Surrogate models for uncertainty quantification in the context of imprecise probability modelling

Report

Author(s):
Schöbi, Roland

Publication date:
2019-09

Permanent link:
https://doi.org/10.3929/ethz-b-000373587

Originally published in:
IBK Bericht 505

Institut für Baustatik und Konstruktion, ETH Zürich

Surrogate models for uncertainty quantification in the context of imprecise probability modelling

Roland SCHÖBI

IBK Bericht Nr. 505, September 2019

KEYWORDS: Uncertainty quantification, structural reliability, Gaussian process modelling, polynomial chaos expansions, PC-Kriging, aleatory/epistemic uncertainty, probability boxes, geomechanics

Roland SCHÖBI:
Surrogate models for uncertainty quantification in the context of imprecise probability modelling

Bericht IBK Nr. 505, September 2019

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.

SURROGATE MODELS FOR UNCERTAINTY QUANTIFICATION IN THE CONTEXT OF IMPRECISE PROBABILITY MODELLING

Roland SCHÖBI

Institute of Structural Engineering
ETH Zurich

September 2019

Foreword

Computer simulation is used nowadays in virtually all civil and mechanical engineering applications, either for designing or operating complex systems. However, the analyst rarely has a complete knowledge of the parameters of the system under consideration, either because they are naturally variable (aleatory uncertainty, such as environmental conditions) or because of lack of knowledge (epistemic uncertainty). Modelling the sources of uncertainty and then propagate them through the computational model to estimate statistical properties of the predictions is becoming an ubiquitous task for the engineer.

In this report the two types of uncertainties are considered simultaneously, in the context where a probabilistic model is built from scarce data. The formalism of choice for this representation is that of "imprecise probability theory", i.e. p-boxes, which are a natural extension of the well-known cumulative distribution function of an uncertain parameter. P-boxes allow the analyst to conveniently represent the two sources of uncertainty as a whole.

This research report was written as a PhD thesis (ETH Dissertation Nr. 24078) by Roland SCHÖBI and presents original methods and algorithms to carry out uncertainty propagation, sensitivity analysis and structural reliability in the context of imprecise probability modelling. The applicability of these approaches has been demonstrated in realistic case studies from geotechnical and aerospace engineering. I would like to thank Roland SCHÖBI for his high commitment all along this research project, and the outstanding quality of his contributions.

Zurich, September 2019
Bruno Sudret

Summary

Nowadays, computer simulations are a popular engineering tool to design systems of even increasing complexity and to determine their performance. Theses simulations aim at reproducing the physical process at hand and hence provide a solution to the underlying governing equations. As an example, finite element models have become a standard tool in modern civil and mechanical engineering. Such models exploit the available computer power, meaning that a single run of the model can take up to hours or days of computing.

At the same time, uncertainties are intrinsic to theses processes and simulation methods. Typically model parameters are not known deterministically, but inferred from data and then modelled probabilistically. However, a common situation in practice is to have a limited budget for data acquisition and hence to end up with sparse datasets. This introduces epistemic uncertainty (lack of knowledge) alongside aleatory uncertainty (natural variability). The mix of these two sources of uncertainties is often referred to as imprecise probabilities.

Among the various concepts to describe imprecise probabilities, there are probability-boxes (p-boxes), which model the uncertainty by an interval-valued cumulative distribution function. Hence, p-boxes are a generalization of the popular probability theory. P-boxes provide an intuitive framework that is more general than conventional probability theory and that allows for a clear distinction between aleatory and epistemic uncertainty. Due to its intuitive setup, p-boxes are easy to interpret and hence easy to apply to practical problem settings.

In this context, engineers are concerned with uncertainty quantification. In other words, engineers analyse how the uncertainty in the model parameters affects the uncertainty in the quantity of interest. In this sense, uncertainty quantification includes uncertainty propagation, structural reliability analysis, sensitivity analysis, and design optimization.

Due to the complexity of the computer simulations and the use of p-boxes, however, these analyses may become intractable. In order to reduce the computational costs and to make uncertainty quantification analyses tractable, meta-models are used throughout this thesis. Kriging (a.k.a Gaussian process modelling) and polynomial chaos expansions (PCE) are two state-of-the-art meta-modelling algorithms which approximate the complex computer simulations with an easy-to-evaluate function. Furthermore, the proposed Polynomial-Chaos-Kriging approach further increases the accuracy of the approximation when comparing to Kriging and PCE.

Depending on the uncertainty quantification analysis, the meta-modelling approaches are modified to fit the needs of the specific analysis in the context of p-boxes. For uncertainty propagation and structural reliability analysis, two-level meta-modelling approaches are proposed. The use of meta-models at different stages of the analysis allows for an efficient, *i.e.* accurate and inexpensive, estimation of the quantity of interest. Emphasis is taken on the limitation of computational resources, *i.e.* of the number of runs of the complex computer simulations.

The approaches are validated on a number of benchmark function, ranging from purely analytical function to finite element models. Further, case studies from the field of aeronautics and geotechnical engineering show the applicability of the proposed algorithm in realistic and complex engineering settings. The variety of examples shows the flexibility and versatility of the proposed algorithms. Hence, the proposed approaches are of importance for the engineering practice when the standard probabilistic approach fails to characterize the uncertainty.

Zusammenfassung

Heutzutage sind Computersimulationen ein beliebtes Hilfsmittel für Ingenieure um immer komplexer werdende Systeme zu untersuchen. Diese Simulationen beabsichtigen es physikalische Prozesse abzubilden und bieten eine Lösung zu den zugrundeliegenden Gleichungen. Zum Beispiel ist das Finite-Elemente-Model zu einem Standardwerkzeug des modernen Bau- und Maschinenbauingenieurs geworden. Diese Modelle nutzen die verfügbaren Computerkapazitäten, so dass eine einzelne Ