters

This section describes the optimization problem used for determining the parameters k and ϕ of the time-warping process. We first propose a function to measure the similarity between two trajectories $y_1(t)$ and $y_2(t)$:

$$g(y_1(t), y_2(t)) = \frac{\left| \int\limits_0^T y_1(t) y_2(t) \mathrm{d}t \right|}{\|y_1(t)\| \|y_2(t)\|}, \tag{4.11}$$

in which $\int\limits_0^T y_1(t)y_2(t)\mathrm{d}t$ is the inner product of the two considered time histories and $\|\cdot\|$ is the associated L^2-norm. In practice, the trajectories are discretized and thus, the inner product (resp. the L^2-norm) becomes the classical dot product between two vectors (resp. the Euclidean norm). By the Cauchy-Schwarz inequality, this similarity measure always takes values in the interval $[0, 1]$. It attains its maximum when the considered trajectories have the same frequency and phase content, in other words the trajectories are "in-phase". Note that the proposed similarity measure of in-phase vibrations will equal 1 independently of the differences in their amplitudes.

The parameters $(k_i,\, \phi_i), i = 1, \ldots, n$ are determined as the maximizers of the similarity measure between $y_i(\tau)$ and $y_r(t)$. The objective function reads:

$$g(k_i, \phi_i) = \frac{\left|\int\limits_0^T y_i(k_i\, t + \phi_i) y_r(t)\mathrm{d}t\right|}{\|y_i(k_i\, t + \phi_i)\| \|y_r(t)\|}. \tag{4.12}$$

Note that the optimal warping parameters $(k_i,\, \phi_i)$ are different for each trajectory. This results in varying total durations of the trajectories after the warping process. This also occurred in the intrusive time transform approach (Le Maître et al., 2010, Figure 4) The objective function is therefore computed on the overlapped duration between the warped trajectory and the reference one.

Let us now examine the solution (k_i, ϕ_i) of the proposed optimization problem. The constraint that τ is a strictly monotonically increasing function of t requires that $k_i > 0$. In case $y_r(t)$ and $y_i(t, \boldsymbol{\xi}_i)$ are both monochromatic signals, the value of k_i that maximizes their similarity in frequency is unique. However, there are multiple values for the shifting factor ϕ that make the considered trajectories in phase. This will be investigated in the next paragraph.

Figure 4.2 depicts the objective function $g(k, \phi)$ as a similarity measure between the reference trajectory $y_r(t) = \sin(\pi\, t)$ and a response $y(t) = \sin(2\, \pi\, t)$. The two trajectories are chosen in such a way that $(k, \phi) = (2, 0)$ is the maximizer of $g(k, \phi)$. However, there are three global maxima in the depicted interval $[-1.5,\, 1.5]$ of ϕ. This is due to the fact that in the virtual time line τ, if the transformed trajectory $y(\tau)$ is shifted (whether to the left or to the right) a distance equal to 1/2 of the period $T_r = 2\ s$ of the reference counterpart, the similarity measure reaches another global maximum. In fact, if $T_r/4 \leq \phi \leq T_r/2$ (resp. $-T_r/2 \leq \phi \leq -T_r/4$) maximizes the similarity measure, then $\phi - T_r/2$ (resp. $\phi + T_r/2$) in the interval $[-T_r/4,\, T_r/4]$ is also a maximizer. In addition, for the sake of simplicity, it is preferable that ϕ is as close to 0 as possible, *i.e.* the time line of the scaled trajectory is shifted as least as possible. Therefore, the selected value of ϕ needs to satisfy the condition that the shifted distance (in time) is not larger than 1/4 of the period T_r of the reference trajectory $y_r(t)$, *i.e.* $|\phi| \leq T_r/4$. This constraint ensures that the solution is unique. By adopting the constraint on ϕ, one finds the solution $(k,\, \phi) = (2,\, 0)$ for the considered example.

Finally, one can set up the global optimization problem for determining the time-

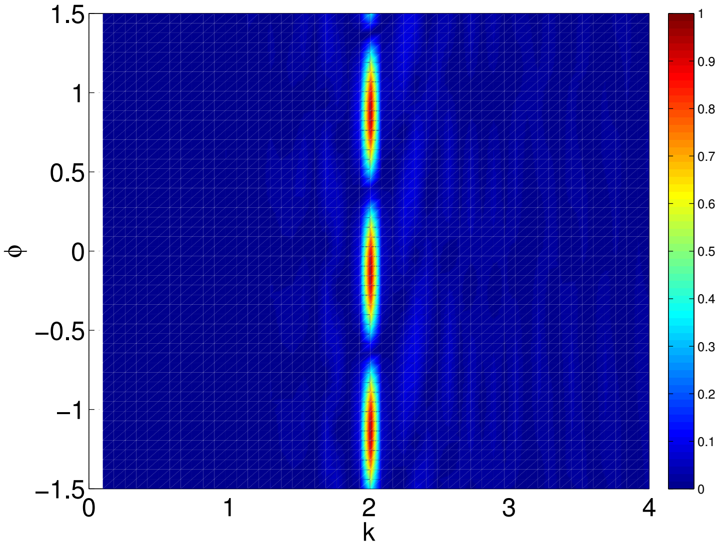


Figure 4.2 – Similarity measure as a function of k and ϕ

warping parameters as follows:

$$(k_i, \phi_i) = \arg \max_{\substack{k_i \in \mathbb{R}^+ \\ |\phi_i| \leqslant T_r/4}} g(k_i, \phi_i). \tag{4.13}$$

This problem can be solved by means of global optimization methods.

4.1.3 Principal component analysis and time-warping polynomial chaos expansions

Blatman and Sudret (2013) proposed a two-step approach which combines principal component analysis (PCA) and PCEs to handle problems involving vector-valued responses, in which case the instant-wise application of PCEs might lead to an important computational burden. The first step consists in conducting PCA to capture the stochastic features of the random vector-valued response with a small number of deterministic principal components and the associated non-physical random variables. The second step relies on representing the resulting random variables with adaptive sparse PCEs. In this section, PCA is first reviewed then the PCA-PCE approach is presented to compute the surrogate model of the transformed responses $y(\tau, \boldsymbol{\xi})$.

4.1.3.1 Principal component analysis

Consider a random vector $\boldsymbol{Y} = (Y_1, \ldots, Y_K)$ in which $Y_i, i = 1, \ldots, K$ is a finite-variance random variable. The covariance matrix $\boldsymbol{\Sigma}$ of the random vector $\boldsymbol{Y}$ reads:

$$\boldsymbol{\Sigma} = \mathbb{E}\left[(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}])(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}])^{\mathsf{T}}\right], \tag{4.14}$$

in which the (i, j)-element is the covariance between random variables Y_i and Y_j:

$$\Sigma_{ij} = \operatorname{Cov}\left[Y_i, Y_j\right] = \mathbb{E}\left[\left(Y_i - \mu_i\right)\left(Y_j - \mu_j\right)\right]. \tag{4.15}$$

The eigen-decomposition of the covariance matrix reads:

$$\boldsymbol{\Sigma} = \boldsymbol{V}\,\boldsymbol{\Lambda}\,\boldsymbol{V}^{\mathsf{T}}, \tag{4.16}$$

in which $\boldsymbol{V} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K\}^{\mathsf{T}}$ is a $K \times K$ matrix with the i^{th} column $\boldsymbol{v}_i$ being an eigenvector of the covariance matrix and $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_K \end{bmatrix}$ is a diagonal matrix whose element λ_i is the eigenvalue corresponding to $\boldsymbol{v}_i$. Note that the eigenvectors are arranged in $\boldsymbol{V}$ according to the descending order of eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_K$.

The principal component decomposition of the random vector $\boldsymbol{Y}$ reads (Jolliffe, 2002):

$$\boldsymbol{Y} = \mathbb{E}\left[\boldsymbol{Y}\right] + \sum_{i=1}^{K} A_i\,\boldsymbol{v}_i, \tag{4.17}$$

where A_i is a random variable which can be obtained as:

$$A_i = \boldsymbol{v}_i^{\mathsf{T}}\left(\boldsymbol{Y} - \mathbb{E}\left[\boldsymbol{Y}\right]\right). \tag{4.18}$$

Given that most of the stochastic features of the random vector $\boldsymbol{Y}$ can be captured by the first few eigenvectors, an approximative representation can be obtained by retaining only first K' eigenvectors in the decomposition as follows:

$$\boldsymbol{Y}^{(K')} = \mathbb{E}\left[\boldsymbol{Y}\right] + \sum_{i=1}^{K'} A_i\,\boldsymbol{v}_i. \tag{4.19}$$

The squared L_2-norm of the truncation error reads (Blatman and Sudret, 2013):

$$\parallel \boldsymbol{Y} - \boldsymbol{Y}^{(K')} \parallel^2 = \sum_{i=K'+1}^{K} \lambda_i. \tag{4.20}$$

A relative error measure can be defined as follows:

$$\frac{\parallel \boldsymbol{Y} - \boldsymbol{Y}^{(K')} \parallel^2}{\parallel \boldsymbol{Y} \parallel^2} = 1 - \frac{\sum\limits_{i=1}^{K'} \lambda_i}{\sum\limits_{i=1}^{K} \lambda_i}. \tag{4.21}$$

This error measure indicates the percentage of the variability in the original random vector $\boldsymbol{Y}$ that is not explained by the set of first K' eigenvectors.

In practice, the mean vector and covariance matrix of $\boldsymbol{Y}$ can only be determined *empirically* using a set of realizations $\mathcal{Y} = \left\{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(N)}\right\}$. The $1 \times K$ empirical mean $\bar{\boldsymbol{y}} = \{\bar{y}_1, \ldots, \bar{y}_K\}$ is calculated by:

$$\bar{\boldsymbol{y}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{y}^{(i)}. \tag{4.22}$$

The (i, j)-element of the empirical covariance matrix reads:

$$\tilde{\Sigma}_{i,j} = \frac{1}{N-1} \sum_{k=1}^{N} \left(y_i^{(k)} - \bar{y}_i\right) \left(y_j^{(k)} - \bar{y}_j\right). \tag{4.23}$$

The eigen-decomposition of the sample-based covariance matrix is given by:

$$\tilde{\boldsymbol{\Sigma}} = \tilde{\boldsymbol{V}} \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{V}}^{\mathsf{T}}, \tag{4.24}$$

where $\tilde{\boldsymbol{V}}$ is the $K \times K$ eigenvectors matrix $\tilde{\boldsymbol{V}} = \{\tilde{\boldsymbol{v}}_1, \ldots, \tilde{\boldsymbol{v}}_K\}$ and $\tilde{\boldsymbol{\Lambda}} = \begin{bmatrix} \tilde{\lambda}_1 & 0 & 0 & \ldots & 0 \\ 0 & \tilde{\lambda}_2 & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & \tilde{\lambda}_K \end{bmatrix}$ is a diagonal matrix with $\tilde{\lambda}_i$ being the eigenvalue associated with the $K \times 1$ empirical eigenvector $\tilde{\boldsymbol{v}}_i$.

The approximate principal component decomposition of $\boldsymbol{Y}$ using empirical mean and eigenvectors is written as follows:

$$\boldsymbol{Y} = \bar{\boldsymbol{y}} + \sum_{i=1}^{K} A_i \, \tilde{\boldsymbol{v}}_i, \tag{4.25}$$

where the random variable A_i is obtained with:

$$A_i = \tilde{\boldsymbol{v}}_i^{\mathsf{T}} \, (\boldsymbol{Y} - \bar{\boldsymbol{y}}). \tag{4.26}$$

The sample matrix $\mathcal{Y}$ can be represented by PCA as follows:

$$\mathcal{Y} = \bar{\mathcal{Y}} + \sum_{i=1}^{K} \boldsymbol{a}_i \, \tilde{\boldsymbol{v}}_i^{\mathsf{T}}, \tag{4.27}$$

in which $\bar{\mathcal{Y}} = \{\bar{\boldsymbol{y}}, \ldots, \bar{\boldsymbol{y}}\}^{\mathsf{T}}$ is a $N \times K$ matrix obtained by replicating the empirical mean N times, $\boldsymbol{a}_i$ is the $N \times 1$ vector of projection coefficients and $\tilde{\boldsymbol{v}}_i$ is the $K \times 1$ empirical eigenvector. The coefficients vector $\boldsymbol{a}_i$ which contains N samples of the random variable A_i is calculated by:

$$\boldsymbol{a}_i = (\mathcal{Y} - \bar{\mathcal{Y}}) \, \tilde{\boldsymbol{v}}_i. \tag{4.28}$$

4.1.3.2 Use of PCA with time-warping PCEs

Having reviewed the construction of PCA, let us now apply this technique to the problem of representing the time dependent response quantity $y(\tau, \boldsymbol{\xi})$. The latter is considered a random vector with the response at each time instant τ being a finite-variance random variable. A sample set $\mathcal{Y}$ of $y(\tau, \boldsymbol{\xi})$ is obtained after the time-warping process. Due to the time scaling and shifting, the durations of the different trajectories are not equal. In other words, in the matrix $\mathcal{Y}$ there are missing values at the end of certain rows. Similarly to Eq. (4.22), the empirical mean $\bar{\boldsymbol{y}}(\tau)$ of the response can be computed at each instant as:

$$\bar{y}(\tau_i) = \frac{1}{\mathrm{card}(\mathcal{K}_i)} \sum_{k \in \mathcal{K}_i} y(\tau_i, \boldsymbol{\xi})^{(k)}, \tag{4.29}$$

in which $\mathcal{K}_i$ is the set of observations where the value of $y(\tau_i, \boldsymbol{\xi})$ is available. Similarly to Eq. (4.23), the empirical covariance matrix $\tilde{\boldsymbol{\Sigma}}$ can be computed element-wise:

$$\tilde{\Sigma}_{i,j} = \frac{1}{\mathrm{card}(\mathcal{K}_{ij}) - 1} \sum_{k \in \mathcal{K}_{ij}} \left(y(\tau_i, \boldsymbol{\xi})^{(k)} - \bar{y}(\tau_i) \right) \left(y(\tau_j, \boldsymbol{\xi})^{(k)} - \bar{y}(\tau_j) \right), \tag{4.30}$$

in which $\mathcal{K}_{ij}$ is the set of observations where the realizations of $y(\tau_i, \boldsymbol{\xi})$ and $y(\tau_j, \boldsymbol{\xi})$ are both available. $\tilde{\boldsymbol{\Sigma}}$ is then used to compute the empirical eigenvectors $\tilde{\boldsymbol{v}}_i$ (Eq. (4.24)).

The response can be represented by means of PCA as follows:

$$y(\tau, \boldsymbol{\xi}) = \bar{\boldsymbol{y}}(\tau) + \sum_{i=1}^{K} A_i(\boldsymbol{\xi})\, \tilde{\boldsymbol{v}}_i(\tau), \tag{4.31}$$

where $\bar{\boldsymbol{y}}(\tau)$ is the empirical mean vector, $\tilde{\boldsymbol{v}}_i(\tau)$ is an empirical eigenvector determined with a sample set of $y(\tau, \boldsymbol{\xi})$ and $A_i(\boldsymbol{\xi})$ is a finite variance random variable. Only a few eigenvectors are retained in the decomposition, which leads to:

$$y(\tau, \boldsymbol{\xi}) = \bar{\boldsymbol{y}}(\tau) + \sum_{i=1}^{K'} A_i(\boldsymbol{\xi})\, \tilde{\boldsymbol{v}}_i(\tau) + \epsilon_1(\tau). \tag{4.32}$$

The number of principal components is selected so that the relative error $1 - \frac{\sum_{i=1}^{K'} \lambda_i}{\sum_{i=1}^{K} \lambda_i}$ is smaller than a prescribed threshold, *e.g.* $\epsilon = 0.01$. The samples of the random coefficient $A_i(\boldsymbol{\xi})$ can be obtained using Eq. (4.28), which are then used as experimental design to compute the PCE of this random coefficient:

$$A_i(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} c_{i,\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) + \epsilon_{2,i}. \tag{4.33}$$

Finally, the response in the transformed time scale is represented by PCA and PCEs as follows:

$$y(\tau,\boldsymbol{\xi}) = \bar{\boldsymbol{y}}(\tau) + \sum_{i=1}^{K'} \sum_{\boldsymbol{\alpha}\in\mathcal{A}} c_{i,\boldsymbol{\alpha}}\, \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\, \tilde{\boldsymbol{v}}_i(\tau) + \epsilon(\tau). \tag{4.34}$$

Note that Blatman and Sudret (2013) introduced a measure of the *upper bound* of the total error induced by the truncation of the principal component analysis and the approximation of the random coefficients $A_i(\boldsymbol{\xi})$ by PCEs. The reader is referred to the mentioned paper for more details. Herein this error measure can be used as an indicator of the accuracy of the computed surrogate models.

4.1.4 Predicting random oscillations with time-warping polynomial chaos expansions

Let us now demonstrate the use of time-warping PCEs to predict responses of the model given a new set of input parameters $\boldsymbol{\xi}'$. For the sake of clarity, the procedure is depicted in Figure 4.3 and explained in two steps as follows:

- One predicts $k(\boldsymbol{\xi}')$, $\phi(\boldsymbol{\xi}')$ and $y(\tau,\boldsymbol{\xi}')$ using the computed PCEs in equations (4.8), (4.9) and (4.34).
- One maps $y(\tau,\boldsymbol{\xi}')$ into $y(t,\boldsymbol{\xi}')$ using the inverse time-warping $t = \dfrac{\tau - \phi(\boldsymbol{\xi}')}{k(\boldsymbol{\xi}')}$ when $t > t_0$ and $t = \dfrac{t_0}{k(\boldsymbol{\xi}')\, t_0 + \phi(\boldsymbol{\xi}')}\, \tau$ when $t \leq t_0$.

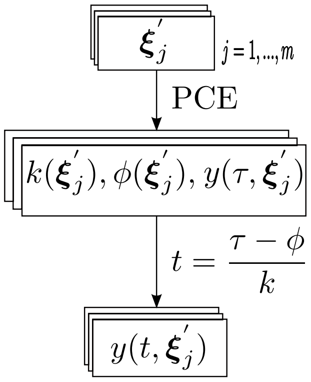


Figure 4.3 – Stochastic time-warping approach: prediction of the response trajectories using PCEs

4.2 Numerical applications

The time-warping-based polynomial chaos expansions (PCEs) developed in this chapter are now applied to various engineering problems, namely a model of rigid body dynamics, a nonlinear Duffing oscillator, the so-called Oregonator model describing the chemical reaction between three species, a Bouc-Wen oscillator subject to a stochastic sinusoidal excitation and the so-called Kraichnan-Orszag model. In each case, time-frozen sparse PCEs[2] are applied first to show the degradation of the prediction accuracy after a certain time. Time-warping PCEs with simple linear time transforms are then investigated. The PCE surrogate models are computed using a small number of numerical simulations of the original model as experimental design, then validated on a large independent validation set of size $N_{val} = 10,000$. The accuracy of the time-frozen and time-warping PCE models are judged on the basis of predicting the responses to specific sets of input parameters and estimating the time histories of first- and second-order statistics of the responses.

More precisely, the relative error of the prediction $\#i$ reads:

$$\epsilon_{val,i} = \frac{\sum\limits_{t=1}^{T} \left(y(t,\boldsymbol{\xi}_i) - \hat{y}(t,\boldsymbol{\xi}_i)\right)^2}{\sum\limits_{t=1}^{T} \left(y(t,\boldsymbol{\xi}_i) - \bar{y}(t,\boldsymbol{\xi}_i)\right)^2}, \tag{4.35}$$

where $\hat{y}(t,\boldsymbol{\xi}_i)$ is the output trajectory predicted by PCEs and $\bar{y}(t,\boldsymbol{\xi}_i)$ is the mean value of the actual response time series $y(t,\boldsymbol{\xi}_i)$ which is obtained with the original numerical solver. The above formula is also used to assess the accuracy of the predicted time-dependent statistics (*i.e.* mean, standard deviation).

These problems are solved in the UQLab framework (Marelli and Sudret, 2014b), more specifically using the least angle regression algorithm implemented in the polynomial chaos expansion module (Marelli and Sudret, 2015).

4.2.1 Rigid body dynamics

We first consider the rotation of a rigid body described by Euler's equations (Peraire and Widnall, 2009) (Figure 4.4). The conservation of angular momentum reads:

$$\left\{ \begin{array}{l} M_x = I_{xx}\,\dot{x} - (I_{yy} - I_{zz})\,y\,z, \\ M_y = I_{yy}\,\dot{y} - (I_{zz} - I_{xx})\,z\,x, \\ M_z = I_{zz}\,\dot{z} - (I_{xx} - I_{yy})\,x\,y, \end{array} \right. \tag{4.36}$$

in which M_x, M_y, M_z are the external moments, I_{xx}, I_{yy}, I_{zz} are the moments of inertia and x, y, z are the angular velocities about the principal axes. In the case when the rigid body rotates freely under no external excitation, *i.e.* $M_x = M_y = M_z = 0$ and $I_{xx} = \dfrac{1-\xi}{2} I_{yy}$, $I_{zz} = \dfrac{1+\xi}{2} I_{yy}$, one obtains the following set of reduced equations:

$$\left\{ \begin{array}{l} \dot{x}(t) = y(t)\,z(t), \\ \dot{y}(t) = \xi\,x(t)\,z(t), \\ \dot{z}(t) = -x(t)\,y(t). \end{array} \right. \tag{4.37}$$

---

[2] For the sake of simplicity, the term "time-frozen PCEs", which refers to the instantaneous sparse adaptive PCEs presented in Chapter 3, will be used throughout the manuscript.

The initial conditions are set equal to $x(0) = 0$, $y(0) = 1$, $z(0) = 1$. Assume that ξ is modelled by a random variable with uniform distribution: $\xi \sim \mathcal{U}(-1, 1)$. Suppose a solver of the coupled ordinary differential equations is available. For any realization of ξ, this solver provides discretized trajectories $\{\{x(t_i), y(t_i), z(t_i)\}, t_i = 0, \Delta_t, \ldots, n\,\Delta_t \equiv T\}$. In this example, the equations are solved using the Matlab ordinary differential equation solver `ode45` (Runge-Kutta method, total duration $T = 50$ s, time step $\Delta_t = 0.01$). We aim at building PCEs of the angular velocity $x(t)$ as a function of the random variable ξ. Note that the corresponding polynomial functions are from the family of orthonormal Legendre polynomials since ξ is uniformly distributed.

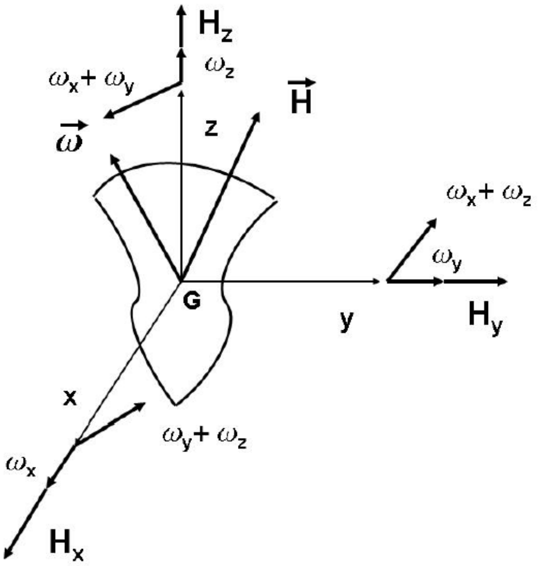


Figure 4.4 – Rotation of a rigid body (after Peraire and Widnall (2009)).

Figure 4.5 depicts a set of 50 trajectories of $x(t)$ obtained for different realizations of the random variable ξ. This set is used as the experimental design for fitting the time-frozen PCEs. $x(t)$ are oscillatory trajectories which fluctuate around zero at different frequencies. This is a typical example of the problem of stochastic oscillation with uncertain frequencies (Wan and Karniadakis, 2005, 2006). At the early instants ($t < 10$ s), one can differentiate between the distinct trajectories, whereas this is hardly the case at later instants, since the patterns are mixed up completely. Due to the growing difference in frequency and phase, $x(t, \xi)$ is more and more non-linear as a function of ξ for increasing t and subsequently, the probability density function of $X(t)$ becomes bi-modal at late instants (see Figure 4.6). This explains why increasing-degree time-frozen PCEs are required in order to represent $x(t)$. As analyzed previously, this is not a sustainable approach since the order of PCEs will certainly become overly high at some point.

Time-frozen PCEs are now utilized to model the variability of the response trajectories, and exemplify the deficiency of such an approach. At each instant t, an adaptive PCE scheme with total polynomial degree increasing from 1 to 20 is used (Eq. (4.2)) based on the available 50 data points from the experimental design made of the 50 trajectories. The PCE model which results in the smallest leave-one-out (LOO) error is retained. Figure 4.7

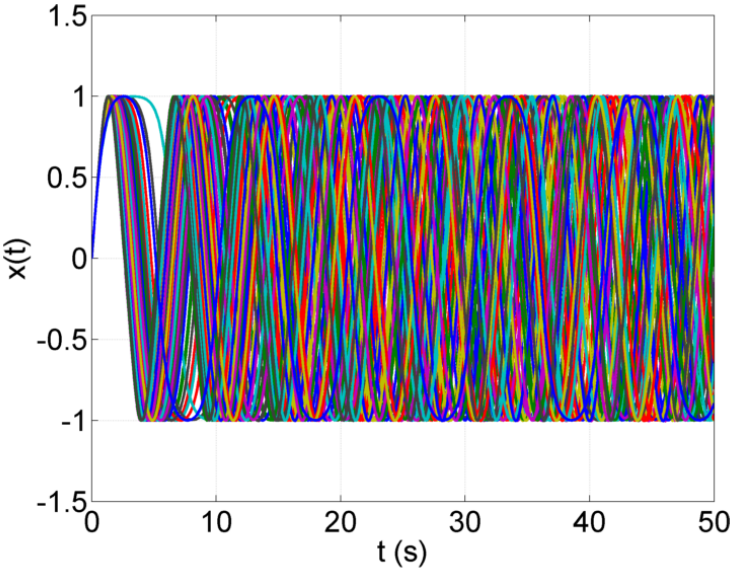
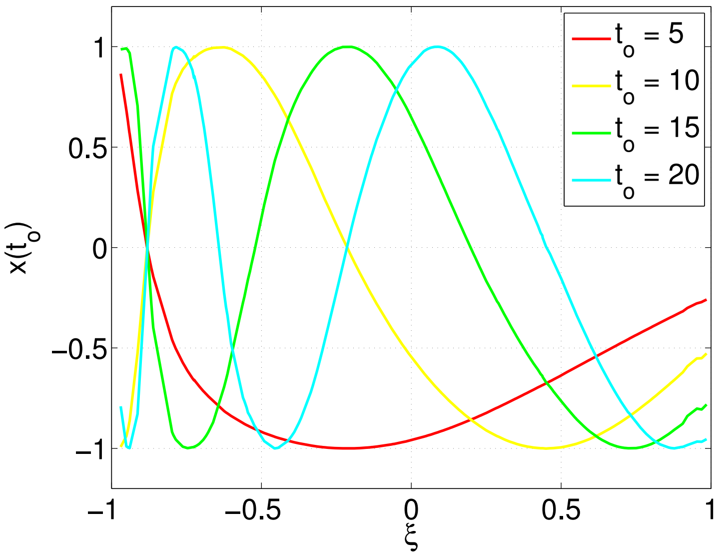


Figure 4.5 – Rigid body dynamics – $N = 50$ different trajectories $x(t)$ in the original time scale t.

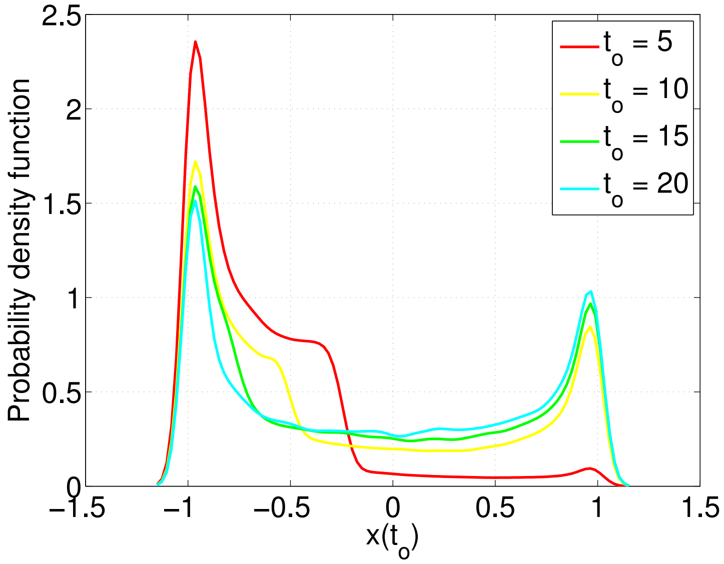
depicts the LOO error of these time-frozen PCEs, which is increasing in time, showing that the accuracy of the PCE model degenerates.

For validation purpose, a set of $10,000$ trajectories is computed using the *ode45* Matlab solver. Figure 4.8 depicts two particular response trajectories predicted by time-frozen PCEs versus the actual responses obtained by numerically solving the system of ordinary differential equations (4.37). After 15 s (when the LOO error is approximately 10^{-2}) the PCE prediction deviates significantly from the actual trajectory. In particular, there are signs of instability in the PCE model, *e.g.* the PCE-based prediction for consecutive instants differ noticeably in terms of accuracy. We now consider the time-dependent statistics of the response $x(t)$. Figure 4.9 represents its mean and standard deviation. In the early time instants ($t < 15$ s), time-frozen PCEs represent the statistics with relatively small error compared to Monte Carlo simulation (MCS). However, after 15 s, the accuracy declines quickly. In particular, PCEs cannot capture the oscillatory behavior of the standard deviation. Another interpretation is that even degree-20 time-frozen PCEs cannot capture the complex distribution of the response at late time instants.

Let us now apply the time-warping approach to pre-process the trajectories $x(t)$. Provided that the initial condition is equal to 0, it suffices to use a linear time-warping $\tau = k\,t$. For each computed realization of the angular velocity $x(t, \xi_i), i = 1, \dots, 50$, the parameters k_i is estimated as the maximizer of the similarity measure described in Eq. (4.12). Note that the same 50 trajectories are used as the experimental design for this approach and the reference trajectory is obtained with the mean value of the input parameter. The optimization problem is solved by means of the global optimization toolbox in Matlab. The function `fmincon` based upon an interior-point algorithm is used while allowing for a maximum of $2,000$ function evaluations. Adaptive sparse PCEs of degree up to 20 are used to represent the parameter k. The relative LOO error is 3.82×10^{-4}, which indicates



(a) $x(t_o)$ as a function of ξ at instants $t_o = 5$, 10, 15, 20 s



(b) Probability density function of $x(t_o)$ at $t_o = 5$, 10, 15, 20 s

Figure 4.6 – Rigid body dynamics – $x(t, \xi)$ as a function of ξ for particular instants and its probability density function.

a high accuracy of the PCE model.

The time-warping is carried out using the estimated parameters and the responses are interpolated into the transformed time line τ, leading to in-phase trajectories $x(\tau)$ (see Figure 4.10). As expected, $x(\tau)$ are smooth functions of ξ at all instants, which allows the effective use of PCEs.

Then principal component analysis (PCA) is conducted on the obtained transformed trajectories. 18 first principal components are retained in order to obtain a PCA trunca-

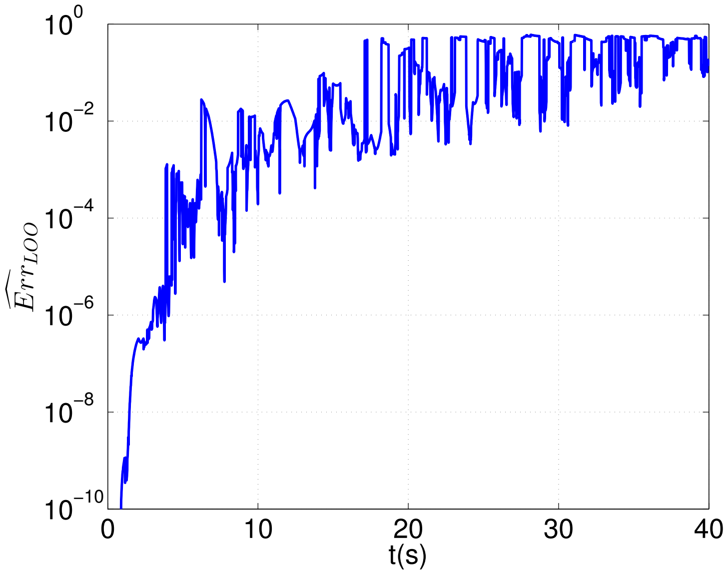
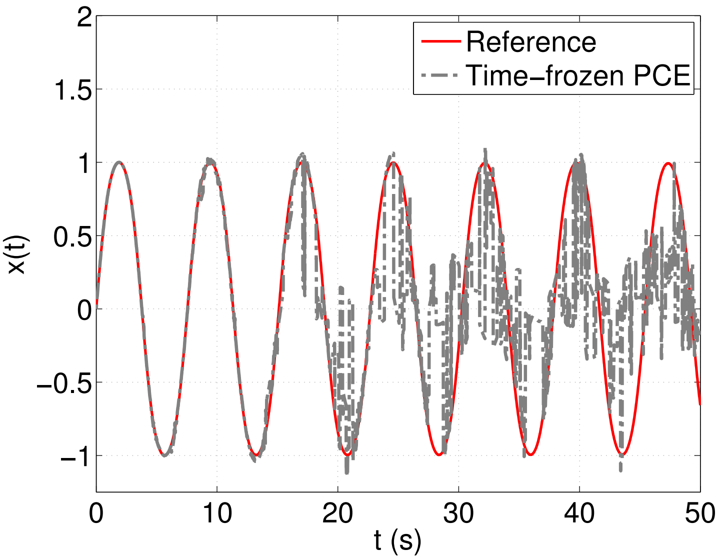


Figure 4.7 – Rigid body dynamics – Leave-one-out error of time-frozen PCEs.

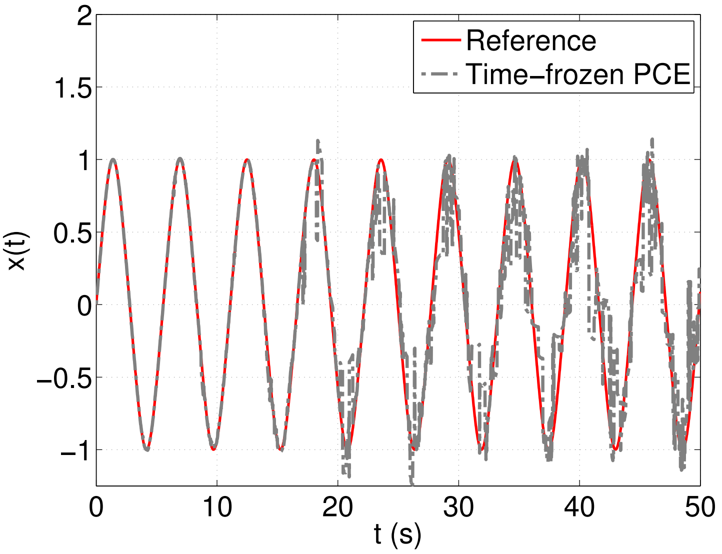
tion error $\epsilon_1 = \sum\limits_{i=K'+1}^{K} \lambda_i / \sum\limits_{i=1}^{K} \lambda_i$ smaller than 1×10^{-3}. The first eight principal components are depicted in Figure 4.11. Figure 4.12 depicts the PCA truncation error ϵ_1 as a function of the number of retained principal components, the LOO error ϵ_2 of the PCE for the coefficient of each principal component and the upper bound of the total error of the PCA-PCE model. It shows that the PCA truncation error ϵ_1 decreases exponentially with the number of retained principal components. Using PCE to represent the first PCA coefficient, the obtained relative LOO error is 7.7×10^{-3}. It is also clear that it is harder to represent the higher mode PCA coefficients by means of PCEs, which was observed by Blatman and Sudret (2013). However, it is worth noting that most of the stochastic features of the response can be captured by the first few components.

Figure 4.13 depicts two specific realizations of the angular velocity $x(t)$ predicted by time-warping PCEs, which are plotted together with the predictions by time-frozen sparse PCEs and the actual responses obtained by the numerical solver. As mentioned previously, one observes that starting from 15 s, the direct approach encounters instability, which results in inaccurate predictions. The time-warping approach allows one to improve notably the quality of the PCE prediction. The predictions by time-warping PCEs are in excellent agreement with the actual responses. A relative error exceeding 0.1 is recorded in only 79 simulations among $10,000$ validations.

In Figure 4.14, the time-dependent mean and standard deviation of the response are plotted. Time-frozen PCEs allow one to represent the mean trajectory with relatively small discrepancy compared to the trajectory obtained with the MCS. It can faithfully predict the standard deviation at the early instants $t < 15$ s, however becomes suddenly unstable afterwards. In contrast, time-warping PCEs provide estimates of the statistics that are almost indistinguishable from the MCS estimates. The relative errors between the reference and predicted mean and standard deviation are 7.31×10^{-4} and 7.19×10^{-4}.

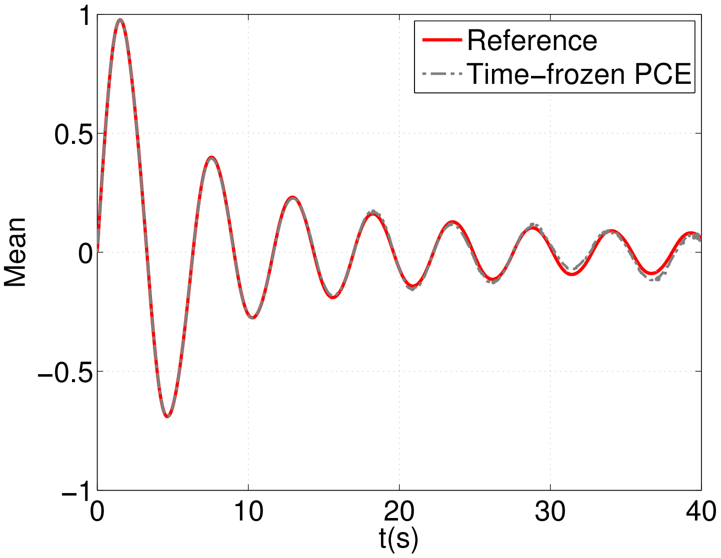


(a) $\xi = -0.5385$

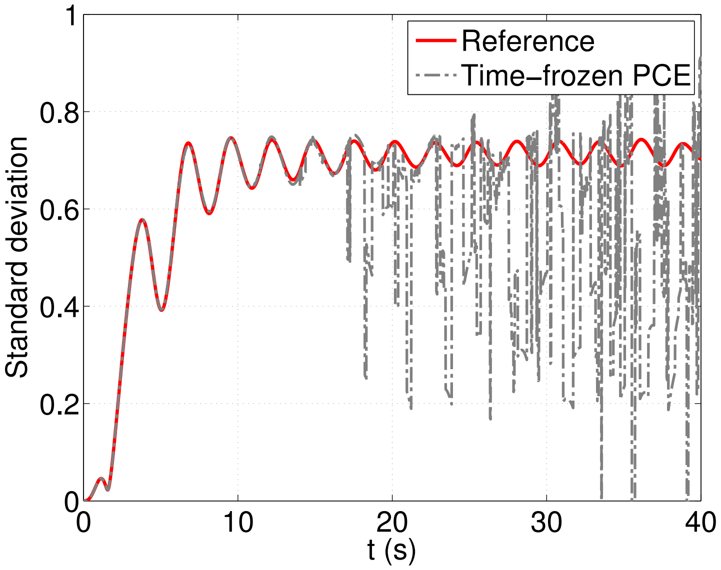


(b) $\xi = 0.6256$

Figure 4.8 – Rigid body dynamics – Two particular trajectories and their prediction by time-frozen PCEs.

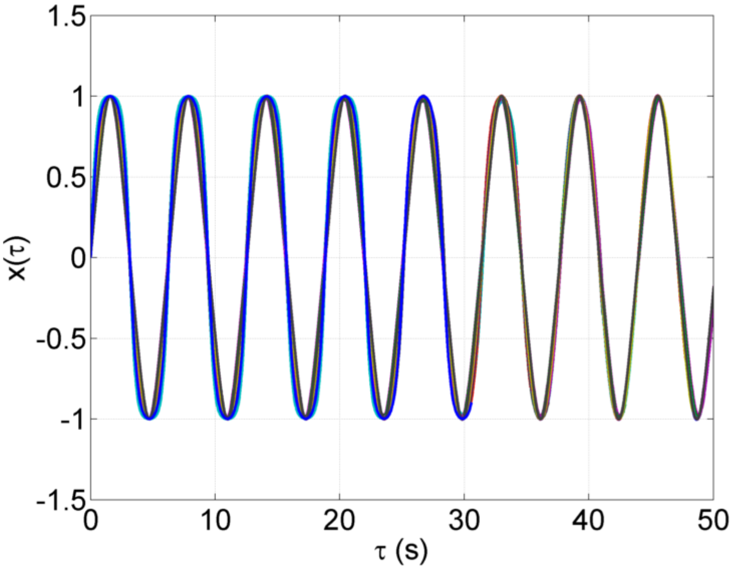


(a) Mean trajectory

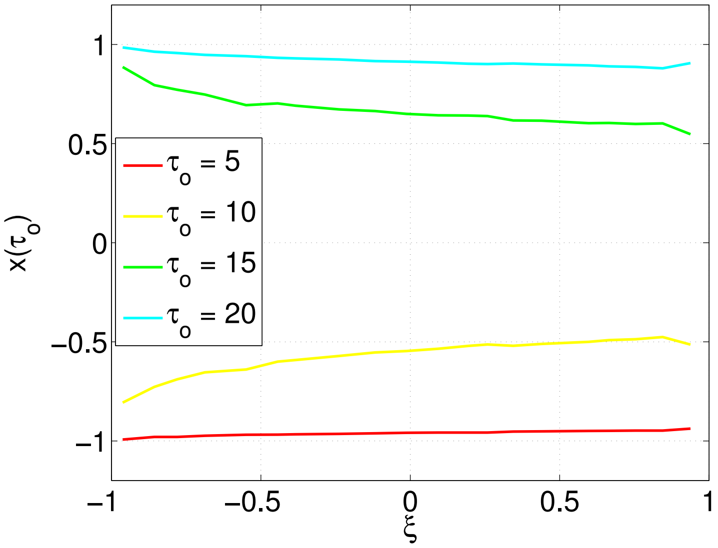


(b) Standard deviation

Figure 4.9 – Rigid body dynamics – Mean and standard deviation of the trajectories predicted by time-frozen PCEs.

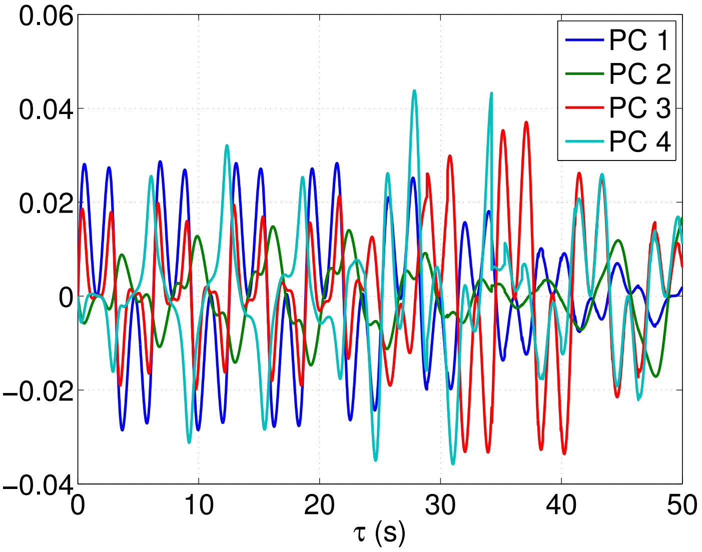


(a) $N = 50$ different trajectories $x(\tau)$ in the warped time scale τ

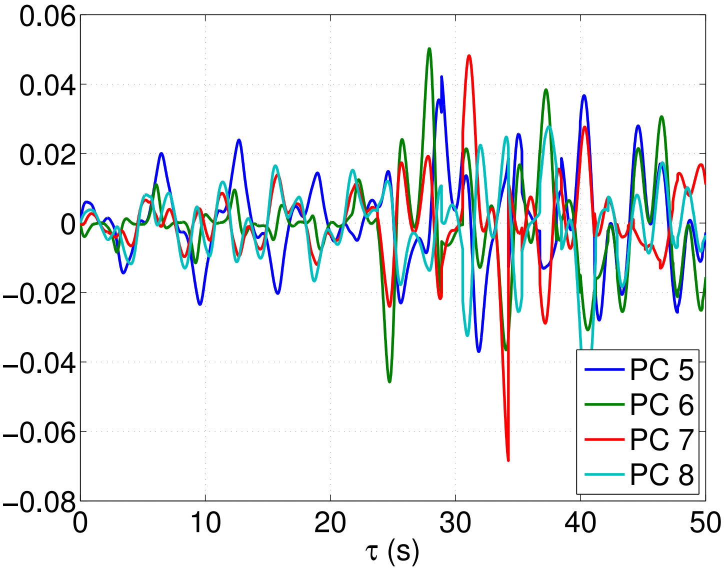


(b) Relationship between the response $x(\tau)$ and the random variable ξ in the warped time scale τ

Figure 4.10 – Rigid body dynamics – Different trajectories $x(\tau)$ in the warped time scale τ and $x(\tau)$ as a function of the random variable ξ.



(a) Principal components 1-4



(b) Principal components 5-8

Figure 4.11 – Rigid body dynamics – The first eight principal components.

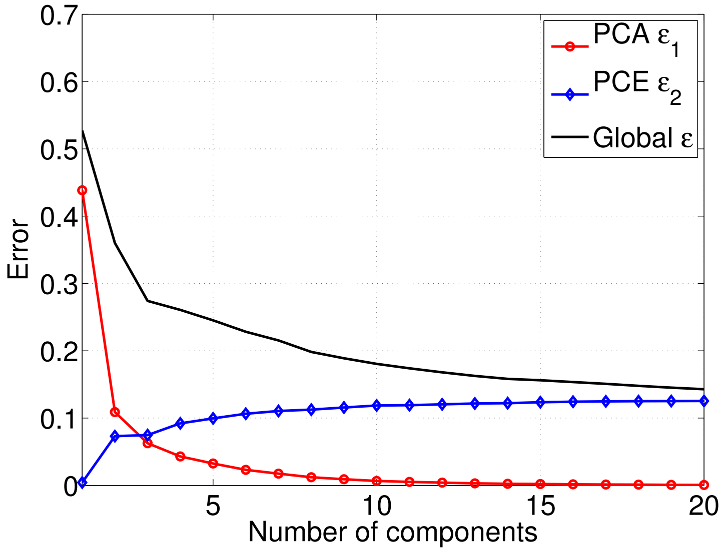
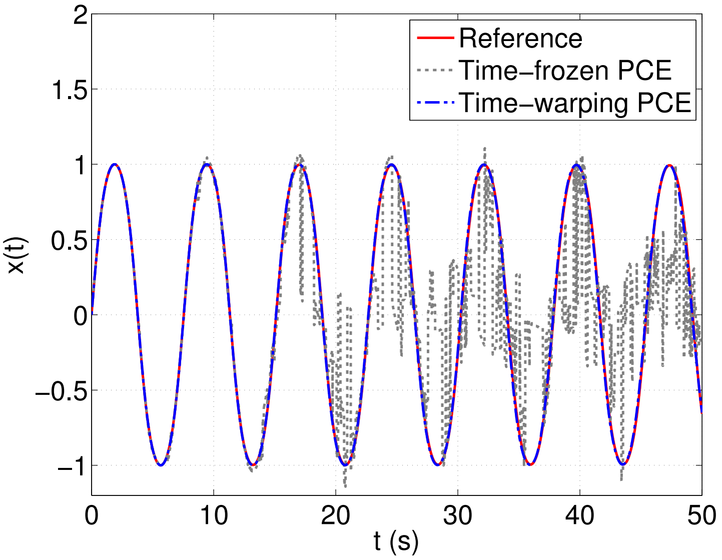
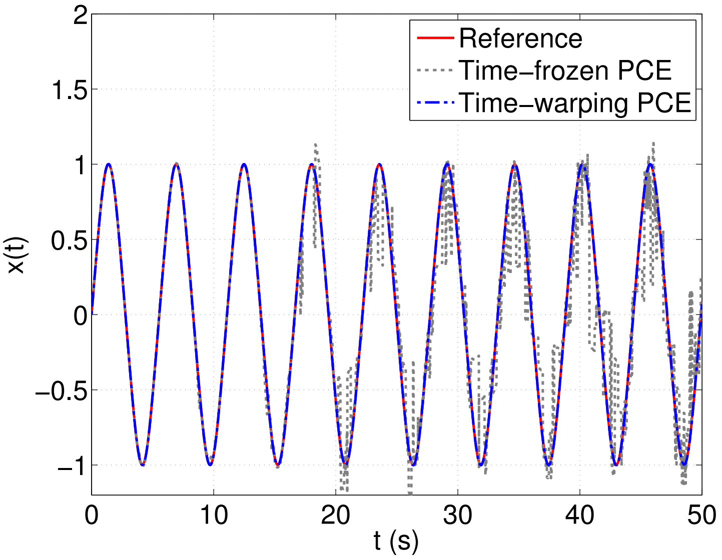


Figure 4.12 – Rigid body dynamics – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace($\tilde{\boldsymbol{\Sigma}}$) and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

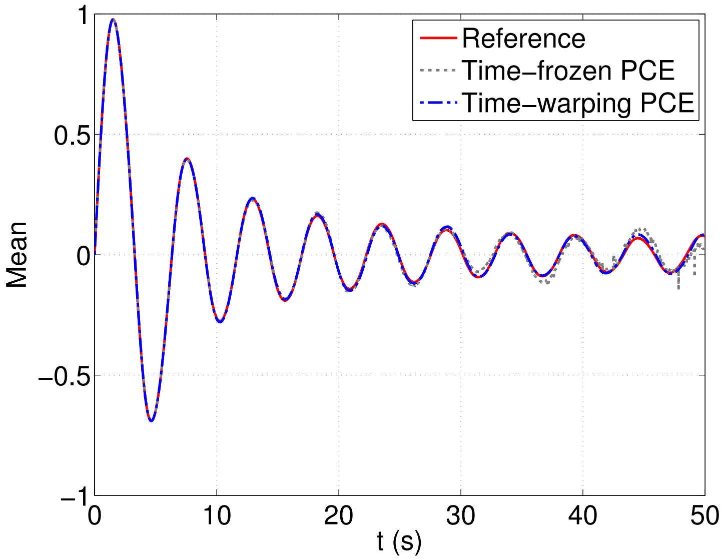


(a) $\xi = -0.5385$

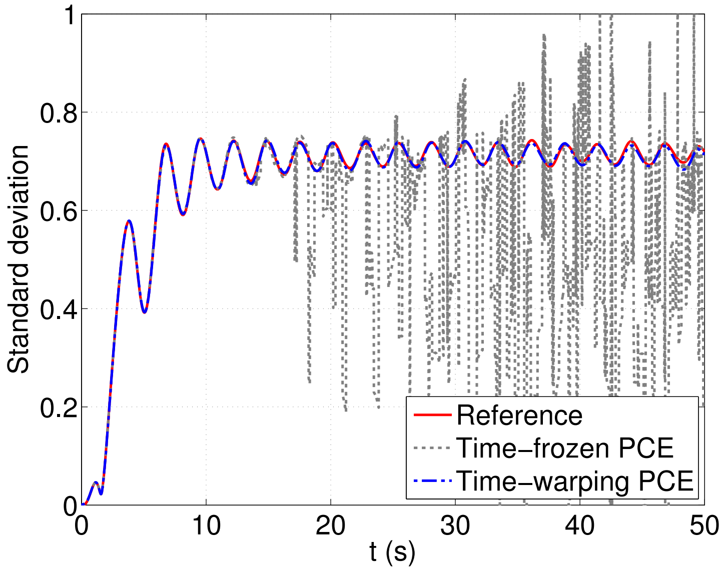


(b) $\xi = 0.6256$

Figure 4.13 – Rigid body dynamics – Two particular trajectories and their predictions by time-frozen and time-warping PCEs.



(a) Mean trajectory



(b) Standard deviation

Figure 4.14 – Rigid body dynamics – Mean and standard deviation of the trajectories: comparison of the two approaches.

4.2.2 Duffing oscillator

Let us consider a non-linear damped single-degree-of-freedom Duffing oscillator under free vibration, which is described by the following equation of motion:

$$\ddot{y}(t) + 2\,\omega\,\zeta\,\dot{y}(t) + \omega^2\left(y(t) + \epsilon\,y^3(t)\right) = 0. \tag{4.38}$$

The oscillator is driven by uncertain parameters described in Table 4.1 and the initial conditions are $y(t=0) = 1$ and $\dot{y}(t=0) = 0$. Note that a simplified form of this equation which represents an undamped linear oscillator was used in other publications for illustrating the time-dependent generalized polynomial chaos (Gerritsma et al., 2010), the intrusive time-transform approach (Le Maître et al., 2010) and the flow map composition PCEs (Luchtenburg et al., 2014).

Table 4.1 – Uncertain parameters of the Duffing oscillator

Input parameters	Distribution	Values
ζ	Uniform	$\mathcal{U}[0.015,\,0.045]$
ω	Uniform	$\mathcal{U}[\pi,\,3\,\pi]$
ϵ	Uniform	$\mathcal{U}[-0.25,\,-0.75]$

Hereafter, we aim at building PCEs of the displacement $y(t)$ as a function of the random variables $(\zeta,\,\omega,\,\epsilon)$. First, we use 200 trajectories of $y(t)$ as experimental design to compute time-frozen PCEs of degree up to 20. The corresponding LOO error (Figure 4.15) start showing insufficient accuracy at $t = 3$ s.

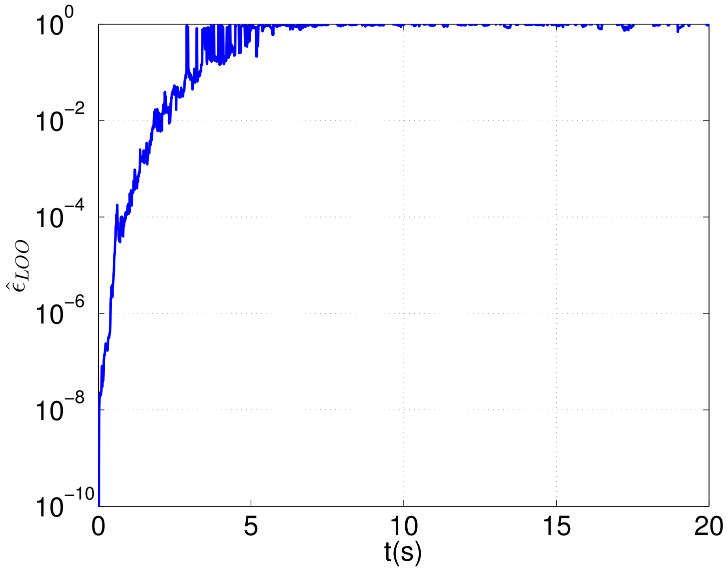
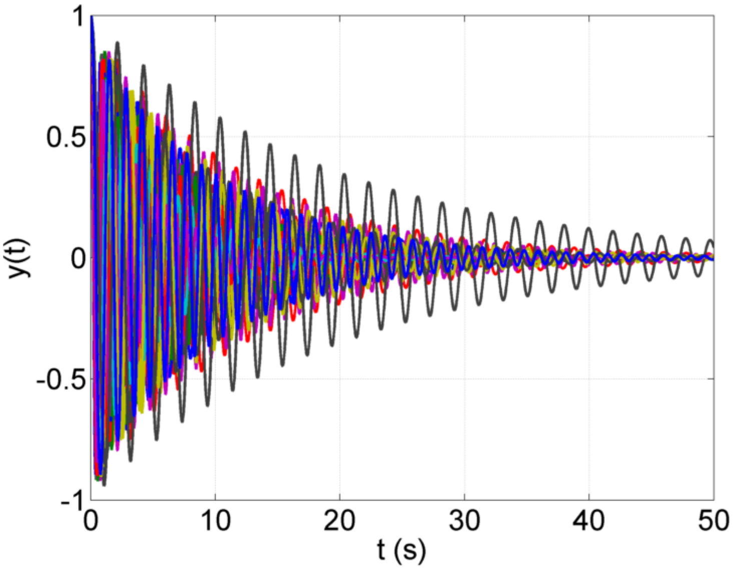
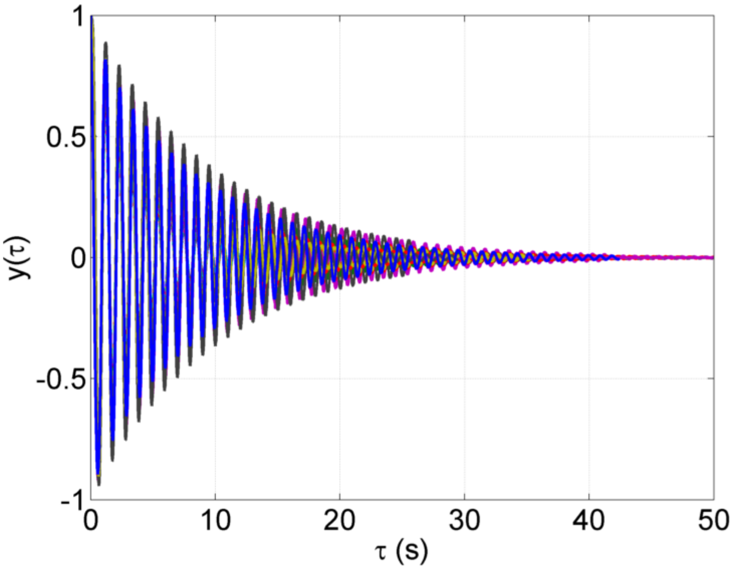


Figure 4.15 – Duffing oscillator – Leave-one-out error $\hat{\epsilon}_{LOO}$ of time-frozen PCEs.

We now use the time-warping approach, which requires only 50 trajectories $y(t)$ as experimental design. The 50 trajectories in the original time scale are plotted in Figure 4.16(a). The same trajectories after time-warping are plotted in Figure 4.16(b). A linear time-warping with two parameters, *i.e.* $\tau = k\,t + \phi$, is used for each trajectory. Using PCEs of degree up to 20, the metamodels of k and ϕ are obtained with relative LOO error 1.87×10^{-5} and 2.08×10^{-4} respectively, which indicates a high level of accuracy.



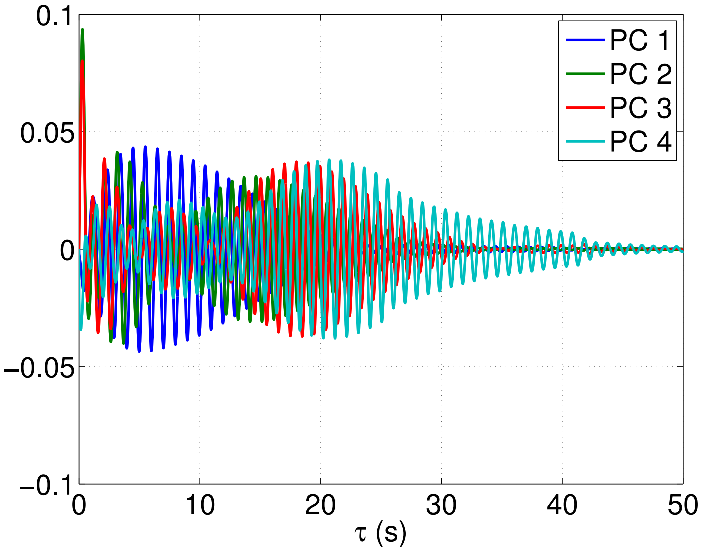
(a) Original time scale t



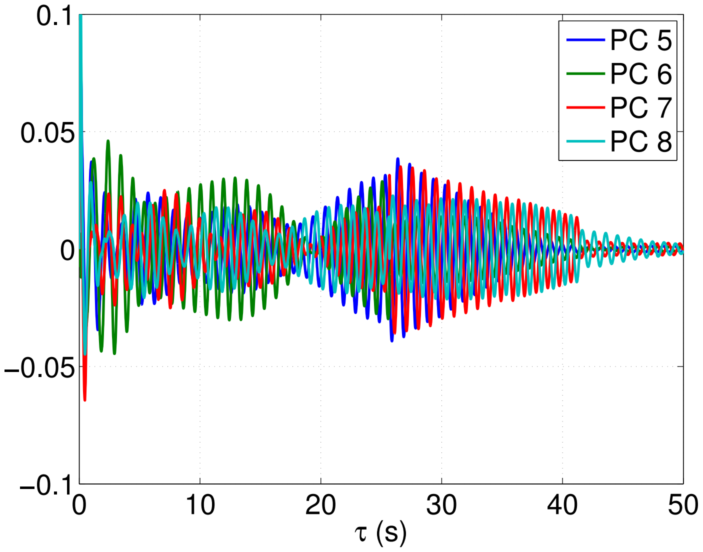
(b) Warped time scale τ

Figure 4.16 – Duffing oscillator – $N = 50$ different trajectories of the response in the original and warped time scales.

PCA is applied to retrieve eight principal components that results in the PCA truncation error smaller than 1×10^{-3}. Figure 4.17 depicts the retained principal components. Figure 4.18 shows the upper bound of the surrogate model versus the number of principal components, which starts converging at the eighth component. The relative LOO errors of PCE models for the first two components are 8×10^{-4} and 4×10^{-3}, respectively.



(a) Principal components 1-4



(b) Principal components 5-8

Figure 4.17 – Duffing oscillator – The first eight principal components.

An independent validation set of $10,000$ runs is used to judge the accuracy of the PCE models. Figure 4.19 presents two specific realizations of the displacement $y(t)$ obtained with two distinct sets of parameters $(\zeta, \omega, \epsilon)$. Without time-warping, PCEs are capable of predicting the response at the early time instants ($t < 3\ s$), then their accuracies degenerate with time, resulting in incorrect predictions. By introducing the time-warping of the trajectories, PCEs can faithfully capture the damped oscillatory behaviour. Only 0.18% of predictions exhibits a relative error exceeding 0.1. Note that an experimental design of size 200 is used for time-frozen PCEs, whereas only 50 trajectories are used for computing time-warping PCEs. This emphasizes the fact that the time-warping pre-processing of the response allows one to build accurate PCEs at a relatively small computational cost.

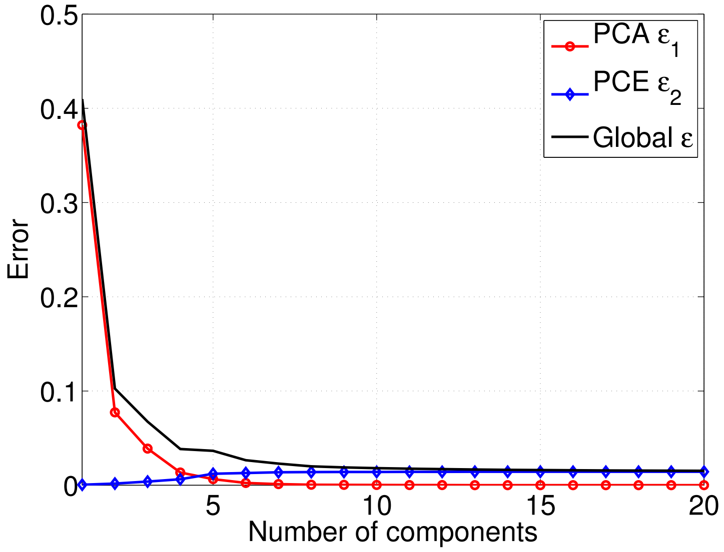
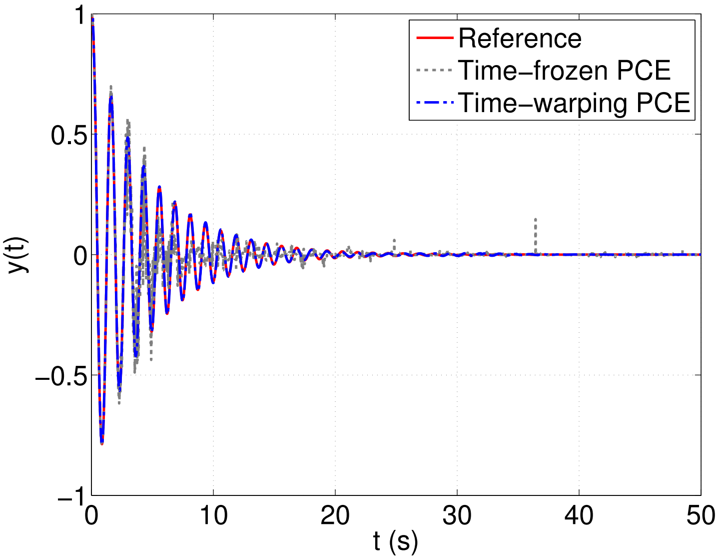
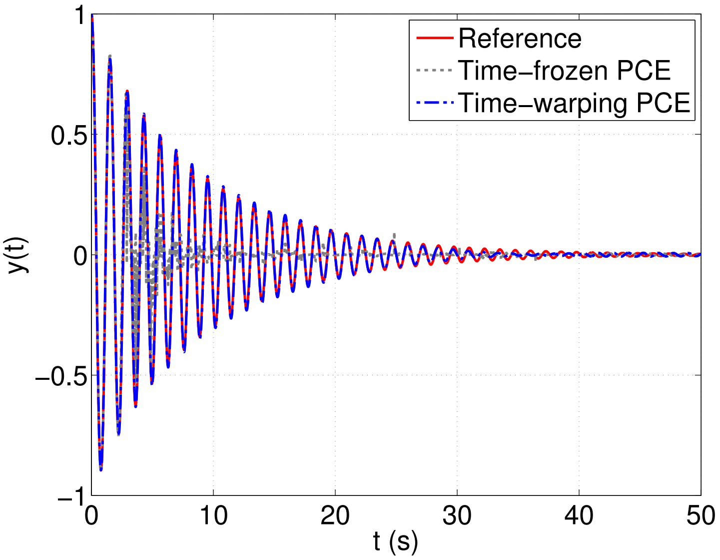


Figure 4.18 – Duffing oscillator – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace$(\tilde{\boldsymbol{\Sigma}})$ and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

In terms of time dependent statistics (Figure 4.20), time-frozen PCEs can predict rather well the mean trajectory, however fail to represent the standard deviation after early instants ($t > 3\ s$). In contrast the time-warping approach provides excellent accuracy on the mean and standard deviation time histories. The relative discrepancies between mean and standard deviation time histories predicted by time-warping PCEs with the reference trajectories are respectively 3.27×10^{-5} and 3.47×10^{-4}.

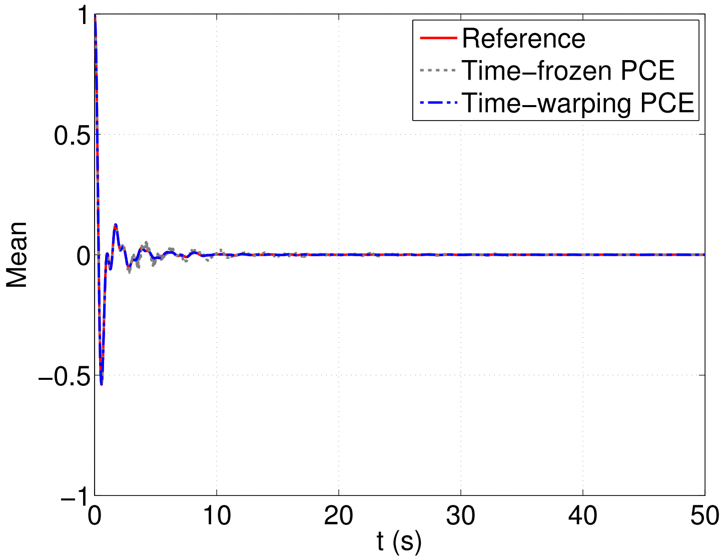


(a) $(\zeta, \omega, \epsilon) = (0.0403, 5.0455, -0.7186)$

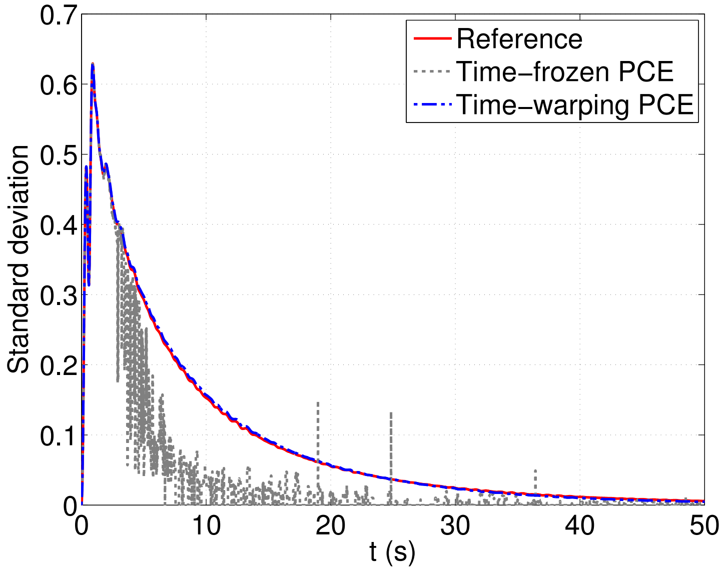


(b) $(\zeta, \omega, \epsilon) = (0.0222, 4.9974, -0.5007)$

Figure 4.19 – Duffing oscillator – Two particular trajectories and their predictions by time-frozen and time-warping PCEs.



(a) Mean trajectory



(b) Standard deviation

Figure 4.20 – Duffing oscillator – Mean and standard deviation of the trajectories: comparison of the two approaches.

4.2.3 Oregonator model

We consider now the Oregonator model which describes the dynamics of a well-stirred, homogeneous chemical system governed by a three species coupled mechanism. Note that this benchmark problem was used by Le Maître et al. (2010) for illustrating the intrusive time-transform approach. This chemical system undergoes an oscillation governed by the following ordinary differential equations (ODEs):

$$\begin{cases} \dot{x}(t) = k_1\, y(t) - k_2\, x(t)\, y(t) + k_3\, x(t) - k_4\, x(t)^2, \\ \dot{y}(t) = -k_1\, y(t) - k_2\, x(t)\, y(t) + k_5\, z(t), \\ \dot{z}(t) = k_3\, x(t) - k_5\, z(t), \end{cases} \tag{4.39}$$

in which (x, y, z) denotes the three species concentration and the coefficients k_i, $i = 1, \ldots, 5$ are the reaction parameters. Hereafter, all the reaction parameters are considered independent random variables with uniform and normal distributions (see Table 4.2). It is worth noting that Le Maître et al. (2010) considered only k_4 and k_5 as uniform random variables while fixing the remaining parameters ($k_1 = 2$, $k_2 = 0.1$, $k_3 = 104$). The initial condition is $(x_0, y_0, z_0) = (6000, 6000, 6000)$, which corresponds to a deterministic mixture. We aim at building PCEs of the concentration $x(t)$ as a function of the random parameters (k_1, k_2, k_3, k_4, k_5).

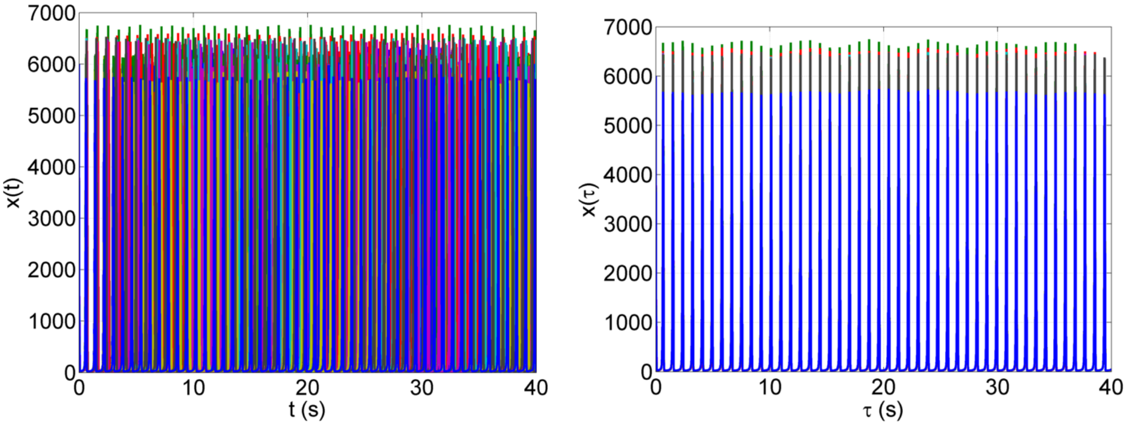
Table 4.2 – Reaction parameters of the Oregonator model

Input parameters	Distribution	Values
k_1	Uniform	$\mathcal{U}[1.8,\ 2.2]$
k_2	Uniform	$\mathcal{U}[0.095,\ 0.1005]$
k_3	Gaussian	$\mathcal{N}(104, 1.04)$
k_4	Uniform	$\mathcal{U}[0.0076, 0.0084]$
k_5	Uniform	$\mathcal{U}[23.4, 28.6]$

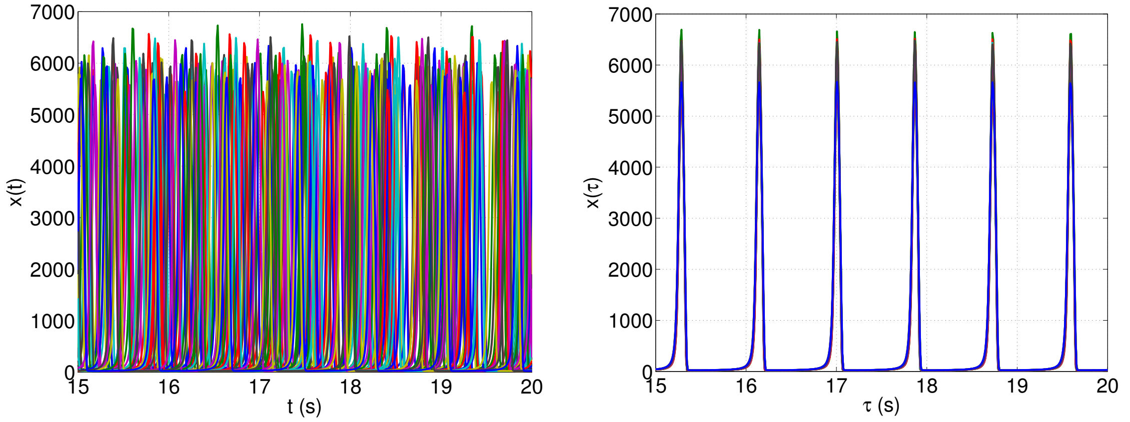
Figure 4.21 depicts 50 trajectories among 500 realizations of $x(t)$, which are used as the experimental design for fitting time-frozen PCEs. One notices that after 5 seconds, the different trajectories are completely out-of-phase. Time-frozen PCEs of total degree up to 20 are used. The corresponding LOO error (Figure 4.22) exhibits a relatively fast escalation. The PCE model actually starts degenerating at $t = 3\ s$. In particular, time-frozen PCE predictions are negative, *i.e.* non physical, at some instants (see Figure 4.25).

We now apply the proposed non-intrusive time-warping approach to this problem. Note that only 50 trajectories of $x(t)$ are used as an experimental design. A linear time-transform $\tau = k\,t + \phi$ is again utilized. The parameters k and ϕ are determined and adaptive PCEs of k and ϕ are then computed. The relative LOO errors of the PCE models for k and ϕ are respectively 4.42×10^{-5} and 4.8×10^{-2}, which indicate a high accuracy. The response trajectories are interpolated into the transformed time line τ (Figure 4.21) and adaptive sparse PCEs of degree up to 20 combined with PCA are then used.

Using PCA, 18 components are retained to obtain a truncation error ϵ_1 smaller than



(a) Original time scale t (b) Warped time scale τ



(c) Original time scale t (zoom) (d) Warped time scale τ (zoom)

Figure 4.21 – Oregonator model – $N = 50$ different trajectories of the response. The figures in the first row are zoomed in the range $[15, 20]$ to obtain the figures in the second row.

1×10^{-2}. The first eight components are depicted in Figure 4.23. Figure 4.24 shows the convergence of the errors with respect to the number of principal components. The PCEs for the first two coefficients have relative errors 7.57×10^{-4} and 1.5×10^{-3}, respectively.

Figure 4.25 depicts two particular realizations computed by the numerical solver (Matlab ordinary differential equation solver `ode45`, using a time step $\Delta_t = 0.01$ for the total duration $T = 40$ s) and predictions by PCEs with and without time-warping. It is shown that without time-warping, PCEs fail to capture the oscillatory behaviour of the response. In contrast, the use of a time-warping allows PCE to predict the response with great accuracy. 1.24% of the predictions (among $10,000$ samples) has a relative error larger than 0.1.

A validation set of $10,000$ trajectories is used to get reference values of mean and standard deviation of the concentration $x(t)$. Figure 4.26 depicts the statistics of $x(t)$ predicted by time-frozen and time-warping PCEs. Without time-warping, the estimates by PCEs differ significantly from the reference trajectories already from 3 s. The discrep-

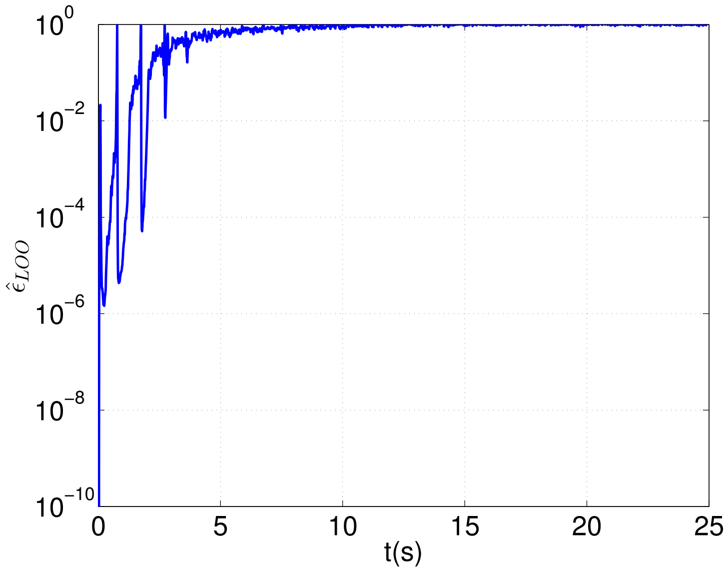
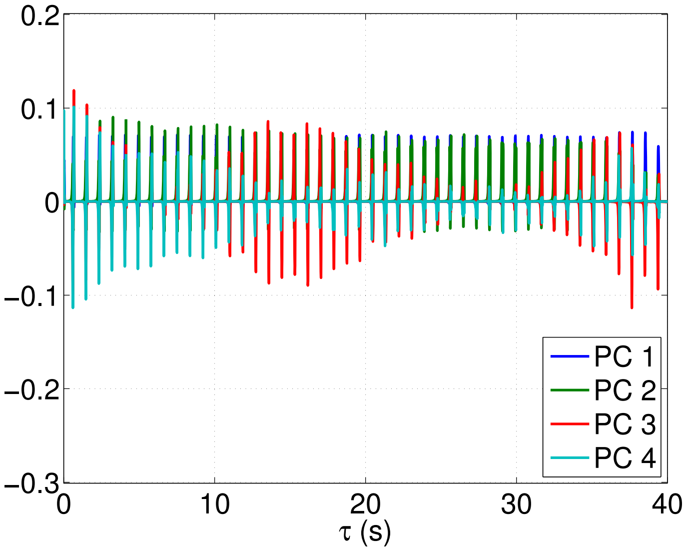


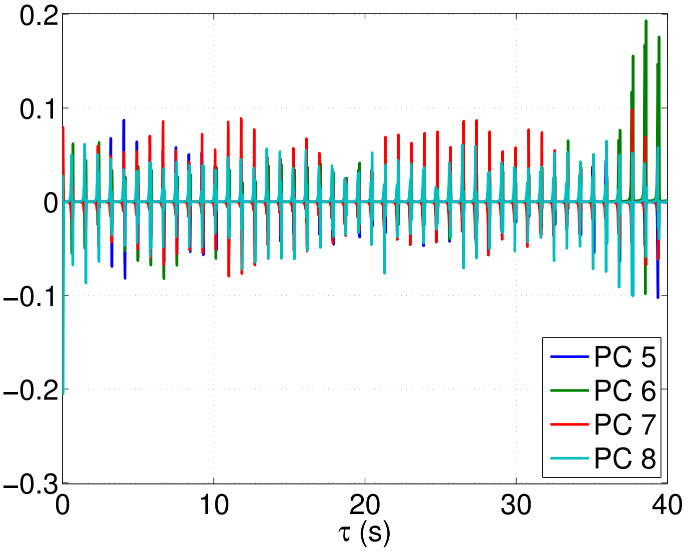
Figure 4.22 – Oregonator model – Leave-one-out error $\widehat{\epsilon}_{LOO}$ of time-frozen PCEs.

ancies then quickly increase in time. For instance, PCEs without time-warping estimate a decreasing trend in time for the standard deviation, whereas the latter actually oscillates around a constant value (around 1400) with high frequency. By introducing the time-warping pre-processing, one can use sparse PCEs to capture the complex behaviour of the time dependent statistics of the response all along the trajectories. The relative error for the mean and standard deviation trajectories are 3.11×10^{-4} and 3.6×10^{-3}, respectively.

Finally, the time-warping PCE scheme is applied to surrogate the responses $y(t)$ and $z(t)$ of the system using the same experimental design of size 50 and the same procedure. Figure 4.27 shows a great agreement between two specific trajectories, the mean and standard deviation of (x, y, z) in the state-space predicted by time-warping PCEs and the reference functions.



(a) Principal components 1-4



(b) Principal components 5-8

Figure 4.23 – Oregonator model – The first eight principal components.

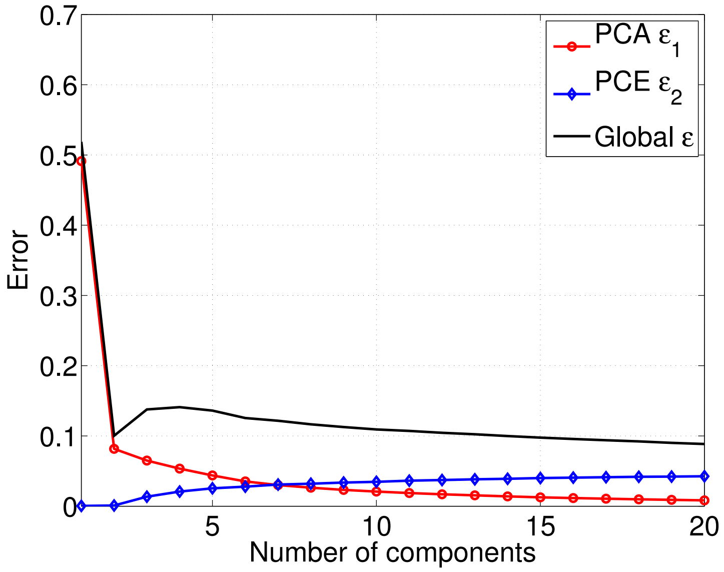
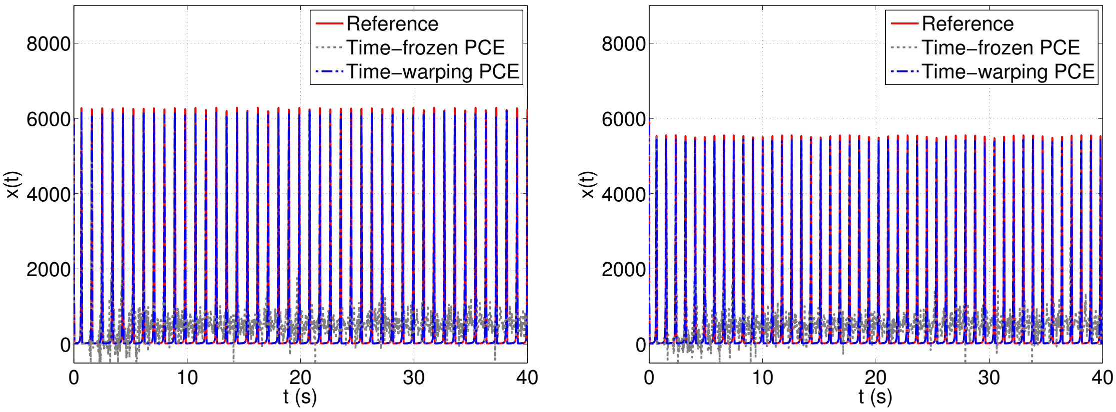
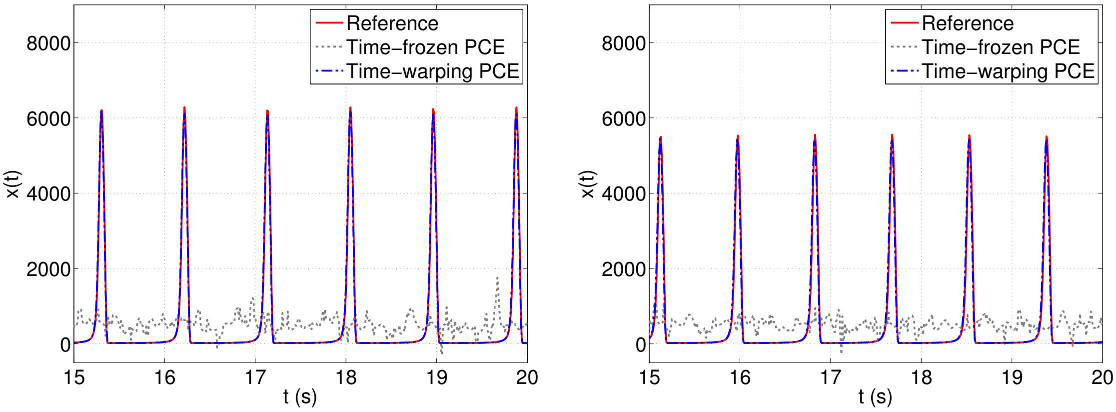


Figure 4.24 – Oregonator model – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace($\tilde{\mathbf{\Sigma}}$) and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

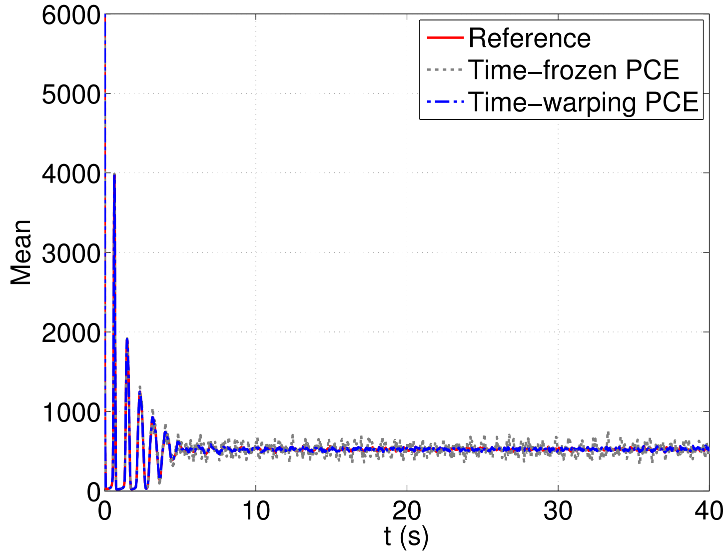


(a) $\boldsymbol{\xi} = (1.8970, 0.1001, 104.3676, 0.0077, 25.5417)$ (b) $\boldsymbol{\xi} = (1.9481,\ 0.0999,\ 102.7929,\ 0.0080,\ 27.6482)$

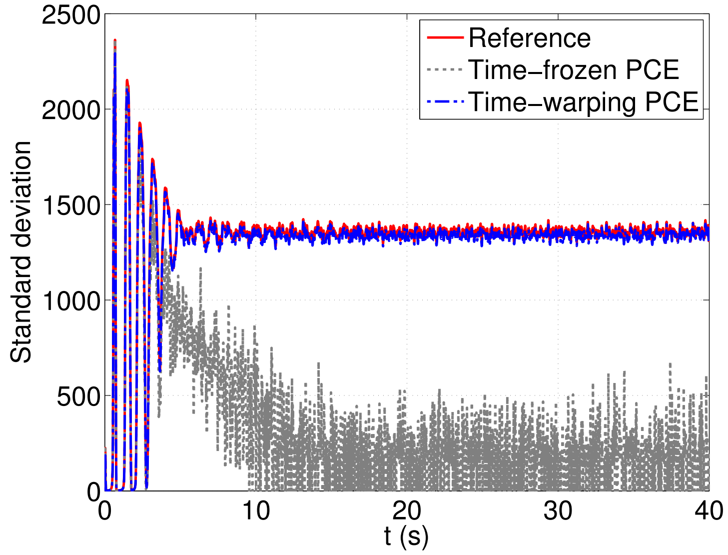


(c) $\boldsymbol{\xi} = (1.8970, 0.1001, 104.3676, 0.0077, 25.5417)$ (d) $\boldsymbol{\xi} = (1.9481,\ 0.0999,\ 102.7929,\ 0.0080,\ 27.6482)$

Figure 4.25 – Oregonator model – Two particular trajectories and their predictions by time-frozen and time-warping PCEs. The figures in the first row are zoomed in the range $[15, 20]$ to obtain the figures in the second row.

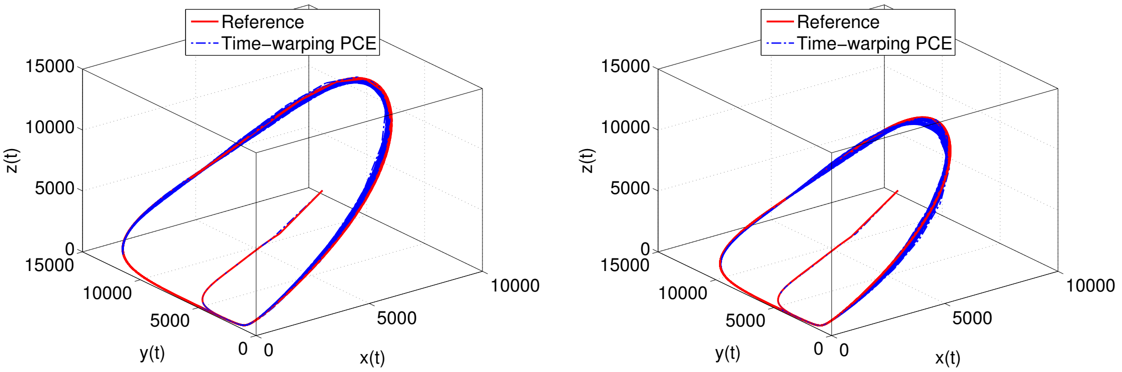


(a) Mean trajectory

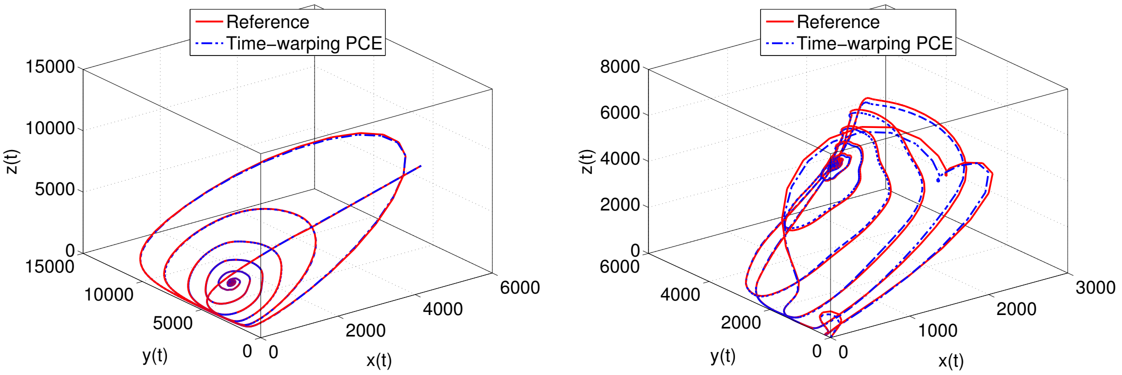


(b) Standard deviation

Figure 4.26 – Oregonator model – Mean and standard deviation of the trajectories: comparison of the two approaches.



(a) $\boldsymbol{\xi} = (1.8970, 0.1001, 104.3676, 0.0077, 25.5417)$ (b) $\boldsymbol{\xi} = (1.9481,\ 0.0999,\ 102.7929,\ 0.0080,\ 27.6482)$



(c) Mean trajectory (d) Standard deviation trajectory

Figure 4.27 – Oregonator model – Trajectories of $(x(t), y(t), z(t))$ predicted by time-warping PCEs *vs.* the reference trajectories.

4.2.4 Forced vibration of a Bouc-Wen oscillator

In the previous case studies, self-oscillating systems were considered. In general, the proposed approach is also applicable to forced vibration systems. Let us now consider the SDOF Bouc-Wen oscillator (Kafali and Grigoriu, 2007) subject to a stochastic excitation. The equation of motion of the oscillator reads:

$$\begin{cases} \ddot{y}(t) + 2\,\zeta\,\omega\,\dot{y}(t) + \omega^2(\rho\,y(t) + (1-\rho)\,z(t)) = -x(t), \\ \dot{z}(t) = \gamma\dot{y}(t) - \alpha\,|\dot{y}(t)|\,|z(t)|^{n-1}\,z(t) - \beta\,\dot{y}(t)\,|z(t)|^{n}, \end{cases} \tag{4.40}$$

in which ζ is the damping ratio, ω is the fundamental frequency, ρ is the post- to pre-yield stiffness ratio, γ, α, β, n are parameters governing the hysteretic loops and the excitation $x(t)$ is a sinusoidal function given by $x(t) = A\,\sin(\omega_x\,t)$.

Deterministic values are used for the following parameters of the Bouc-Wen model: $\rho = 0$, $\gamma = 1$, $n = 1$, $\beta = 0$. The remaining parameters are considered independent random variables with associated distributions given in Table 4.3.

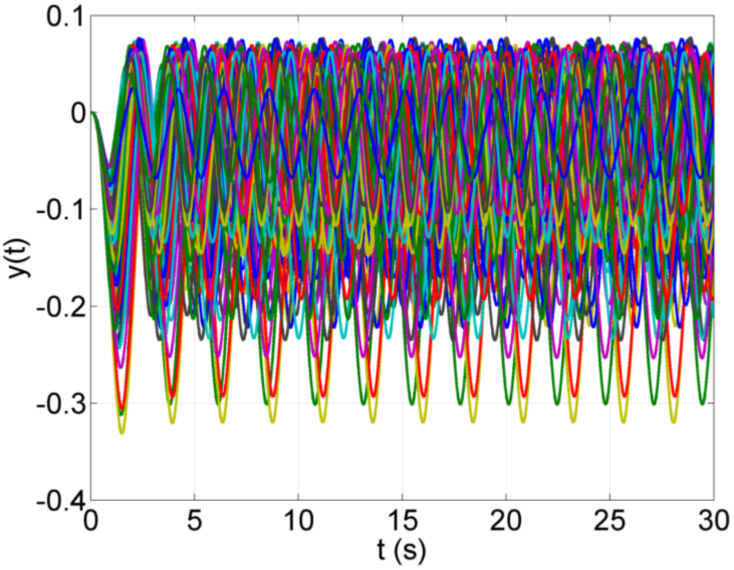
Table 4.3 – Uncertain parameters of the Bouc-Wen model

Input parameters	Distribution	Mean	Standard deviation
ζ	Uniform	0.02	0.002
ω	Uniform	$2\,\pi$	$0.2\,\pi$
α	Uniform	50	5
A	Uniform	1	0.1
ω_x	Uniform	π	$0.1\,\pi$

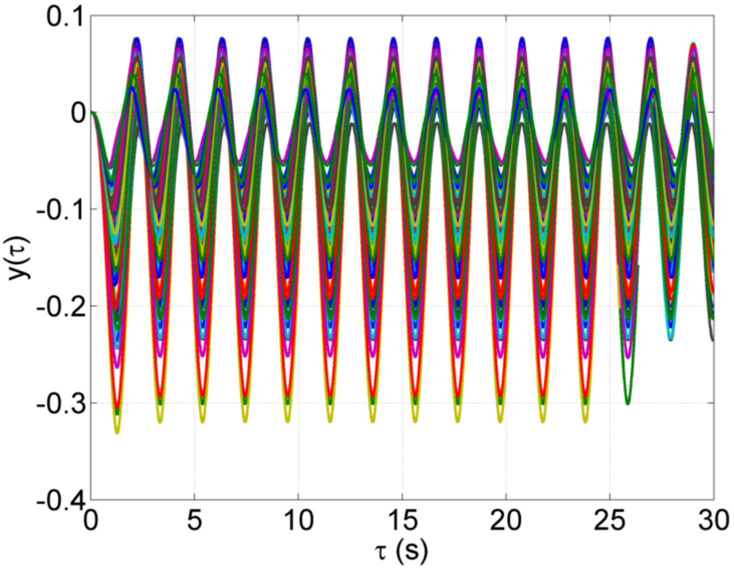
One aims at representing the oscillator displacement $y(t)$ as a function of the uncertain input parameters using time-warping PCEs. To this end, 100 simulations of the oscillator are carried out using the Matlab solver `ode45` with time increment $\Delta_t = 0.005$ s for the total duration $T = 30$ s and initial condition $y(t=0) = 0$, $\dot{y}(t=0) = 0$. The displacement trajectories are depicted in Figure 4.28(a).

For this case study, a time-warping scheme $\tau = k\,t$ with only one parameter is used. After the time-warping process, the trajectories become in-phase as depicted in Figure 4.28(b). Adaptive sparse PCE representing k has the relative LOO error 5×10^{-5}. In order to achieve a truncation error ϵ_1 smaller than 1×10^{-3}, 13 first principal components are retained in PCA. Figure 4.29 depicts the first eight principal components. The relative LOO errors of PCEs for the first two components are 6×10^{-3} and 6.21×10^{-2}, respectively. Figure 4.30 shows the errors induced by PCA and PCE as a function of the number of retained principal components.

Let us validate the accuracy of the time-warping PCE model. In Figure 4.31, two specific predictions of the PCE model are plotted against the actual responses obtained with the original Matlab solver. A remarkable agreement can be observed. Among 10,000 validations, only 4.87% has a relative error larger than 0.1. Regarding the time dependent mean and standard deviation of the oscillator, time-warping PCE-based estimates



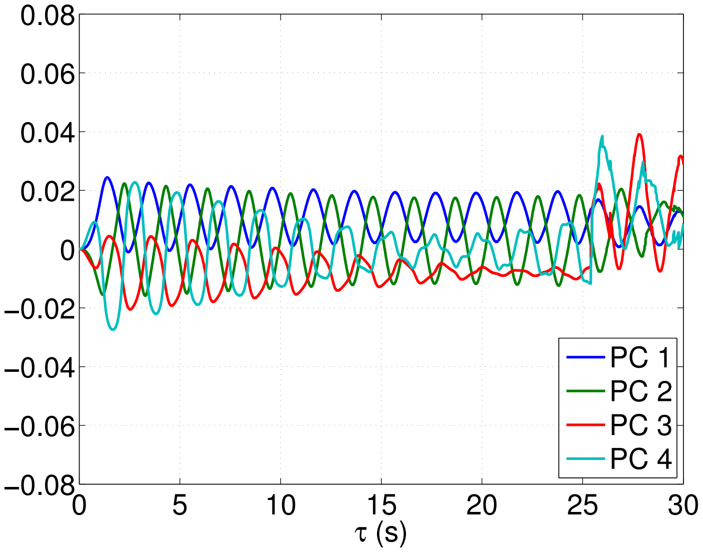
(a) Original time scale t



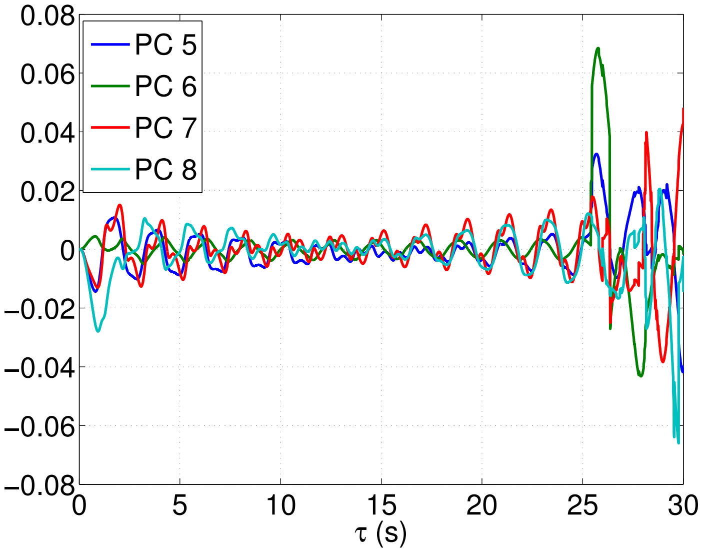
(b) Warped time scale τ

Figure 4.28 – Bouc-Wen oscillator – $N = 100$ different trajectories of the solution in the original time scale t and in the transformed time line τ.

outstandingly match the reference trajectories (Figure 4.32). Only a minor discrepancy can be observed at the end of the considered time duration $T = 30$ s, which is due to the modest number of simulations used as the experimental design. The corresponding relative errors are both 2.4×10^{-3}, respectively



(a) Principal components 1-4



(b) Principal components 5-8

Figure 4.29 – Bouc-Wen oscillator – The first eight principal components.

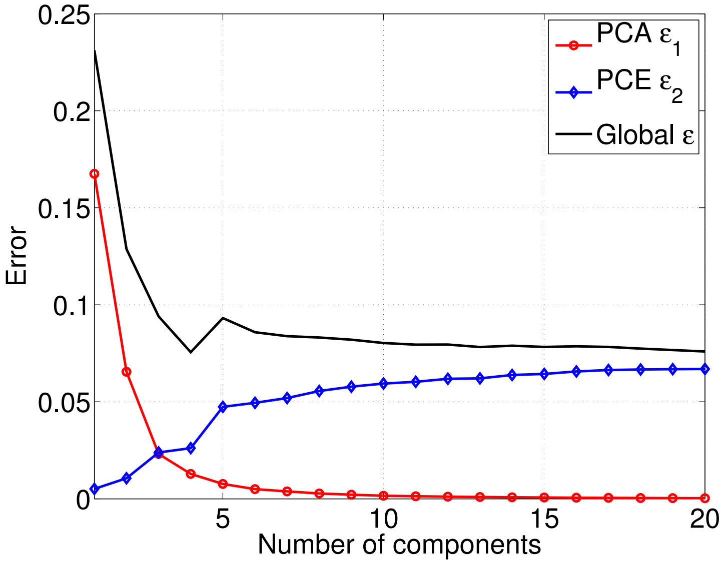
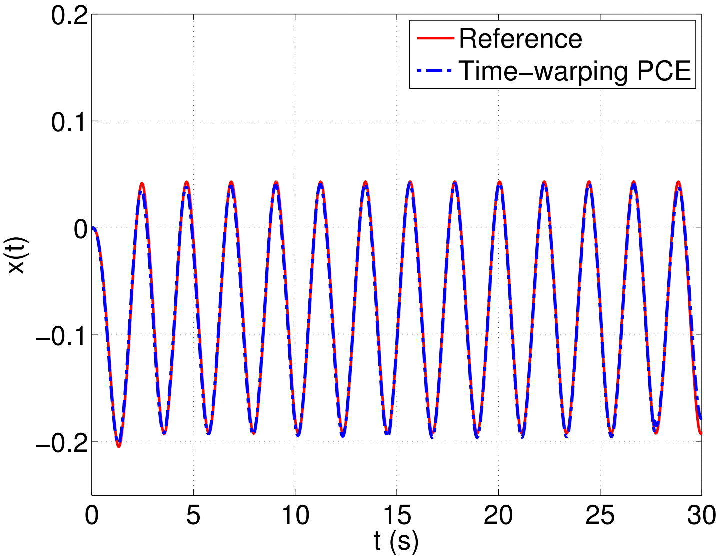
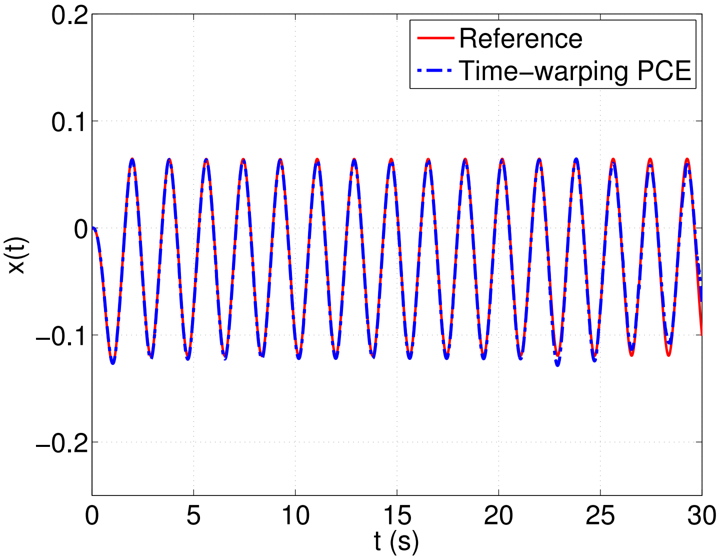


Figure 4.30 – Bouc-Wen oscillator – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace($\tilde{\boldsymbol{\Sigma}}$) and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

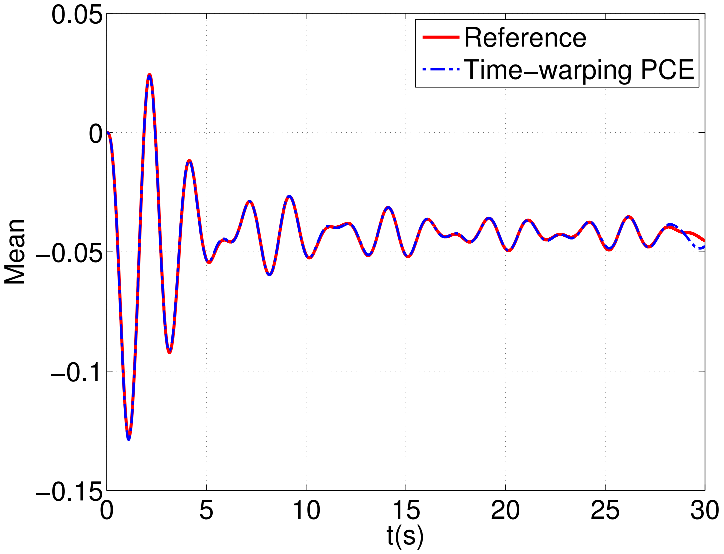


(a) $\boldsymbol{\xi} = (0.0191, 5.6208, 57.3581, 0.9401, 2.8577)$

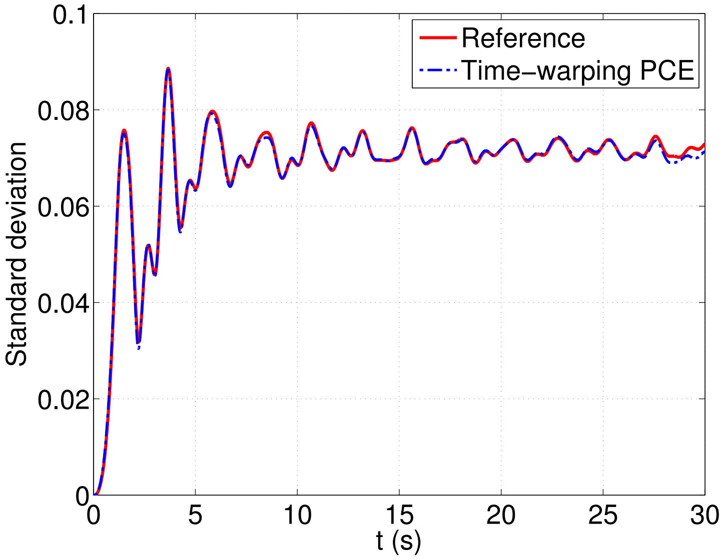


(b) $\boldsymbol{\xi} = (0.0196, 6.1226, 46.9916, 1.0291, 3.4542)$

Figure 4.31 – Bouc-Wen oscillator – Two particular trajectories and their predictions by time-warping PCEs.



(a) Mean trajectory



(b) Standard deviation

Figure 4.32 – Bouc-Wen oscillator – Mean and standard deviation of the trajectories: comparison of time-warping PCE estimates and the reference trajectories.

4.2.5 Kraichnan-Orszag model: 1D case

In the previous numerical applications, we considered the systems with uncertain governing parameters subject to deterministic initial conditions. Hereafter, let us investigate the dynamical systems with random initial conditions, *e.g.* the so-called Kraichnan-Orszag three-mode problem. It was introduced by Kraichnan (1963) to model a system of several interacting shear waves and was studied later by Orszag (1967) in the case of Gaussian initial conditions. This model is described by the following system of ODEs:

$$\begin{cases} \dot{x}(t) = y(t)\,z(t), \\ \dot{y}(t) = z(t)\,x(t), \\ \dot{z}(t) = -2\,x(t)\,y(t). \end{cases} \tag{4.41}$$

In the first place, we consider the 1D case which is a benchmark problem used to illustrate the time-dependent PCE approach (Gerritsma et al., 2010). The initial condition of $x(t)$ is considered stochastic, *i.e.* $x(t=0) = \alpha + 0.01\,\xi$ with $\xi \sim \mathcal{U}[-1,1]$ whereas $y(t=0) = 1.0$, $z(t=0) = 1.0$. Note that when α is in the range $[0, 0.9]$, the responses are insensitive to the initial conditions. For $\alpha \in [0.9, 1]$, there is a strong dependence of the responses on the initial conditions. Figure 4.33 depicts the large discrepancies between time-histories of $x(t)$ due to a minor variability of the initial condition $x(t=0)$. Herein, we consider $\alpha = 0.99$ as in Gerritsma et al. (2010).

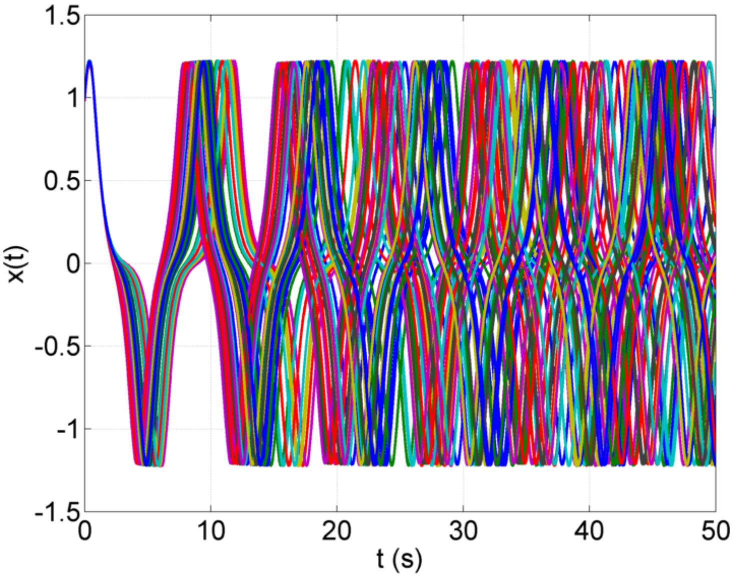


Figure 4.33 – Kraichnan-Orszag model – $N = 50$ different trajectories $x(t)$ in the original time scale t.

The surrogate model of the response $x(t)$ is computed with time-warping PCEs using an experimental design of size $N = 50$ (Figure 4.33). Herein, a time-warping scheme $\tau = k\,t$ with one governing parameter is used. The trajectories resulting from the time-warping process are depicted in Figure 4.34. The adaptive sparse PCE representing k has a relative LOO error equal to 2.2×10^{-6}. 13 first principal components are retained so that 99.9% of the response's variance is explained. Figure 4.35 depicts the first four principal components. The relative LOO errors of PCEs for the first two components are 9.4×10^{-5} and 7×10^{-3}, respectively. Figure 4.36 shows the errors induced by PCA and PCE as a function of the number of retained principal components.

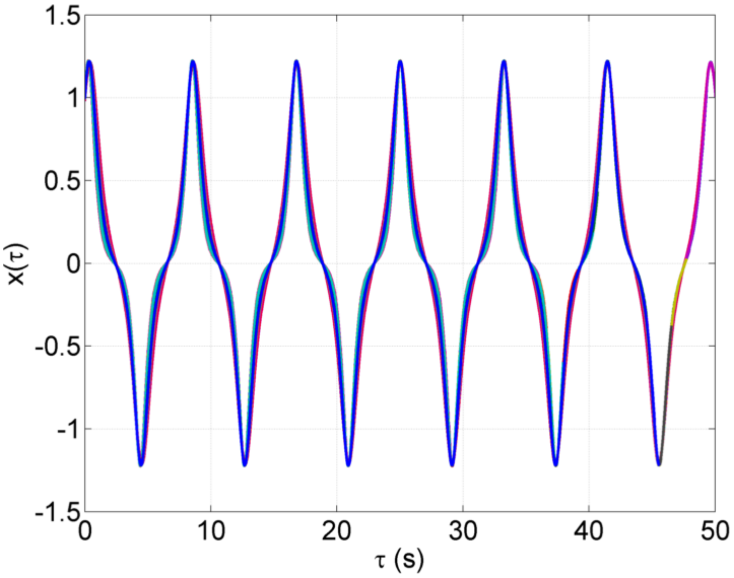


Figure 4.34 – Kraichnan-Orszag model – $N = 50$ different trajectories $x(\tau)$ in the warped time scale τ

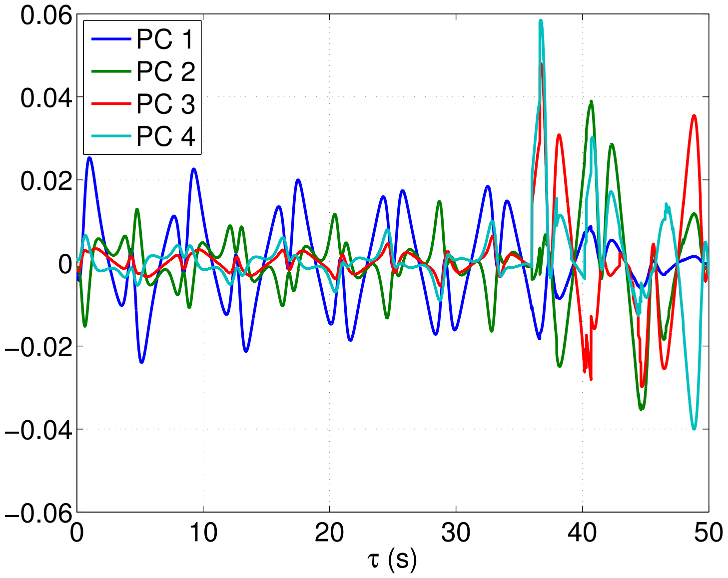


Figure 4.35 – Kraichnan-Orszag model – The first four principal components.

The time-warping PCE model is then validated by accessing the accuracy of its predictions. Figure 4.37 plots two specific predictions of the surrogate model which are graphically indistinguishable from the actual time-histories obtained with the original

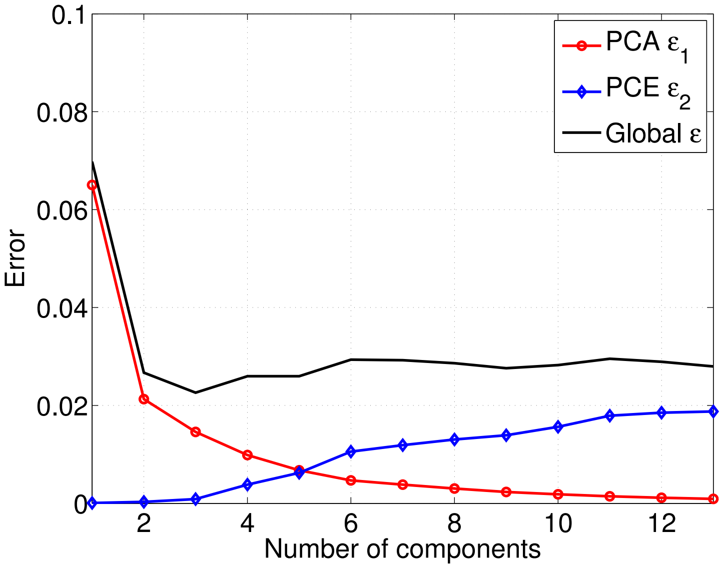
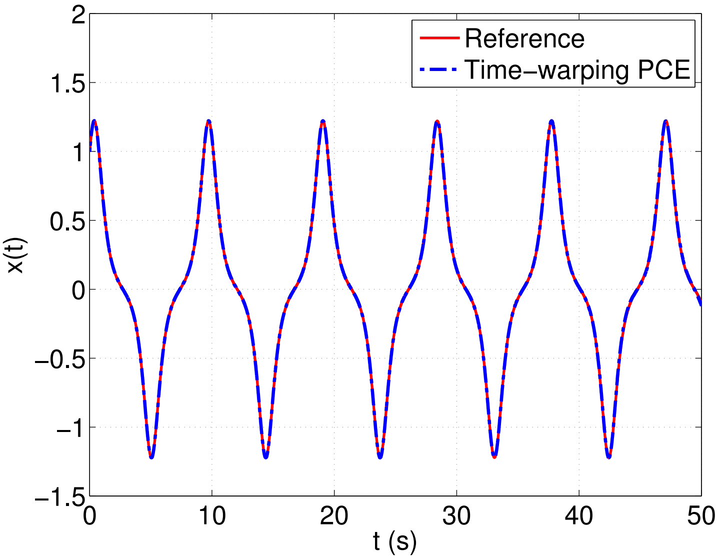


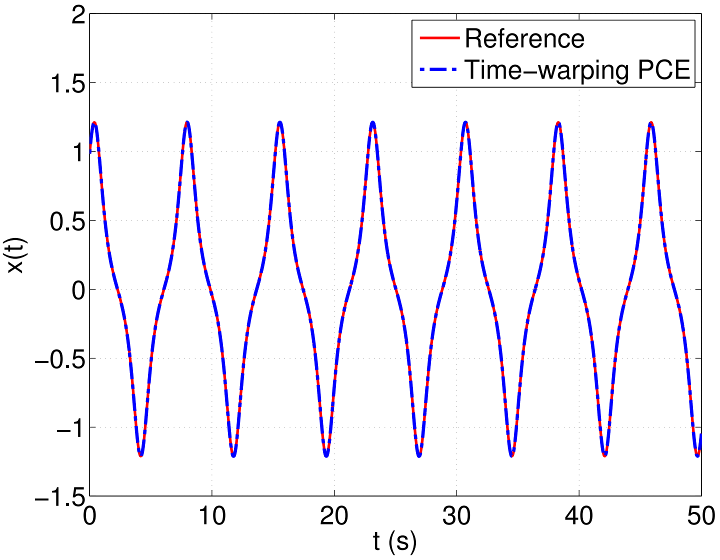
Figure 4.36 – Kraichnan-Orszag model – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace($\tilde{\mathbf{\Sigma}}$) and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

Matlab solver. Only 1.27% of the predictions experience a relative error larger than 0.1. Regarding the mean and standard deviation trajectories (Figure 4.38), the time-warping approach leads to respective relative errors equal to 2.1×10^{-4} and 5.3×10^{-4}, which shows an excellent agreement between the predictions and the true functions.

This numerical example illustrates the potential application of the proposed time-warping approach to problems with uncertain initial conditions. The excellent performance of the approach is even more impressive given the chaotic behaviour of the considered system, *i.e.* the responses are strongly sensitive with respect to a minor variability of the initial condition.

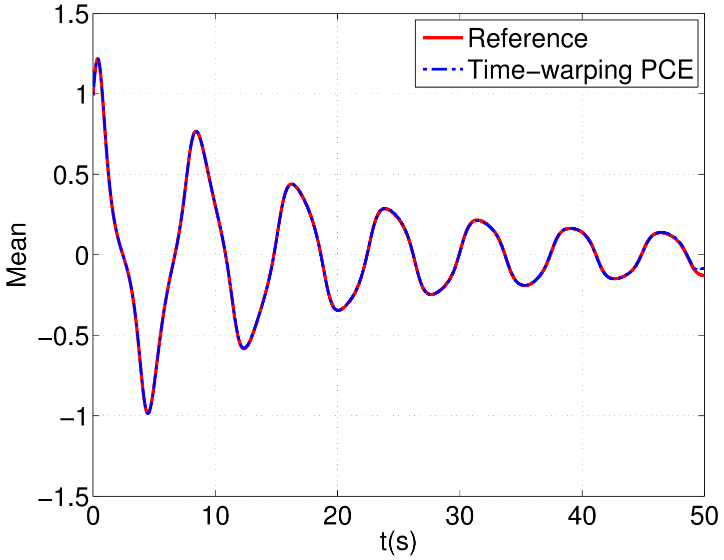


(a) $\xi = 0.6294$

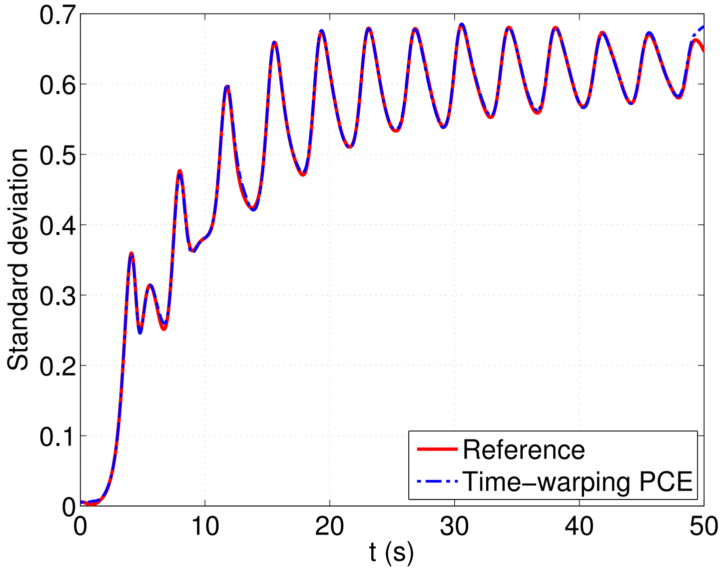


(b) $\xi = -0.7460$

Figure 4.37 – Kraichnan-Orszag model – Two particular trajectories and their predictions by time-warping PCEs.



(a) Mean trajectory



(b) Standard deviation

Figure 4.38 – Kraichnan-Orszag model – Mean and standard deviation of the trajectories: comparison of time-warping PCE estimates and the reference trajectories.

4.2.6 Kraichnan-Orszag model: 2D case

Let us consider the transformed Kraichnan-Orszag model which was investigated by Wan and Karniadakis (2005) and Guo et al. (2016). the system of ODEs reads:

$$\begin{cases} \dot{x}(t) = x(t)\, z(t), \\ \dot{y}(t) = -y(t)\, z(t), \\ \dot{z}(t) = -x(t)^2 + y(t)^2, \end{cases} \tag{4.42}$$

in which the initial conditions are considered stochastic: $x(t=0) = 1$, $y(t=0) = 0.1\,\xi_1$ and $z(t=0) = \xi_2$ with $\xi_i \sim \mathcal{U}[-1,1]$, $i = 1,2$. For the present system, the stochastic simulation of the responses is challenging since the solution is not only sensitive to the initial conditions but also exhibits bifurcation on the parameters ξ_1 and ξ_2 (see Figure 4.39). To overcome the singularity associated with the bifurcation of the solution, Wan and Karniadakis (2005) proposed the use of multi-element PCEs whereas Guo et al. (2016) recently introduced the weighted essentially non-oscillatory collocation method. Herein, we aim at simulating the response $y(t)$ by time-warping PCEs using a small experimental design of size $N = 200$ (Figure 4.40).

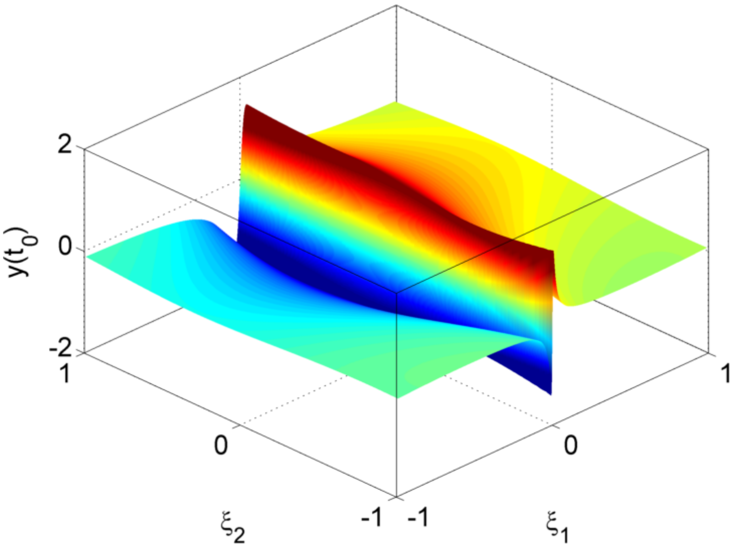


Figure 4.39 – Kraichnan-Orszag model – The solution $y(t_o)$ at $t_o = 6$ s on the random space. The bifurcation at $\xi_1 = 0$ is visible.

For this problem, the time-warping scheme $\tau = k\,t$ is utilized. Figure 4.41 depicts the resulting in-phase trajectories. It shows two ensembles of trajectories, which are symmetric with respect to the line $y(t) = 0$. They correspond to the initial condition $y(t=0) < 0$ and $y(t=0) > 0$, thus being a consequence of the bifurcation under study. The parameter k is modelled by an adaptive sparse PCE with a relative LOO error 2×10^{-3}. To explain 99.9% of the response's variance, we retained the first seven principal components, four of them are depicted in Figure 4.42. The relative LOO errors of PCEs for the first two components are equal to 1.59×10^{-2} and 3.91×10^{-2}, respectively. The errors induced by combining PCA with time-warping PCE are plotted in Figure 4.43

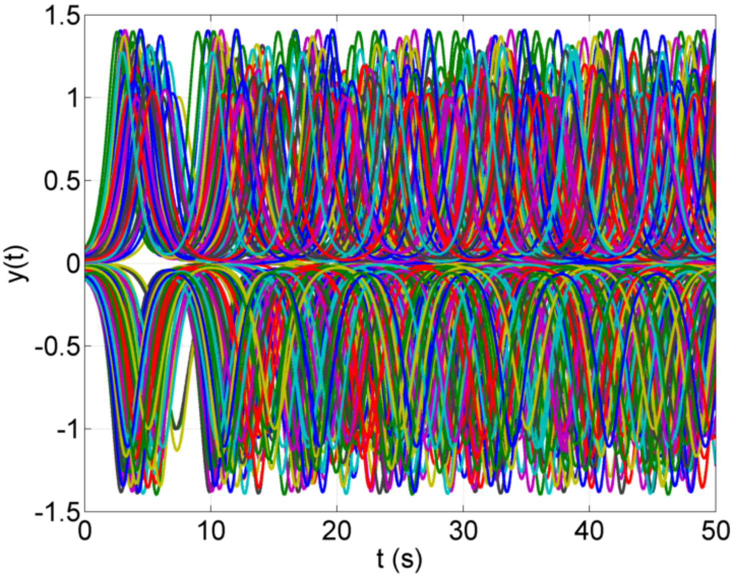


Figure 4.40 – Kraichnan-Orszag model – $N = 200$ different trajectories $y(t)$ in the original time scale t.

as a function of the number of retrieved principal components.

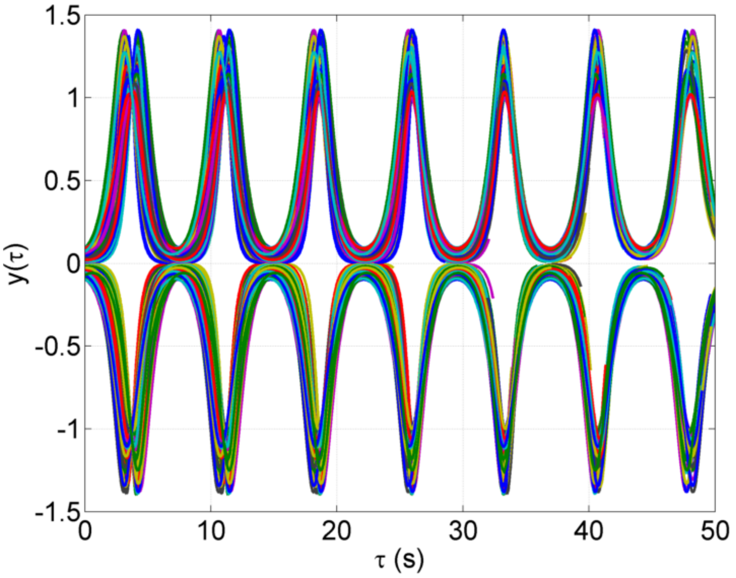


Figure 4.41 – Kraichnan-Orszag model – $N = 200$ different trajectories $y(\tau)$ in the warped time scale τ

The time-warping PCE model is then used to predict the solutions to new values of the initial conditions. Two specific predictions of the surrogate model are plotted against the actual time-histories obtained with the original Matlab solver in Figure 4.44. There are non-negligible discrepancies between the predictions and the actual solutions, with the corresponding relative errors 1.31×10^{-2} and 5.6×10^{-2}. Among $10,000$ runs, 26.22% has a relative error larger than 0.2.

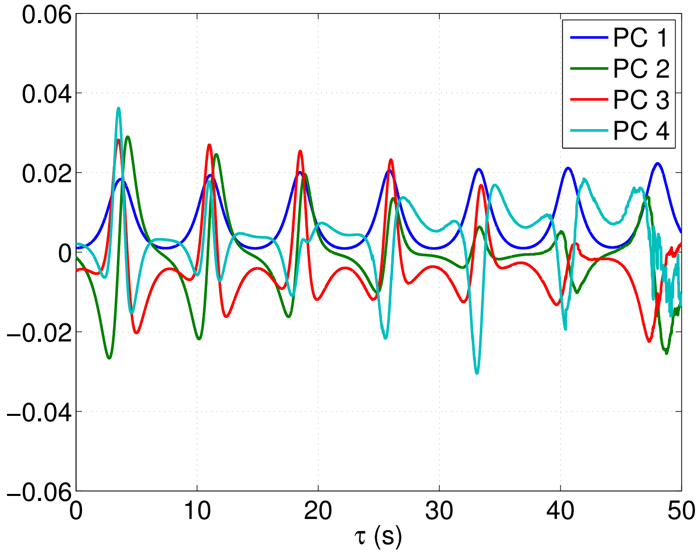


Figure 4.42 – Kraichnan-Orszag model – The first four principal components.

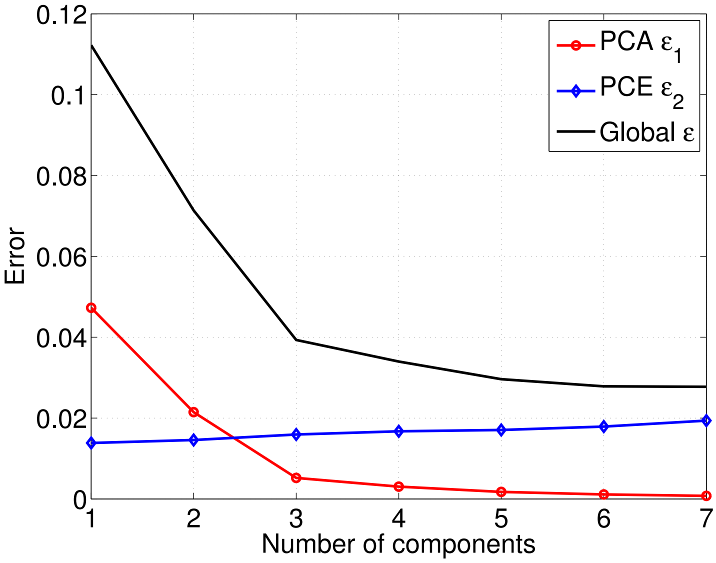
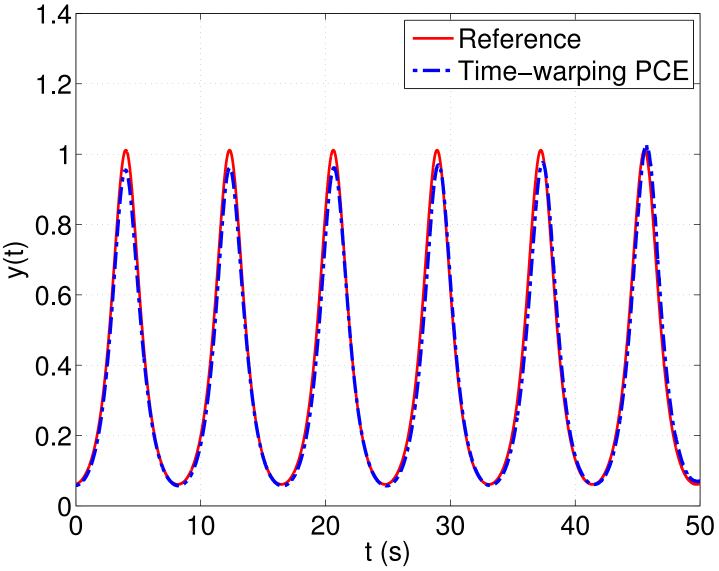
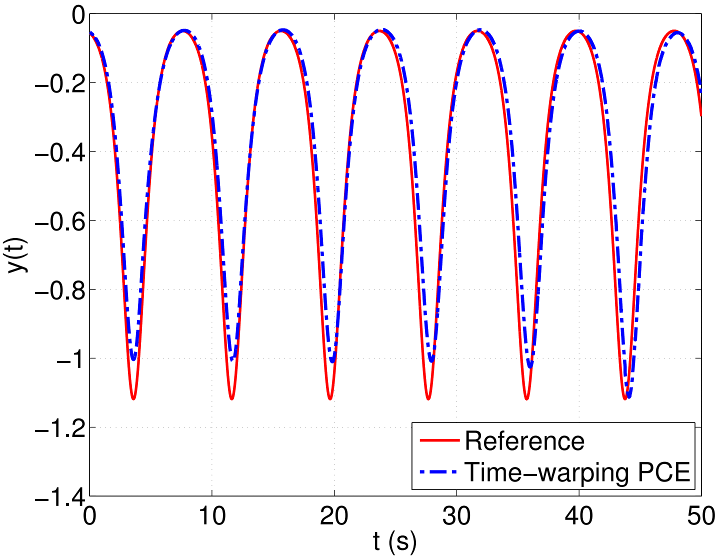


Figure 4.43 – Kraichnan-Orszag model – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace($\tilde{\boldsymbol{\Sigma}}$) and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

Figure 4.45 presents the predicted mean and standard deviation trajectories. Due to the bifurcation, the solution exhibits two different behaviours characterized by trajectories symmetric with respect to the line $y(t) = 0$, thus the mean trajectory is close to zero everywhere. The predicted standard deviation differ from the reference trajectory by a relative error 3.42×10^{-2}, which indicates a relatively good agreement. In order to obtain more accurate predictions with time-warping PCEs, one might tackle independently the two dynamical behaviours of the response associated with the initial condition $y(t = 0) < 0$ and $y(t = 0) > 0$. It is worth noting that herein the prediction is made for a long time period, *i.e.* $T = 50\ s$. In the works conducted by Wan and Karniadakis (2005) and Guo et al. (2016) only the first 10 seconds were taken into account.



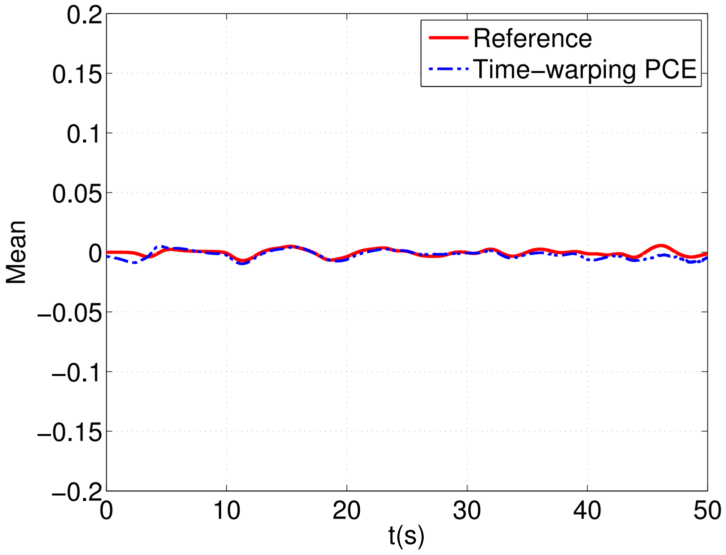
(a) $\boldsymbol{\xi} = (0.6142,\ -0.1502)$



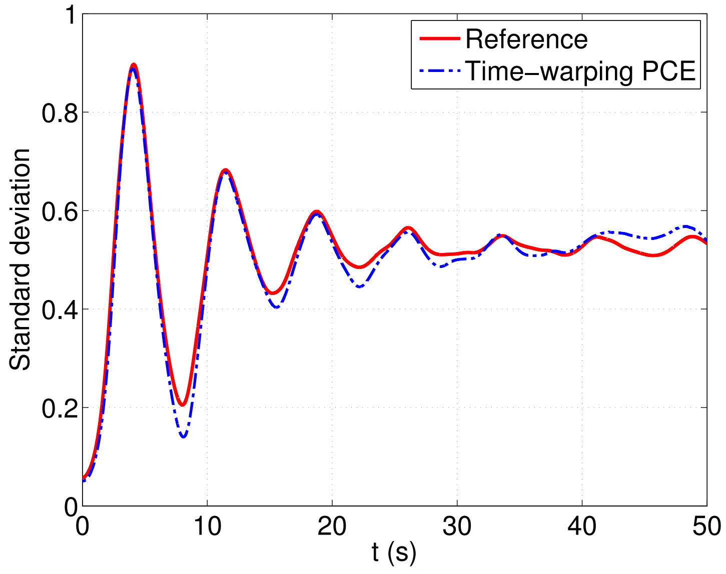
(b) $\boldsymbol{\xi} = (-0.5607,\ -0.5002)$

Figure 4.44 – KO model – Two particular trajectories and their predictions by time-warping PCEs.

As a summary, time-warping PCEs can potentially be applied to any problem with similar features, *i.e.* systems subject to uncertain initial conditions in which the solutions exhibit a bifurcation and/or chaotic behaviour. The resulting accuracy might not be outstanding, however, it remains acceptable in the context of engineering applications.



(a) Mean trajectory



(b) Standard deviation

Figure 4.45 – Kraichnan-Orszag model – Mean and standard deviation of the trajectories: comparison of time-warping PCE estimates and the reference trajectories.

4.2.7 Quarter car model

In the last numerical application, we investigate a quarter car model of a vehicle suspension whose behaviour is modelled by a nonlinear two DOF system (Kewlani et al., 2012) (Figure 4.46). The sprung mass m_s and the unsprung mass m_u are connected by a nonlinear spring of stiffness k_s and a linear damper of damping coefficient c. The forcing function $x(t)$ is applied to m_u through a linear spring of stiffness k_u. $y_1(t)$ and $y_2(t)$ are respectively the displacements of m_s and m_u, which are governed by the following system of ODEs:

$$\begin{cases} m_s\,\ddot{y}_1(t) = -k_s\,(y_1(t) - y_2(t))^3 - c\,(\dot{y}_1(t) - \dot{y}_2(t)), \\ m_u\,\ddot{y}_2(t) = k_s\,(y_1(t) - y_2(t))^3 + c\,(\dot{y}_1(t) - \dot{y}_2(t)) + k_u\,(x(t) - y_2(t)). \end{cases} \tag{4.43}$$

Herein, a sinusoidal function road profile with amplitude A and frequency ω is considered:

$$x(t) = A\,\sin(\omega\,t). \tag{4.44}$$

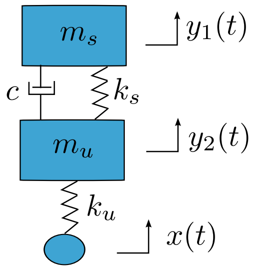


Figure 4.46 – Quarter car model – Mechanical system.

All the parameters of the quarter car model and of the excitation $\boldsymbol{\xi} = \{k_s, k_u, m_s, m_u, c, A, \omega\}$ are considered uncertain. They are modelled by independent random variables with marginal distributions given in Table 4.4. Note that Kewlani et al. (2012) addressed this numerical example with the multi-element PCE approach in which only two parameters k_s, k_u were considered uncertain and the solution was computed only for the first 6 seconds.

Table 4.4 – Parameters of the quarter car model and of the road excitation

Parameter	Distribution	Mean	Standard deviation
k_s (N/m^3)	Gaussian	2000	200
k_u (N/m)	Gaussian	2000	200
m_s (kg)	Gaussian	20	2
m_u (kg)	Gaussian	40	4
c (N s /m)	Gaussian	600	60
A (m)	Uniform	0.1	$0.01/\sqrt{3}$
ω (rad/s)	Uniform	$2\,\pi$	$0.2\,\pi/\sqrt{3}$

We aim at computing the surrogate model of the displacement $y_1(t)$ with time-warping PCEs using a small experimental design of size $N = 200$ (Figure 4.47). A time-warping scheme $\tau = k\,t$ with one governing parameter is specified, which results in trajectories in the transformed time line depicted in Figure 4.48. Parameter k is then modelled by an adaptive sparse PCE with the relative LOO error of 4.9×10^{-3}. Next, 26 first principal components are retained so that the truncation leads to a relative error smaller than 1×10^{-3}. The first four principal components are depicted in Figure 4.49. The relative LOO errors of PCEs for the first two components are respectively 0.40 and 0.224, which indicates indeed a relatively low level of accuracy. Figure 4.50 shows the errors induced by the truncation of PCA and the approximation with PCE as a function of the number of principal components kept by the algorithm.

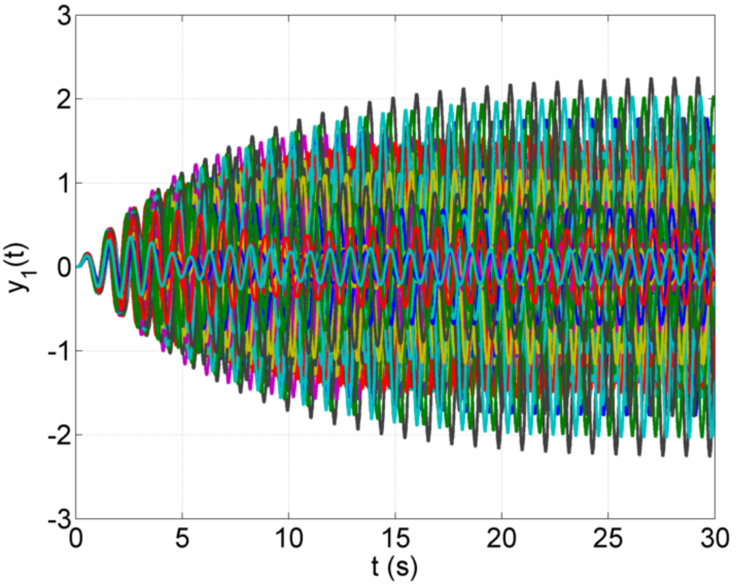


Figure 4.47 – Quarter car model – $N = 200$ different trajectories $y_1(t)$ in the original time scale t.

The time-warping PCE model is then validated by accessing the accuracy of its predictions. Two specific predictions of the surrogate model are plotted against the actual time-histories obtained with the original Matlab solver in Figure 4.51. The amplitudes of the predictions are in good agreement with the true amplitudes although the peaks are not perfectly captured. This is due to the fact that PCEs in the transformed time scale did not achieve a high accuracy. However, the predicted trajectories are strongly in-phase with the reference trajectories, which can be explained by the appropriate stochastic time-warping model obtained with PCEs. Among $10,000$ validations, 19.98% has a relative error larger than 0.5. Regarding the mean and standard deviation trajectories (Figure 4.52), the time-warping approach leads to corresponding relative errors of 5.7×10^{-3} and 8.4×10^{-3} which shows an excellent agreement between the predictions and the reference. Minor inaccuracy of the standard deviation is observed at the late instants, which can be removed by introducing additional samples in the experimental design.

This example showcases the application of the time-warping approach in high-

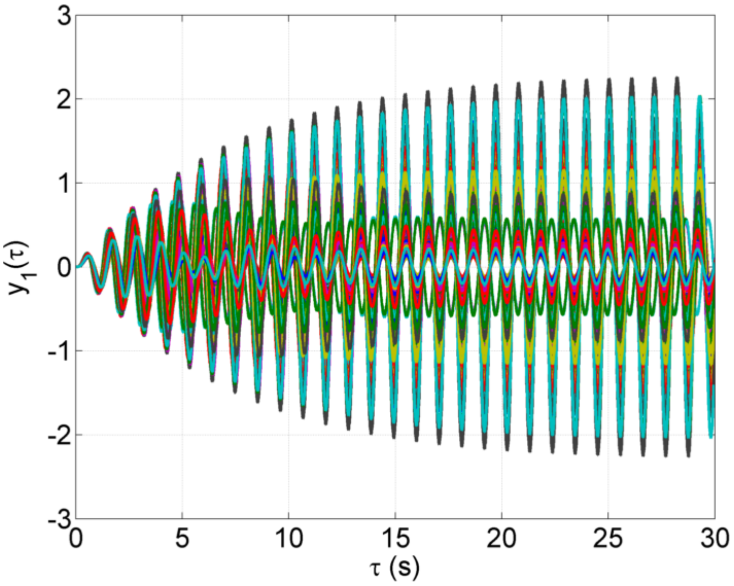


Figure 4.48 – Quarter car model – $N = 200$ different trajectories $y_1(\tau)$ in the warped time scale τ

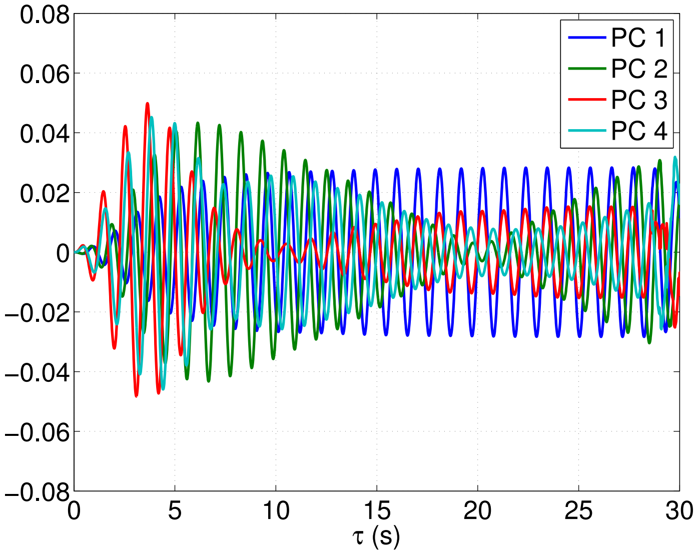


Figure 4.49 – Quarter car model – The first four principal components.

dimensional complex engineering problems, in which uncertainties from the mechanical properties and the excitations are accounted for simultaneously. It is worth noting that the resulting surrogate model is capable of capturing the time-dependent statistics (mean and standard deviation) of the response although it can only predict specific solutions with limited accuracy.

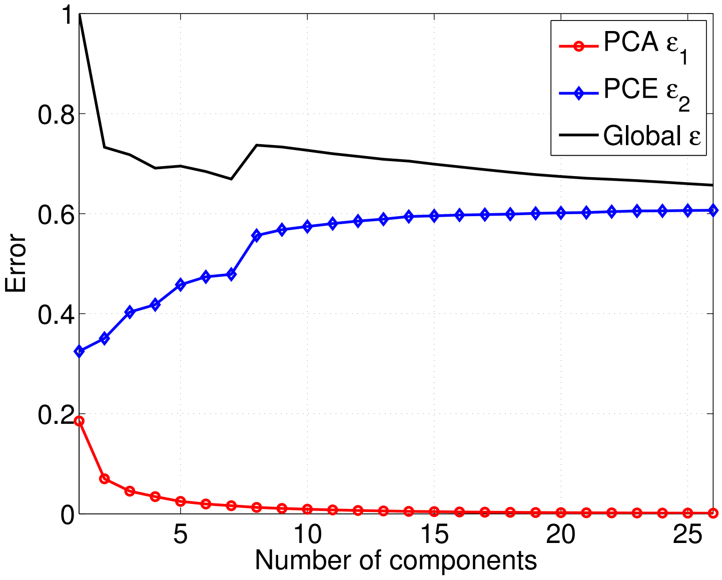
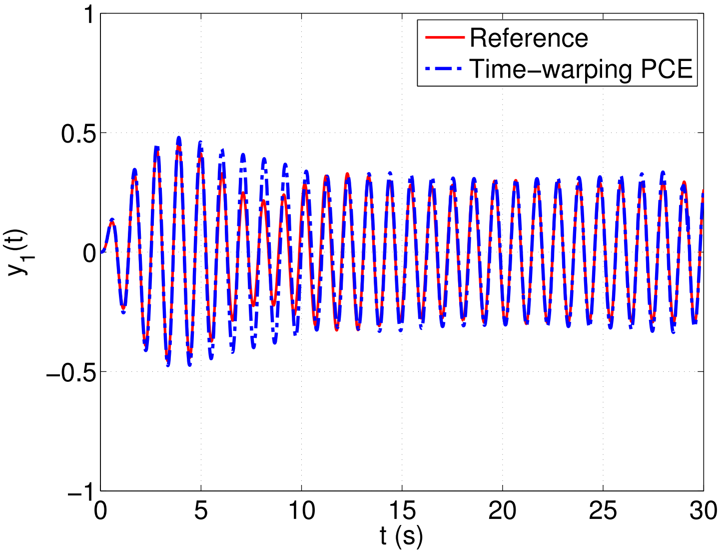
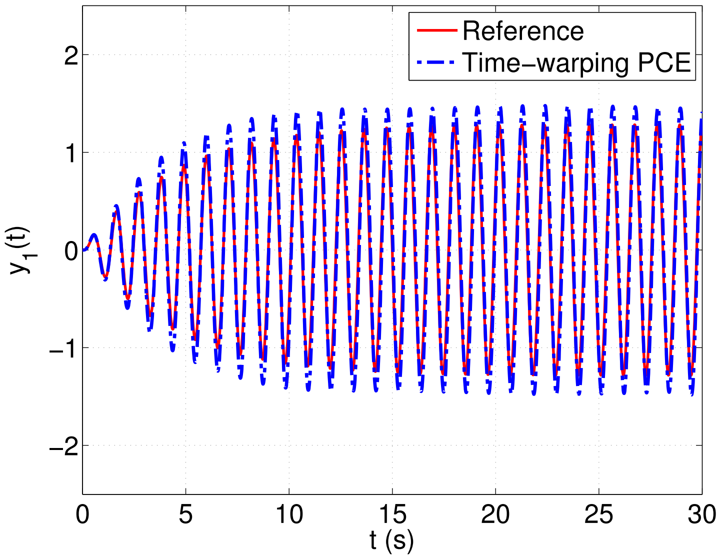


Figure 4.50 – Quarter car model – PCA truncation-induced error ϵ_1, PCE approximation error ϵ_2 normalized by trace($\tilde{\boldsymbol{\Sigma}}$) and the upper bound $\epsilon = \left(\sqrt{\epsilon_1} + \sqrt{\epsilon_2}\right)^2$ of the total error.

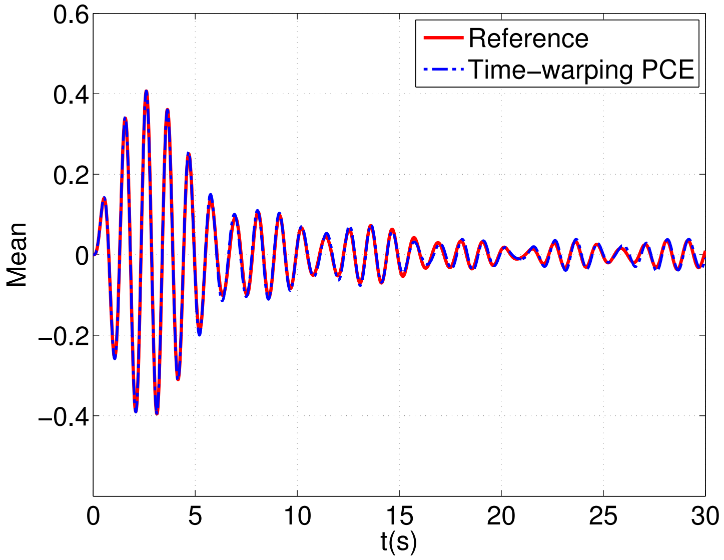


(a) $\boldsymbol{\xi} = (1897.02, 1771.4, 22.7, 42.0, 601.8, 0.09, 6.00)$

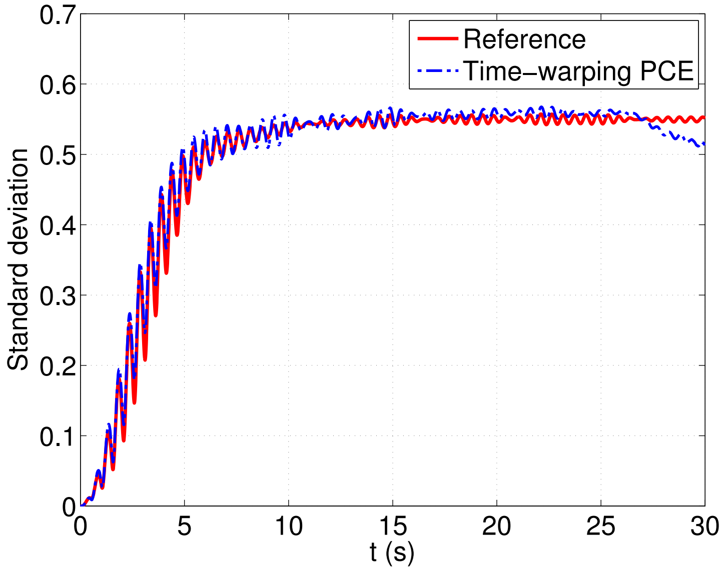


(b) $\boldsymbol{\xi} = (2087.88, 2184.06, 21.57, 43.43, 550.57, 0.10, 5.75)$

Figure 4.51 – Quarter car model – Two particular trajectories and their predictions by time-warping PCEs.



(a) Mean trajectory



(b) Standard deviation

Figure 4.52 – Quarter car model – Mean and standard deviation of the trajectories: comparison of time-warping PCE estimates and the reference trajectories.

4.3 Discussion

In this section, we discuss the reasons why a simple time-warping process allows to reduce significantly the complexity of the considered problems, in such a way that sparse PCEs can be applied successfully. The apparent explanation relies on the fact that the time-warping process maximizes the similarities in frequency and phase content of the different response trajectories by making them in-phase with a reference trajectory. As a result, the transformed responses vary around a reference trajectory. It is not mandatory that the responses in the transformed time line stay in a small neighbourhood of the reference counterpart in contrast to what is suggested by Le Maître et al. (2010). This is showcased by the numerical example of the Bouc-Wen oscillator, when variabilities in the responses after time-warping remain large, though they are smaller than in the original time scale.

The effectiveness of the time-warping PCEs approach may be further explained in a more general manner as follows. It was observed that when presented on the temporal variable t, the system's responses are increasingly non-linear functions of the uncertain parameters. When projecting the responses onto a *suitable* time scale, in this case the transformed time line τ, the resulting trajectories become smooth functions of the uncertain input parameters, whose complexity does hardly increase with time. Therefore, PCEs can be applied effectively to the projected responses and represent well the solutions at late instants. In this chapter, a measure of similarity was proposed to define a suitable space for projecting the responses, which exploits the periodicity of the trajectories. Further investigations are required to clearly determine such a suitable space in a more general case.

In the proposed approach, the virtual time τ is a function of the uncertain parameters $\boldsymbol{\xi}$. In other words, the time scale τ onto which the responses are projected is not deterministic. This differs significantly from approaches commonly used in the literature, in which the response trajectories are first projected onto a set of deterministic reduced basis determined a priori using a set of numerical simulations of the system. This is usually done with a simple *linear* transform in which the basis functions are *deterministic*, for instance the use of data compression techniques such as principal component analysis or wavelet decomposition.

When analysing further, one discovers a particular feature which constitutes a major difference between the classical time-frozen PCE approach and the proposed time-warping method. The PC coefficients $y_{\beta}(\tau)$ in the time-warping representation (Eq. (4.6)) are functions of τ, therefore being dependent on $\boldsymbol{\xi}$. This contradicts the representation of time-frozen PCEs (Eq. (4.2)), in which t and $\boldsymbol{\xi}$ intervene in the solution in a separated manner.

From a more general perspective, the effectiveness of the approach can be explained by analysing the functionalities of time-warping and PCEs. The most important feature of an oscillatory trajectory consists in its spectral content, which is characterized by the vibration periodicity. The other feature is the temporal content, which is characterized by the vibration amplitude. The pre-processing step handles partially the dynamics of the system by dealing with the frequency content. By means of the time-warping process, the trajectories have similar frequencies and phases. In other words, in terms of frequencies, the transformed trajectories exhibit a similar dynamical behaviour, which is close to that of the reference trajectory. The other aspect of the dynamics, *i.e.* the random temporal amplitude of the trajectories, is handled by means of sparse PCEs. As a summary, the dynamic is captured by the time-warping process, whereas the uncertainties are represented by PCEs.

As explained, sparse PCEs alone are not capable of dealing with the dynamics. The proposed approach illustrates a novel way to solve the uncertain dynamical problems, in which a specialized technique might be used to capture the dynamical aspect whereas sparse PCEs are used to propagate uncertainties. This principle is further developed in the next chapter to tackle more complex problems in which non-linear uncertain structures subject to stochastic motions are of interest and where the response trajectories are non-stationary, *i.e.* they do not show pseudo-periodic oscillations. The projection of the responses onto a special basis made of auto-regressive functions will allow us to represent the non-linear dynamical behaviour of the systems.

Finally, it is worthwhile mentioning that the proposed methodology is fully non-intrusive, *i.e.* the surrogate models of the systems' response trajectories are obtained by using a pre-computed set of trajectories related to an experimental design. In this respect, the methodology is readily applicable to any other problems featuring randomized limit cycle oscillations.

4.4 Summary

In this chapter, a non-intrusive sparse PCE approach based on a stochastic time-warping process is introduced. The proposed approach focuses on solving the problem of stochastic oscillations with random frequencies, which has been of long-term interest in the uncertainty quantification community. The practical application of the approach is explained in detail, including the determination of parameters governing the time-warping process, the coupling of principal component analysis and polynomial chaos expansions to represent the responses in the transformed time space and the use of the resulting surrogate model for predicting responses of the system and conducting statistical analyses. The chapter is concluded with a discussion on the reasons for the effectiveness of the proposed approach, which differs from the classical approach in terms of the philosophy employed to solve the stochastic dynamical problems.

5

Polynomial chaos nonlinear autoregressive with exogenous input model

The stochastic time-warping PCE approach introduced in the previous chapter allows one to tackle a class of problems involving oscillatory dynamics. It leads to a significant improvement compared to the time-frozen PCE approach. However, the use of the time-warping method in a more complicated case such as earthquake engineering is not straightforward. In this situation, a different strategy needs to be developed.

Recently, Spiridonakos and Chatzi (2012, 2013, 2015a,b) introduced a numerical approach that is based on the combination of two techniques, namely PCEs and nonlinear autoregressive with exogenous input (NARX) modelling, which is a universal tool in the field of system identification. In the proposed approach, the NARX model is used to represent the dynamical behaviour of the system, whereas PCEs tackle the uncertainty part. A two phase scheme is employed. First, a *stochastic* NARX model is identified to represent the dynamical system. It is characterized by a set of specified NARX model terms and associated *random coefficients*. Second, the latter are represented as PCEs of the random input parameters which govern the uncertainties in the considered system. In the two phases, both the NARX terms and the polynomial functions are selected by means of the heuristic genetic algorithm, which evolves randomly generated candidates toward better solutions using techniques inspired by natural evolution, *e.g.* mutation and crossover (Goldberg, 1989). The PC-NARX model is distinguished from conventional deterministic system identification tools in that it allows one to account for uncertainties arising both from the system properties, *e.g.* stiffness, hysteretic behaviour and energy dissipation, and from the stochastic excitations, *e.g.* ground motions in structural analysis. The approach proved its effectiveness in several case studies in structural dynamics (Spiridonakos and Chatzi, 2015a,b). It is worth mentioning that early combinations of system identification tools with polynomial chaos expansions can be found in the literature. Ghanem et al. (2005) regressed the restoring force of an oscillator on the Chebychev polynomials of state variables of the system, then used PCEs to represent the polynomial coefficients. Wagner and Ferris (2007) used PC-ARIMA models with a-priori known deterministic coefficients for characterizing terrain topology. Linear ARX-PCE models were also used by Kopsaftopoulos and Fassois (2013), Samara et al. (2013) and Sakellariou and Fassois (2016). However, in those studies the input parameters are not characterized by known probability density functions, thus the bases are constructed from arbitrarily selected families of orthogonal polynomials. Jiang and Kitagawa (1993) and Poulimenos and Fassois (2006) used time-dependent autoregressive moving average model with stochastic parameter evolution, in which the model parameters are random variables that change with time under stochastic smoothness constraints. Lately, Kumar and Budman (2014) represented the coefficients of a Volterra series with PCEs. Most recently, Ozen and Bal (2016) introduced the dynamical PCEs, which is also based on the idea that the future evolution of the response depends on the present solution. The above mentioned approaches were carried out in the time domain. In the frequency domain, Pichler et al. (2009) used linear regression-based metamodels to represent the frequency response functions of linear structures. In particular, the same strategy was recently used with PCEs by Yang et al. (2015), Jacquelin et al. (2015) and Yaghoubi et al. (2016).

In summary, the PC-NARX model consists of two components, namely a NARX and a PCE model. Spiridonakos and Chatzi (2015a,b) computed the two components with the heuristic genetic algorithm. The current chapter aims at introducing the so-called least angle regression (LARS) technique (Efron et al., 2004) for the computation of both NARX and PCE which are merely linear regression models. Indeed LARS has proven to be efficient in computing adaptive sparse PCEs at a relatively low computational cost (Blatman and Sudret, 2011). Note that LARS has been recently used for selecting the NARX terms in the context of system identification (Zhang and Li, 2015). Yet the original contribution of this chapter is to use LARS as a selection scheme for both NARX terms and PCE basis terms. This way we provide a new fully non-intrusive surrogate modelling technique that allows to tackle nonlinear dynamical systems with uncertain parameters.

The chapter is organized as follows: in the first section, the theory of NARX model is briefly reviewed. Next, the proposed LARS-based PC-NARX approach is introduced. The methodology is illustrated with four benchmark engineering case studies. Finally, a detailed discussion on the approach is given. The content of this chapter was presented by Mai and Sudret (2016b); Mai et al. (2016c) and has been submitted for publication in the International Journal for Uncertainty Quantification (Mai et al., 2016b).

5.1 Nonlinear autoregressive with exogenous input model

Let us consider a computational model $y(t) = \mathcal{M}(x(t))$ where $x(t)$ is the time-dependent input excitation and $y(t)$ is the response time history of interest. System identification aims at building a mathematical model describing $\mathcal{M}$ using the observed data of the input and output signals. Herein, we focus on system identification in the time domain. One discretizes the time duration under investigation in T discrete instants $t = 1, \ldots, T$. A nonlinear autoregressive with exogenous input (NARX) model allows one to represent the output quantity at a considered time instant as a function of its past values and values of the input excitation at the current or previous instants (Chen et al., 1989; Billings, 2013):

$$y(t) = \mathcal{F}\left(\boldsymbol{z}(t)\right) + \epsilon_t = \mathcal{F}\left(x(t), \ldots, x(t-n_x), y(t-1), \ldots, y(t-n_y)\right) + \epsilon_t, \tag{5.1}$$

where $\mathcal{F}(\cdot)$ is the underlying mathematical model to be identified, $\boldsymbol{z}(t) = (x(t), \ldots, x(t-n_x), y(t-1), \ldots, y(t-n_y))^{\mathsf{T}}$ is the vector of current and past values, n_x and n_y denote the maximum input and output time lags, ϵ_t is the residual error of the NARX model. In standard NARX models, the residual is assumed to be a normally independently distributed process. There are multiple options for the mapping function $\mathcal{F}(\cdot)$, *e.g.* wavelet, polynomial, sigmoid functions, radial basis functions and neural networks, which commonly specify an explicit function form with a finite number of parameters to describe the underlying relationship. Such a function is classified as parametric time series forecasting model (Cheng et al., 2015). In the literature, the following linear-in-the-parameters form is commonly used:

$$y(t) = \sum_{i=1}^{n_g} \vartheta_i \, g_i(\boldsymbol{z}(t)) + \epsilon_t, \tag{5.2}$$

in which n_g is the number of model terms $g_i(\boldsymbol{z}(t))$ that are functions of the regression vector $\boldsymbol{z}(t)$ and ϑ_i are the coefficients of the NARX model.

NARX models allow one to capture the dynamical behaviour of the system which follows the principle of causality, *i.e.* the current output quantity (or state of the system) $y(t)$ is affected by its previous states $\{y(t-1), \ldots, y(t-n_y)\}$ and the external excitation $\{x(t), \ldots, x(t-n_x)\}$. Note that the cause-consequence effect tends to fade away as time evolves, therefore it suffices to consider only a limited number of time lags before the current time instant. It is worth emphasizing that the model terms may be constructed from a variety of global or local basis functions. For instance the use of polynomial NARX model with $g_i(\boldsymbol{z}(t))$ being polynomial functions is popular in the literature, see *e.g.* Leontaritis and Billings (1985); Spinelli et al. (2006); Cantelmo and Piroddi (2010); Cheng et al. (2011).

The identification of a NARX model for a system consists of two major steps. The first one is *structure selection*, *i.e.* determining which NARX terms $g_i(\boldsymbol{z}(t))$ are in the model. The second step is parameter estimation, *i.e.* determining the associated model coefficients. Note that structure selection, particularly for systems involving nonlinearities, is critically important and difficult. Including spurious terms in the model leads to numerical and computational problems, therefore Billings (2013) suggests to identify the simplest model to represent the underlying dynamics of the system, which can be achieved by using the orthogonal least squares algorithm and its derivatives to select the relevant model terms one at a time. Different approaches for structure selection include trial and error methods, see *e.g.* Chen and Ni (2011); Piroddi (2008), and correlation-based methods, see *e.g.* Wei and Billings (2008); Cheng et al. (2011).

The identified model can be used for several purposes. First, it helps the analysts reveal the mechanism and behaviour of the underlying system. Understanding how a system operates offers the possibility to control it better. Second, the identified mathematical model can be utilized for predicting future responses of the system. From this point of view, it can be considered a metamodel (or approximate model) of the original $\mathcal{M}$.

5.2 Polynomial chaos - nonlinear autoregressive with exogenous input model

Consider a computational model $y(t,\boldsymbol{\xi}) = \mathcal{M}(x(t,\boldsymbol{\xi}_x),\boldsymbol{\xi}_s)$ where $\boldsymbol{\xi} = (\boldsymbol{\xi}_x,\, \boldsymbol{\xi}_s)^{\mathsf{T}}$ is the vector of uncertain parameters, $\boldsymbol{\xi}_x$ and $\boldsymbol{\xi}_s$ respectively represent the uncertainties in the input excitation $x(t,\boldsymbol{\xi}_x)$ and in the system itself. For instance, $\boldsymbol{\xi}_x$ can contain parameters governing the amplitude and frequency content of the excitation time series, while $\boldsymbol{\xi}_s$ can comprise parameters determining the system properties such as geometries, stiffness, damping and hysteretic behaviour.

Spiridonakos and Chatzi (2015a,b) proposed a numerical approach based on PCEs and the NARX model to identify the metamodel of such a dynamical system with uncertainties arising from both the excitation and the system properties. The time-dependent output quantity is first represented by means of a NARX model:

$$y(t,\boldsymbol{\xi}) = \sum_{i=1}^{n_g} \vartheta_i(\boldsymbol{\xi})\, g_i(\boldsymbol{z}(t)) + \epsilon_g(t,\boldsymbol{\xi}), \tag{5.3}$$

in which the model terms $g_i\left(\boldsymbol{z}(t)\right)$ are functions of the regression vector $\boldsymbol{z}(t) = (x(t),\, \dots, x(t-n_x),\, y(t-1),\, \dots, y(t-n_y))^{\mathsf{T}}$, n_x and n_y denote the maximum input and output time lags, $\vartheta_i(\boldsymbol{\xi})$ are the coefficients of the NARX model, $\epsilon_g(t,\boldsymbol{\xi})$ is the residual error, with zero mean Gaussian distribution and variance $\sigma_\epsilon^2(t)$. The proposed NARX model differs from the classical NARX model in the fact that the coefficients $\vartheta_i(\boldsymbol{\xi})$ are functions of the uncertain input parameters $\boldsymbol{\xi}$ instead of being deterministic. The stochastic coefficients $\vartheta_i(\boldsymbol{\xi})$ of the NARX model are then represented by means of truncated PCEs as follows (Soize and Ghanem, 2004):

$$\vartheta_i(\boldsymbol{\xi}) = \sum_{j=1}^{n_\psi} \vartheta_{i,j}\, \psi_j(\boldsymbol{\xi}) + \epsilon_i, \tag{5.4}$$

in which $\{\psi_j(\boldsymbol{\xi}), j = 1, \ldots, n_\psi\}$ are multivariate orthonormal polynomials of $\boldsymbol{\xi}$, $\{\vartheta_{i,j}, i = 1, \ldots, n_g, j = 1, \ldots, n_\psi\}$ are associated PC coefficients and ϵ_i is the truncation error. Finally, by substituting for Eq. (5.4) in Eq. (5.3), the PC-NARX model reads:

$$y(t, \boldsymbol{\xi}) = \sum_{i=1}^{n_g} \sum_{j=1}^{n_\psi} \vartheta_{i,j} \, \psi_j(\boldsymbol{\xi}) \, g_i(\boldsymbol{z}(t, \boldsymbol{\xi})) + \epsilon(t, \boldsymbol{\xi}), \tag{5.5}$$

where $\epsilon(t, \boldsymbol{\xi})$ is the total error time series due to the truncations of NARX and PCE models. In the proposed approach, the NARX model is used to capture the dynamics of the considered system, whereas PCEs are used to propagate uncertainties.

Let us discuss the difference between the PC-NARX model and the conventional time-dependent PCE formulation in Eq. (3.2). For the sake of clarity, Eq. (5.5) can be rewritten as follows:

$$y(t, \boldsymbol{\xi}) = \sum_{j=1}^{n_\psi} \left(\sum_{i=1}^{n_g} \vartheta_{i,j} \, g_i(\boldsymbol{z}(t, \boldsymbol{\xi})) \right) \psi_j(\boldsymbol{\xi}) \; + \epsilon(t, \boldsymbol{\xi}). \tag{5.6}$$

At a particular instant t, the polynomial coefficients $y_j(t) \overset{\text{def}}{=} \sum_{i=1}^{n_g} \vartheta_{i,j} \, g_i(\boldsymbol{z}(t, \boldsymbol{\xi}))$ are represented as functions of the past values of the excitation and the output quantity of interest. In other words, the polynomial coefficients follow certain dynamical behaviours, which constitutes a fundamental difference when being compared with the coefficients in the conventional model in Eq. (3.2). There the polynomial coefficients at time t are *deterministic* and are determined *independently*. As a consequence, a *high and increasing* polynomial order is required to maintain an accuracy level and properly capture the dynamics as time evolves (Wan and Karniadakis, 2006). In contrast, when a functional form is used to relate the coefficients $y_j(t)$ with the excitation and output time series, a *low and constant* polynomial order suffices. Spiridonakos and Chatzi (2015a) used PC-NARX models with fourth order PCEs to obtain remarkable results in the considered structural dynamics case studies. In the literature, Gerritsma et al. (2010) showed that when applying time-dependent PCEs, *i.e.* adding previous responses to the set of random variables to represent the current response, low-order polynomials could also be used effectively. From a similar perspective, the PC-flow map composition scheme proposed by Luchtenburg et al. (2014) was also proven efficient in solving the problems with low polynomial order, which was impossible with PCEs alone. At this point, one can recognize a similarity between the time-dependent PCE, the PC-flow map composition and the PC-NARX model. These methods rely on a recursive formulation which allows the prediction of future output values using *orthogonal* polynomials of its current state, thus making use of the causality effect in the dynamical system. In the recursive formulation of time-dependent PCEs and the PC-flow map composition, only one lagged output is utilized as historical data. PC-NARX model takes into consideration a series of lagged output and external excitations, therefore it can be considered as an extension of time-dependent PCEs and the PC-flow map composition that can be used in a more general case.

Indeed, not all the NARX and PC terms originally specified are relevant, as commonly observed in practice. The use of redundant NARX or PC terms might lead to large inaccuracy. Therefore, it is of utmost importance to identify the correct structure of NARX and PC models, *i.e.* to select appropriate NARX terms and PC bases. To this end, Spiridonakos and Chatzi (2015a) proposed a two-phase approach, in which the NARX terms and PC functions are subsequently selected by means of a genetic algorithm (Goldberg, 1989). However, due to the linear-in-parameters formulations of the NARX model (Eq. (5.3)) and the PC expansions (Eq. (5.4)), the question of selecting NARX and PC terms boils down to solving two linear regression problems. To this end, it appears that one can use techniques that are specially designed for linear regression analysis, for instance the least angle regression (LARS) technique proposed by Efron et al. (2004). In the field of system identification, LARS has been recently used by Zhang and Li (2015) for selecting the NARX terms. Before that, Cantelmo and Piroddi (2010) used the least absolute shrinkage and selection operator method (LASSO) (Tibshirani, 1996) for adaptive selection of a polynomial NARX model. Efron et al. (2004) showed that with one modification, the LARS procedure provides the entire paths of LASSO solution coefficients. The use of LARS in system identification can be classified as a correlation-based method, which selects the NARX terms that make significant contribution to the output using correlation analysis, see *e.g.* Billings and Wei (2008); Wei and Billings (2008); Cheng et al. (2011). LARS has also been used in the adaptive sparse PCE scheme (Blatman and Sudret, 2011) and showed great advantages compared to the other predictor selection methods, *i.e.* fast convergence and high accuracy with an ED of limited size.

5.2.1 Least angle regression-based approach

In this section, we introduce least angle regression (LARS) for the selection of appropriate NARX and PCE models. A two-phase approach is used, which sequentially selects NARX and PCE models as follows:

- Phase 1: Selection of the appropriate NARX model among a set of candidates.
    - Step 1.1: One specifies general options for the NARX model (model class and related properties), *e.g.* type of basis functions (polynomials, wavelet, sigmoid functions, neural network, *etc.*), maximum time lags of input and output, properties of the basis functions (*e.g.* maximum polynomial order). Note that it is always preferable to start with simple models having a reasonable number of terms. In addition, any available knowledge on the system, *e.g.* number of degrees of freedom, type of non-linear behaviour, should be used in order to obtain useful options for the general NARX structure. This leads to a full NARX model which usually contains more terms than actually needed for a proper representation of the considered dynamical system. At this stage, one assumes that the specified full NARX model contains the terms that can sufficiently describe the system. This assumption will be verified in the final step of this phase.
    - Step 1.2: One selects some candidate NARX models that are subsets of the specified full model. To this end, one considers the experiments exhibiting a

high level of non-linearity. For instance those experiments can be chosen with measures of nonlinearity or by inspection of the simulations with maximum response values exceeding a specified threshold. For each of the selected experiments, one determines a candidate NARX model containing a subset of the NARX terms specified by the full model. This is done using LARS and the input-output time histories of the considered experiment.

For experiment #k, the one-step-ahead prediction of the response reads:

$$\hat{y}_p(t,\boldsymbol{\xi}_k) = \sum_{i=1}^{n_g} \vartheta_i(\boldsymbol{\xi}_k)\, g_i(\hat{\boldsymbol{z}}_p(t,\boldsymbol{\xi}_k)), \tag{5.7}$$

in which

$$\hat{\boldsymbol{z}}_p(t,\boldsymbol{\xi}_k) = \Big(x(t,\boldsymbol{\xi}_{x,k}),\, \ldots\, , x(t-n_x,\boldsymbol{\xi}_{x,k}), y(t-1,\boldsymbol{\xi}_k),\, \ldots\, , y(t-n_y,\boldsymbol{\xi}_k)\Big)^{\mathsf{T}}, \tag{5.8}$$

where $y(t)$ is the recorded data used for training the NARX model. Denoting $\boldsymbol{\phi}(t) = \{g_i(\hat{\boldsymbol{z}}_p(t),\, i = 1,\, \ldots\, , n_g)\}^{\mathsf{T}}$ and $\boldsymbol{\vartheta} = \{\vartheta_i,\, i = 1,\, \ldots\, , n_g\}^{\mathsf{T}}$, the residual time series reads:

$$\epsilon_p(t,\boldsymbol{\xi}_k) = y(t,\boldsymbol{\xi}_k) - \hat{y}_p(t,\boldsymbol{\xi}_k) = y(t,\boldsymbol{\xi}_k) - \boldsymbol{\phi}^{\mathsf{T}}(t)\, \boldsymbol{\vartheta}(\boldsymbol{\xi}_k). \tag{5.9}$$

Thus, the sum of squared errors is given by:

$$\sum_{t=1}^{T} [\epsilon_p(t,\boldsymbol{\xi}_k)]^2 = \sum_{t=1}^{T} \left[y(t,\boldsymbol{\xi}_k) - \boldsymbol{\phi}^{\mathsf{T}}(t)\, \boldsymbol{\vartheta}(\boldsymbol{\xi}_k)\right]^2. \tag{5.10}$$

Using Eq. (5.9) and assembling all time instants in the k-th experiment, one obtains:

$$\begin{bmatrix} y(1,\boldsymbol{\xi}_k) \\ \vdots \\ y(T,\boldsymbol{\xi}_k) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}^{\mathsf{T}}(1) \\ \vdots \\ \boldsymbol{\phi}^{\mathsf{T}}(T) \end{bmatrix} \boldsymbol{\vartheta}(\boldsymbol{\xi}_k) + \begin{bmatrix} \epsilon(1,\boldsymbol{\xi}_k) \\ \vdots \\ \epsilon(T,\boldsymbol{\xi}_k) \end{bmatrix} \tag{5.11}$$

The above equation can be rewritten in matrix notations as follows:

$$\boldsymbol{y}_k = \boldsymbol{\Phi}_k\, \boldsymbol{\vartheta}(\boldsymbol{\xi}_k) + \boldsymbol{\epsilon}_k, \tag{5.12}$$

where $\boldsymbol{y}_k$ is the $T \times 1$ vector of output time-series, $\boldsymbol{\Phi}_k$ is the $T \times n_g$ information matrix with the i^{th} row containing the evaluations of NARX terms $\boldsymbol{\phi}(t)$ at instant $t = i$ and $\boldsymbol{\epsilon}_k$ is the residual vector. This is typically the equation of a linear regression problem, for which the relevant NARX regressors among the NARX candidate terms $\boldsymbol{\phi}(t)$ can be selected by LARS.

Note that the same candidate model might be obtained from different selected experiments. In theory all the available experiments can be considered, *i.e.* the number of candidate models is at most the size of the ED. Herein, we search for the appropriate NARX model among a limited number of experiments which are exhibiting strong non-linearity.

- Step 1.3: For each candidate NARX model, the corresponding NARX coefficients are computed for each of the experiments by ordinary least-squares. The parameters $\boldsymbol{\vartheta}(\boldsymbol{\xi}_k)$ minimizing the total errors in Eq. (5.10) is the least-squares solution of Eq. (5.12), *i.e.* :

$$\boldsymbol{\vartheta}(\boldsymbol{\xi}_k) = \arg\min_{\boldsymbol{\vartheta}}(\boldsymbol{\epsilon}_k^{\mathsf{T}}\, \boldsymbol{\epsilon}_k) = \left[\boldsymbol{\Phi}_k^{\mathsf{T}}\, \boldsymbol{\Phi}_k\right]^{-1} \boldsymbol{\Phi}_k^{\mathsf{T}}\, \boldsymbol{y}_k, \tag{5.13}$$

with the information matrix $\boldsymbol{\Phi}_k$ containing only the NARX regressors specified in the NARX candidate. Having at hand the NARX coefficients, the free-run reconstruction for each experiment output is conducted as follows:

$$\hat{y}_s(t, \boldsymbol{\xi}_k) = \sum_{i=1}^{n_g} \vartheta_i(\boldsymbol{\xi}_k)\, g_i(\hat{\boldsymbol{z}}_s(t, \boldsymbol{\xi}_k)), \tag{5.14}$$

in which

$$\hat{\boldsymbol{z}}_s(t, \boldsymbol{\xi}_k) = \Big(x(t, \boldsymbol{\xi}_{x,k}), \ldots, x(t - n_x, \boldsymbol{\xi}_{x,k}), \hat{y}_s(t-1, \boldsymbol{\xi}_k), \ldots, \hat{y}_s(t - n_y, \boldsymbol{\xi}_k)\Big)^{\mathsf{T}}. \tag{5.15}$$

It is worth underlining that the free-run reconstruction of the response is obtained using only the excitation time series $x(t)$ and the response initial condition y_0. The response is reconstructed recursively, *i.e.* its estimate at one instant is used to predict the response at later instants. This differs from Eq. (5.7) where the recorded response was used in the recursive formulation. The relative error for simulation #k reads:

$$\epsilon_k = \frac{\sum\limits_{t=1}^{T} (y(t, \boldsymbol{\xi}_k) - \hat{y}_s(t, \boldsymbol{\xi}_k))^2}{\sum\limits_{t=1}^{T} (y(t, \boldsymbol{\xi}_k) - \bar{y}(t, \boldsymbol{\xi}_k))^2}, \tag{5.16}$$

in which $\hat{y}_s(t, \boldsymbol{\xi}_k)$ is the output trajectory reconstructed by the NARX model and $\bar{y}(t, \boldsymbol{\xi}_k)$ is the mean value of the response time series $y(t, \boldsymbol{\xi}_k)$.

- Step 1.4: One selects the most appropriate NARX model among the candidates. Herein, the error criterion of interest is the mean value of the relative errors:

$$\bar{\epsilon} = \frac{1}{K} \sum_{k=1}^{K} \epsilon_k. \tag{5.17}$$

We propose to choose the NARX candidate that achieves a small mean error over the entire set of conducted simulations in the experimental design, *e.g.* $\bar{\epsilon} < 1 \times 10^{-3}$, with the smallest number of NARX terms. For instance, if two NARX candidates lead to the same level of mean error over the experimental design, the one comprising less NARX terms will be selected. In other words, the appropriate model is the simplest candidate that is capable of capturing the underlying system dynamical behaviour.

To refine the estimated coefficients, a nonlinear optimization for minimizing the simulation error (Eq. (5.16)) may be conducted afterwards (Spiridonakos and Chatzi, 2015a). However, this is not used in the current work due to the fact that LARS allows one to detect the appropriate NARX terms, therefore the models estimated by ordinary least-squares appear sufficiently accurate.

If no appropriate NARX model is obtained, *i.e.* the initial assumption that the full NARX model includes an appropriate candidate is not satisfied, the process is re-started from Step 1.1 (choice of model class), when different options for the full NARX model should be considered. For instance, one may use more complex models with larger time lags, different basis functions, etc. This basically introduces a loop for selecting the best NARX model.

- Phase 2: Representation of the NARX coefficients by means of PCEs using the sparse adaptive PCE scheme which is based on LARS (see Section 2.2). The NARX coefficients obtained from Phase 1 are used together with the corresponding sample set of the uncertain input parameters for training the PC expansions. Note that the marginal distributions and the correlation between the input parameters have been defined.

The proposed approach for computing a PC-NARX model is summarized by the flowchart in Figure 5.1.

5.2.2 Use of the surrogate model for prediction

The PC-NARX surrogate model can be used for the prediction[1] of the response for a set of input parameters $\boldsymbol{\xi}'$. Given the excitation $x(t, \boldsymbol{\xi}'_x)$ and the initial conditions of the response $y(t = 1, \boldsymbol{\xi}') = y_0$, the output time history of the system can be recursively obtained as follows:

$$\hat{y}(t, \boldsymbol{\xi}') = \sum_{i=1}^{n_g} \sum_{j=1}^{n_\psi} \vartheta_{i,j} \psi_j(\boldsymbol{\xi}') \, g_i(\hat{\boldsymbol{z}}(t, \boldsymbol{\xi}')), \; t = 2, \, \ldots \, , T, \tag{5.18}$$

in which

$$\hat{\boldsymbol{z}}(t, \boldsymbol{\xi}') = \left(x(t, \boldsymbol{\xi}_x^{'}), \, \ldots \, , x(t - n_x, \boldsymbol{\xi}_x^{'}), \hat{y}(t-1, \boldsymbol{\xi}'), \, \ldots \, , \hat{y}(t - n_y, \boldsymbol{\xi}') \right)^{\mathsf{T}}. \tag{5.19}$$

Currently, no close-form formulation for computing the time-dependent statistics of the output quantity is available as opposed to time-frozen PCEs. However, the evolutionary response statistics can be straightforwardly obtained by means of Monte Carlo simulation using the PC-NARX model.

---

[1] In what follows, the term “prediction” is employed to refer to the NARX model’s so-called “simulation mode” as addressed in signal processing literature, which stands for the estimation of the response relying only on its initial condition and feedback of the excitation. The term “prediction” is used however because it is the standard wording in the surrogate modelling community.

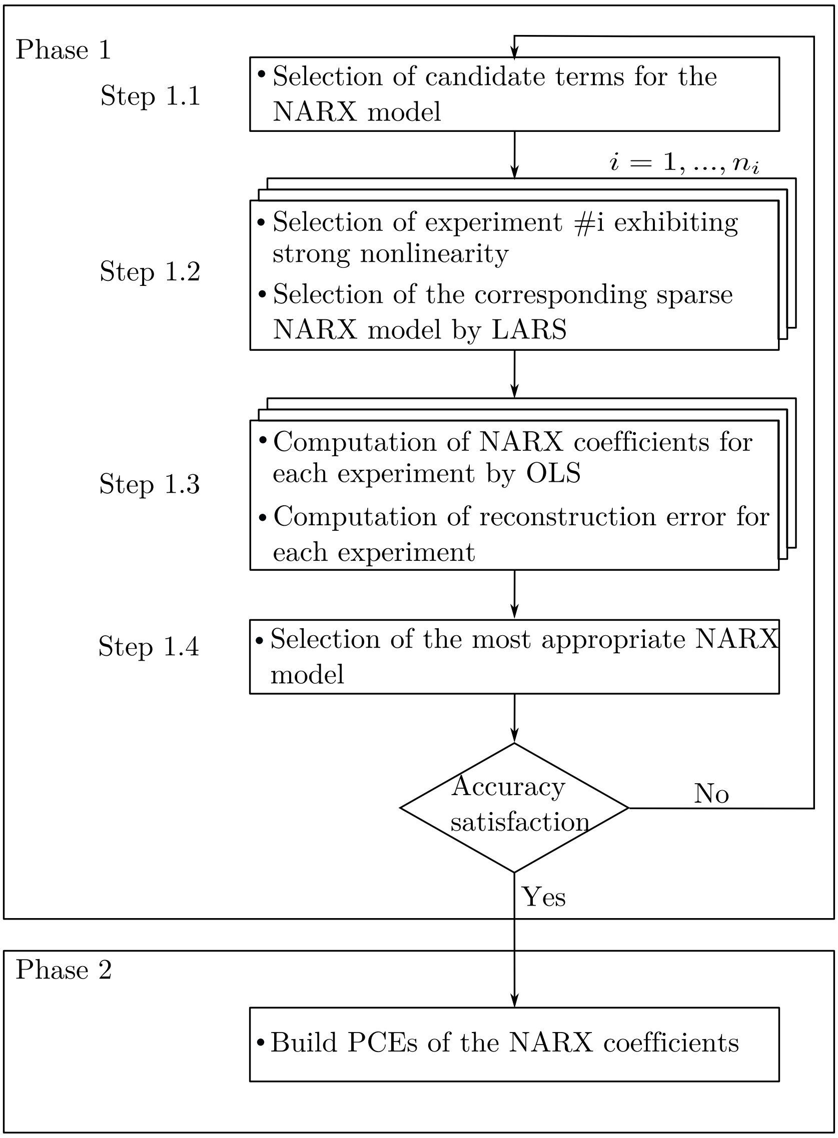


Figure 5.1 – Computation of LARS-based PC-NARX model

5.2.3 Validation of the surrogate model

The PC-NARX model is computed using an ED of limited size. The validation process is conducted with a validation set of large size which is independent of the ED. A large number, *e.g.* $n_{val} = 10^4$, of input parameters and excitations is generated. One uses the numerical solver to obtain the response time histories sampled at the discrete time instants $t = 1, \ldots, T$. Then the PC-NARX model (Eq. (5.18)) is used to predict the time dependent responses for the excitations and uncertain parameters of the validation set. The accuracy of the computed PC-NARX model is validated by comparing its predictions with the actual responses in terms of the relative errors and the evolutionary statistics of the response. For prediction #i, the relative error reads:

$$\epsilon_{val,i} = \frac{\sum\limits_{t=1}^{T} (y(t,\boldsymbol{\xi}_i) - \hat{y}(t,\boldsymbol{\xi}_i))^2}{\sum\limits_{t=1}^{T} (y(t,\boldsymbol{\xi}_i) - \bar{y}(t,\boldsymbol{\xi}_i))^2}, \tag{5.20}$$

where $\hat{y}(t,\boldsymbol{\xi}_i)$ is the output trajectory predicted by PC-NARX and $\bar{y}(t,\boldsymbol{\xi}_i)$ is the mean value of the actual response time series $y(t,\boldsymbol{\xi}_i)$. The above formula is also used to calculate the accuracy of the time-dependent statistics (*i.e.* mean, standard deviation) predicted by PC-NARX. The mean value of the relative errors over n_{val} predictions reads:

$$\bar{\epsilon}_{val} = \frac{1}{n_{val}} \sum_{i=1}^{n_{val}} \epsilon_{val,i}. \tag{5.21}$$

The relative error for a quantity y, *e.g.* the maximal value of the response (resp. the response at a specified instant) is given by:

$$\epsilon_{val,y} = \frac{\sum\limits_{i=1}^{n_{val}} (y_i - \hat{y}_i)^2}{\sum\limits_{i=1}^{n_{val}} (y_i - \bar{y})^2}, \tag{5.22}$$

where y_i is the actual response, $\hat{y}_i$ is the prediction by PC-NARX and $\bar{y}$ is the mean value defined by $\bar{y} = \frac{1}{n_{val}} \sum\limits_{i=1}^{n_{val}} y_i$.

5.3 Numerical applications

The use of the LARS-based PC-NARX model is now illustrated with four nonlinear dynamical systems with increasing complexity, namely a quarter car model subject to a stochastic sinusoidal road profile, a single degree-of-freedom (SDOF) Duffing, a SDOF Bouc-Wen oscillator and a three-story steel building subject to *non-stationary stochastic* excitation. In all considered numerical examples, uncertainties arising from the system properties and from the excitation are taken into account. PC-NARX models are computed using a small number of numerical simulations as experimental design. The validation is conducted by comparing their response predictions with the reference values obtained by using Monte Carlo simulation (MCS) on the numerical solvers.

5.3.1 Quarter car model

5.3.1.1 Problem statement

In the first numerical example, a quarter car model of a vehicle suspension represented by a nonlinear two DOF system (Kewlani et al., 2012) (Figure 5.2) is considered. The displacements of the masses are governed by the following system of ordinary differential equations (ODEs):

$$\begin{cases} m_s\, \ddot{y}_1(t) = -k_s\,(y_1(t) - y_2(t))^3 - c\,(\dot{y}_1(t) - \dot{y}_2(t)), \\ m_u\, \ddot{y}_2(t) = k_s\,(y_1(t) - y_2(t))^3 + c\,(\dot{y}_1(t) - \dot{y}_2(t)) + k_u\,(x(t) - y_2(t)), \end{cases} \tag{5.23}$$

in which the sprung mass m_s and the unsprung mass m_u are connected by a nonlinear spring of stiffness k_s and a linear damper of damping coefficient c. The forcing function $x(t)$ is applied to m_u through a linear spring of stiffness k_u. $y_1(t)$ and $y_2(t)$ are the displacements of m_s and m_u respectively. A sinusoidal function road profile with amplitude A and frequency ω is considered:

$$x(t) = A\,\sin(\omega\, t). \tag{5.24}$$

The parameters of the quarter car model and of the excitation $\boldsymbol{\xi} = \{k_s,\, k_u,\, m_s,\, m_u,\, c,\, A,\, \omega\}$ are modelled by independent random variables with marginal distributions given in Table 5.1. Gaussian distributions are used as in the original publication, although it would be more appropriate to use *e.g.* lognormal variables to ensure the positivity of mass and stiffness parameters. Kewlani et al. (2012) addressed this numerical example with the multi-element PCE approach.

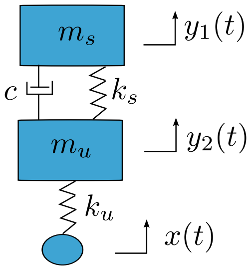


Figure 5.2 – Quarter car model – Mechanical system.

Table 5.1 – Parameters of the quarter car model and of the road excitation

Parameter	Distribution	Mean	Standard deviation
k_s (N/m^3)	Gaussian	2000	200
k_u (N/m)	Gaussian	2000	200
m_s (kg)	Gaussian	20	2
m_u (kg)	Gaussian	40	4
c (N s /m)	Gaussian	600	60
A (m)	Uniform	0.1	$0.01/\sqrt{3}$
ω (rad/s)	Uniform	$2\,\pi$	$0.2\,\pi/\sqrt{3}$

5.3.1.2 PC-NARX model

We now aim at building the metamodel for representing the displacement time histories $y_1(t)$ of the sprung mass m_s. For this purpose, $N = 100$ analyses of the system are conducted with 100 samples of the uncertain input parameters generated by Latin hypercube sampling. The system of ODEs are solved with the Matlab solver `ode45` (explicit Runge-Kutta method) for the total duration $T = 30$ s and the time step $\mathrm{d}t = 0.01$ s. In the first place, a NARX model structure is chosen, in which the model terms are polynomial functions of past values of the output and excitation $g_i(t) = y_1^l(t-j)\, x^m(t-k)$ with $l+m \leq 3$, $0 \leq l \leq 3$, $0 \leq m \leq 1$, $j = 1, \ldots, 4$, $k = 0, \ldots, 4$. The specified full NARX model contains 86 terms. It is worth noticing that the initial choice of the NARX structure was facilitated by the knowledge of the dynamical nonlinear behaviour of the system of interest. For instance, polynomial functions of order up to 3 are used because of the cubic nonlinear behaviour in Eq. (5.23). As a rule of thumb, the maximum time lags $n_x = n_y = 4$ are chosen equal to twice the number of degrees of freedom of the considered system.

Next, the candidate NARX models were computed. To this end, we selected the simulations with maximum displacement exceeding a large threshold, *i.e.* $\max(|y_1(t)|) > 1.2$ m and retained 15 experiments. For each selected simulation, LARS was applied to the initial full NARX model to detect the most relevant NARX terms constituting a candidate NARX model. This procedure resulted in 10 different candidate NARX models in total.

For each candidate NARX model, we computed the NARX coefficients for all simulations by ordinary least-squares to minimize the sum of squared errors (Eq. (5.13)). We then reconstructed all the output time histories with the computed coefficients and calculated the relative errors ϵ_k of the reconstruction (Eq. (5.16)).

Among the 10 candidates, we selected the most appropriate NARX model which achieves a sufficiently small overall error with the smallest number of terms. The selected model results in a mean relative error $\bar{\epsilon} = 3.56 \times 10^{-4}$ for 100 simulations in the ED and contains 6 terms, namely the constant term, $x(t-4)$, $y_1(t-4)$, $y_1(t-1)$, $y_1^3(t-1)$, $y_1^2(t-4)\, x(t-4)$. LARS proves effective in selecting the appropriate NARX model by retaining only 6 among the 86 candidate terms available to describe the system.

In the next step, we expanded the 6 NARX coefficients by adaptive sparse PCEs of order $p \leq 20$ with maximum interaction rank $r = 2$ and truncation parameter $q = 1$. The NARX coefficients computed in the previous step were used for training the metamodel. PCE models that minimize the LOO errors were selected. This led to LOO errors smaller than 10^{-7}. The optimal PCE order selected by the adaptive scheme is up to 6.

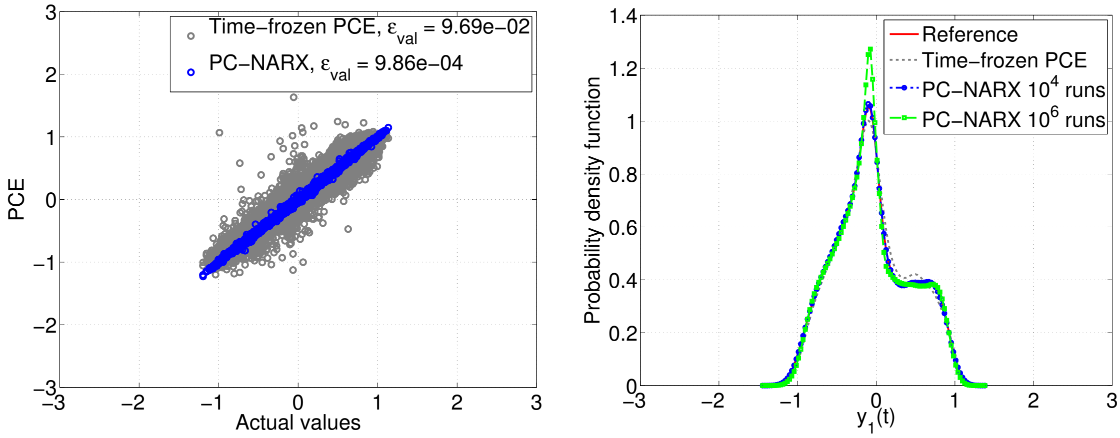
For the sake of comparison, we represent the response $y_1(t)$ by means of time-frozen sparse PCEs. For this purpose, adaptive sparse PCEs are used with an ED of size $N = 500$. The best PCE with maximum degree $1 \leqslant p \leqslant 20$, maximum interaction rank $r = 2$ and truncation parameter $q = 1$ is selected The PC-NARX and time-frozen PCEs are used to predict the output time histories for an independent validation set of size $n_{val} = 10^4$ which is pre-computed by the numerical solver (Matlab `ode45` algorithm). The accuracy of the two PCE approaches are compared in the following.

5.3.1.3 Prediction of instantaneous and maximal displacements

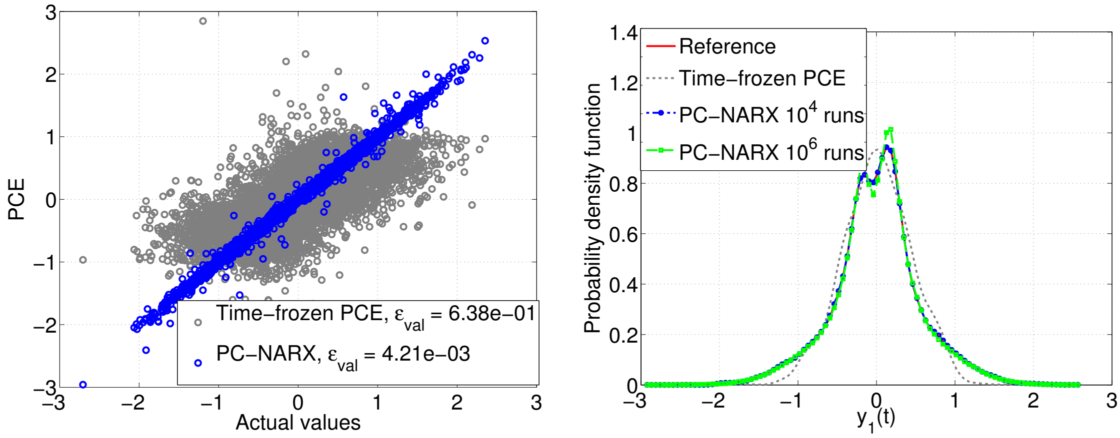
We compare the predictions of the mass displacement at an early time instant $t = 5$ s and a late instant $t = 30$ s (Figure 5.3). The instantaneous responses $y_1(t)$ predicted by two PCE approaches are plotted in Figure 5.3(a) and Figure 5.3(c) versus the actual responses obtained by numerically solving the considered system of ODEs. One observes that PC-NARX outperforms time-frozen PCEs. At $t = 30$ s, PC-NARX is still capable of providing highly accurate predictions ($\epsilon_{val} = 4.21 \times 10^{-3}$) whereas time-frozen PCEs show large inaccuracies already from $t = 5$ s. In terms of statistics, the probability density functions (PDFs) of the instantaneous responses are depicted in Figure 5.3(b) and Figure 5.3(d). The reference PDF is obtained from MCS using the validation set of size $n_{val} = 10^4$. One notices that the instantaneous PDFs computed by time-frozen PCEs do not differ significantly from the reference function, whereas the predictions are actually of poor quality. Indeed, it is always possible that two sets of samples provide identical PDFs while their pair-wise comparison shows large discrepancies. This example shows a problem that is overlooked in the literature on surrogate models for dynamical systems, when conclusions are commonly drawn only based on the statistics (for instance, PDFs, mean and standard deviation) of the predicted quantities. This might be misleading for judging the accuracy of the metamodel under investigation in predicting *specific* output values. The PDFs computed by PC-NARX show perfect agreement with the reference functions, which is obvious because every single PC-NARX-based prediction is highly accurate.

We investigate now the effect of the sample size on the MCS estimates. PC-NARX is used to predict the responses for a different validation set of size $n_{val} = 10^6$. Figure 5.3(b) and Figure 5.3(d) show that the PDFs obtained with 10^6 runs differ slightly from the PDFs obtained with 10^4 runs. For instance, the peaks of the 10^6 runs-based PDFs become marginally higher. It is believed that the PDF obtained with 10^6 runs of the PC-NARX surrogate is the most accurate, although this cannot be validated by running 10^6 times the computationally expensive model. In general, it is prohibitive to obtain the true PDFs with MCS on the original numerical model, however it is feasible using PC-NARX model.

Next, we compare the maximal displacements $\max(|y_1(t)|)$ predicted by the two PCE approaches with the actual values. Note that the maximal displacements from the ED are retrieved, then used as ED to directly compute the sparse adaptive PCE of that quantity. The same options as in time-frozen PCEs are used. Figure 5.4 clearly show that PC-NARX outperforms the usual PCE approach in predicting the maximal response. The former provides predictions with great accuracy indicated by the validation error $\epsilon_{val} = 3.12 \times 10^{-3}$, resulting in a PDF that is consistent with the reference one. In contrast, the PCE of the maximal displacement gives poor results, as it is expected. It has already been observed that the instantaneous responses are increasingly complex functions of the input parameters as time evolves. Consequently, the maximal value, which does not occur at the same time instant for different trajectories, is not a smooth function of the input random variable and cannot be approximated accurately with regular PCEs. As shown in Figure 5.4(b), the PC-NARX technique provides the PDF of this maximum



(a) $t = 5$ s (b) $t = 5$ s



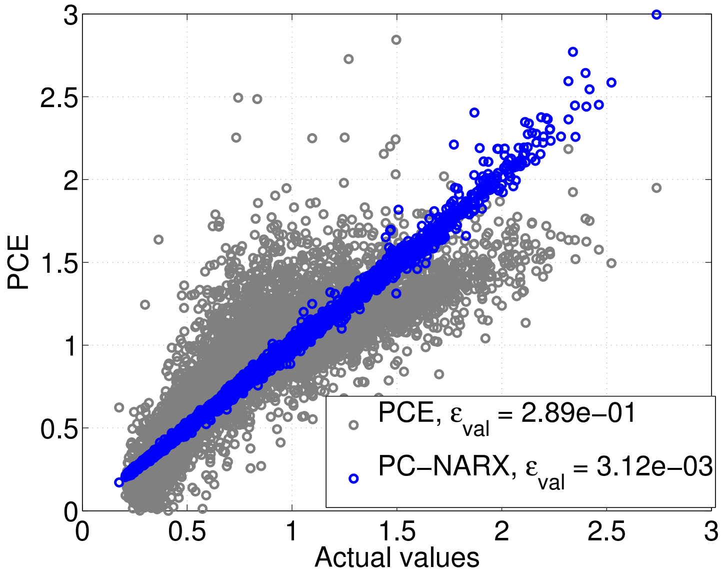
(c) $t = 30$ s (d) $t = 30$ s

Figure 5.3 – Quarter car model – Instantaneous displacements: comparison of the two approaches.

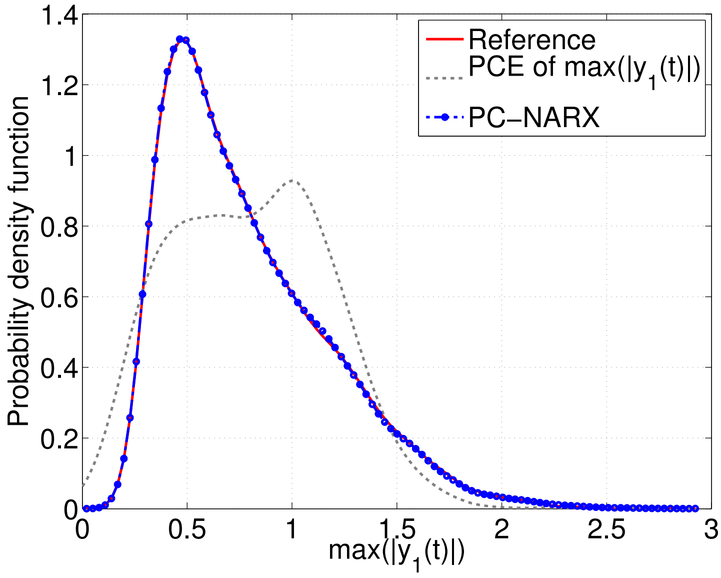
with remarkable accuracy.

5.3.1.4 Prediction of new trajectories and statistical moments

Let us evaluate the overall performance of the two PCE approaches in predicting the entire response time histories. Figure 5.5 depicts two specific time dependent response trajectories for two distinct samples of the uncertain input parameters. The predictions by the two PCE metamodels are compared with the reference responses obtained by the numerical solver. It is shown that the accuracy of time-frozen PCEs degenerates relatively quickly as time progresses. Around $t = 5$ s, time-frozen PCEs start showing signs of instability and the predictions become inaccurate. In contrast, PC-NARX provides predictions that are indistinguishable from the actual responses. Over 10^4 validation trajectories, PC-NARX leads to a mean relative error $\bar{\epsilon}_{val} = 0.17 \times 10^{-2}$, and only 5 simulations among them exhibit a relative error $\epsilon_{val,i}$ exceeding 0.1.



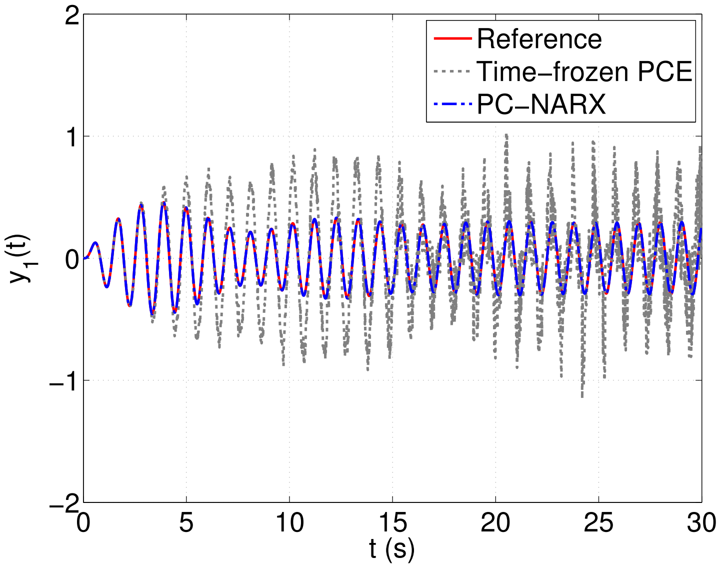
(a) Maximal displacement $\max(|y_1(t)|)$



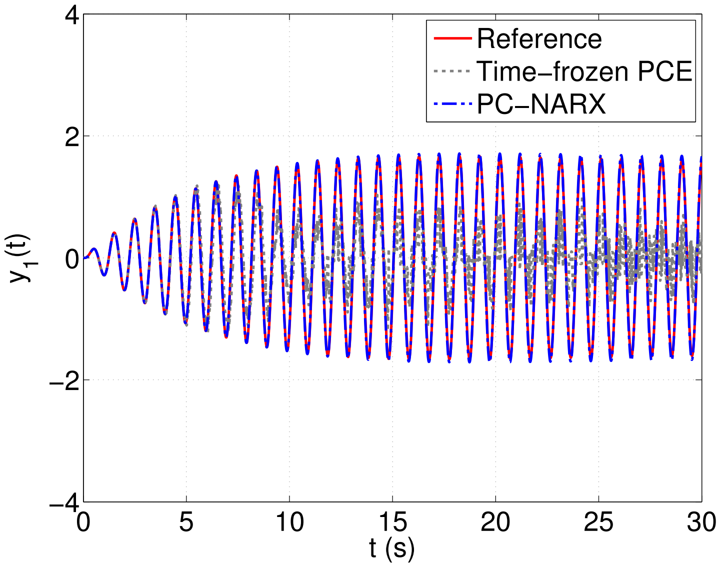
(b) PDF of $\max(|y_1(t)|)$

Figure 5.4 – Quarter car model – Maximal displacement: comparison of the two approaches.

The two PCE approaches are now compared in terms of predicting the evolutionary response statistics. Figure 5.6 shows that time-frozen PCEs can represent relatively well the mean trajectory, except for the late instants ($t > 20$ s) where the discrepancy is noticeable. However, time-frozen PCEs fail to capture the oscillatory behaviour of the standard deviation when the prediction starts deviating significantly from the actual trajectory at $t = 5$ s. The improvement in accuracy of PC-NARX is outstanding, in particular because it can represent in detail the oscillatory response statistics. The relative errors are respectively $\epsilon_{val,Mean} = 0.73 \times 10^{-2}$ and $\epsilon_{val,Std} = 0.91 \times 10^{-2}$ for the mean and standard



(a) $\boldsymbol{\xi} = (1897.02, 1771.4, 22.7, 42.0, 601.8, 0.09, 6.00)$



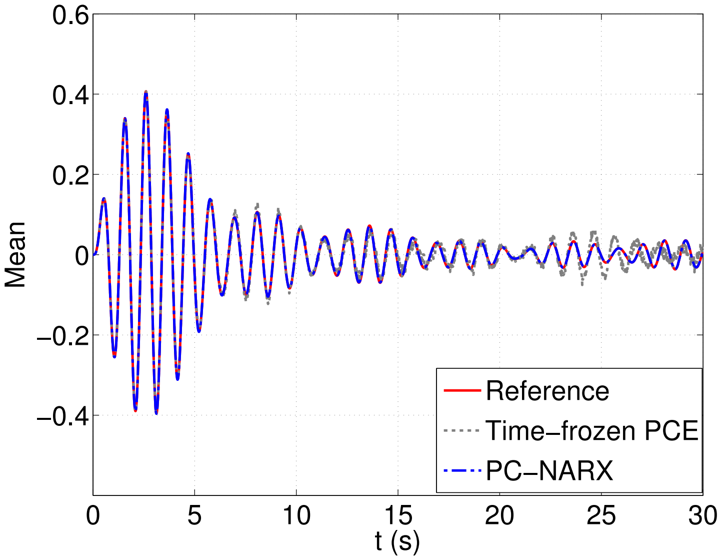
(b) $\boldsymbol{\xi} = (2082.94, 2187.10, 15.46, 40.08, 564.26, 0.10, 6.41)$

Figure 5.5 – Quarter car model – Two particular trajectories and their predictions by means of time-frozen PCEs and PC-NARX.

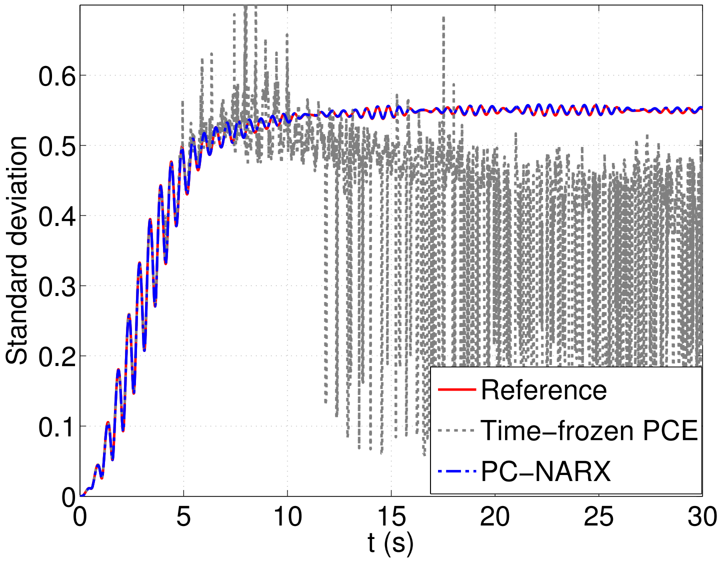
deviation time-histories.

5.3.1.5 Effect of the size of the experimental design

It has been shown that using an experimental design of size 100 led to remarkably accurate PC-NARX model. Let us now investigate the case when only few simulations are available, for instance assume that an ED of size 30 is given. The same procedure as presented above



(a) Mean trajectory

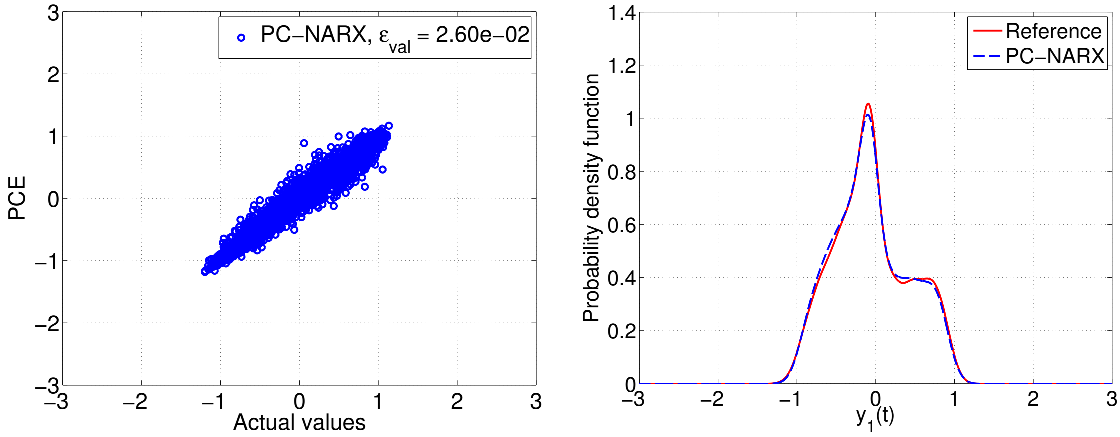


(b) Standard deviation

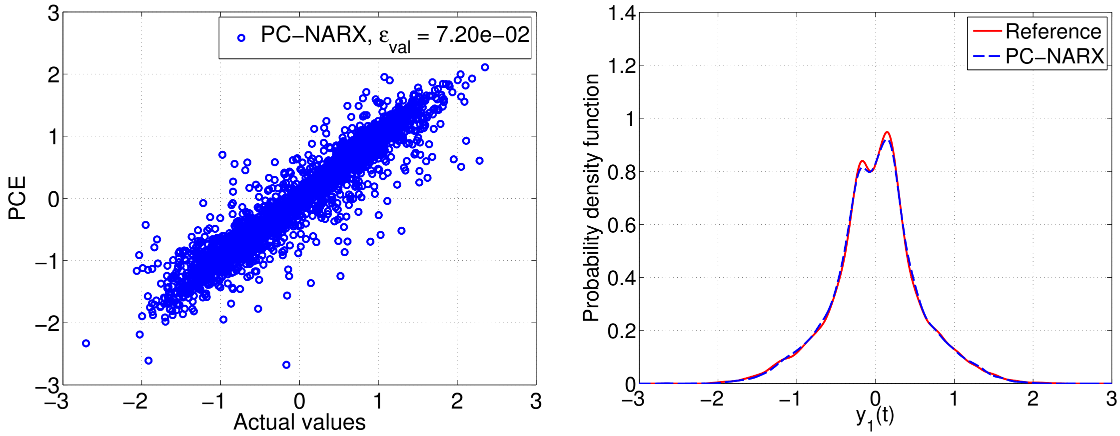
Figure 5.6 – Quarter car model – Mean and standard deviation of the trajectories: comparison of the two approaches.

is applied to compute the PC-NARX model, which is validated on the same independent validation set of size $10,000$. Among the validation set, 876 predictions exhibit a relative error larger than 0.1.

In Figure 5.7, we compare the displacements predicted by PC-NARX with the actual values at two instants $t = 5$ and $t = 30$ s. One notices a clear deterioration in terms of the accuracy of PC-NARX when compared to the previous study conducted with a larger experimental design. The relative error of the predictions increases by one order of magnitude. The effect of the inaccuracy on the PDFs is, however, not pronounced. The predicted PDFs can even be considered highly accurate as conventionally accepted in the literature.



(a) $t = 5$ s (b) $t = 5$ s

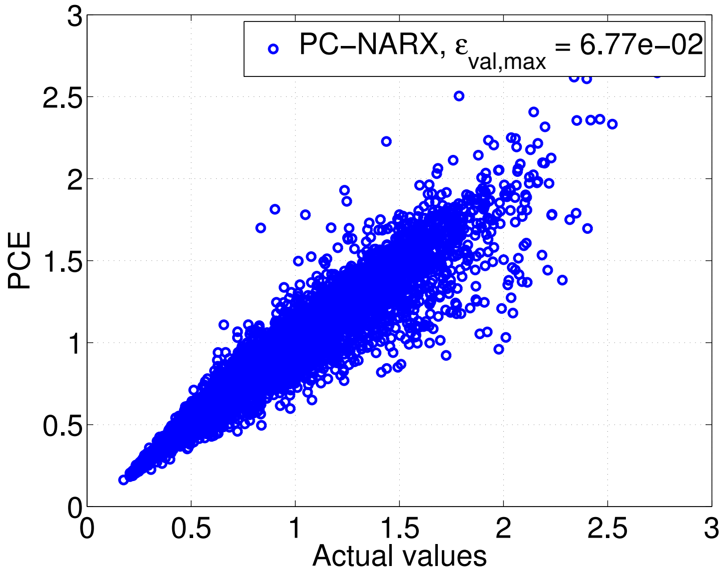


(c) $t = 30$ s (d) $t = 30$ s

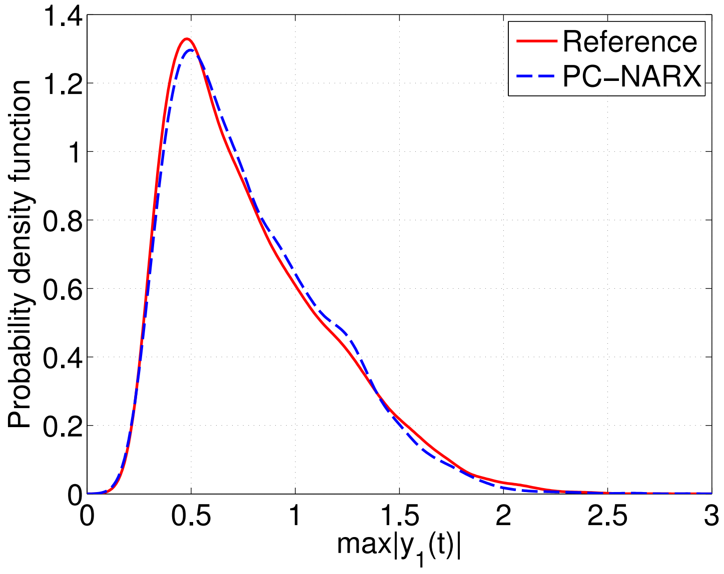
Figure 5.7 – Quarter car model – Instantaneous displacements.

Figure 5.8(a) depicts the maximal displacements predicted by the PC-NARX model in comparison with the actual values. For large values, the predictions exhibit considerable discrepancies from the actual responses, which is indicated by the large error $\epsilon_{val} = 6.77 \times 10^{-2}$. Therefore, the PDF predicted by the PC-NARX model differs from the reference function (Figure 5.8(b)). The discrepancy is, however, not large. It showcases again the problem that specific predictions by metamodels might be not sufficiently accurate, however the resulting PDF does not deviate significantly from the reference function. Regarding the evolutionary statistics of the response, Figure 5.9 shows that the mean trajectory is well captured by the PC-NARX model. In addition, only a minor discrepancy is observed for the standard deviation curve.

This numerical case study illustrates the effectiveness of the LARS-based PC-NARX model with respect to the usual time-frozen sparse PCE approach. It is worth noting that



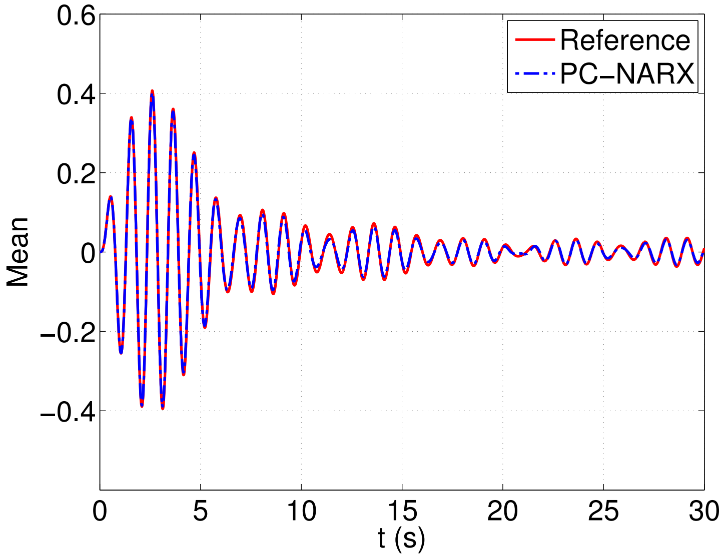
(a) Maximal displacement $\max(|y_1(t)|)$



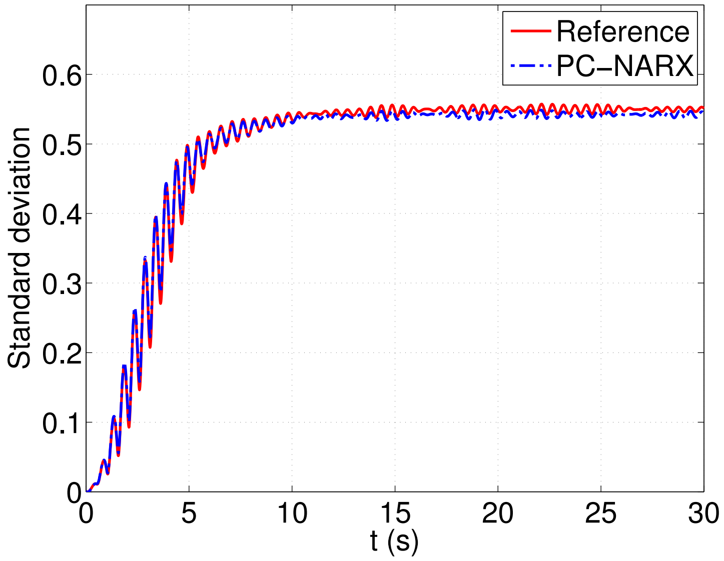
(b) PDF of $\max(|y_1(t)|)$

Figure 5.8 – Quarter car model – Maximal displacement: comparison of the two approaches.

the process was conducted using the uncertainty quantification software UQLab (Marelli and Sudret, 2014a), more specifically the polynomial chaos expansion toolbox (Marelli and Sudret, 2015). The following case studies will be more complicated, in particular because the excitations are *non-stationary stochastic processes* with varying intensity and frequency content that are generated by means of a probabilistic model presented in the following section.



(a) Mean trajectory



(b) Standard deviation

Figure 5.9 – Quarter car model – Mean and standard deviation of the trajectories.

5.3.2 Probabilistic ground motion model

In earthquake engineering, it is common practice to use recorded ground motions for structural dynamic analysis. Although the databases of recorded accelerograms are constantly enriched with recent events occurring worldwide, the number of motions with characteristics corresponding to a design scenario under investigation is commonly not sufficient for statistical analysis. This led to the development of synthetic ground motions models in the last few decades, which aim at generating artificial accelerograms with characteristics similar to seismic records.

Various stochastic ground motion models in the literature can be classified in three types (Douglas and Aochi, 2008): record-based parameterized models that are fit to recorded motions (Pousse et al., 2006; Rezaeian and Der Kiureghian, 2010a; Yamamoto and Baker, 2013), source-based models that consider the physics of the source mechanism and wave travel-path (Pitarka and Irikura, 1998; Boore, 2003), and hybrid models that combine elements from both source- and record-based models (Graves and Pitarka, 2010). The record-based stochastic models do not allow to take into consideration physical phenomena such as surface waves, travel path of the earthquake. This is, however, not a major issue in the context of structural engineering, when the analysts usually do not have knowledge about the source, path and site, and are more interested in the acceleration time-histories that excite the structures. Moreover, the record-based stochastic models are usually straightforward to use with only a handful of required input parameters.

Rezaeian and Der Kiureghian (2010a) proposed a probabilistic model depending on 6 parameters that govern the spectral and temporal contents of the motions. This model is employed in the present work to generate a large suite of synthetic ground motions used for non-linear dynamic analysis of the structures. The ground motion acceleration is represented as a non-stationary process by means of a modulated filtered white noise process as follows:

$$x(t) = q(t, \boldsymbol{\alpha}) \left\{ \frac{1}{\sigma_h(t)} \int_{-\infty}^{t} h[t-\tau, \boldsymbol{\lambda}(\tau)] \omega(\tau) \,\mathrm{d}\tau \right\}. \tag{5.25}$$

The white-noise process denoted by $\omega(\tau)$ passes a filter $h[t-\tau, \boldsymbol{\lambda}(\tau)]$ which is selected as an impulse-response function:

$$\begin{aligned} h[t-\tau, \boldsymbol{\lambda}(\tau)] &= \frac{\omega_f(\tau)}{\sqrt{1-\zeta_f^2(t)}} \exp\left(-\zeta_f(\tau)\omega_f(\tau)(t-\tau)\right) \\ &\quad \times \sin\left(\omega_f(\tau)\sqrt{1-\zeta_f^2(\tau)}(t-\tau)\right) \quad \text{for} \quad \tau \leq t, \\ h[t-\tau, \boldsymbol{\lambda}(\tau)] &= 0 \quad \text{for} \quad \tau > t, \end{aligned} \tag{5.26}$$

where $\boldsymbol{\lambda}(\tau) = (\omega_f(\tau), \zeta_f(\tau))$ is the vector of time-varying parameters of the filter h. $\omega_f(\tau)$ and $\zeta_f(\tau)$ are respectively the filter's frequency and bandwidth at time instant τ. They represent the evolving predominant frequency and bandwidth of the ground motion. A linear model is assumed for $\omega_f(\tau)$ and $\zeta_f(\tau)$ is constant during the entire signal duration:

$$\omega_f(\tau) = \omega_{mid} + \omega'(\tau - t_{mid}) \quad \text{and} \quad \zeta_f(\tau) = \zeta_f. \tag{5.27}$$

In the above equation, t_{mid} is the instant at which 45% of the expected Arias intensity I_a is reached, ω_{mid} is the filter frequency at instant t_{mid} and ω' is the slope of linear evolution of $\omega_f(\tau)$. After being normalized by the standard deviation $\sigma_h(t)$, the integral in Eq. (5.25) becomes a unit variance process with time-varying frequency and constant bandwidth, which represents the spectral non-stationarity of the ground motion.

The non-stationarity in intensity is then captured by the modulation function $q(t, \boldsymbol{\alpha})$.

This time-modulating function determines the shape, intensity and duration of the motion as follows:

$$q(t, \boldsymbol{\alpha}) = \alpha_1 t^{\alpha_2 - 1} \exp(-\alpha_3\, t). \tag{5.28}$$

The vector of parameters $\boldsymbol{\alpha} = (\alpha_1,\ \alpha_2,\ \alpha_3)$ is directly related to the physical characteristics of the ground motion, namely the expected Arias intensity I_a, the time interval D_{5-95} between the instants at which the 5% and 95% of I_a are reached and the instant t_{mid}.

In the discrete time domain, the synthetic ground motion in Eq. (5.25) becomes:

$$\hat{x}(t) = q(t, \boldsymbol{\alpha}) \sum_{i=1}^{n} s_i(t, \lambda(t_i)) U_i, \tag{5.29}$$

where the standard normal random variable U_i represents an impulse at instant t_i and $s_i(t, \lambda(t_i))$ is given by:

$$\begin{aligned} s_i(t, \lambda(t_i)) &= \frac{h[t - t_i, \lambda(t_i)]}{\sqrt{\sum_{j=1}^{k} h^2[t - t_j, \lambda(t_j)]}} \text{ for } t_i < t_k,\ t_k \leq t < t_{k+1}, \\ &= 0 \quad \text{for} \quad t \leq t_i. \end{aligned} \tag{5.30}$$

Table 5.2 lists the probabilistic distributions associated with the uncertain parameters. The six parameters describing the ground motion are considered dependent with a Nataf distribution (a.k.a Gaussian copula) (Liu and Der Kiureghian, 1986; Lebrun and Dutfoy, 2009). The correlation matrix is given in Table 5.3.

Table 5.2 – Marginal distributions of the stochastic ground motion model (after Rezaeian and Der Kiureghian (2010a)).

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
ζ_f (.)	Beta	[0.02, 1]	0.213	0.143

Table 5.3 – Correlation matrix of the Nataf distribution of the stochastic ground motion model (after Rezaeian and Der Kiureghian (2010a)).

	I_a	D_{5-95}	t_{mid}	ω_{mid}	ω'	ζ_f
I_a	1	-0.36	0.01	-0.15	0.13	-0.01
D_{5-95}	-0.36	1	0.67	-0.13	-0.16	-0.2
t_{mid}	0.01	0.67	1	-0.28	-0.2	-0.22
ω_{mid}	-0.15	-0.13	-0.28	1	-0.2	0.28
ω'	0.13	-0.16	-0.2	-0.2	1	-0.01
ζ_f	-0.01	-0.2	-0.22	0.28	-0.01	1

5.3.3 Bouc-Wen oscillator subject to Northridge ground motions

5.3.3.1 Problem statement

Let us consider the SDOF Bouc-Wen oscillator (Kafali and Grigoriu, 2007) subject to stochastic excitation. The equation of motion of the oscillator reads:

$$\begin{cases} \ddot{y}(t) + 2\,\zeta\,\omega\,\dot{y}(t) + \omega^2(\rho\,y(t) + (1-\rho)\,z(t)) = -x(t), \\ \dot{z}(t) = \gamma\dot{y}(t) - \alpha\,|\dot{y}(t)|\,|z(t)|^{n-1}\,z(t) - \beta\,\dot{y}(t)\,|z(t)|^{n}, \end{cases} \tag{5.31}$$

in which ζ is the damping ratio, ω is the fundamental frequency, ρ is the post- to pre-yield stiffness ratio, γ, α, β, n are parameters governing the hysteretic loops. In this example, the excitation $x(t)$ is a ground motion generated by means of the probabilistic model presented in Section 5.3.2 with deterministic properties extracted from the component 090 of the Northridge earthquake recorded at the LA 00 station, *i.e.* $(I_a, D_{5-95}, t_{mid}, \omega_{mid}, \omega', \zeta_f) = (0.109,\ 7.96,\ 7.78,\ 4.66 \times 2\,\pi, -0.09 \times 2\,\pi, 0.24)$ (Rezaeian and Der Kiureghian, 2010b). More precisely, this set of parameters is used together with random white-noise processes to generate a series of artificial motions.

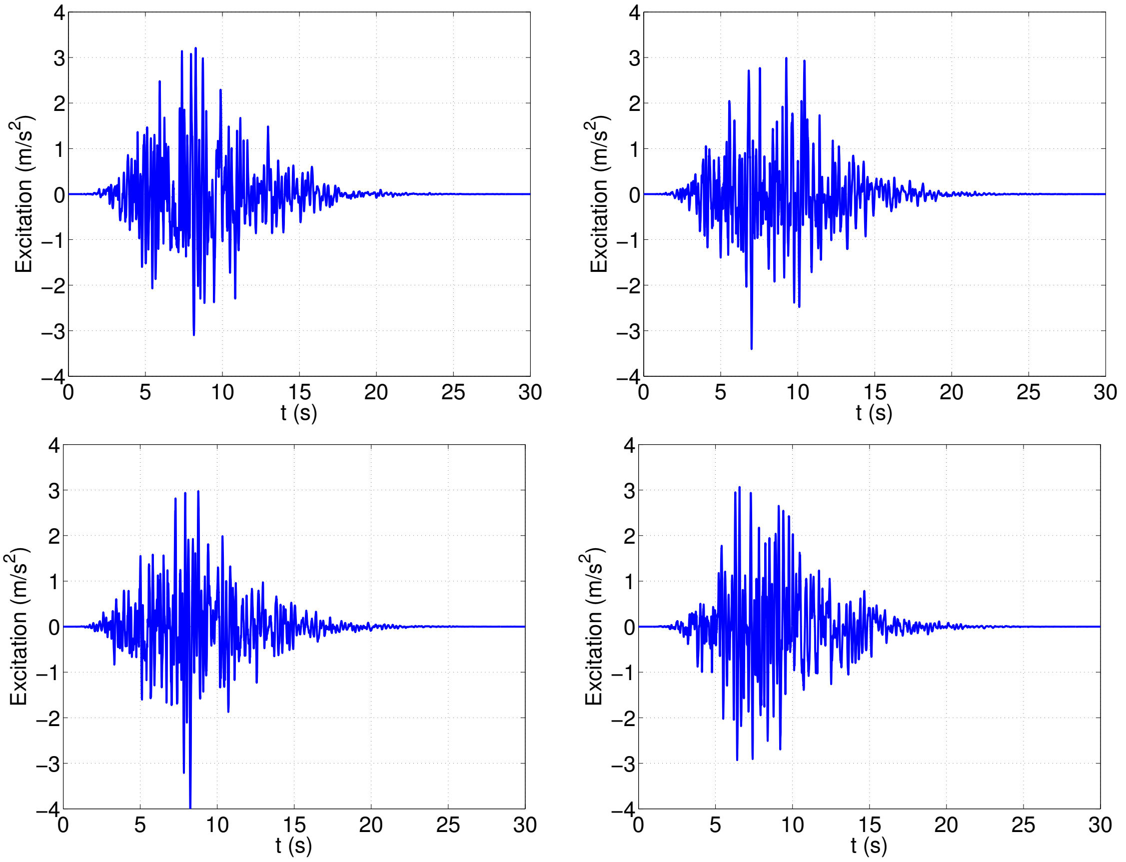


Figure 5.10 – Bouc Wen oscillator – Example synthetic motions generated using the parameters of a single recorded motion (Northridge earthquake, LA 00 station).

Deterministic values are used for the following parameters of the Bouc-Wen model: $\zeta = 0.02$, $\rho = 0$, $\gamma = 1$, $n = 1$, $\beta = 0$. The remaining parameters are considered independent random variables with associated distributions given in Table 5.4, which constitute the vector of uncertain parameters $\boldsymbol{\xi} = (\omega, \alpha)$.

Table 5.4 – Marginal distributions of the Bouc-Wen model parameters

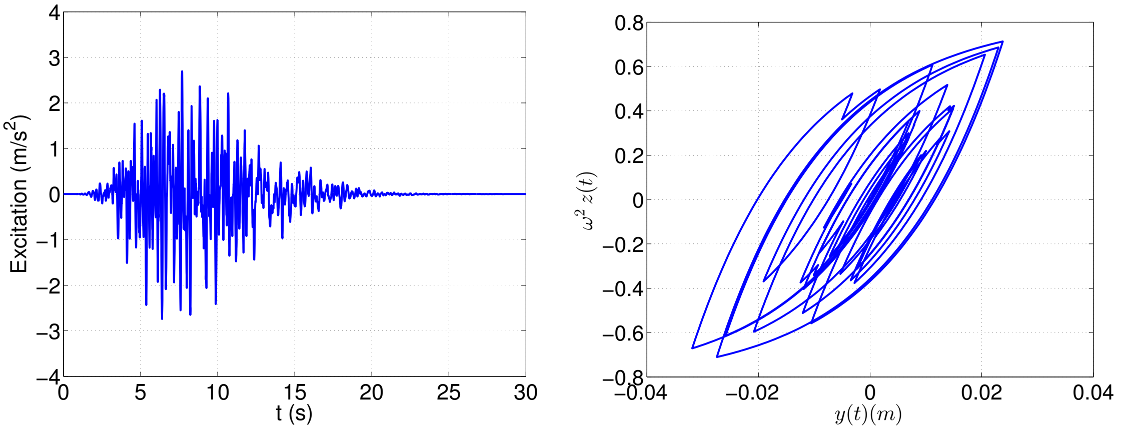
Parameter	Distribution	Support	Mean	Standard deviation
ω (rad/s)	Uniform	[5.373, 6.567]	5.97	0.3447
α (1/m)	Uniform	[45, 55]	50	2.887

5.3.3.2 PC-NARX model

We first build the metamodel for representing the velocity time histories $v(t)$ of the oscillator. 50 simulations are conducted with 50 samples of the input parameters generated by Latin hypercube sampling. The system of ODEs are solved with the Matlab solver `ode45` (explicit Runge-Kutta method with relative error tolerance 1×10^{-5}) with the total duration $T = 30$ s and time step $\mathrm{d}t = 0.005$ s. In the first place, a NARX model structure is chosen, in which the model terms are $g_i(t) = x(t-k)^l \, |v(t-1)|^m$ and $g_i(t) = v(t-j)^l \, |v(t-1)|^m$ with $l = 0, 1$, $m = 0, 1$, $j = 1, \ldots, 4$, $k = 0, \ldots, 4$. The use of absolute terms has proven effective in capturing the hysteretic behaviour of nonlinear systems in Spiridonakos and Chatzi (2015a). The initial NARX model contains 19 terms in total.

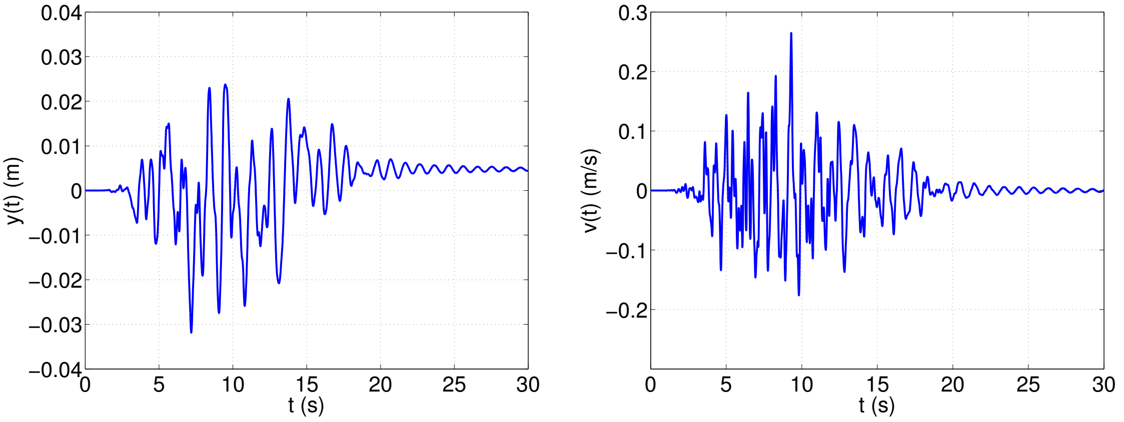
Next, the candidate NARX models are computed. For this purpose, we selected the simulations with maximum velocity exceeding a large threshold, *i.e.* $\max(|v(t)|) > 0.25$ m/s and obtained 17 experiments. LARS was applied to the initial full NARX model to detect the most relevant NARX terms constituting a candidate NARX model from each simulation previously selected. This procedure resulted in 9 candidates in total, which is due to the fact that the same NARX candidate is obtained from different simulations. Ordinary least squares (Eq. (5.13)) is used to determine the NARX coefficients corresponding to each NARX candidate model for all the simulations. To evaluate the accuracy of the NARX candidate, Eq. (5.16) is used to compute the error indicators. The most appropriate NARX model achieves a mean relative error $\bar{\epsilon} = 4.1 \times 10^{-4}$ over 50 experiments and contains 12 terms, namely constant term, $x(t-4)$, $x(t-3)$, $x(t-2)$, $x(t-2)\,|v(t-1)|$, $x(t-1)$, $x(t-1)\,|v(t-1)|$, $x(t)$, $v(t-4)$, $v(t-4)\,|v(t-1)|$, $v(t-1)\,|v(t-1)|$, $v(t-1)$. Figure 5.11 depicts the experiment from which the most appropriate NARX model is selected. Note that the nonlinear behaviour is noticeable and the oscillator exhibits a residual displacement after entering the domain of nonlinearity.

Then we represented the NARX coefficients by sparse adaptive PCEs with degree up to 20, maximum interaction order $r = 2$ and truncation parameter $q = 1$. The PCEs of the NARX coefficients have LOO errors smaller than 5×10^{-6}. The PC-NARX model of the velocity was obtained and used for predicting the velocity on the validation set. The displacement time history is then obtained by integration.



(a) Excitation

(b) Nonlinear behavior



(c) Displacement

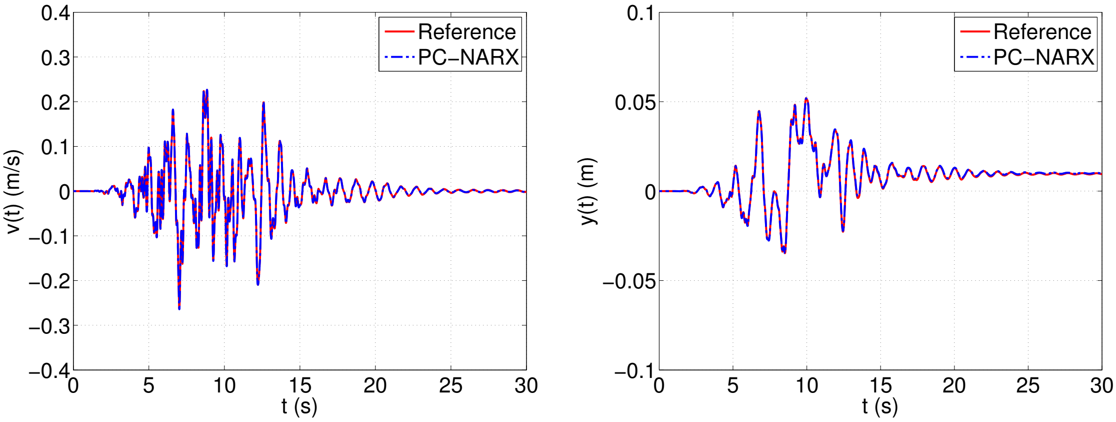
(d) Velocity

Figure 5.11 – Bouc Wen oscillator – Experiment leading to the selected NARX structure. $\boldsymbol{\xi} = (6.486, 50.9247)$

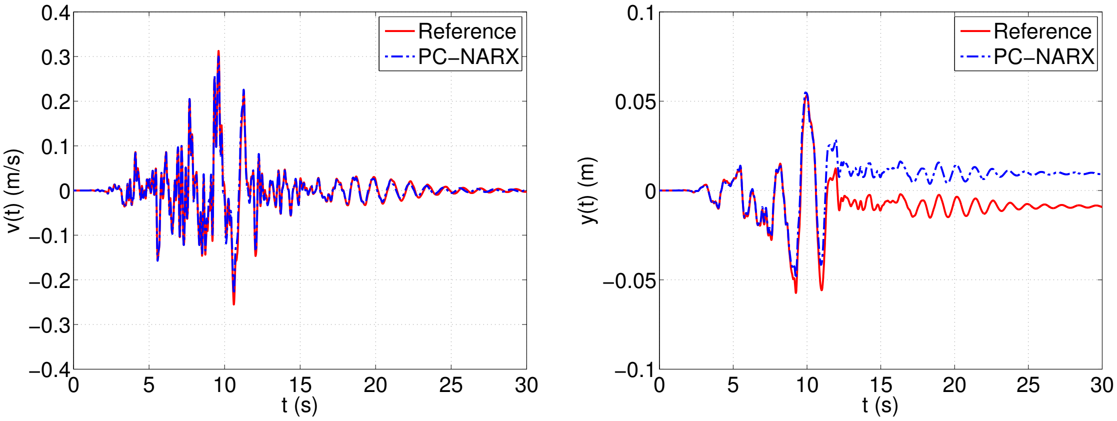
5.3.3.3 Prediction of new trajectories and statistical moments

Figure 5.12 depicts two specific velocity and displacement trajectories for distinct validation sets of parameters. One observes that the velocity trajectories are perfectly predicted by PC-NARX. Indeed, the mean relative error over 10^4 validations is $\bar{\epsilon}_{val} = 1.1 \times 10^{-3}$. Despite the high accuracy of the PC-NARX model for the velocity, the displacements obtained with the integration might exhibit some slight discrepancies with respect to the actual trajectories. The inaccuracies occur when the peak values of the velocity are not well predicted and the error is accumulated through integration in time. Over the validation set, the mean relative error is $\bar{\epsilon}_{val} = 7.7 \times 10^{-3}$. Note that predicting the response of a mechanical system subject to *nonstationary excitation* is never an easy task. From our experience, it is thus of no interest to apply time-frozen PCEs to this type of problems.

Figure 5.13 compares the time-dependent standard deviation of the two response quantities predicted by PC-NARX with those obtained from Monte Carlo simulation. The standard deviation of the velocity and the displacement is remarkably well captured by



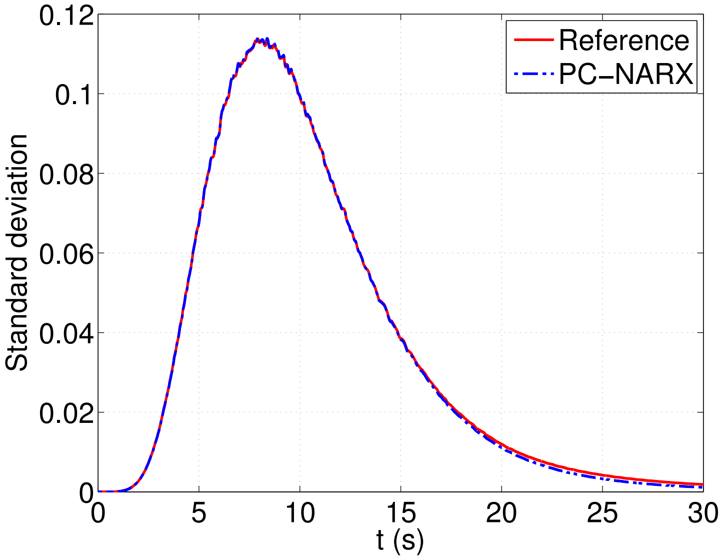
First example trajectory $\boldsymbol{\xi} = (6.5164, 47.9249)$



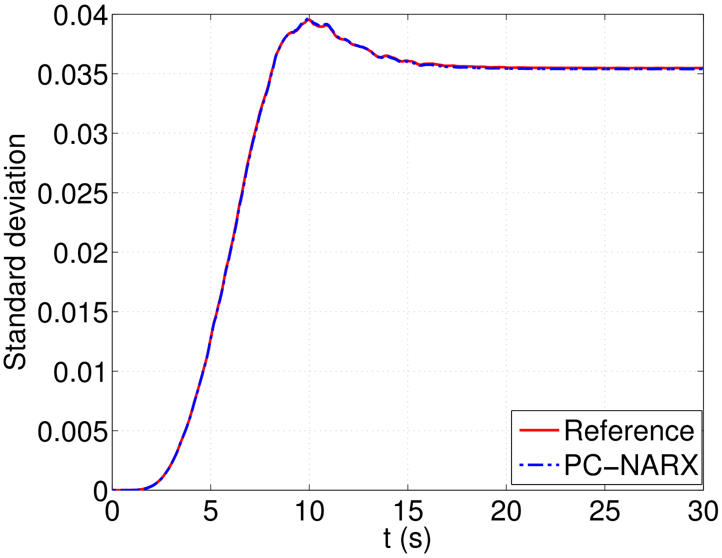
Second example trajectory $\boldsymbol{\xi} = (5.38, 45.0578)$

Figure 5.12 – Bouc-Wen oscillator – Two particular trajectories of velocity $v(t)$ and displacement $y(t)$ and their predictions by means of PC-NARX.

PC-NARX, with relative errors being $\epsilon_{std,vel} = 2.43 \times 10^{-4}$ and $\epsilon_{std,disp} = 3.84 \times 10^{-5}$, respectively. For the displacement, the standard deviation increases in time, reaches its peak before stabilizing in a plateau, which is related to residual displacements. This plateau at the late instants is not observed for the standard deviation of the velocity, which always returns to zero after the shaking. In addition, the displacement's standard deviation reaches its maximum 2 seconds later than the velocity. It can be concluded that the different response quantities of a system might experience their largest variabilities at different instants. Given the time-dependent statistics of a response quantity, one cannot derive the time instant at which the other response varies the most. Most interestingly, the standard deviation trajectories for both the velocity and the displacement are relatively smooth with only minor perturbations recorded around their peaks.



(a) Standard deviation of velocity $v(t)$

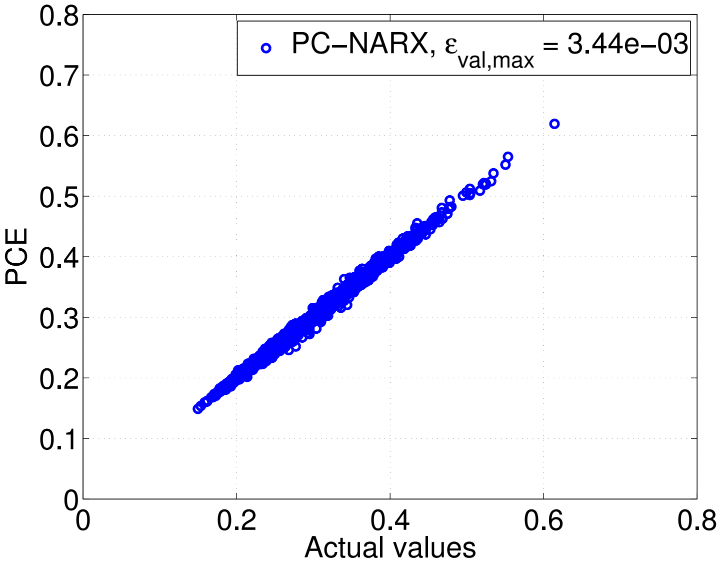


(b) Standard deviation of displacement $y(t)$

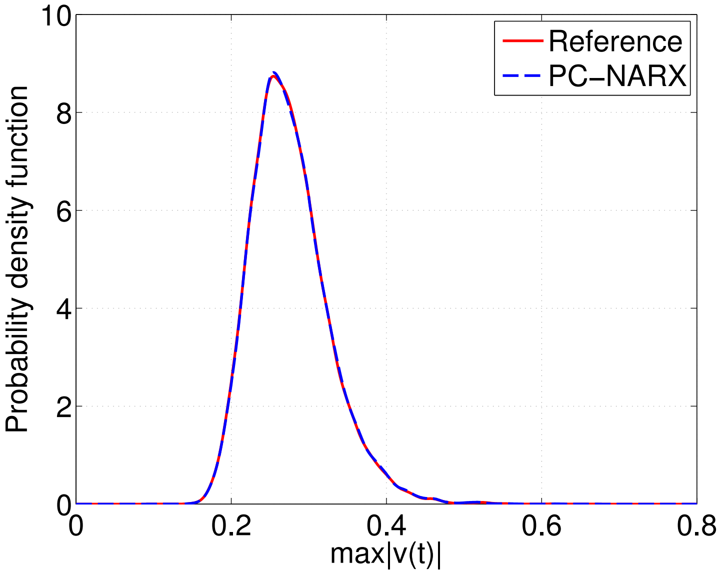
Figure 5.13 – Bouc-Wen oscillator – Standard deviation trajectories of the responses.

5.3.3.4 Prediction of maximal values of the responses

It has been shown that the PC-NARX surrogate model provide excellent predictions of the oscillator's response time histories. Consequently, it is expected that the maximal values of the responses can be captured at a high level of accuracy. Figure 5.14 (resp. Figure 5.15) plots the maximal velocities (resp. displacements) predicted by PC-NARX against the actual values obtained with the numerical solver. The figures show an excellent agreement between the predictions and the true values. As a result, the predicted PDFs of the maximal values perfectly match the reference functions.



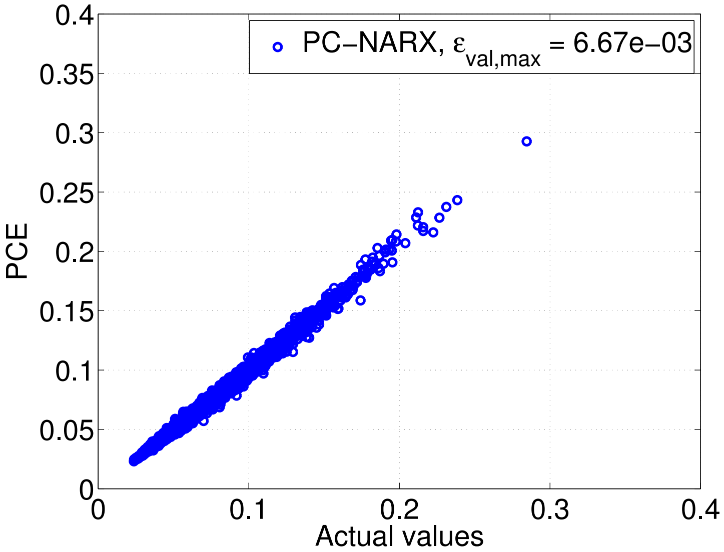
(a) Maximal values of velocity



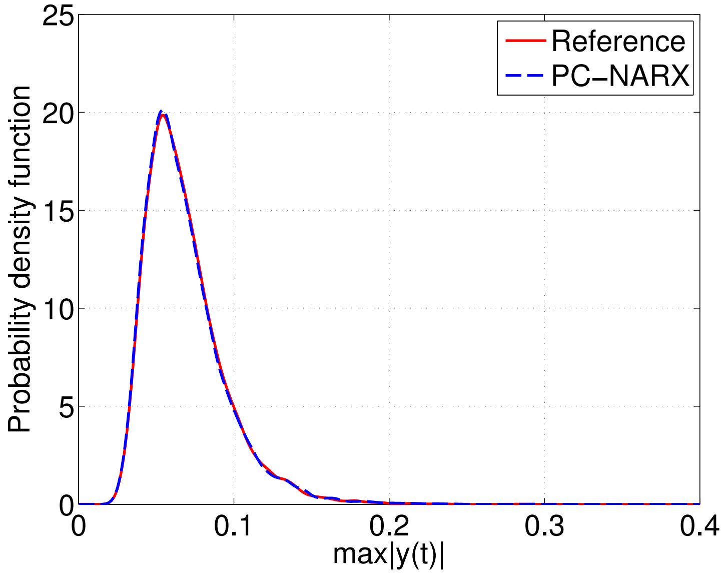
(b) Probability density function of maximal values of velocity

Figure 5.14 – Bouc Wen oscillator – Maximal responses.

In this example, we considered an oscillator with random mechanical properties subject to seismic motions mimicking a real accelerogram. Therein, the time histories of the excitations are not exactly the same, however they incorporate similar features regarding the evolution of frequency, the energy content, etc. In practical applications, it is of utmost importance to also take into account the effect of uncertainties from the excitations corresponding to various earthquake scenarios. This can be obtained by randomizing the parameters of the probabilistic seismic ground motion model in accordance with their distributions obtained from statistical analysis on real events. This topic will be investigated in the following applications.



(a) Maximal values of displacement



(b) Probability density function of maximal values of displacement

Figure 5.15 – Bouc Wen oscillator – Maximal responses.

5.3.4 Duffing oscillator subject to synthetic ground motions

5.3.4.1 Problem statement

For the sake of variety, let us consider a SDOF Duffing oscillator (Kafali and Grigoriu, 2007). Note that the vibration of a Duffing oscillator is a commonly used benchmark in the literature, see *e.g.* Orabi and Ahmadi (1987); Lucor and Karniadakis (2004); Kougioumtzoglou and Spanos (2009); Spiridonakos and Chatzi (2015a); Chouvion and

Sarrouy (2016) in which an oscillator subject to either periodic or white-noise external force is considered. The dynamics of the oscillator can be described by the following equation of motion:

$$\ddot{y}(t) + 2\,\zeta\,\omega\,\dot{y}(t) + \omega^2\,(y(t) + \epsilon\,y(t)^3) = -x(t), \tag{5.32}$$

in which $y(t)$ is the oscillator displacement, ζ is the damping ratio, ω is the fundamental frequency, ϵ is the parameter governing the nonlinear behaviour and $x(t)$ is the excitation which is herein generated by the probabilistic ground motion model proposed by Rezaeian and Der Kiureghian (2010a).

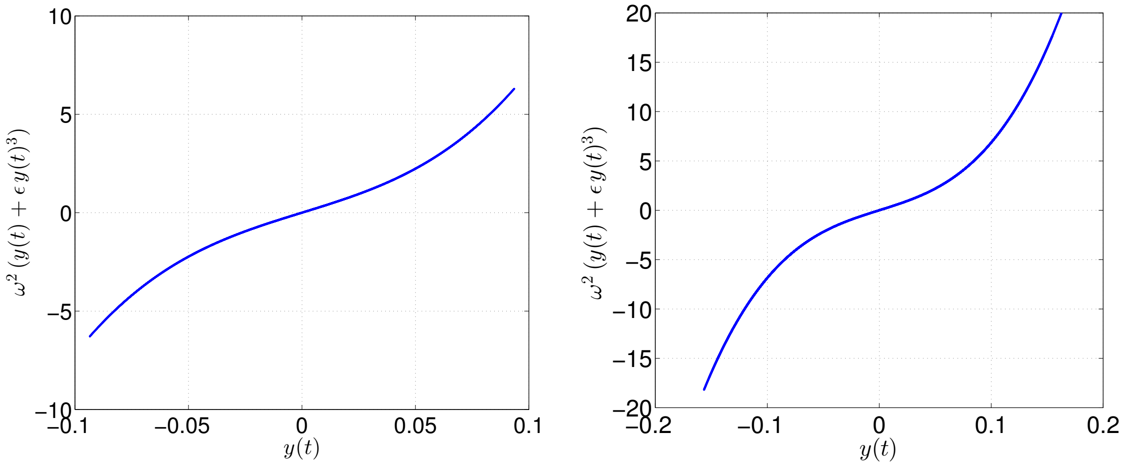
The parameters ζ and ω of the SDOF oscillator are considered deterministic with values $\zeta = 0.02$ and $\omega = 5.97$ rad/s whereas ϵ is modelled by a uniform random variable on the support domain $[90, 110]$. The uncertain input vector contains parameters of the oscillator and parameters representing the intensity and frequency features of the ground motion model $\boldsymbol{\xi} = (\epsilon, I_a, D_{5-95}, t_{mid}, \omega_{mid}, \omega', \zeta_f)$, as described in the previous section.

5.3.4.2 PC-NARX model

We now aim at building the metamodel for representing the displacement time histories $y(t)$. For this purpose, $N = 200$ analyses of the system were conducted with the input parameters generated by Latin hypercube sampling. The equation of motion was solved with the Matlab solver `ode45` (explicit Runge-Kutta method). Note that for the sake of consistency, all the synthetic motions were generated with the total duration $T = 30$ s and time step $\mathrm{d}t = 0.005$ s. In the first place, a NARX model structure was chosen in which the model terms are polynomial functions of past values of the output and excitation $g_i(t) = y^l(t-j)\,x^m(t-k)$ with $l+m \leq 3$, $0 \leq l \leq 3$, $0 \leq m \leq 1$, $j = 1, 2$, $k = 0, 1, 2$. The chosen full NARX model contains 10 terms.

Next, candidate NARX models were computed. To this end, we selected the simulations with maximum displacement exceeding a large threshold, *i.e.* $\max(|y(t)|) > 0.07$ m, leading to 19 selected experiments. Figure 5.16 depicts the nonlinear behaviour of the oscillator in two experiments selected for extracting NARX candidates. For each simulation previously selected, LARS was applied to the initial full NARX model to detect the most relevant NARX terms constituting a candidate NARX model. This procedure resulted in 12 candidates in total. The NARX coefficients corresponding to each candidate model are then computed for each simulation in the ED by means of ordinary least squares (Eq. (5.13)). The responses are reconstructed using the computed coefficients, leading to the errors ϵ_k (Eq. (5.16)). The best NARX model achieves a mean relative error $\bar{\epsilon} = 7.4 \times 10^{-4}$ over 200 experiments and contains 7 terms, namely the constant term, $x(t-2)$, $x(t)$, $y(t-2)$, $y(t-1)$, $y^2(t-2)$, $y^3(t-1)$.

In the next step, we represented the NARX coefficients by adaptive sparse PCEs of order up to 20 with maximum interaction rank $r = 2$ and truncation parameter $q = 1$. The PCEs of the NARX coefficients have LOO errors smaller than 7.34×10^{-4}. The optimal PCE selected by the adaptive scheme is of total degree 3.



(a) Experiment leading to the selected NARX structure (b) Nonlinear behaviour exhibited in a different experiment

Figure 5.16 – Duffing oscillator – Nonlinear behaviour of the oscillator in the selected experiments.

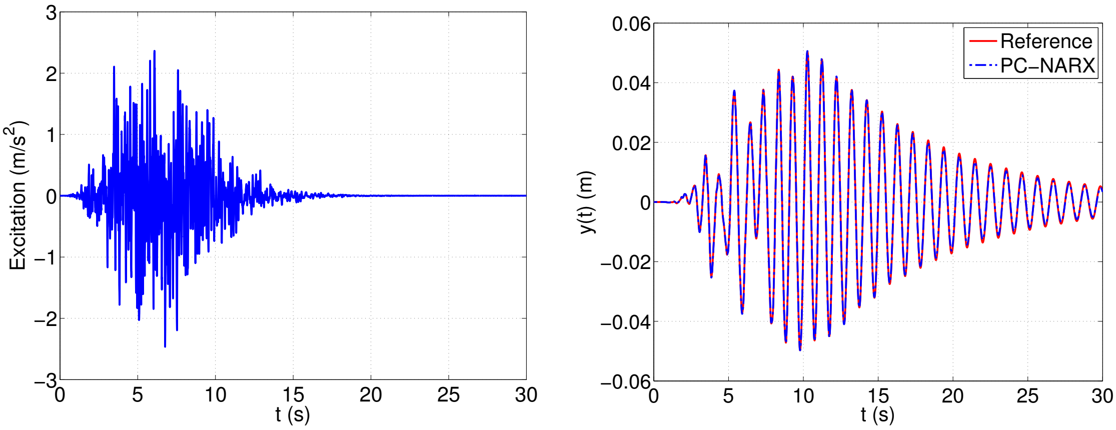
5.3.4.3 Prediction of new trajectories and statistical moments

A validation set of size $n_{val} = 10^4$ was pre-computed by the numerical Matlab solver. The displacements were then predicted by the PC-NARX model. Figure 5.17 depicts three specific excitations and the corresponding response time histories. PC-NARX provides predictions that are in remarkable agreement with the actual responses. Over 10^4 validation trajectories, the mean relative error is $\bar{\epsilon}_{val} = 3.53 \times 10^{-2}$. Less than 5% of those simulations exhibit a relative error $\epsilon_{val,i}$ exceeding 0.1.

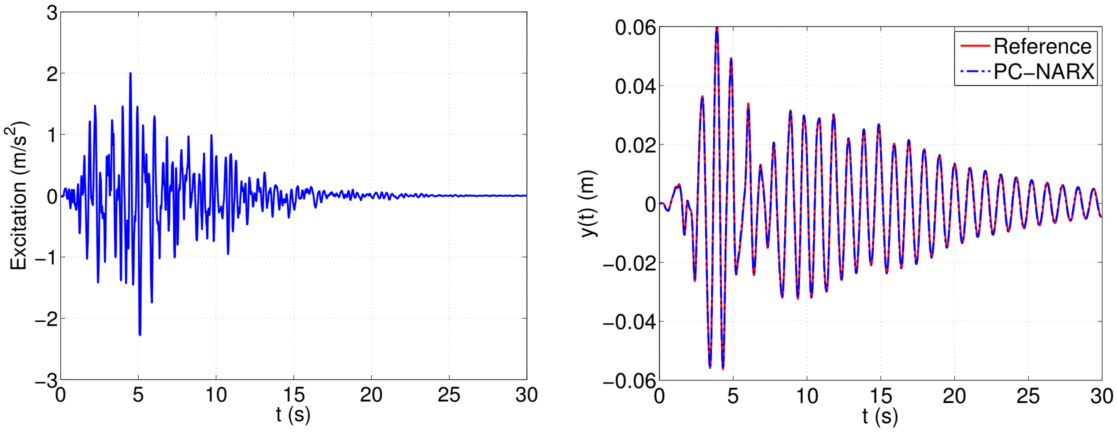
Figure 5.18 presents the evolutionary standard deviation of the displacement. It is noticeable that this function is no longer smooth, especially the high frequency content is manifest from early instants until the end of the considered time range. This might be related to the varying properties of the excitations, which lead to strong differences between the motions and consequently variabilities between the responses regarding their important phases, their amplitudes, etc. In addition, the Duffing oscillator does not exhibit residual displacement, thus after reaching its peak the standard deviation tends to decrease constantly and does not show a plateau as the Bouc-Wen oscillator. In general, the curve computed by PC-NARX is in good agreement with the reference one with a relative error $\epsilon_{val,Std} = 0.5 \times 10^{-2}$. The mean trajectory, which is slightly fluctuating around zero, is not informative, therefore not presented herein.

5.3.4.4 Prediction of maximal values of the response

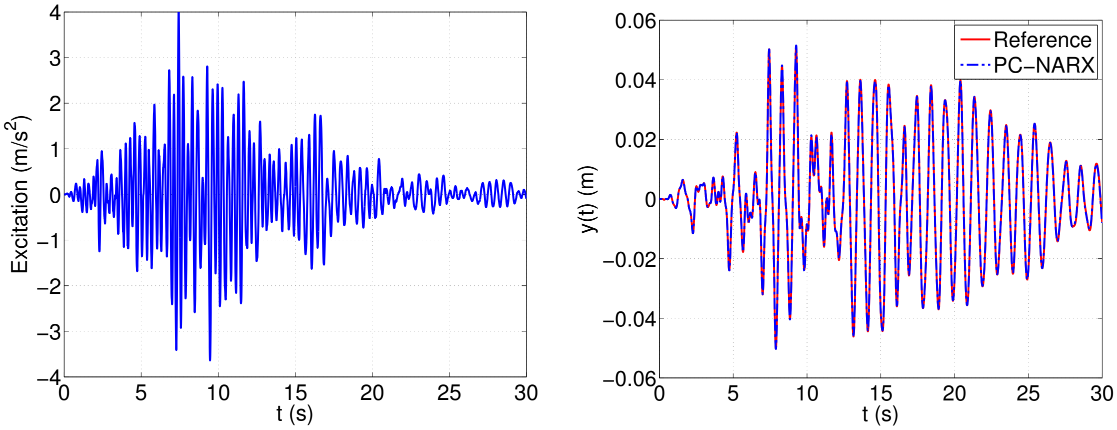
Finally, we compare the maximal displacements predicted by PC-NARX with the actual values computed by the numerical solver. Figure 5.19 shows that overall the predictions are quite accurate with a relative error $\epsilon_{val,max} = 0.7 \times 10^{-2}$. The simulations with large



First example trajectory $\boldsymbol{\xi} = (106.30, 0.05, 6.53, 5.47, 55.16, -1.02, 0.31)$



Second example trajectory $\boldsymbol{\xi} = (105.84, 0.05, 9.08, 4.53, 19.76, -0.11, 0.28)$



Third example trajectory $\boldsymbol{\xi} = (90.68, 0.18, 15.86, 7.98, 19.5376, -0.37, 0.07)$

Figure 5.17 – Duffing oscillator – Three particular excitations (left column), associated response trajectories and their prediction by means of PC-NARX (right column).

discrepancies between the prediction and the actual values correspond to large displacements, *i.e.* they belong to the domain of rare events (the upper tail of the distribution) that was not well represented in the small training set of size 200.

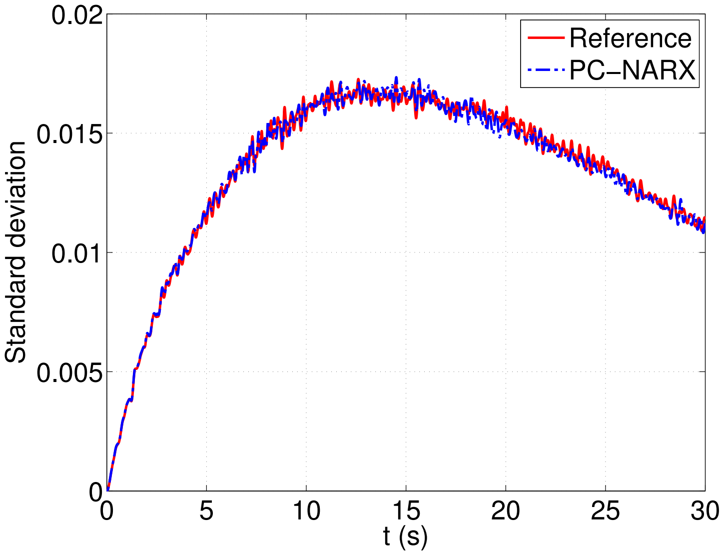
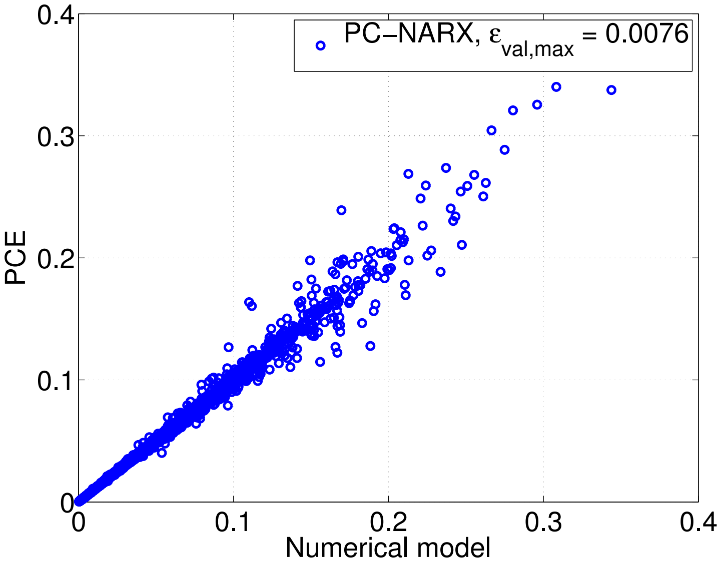
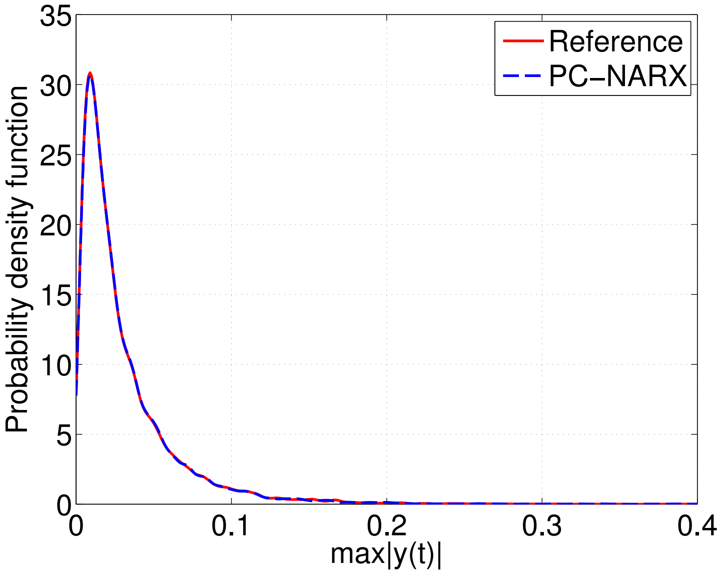


Figure 5.18 – Duffing oscillator – Standard deviation trajectory of the displacement.



(a) PCE-based predictions *vs.* actual values



(b) Probability density function

Figure 5.19 – Duffing oscillator – Maximal displacements.

5.3.5 Bouc-Wen oscillator subject to synthetic ground motions

5.3.5.1 Problem statement

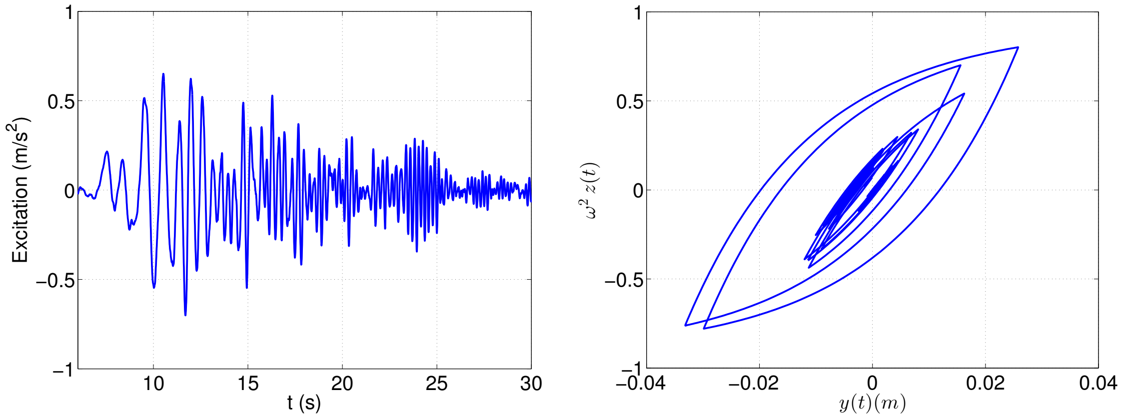
In this application, we re-investigate the SDOF Bouc-Wen oscillator (Kafali and Grigoriu, 2007). The difference compared to the previous study is that, the stochastic excitations are generated by the probabilistic ground motion model with the governing parameters being randomly sampled in accordance with their distributions reported by Rezaeian and Der Kiureghian (2010a). One assumes the same random properties of the oscillator, *i.e.* $\omega \sim \mathcal{U}[5.373, 6.567]$ and $\alpha \sim \mathcal{U}[45, 55]$. The uncertainties in the considered system is therefore characterized by means of the vector of uncertain parameters $\boldsymbol{\xi} = (\omega, \alpha, I_a, D_{5-95}, t_{mid}, \omega_{mid}, \omega', \zeta_f)$.

5.3.5.2 PC-NARX model

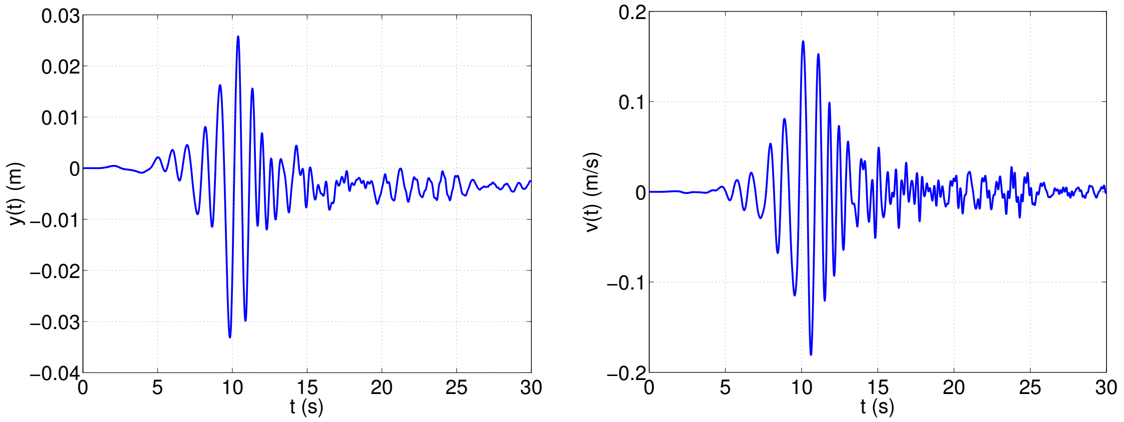
We first build the metamodel for representing the velocity time histories $v(t)$ of the oscillator. 200 simulations are conducted with 200 samples of the input parameters generated by Latin hypercube sampling. The system of ODEs are solved with the Matlab solver *ode45* (explicit Runge-Kutta method with relative error tolerance 1×10^{-3}) with the total duration $T = 30$ s and time step $\mathrm{d}t = 0.005$ s as in the previous example. In the first place, a NARX model structure is chosen, in which the model terms are $g_i(t) = x(t-k)^l \, |v(t-1)|^m$ and $g_i(t) = v(t-j)^l \, |v(t-1)|^m$ with $l = 0, 1$, $m = 0, 1$, $j = 1, \ldots, 4$, $k = 0, \ldots, 4$. The initial NARX model contains 19 terms in total.

Next, the candidate NARX models are computed. For this purpose, we selected the simulations with maximum velocity exceeding a large threshold, *i.e.* $\max(|v(t)|) > 0.25$ m/s and obtained 15 experiments. LARS was applied to the initial full NARX model to detect the most relevant NARX terms constituting a candidate NARX model from each simulation previously selected. This procedure resulted in 11 candidates in total. Ordinary least squares (Eq. (5.13)) is used to determine the NARX coefficients corresponding to each NARX candidate model for all the simulations. To evaluate the accuracy of the NARX candidate, Eq. (5.16) is used to compute the error indicators. The most appropriate NARX model achieves a mean relative error $\bar{\epsilon} = 6.27 \times 10^{-4}$ over 200 experiments and contains 12 terms, namely constant term, $x(t-4)$, $x(t-4)\,|v(t-1)|$, $x(t-3)$, $x(t-3)\,|v(t-1)|$, $x(t-2)$, $x(t-1)$, $x(t)$, $v(t-4)$, $v(t-4)\,|v(t-1)|$, $v(t-3)\,|v(t-1)|$, $v(t-1)$. Figure 5.20 depicts the experiment from which the most appropriate NARX model is selected. Note that the nonlinear behaviour is noticeable and the oscillator exhibits a residual displacement after entering the domain of nonlinearity.

Then we represented the NARX coefficients by sparse PCEs. The optimal polynomial of order $p = 2$ was found adaptively with maximum interaction order $r = 2$ and truncation parameter $q = 1$ so that the resulting PC-NARX model led to the smallest error when reconstructing the responses in the ED. The PCEs of the NARX coefficients have LOO errors smaller than 1.68×10^{-4}. The PC-NARX model of the velocity was obtained and used for predicting the velocity on the validation set. The displacement time history is then obtained by integration.



(a) $\boldsymbol{\xi} = (6.24, 53.64, 0.04, 6.74, 7.54, 17.22, 0.08, 0.26)$ (b) Nonlinear behavior



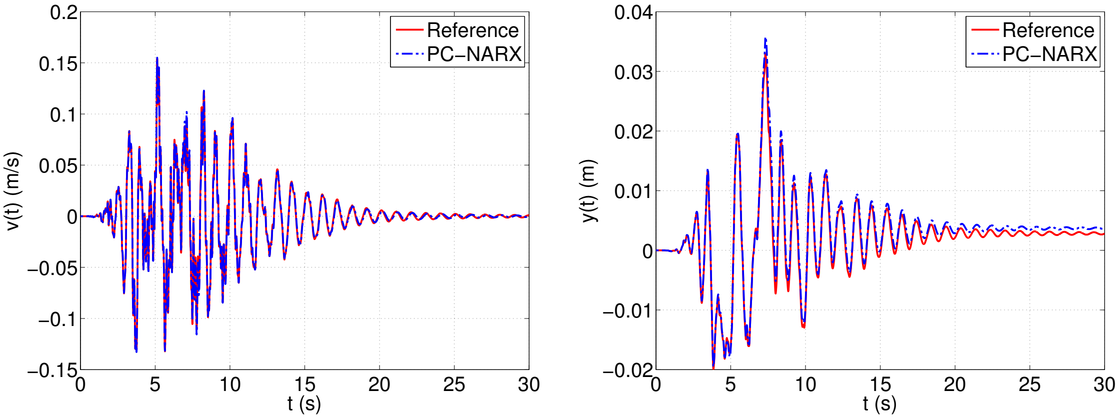
(c) Displacement (d) Velocity

Figure 5.20 – Bouc-Wen oscillator – Experiment leading to the selected NARX structure.

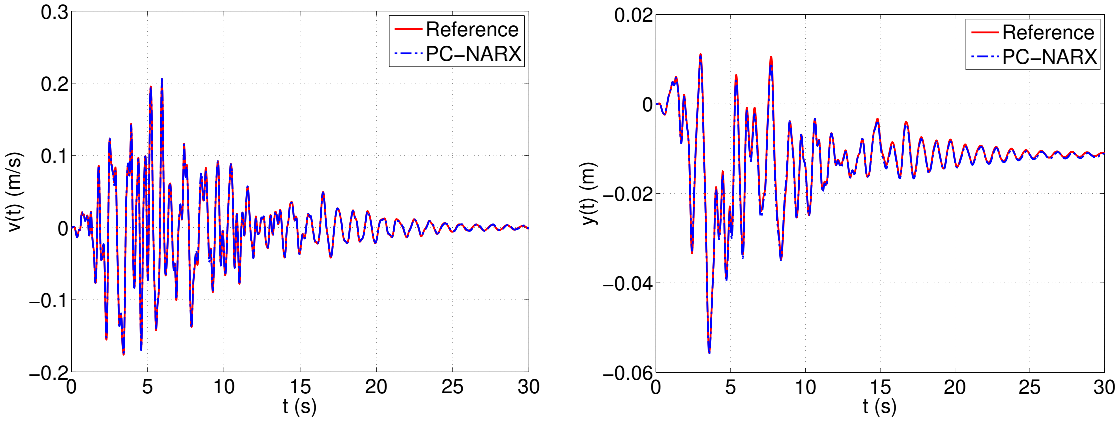
5.3.5.3 Prediction of new trajectories and statistical moments

Figure 5.21 depicts two specific velocity and displacement trajectories for distinct validation sets of parameters. One observes that the velocity trajectories are relatively well predicted by PC-NARX. Indeed, the mean relative error over 10^4 validations is $\bar{\epsilon}_{val} = 1.82 \times 10^{-2}$ and only 3% of those time-histories have a relative error $\epsilon_{val,i}$ exceeding 0.1. Despite the high accuracy of the PC-NARX model for the velocity, the predicted displacements exhibit some slight discrepancies with respect to the actual trajectories. The inaccuracies are slightly large in the later phase of the considered time domain due to the occurrence of residual displacements, but remain acceptable.

Figure 5.22 compares the time-dependent standard deviation of the two response quantities predicted by PC-NARX with those obtained from Monte Carlo simulation. The high frequency content is observed in the evolutionary response statistics. This phenomenon was also reported in the literature when applying Monte Carlo simulation to a nonlinear structure (Kougioumtzoglou and Spanos, 2013). It shows how the varying features of the seismic excitations can strongly affect the systems' responses. The standard deviation of the velocity is remarkably well captured by PC-NARX (relative error $\epsilon_{std,vel} = 0.39 \times 10^{-2}$). For the displacement, the standard deviation tends to increase in time, which is different from the Duffing oscillator that does not exhibit residual displacement. The discrepancy between the prediction and the actual time histories is also increasing in time. However, the resulting relative error remains rather small ($\epsilon_{std,disp} = 1.57 \times 10^{-2}$).



First example trajectory $\boldsymbol{\xi} = (6.35, 46.54, 0.05, 6.53, 5.47, 55.16, -1.02, 0.31)$

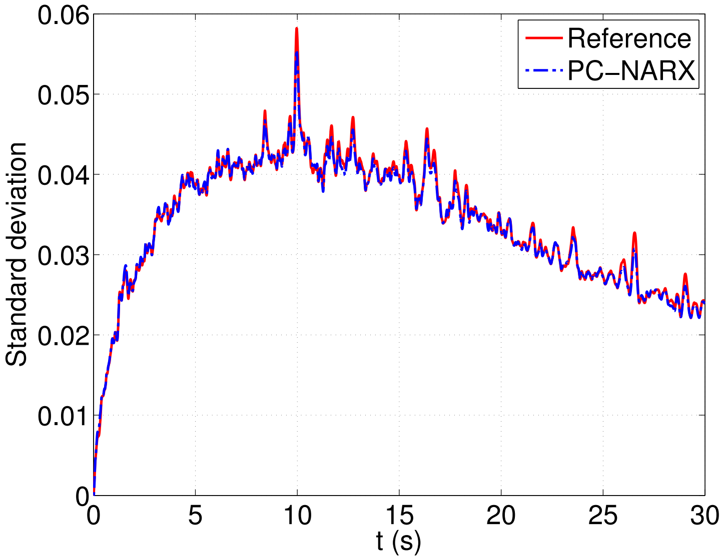


Second example trajectory $\boldsymbol{\xi} = (6.32, 48.24, 0.05, 9.08, 4.53, 19.76, -0.11, 0.28)$

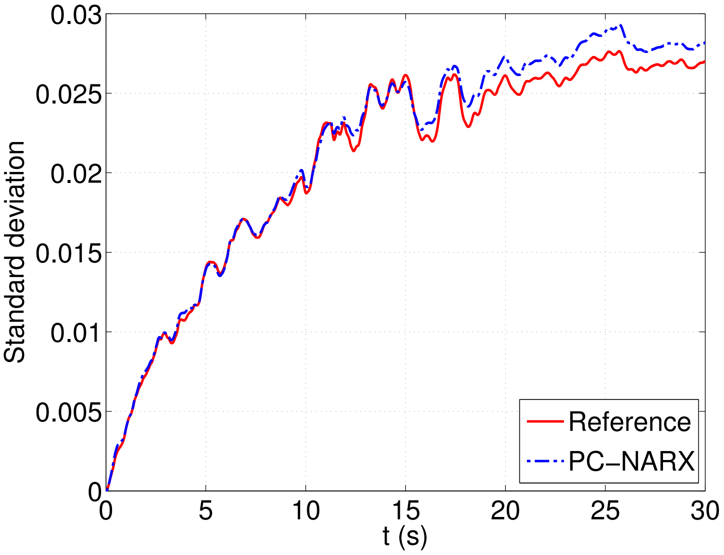
Figure 5.21 – Bouc-Wen oscillator – Two particular trajectories of velocity $v(t)$ and displacement $y(t)$ and their predictions by means of PC-NARX.

5.3.5.4 Prediction of maximal values of the responses

Figure 5.23 and Figure 5.24 compare the maximum values of velocity and displacement predicted by PC-NARX with those values computed by the numerical solver. Despite the complexity of the problem, the predictions are remarkably consistent with the true values, with sufficiently small validation errors of $\epsilon_{val,\max(|v(t)|)} = 9 \times 10^{-3}$ and $\epsilon_{val,\max(|y(t)|)} =$



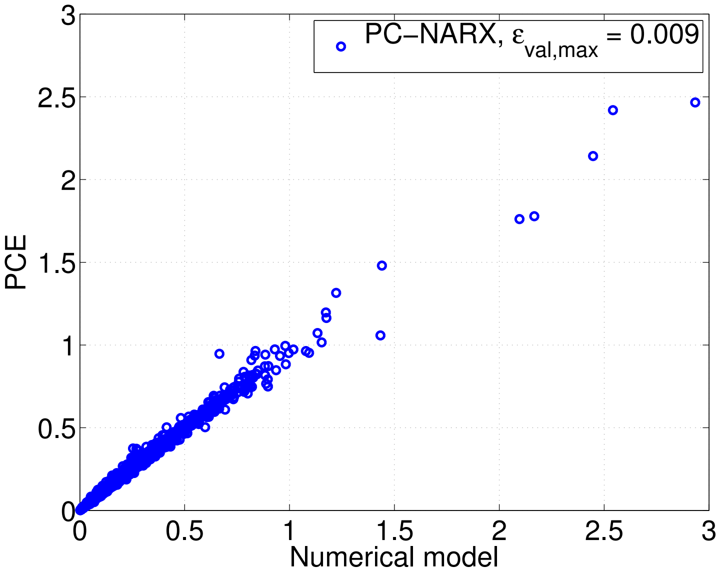
(a) Standard deviation of velocity $v(t)$



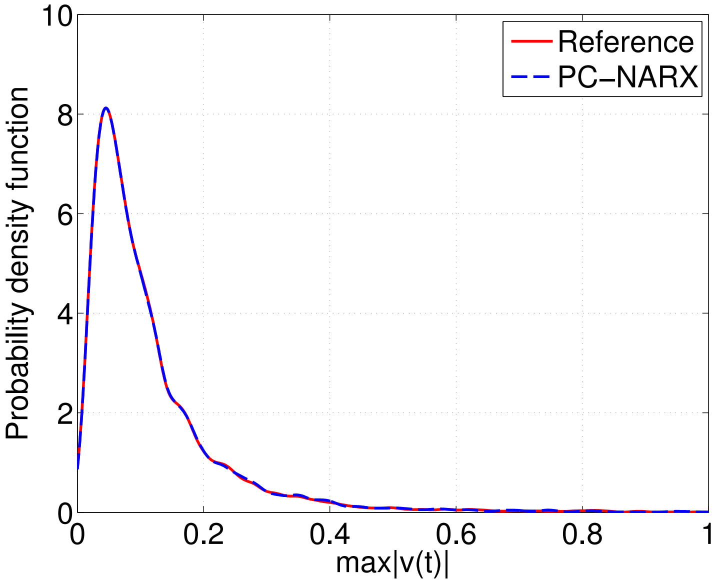
(b) Standard deviation of displacement $y(t)$

Figure 5.22 – Bouc-Wen oscillator – Standard deviation trajectories of the responses.

1.9×10^{-2}. The accurate predictions allow one to obtain the probability density functions of the maximum responses that are in good agreement with the reference functions, as shown in Figure 5.23(b) and Figure 5.24(b).

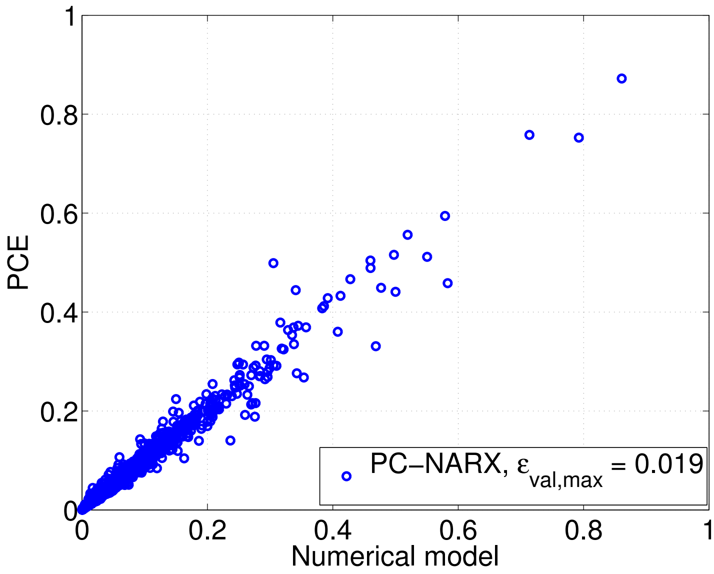


(a) Maximal values of velocity

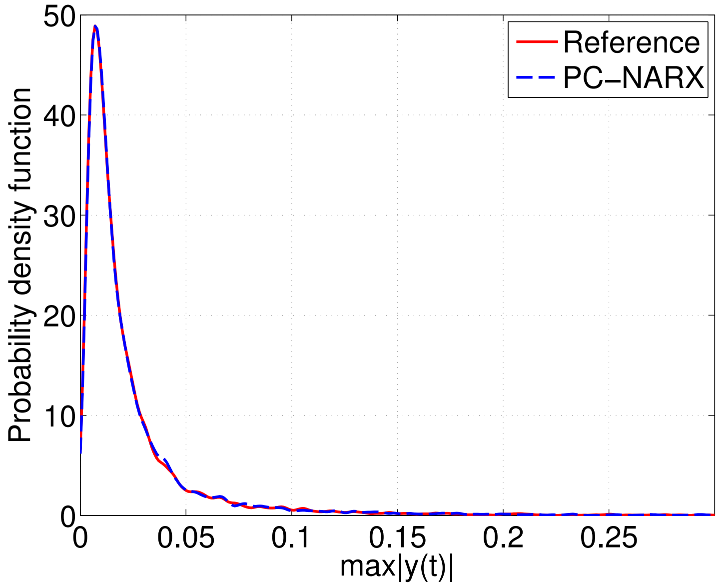


(b) Probability density function of maximal values of velocity

Figure 5.23 – Bouc-Wen oscillator – Maximal responses computed by numerical solver and predicted by PC-NARX model.



(a) Maximal values of displacement



(b) Probability density function of maximal values of displacement

Figure 5.24 – Bouc-Wen oscillator – Maximal responses computed by numerical solver and predicted by PC-NARX model.

5.3.6 Steel frame subject to synthetic ground motions

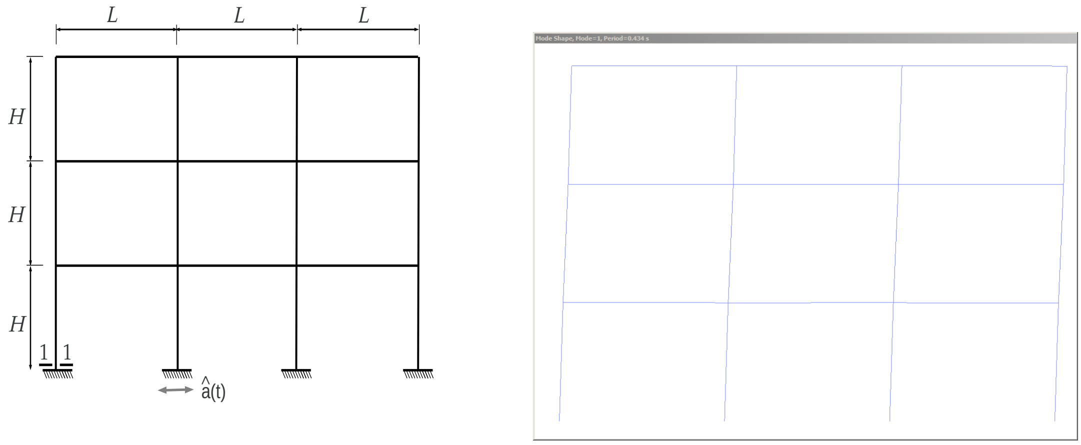
5.3.6.1 Problem statement

We consider a 3-storey 3-span steel frame structure with the following dimensions: storey-height $H = 3$ m, span-length $L = 5$ m (see Figure 5.25). The distributed load on the beams is $q = 10$ kN/m. The standard European I beams with designation IPE 300 A and IPE 330 O are chosen respectively for beams and columns. The steel material has non-linear isotropic hardening behavior following the uniaxial Giuffre-Menegotto-Pinto steel model as implemented in the finite element software OpenSees (Pacific Earthquake Engineering and Research Center, 2004). The Young's modulus E_0 and the yield strength f_y of the steel S235 are considered uncertain parameters with lognormal marginal distributions specified in Table 5.5 and are uncorrelated as recommended by the Joint Committee on Structural Safety (2001b). The strain hardening ratio, *i.e.* ratio of the post-yield tangent to the initial value, is $b = 0.01$. The parameters governing the shape of the hysteretic behavior are $R0 = 18$, $cR1 = 0.925$, $cR2 = 0.15$ (Filippou et al., 1983). The structural components are modelled with nonlinear force-based beam-column elements which allow distributed plasticity along their lengths. The connections between structural elements are modeled by rigid nodes. Rayleigh damping model is used for the structure, in which a damping ratio equal to 2% is assigned to the first two modes of the frame. When considering the material properties equal to their mean values, the vibration periods of the frame for the first three modes are respectively $T_1 = 0.434\ s$, $T_2 = 0.129\ s$ and $T_3 = 0.070\ s$. Note that the empirical formula for estimating the fundamental vibration period of steel moment-resistant buildings in Eurocode 8 (2004) and ASCE 7-98 (2000) reads $T = C_t H^{3/4}$ with $C_t = 0.035$ and $H = 9\ m = 29.52\ ft$. This leads to the empirical fundamental period for the considered frame $T = 0.443\ s$ which is approximately the value of T_1. In addition, one obtains $T_1/T_2 = 3.4$ and $T_1/T_3 = 6.2$. The periods of vibration of higher modes roughly decreases as the inverse of the odd integers, *i.e.* 1/3 and 1/5, as suggested by Housner and Jennings (1982). The first three eigenmodes of the structure are depicted in Figure 5.25. The frame is subject to in-plane stochastic excitation modelled by the probabilistic ground motion model described in Section 5.3.2 (Rezaeian and Der Kiureghian, 2010a).

Table 5.5 – Marginal distributions of the steel material properties(Joint Committee on Structural Safety, 2001b)

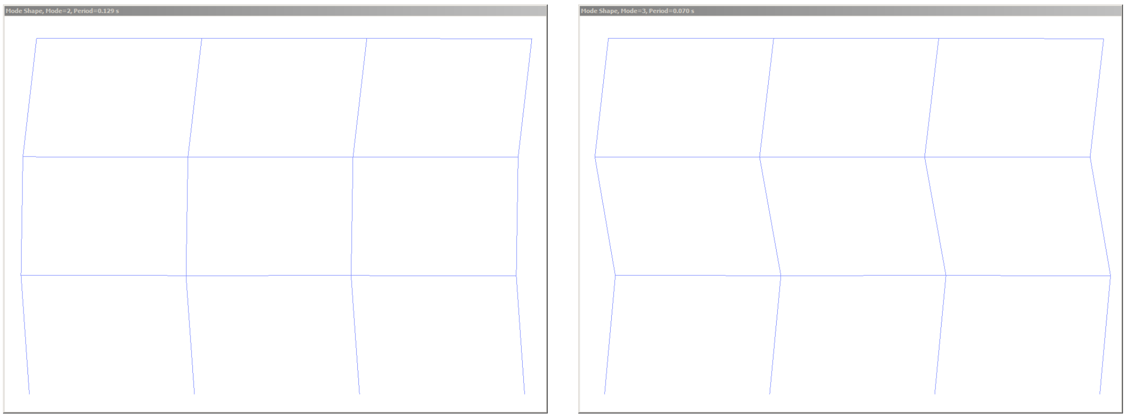
Parameter	Distribution	Mean	Standard deviation	Coefficient of variation
f_y (MPa)	Lognormal	264.2878	18.5	0.07
E_0 (MPa)	Lognormal	$210,000$	630	0.03

A sample set of size $N = 300$ of the parameters $\boldsymbol{\xi} = (f_y, E_0, I_a, D_{5-95}, t_{mid}, \omega_{mid}, \omega', \zeta)$ governing the uncertainties in the structure and excitation is obtained by means of Latin hypercube sampling. The corresponding synthetic motions are generated with the total duration $T = 30$ s and time step $\mathrm{d}t = 0.01$ s. Dynamic analyses of the finite element model subject to 300 synthetic motions are conducted and the structural responses are



(a) Steel frame structure

(b) First eigenmode, $T_1 = 0.434\ s$



(c) Second eigenmode, $T_2 = 0.129\ s$

(d) Third eigenmode, $T_3 = 0.070\ s$

Figure 5.25 – Steel frame structure and its first three eigenmodes computed with OpenSees

used for training the metamodel. We aim at computing the metamodel for the first floor displacement.

5.3.6.2 PC-ARX model

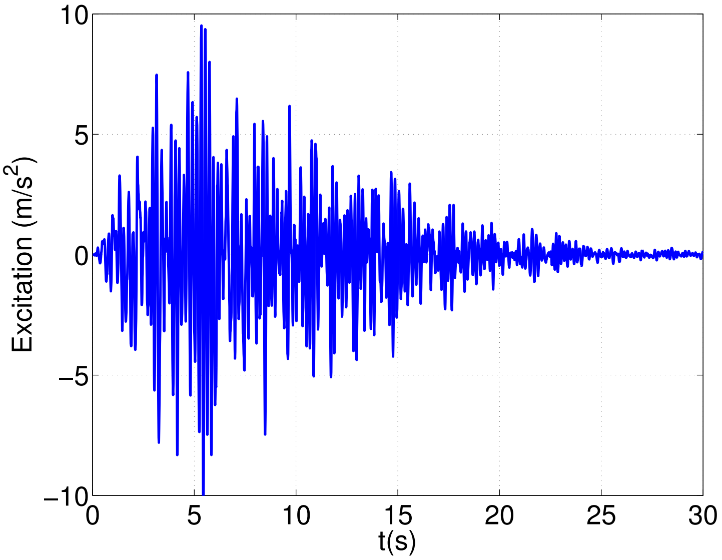
For this example, a linear ARX model structure is selected, which includes the past values of the excitation and the response $\{1, x(t), x(t-1), \dots, x(t-15), y(t-1), \dots, y(t-15)\}$. The chosen full ARX model contains 32 terms in total. Note that all the ARX terms are linear functions, while the specified time lags are large $n_x = n_y = 15$. In this case study, the use of more complicated NARX functions, *e.g.* with absolute terms or higher polynomial degrees, unfortunately did not lead to good results. This means that the non-linear behavior of this system should be represented by a different type of NARX functions, which is so far not known, thus requiring further investigations. It is worth noting that linear ARX models with large time lags were also used by Spiridonakos and Chatzi (2012);

Spiridonakos et al. (2015) to surrogate non-linear structures. This practice is, however, only acceptable when the considered structure mainly exhibits slight non-linearity. In the current numerical application, the steel frame responses remain mainly in the linear and slightly nonlinear range with the inter-story drifts hardly exceeding 1.5%.

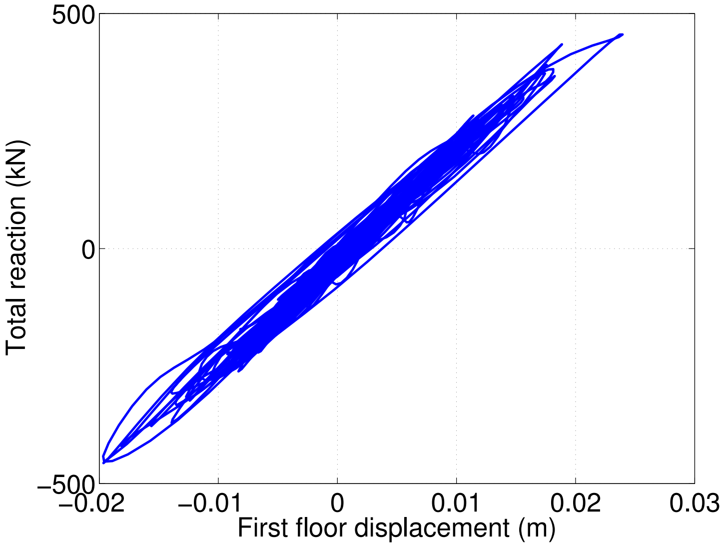
We selected 15 simulations with maximum displacement exceeding a large threshold, *i.e.* $\max(|y(t)|) > 0.015$ m. Candidate ARX models were obtained by LARS from each experiment. The ARX coefficients are computed using OLS, then used to reconstruct the responses and compute the relative error ϵ_k. The best candidate contains 18 terms namely: constant term, $x(t)$, $x(t-1)$, ..., $x(t-13)$, $y(t-1)$, $y(t-5)$, $y(t-6)$, $y(t-15)$, which leads to the mean relative error $\bar{\epsilon} = 4.61 \times 10^{-2}$. Figure 5.26 depicts the excitation and corresponding response from the simulation which led to the selected ARX model. Different levels of nonlinearity that the frame structure exhibits during the earthquake excitations are illustrated in Figure 5.27. We represented the ARX coefficients by adaptive OLS-based PCEs with optimal total degree p equal to 2.

5.3.6.3 Prediction of new trajectories and statistical moments

The PC-ARX model is then validated by MCS (10^4 independent simulations of the finite element model). Two specific validations are shown in Figure 5.28, where the responses computed by FEM are compared with those predicted by the PC-ARX model. Over 10^4 validations, 1049 simulations have the error $\epsilon_{val,i}$ exceeding 0.1 and the mean error is $\bar{\epsilon}_{val} = 4.78 \times 10^{-2}$. The time-dependent standard deviation of the response computed by FEM is plotted versus the trajectory predicted by the PC-ARX model in Figure 5.29 with the relative error $\epsilon_{val,std} = 10.19 \times 10^{-2}$. The maximal displacement is depicted in Figure 5.30, in which discrepancies between FEM-based and PC-ARX-based values are observed in the range of large values. In particular, the relative error of the maximal displacement prediction is $\epsilon_{val,max} = 3.8 \times 10^{-2}$. The figure also shows that the PDF of the maximal displacement is predicted quite accurately with the PC-ARX model.

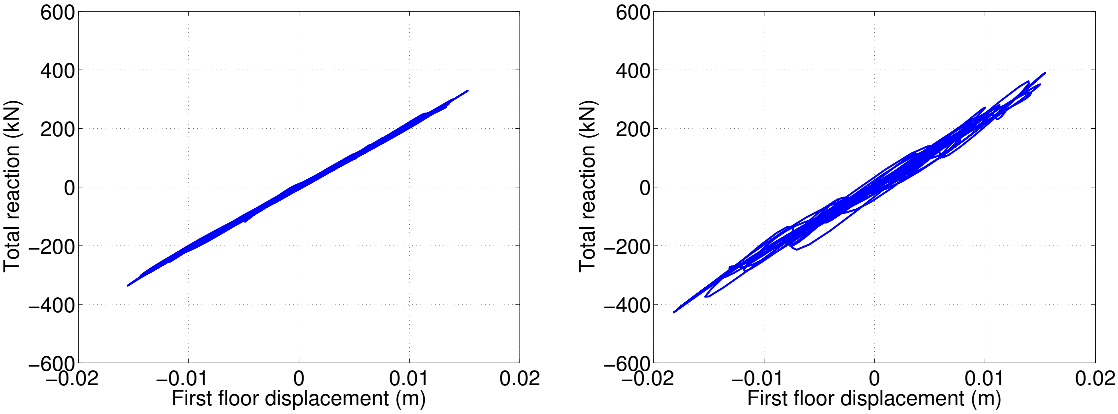


(a) Seismic ground acceleration

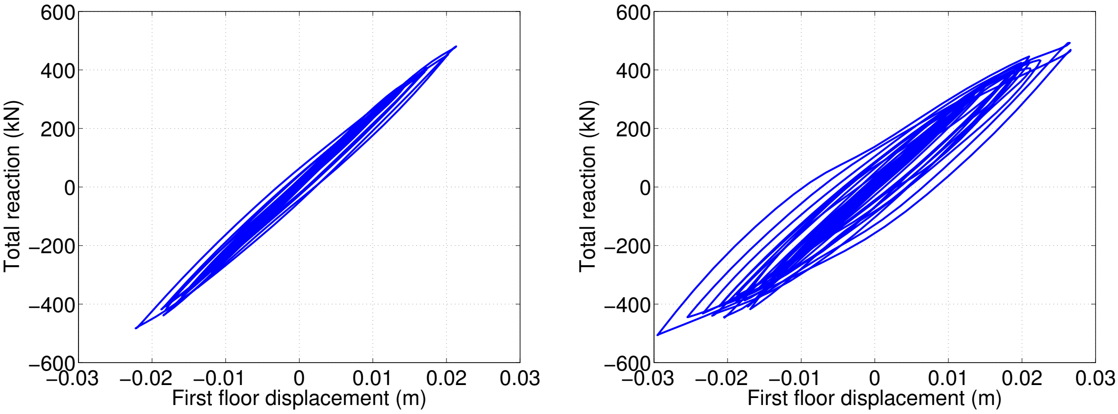


(b) Nonlinear behavior

Figure 5.26 – Steel frame – Experiment leading to the selected ARX model.

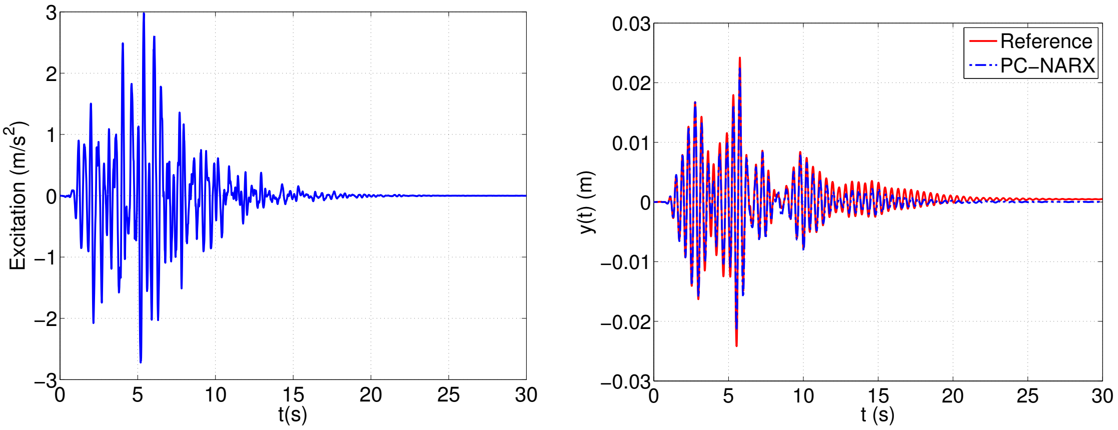


(a) Experiment 1 (b) Experiment 2

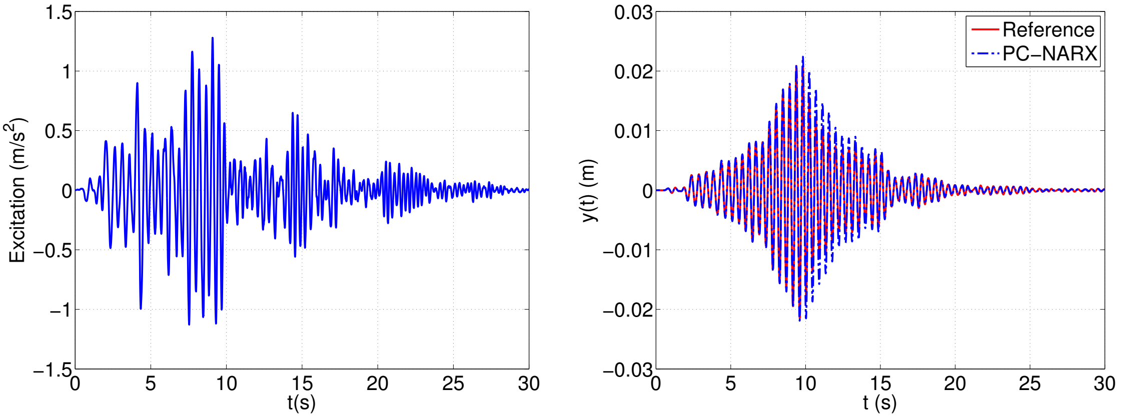


(c) Experiment 3 (d) Experiment 4

Figure 5.27 – Steel frame – Force-displacement responses for the first story illustrating different levels of nonlinearity.



First example trajectory $\boldsymbol{\xi} = (286.62, 205901.1, 0.09, 7.43, 4.63, 17.95, 0.25, 0.15)$



Second example trajectory $\boldsymbol{\xi} = (271.1, 210116.7, 0.024, 14.14, 6.99, 14.2285, 0.62, 0.054)$

Figure 5.28 – Steel frame – Two particular excitations (left column), associated response trajectories and their prediction by means of PC-ARX (right column).

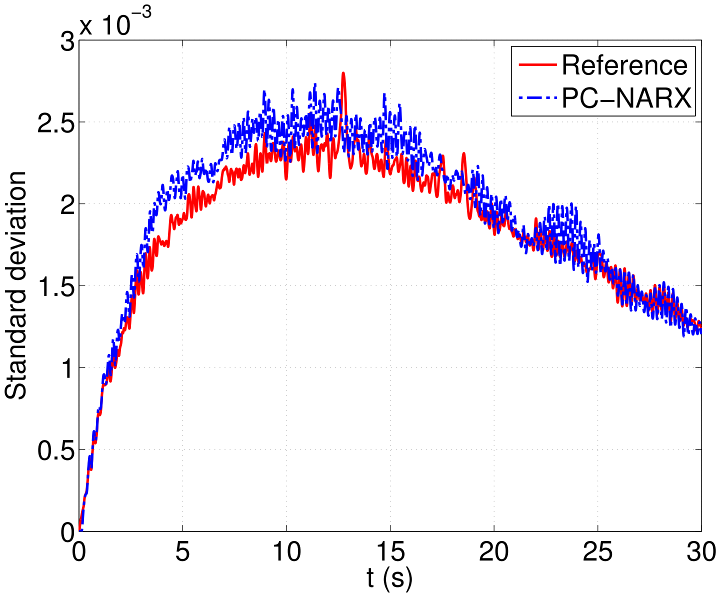
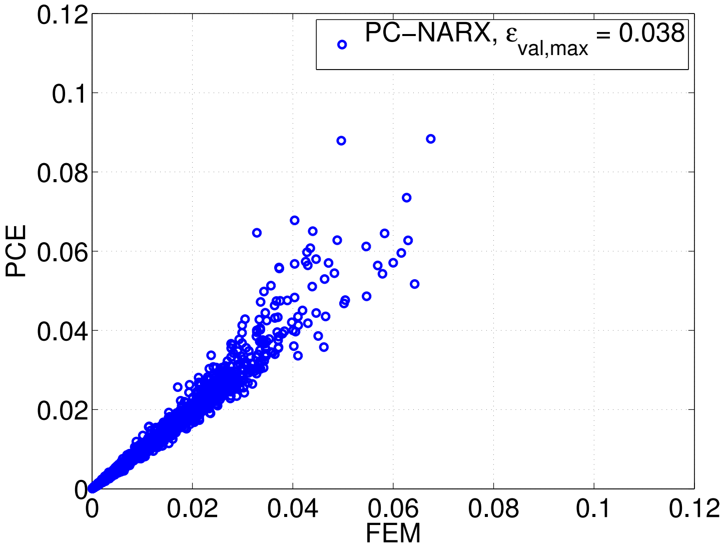
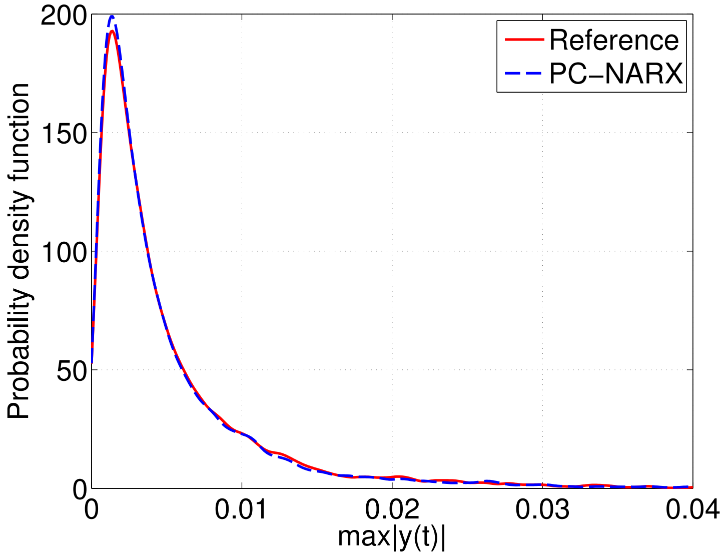


Figure 5.29 – Steel frame – Standard deviation of the first floor displacement trajectory.



(a) PCE-based predictions *vs.* actual values



(b) Probability density function

Figure 5.30 – Steel frame – Maximal displacements.

5.4 Discussion

In this section, we discuss certain features of the PC-NARX approach. It is clear that PC-NARX outperforms traditional time-frozen PCEs. The latter are found to fail in representing the random response of dynamical systems. The reason is that PCEs are designed for propagating uncertainties, but not for capturing the evolutionary behaviour in time domain. In order to use PCEs in this context, one should utilize a specific tool to tackle the dynamical aspect of the problem. For instance, one can use the available system of equations describing the model in an intrusive way. This approach is, however, also known to deteriorate in long term integration due to the accumulation of error in time (Wan and Karniadakis, 2006; Ghosh and Iaccarino, 2007). One also apprehends the remarkable performance of PC-NARX approach, which handles separately the two aspects of the problem, namely dynamics and uncertainties with the specific tools of NARX modelling and PCEs, respectively.

PC-NARX can also be considered to belong to a class of approaches where a pre-processing of the response time histories is conducted before applying PCEs. The response trajectories are first projected onto a different space, for instance phase-space (Desai et al., 2013), transformed time space (Le Maître et al., 2010) or space of NARX functions (Spiridonakos and Chatzi, 2015a). In the case of PC-NARX, the NARX functions form a "coordinate system" that moves along the trajectory. Due to the projection of the trajectories on the *suitably selected* set of basis functions, the projection coefficients become smooth functions of the uncertain parameters, therefore they can be represented effectively by low order PCEs. It is worth emphasizing that a linear projection of the trajectories on *deterministic* basis functions, namely principal components, wavelet, trigonometric functions will not help to overcome the problems that face time-frozen PCEs. Such a linear projection certainly allows one to reduce the computational cost associated with computing PCEs at each time instant by representing only a few relevant projection coefficients with PCEs. However, the complexity of the problem remains the same after this step of data compression.

The PC-NARX approach also has its own limitations. It uses data observed in discrete time. When the recorded data is sparse in the time domain, *i.e.* the time step is too large, the mechanism that relates the current state with the previous state and excitation cannot be revealed. In addition, the recursive form of the PC-NARX formulation renders it more difficult for post-processing. For instance, closed-form solutions for time-dependent statistics and sensitivity analysis are currently not available. Most importantly, the problem of specifying and selecting appropriate NARX models for a non-linear structure is so far a major challenge in system identification. For the considered numerical applications, we could specify appropriate nonlinear functions, *e.g.* cubic polynomials and absolute functions, for the candidate NARX models given the knowledge of the nonlinear behaviour of the mechanical systems. However, this might not be so straightforward for many other applications. Nonetheless, the use of versatile polynomial chaos functions for NARX terms is potentially a remedy for the issue of specifying appropriate NARX models. Luchtenburg et al. (2014) have recently proved the effectiveness of using of polynomial chaos functions for constructing flow maps in computational fluid dynamics. Regarding the selection of relevant NARX terms, for the considered numerical applications, the LARS-based technique performs comparably well in comparison with the genetic algorithm formerly used by Spiridonakos and Chatzi (2015a) in terms of the required computational effort and the accuracy of the resulting metamodels. It proves that correlation-based techniques can be used for selecting NARX models. However, given a complex engineering system, it is not always obvious that one can easily find the appropriate form of NARX functions and retain only the relevant terms. This requires a strong expertise of the investigator. The steel frame case study illustrates the difficulties that one might have to face.

In general, we are convinced that approaches which make use of interdisciplinary tools such as PC-NARX are promising. In other words, PCEs should be combined with tools designed for predicting non-linear time series in uncertainty quantification of stochastic dynamical systems. In particular, the use of generalized frequency response function for non-linear structures, which is the equivalent of NARX model in the frequency domain, deserves investigation in the future.

5.5 Summary

In this section the least angle regression (LARS) based polynomial chaos non-linear autoregressive with exogenous (PC-NARX) models were introduced as metamodels of non-linear systems with uncertainties subject to stochastic excitations. In particular, the approach relies on solving two linear regression problems, namely selecting the relevant NARX functions and the important PC basis. To this end, LARs proves effective in handling both regressors selection problems. The LARS-based PC-NARX approach is illustrated in predicting the response time histories of four benchmark engineering systems subject to non-stationary stochastic excitation. In the next chapter, PC-NARX surrogate models will be used for computing seismic fragility curves of structures.

6

Applications of surrogate modelling to earthquake engineering

In the previous chapter, the use of polynomial chaos-based surrogate models for representing time-history responses of structures subject to earthquake excitations was introduced. The objective of the current chapter is to explore the application of surrogate models in seismic risk assessment.

In modern engineering, seismic risk assessment is a key element not only in the design process of structures but also in the decision-making process regarding risk-mitigation strategies. This allows engineers and managers to determine the potential life safety and socio-economic losses due to damages of components, structures and infra-structures caused by earthquake scenarios. It is nowadays obligatory to conduct seismic risk assessment for facilities of major importance, for instance nuclear power plants and electric power systems. In the context of civil engineering, the assessment of seismic risk for structures is also an important issue which has attracted a large attention from researchers and engineers for several decades. This led to the development of the so-called performance-based earthquake engineering (PBEE) framework (Cornell and Krawinkler, 2000; Mackie and Stojadinovic, 2001), which is the most advanced and thorough approach for risk assessment in civil engineering. This framework is the aggregate of several elements namely seismic hazards analysis, structural analysis, damage and loss analysis which are gathered using the total probability theorem. In PBEE, a significant element is the computation of seismic fragility curves, which represent the likelihood that the considered components or structures fail to fulfil a prescribed earthquake-resistance criterion. Such curves are commonly obtained using data from structural analysis or post-earthquake survey with simplifying assumptions, for instance the curves are presented by lognormal cumulative distribution functions. This chapter aims at clarifying the effect of this assumption by using surrogate modelling.

The chapter is organized into 6 sections. Section 6.1 introduces the framework of performance-based earthquake engineering. Section 6.2 will be devoted to the commonly used parametric approaches for computing seismic fragility curves, which are based on simplifying assumptions. In Section 6.3, two non-parametric approaches are introduced to establish the fragility curves. The proposed methods are not utilizing assumptions on the shape of the curves. In Section 6.4, the assessment of epistemic uncertainties of fragility curves by means of the bootstrap resampling method is described. Section 6.5 presents three numerical case studies in which parametric and non-parametric curves are obtained with data from numerical simulations of the actual model or from predictions by the surrogate model. Based on the obtained results, discussions on the validity of parametric curves and the applicability of non-parametric approaches are given in Section 6.6. Finally, Section 6.7 concludes the chapter with a summary. Note that the review of the classical lognormal fragility curves and the theory of non-parametric curves were partially presented by Mai et al. (2014) and constitute a report recently submitted for publication (Mai et al., 2016a).

6.1 Seismic risk assessment

6.1.1 Performance-based earthquake engineering

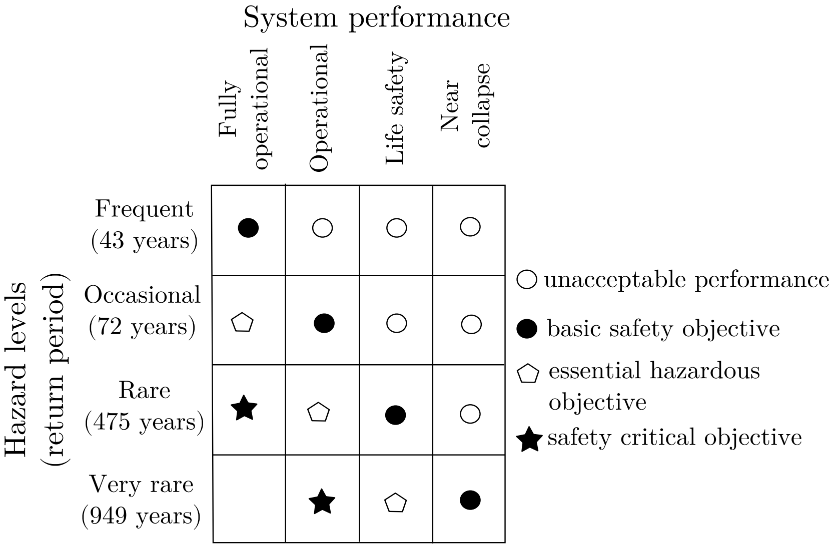
The development of seismic design codes is closely related to the history of major earthquakes that occurred worldwide. In the early 1990's, after the 1908 Messina earthquake (Italy) and the 1923 Great Kanto earthquake (Tokyo, Japan), guidelines were introduced for civil engineers to design buildings taking into account the effects of horizontal excitations. The 1925 Santa Barbara earthquake (California, USA) led to the first seismic regulations as a voluntary appendix in the 1927 Uniform Building Code. Since then, seismic deign codes are under constant evolution. Every time a major earthquake occurs, lessons are learnt through post-earthquake surveys and analyses, thus the codes are updated accordingly.

The scope of the modern seismic design codes are clearly defined. For instance, Eurocode 8 requires that:"In the event of earthquakes, human lives are protected, damage is limited and structures important for civil protection remain operational". In other words, the current seismic design philosophy focuses on assuring that no damage occurs during minor earthquakes, limiting damages during moderate ground motions, preventing structures from collapse during strong earthquakes, thus assuring life safety and maintaining the operation of important structures.

From a technical point of view, the current codes are based on the so-called load and resistance factor design (LRFD), which aims at assuring the strength of structural elements such as columns, beams, slabs. This constitutes a limitation of the codes, which was revealed in the 1994 Northridge earthquake in California and in Kobe (Japan, 1995). After the events, it was observed that structures designed in accordance with the codes survived the earthquakes, however, the socio-economic consequences due to reparation and loss of services were excessive (Lee and Mosalam, 2006). In particular, after recent events like the L'Aquila (Italy, 2009) and Chile (2010) earthquakes, those observations were once again confirmed. Some traditionally designed hospitals, which are extremely important for saving human lives after natural disasters, had to be evacuated after the earthquakes due to the non-structural damages (Holmes, 2010; Günay and Mosalam, 2013). In addition, many residents were not willing to continue living in their houses after the earthquake due to the visible damages, although the seismic performance of those buildings was satisfactory according to the available codes (Holmes, 2010). In summary, current codes exhibit shortcomings when focusing only on the strength at the element level.

Therefore, it is of utmost importance to design structures with emphasis on the global performance in terms of casualty prevention, economy and post-earthquake operability. In other words, this highlights the need for proper performance-based earthquake engineering (PBEE) as a basis for robust structural design, efficient decision making on mitigation actions and disaster planning.

For that reason, the first-generation PBEE framework was investigated with the Vision 2000 report (SEAOC Vision 2000 Committee, 1995) being one of the earliest documents. In this report, the system's performance is classified into four levels, namely fully operational, operational, life safety and near collapse. The performance levels are defined by the element force and deformation acceptability criteria, which are deterministic. The hazard levels are classified as frequent, occasional, rare and very rare events with return periods equal to 43, 72, 475 and 949 years respectively. Based on the public resilience requirements or the private property owners expectations, the intended performance level of the structure corresponding to a hazard level is determined (see Figure 6.1) and used as the design criterion. The performance-based design process is depicted in Figure 6.2. Following Vision 2000, subsequent documents namely ATC-40 (Applied Technology Council, 1996), FEMA-273 (Building Seismic Safety Council, 1997) and FEMA-356 (American Society of Civil Engineers, 2000) introduced similar frameworks in which the descriptions of performance levels and seismic hazard levels are slightly different. As one can see, the first generation PBEE approaches do not take into account uncertainties in the evaluation of performances. Moreover, the methods provide only criteria at element level, *e.g.* displacement, deformation, which are rarely meaningful for the stakeholders. These are the shortcomings of the first generation PBEE framework.



System performance

Hazard levels (return period)

	Fully operational	Operational	Life safety	Near collapse
Frequent (43 years)	●	○	○	○
Occasional (72 years)	⬠	●	○	○
Rare (475 years)	★	⬠	●	○
Very rare (949 years)		★	⬠	●

○ unacceptable performance
● basic safety objective
⬠ essential hazardous objective
★ safety critical objective

Figure 6.1 – Recommended seismic performance objectives for building (after SEAOC Vision 2000 Committee (1995))

To overcome the limitations of the first generation approaches, the Pacific Earthquake Engineering Research Center (PEER) has been developing a *probabilistic* PBEE framework (Cornell and Krawinkler, 2000), which allows explicit evaluation of perfor-

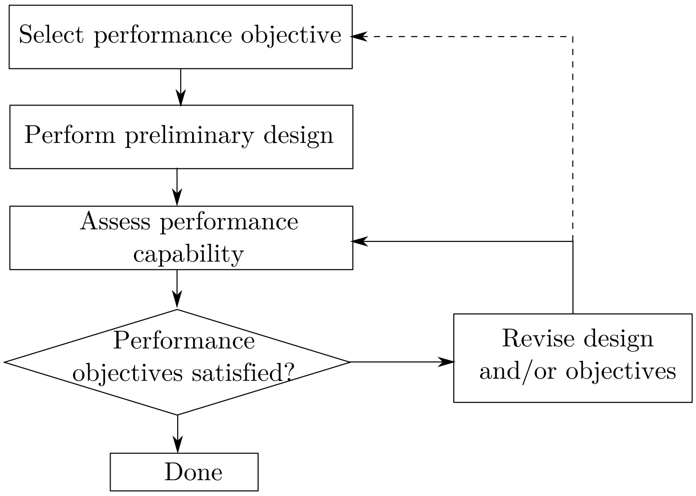


Figure 6.2 – First generation performance-based earthquake engineering design process (after Günay and Mosalam (2013))

mance measures that serve as decision variables (DV) (*e.g.* monetary losses, casualties, downtime) accounting for the prevailing uncertainties (*e.g.* arising from ground motion characteristics, structural properties, damage occurrence). The key steps in the PBEE framework comprise seismic hazard identification, structural response evaluation, damage analysis and eventually, consequence assessment. The four steps allow one to take into account the uncertainties, thus the outcome of each step is characterized by a probabilistic quantity. Details on the four steps of the framework are given below (Mackie and Stojadinovic, 2003; Günay and Mosalam, 2013):

- Hazard analysis: this aims at computing a seismic hazard curve representing the mean annual frequency of exceedance versus the ground motion intensity measure (IM). This analysis takes into account the uncertainties from the seismic activities, *e.g.* the fault location, magnitude-reoccurance rate, etc. At the end of the analysis, a set of ground motions compatible with the hazard curves is selected.

- Structural analysis: in this step, nonlinear time-history simulations of the structure subject to previously selected ground motions are conducted. This step aims at representing the structural responses, *a.k.a.* engineering demand parameters (EDP), *e.g.* displacement, interstory drift, as a probabilistic quantity conditioned on the motion IM. This allows one to take into consideration the uncertainties in the structure, *e.g.* material properties, geometry, damping ratio and also characteristics of the ground motions, *e.g.* temporal and frequency content.

- Damage analysis: in this step, one represents the damage measures (DM) as a probabilistic quantity conditioned on the different levels of EDP. The damage measures are commonly associated with the repair measures required to rehabilitate the structure. This step allows one to consider uncertainties from the pattern and history of the structural responses.

- Loss analysis: this step is done to convert the damage measures from the previous step to decision variables (DV), such as number of injuries and fatalities, monetary loss due to repair and downtime.

In the step of loss analysis, the mean annual frequency of exceedance of a DV is evaluated as (Mackie and Stojadinovic, 2005; Baker and Cornell, 2008):

$$\lambda(DV) = \int\int\int P(DV|DM)\,\mathrm{d}P(DM|EDP)\,\mathrm{d}P(EDP|IM)\;|\mathrm{d}\lambda(IM)|\,, \tag{6.1}$$

in which $P(x|y)$ is the conditional probability of x given y, DM is a damage measure typically defined according to repair measures (*e.g.* light, moderate or severe damage), EDP is an engineering demand parameter obtained from structural analysis (*e.g.* force, displacement, drift ratio), IM is an intensity measure characterizing the ground motion severity (*e.g.* peak ground acceleration, spectral acceleration) and $\lambda(IM)$ is the annual frequency of exceedance of the IM. Determination of the probabilistic model $P(EDP|IM)$ constitutes a major challenge in the PBEE framework. The present chapter is concerned with this step of the analysis.

It is worth noting that in the nuclear engineering, there is a framework named probabilistic seismic risk assessment (PSRA) which shares important features with PBEE in civil engineering. First used in the late 1970's for assessing the seismic risk of nuclear facilities, some technical details of the framework were then published by Kennedy et al. (1980). In 1981, the first complete PSRA study was conducted on a commercial nuclear power plant (Pickard Lowe and Garrick Inc., 1981). Nowadays PSRA still plays an important role in the nuclear industry. The PSRA framework provides the following outcomes which are meaningful for stakeholders, *e.g.* seismic fragilities of components, seismic accident sequence frequencies and the risk curves showing the annual frequency of exceeding certain damage levels, *e.g.* number of deaths, cancer fatalities, damage. The PSRA framework includes the following steps:

- Seismic hazard analysis which results in a family of seismic hazard curves relating the frequency of exceedance to different levels of ground motion intensity measure.

- Seismic fragility evaluation, which evaluates the vulnerability of each component under ground motions, *e.g.* the failure probabilities with respect to some prescribed safety criteria. The analysis leads to a family of *seismic fragility curves* for each component.

- System/accident sequence analysis, which models all the possible combinations of component failures, in other words, all the scenarios that may result in the global failure of the system. The analysis may be carried out by means of fault trees using logical structures.

- Risk quantification, which assembles seismic hazard analysis, seismic fragility analysis and system analysis to provide information such as the annual frequency of exceedance of a prescribed damage level (*e.g.* number of deaths, repair costs).

One can observe that PSRA does not differ significantly from PBEE, except that the former is used for assessing the seismic risk of nuclear facilities, therefore system/accident sequence analysis is an important step whereas it is not needed in the context of civil engineering. Both frameworks are based on the total probability theorem for estimating the seismic risk of structures. To conclude, the underlying principle of PBEE can be easily generalized and applied to different natural and non-natural disaster scenarios *e.g.* hurricane, flood, tsunami, snow avalanche, rock fall, landslide, fire and terrorist attacks. It is obvious that each disaster scenario has its own characteristics. However, the experiences learnt from the development of PBEE, in particular the probabilistic framework behind it, will be certainly relevant for the other frameworks.

6.1.2 Seismic fragility curves

The conditional probability $P(EDP \geq \overline{edp}|IM)$, where $\overline{edp}$ denotes an acceptable demand threshold, is commonly represented graphically in the shape of the so-called *demand fragility curves* (Mackie and Stojadinovic, 2005). Thus, a demand fragility curve represents the probability that an engineering demand parameter exceeds a prescribed threshold conditioned on the intensity measure of the ground motion. For the sake of simplicity, demand fragility curves are simply denoted fragility curves hereafter, which is also typical in the literature, see *e.g.* Ellingwood and Kinali (2009); Seo et al. (2012). We note however that the term *fragility* may also be used elsewhere to indicate $P(DM \geq \overline{dm}|IM)$ or $P(DM \geq \overline{dm}|EDP)$, *i.e.* the conditional probability of the damage measure exceeding a threshold $\overline{dm}$ given the ground motion intensity (Banerjee and Shinozuka, 2007) or the engineering demand parameter (Baker and Cornell, 2008; Günay and Mosalam, 2013), respectively.

Originally introduced in the early 1980's for nuclear safety evaluation (see *e.g.* Kennedy et al. (1980); Richardson et al. (1980); Kennedy and Ravindra (1984)), fragility curves are nowadays widely used for multiple purposes, *e.g.* loss estimation (Pei and Van De Lindt, 2009), assessment of collapse risk (Eads et al., 2013), design checking (Dukes et al., 2012), evaluation of the effectiveness of retrofit measures (Güneyisi and Altay, 2008), etc. Several novel methodological contributions to fragility analysis have been made in recent years, including the development of multi-variate fragility functions (Seyedi et al., 2010), the incorporation of Bayesian updating (Gardoni et al., 2002) and the consideration of time-dependent fragility (Ghosh and Padgett, 2010). In particular, in the context of seismic risk assessment, fragility curves are a popular tool with relevant applications on various type of structures, such as irregular buildings (Seo et al., 2012), underground tunnels (Argyroudis and Pitilakis, 2012), pile-supported wharfs (Chiou et al., 2011), wind turbines (Quilligan et al., 2012), nuclear power plant equipments (Borgonovo et al., 2013). Estimating fragility curves constitutes a major cornerstone in the PBEE framework.

Fragility curves are typically classified into four categories according to the data sources, namely judgment-based, empirical, analytical and hybrid fragility curves (Rossetto and Elnashai, 2003). Judgment-based curves are estimated by expert panels specialized in the field of earthquake engineering, thus they are strongly dependent on the personal opinions of the expert consulted. Empirical fragility curves are based on the observation of earthquake-induced damage reported in post-earthquake surveys. Due to the rarity of strong ground motions, the damage data tends to be scarce in the range of large IM and clustered in the range of low-damage low-ground motion severity. To circumvent the problem related to limited data set, one can use analytical fragility curves which are derived from data obtained by numerical analyses of structural models. The structure is subject to a series of ground motions, which is chosen in accordance with the seismic hazard of the construction site, then the damages are estimated based on the recorded structural responses. Finally, by combining data from different sources, one obtains hybrid curves. For instance analytical curves can be adjusted using damage data from real events. In this chapter, analytical fragility curves based on data collected from numerical structural analyses are of interest.

The typical approach to compute analytical fragility curves assumes that the curves have the shape of a lognormal cumulative distribution function (Shinozuka et al., 2000; Ellingwood, 2001; Mackie and Stojadinovic, 2003). This approach is therefore considered *parametric*. The parameters of the lognormal distribution are determined either by maximum likelihood estimation (Shinozuka et al., 2000; Zentner, 2010; Seyedi et al., 2010) or by fitting a linear probabilistic seismic demand model in the log-scale (Ellingwood and Kinali, 2009; Gencturk et al., 2008; Jeong et al., 2012; Banerjee and Shinozuka, 2008). The assumption of lognormal fragility curves is almost unanimous in the literature due to the computational convenience as well as due to the ease of combining such curves with other elements of the PBEE framework. The computation of parametric curves will be described in details in the next section.

6.2 Parametric fragility curves

Fragility curves represent the probability of failure of a system, which is associated with a specified criterion, for a given intensity measure (IM) of the earthquake motion. The term “failure” herein indicates that the engineering demand parameter, denoted by Δ, attains or exceeds a prescribed demand limit δ_0. Thus, the fragility function is cast as follows:

$$\text{Frag}(IM;\, \delta_o) = \mathbb{P}[\Delta \geq \delta_o | IM], \tag{6.2}$$

in which $\text{Frag}(IM;\, \delta_o)$ denotes the fragility at the given IM for a threshold δ_o of Δ. In order to establish the fragility curves, a number N of transient finite element analyses of the structure under consideration are used to provide paired values $\{(IM_i, \Delta_i)\,, i = 1, \ldots, N\}$. The classical approach for establishing fragility curves consists in assuming a *lognormal shape* for the curves described in Eq. (6.2). Two techniques are typically used to estimate the parameters of the lognormal fragility curves, namely *maximum likelihood estimation* and *linear regression*. These techniques are presented in the following.

6.2.1 Maximum likelihood estimation-based approach

One assumes that the fragility curves can be written in the following general form:

$$\widehat{\mathrm{Frag}}(IM;\, \delta_o) = \Phi\left(\frac{\ln IM - \ln\alpha}{\beta}\right), \tag{6.3}$$

where $\Phi(\cdot)$ denotes the standard Gaussian cumulative distribution function (CDF), α is the "median" and β is the "log-standard deviation" of the lognormal curve. Shinozuka et al. (2000) proposed the use of maximum likelihood estimation to determine these parameters as follows. Given a threshold δ_0, for each simulation corresponding to IM_i, the event $\{\Delta \geq \delta_0\}$ is modelled by a Bernoulli variable Y_i with $\mathbb{P}\left(Y_i = 1\right) = \mathbb{P}\left(\Delta \geq \delta_0\right) = \widehat{\mathrm{Frag}}(IM_i;\, \delta_o)$ and $\mathbb{P}\left(Y_i = 0\right) = \mathbb{P}\left(\Delta < \delta_0\right) = 1 - \widehat{\mathrm{Frag}}(IM_i;\, \delta_o)$. Considering a set of $i = 1, \ldots, N$ analyses, the likelihood function reads:

$$\mathcal{L}\left(\alpha,\, \beta,\, \{IM_i,\, i = 1,\, \ldots,\, N\}\right) = \prod_{i=1}^{N} \left[\widehat{\mathrm{Frag}}(IM_i;\, \delta_o)\right]^{y_i} \left[1 - \widehat{\mathrm{Frag}}(IM_i;\, \delta_o)\right]^{1-y_i}, \tag{6.4}$$

where IM_i is the intensity measure of the i^{th} seismic motion and y_i represents a realization of the Bernoulli random variable Y. The latter takes the value 1 or 0 depending on whether the structure under the i^{th} ground motion exceeds the demand threshold δ_o or not. The parameters $(\alpha,\, \beta)$ are obtained by maximizing the likelihood function. In practice, a straightforward optimization algorithm is applied on the log-likelihood function:

$$\begin{aligned} \{\alpha^*;\, \beta^*\}^{\mathsf{T}} &= \arg\min_{(\alpha,\beta)} \left(-\ln\mathcal{L}\left(\alpha,\, \beta,\, \{IM_i,\, i = 1,\, \ldots,\, N\}\right)\right) \\ &= \arg\min_{(\alpha,\beta)} -\left(\sum_{i:\Delta_i \geq \delta_0} \ln\Phi\left(\frac{\ln IM_i - \ln\alpha}{\beta}\right) + \sum_{i:\Delta_i < \delta_0} \ln\left(1 - \Phi\left(\frac{\ln IM_i - \ln\alpha}{\beta}\right)\right)\right). \end{aligned} \tag{6.5}$$

6.2.2 Linear regression-based approach

One first assumes a *probabilistic seismic demand model*, which relates a structural response quantity of interest to an intensity measure of the ground motion. Specifically, the demand Δ is assumed to follow a lognormal distribution of which the log-mean value is a *linear function* of $\ln IM$, leading to:

$$\ln\Delta = A\,\ln IM + B + \zeta\, Z, \tag{6.6}$$

where $Z \sim \mathcal{N}(0, 1)$ is a standard normal variable. Parameters A and B are determined by means of ordinary least squares estimation in a log-log scale. Parameter ζ is obtained by:

$$\zeta^2 = \sum_{i=1}^{N} e_i^2 / \left(N - 2\right), \tag{6.7}$$

where e_i the residual between the actual value $\ln \Delta$ and the value predicted by the linear model: $e_i = \ln \Delta_i - A \ln (IM_i) - B$. Then, Eq. (6.2) rewrites:

$$\begin{aligned} \widehat{\mathrm{Frag}}(IM;\, \delta_o) &= \mathbb{P}\left[\ln \Delta \geq \ln \delta_o\right] = 1 - \mathbb{P}\left[\ln \Delta \leq \ln \delta_o\right] \\ &= \Phi\left(\frac{\ln IM - \left(\ln \delta_o - B\right)/A}{\zeta / A}\right). \end{aligned} \tag{6.8}$$

The median and log-standard deviation of the lognormal fragility curve in Eq. (6.8) are $\alpha = \exp\left(\left(\ln \delta_o - B\right)/A\right)$ and $\beta = \zeta / A$ respectively. This approach to fragility estimation is widely employed in the literature, see *e.g.* Ellingwood (2001); Mackie and Stojadinovic (2001); Choi et al. (2004); Padgett et al. (2008); Zareian and Krawinkler (2007) among others.

The two methods described in this section are *parametric* because they impose the shape of the fragility curves (Eq. (6.3) and Eq. (6.8)), which is that of a lognormal CDF when considered as a function of IM. We note that by using the linear-regression approach, one accepts two additional assumptions, namely the linear function for the log-mean value of Δ and the constant dispersion (or homoscedasticity) of the residuals independently of the IM level. Effects of these assumptions have been investigated by Karamlou and Bocchini (2015). In the sequel, we propose two *non-parametric* approaches to compute fragility curves *without* relying on the lognormality assumption.

6.3 Non-parametric fragility curves

6.3.1 Binned Monte Carlo simulation-based approach

Having at hand a large sample set $\{(IM_j, \Delta_j)\,,\, j = 1, \ldots, N\}$, it is possible to use *binned Monte Carlo simulation* (bMCS) to compute the fragility curves, as described next. Let us consider a given abscissa IM_o. Within a small bin surrounding IM_o, say $[IM_o - h, IM_o + h]$ one assumes that the structural response Δ is linearly related to the IM. This assumption is exact in the case of linear structures, but would only be an approximation in the nonlinear case. Therefore, the maximal drift Δ_j, which is related to $IM_j \in [IM_o - h, IM_o + h]$, is converted into the response $\widetilde{\Delta_j}(IM_o)$, which is related to the j^{th} input signal scaled to have an intensity measure equal to IM_o:

$$\widetilde{\Delta_j}(IM_o) = \Delta_j \frac{IM_o}{IM_j}. \tag{6.9}$$

This procedure is illustrated in Figure 6.3. The fragility curve at IM_o is then obtained by a crude Monte Carlo estimator:

$$\widehat{\mathrm{Frag}}(IM_o) = \frac{N_f\left(IM_o\right)}{N_s\left(IM_o\right)}, \tag{6.10}$$

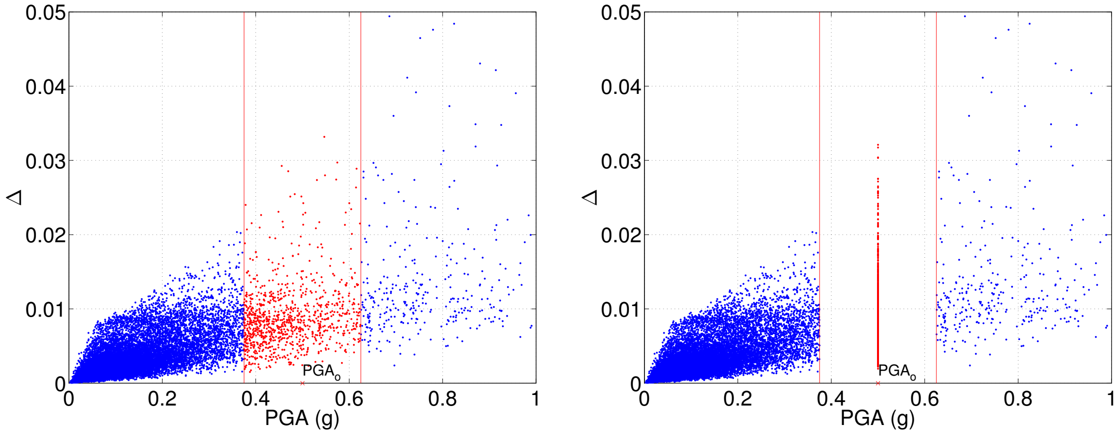


Figure 6.3 – Scaling of ground motions and corresponding responses in binned Monte Carlo simulation (the bin is enlarged to facilitate visualization).

where $N_f\,(IM_o)$ is the number of points in the bin such that $\widetilde{\Delta}_j(IM_o) \geqslant \delta_o$ and $N_s(IM_o)$ is the total number of points that fall into the bin $[IM_o - h, IM_o + h]$.

We note that the bMCS approach bears similarities to the stripe analysis introduced by Shome et al. (1998). However, when using stripe analysis, one scales *all* ground motions to the intensity level of interest. As a result, certain signals are scaled with factors that are considerably larger or smaller than unity, which may lead to gross approximations of the corresponding responses (Luco and Bazzurro, 2007; Cimellaro et al., 2009; Mehdizadeh et al., 2012). The reader is referred to Mehdizadeh et al. (2012) for some illustrations of the effects of the scale factor on the introduced bias, with the latter represented by the ratio of the median nonlinear response of the considered system subject to the scaled motions to the respective median response of the system subject to natural motions with all motions characterized by the same IM level. In general, the bias ratio tends to become larger with the scale factor significantly differs from unity. On the other hand, the scaling in binned MCS is confined in the vicinity of the intensity level IM_o, where the vicinity is defined by the bin width $2h$ chosen so that the scale factors are close to unity. Accordingly, the bias due to ground motion scaling is negligible in bMCS.

Following the above discussion, it should be noted that bias from scaling can be avoided by a proper selection of ground motions. For instance, Shome et al. (1998) showed that the scaling of motions that correspond to a narrow interval of earthquake magnitudes and source-to-site distances does not introduce bias into the nonlinear response estimates. Furthermore, Luco and Bazzurro (2007) concluded that the bias can be reduced by selecting records that have appropriate response spectrum shapes. According to Bazzurro et al. (1998) and Vamvatsikos and Cornell (2002), the existence of scale-induced bias also depends on several other factors, such as the structural characteristics and the considered intensity and damage measures. The topic of ground motion scaling is complex and falls outside the scope of this chapter. It should be underlined that by using the bMCS approach, one avoids introducing bias in the results independently of the ground motion characteristics or other factors. In the following case studies, the resulting fragility curves serve as reference for assessing the accuracy of the various considered techniques for fragility estimation.

6.3.2 Kernel density estimation-based approach

The fragility function defined in Eq. (6.2) may be reformulated using the conditional probability density function (PDF) $f_{\Delta|IM}$ as follows:

$$\mathrm{Frag}(a;\, \delta_o) = \mathbb{P}\left(\Delta \geq \delta_o | IM = a\right) = \int_{\delta_o}^{+\infty} f_{\Delta}(\delta | IM = a)\, \mathrm{d}\delta. \tag{6.11}$$

By definition, this conditional PDF is given as:

$$f_{\Delta}(\delta | IM = a) = \frac{f_{\Delta,IM}(\delta, a)}{f_{IM}(a)}, \tag{6.12}$$

where $f_{\Delta,IM}(\cdot)$ is the joint distribution of the vector $(\Delta,\, IM)$ and $f_{IM}(\cdot)$ is the marginal distribution of the IM. If these quantities were known, the fragility function in Eq. (6.11) would be obtained by a mere integration. In this section, we propose to determine the joint and marginal PDFs from a sample set $\{(IM_i, \Delta_i)\,,\ i = 1,\, \dots\,, N\}$ by means of *kernel density estimation* (KDE).

For a single random variable X for which a sample set $\{x_1,\, \dots\,, x_N\}$ is available, the kernel density estimate of the PDF reads (Wand and Jones, 1995):

$$\hat{f}_X\left(x\right) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right), \tag{6.13}$$

where h is the *bandwidth* parameter and $K(\cdot)$ is the *kernel* function which integrates to one. Classical kernel functions are the Epanechnikov, uniform, normal and triangular functions. Wand and Jones (1995) suggested that the choice of the kernel does not affect strongly the quality of the estimate provided the sample set is large enough. In case a standard normal PDF is adopted for the kernel, the kernel density estimate becomes:

$$\hat{f}_X\left(x\right) = \frac{1}{Nh} \sum_{i=1}^{N} \frac{1}{\left(2\pi\right)^{1/2}} \exp\left(-\frac{1}{2}\left(\frac{x - x_i}{h}\right)^2\right). \tag{6.14}$$

On the other hand, the choice of the bandwidth h is crucial since an inappropriate value of h can lead to an oversmoothed or undersmoothed PDF estimate (Duong, 2004).

Kernel density estimation may be extended to a random vector $\boldsymbol{X} \in \mathbb{R}^d$ given an i.i.d sample $\{\boldsymbol{x}_1,\, \dots\,, \boldsymbol{x}_N\}$ (Wand and Jones, 1995):

$$\hat{f}_{\boldsymbol{X}}\left(\boldsymbol{x}\right) = \frac{1}{N\left|\boldsymbol{H}\right|^{1/2}} \sum_{i=1}^{N} K\left(\boldsymbol{H}^{-1/2}(\boldsymbol{x} - \boldsymbol{x}_i)\right), \tag{6.15}$$

where $\boldsymbol{H}$ is a symmetric positive definite *bandwidth matrix* with determinant denoted by $|\boldsymbol{H}|$. When a multivariate standard normal kernel is adopted, the joint distribution estimate becomes:

$$\hat{f}_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{N\,|\boldsymbol{H}|^{1/2}} \sum_{i=1}^{N} \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_i)^{\mathsf{T}}\boldsymbol{H}^{-1}(\boldsymbol{x}-\boldsymbol{x}_i)\right), \tag{6.16}$$

where $(\cdot)^{\mathsf{T}}$ denotes the transposition. For multivariate problems (*i.e.* $\boldsymbol{X} \in \mathbb{R}^d$), the bandwidth matrix typically belongs to one of the following classes: spherical, ellipsoidal and full matrix, which respectively contain 1, d and $d(d+1)/2$ independent unknown parameters. The matrix $\boldsymbol{H}$ can be computed by means of plug-in or cross-validation estimators. Both estimators aim at minimizing the asymptotic mean integrated squared error (MISE):

$$\mathrm{MISE} = \mathbb{E}\left[\int_{\mathbb{R}^d} \left[\hat{f}_{\boldsymbol{X}}(\boldsymbol{x};\,\boldsymbol{H}) - f_{\boldsymbol{X}}(\boldsymbol{x})\right]^2 \mathrm{d}\boldsymbol{x}\right]. \tag{6.17}$$

However, the two approaches differ in the formulation of the numerical approximation of MISE. For further details, the reader is referred to Duong (2004). In the most general case, when the correlations between the random variables are not known, the full matrix should be used.

Eq. (6.14) is used to estimate the marginal PDF of the IM, namely $\hat{f}_{IM}(a)$, from a sample $\{IM_i,\ i = 1, \ldots, N\}$:

$$\hat{f}_{IM}(a) = \frac{1}{(2\pi)^{1/2}\,N h_{IM}} \sum_{i=1}^{N} \exp\left(-\frac{1}{2}\left(\frac{a - IM_i}{h_{IM}}\right)^2\right). \tag{6.18}$$

Eq. (6.16) is used to estimate the joint PDF $\hat{f}_{\Delta,IM}(\delta, a)$ from the data pairs $\{(IM_i,\ \Delta_i),\ i = 1, \ldots, N\}$:

$$\hat{f}_{\Delta,IM}(\delta, a) = \frac{1}{2\pi N\,|\boldsymbol{H}|^{1/2}} \sum_{i=1}^{N} \exp\left(-\frac{1}{2}\begin{pmatrix}\delta-\Delta_i\\ a-IM_i\end{pmatrix}^{\mathsf{T}} \boldsymbol{H}^{-1} \begin{pmatrix}\delta-\Delta_i\\ a-IM_i\end{pmatrix}\right). \tag{6.19}$$

The conditional PDF $f_\Delta(\delta|IM = a)$ is eventually estimated by plugging the estimations of the numerator and denominator in Eq. (6.12). The proposed estimator of the fragility function eventually reads:

$$\widehat{\mathrm{Frag}}(a;\,\delta_o) = \frac{h_{IM}}{(2\pi\,|\boldsymbol{H}|)^{1/2}} \frac{\int\limits_{\delta_o}^{+\infty} \sum\limits_{i=1}^{N} \exp\left(-\frac{1}{2}\begin{pmatrix}\delta-\Delta_i\\ a-IM_i\end{pmatrix}^{\mathsf{T}} \boldsymbol{H}^{-1} \begin{pmatrix}\delta-\Delta_i\\ a-IM_i\end{pmatrix}\right) \mathrm{d}\delta}{\sum\limits_{i=1}^{N} \exp\left(-\frac{1}{2}\left(\frac{a-IM_i}{h_{IM}}\right)^2\right)}. \tag{6.20}$$

The choice of the bandwidth parameter h_{IM} and the bandwidth matrix $\boldsymbol{H}$ plays a crucial role in the estimation of fragility curves, as seen in Eq. (6.20). In the above formulation, the same bandwidth is considered for the whole range of the IM values.

However, there are typically few observations available corresponding to the upper tail of the distribution of the IM. This is due to the fact that the annual frequency of seismic motions with IM values in the respective range (*e.g.* PGA exceeding $1g$) is low (see *e.g.* Frankel et al. (2000)). This is also the case when synthetic ground motions are used, since these are generated consistently with statistical features of recorded motions. Preliminary investigations have shown that by applying the KDE method on the data in the original scale, the fragility curves for the higher demand thresholds tend to be unstable in their upper tails (Sudret and Mai, 2013b). To reduce effects from the scarcity of observations at large IM values, we propose the use of KDE in the logarithmic scale, as described next.

Let us consider two random variables X, Y with positive supports, and their logarithmic transformations $U = \ln X$ and $V = \ln Y$. One has:

$$\int_{y_0}^{+\infty} f_Y(y|X=x)\,\mathrm{d}y = \int_{y_0}^{+\infty} \frac{f_{X,Y}(x,y)}{f_X(x)}\,\mathrm{d}y = \int_{\ln y_0}^{+\infty} \frac{\dfrac{f_{U,V}(u,v)}{x\,y}}{\dfrac{f_U(u)}{x}}\,y\,\mathrm{d}v = \int_{\ln y_0}^{+\infty} f_V(v|U=u)\,\mathrm{d}v. \tag{6.21}$$

Accordingly, by substituting $X = IM$ and $Y = \Delta$, the fragility function in Eq. (6.11) can be obtained in terms of $U = \ln IM$ and $V = \ln \Delta$ as:

$$\widehat{\mathrm{Frag}}(a;\,\delta_o) = \int_{\delta_o}^{+\infty} \hat{f}_\Delta(\delta|IM=a)\,\mathrm{d}\delta = \int_{\ln \delta_o}^{+\infty} \hat{f}_V(v|U=\ln a)\,\mathrm{d}v. \tag{6.22}$$

The use of a constant bandwidth in the logarithmic scale is equivalent to the use of a varying bandwidth in the original scale, with larger bandwidths corresponding to larger values of IM. The resulting fragility curves are smoother than those obtained by applying KDE with the data in the original scale.

6.4 Epistemic uncertainty of fragility curves

In the context of reliability and risk analysis, uncertainties from multiple sources can be categorized into two classes, namely aleatory and epistemic uncertainties (Der Kiureghian and Ditlevsen, 2009). The first category corresponds to the uncertainties that can not be reduced. In contrast, by gathering more data, one can reduce the epistemic uncertainties. Der Kiureghian and Ditlevsen (2009) showed that neglect or improper characterization of uncertainties could lead to inaccurate reliability assessment, for instance the failure probability of the system might be underestimated by orders of magnitude.

Therefore, it is of major importance in fragility analysis to investigate the effects of both aleatory and epistemic uncertainties. In terms of aleatory randomness, the inherent variabilities of material properties and earthquake excitations must be properly accounted for and propagated to structural responses. This has been intensively investigated in the previous chapters. Hereafter, we are interested in quantifying the variability in the estimated fragility curves which is due to epistemic uncertainty associated with a limited number of finite element runs. A fragility curve is commonly computed based on a limited amount of data, *i.e.* few structural analyses. Large epistemic uncertainties in fragility curves may affect significantly the total variability of the seismic risk assessment outcomes. Characterizing and propagating epistemic uncertainties in seismic loss estimation has also attracted attention from several researchers, see *e.g.* Baker and Cornell (2008); Bradley and Lee (2010); Liel et al. (2009).

The theoretical approach to determine the variability of an estimator relies on repeating the estimation with an ensemble of different random samples. However, this approach is not feasible in earthquake engineering because of the high computational cost which is required. In this context, the *bootstrap resampling* technique originally introduced in statistics by Efron (1979) is deemed an appropriate tool for estimating the effects of epistemic uncertainties (Baker and Cornell, 2008). Given a set of observations $\mathcal{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ of $\boldsymbol{X}$ following an unknown probability distribution, the bootstrap method allows estimation of the statistics of a random variable that depends on $\boldsymbol{X}$ in terms of the observed data $\mathcal{X}$ and their empirical distribution.

To estimate statistics of the fragility curves with the bootstrap method, we first draw M independent random samples *with replacement* from the original data set $\left\{\left(IM_i^{(b)}, \Delta_i^{(b)}\right), i = 1, \ldots, N, \, b = 1, \ldots, M\right\}$. These represent the so-called bootstrap samples. Each bootstrap sample has the same size N as the original sample, but the observations are different: in a particular sample, some of the original observations may appear multiple times while others may be missing. Next, we compute the fragility curves for each bootstrap sample using the approaches previously presented. Finally, we perform statistical analysis of the so-obtained M bootstrap curves. In the subsequent case studies, the above procedure is employed to evaluate the median and 95% confidence intervals of the estimated fragility curves and also, to assess the variability of the IM value corresponding to a 50% probability of failure. To clarify the effects of gathering more data on reducing epistemic uncertainties, two data sets comprising respectively a small and a large number of simulations (respectively 200 and $10,000$) will be used. Note that the final fragility curves will exhibit aleatory uncertainties from the structure and earthquake ground motions as well as epistemic uncertainties due to the limited information. It is also worth emphasizing that epistemic uncertainties can be of different nature. Herein we investigate only the statistical uncertainty of the estimated fragility curves due to the fact that one uses 100 or 200 simulations instead of $10,000$. The epistemic uncertainties from other sources, *e.g.* the computational model and the selection of the form of probabilistic models describing the distribution of input parameters, are not taken into account.

6.5 Applications of surrogate modelling to computation of fragility curves

In this section, three numerical case studies, namely a Duffing, a Bouc-Wen oscillator and a steel frame, are investigated. It is worth noting that use of single-degree-of-freedom (SDOF) system is common in earthquake engineering. For instance, the roof displacement of a multistory building is determined from the earthquake-induced deformation of an inelastic SDOF system derived from the push-over curve (Building Seismic Safety Council, 1997; Applied Technology Council, 1996), which has been further investigated by Gupta and Krawinkler (2000); Chopra et al. (2003). In the context of PBEE, Christopoulos et al. (2003) considered a nonlinear SDOF to clarify the factors with largest influences on the residual deformations of structures. Recently, Kafali and Grigoriu (2007) computed seismic fragility curves of Duffing and Bouc-Wen oscillators, which are investigated in the current work.

For each case study, the following process is applied. First, one computes the fragility curves by means of the parametric lognormal approaches using data from a few simulations, typically 200 or 300 dynamic analyses of the system subject to synthetic ground motions which were used as an experimental design in the previous chapter. Such computation of fragility curves is commonly done in practice when only a few numerical analyses are conducted. Second, one computes the curves by means of parametric and non-parametric approaches using data from a large number of simulations $(10,000)$, which constitute the validation set used in the previous chapter. Third, one computes the curves using KDE with data being predictions of the PC-NARX metamodel that was built using an experimental design of size 200 or 300. Fourth, one estimates the uncertainties of the curves when using small and large numbers of simulations together with different approaches.

The obtained results allow one to validate the accuracy of the commonly used parametric approaches and to clarify the effect of epistemic uncertainties on the resulting curves. Furthermore, one can examine the validity of the proposed non-parametric methods and the applicability of surrogate modelling in this context.

6.5.1 Duffing oscillator

6.5.1.1 Computation of fragility curves

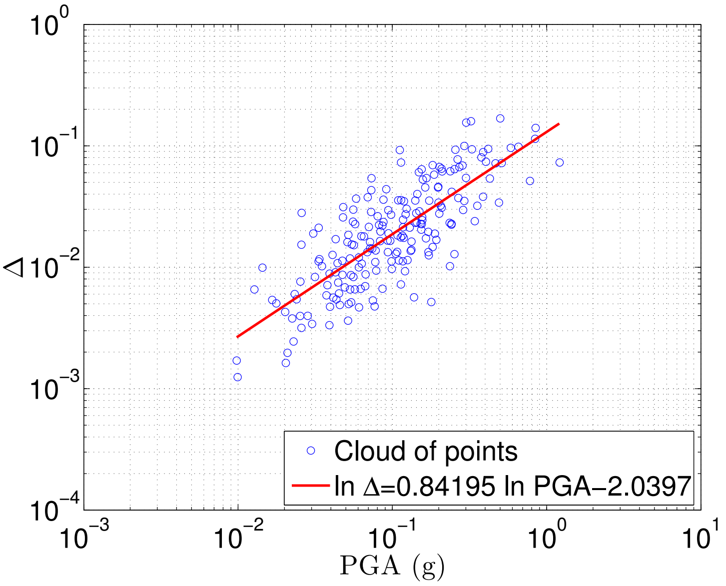
Let us consider the Duffing oscillator which was investigated in the previous chapter. Two limit states for the maximal oscillator's displacement are considered, namely $\delta_0 = 0.03$ m and $\delta_0 = 0.08$ m (Kafali and Grigoriu, 2007). First, one computes the fragility curves by means of the parametric approaches. As described in Section 6.2, the parametric approaches rely on assuming that the fragility curves have the shape of a lognormal CDF. Using the maximum likelihood estimation (MLE) method, the observed failures for each limit state are modeled as outcomes of a Bernoulli experiment and the parameters (α, β)

of the fragility curves are determined by maximizing the respective likelihood function. Using the linear regression (LR) technique, the parameters of the lognormal curves are derived by fitting a linear model to the paired data $(\ln IM, \ln \Delta)$. Figure 6.4(a) and Figure 6.4(b) depicts the paired data together with the fitted linear regression models when using respectively 200 and 10, 000 simulations. The coefficient of determination of the fitted linear model is $R^2 = 0.611$ and $R^2 = 0.665$, respectively. Next, one computes the reference curves by means of bMCS and KDE method using data from 10, 000 simulations. For bMCS method, the bandwidth h_{bMCS} is set equal to $0.25\, IM_o$. The resulting scale factors vary in the range [0.75, 1.25] corresponding to a bias ratio approximately equal to 1. The KDE approach requires estimation of the bandwidth parameter and the bandwidth matrix. Using the cross-validation estimation implemented in `R` (Duong, 2007), these are determined as $h = 0.175\, g$, $\boldsymbol{H} = \begin{bmatrix} 0.0495\, g & 0.0387\, g; & 0.0387\, g & 0.0436\, g \end{bmatrix}$. Finally, the curves are computed with 10, 000 responses Δ predicted by the PC-NARX model which was built with an experimental design of 200 simulations. The bandwidth parameter and matrix are $h = 0.175\, g$, $\boldsymbol{H} = \begin{bmatrix} 0.045\, g & 0.0343\, g; & 0.0343\, g & 0.0387\, g \end{bmatrix}$ respectively.

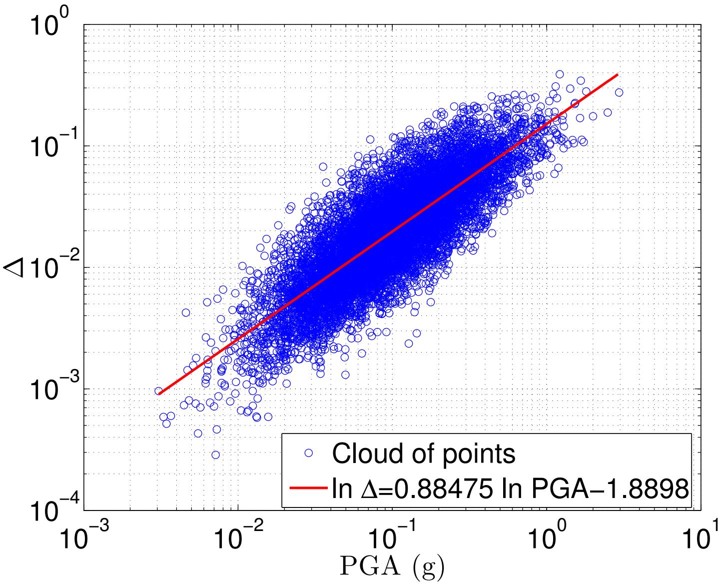
Figure 6.5 shows the lognormal curves obtained with the LR and MLE methods using 200 simulations. They are plotted against the bMCS- and KDE-based non-parametric curves using 10, 000 simulations. One first observes a remarkable consistency between the curves obtained with the two non-parametric approaches despite the distinct differences in the underlying algorithms. This validates the accuracy of the proposed non-parametric methods. In order to investigate the efficiency of the metamodel, the KDE-based curves obtained with 10, 000 predictions of the PC-NARX model are depicted on the same figure. The plot suggests that the PCE-based curves are in great agreement with the curves obtained with 10, 000 analyses of the original model. This is explained by the high accuracy of the metamodel's predictions. On the contrary, when using only 200 simulations without surrogate modelling, the parametric curves are in good agreement with each other but exhibit strong discrepancies in comparison to the non-parametric ones. The dissimilarities are particularly pronounced in the range of large IM ($PGA > 0.8\, g$) and higher failure threshold ($\delta_0 = 0.08$ m). Figure 6.6 indicates that when using the full set of 10, 000 simulations, the lognormal curves tend toward non-parametric ones. Use of large number of simulations helps to reduce drastically the discrepancies between parametric and non-parametric curves.

6.5.1.2 Estimation of epistemic uncertainty by bootstrap resampling

In the following, the bootstrap resampling technique (see Section 6.4) is used to investigate the epistemic uncertainty in the fragility curves estimated with the different approaches. We examine the variability in the estimation by computing bootstrap confidence intervals. Figure 6.7 depicts the median lognormal fragility curves and the 95% confidence intervals obtained by bootstrap resampling with 100 replications together with the respective estimated curves when 200 simulations are used. Large confidence intervals are observed in MLE-based curves, which indicates a high level of epistemic uncertainties. This also suggests that MLE-based curves are considerably sensitive with respect to the change of the *small* data set used for their computation. In other words, when a different small data set



(a) 200 simulations



(b) 10, 000 simulations

Figure 6.4 – Duffing oscillator – Paired data and fitted linear regression model in the logarithmic scale ($R^2 = 0.611$ (a) and $R^2 = 0.665$ (b)).

is used, the resulting MLE-based curves might change largely and might eventually differ from the LR-based ones. The LR-based curves exhibits epistemic uncertainties which are less severe. It is worth noting that significant epistemic uncertainties in estimated fragility curves were also reported by Liel et al. (2009); Zentner et al. (2011); Rajeev and Tesfamariam (2012). In both MLE and LR curves, the confidence interval increases with the demand limit δ_0 and the IM value.

The epistemic uncertainties of the curves can be reduced by using a sufficiently large

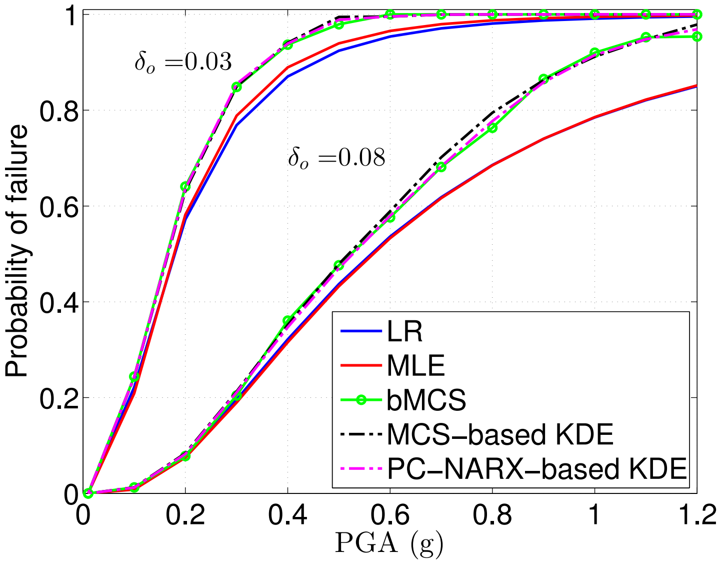


Figure 6.5 – Duffing oscillator – Fragility curves. Lognormal curves (LR, MLE) are computed with 200 simulations. bMCS and MCS-based KDE curves are obtained with $10,000$ simulations. PC-NARX-based KDE curves are calculated with $10,000$ predictions of the surrogate model fitted with 200 simulations.

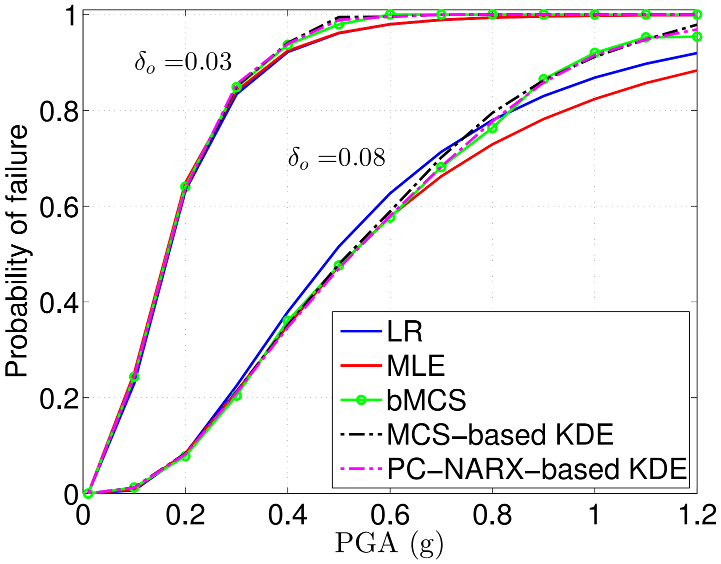
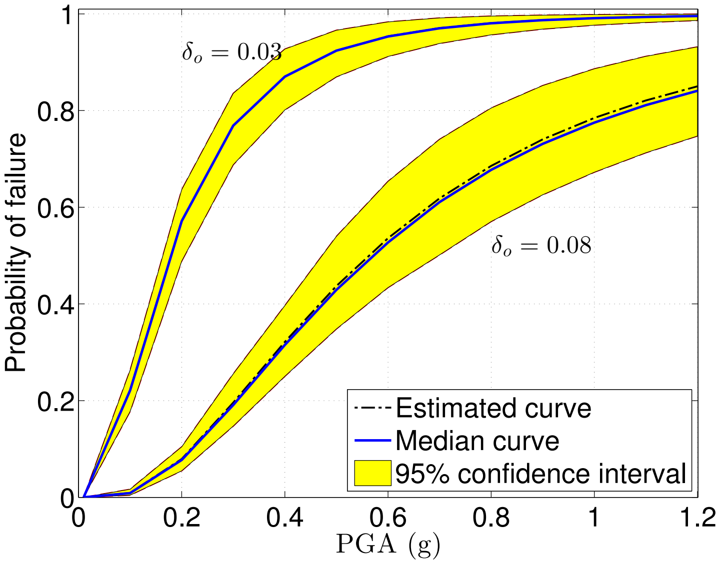
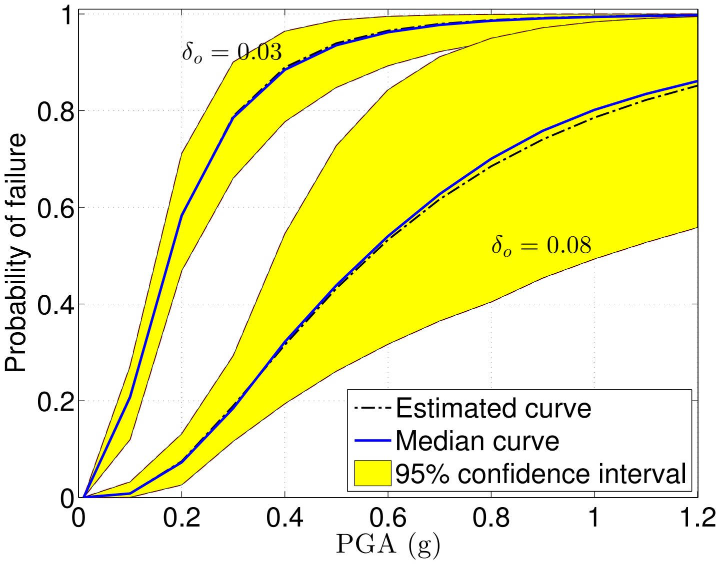


Figure 6.6 – Duffing oscillator – Fragility curves. Lognormal curves (LR, MLE), bMCS and MCS-based KDE curves are obtained with $10,000$ simulations. PC-NARX-based KDE curves are computed with $10,000$ predictions of the surrogate model fitted with an ED of size 200.

data set. Figure 6.8 depicts the median curves and the confidence intervals of LR, MLE, bMCS-based curves when $10,000$ simulations are used. This figure also presents the confidence intervals of the KDE curves when $10,000$ PCE-based predictions are utilized. The graph indicates that with a large data set, the variability of the parametric and non-



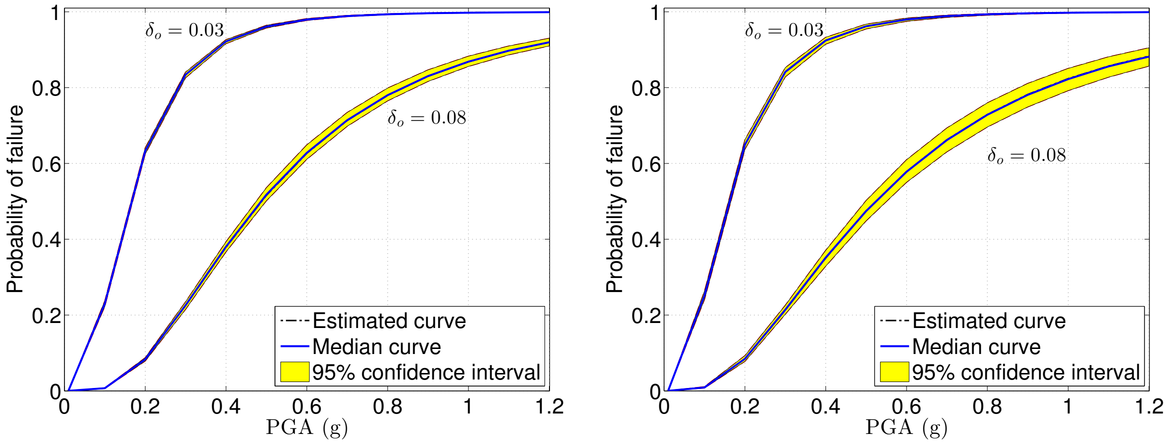
(a) Linear regression



(b) MLE

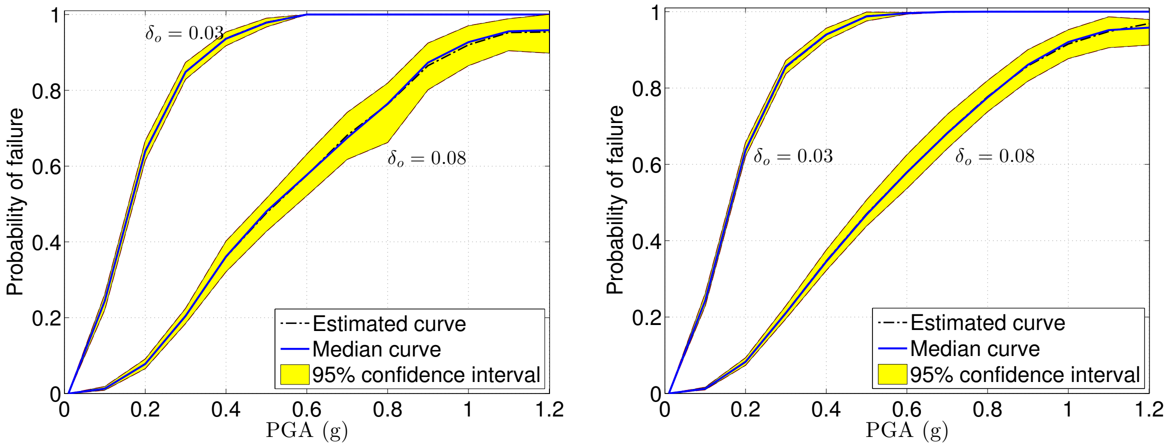
Figure 6.7 – Duffing oscillator – Median bootstrap fragility curves and 95% confidence intervals for the lognormal approaches using 200 simulations.

parametric curves are negligible as expected. For a specified IM and demand limit, the confidence intervals in non-parametric curves have similar width and are slightly larger than in lognormal curves. This is due to the fact that non-parametric approaches do not impose a strict condition on the form of the curves as parametric approaches do. It is also observed that the variability tends to increase with increasing demand limit and IM value.



(a) Linear regression

(b) MLE



(c) Binned Monte Carlo simulation

(d) KDE

Figure 6.8 – Duffing oscillator – Median bootstrap fragility curves and 95% confidence intervals for the lognormal approaches, binned Monte Carlo simulation using $10,000$ runs and for the kernel density estimation approach using $10,000$ predictions by the surrogate model.

In order to quantify the effects of epistemic uncertainty, one can estimate the variability of the median $IM_{50\%}$ defined as the IM value corresponding to 50% probability of exceedance. Assuming that the median $IM_{50\%}$ follows a lognormal distribution (Choun and Elnashai, 2010), the median $IM_{50\%}$ is determined for each bootstrap curve and the statistics of the median is computed. Table 6.1 lists the mean, standard deviation and coefficient of variation of $IM_{50\%}$ when using LR, MLE, bMCS and KDE methods. Note that the bootstrap resampling is applied with the parametric methods using the original experimental design of size 200. For all the methods, the epistemic uncertainty of the estimated medians $IM_{50\%}$ is increasing with the threshold δ_o. The variability of $IM_{50\%}$ resulting from the MLE method is systematically larger for the LR method. The coefficients of variation obtained with the parametric methods are several times higher than those obtained with non-parametric approaches. Using bMCS and KDE approaches, the coefficients of variation for the two levels of δ_o are similar and relatively small (< 0.04.) It is worth noting the KDE curves were obtained at the cost of only 200 simulations which were used as an experimental design for training the PC-NARX metamodel.

Table 6.1 – Duffing oscillator – Statistics of the $IM_{50\%}$ (IM such that the probability of failure is 50%).

δ_o	Approach	Mean	Standard deviation	COV
0.03 m	LR	0.1763 g	0.0111 g	0.0630
	MLE	0.1761 g	0.0153 g	0.0868
	bMCS	0.1646 g	0.0021 g	0.0130
	KDE	0.1651 g	0.0016 g	0.0097
0.08 m	LR	0.5708 g	0.0576 g	0.1008
	MLE	0.6082 g	0.1832 g	0.3012
	bMCS	0.5216 g	0.0194 g	0.0372
	KDE	0.5290 g	0.0162 g	0.0305

6.5.1.3 Summary

The findings in this numerical application can be summarized as follows. First, the usual practice in which only a small number of simulations is used results in fragility curves that are inaccurate, in particular for the high level of the safety threshold. In addition, the epistemic uncertainties in the estimated curves are considerable. In order to obtain more reliable curves with the parametric methods, it is important to have a sufficiently large data set, which is commonly infeasible with the computationally expensive numerical model, however can be achieved with surrogate models. Second, the non-parametric methods provide the same curves with negligible epistemic uncertainty. It is worth noting that the PC-NARX method requires only 200 simulations to attain the same level of accuracy and epistemic uncertainties as the MCS-based method which makes use of $10,000$ simulations. This numerical application showcases the effective use of PCE-based surrogate models to reduce the epistemic uncertainties in the estimation of the quantity under investigation with a significant computational efficiency in a measurable way.

6.5.2 Bouc-Wen oscillator

6.5.2.1 Computation of fragility curves

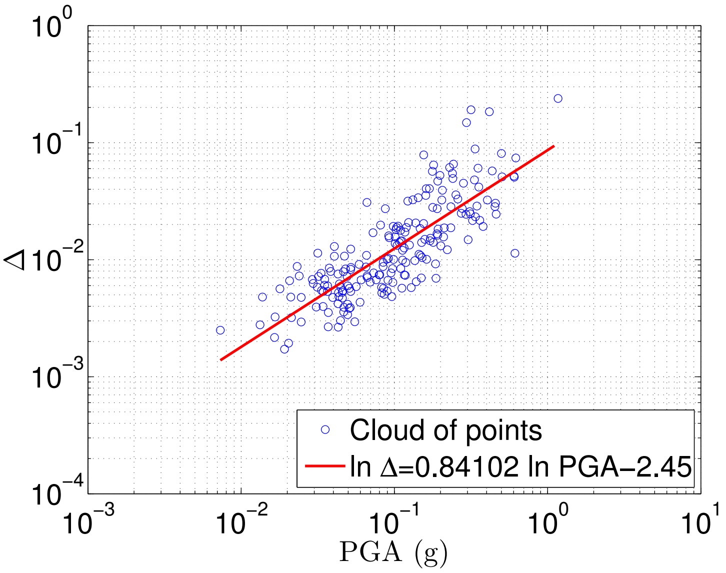
In the second case study, let us consider the Bouc-Wen oscillator which was also investigated in the previous chapter. Two demand limits are considered, namely $\delta_0 = 0.03$ m and $\delta_0 = 0.08$ m (Kafali and Grigoriu, 2007). Two data sets (PGA, Δ) are obtained from the respective 200 and $10,000$ dynamical analyses of the oscillator subject to synthetic ground motions. The paired data are depicted in Figure 6.9 together with the fitted linear regression models which have the coefficient of determination $R^2 = 0.659$ and $R^2 = 0.671$, respectively. The small data set of size $N = 200$ is used to compute the PC-NARX model, which then allows one to predict the $10,000$ responses of the oscillator without running the original numerical model. The fragility curves are computed in the following cases: (i) using parametric approaches with data from 200 simulations, (ii) using parametric and non-parametric approaches with $10,000$ simulations and finally (iii) using KDE method with $10,000$ responses predicted by the PC-NARX model fitted from the ED of size 200. Note that the bandwidth parameter h and matrix H of the KDE method when using $10,000$ simulations (resp. predictions) are $h = 0.175\, g$ and $H = \left[0.0399\, g \quad 0.0309\, g;\ 0.0309\, g \quad 0.0362\, g\right]$ (resp. $h = 0.175\, g$ and $H = \left[0.0423\, g \quad 0.0355\, g;\ 0.0355\, g \quad 0.0431\, g\right]$).

Figure 6.10 plots the parametric curves obtained from 200 simulations together with non-parametric curves using $10,000$ simulations. It is noticeable that the non-parametric curves computed with bMCS and KDE methods are in great agreement with each other. In contrast, the parametric curves differ significantly from the reference non-parametric counterparts, in particular for the limit state $\delta_0 = 0.08$ m.

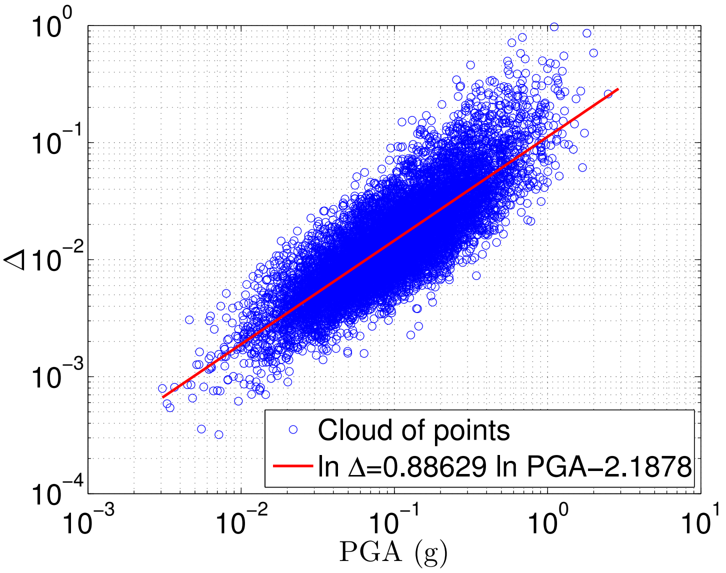
The discrepancies between parametric and non-parametric curves might be associated with the small number (200) of simulations used to compute the parametric curves. By using a large number of simulations, it is expected that the resulting parametric curves will become consistent with the reference ones. Figure 6.11 depicts the parametric curves obtained with $10,000$ simulations, which are compared with the non-parametric counterparts. One observes that the MLE-based curves are now in good agreement with the reference curves, whereas the LR-based curve for the limit state $\delta_0 = 0.08$ m still exhibits slight discrepancies. It is also worth underlying the good accuracy of the KDE-based curves obtained with $10,000$ predictions of the PC-NARX model. The latter was built using only an ED of size 200.

6.5.2.2 Estimation of epistemic uncertainty by bootstrap resampling

The effects of epistemic uncertainties are investigated by means of the bootstrap resampling technique. Figure 6.12 shows the median lognormal curves and the confidence intervals for the parametric approaches using 200 simulations. From this figure it can be seen that at one IM and one demand limit δ_0, the confidence interval of MLE-based curves is significantly larger than that of the LR-based curves. For both methods, the



(a) 200 simulations



(b) 10,000 simulations

Figure 6.9 – Bouc-Wen oscillator – Paired data and fitted linear regression model in the logarithmic scale ($R^2 = 0.659$ (a) and $R^2 = 0.671$ (b)).

variability increases with the value of IM and δ_0.

Figure 6.13 shows the small variabilities of lognormal curves and bMCS curves when using 10,000 simulations. Figure 6.13(d) depicts the confidence interval of KDE-based curves which are computed with 10,000 predictions of the PC-NARX model. The confidence intervals are small, *i.e.* the estimated curves are more reliable and they are not strongly dependent on a specific data set.

In Table 6.2, the statistics of the median $IM_{50\%}$ obtained by bootstrap resampling are

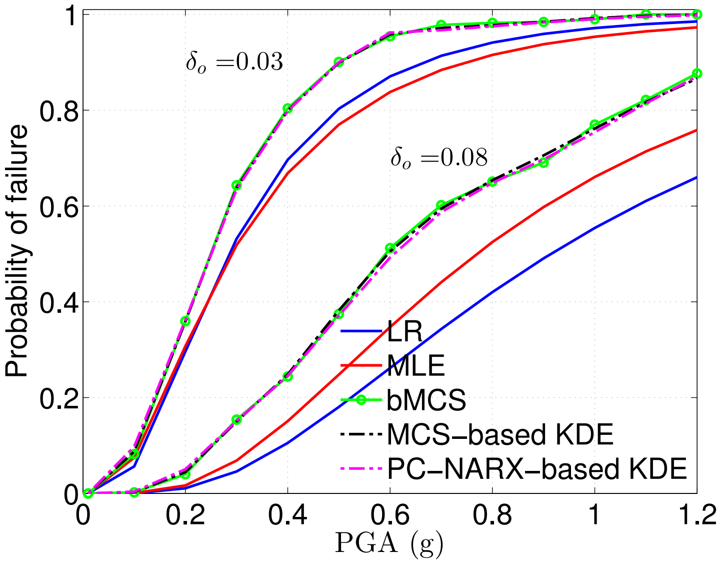


Figure 6.10 – Bouc-Wen oscillator – Fragility curves. Lognormal curves (LR, MLE) are computed with 200 simulations. bMCS and MCS-based KDE curves are obtained with 10,000 simulations. PC-NARX-based KDE curves are calculated with 10,000 predictions of the surrogate model built on the same ED of size 200

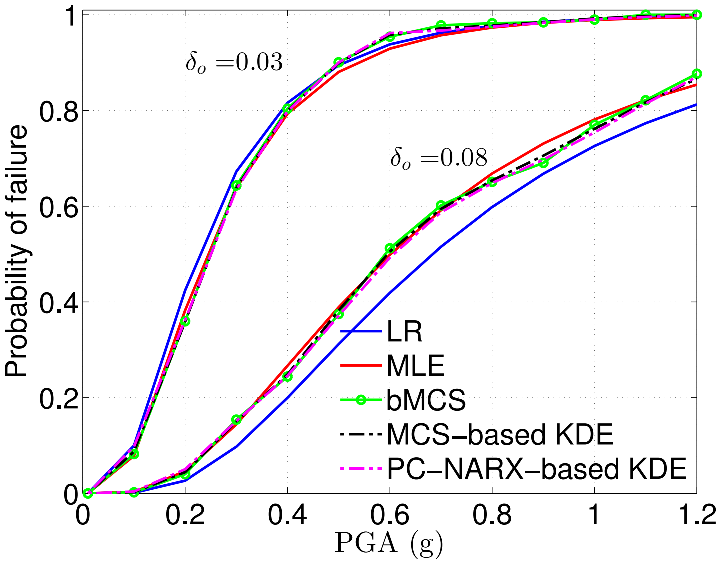
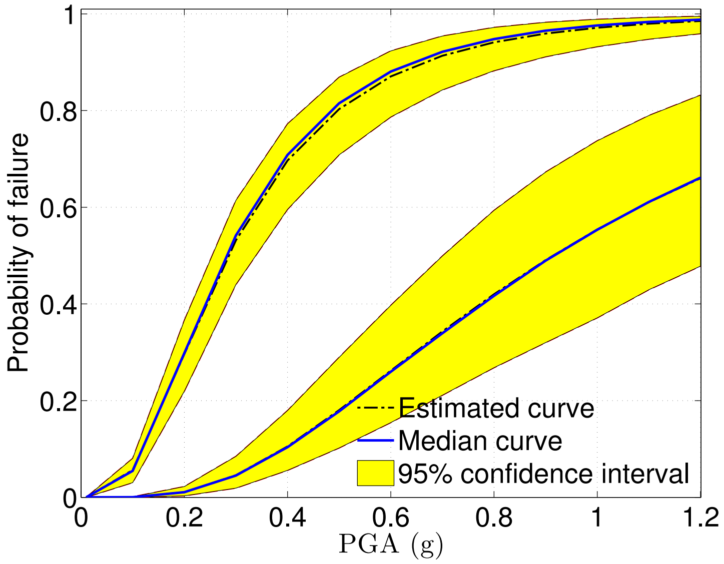
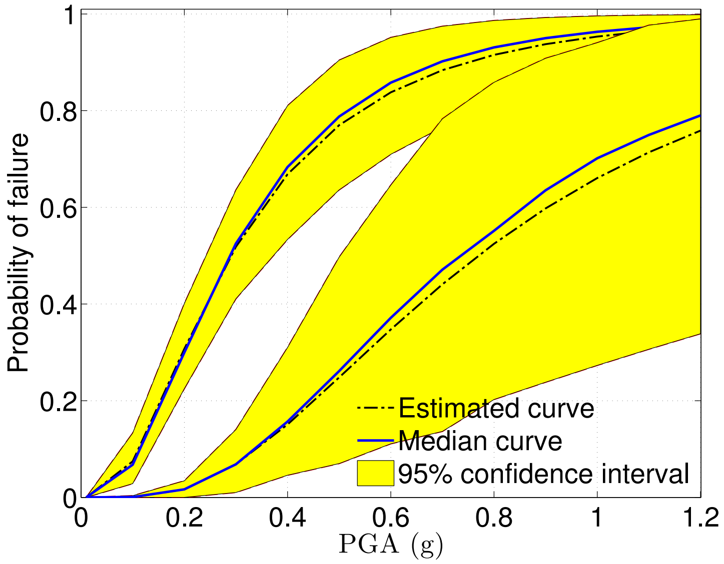


Figure 6.11 – Bouc-Wen oscillator – Fragility curves. Lognormal curves (LR, MLE), bMCS and MCS-based KDE curves are obtained with 10,000 simulations. PC-NARX-based KDE curves are computed with 10,000 predictions of the surrogate model built on an ED of size 200

presented. The remarks are similar to the Duffing oscillator case study. It is noticeable that the parametric methods result in fragility curves corrupted with large uncertainties. The coefficients of variation of the estimated $IM_{50\%}$ reachs 32.51% with MLE for the high level of δ_o. Using the parametric methods, the coefficients of variation for the two levels



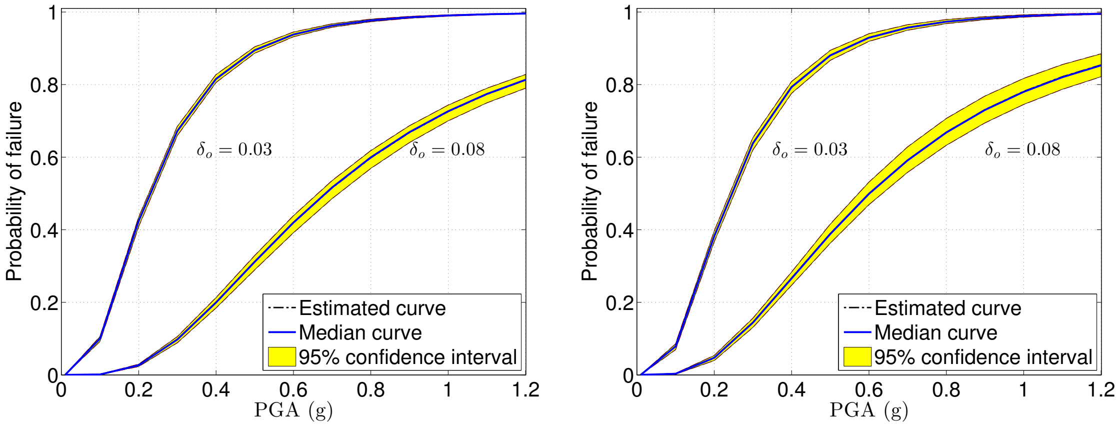
(a) Linear regression



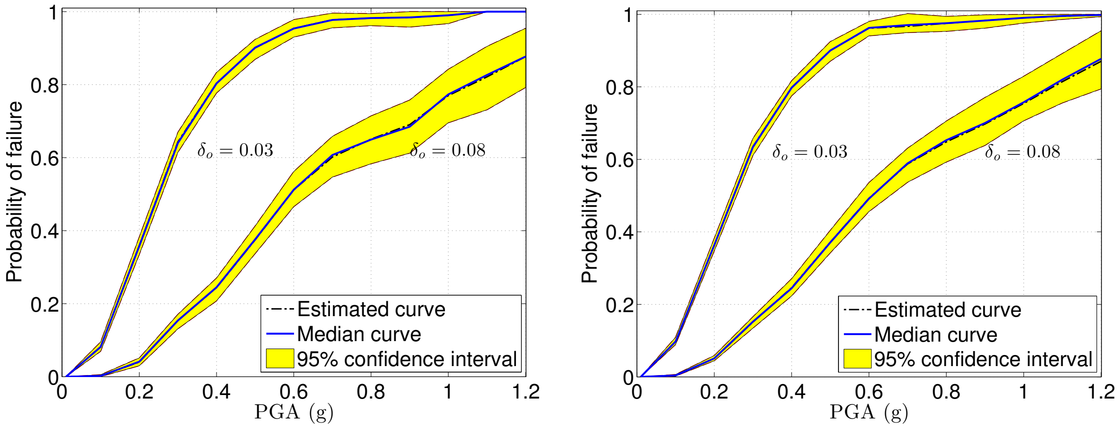
(b) MLE

Figure 6.12 – Bouc-Wen oscillator – Median bootstrap fragility curves and 95% confidence intervals for the lognormal approaches using 200 simulations.

of δ_o are relatively small (< 0.04.) It is worth noting that the PC-NARX metamodel requires only 200 simulations of the original model.



(a) Linear regression (b) MLE



(c) Binned Monte Carlo simulation (d) KDE

Figure 6.13 – Bouc-Wen oscillator – Median bootstrap fragility curves and 95% confidence intervals for the lognormal approaches, binned Monte Carlo simulation using $10,000$ runs and the kernel density estimation approach using $10,000$ predictions by the surrogate model.

6.5.2.3 Summary

The investigation above is concluded as follows. First, the fragility curves obtained by parametric methods with 200 simulations differ significantly from the true curves. Furthermore, they exhibit considerable epistemic uncertainties, in particular with MLE. The inaccuracy of the parametric fragility curves and associated epistemic uncertainties can be reduced at a large computational cost, when $10,000$ simulations are used. The non-parametric methods provide similarly reliable curves. Most importantly, at the cost of only 200 simulations the PC-NARX surrogate model allows us to obtain the estimations comparable to MCS regarding the accuracy and the reduced epistemic uncertainties.

Table 6.2 – Bouc-Wen oscillator – Statistics of the $IM_{50\%}$ (IM such that the probability of failure is 50%).

δ_o	Approach	Mean	Standard deviation	COV
0.03 m	LR	0.2845 g	0.0210 g	0.0737
	MLE	0.2932 g	0.0344 g	0.1172
	bMCS	0.2499 g	0.0033 g	0.0130
	KDE	0.2507 g	0.0031 g	0.0125
0.08 m	LR	0.9201 g	0.1354 g	0.1471
	MLE	0.8321 g	0.2705 g	0.3251
	bMCS	0.5916 g	0.0203 g	0.0343
	KDE	0.6092 g	0.0204 g	0.0335

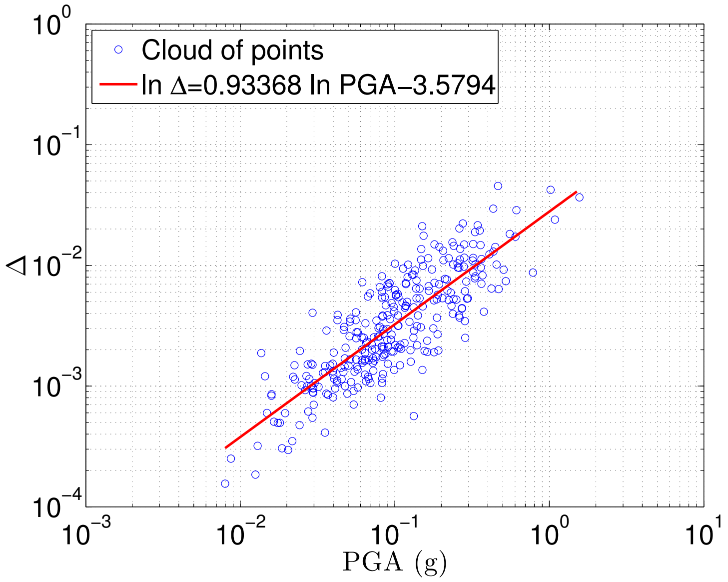
6.5.3 Steel frame

Let us consider the steel frame which has already been investigated in the previous chapter. Two demand limits of the inter-story displacement are of interest, namely $\delta_0 = 0.021$ m and $\delta_0 = 0.045$ m. The two limit states correspond to the damage limitation recommended by American Society of Civil Engineers (2000) and Eurocode 8 (2004). An experimental design of size 300 is used on the one hand. 10,000 simulations of the steel frame subject to synthetic ground motions are conducted for reference solutions. Figure 6.14 depicts the paired $(\ln PGA, \ln \Delta)$ and the resulting linear regressions in the probabilistic seismic demand models, in which the coefficients of determination are respectively $R^2 = 0.695$ and $R^2 = 0.687$.

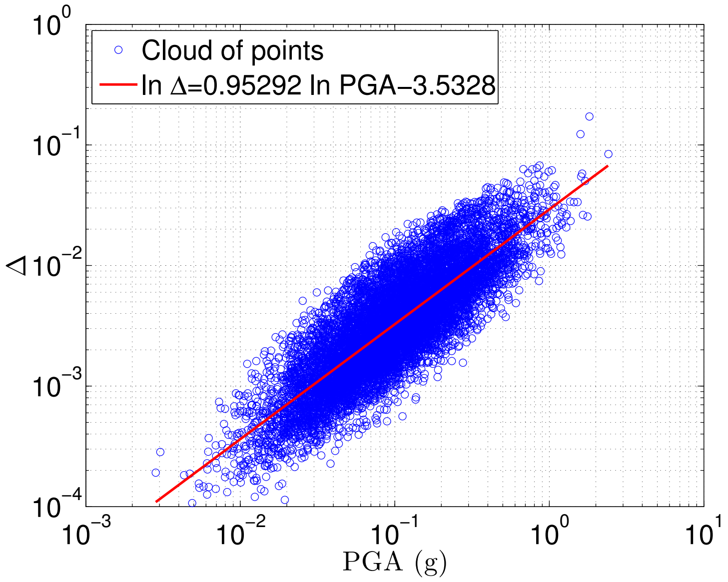
First, one computes the parametric fragility curves using 300 simulations. They are compared with the reference non-parametric curves using 10,000 simulations in Figure 6.15. It is shown that the lognormal curves are in good agreement with the reference ones, except for $\delta_0 = 0.045$ m when LR method (resp. MLE) overestimates (resp. underestimates) the failure probability. Figure 6.16 shows that using 10,000 simulations, the lognormal curves approach the reference ones. In particular, the KDE-based curves using predictions of the PC-NARX model built with 300 simulations are in good match with the reference.

Figure 6.17 depicts the variability of the lognormal curves by bootstrap resampling of the 300 simulations. It is clear that MLE-based curves exhibit excessively large uncertainties for both levels of limit states. This indicates that the resulting lognormal curves are indeed strongly sensitive to the specific data set used for their computation.

By using 10,000 simulations, the variability of lognormal curves is reduced drastically, as shown in Figure 6.18. At the same IM level and demand limit, the MLE curves exhibit larger variability than the LR curves. Regarding non-parametric approaches, the KDE-curves using predictions of the PC-NARX model exhibit similar confidence intervals as the bMCS-based curves using 10,000 simulations.



(a) 300 simulations



(b) $10,000$ simulations

Figure 6.14 – Steel frame – Paired data and fitted linear regression model in the logarithmic scale ($R^2 = 0.695$ (a) and $R^2 = 0.687$ (b)).

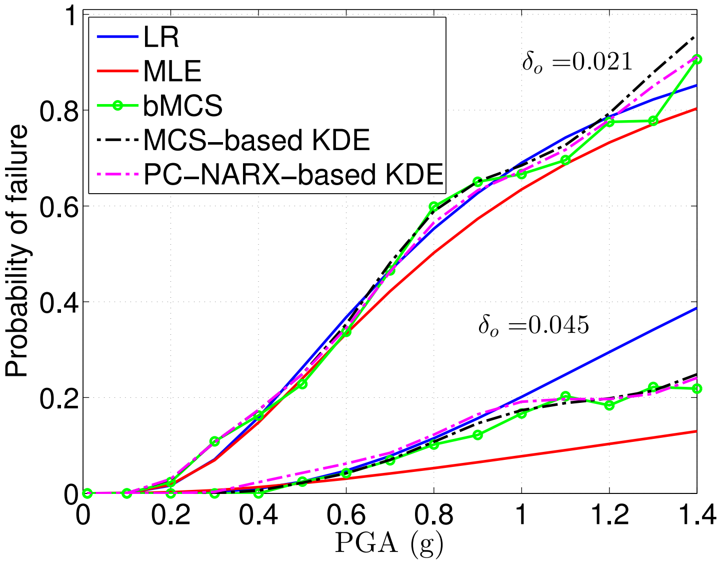


Figure 6.15 – Steel frame – Fragility curves. Lognormal curves (LR, MLE) are computed with 300 simulations. bMCS and MCS-based KDE curves are obtained with $10,000$ simulations. PC-NARX-based KDE curves are calculated with $10,000$ predictions of the surrogate model fitted with the same ED of size 300.

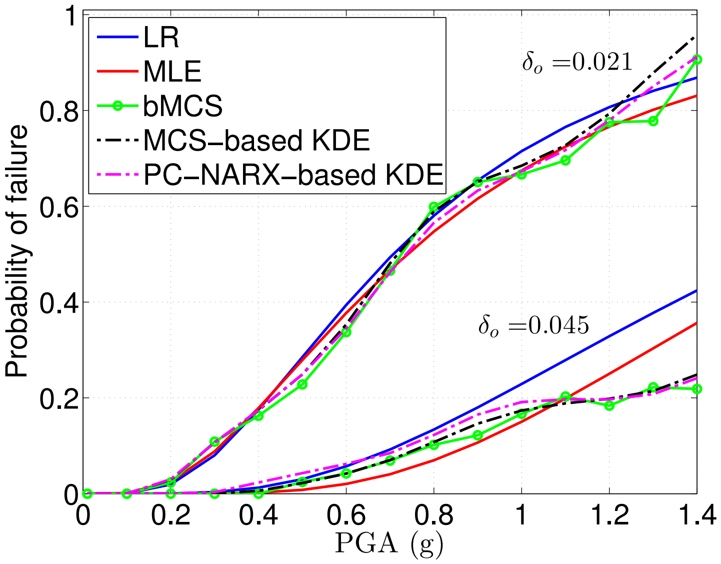
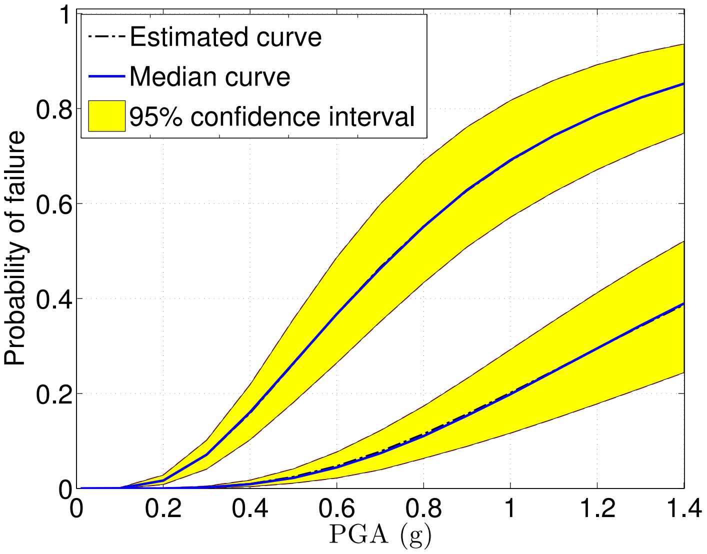
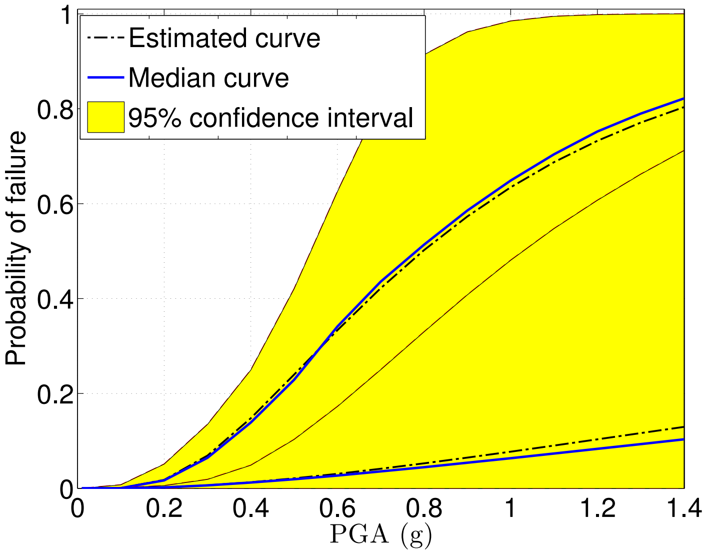


Figure 6.16 – Steel frame – Fragility curves. Lognormal curves (LR, MLE), bMCS and MCS-based KDE curves are obtained with $10,000$ simulations. PC-NARX-based KDE curves are computed with $10,000$ predictions of the surrogate model fitted with the ED of size 300.

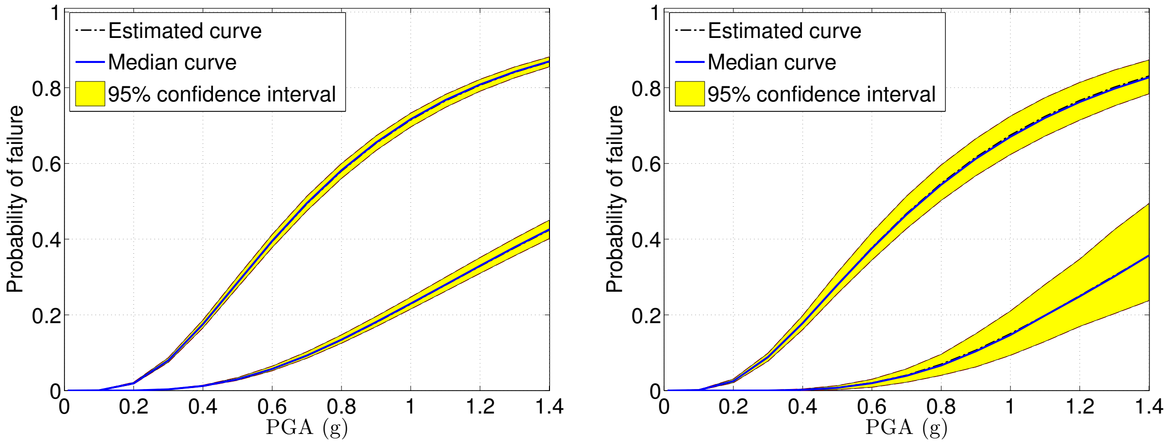


(a) Linear regression

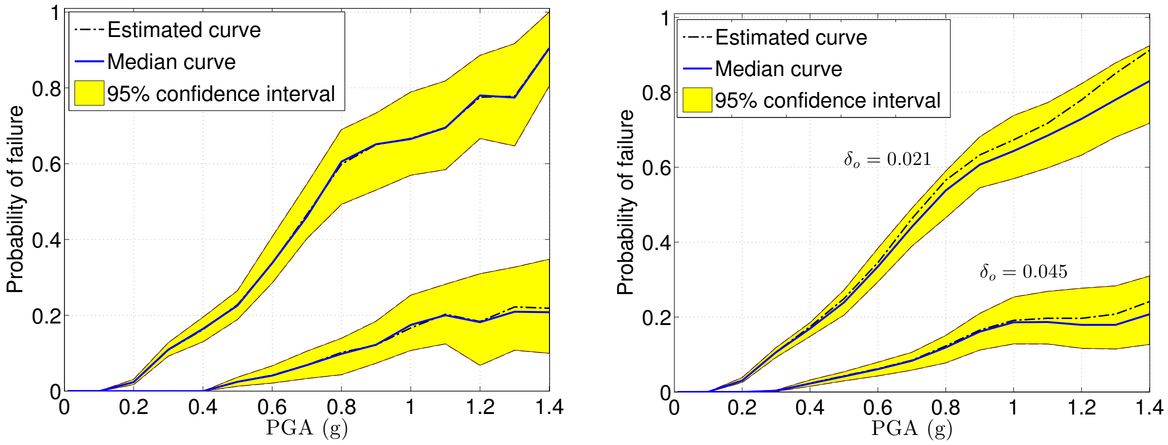


(b) MLE

Figure 6.17 – Steel frame – Median bootstrap fragility curves and 95% confidence intervals for the lognormal approaches using 300 simulations.



(a) Linear regression (b) MLE



(c) Binned Monte Carlo simulation (d) PC-NARX using KDE

Figure 6.18 – Steel frame – Median bootstrap fragility curves and 95% confidence intervals for the lognormal approaches, binned Monte Carlo simulation using 10, 000 runs and the kernel density estimation approach using 10, 000 predictions by the surrogate model.

6.6 Discussion

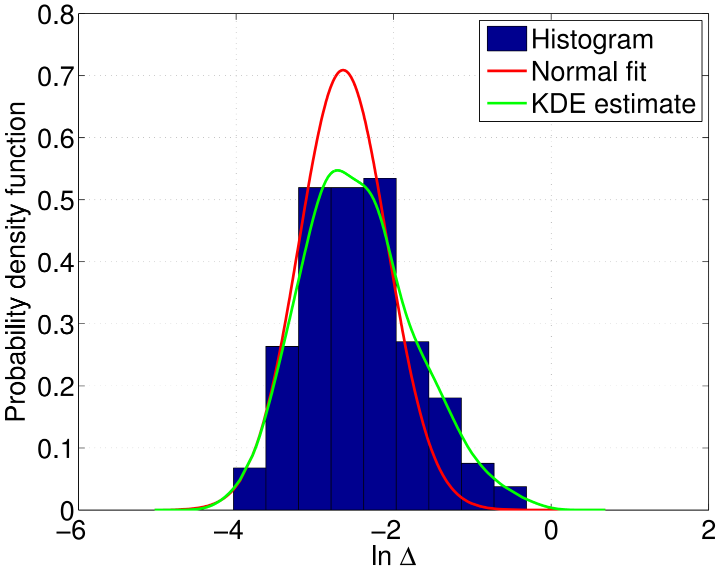
Based on the obtained results, one can draw the following conclusions on the accuracy of the parametric approaches and the applicability of non-parametric approaches for computing seismic fragility curves.

The results suggest that the lognormal curves built with a few hundred simulations (200 or 300) might differ significantly from reference non-parametric curves obtained with a large number of simulations $(10,000)$. Most importantly, with a small data set, the lognormal curves exhibit large epistemic uncertainties, which are more pronounced for MLE than LR method. It indicates that the lognormal curves are strongly dependent on a small data set. This conclusion on the effects of epistemic uncertainties is consistent with findings reported by Liel et al. (2009); Zentner et al. (2011); Rajeev and Tesfamariam (2012). This highlights the need to use a sufficiently large data set even when using parametric approaches.

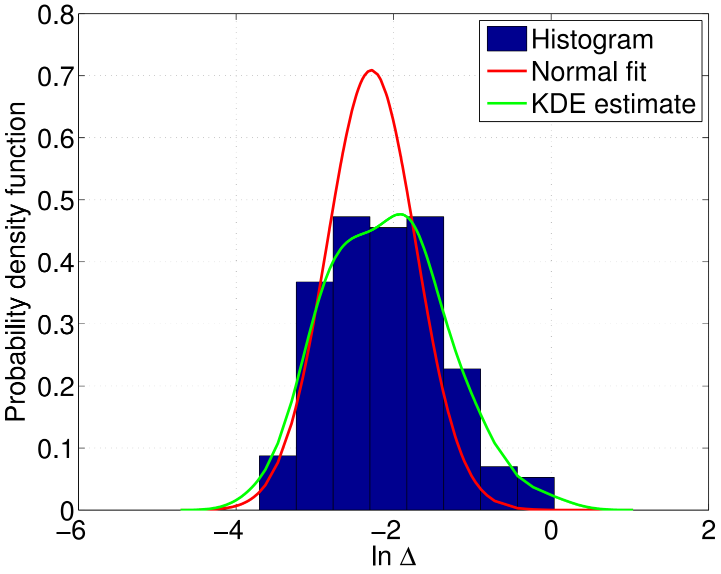
When using a larger data set, the lognormal curves tend to approach the non-parametric curves. However, slight discrepancies can still be observed in some cases. In general, given a sufficiently large data set, the MLE-based lognormal approach yields fragility curves that are overall close to the non-parametric ones; however, it smooths out some details of the curves that can be obtained with the non-parametric approaches. As noted in Section 6.2.2, the LR approach assumes that the residuals of the fitted model in the log-scale (Eq. (6.6)) follow a normal distribution with a constant standard deviation independently of the IM level. Figure 6.19 shows histograms and KDE-based probability density functions of $\ln \Delta$ at two example levels of PGA together with the fitted normal distributions according to Eq. (6.6). The responses Δ at each IM level are obtained consistently with the bMCS approach. Obviously, the assumption of a normal distribution is not exactly correct. However, in the numerical examples considered herein, the effects of the assumptions on homoscedasticity and constant error are not pronounced.

As a summary, using the non-parametric fragility curves as reference, the accuracy of the lognormal curves is found to depend on the method used to estimate the parameters of the underlying CDF and the drift threshold of interest. In most cases, given sufficient data the MLE-based curves are fairly close to the non-parametric ones, whereas the LR-based curves might exhibit non-negligible deviations. The lognormal curves tend to deviate more from the non-parametric ones for larger drift limits.

As noted in Section 6.3.1, the bMCS approach bears similarities with the so-called stripe analysis (Shome et al., 1998; Bazzurro et al., 1998; Jalayer and Cornell, 2009). To some extents, the KDE approach can be considered as a cloud analysis method. This is due to the fact that KDE method also uses the entire data set to compute the required probability density functions. The difference with the classical cloud approach is that no assumption on the probability density functions is taken. A comparison between the stripe and cloud analyses, where the latter corresponds to the LR-based lognormal approach (Cornell et al., 2002; Baker, 2007; Jalayer et al., 2014), was carried out by Celik and Ellingwood (2010). In the mentioned study, concrete structures were subject to 40 synthetic ground motions. Differences in the response statistics obtained with the two



(a) $PGA = 0.6\,g$



(b) $PGA = 0.9\,g$

Figure 6.19 – Bouc-Wen oscillator – Histograms, fitted normal distributions and kernel density estimates for $\ln \Delta$ at two levels of PGA.

methods for three IM levels were found insignificant and hence, use of the cloud analysis was justified. In contrast, Baker (2007) showed that cloud analysis can significantly underestimate the mean annual rate of exceeding a large maximum interstory drift. Karamlou and Bocchini (2015) recently conducted large-scale simulations on bridge structures in order to investigate the underlying assumptions of the cloud analysis. Their results indicated that, in general, the conditional distribution of a demand parameter for a given IM level is not lognormal. In addition, it was found that the assumptions of a linear function for the probabilistic seismic demand model in the log-scale (power function in the normal scale) and of constant dispersion of the respective errors can lead to significant errors in fragility analysis. The limitations of the LR-based approach have also been mentioned by Jalayer et al. (2014). All those findings corroborate our current results.

Based on the results of our case studies and the above discussion, we recommend the use of the MLE approach if fragility curves are developed in a parametric manner and most importantly, a sufficiently large data set is available. In any case, it is recommended that the epistemic uncertainties of fragility curves are quantified and used as an indicator of the reliability of the curves. However, when a detailed description of the fragility function is important, a non-parametric approach should be used. The non-parametric bMCS and KDE methods require a large number of simulations, which can be typically obtained by conducting structural analyses with synthetic ground motions. Note that thanks to the PC-NARX metamodelling techniques, this can be obtained at a reduced computational cost of 200-300 simulations. The metamodels naturally take into account the aleatory uncertainties induced by the structures properties and earthquake excitations. In addition, the use of metamodels offers the possibility to reduce significantly the epistemic uncertainties in the resulting fragility curves with a considerable measurable computational efficiency. We again emphasize that the two proposed non-parametric approaches lead to almost identical curves in the case when they could be applied independently with the same large dataset.

6.7 Summary

The current chapter provided a brief review on the use of seismic fragility curves in the context of performance-based earthquake engineering. Different methods for computing these curves were presented. First, two commonly used parametric approaches were reviewed. To provide an alternative and a reference to validate the parametric curves, two non-parametric approaches were introduced. The analysis on three numerical case studies indicated that parametric curves might exhibit significant inaccuracies when only a small number of dynamical analyses of the system is performed. The applications of surrogate modelling in this context was investigated. It was shown that use of surrogate models allows one to overcome the problem related to a limited number of dynamical analyses, thus accurate non-parametric curves can be obtained.

7

Conclusions and outlook

7.1 Summary and scientific achievements

The present manuscript aims at developing polynomial chaos expansions-based methods to solve selected problems of stochastic dynamics. The targeted applications include systems with uncertain governing parameters subject to stochastic excitations, namely chemical systems and structures under seismic motions. This is by nature a challenging task because one has to deal at once with two tough topics, namely the dynamics (in most cases, nonlinear dynamics) and uncertainty quantification.

To this end, it is important to conduct in the first place a thorough literature review of the methods proposed to handle the uncertain dynamics. Chapter 3 has focussed on presenting these methods in an unprecedented way. Instead of putting emphasis on the detailed formulations and equations, the review shed light on the ideas behind the utilized techniques. Surprising methodological similarities between methods that look apparently different at first glance were revealed. For the sake of comprehensiveness and clarity, these methods were classified into two groups according to their closeness.

The first group is characterized purely by the direct consideration of the visible effects of the dynamics. More precisely, the increasingly strong nonlinearity in the random space, which is the consequence of the dynamics, is handled directly by different techniques, namely using high-degree polynomial chaos expansions (PCEs) or local functions. Therefore, the stochastic dynamics is handled uniquely by the uncertainty propagation tool. The second group gathers the methods which clearly distinguish the two aspects of the problems. The dynamics is taken into account by specific representations. These techniques usually rely on representing future responses as functions of its past values, *e.g.* adding random variables on-the-fly and using autoregressive models. In special cases involving oscillatory behaviours, the dynamics is captured by a time transformation scheme or a nonlinear pre-processing of the responses. With the approach of the second group, the consideration of uncertainties tends to be less complex than with methods in the first group where effects of dynamics and uncertainties are mixed up. After a careful consideration of the existing methods, the conclusion was that it is more promising to follow the approach advocated by the latter group. Inspired by two original ideas, two non-intrusive PC-based methods were investigated.

Chapter 4 considers the use of a PC-based stochastic time-warping method which aims at solving problems of random responses of oscillatory systems that are common in mechanical and chemical engineering. The time-warping approach exploits the oscillatory feature of the response. In this approach, first a similarity measure, which indicates the closeness between different time-histories in frequency and phase contents, is introduced. The random response time-histories are then projected on an auxiliary time scale by a stochastic time-warping process, which is determined by maximizing the similarities between the responses and a predefined reference trajectory. In the auxiliary time scale, the responses exhibit the dynamics which is close to a referent mechanism. The uncertain magnitudes of the responses are finally represented by PCEs. The approach was applied successfully to various chemical and mechanical systems with uncertain governing parameters, initial conditions and stochastic periodic excitations. Some of the applications exhibit complex physical phenomena like hysteretic behaviour, chaotic behaviour and bifurcation.

Chapter 5 aims at solving the stochastic responses of mechanical systems with uncertain properties subject to non-stationary excitations with varying temporal and spectral contents. To this end, an original approach is investigated which tackles step-by-step the two aspects of dynamics and uncertainties. First, the dynamical behaviour of the system is captured by a stochastic nonlinear autoregressive with exogenous input (NARX) model which consists in representing the future responses as functional forms of past values of excitations and responses. Second, the stochastic NARX parameters are represented with PCEs. The originality is that a regression-based technique is used to select both NARX terms and polynomial functions. The approach was used to solve various structural dynamic problems in earthquake engineering where uncertainties from the structures and the excitations are accounted for. In particular, in the considered applications the excitations are synthetic stochastic non-stationary ground motions generated by a probabilistic model. The obtained PC-NARX surrogate models were used to predict responses of the systems and study their statistics (*i.e.* instantaneous and maximal responses, evolutionary statistics). The obtained results were superior to those reported so far in the literature.

The conducted investigations in Chapters 3, 4 and 5 lead to the following conclusions. First, the traditional approach to handle the uncertain dynamics with PCEs might not be the optimal solution. The classical formulation of PCEs with polynomial functions and deterministic time-dependent coefficients, which is convenient for computation and post-processing, is indeed not suitable for representing the stochastic dynamics. The reader is referred to Chapter 3 for longer discussions. Second, a different strategy should be used instead of the traditional approach, in which the constituents of the problems, *i.e.* the dynamics and uncertainties, are handled by corresponding tools which are designed for that particular purpose. The time-warping and NARX modelling techniques are good options to capture the dynamics. There are, however, various techniques in this class to be investigated.

Using PC-NARX-based surrogate models, one can obtain the results of a large number of structural simulations at a reduced computational cost, which is the cost of computing the experimental design of limited size. This allows the use of recently proposed non-parametric methods for computing seismic fragility curves as shown in Chapter 6. The commonly used assumption of the lognormal shape of the curves was assessed. It was shown that the lognormality assumption has significant effects on the resulting curves, which might differ largely from non-parametric counterparts in some cases, in particular when only a few hundred simulations are used to compute the lognormal curves. This emphasizes the need for metamodels, which can be used as substitutes of the original models to obtain accurate fragility curves and reduce significantly their epistemic uncertainties.

7.2 Limitations and outlook

It was shown through theoretical discussions and various numerical applications that the proposed methods outperform the classical PCEs in handling stochastic dynamics. The difference lies on the formulations of the utilized time-dependent PCEs. In the classical representation, the PC time-dependent coefficients are deterministic instant-wise whereas a different strategy is advocated by the proposed methods. In the PC-NARX approach, the PC coefficients are not considered deterministic but functions of the responses' past values. In the time-warping approach, the PC representation takes deterministic coefficients in the transformed time scale. However, the stochastic inverse time warping process imposes that the PC coefficients in the physical time scale are dependent on the uncertain parameters. In summary, we endorse the use of PCEs that are different from the classical representation to solve stochastic dynamics. This remedy constitutes, however, a limitation for the proposed methods. By discarding the classical time-dependent PC representation, one can no longer make use of the associated straightforward post-processing. Instead of using analytical formulations, one currently has to perform Monte Carlo simulation on the surrogate model to obtain the responses' statistical moments. For the sake of completeness of the proposed methods, it would be relevant to derive rigorous analytical solutions for post-processing the time-warping and PC-NARX models. It is worth noting that in Chapter 2, a derivative-based sensitivity measure was introduced as a mere post-processing of PCEs. However, the technique was not applied to case studies in the following chapters featuring PCE techniques designed for uncertain dynamics. This is also due to the fact that the proposed PCE schemes rely on formulations that are different from conventional approach. The missing link constitutes a limitation for the conducted work.

The effectiveness of the proposed methods were demonstrated with various engineering applications, which are the most complicated benchmark problems found in the literature. It is worth noting that modifications were made so as to increase the complexity of the case studies. For instance, longer time ranges and additional uncertain parameters were considered, non-trivial nonlinear dynamical behaviours were investigated and non-stationary excitations were used instead of white-noise time-series. For the considered case studies, a remarkable performance of the proposed methods was recorded. However, the author is well aware that it remains challenging to apply these methods to large numerical models in engineering, as seen in the last applications in Chapters 4 and 5. Future adaptation and improvement are expected to improve the performance of the proposed methods. For instance, when a large DOF system is under consideration, it might be advantageous to couple the proposed methods with model order reduction techniques which can help lower the computational complexity of the problem. When complex nonlinear behaviours are of interest, it is of utmost importance to understand the characteristics of the dynamical phenomenon, thus being capable of mimicking its evolution.

Furthermore, the algorithms proposed in this manuscript are designed only to solve problems in the time domain. It is worth noting that considering a dynamical system in the frequency domain is also of common practice. The approach advocated herein, *i.e.* representing the dynamics (resp. the uncertainties) by a specific tool (resp. PCEs), can be extended easily to the investigations in the frequency domain.

In a general case, *e.g.* in stochastic fluid dynamics, the random responses are functions of not only the time variable but also the spatial coordinates. This is actually an extension of the problem investigated in the current manuscript, when the spatial variables are additionally taken into account. The dynamics of the system is therefore dependent on both time and space, which makes the problem even more challenging. Based on the current understanding, the author is convinced that the commonly used approach, which relies on projecting the responses on deterministic time- and/or space-dependent basis functions, is not effective except for very simple case studies. A novel approach, which accounts for the dynamics in time and space and handles the uncertainties in a separated manner, is needed.

To close the discussions, the author would like to emphasize once again the strategy promoted in the manuscript: give back to PCEs *only* what are uncertainties', and to another specific tool what are dynamics'. May this strategy be useful for future researches.

Appendices

A

First order derivative of polynomial chaos functions

A.1 Hermite polynomials

The classical Hermite polynomials $\{He_n, n \in \mathbb{N}\}$, where n determines the degree of the polynomial, are defined on the set of real numbers $\mathbb{R}$ so as to be orthogonal with respect to the Gaussian probability measure and associated inner product:

$$\langle He_m,\, He_n \rangle \overset{\text{def}}{=} \int_{\mathbb{R}} He_m(z) He_n(z) \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z = n!\,\delta_{mn} \tag{A.1}$$

The Hermite polynomials satisfy the following differential equation (Abramowitz and Stegun, 1970, Chap. 22):

$$\frac{\mathrm{d}}{\mathrm{d}z} He_n(z) = n He_{n-1}(z) \tag{A.2}$$

From Eq. (A.1) the norm of Hermite polynomials reads:

$$\langle He_n,\, He_n \rangle = n! \tag{A.3}$$

so that the *orthonormal Hermite polynomials* are defined by:

$$\tilde{He}_n(z) = \frac{1}{\sqrt{n!}} He_n(z) \tag{A.4}$$

Substituting for Eq. (A.4) in Eq. (A.2), one gets the derivative of orthonormal Hermite polynomial $\tilde{He}_n^{'}(z) \overset{\text{def}}{=} \dfrac{\mathrm{d}\tilde{He}(z)}{\mathrm{d}z}$:

$$\tilde{He}_n^{'}(z) = \sqrt{n}\,\tilde{He}_{n-1}(z) \tag{A.5}$$

For computational purposes the following matrix notation is introduced:

$$\left\{\tilde{He}_1^{'}(z),\, \ldots\, , \tilde{He}_n^{'}(z)\right\}^{\mathsf{T}} = \boldsymbol{C}^{\mathcal{H}} \cdot \{\tilde{He}_0(z),\, \ldots\, , \tilde{He}_{n-1}(z)\}^{\mathsf{T}} \tag{A.6}$$

which allows one to cast the derivative of the orthonormal Hermite polynomials in the initial basis. From Eq. (A.5), $\boldsymbol{C}^{\mathcal{H}}$ is obviously diagonal:

$$\boldsymbol{C}^{\mathcal{H}}_{i,j} = \sqrt{i}\,\delta_{ij} \tag{A.7}$$

A.2 Legendre polynomials

The classical Legendre polynomials $\{Le_n, n \in \mathbb{N}\}$ are defined over $[-1, 1]$ so as to be orthogonal with respect to the uniform probability measure and associated inner product:

$$\langle Le_m,\, Le_n\rangle \overset{\text{def}}{=} \int_{-1}^{1} Le_m(z) Le_n(z) \frac{\mathrm{d}z}{2} = \frac{1}{2\,n+1}\,\delta_{mn} \tag{A.8}$$

They satisfy the following differential equation (Abramowitz and Stegun, 1970, Chap. 22):

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[Le_{n+1}(z) - Le_{n-1}(z)\right] = (2n+1) Le_n(z) \tag{A.9}$$

Using the notation $Le^{'}_n(z) \overset{\text{def}}{=} \dfrac{\mathrm{d}Le_n(z)}{\mathrm{d}z}$ one can transform Eq. (A.9) into the equation:

$$\begin{aligned} Le^{'}_{n+1}(z) &= (2n+1)Le_n(z) + Le^{'}_{n-1}(z) \\ &= (2n+1)Le_n(z) + (2(n-2)+1)Le_{n-2}(z) + Le^{'}_{n-3}(z) \\ &= \cdots \end{aligned} \tag{A.10}$$

From Eq. (A.8), the norm of Legendre polynomials reads:

$$\langle Le_n,\, Le_n\rangle = \frac{1}{2\,n+1} \tag{A.11}$$

so the that the *orthonormal Legendre polynomials* read:

$$\tilde{Le}_n(z) = \sqrt{2n+1}\, Le_n(z) \tag{A.12}$$

Substituting for Eq. (A.12) in Eq. (A.10) one obtains:

$$\begin{aligned} \tilde{Le}^{'}_{n+1}(z) = \sqrt{2n+3}\,\Big[&\sqrt{2n+1}\,\tilde{Le}_n(z) + \sqrt{2(n-2)+1}\,\tilde{Le}_{n-2}(z) \\ &+ \sqrt{2(n-4)+1}\,\tilde{Le}_{n-4}(z) + \dots\Big] \end{aligned} \tag{A.13}$$

Introducing the matrix notation:

$$\left\{\tilde{Le}^{'}_1(z),\, \dots\, , \tilde{Le}^{'}_n(z)\right\}^{\mathsf{T}} = \boldsymbol{C}^{\mathcal{L}e} \cdot \{\tilde{Le}_0(z),\, \dots\, , \tilde{Le}_{n-1}(z)\}^{\mathsf{T}} \tag{A.14}$$

the matrix $\boldsymbol{C}^{Le}$ reads:

$$\boldsymbol{C}^{Le} = \begin{bmatrix} \sqrt{3} & 0 & 0 & 0 & \dots & \\ 0 & \sqrt{5}\sqrt{3} & 0 & 0 & \dots & \\ \sqrt{7}\cdot 1 & 0 & \sqrt{7}\sqrt{5} & 0 & \dots & \\ \vdots & & & & & \\ 0 & \sqrt{4p+1}\sqrt{3} & 0 & \sqrt{4p+1}\sqrt{7} & \dots & \sqrt{4n+1}\sqrt{4n-1} \end{bmatrix} \tag{A.15}$$

when $n = 2p$ is even and

$$\boldsymbol{C}^{Le} = \left[\begin{array}{ccccccc} \sqrt{3} & 0 & 0 & 0 & \dots & & \\ 0 & \sqrt{5}\sqrt{3} & 0 & 0 & \dots & & \\ \sqrt{7}\cdot 1 & 0 & \sqrt{7}\sqrt{5} & 0 & \dots & & \\ \vdots & & & & & & \\ \sqrt{4p+3}\cdot 1 & 0 & \sqrt{4p+3}\sqrt{5} & 0 & \dots & 0 & \sqrt{4p+3}\sqrt{4p+1} \end{array}\right] \tag{A.16}$$

when $n = 2p + 1$ is odd.

A.3 Generalized Laguerre polynomials

Consider a model $Y = \mathcal{M}(\boldsymbol{X})$ where the input vector $\boldsymbol{X}$ contains M independent random variables with Gamma distribution $X_i \sim \Gamma(\alpha_i, \beta_i)$, $(\alpha_i, \beta_i > 0)$ with prescribed probability density functions:

$$f_{X_i}(x_i) = {\beta_i}^{\alpha_i} \, \frac{1}{\Gamma(\alpha_i)} x^{\alpha_i - 1} e^{-\beta_i x_i} \tag{A.17}$$

where $\Gamma(\cdot)$ is the Gamma function. We first use an isoprobabilistic transform to convert the input factors into a random vector $\boldsymbol{Z} = \{Z_i, \, \dots \, , Z_M\}$ as follows:

$$Z_i = \beta_i \, X_i \tag{A.18}$$

One can prove that:

$$f_{Z_i}(z_i) = \left|\frac{\mathrm{d}x_i}{\mathrm{d}z_i}\right| f_{X_i}(x_i) = \frac{1}{\Gamma(\alpha)} {z_i}^{\alpha - 1} e^{-z_i} \tag{A.19}$$

which means $Z_i \sim \Gamma(\alpha_i, 1)$.

By definition, the generalized Laguerre polynomials $\left\{L_n^{(\alpha-1)}(z), \, n \in \mathbb{N}\right\}$, where n is the degree of the polynomial, are orthogonal with respect to the weight function $w(z) = z^{\alpha-1} e^{-z}$ over $(0, \infty)$:

$$\langle L_n^{(\alpha-1)}(z), L_m^{(\alpha-1)}(z) \rangle \overset{\text{def}}{=} \int\limits_0^{+\infty} z^{\alpha-1} e^{-z} L_n^{(\alpha-1)}(z) L_m^{(\alpha-1)}(z) \mathrm{d}z = \frac{\Gamma(n+\alpha)}{n!} \delta_{mn} \tag{A.20}$$

The derivative of $L_n^{(\alpha-1)}$ reads:

$$L_n^{\prime(\alpha-1)}(z) = -\sum_{k=0}^{n-1} L_k^{(\alpha-1)}(z) \tag{A.21}$$

Recall that one obtains the Gamma distribution by scaling the weight function $w(z)$ by $1/\Gamma(\alpha)$. Therefore in the context of PCE, we use the generalized Laguerre polynomials functions orthonormalized as follows:

$$\tilde{L}_n^{(\alpha-1)}(z) = \sqrt{\frac{n!\Gamma(\alpha)}{\Gamma(n+\alpha)}} L_n^{(\alpha-1)}(z) = \sqrt{n\,B(n,\alpha)}\,L_n^{(\alpha-1)}(z) \tag{A.22}$$

where $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the beta function. Substituting for Eq. (A.22) in Eq. (A.21) one obtains:

$$\tilde{L}_n'^{(\alpha-1)}(z) = -\sum_{k=0}^{n-1} \sqrt{\frac{\Gamma(k+\alpha+1)\,n!}{\Gamma(n+\alpha+1)\,k!}}\,\tilde{L}_k^{(\alpha-1)}(z) = -\sum_{k=1}^{n} \sqrt{\frac{B(n+1,\alpha)}{B(k,\alpha)}}\,\tilde{L}_{k-1}^{(\alpha-1)}(z) \tag{A.23}$$

Introducing the matrix notation:

$$\left\{\tilde{L}_1'(z), \ldots, \tilde{L}_n'(z)\right\}^\mathsf{T} = \boldsymbol{C}^{\mathcal{L}a} \cdot \{\tilde{L}_0(z), \ldots, \tilde{L}_{n-1}(z)\}^\mathsf{T} \tag{A.24}$$

the constant matrix $\boldsymbol{C}^{\mathcal{L}a}$ is a lower triangular matrix whose generic term reads:

$$\boldsymbol{C}_{i,j}^{\mathcal{L}a} = -\sqrt{\frac{B(i+1,\alpha)}{B(j,\alpha)}} \tag{A.25}$$

Bibliography

Abramowitz, M. and Stegun, I. *Handbook of mathematical functions.* Dover Publications, Inc., 1970.

American Society of Civil Engineers. *Commentary for the seismic rehabilitation of buildings.* Washington D.C., 2000.

Applied Technology Council. Seismic Evaluation and Retrofit of Concrete Buildings. Technical report, Redwood, California, 1996.

Argyroudis, S. and Pitilakis, K. Seismic fragility curves of shallow tunnels in alluvial deposits. *Soil Dyn. Earthq. Eng.*, 35:1–12, 2012.

ASCE 7-98. Minimum design loads for buildings and other structures. American Society of Civil Engineers, 2000.

Au, S. and Beck, J. Estimation of small failure probabilities in high dimensions by subset simulation. *Prob. Eng. Mech.*, 16(4):263–277, 2001.

Au, S. and Beck, J. Subset simulation and its application to seismic risk based on dynamic analysis. *J. Eng. Mech.*, 129(8):901–917, 2003.

Baker, J. W. and Cornell, C. A. Uncertainty propagation in probabilistic seismic loss estimation. *Structural Safety*, 30(3):236–252, 2008.

Baker, J. W. Probabilistic structural response assessment using vector-valued intensity measures. *Earthq. Eng. Struct. Dyn.*, 36(13):1861–1883, 2007.

Banerjee, S. and Shinozuka, M. Nonlinear static procedure for seismic vulnerability assessment of bridges. *Comput-Aided Civ. Inf.*, 22(4):293–305, 2007.

Banerjee, S. and Shinozuka, M. Mechanistic quantification of RC bridge damage states under earthquake through fragility analysis. *Prob. Eng. Mech.*, 23(1):12–22, 2008.

Barbato, M. and Conte, J. P. Finite element structural response sensitivity and reliability analyses using smooth versus non-smooth material constitutive models. *Int. J. Reliab. Saf.*, 1(2):3–39, 2006.

Bartlett, F. M. and Macgregor, J. G. Statistical analysis of the compressive strength of concrete in structures. *Mater. J.*, 93(2):158–168, 1996.

Bazzurro, P., Cornell, C. A., Shome, N., and Carballo, J. E. Three proposals for characterizing MDOF nonlinear seismic response. *J. Struct. Eng.*, 124(11):1281–1289, 1998.

Benaroya, H., Han, S. M., and Nagurka, M. *Probability models in engineering and science.* 2005.

Beran, P. S., Pettit, C. L., and Millman, D. R. Uncertainty quantification of limit-cycle oscillations. *J. Comp. Phys.*, 217(1):217–247, 2006a.

Beran, P., Pettit, C., and Millman, D. Uncertainty quantification of limit-cycle oscillations. *J. Comput. Phys.*, 217:217–247, 2006b.

Berveiller, M., Sudret, B., and Lemaire, M. Stochastic finite elements: a non intrusive approach by regression. *Eur. J. Comput. Mech.*, 15(1-3):81–92, 2006.

Billings, S. A. *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains.* John Wiley & Sons, 2013.

Billings, S. A. and Wei, H. L. An adaptive orthogonal search algorithm for model subset selection and non-linear system identification. *Int. J. Control*, 81(5):714–724, 2008.

Blatman, G. *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis.* PhD thesis, Université Blaise Pascal, Clermont-Ferrand, 2009.

Blatman, G. and Sudret, B. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Prob. Eng. Mech.*, 25:183–197, 2010.

Blatman, G. and Sudret, B. Adaptive sparse polynomial chaos expansion based on Least Angle Regression. *J. Comput. Phys*, 230:2345–2367, 2011.

Blatman, G. and Sudret, B. Sparse polynomial chaos expansions of vector-valued response quantities. In Deodatis, G., editor, *Proc. 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2013), New York, USA*, 2013.

Bogdanoff, J. L., Goldberg, J. E., and Bernard, M. C. Response of a simple structure to a random earthquake-type disturbance. *Bull. Seismol. Soc. Am.*, 51(2):293–310, 1961.

Bookstein, F. L. *Morphometric tools for landmark data: geometry and biology.* Cambridge University Press, 1997.

Boore, D. M. Simulation of ground motion using the stochastic method. *Pure Appl. Geophys.*, 160(3):635–676, 2003.

Borgonovo, E., Zentner, I., Pellegri, A., Tarantola, S., and de Rocquigny, E. On the importance of uncertain factors in seismic fragility assessment. *Reliab. Eng. Sys. Safety*, 109(0):66–76, 2013.

Bradley, B. A. and Lee, D. S. Accuracy of approximate methods of uncertainty propagation in seismic loss estimation. *Structural Safety*, 32(1):13–24, 2010.

Building Seismic Safety Council. NEHRP Guidelines for the Seismic Rehabilitation of Buildings, 1997.

Bunton, R. W. and Denegri, C. M. Limit cycle oscillation characteristics offighter aircraft. *J. Aircraft*, 37(5):916–918, 2000.

Campolongo, F., Cariboni, J., and Saltelli, A. An effective screening design for sensitivity analysis of large models. *Environ. Modell. Softw.*, 22(10):1509–1518, 2007.

Cantelmo, C. and Piroddi, L. Adaptive model selection for polynomial NARX models. *IET Control Theory A.*, 4(12):2693–2706, 2010.

Caughey, T. K. Equivalent linearization techniques. *J. Acoust. Soc. Am.*, 35(11):1706, 1963.

Celik, O. C. and Ellingwood, B. Seismic fragilities for non-ductile reinforced concrete frames - role of aleatoric and epistemic uncertainties. *Structural Safety*, 32(1):1–12, 2010.

Chen, S., Billings, S. A., and Luo, W. Orthogonal least squares methods and their application to non-linear system identification. *Int. J. Control*, 50(5):1873–1896, 1989.

Chen, Y., Jakeman, J., Gittelson, C., and Xiu, D. Local polynomial chaos expansion for linear differential equations with high dimensional random inputs. *SIAM J. Sci. Comput.*, 37(1):A79–A102, 2015.

Chen, Z. H. and Ni, Y. Q. On-board identification and control performance verification of an MR damper incorporated with structure. *J. Intel. Mat. Syst. Str.*, 22(14):1551–1565, 2011.

Cheng, C., Sa-Ngasoongsong, A., Beyca, O., Le, T., Yang, H., Kong, Z., and Bukkapatnam, S. Time series forecasting for nonlinear and non-stationary processes: a review and comparative study. *IIE Transactions*, 47(10):1053–1071, 2015.

Cheng, M., Hou, T. Y., and Zhang, Z. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations I: Derivation and algorithms. *J. Comp. Phys.*, 242:843–868, 2013.

Cheng, Y., Wang, L., Yu, M., and Hu, J. An efficient identification scheme for a nonlinear polynomial NARX model. *Artificial Life and Robotics*, 16(1):70–73, 2011.

Chiou, J., Chiang, C., Yang, H., and Hsu, S. Developing fragility curves for a pile-supported wharf. *Soil Dyn. Earthq. Eng.*, 31:830–840, 2011.

Choi, E., DesRoches, R., and Nielson, B. Seismic fragility of typical bridges in moderate seismic zones. *Eng. Struct.*, 26(2):187–199, 2004.

Choi, M., Sapsis, T. P., and Karniadakis, G. E. On the equivalence of dynamically orthogonal and bi-orthogonal methods: Theory and numerical simulations. *J. Comp. Phys.*, 270:1–20, 2014.

Choi, S., Grandhi, R., Canfield, R., and Pettit, C. Polynomial chaos expansion with Latin Hypercube sampling for estimating response variability. *AIAA Journal*, 45:1191–1198, 2004.

Chopra, A. K. *Dynamics of structures*, volume 3. Prentice Hall New Jersey, 1995.

Chopra, A. K., Goel, R. K., and Chintanapakdee, C. Statistics of single-degree-of-freedom estimate of displacement for pushover analysis of buildings. *J. Struct. Eng.*, 129 (4):459–469, 2003.

Choun, Y. S. and Elnashai, A. S. A simplified framework for probabilistic earthquake loss estimation. *Prob. Eng. Mech*, 25(4):355–364, 2010.

Chouvion, B. and Sarrouy, E. Development of error criteria for adaptive multi-element polynomial chaos approaches. *Mech. Syst. Signal Pr.*, 66-67:201–222, 2016.

Christopoulos, C., Pampanin, S., and Priestley, M. J. Performance-based seismic response of frame structures including residual deformations. Part I: Single-degree of freedom systems. *J. Earthq. Eng.*, 7(1):97–118, 2003.

Cimellaro, G. P., Reinhorn, A. M., D'Ambrisi, A., and De Stefano, M. Fragility analysis and seismic record selection. *J. Struct. Eng.*, 137(3):379–390, 2009.

Conte, J. P. and Peng, B. F. Fully nonstationary analytical earthquake ground-motion model. *J. Eng. Mech.*, 123(1):15–24, 1997.

Cornell, C., Jalayer, F., Hamburger, R., and Foutch, D. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *J. Struct. Eng. (ASCE)*, 128(4):526–533, 2002.

Cornell, C. A. and Krawinkler, H. Progress and challenges in seismic performance assessment. *PEER Center News*, 3(2):1–3, 2000.

Denegri, C. Limit cycle oscillation flight test results of a fighter with external stores. *J. Aircraft*, 37(5):761–769, 2000.

Der Kiureghian, A. Structural reliability methods for a seismic safety assessment: a review. *Eng. Struct.*, 18(6):412–424, 1996.

Der Kiureghian, A. The geometry of random vibrations and solutions by FORM and SORM. *Prob. Eng. Mech.*, 15(1):81–90, 2000.

Der Kiureghian, A. and Ditlevsen, O. Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2):105–112, 2009.

Der Kiureghian, A. and Fujimura, K. Nonlinear stochastic dynamic analysis for performance-based earthquake engineering. *Earthq. Eng. Struct. Dyn.*, 38(5):719–738, 2009.

Desai, A., Witteveen, J. A. S., and Sarkar, S. Uncertainty quantification of a nonlinear aeroelastic system using polynomial chaos expansion with constant phase interpolation. *J. Vib. Acoust.*, 135(5):51034, 2013.

Ditlevsen, O. and Madsen, H. *Structural reliability methods.* J. Wiley and Sons, Chichester, 1996.

Doostan, A., Validi, A., and Iaccarino, G. Non-intrusive low-rank separated approximation of high-dimensional stochastic models. *Comput. Methods Appl. Mech. Eng.*, 263: 42–55, 2013.

Douglas, J. and Aochi, H. A survey of techniques for predicting earthquake ground motions for engineering purposes. *Surveys in Geophysics*, 29(3):187–220, 2008.

Duenas-Osorio, L., Craig, J. I., and Goodno, B. J. Seismic response of critical interdependent networks. *Earthq. Eng. Struct. Dyn.*, 36(2):285–306, 2007.

Dukes, J., DesRoches, R., and Padgett, J. E. Sensitivity study of design parameters used to develop bridge specific fragility curves. In *Proc. 15th World Conf. Earthquake Eng.*, 2012.

Duong, T. *Bandwidth selectors for multivariate kernel density estimation.* PhD thesis, School of mathematics and Statistics, University of Western Australia, 2004.

Duong, T. ks: kernel density estimation and kernel discriminant analysis for multivariate data in R. *J. Stat. Softw.*, 21(7):1–16, 2007.

Eads, L., Miranda, E., Krawinkler, H., and Lignos, D. G. An efficient method for estimating the collapse risk of structures in seismic regions. *Earthq. Eng. Struct. Dyn.*, 42(1):25–41, 2013.

Efron, B. Bootstrap methods: another look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. Least angle regression. *Annals of Statistics*, 32:407–499, 2004.

Elishakoff, I. and Zhang, R. Comparison of the new energy-based versions of the stochastic linearization technique. In Bellomo, N. and Casciati, F., editors, *Nonlinear Stochastic Mechanics: IUTAM Symposium*, pages 201–212. Springer, 1992.

Ellingwood, B. and Kinali, K. Quantifying and communicating uncertainty in seismic risk assessment. *Structural Safety*, 31(2):179–187, 2009.

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

Eurocode 8. *Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings*, 2004.

Faravelli, L. Response surface approach for reliability analysis. *J. Eng. Mech.*, 115(12): 2763–2781, 1989.

Filippou, F. C., Popov, E. P., and Bertero, V. V. Effects of bond deterioration on hysteretic behavior of reinforced concrete joints. Technical report, Earthquake Engineering Research Center, University of California, Berkeley, 1983.

Forrester, A., Sobester, A., and Keane, A. *Engineering design via surrogate modelling: a practical guide.* Wiley, 2008.

Franchin, P. Reliability of uncertain inelastic structures under earthquake excitation. *J. Eng. Mech.*, 130(2):180–191, 2004.

Frankel, A. D., Mueller, C. S., Barnhard, T. P., Leyendecker, E. V., Wesson, R. L., Harmsen, S. C., Klein, F. W., Perkins, D. M., Dickman, N. C., and Hanson, S. L. USGS national seismic hazard maps. *Earthquake spectra*, 16(1):1–19, 2000.

Fujimura, K. and Der Kiureghian, A. Tail-equivalent linearization method for nonlinear random vibration. *Prob. Eng. Mech.*, 22(1):63–76, 2007.

Gardoni, P., Der Kiureghian, A., and Mosalam, K. M. Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations. *J. Eng. Mech.*, 128(10):1024–1038, 2002.

Gencturk, B., Elnashai, A., and Song, J. Fragility relationships for populations of woodframe structures based on inelastic response. *J. Earthq. Eng.*, 12:119–128, 2008.

Gerritsma, M., van der Steen, J.-B., Vos, P., and Karniadakis, G. Time-dependent generalized polynomial chaos. *J. Comput. Phys*, 229(22):8333–8363, 2010.

Ghanem, R., Masri, S., Pellissetti, M., and Wolfe, R. Identification and prediction of stochastic dynamical systems in a polynomial chaos basis. *Comput. Methods Appl. Mech. Engrg.*, 194:1641–1654, 2005.

Ghanem, R. and Spanos, P. *Stochastic finite elements – A spectral approach.* Springer Verlag, New York, 1991. (Reedited by Dover Publications, Mineola, 2003).

Ghosh, D. and Ghanem, R. Stochastic convergence acceleration through basis enrichment of polynomial chaos expansions. *Int. J. Numer. Meth. Engng.*, 73(2):162–184, 2008.

Ghosh, D. and Iaccarino, G. Applicability of the spectral stochastic finite element method in time-dependent uncertain problems. *Annual Research Briefs of Center for Turbulence Research*, pages 133–141, 2007.

Ghosh, J. and Padgett, J. E. Aging considerations in the development of time-dependent seismic fragility curves. *J. Struct. Eng.*, 136(12):1497–1511, 2010.

Gidaris, I., Taflanidis, A. A., and Mavroeidis, G. P. Kriging metamodeling in seismic risk assessment based on stochastic ground motion models. *Earthq. Eng. Struct. Dyn.*, 44(14):2377–2399, 2015.

Goldberg, D. E. *Genetic algorithms in search, optimization and machine learning.* Addison-Wesley Longman Publishing Co., Inc., 1989.

Graves, R. W. and Pitarka, A. Broadband ground-motion simulation using a hybrid approach. *Bull. Seismol. Soc. Am.*, 100(5A):2095–2123, 2010.

Günay, S. and Mosalam, K. M. PEER performance-based earthquake engineering methodology, revisited. *J. Earthq. Eng.*, 17(6):829–858, 2013.

Günay, S. and Mosalam, K. M. PEER performance-based earthquake engineering methodology, revisited. *J. Earthq. Eng.*, 17(6):829–858, 2013.

Güneyisi, E. M. and Altay, G. Seismic fragility assessment of effectiveness of viscous dampers in R/C buildings under scenario earthquakes. *Structural Safety*, 30(5):461–480, 2008.

Guo, W., Lin, G., Christlieb, A., and Qiu, J. An adaptive WENO collocation method for differential equations with random coefficients. *Mathematics*, 4(2):29, 2016.

Gupta, A. and Krawinkler, H. Estimation of seismic drift demands for frame structures. *Earthquake Eng. Struct. Dyn.*, 29(9):1287–1305, 2000.

Hampton, J. and Doostan, A. Compressive sampling of polynomial chaos expansions: Convergence analysis and sampling strategies. *J. Comp. Phys.*, 280:363–386, jan 2015.

Harbitz, A. Efficient and accurate probability of failure calculation by use of the importance sampling technique. In *Proc. 4th Int. Conf. on Appl. of Statistics and Probability in Soil and Structural Engineering (ICASP4), Firenze, Italy*, pages 825–836. Pitagora Editrice, 1983.

Heuveline, V. and Schick, M. A hybrid generalized polynomial chaos method for stochastic dynamical systems. *Int. J. Uncertainty Quantification*, 4(1):37–61, 2014.

Hilbert, M. and López, P. The world's technological capacity to store, communicate, and compute information. *Science (New York, N.Y.)*, 332(6025):60–5, 2011.

Holmes, W. Reconnaissance reports on hospitals, Chile EERI/PEER reconnaissance brefing at UC Berkeley. Technical report, Berkeley, California, 2010.

Housner, G. W. and Jennings, P. C. Earthquake design criteria. Technical report, Earthquake Engineering Research Institute, Berkeley, CA, 1982.

Huh, J. and Haldar, A. Stochastic finite-element-based seismic risk of nonlinear structures. *J. Struct. Eng.*, 127(3):323–329, 2001.

Hurtado, J. E. and Barbat, A. H. Monte Carlo techniques in computational stochastic mechanics. *Arch. Comput. Methods Eng.*, 5(1):3–29, 1998.

Hwang, H. H. M., Lin, H., and Shinozuka, M. Seismic performance assessment of water delivery systems. *J. Infrastruct. Syst.*, 4(3):118–125, 1998.

Ichimura, T., Fujita, K., Tanaka, S., Hori, M., Lalith, M., Shizawa, Y., and Kobayashi, H. Physics-based urban earthquake simulation enhanced by 10.7 BlnDOF x 30 K time-step unstructured FE non-linear seismic wave simulation. In Damkroger, T. and Dongarra, J., editors, *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 15–26. IEEE, 2014.

Iervolino, I., Fabbrocino, G., and Manfredi, G. Fragility of standard industrial structures by a response surface-based method. *J. Earthq. Eng.*, 08(06):927—-945, 2004.

Iyengar, R. N. and Iyengar, K. T. S. R. A nonstationary random process model for earthquake accelerograms. *Bull. Seismol. Soc. Amer.*, 59(3):1163–1188, 1969.

Jacquelin, E., Adhikari, S., Sinou, J.-J., and Friswell, M. I. Polynomial chaos expansion and steady-state response of a class of random dynamical systems. *J. Eng. Mech.*, 141 (4):04014145, 2015.

Jakeman, J., Eldred, M., and Sargsyan, K. Enhancing ℓ_1-minimization estimates of polynomial chaos expansions using basis selection. *J. Comput. Phys.*, 289:18–34, 2015.

Jalayer, F. and Cornell, C. A. Alternative non-linear demand estimation methods for probability-based seismic assessments. *Earthq. Eng. Struct. Dyn.*, 38(8):951–972, 2009.

Jalayer, F., De Risi, R., and Manfredi, G. Bayesian cloud analysis: efficient structural fragility assessment using linear regression. *Bull. Earthquake Eng.*, 13(4):1183–1203, 2014.

Jensen, H. and Iwan, W. D. Response of systems with uncertain parameters to stochastic excitation. *J. Eng. Mech.*, 118(5):1012–1025, 1992.

Jeong, S. H., Mwafy, A. M., and Elnashai, A. S. Probabilistic seismic performance assessment of code-compliant multi-story RC buildings. *Eng. Struct.*, 34:527–537, 2012.

Jiang, X. Q. and Kitagawa, G. A time varying coefficient vector AR modeling of non-stationary covariance time series. *Signal Processing*, 33(3):315–331, 1993.

Joint Committee on Structural Safety. Probabilistic model code - Part 2 : Load models, 2001a.

Joint Committee on Structural Safety. Probabilistic model code - Part 3 : Resistance models, 2001b.

Jolliffe, I. *Principal component analysis*. Springer Verlag (Second Edition), 2002.

Kafali, C. and Grigoriu, M. Seismic fragility analysis: Application to simple linear and nonlinear systems. *Earthq. Eng. Struct. Dyn.*, 36(13):1885–1900, 2007.

Kahan, M., Gibert, R. J., and Bard, P. Y. Influence of seismic waves spatial variability on bridges : A sensitivity analysis. *Earthq. Eng. Struct. Dyn.*, 25(8):795–814, 1996.

Kanai, K. Semi-empirical formula for the seismic characteristics of the ground. Technical report, Earthquake Research Institute,University of Tokyo, 1957.

Karamlou, A. and Bocchini, P. Computation of bridge seismic fragility by large-scale simulation for probabilistic resilience analysis. *Earthq. Eng. Struct. Dyn.*, 44(12):1959–1978, 2015.

Kennedy, R. and Ravindra, M. Seismic fragilities for nuclear power plant risk studies. *Nuc. Eng. Des.*, 79(1):47–68, 1984.

Kennedy, R., Cornell, C., Campbell, R., Kaplan, S., and Perla, H. Probabilistic seismic safety study of an existing nuclear power plant. *Nuc. Eng. Des.*, 59(2):315–338, 1980.

Kewlani, G., Crawford, J., and Iagnemma, K. A polynomial chaos approach to the analysis of vehicle dynamics under uncertainty. *Vehicle Syst. Dyn.*, 50(5):749–774, 2012.

Konakli, K. and Der Kiureghian, A. Simulation of spatially varying ground motions including incoherence, wave-passage and differential site-response effects. *Earthq. Eng. Struct. Dyn.*, 41(3):495–513, 2012.

Kopsaftopoulos, F. and Fassois, S. A functional model based statistical time series method for vibration based damage detection, localization, and magnitude estimation. *Mech. Syst. Signal Pr.*, 39(1-2):143–161, 2013.

Kougioumtzoglou, I. A. and Spanos, P. D. An approximate approach for nonlinear system response determination under evolutionary stochastic excitation. *Current science*, 97(8): 1203–1211, 2009.

Kougioumtzoglou, I. A. and Spanos, P. D. Nonlinear MDOF system stochastic response determination via a dimension reduction approach. *Comput. Struct.*, 126:135–148, 2013.

Kraichnan, R. H. Direct-interaction approximation for a system of several interacting simple shear waves. *Phys. Fluids*, 6(1963):1603, 1963.

Kucherenko, S., Rodriguez-Fernandez, M., Pantelides, C., and Shah, N. Monte Carlo evaluation of derivative-based global sensitivity measures. *Reliab. Eng. Sys. Safety*, 94 (7):1135–1148, 2009.

Kumar, D. and Budman, H. Robust nonlinear MPC based on Volterra series and polynomial chaos expansions. *J. Process Contr.*, 24(1):304–317, 2014.

Kwon, O. S. and Elnashai, A. The effect of material and ground motion uncertainty on the seismic vulnerability curves of RC structure. *Eng. Struct.*, 28(2):289–303, 2006.

Lamboni, M., Iooss, B., Popelin, A.-L., and Gamboa, F. Derivative-based global sensitivity measures: general links with Sobol' indices and numerical tests. *Math. Comput. Simul.*, 87:44–54, 2013.

Le Gratiet, L., Iooss, B., Blatman, G., Browne, T., Cordeiro, S., and Goursaud, B. Model assisted probability of detection curves: new statistical tools and progressive methodology. *arXiv:1601.05914*, 2016.

Le Maître, O. and Knio, O. *Spectral Methods for Uncertainty Quantification.* Springer, 2010. 552 pages.

Le Maître, O. and Mathelin, L. Equation-free model reduction for complex dynamical systems. *Int. J. Numer. Methods Fluids*, 63(2):163–184, 2010.

Le Maître, O., Najm, H., Pébay, P., Ghanem, R., and Knio, O. Multi-resolution analysis scheme for uncertainty quantification in chemical systems. *SIAM J. Sci. Comput.*, 29 (2):864–889, 2007.

Le Maître, O., Mathelin, L., Knio, O., and Hussaini, M. Asynchronous time integration for polynomial chaos expansion of uncertain periodic dynamics. *Discret. Contin. Dyn. Sys. - Series A (DCDS-A)*, 28(1):199–226, 2010.

Le Maître, O., Knio, O., Najm, H., and Ghanem, R. A stochastic projection method for fluid flow – I. Basic formulation. *J. Comput. Phys.*, 173:481–511, 2001.

Le Meitour, J., Lucor, D., and Chassaing, J. C. Prediction of stochastic limit cycle oscillations using an adaptive polynomial chaos method. *J. Aero. Struct. Dyn.*, 2(1), 2010.

Lebrun, R. and Dutfoy, A. An innovating analysis of the Nataf transformation from the copula viewpoint. *Prob. Eng. Mech.*, 24(3):312–320, 2009.

Lee, T. H. and Mosalam, K. M. Probabilistic Seismic Evaluation of Reinforced Concrete Structural Components and Systems. Technical report, PEER, 2006.

Lemaire, M. *Structural reliability.* Wiley, 2009.

Leontaritis, I. J. and Billings, S. A. Input-output parametric models for non-linear systems. Part II: stochastic non-linear systems. *Int. J. Control*, 41(2):329–344, 1985.

Li, C. and Der Kiureghian, A. Mean out-crossing rate of nonlinear response to stochastic input. In Lemaire, M., Favre, J., and Mebarki, A., editors, *Proc. 7th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP7), Paris, France*, pages 1135–1141. Balkema, Rotterdam, 1995.

Li, R. and Ghanem, R. G. Adaptive polynomial chaos expansions applied to statistics of extremes in nonlinear random vibration. *Prob. Eng. Mech.*, 13(2):125–136, 1998.

Liel, A. B., Haselton, C. B., Deierlein, G. G., and Baker, J. W. Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. *Structural Safety*, 31(2):197–211, 2009.

Liu, P.-L. and Der Kiureghian, A. Multivariate distribution models with prescribed marginals and covariances. *Prob. Eng. Mech.*, 1(2):105–112, 1986.

Liu, W., Belytschko, T., and Mani, A. Probabilistic finite elements for non linear structural dynamics. *Comput. Methods Appl. Mech. Eng.*, 56:61–86, 1986.

Luchtenburg, D. M., Brunton, S. L., and Rowley, C. W. Long-time uncertainty propagation using generalized polynomial chaos and flow map composition. *J. Comp. Phys.*, 274:783–802, 2014.

Luco, N. and Bazzurro, P. Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses? *Earthq. Eng. Struct. Dyn.*, 36(13):1813–1835, 2007.

Lucor, D. and Karniadakis, G. Adaptive generalized polynomial chaos for nonlinear random oscillators. *SIAM J. Sci. Comput.*, 26(2):720–735, 2004.

Mack, C. A. Fifty years of Moore's law. *IEEE Transactions on Semiconductor Manufacturing*, 24(2):202–207, 2011.

Mackie, K. and Stojadinovic, B. Probabilistic seismic demand model for California highway bridges. *J. Bridge Eng.*, 6(6):468–481, 2001.

Mackie, K. and Stojadinovic, B. Seismic demands for performance-based design of bridges. Technical report, Pacific Earthquake Engineering Research Center, 2003.

Mackie, K. and Stojadinovic, B. *Fragility basis for California highway overpass bridge seismic decision making.* Pacific Earthquake Engineering Research Center, College of Engineering, University of California, Berkeley, 2005.

Mai, C. V. and Sudret, B. Polynomial chaos expansions for non-linear damped oscillators. In Haukaas, T., editor, *12th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP12)*, Vancouver, Canada, 2015a.

Mai, C. V. and Sudret, B. Polynomial chaos expansions for time-dependent problems. In *MascotNum*, Saint Étienne, France, 2015b.

Mai, C. V. and Sudret, B. Hierarchical adaptive polynomial chaos expansions. In Papadrakakis, M., Papadopoulos, V., and Stefanou, G., editors, *1st Int. Conf. on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP)*, Creta, Greece, 2015c.

Mai, C. V. and Sudret, B. Surrogate models for oscillatory systems using polynomial chaos expansions and stochastic time warping. *SIAM Int. J. Uncertainty Quantification (submitted)*, 2016a.

Mai, C. V. and Sudret, B. Mixing auto-regressive models and sparse polynomial chaos expansions for time-variant probems. In *Siam Conference on Uncertainty Quantification*, Lausanne, Switzerland, 2016b.

Mai, C. V., Sudret, B., Mackie, K., Stojadinovic, B., and Konakli, K. Non parametric fragility curves for bridges using recorded ground motions. In Cunha, A., Caetano, E., Ribeiro, P., and Müller, G., editors, *IX Int. Conf. on Structural Dynamics*, Porto, Portugal, 2014.

Mai, C. V., Konakli, K., and Sudret, B. Assessment of the lognormality assumption of seismic fragility curves using non-parametric representation. *Frontiers of Structural and Civil Engineering (submitted)*, 2016a.

Mai, C. V., Spiridonakos, M., Chatzi, E. N., and Sudret, B. Surrogate modelling for stochastic dynamical systems by combining NARX models and polynomial chaos expansions. *Int. J. Uncertainty Quantification (submitted)*, 2016b.

Mai, C. V., Spiridonakos, M. D., Chatzi, E. N., and Sudret, B. LARS-based ARX PCE metamodel for computing seismic fragility curves. In *Probabilistic mechanics and reliability conference*, Vanderbilt, USA, 2016c.

Marelli, S. and Sudret, B. UQLab: A framework for uncertainty quantification in Matlab. In *Vulnerability, Uncertainty, and Risk (Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom)*, pages 2554–2563, 2014a.

Marelli, S. and Sudret, B. UQLab: a framework for uncertainty quantification in MATLAB. *Advances in Engineering Software*, 2014b. in preparation.

Marelli, S. and Sudret, B. UQLab user manual – Polynomial chaos expansions. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, 2015. Report # UQLab-V0.9-104.

Marelli, S., Lamas, C., and Sudret, B. UQLab user manual – Sensitivity analysis. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, 2015. Report # UQLab-V0.9-106.

Matthies, H. and Keese, A. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194:1295–1331, 2005.

Mehdizadeh, M., Mackie, K. R., and Nielson, B. G. Scaling bias and record selection for fragility analysis. In *Proc. 15th World Conf. Earthquake Eng.*, 2012.

Melchers, R. Importance sampling in structural systems. *Structural Safety*, 6:3–10, 1989.

Mirza, S. A. and MacGregor, J. G. Variability of mechanical properties of reinforcing bars. *J. Struct. Div.*, 105(5):921–937, 1979.

Moore, G. Moore's law. *Electronics Magazine*, 38(8), 1965.

Morris, M. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.

Nelsen, R. *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2nd edition, 2006.

Nobile, F., Tempone, R., and Webster, C. A sparse grid stochastic collocation method for elliptic partial differential equations with random input data. Technical Report MOX Report 85, Politecnico di Milano, 2006.

Nouy, A. Identification of multi-modal random variables through mixtures of polynomial chaos expansions. *Comptes Rendus Mécanique*, 338(12):698–703, 2010.

Oakley, J. and O'Hagan, A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. *J. Royal Stat. Soc., Series B*, 66:751–769, 2004.

Orabi, I. I. and Ahmadi, G. Nonstationary response analysis of a Duffing oscillator by the Wiener-Hermite expansion method. *J. Appl. Mech.*, 54(2):434, 1987.

Orszag, S. A. Dynamical properties of truncated Wiener-Hermite expansions. *Phys. Fluids*, 10(12):2603, 1967.

Owen, A. A central limit theorem for Latin hypercube sampling. *J. Royal Stat. Soc., Series B*, 54:541–551, 1992.

Ozen, H. C. and Bal, G. Dynamical polynomial chaos expansions and long time evolution of differential equations with random forcing. *SIAM/ASA J. Uncertainty Quantification*, 4(1):609–635, 2016.

Pacific Earthquake Engineering and Research Center. *OpenSees: The Open System for Earthquake Engineering Simulation*, 2004.

Padgett, J., Nielson, B., and DesRoches, R. Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earthq. Eng. Struct. Dyn.*, 37(5):711–725, 2008.

Paffrath, M. and Wever, U. Adapted polynomial chaos expansion for failure detection. *J. Comput. Phys*, 226(1):263–281, 2007.

Papadrakakis, M. and Papadopoulos, V. Robust and efficient methods for stochastic finite element analysis using Monte Carlo simulation. *Comput. Methods Appl. Mech. Eng.*, 134(3-4):325–340, 1996.

Pei, S. and Van De Lindt, J. Methodology for earthquake-induced loss estimation: An application to woodframe buildings. *Structural Safety*, 31(1):31–42, 2009.

Peng, J., Hampton, J., and Doostan, A. On polynomial chaos expansion via gradient-enhanced l_1-minimization. *J. Comp. Phys.*, 310:440–458, 2016.

Peraire, J. and Widnall, S. 3d rigid body dynamics: equations of motion. *MIT OpenCourseWare*, 2009.

Pettit, C. and Beran, P. Spectral and multiresolution Wiener expansions of oscillatory stochastic processes. *J. Sound. Vib.*, 294(4-5):752–779, 2006.

Pichler, L., Pradlwarter, H. J., and Schuëller, G. I. A mode-based meta-model for the frequency response functions of uncertain structural systems. *Comput. Struct.*, 87(5-6): 332–341, 2009.

Pickard Lowe and Garrick Inc. Zion Probabilistic Safety Study, prepared for Commonwealth Edison Company. Technical report, Westing house Electric Corporation, and Fauske & Associates LLC, Chicago, 1981.

Piroddi, L. Simulation error minimisation methods for NARX model identification. *Int. J. Model. Ident. Control*, 3(4):392–403, 2008.

Pitarka, A. and Irikura, K. Three-dimensional simulation of the near-fault ground motion for the 1995 Hyogo-ken Nanbu (Kobe), Japan, earthquake. *Bulletin of the . . .*, 1998.

Porter, K. A., Beck, J. L., and Shaikhutdinov, R. V. Sensitivity of building loss estimates to major uncertain variables. *Earthquake Spectra*, 18(4):719–743, 2002.

Poulimenos, A. G. and Fassois, S. D. Parametric time-domain methods for non-stationary random vibration modelling and analysis - A critical survey and comparison. *Mech. Syst. Signal Pr.*, 20(4):763–816, 2006.

Pousse, G., Bonilla, L. F., Cotton, F., and Margerin, L. Nonstationary stochastic simulation of strong ground motion time histories including natural variability: application to the K-Net Japanese database. *Bull. Seismol. Soc. Am.*, 96(6):2103–2117, 2006.

Proppe, C., Pradlwarter, H., and Schuëller, G. Equivalent linearization and Monte Carlo simulation in stochastic dynamics. *Prob. Eng. Mech.*, 18(1):1–15, 2003.

Pujol, G., Iooss, B., and Janon, A. *sensitivity: Sensitivity Analysis*, 2013. URL `http://cran.r-project.org/package=sensitivity`.

Quilligan, A., O Connor, A., and Pakrashi, V. Fragility analysis of steel and concrete wind turbine towers. *Eng. Struct.*, 36:270–282, 2012.

Rajeev, P. and Tesfamariam, S. Seismic fragilities for reinforced concrete buildings with consideration of irregularities. *Structural Safety*, 39:1–13, 2012.

Ramsay, J. O. and Li, X. Curve registration. *J. Roy. Stat. Soc. B.*, 60(2):351–363, 1998.

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

Rezaeian, S. and Der Kiureghian, A. Stochastic modeling and simulation of ground motions for performance-based earthquake engineering. Technical report, Pacific Earthquake Engineering Research Center, 2010b.

Richardson, J. E., Bagchi, G., and Brazee, R. J. The seismic safety margins research program of the U.S. Nuclear Regulatory Commission. *Nuc. Eng. Des.*, 59(1):15–25, 1980.

Rosenblatt, M. Remarks on a multivariate transformation. *Ann. Math. Stat.*, 23:470–472, 1952.

Rosenblueth, E. Probabilistic design to resist earthquakes. *J. Eng. Mech. Div.-ASCE*, 90(5):189–220, 1964.

Rossetto, T. and Elnashai, A. Derivation of vulnerability functions for European-type RC structures based on observational data. *Engi. Struct.*, 25(10):1241–1263, 2003.

Rubinstein, R.-Y. *Simulation and the Monte Carlo methods.* John Wiley & Sons, 1981.

Saha, S. K., Sepahvand, K., Matsagar, V. A., Jain, A. K., and Marburg, S. Stochastic analysis of base-isolated liquid storage tanks with uncertain isolator parameters under random excitation. *Eng. Struct.*, 57(0):465–474, 2013.

Sakellariou, J. S. and Fassois, S. D. Functionally pooled models for the global identification of stochastic systems under different pseudo-static operating conditions. *Mech. Syst. Signal Pr.*, 72-73:785–807, 2016.

Sakoe, H. and Chiba, S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. Acoust. Speech Signal Process.*, 26(1):43–49, 1978.

Saltelli, A., Chan, K., and Scott, E., editors. *Sensitivity analysis.* J. Wiley & Sons, 2000.

Saltelli, A., Tarentola, S., Campolongo, F., and Ratto, M. *Sensitivity analysis in practice – A guide to assessing scientific models.* J. Wiley & Sons, 2004.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. *Global Sensitivity Analysis – The Primer.* Wiley, 2008.

Saltelli, A., Annoni, P., Azzini, V., Campolongo, F., Ratto, M., and Tarantola, S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput. Phys. Comm.*, 181:259–270, 2010.

Samara, P. A., Sakellariou, J. S., Fouskitakis, G. N., Hios, J. D., and Fassois, S. D. Aircraft virtual sensor design via a time-dependent functional pooling NARX methodology. *Aerosp. Sci. Technol.*, 29(1):114–124, 2013.

Sapsis, T. P. and Lermusiaux, P. F. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D: Nonlinear Phenomena*, 238(23-24): 2347–2360, 2009.

Sarrouy, E., Dessombz, O., and Sinou, J.-J. Stochastic study of a non-linear self-excited system with friction. *Eur. J. Mech. A. Solids*, 40:1–10, 2013.

Schick, M., Heuveline, V., and Le Maître, O. P. A Newton–Galerkin method for fluid flow exhibiting uncertain periodic dynamics. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):153–173, 2014.

Schuëller, G. I. Computational stochastic mechanics - recent advances. *Comput. Struct.*, 79(22–25):2225–2234, 2001.

SEAOC Vision 2000 Committee. Performance-based Seismic Engineering, 1995.

Seo, J., Duenas-Osorio, L., Craig, J. I., and Goodno, B. J. Metamodel-based regional vulnerability estimate of irregular steel moment-frame structures subjected to earthquake events. *Eng. Struct.*, 45:585–597, 2012.

Seyedi, D. M., Gehl, P., Douglas, J., Davenne, L., Mezher, N., and Ghavamian, S. Development of seismic fragility surfaces for reinforced concrete buildings by means of nonlinear time-history analysis. *Earthq. Eng. Struct. Dyn.*, 39(1):91–108, 2010.

Shinozuka, M. Monte Carlo solution of structural dynamics. *Comput. Struct.*, 2(5-6): 855–874, 1972.

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

Shome, N., Cornell, C. A., Bazzurro, P., and Carballo, J. E. Earthquakes, Records, and Nonlinear Responses. *Earthquake Spectra*, 14(3):469–500, 1998.

Simpson, T. W., Poplinski, J. D., Koch, P. N., and Allen, J. K. Metamodels for computer-based engineering design: survey and recommendations. *Eng. Comput.*, 17(2):129–150, 2001.

Sobol', I. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.*, 55(1-3):271–280, 2001.

Sobol', I. and Kucherenko, S. Derivative based global sensitivity measures and their link with global sensitivity indices. *Math. Comput. Simul.*, 79(10):3009–3017, 2009.

Soize, C. Polynomial chaos expansion of a multimodal random vector. *SIAM/ASA J. Uncertainty Quantification*, 3(1):34–60, 2015.

Soize, C. and Ghanem, R. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM J. Sci. Comput.*, 26(2):395–410, 2004.

Spinelli, W., Piroddi, L., and Lovera, M. A two-stage algorithm for structure identification of polynomial NARX models. In *2006 American Control Conference*, page 6 pp. IEEE, 2006.

Spiridonakos, M., Chatzi, E., and Sudret, B. Polynomial chaos expansion models for the monitoring of structures under operational variability. *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng.*, 2016. Accepted.

Spiridonakos, M. D. and Chatzi, E. N. Metamodeling of structural systems through polynomial chaos ARX models. In *4th International Conference on Uncertainty in Structural Dynamics (USD), Leuven, Belgium*, 2012.

Spiridonakos, M. D. and Chatzi, E. N. Metamodeling of structural systems with parametric uncertainty subject to stochastic excitation. In Deodatis, G., editor, *Proc. 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2013), New York, USA*, 2013.

Spiridonakos, M. D. and Chatzi, E. N. Metamodeling of dynamic nonlinear structural systems through polynomial chaos NARX models. *Comput. Struct.*, 157:99–113, 2015a.

Spiridonakos, M. D. and Chatzi, E. N. Metamodeling of nonlinear structural systems with parametric uncertainty subject to stochastic dynamic excitation. *Earthquakes and Structures*, 8(4):915–934, 2015b.

Spiridonakos, M. D., Giagopoulos, D., Chatzi, E. N., and Dertimanis, V. Finite element metamodeling for nonlinear systems with uncertain properties. In *EMI 2015 - Engineering Mechanics Institute Conference*, Stanford University, 2015.

Stefanou, G. The stochastic finite element method: Past, present and future. *Comput. Methods Appl. Mech. Engrg.*, 198:1031–1051, 2009.

Sudret, B. Global sensitivity analysis using polynomial chaos expansions. In Spanos, P. and Deodatis, G., editors, *Proc. 5th Int. Conf. on Comp. Stoch. Mech (CSM5), Rhodos, Greece*, 2006.

Sudret, B. *Uncertainty propagation and sensitivity analysis in mechanical models – Contributions to structural reliability and stochastic spectral methods.* Université Blaise Pascal, Clermont-Ferrand, France, 2007. Habilitation à diriger des recherches, 173 pages.

Sudret, B. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Sys. Safety*, 93:964–979, 2008.

Sudret, B. *Polynomial chaos expansions and stochastic finite element methods*, chapter 6, pages 265–300. Risk and Reliability in Geotechnical Engineering. Taylor and Francis, 2015.

Sudret, B. and Mai, C. V. Computing seismic fragility curves using polynomial chaos expansions. In Deodatis, G., editor, *Proc. 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2013), New York, USA*, 2013a.

Sudret, B. and Mai, C. V. Calcul des courbes de fragilité par approches non-paramétriques. In *Proc. 21ème Congrès Français de Mécanique (CFM21)*, Bordeaux, France, 2013b.

Sudret, B. and Mai, C. Derivative-based sensitivity indices based on polynomial chaos expansions. In *Proc. 7th Int. Conf. Sensitivity Anal. Model Output (SAMO'2013), Nice, France*, 2013c.

Sudret, B. and Mai, C. Computing derivative-based global sensitivity measures using polynomial chaos expansions. *Reliab. Eng. Sys. Safety*, 134:241–250, 2015.

Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. Royal Stat. Soc., Series B*, 58:267–288, 1996.

Towashiraporn, P., Duenas-Osorio, L., Craig, J. I., and Goodno, B. J. An application of the response surface metamodel in building seismic fragility estimation. In *Proc. 14th World Conf. Earthquake Eng.*, pages 12–17, 2008.

Vamvatsikos, D. and Cornell, C. A. Incremental dynamic analysis. *Earthq. Eng. Struct. Dyn.*, 31(3):491–514, 2002.

Wagner, S. M. and Ferris, J. B. A polynomial chaos approach to ARIMA modeling and terrain characterization. In Schum, K. and Trevisani, D. A., editors, *SPIE 6564, Modeling and Simulation for Military Operations II*, pages 65640M 1–10, 2007.

Wall, F. and Bucher, C. Sensitivity of expected exceedance rate of SDOF-system response to statistical uncertainties of loading and system parameters. *Prob. Eng. Mech.*, 2(3):138–146, 1987.

Wan, X. and Karniadakis, G. Long-term behavior of polynomial chaos in stochastic flow simulations. *Comput. Methods Appl. Mech. Eng.*, 195:5582–5596, 2006.

Wan, X. and Karniadakis, G. E. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J. Comput. Phys.*, 209:617–642, 2005.

Wand, M. and Jones, M. C. *Kernel smoothing.* Chapman and Hall, 1995.

Wang, K. and Gasser, T. Alignment of curves by dynamic time warping. *Ann. Stat.*, 25 (3):1251–1276, 1997.

Wei, H. L. and Billings, S. A. Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information. *Int. J. Model. Ident. Control*, 3(4):341–356, 2008.

Witteveen, J. A. and Bijl, H. An alternative unsteady adaptive stochastic finite elements formulation based on interpolation at constant phase. *Comput. Methods Appl. Mech. Eng.*, 198(3):578–591, 2008.

Wong, F.-A. Slope reliability and response surface method. *J. Geotech.*, 111(1):32–53, 1985.

Xiu, D. and Hesthaven, J. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.

Xiu, D. and Karniadakis, G. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.

Yaghoubi, V., Marelli, S., Sudret, B., and Abrahamsson, T. Sparse polynomial chaos expansions of frequency response functions using stochastic frequency transformation. *submitted*, 2016.

Yamamoto, Y. and Baker, J. W. Stochastic model for earthquake ground motion using wavelet packets. *Bull. Seismol. Soc. Am.*, 103(6):3044–3056, 2013.

Yang, J., Faverjon, B., Peters, H., and Kessissoglou, N. Application of polynomial chaos expansion and model order reduction for dynamic analysis of structures with uncertainties. *Procedia IUTAM*, 13:63–70, 2015.

Yang, X., Lei, H., Baker, N. A., and Lin, G. Enhancing sparsity of Hermite polynomial expansions by iterative rotations. *J. Comp. Phys.*, 307:94–109, 2016.

Yuan, M., Joseph, V. R., and Lin, Y. An efficient variable selection approach for analyzing designed experiments. *Technometrics*, 49(4):430–439, 2007.

Zareian, F. and Krawinkler, H. Assessment of probability of collapse and design for collapse safety. *Earthq. Eng. Struct. Dyn.*, 36(13):1901–1914, 2007.

Zentner, I. Numerical computation of fragility curves for NPP equipment. *Nuc. Eng. Des*, 240:1614–1621, 2010.

Zentner, I., Humbert, N., Ravet, S., and Viallet, E. Numerical methods for seismic fragility analysis of structures and components in nuclear industry - Application to a reactor coolant system. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 5(2):99–109, 2011.

Zhang, J. and Ellingwood, B. Effects of uncertain material properties on structural stability. *J. Struct. Eng.*, 121(4):705–716, 1995.

Zhang, L. and Li, K. Forward and backward least angle regression for nonlinear system identification. *Automatica*, 53:94–102, 2015.

Zhang, Y. and Der Kiureghian, A. Finite element reliability methods for inelastic structures. Technical Report UCB/SEMM-97/05, University of California, Berkeley, 1997.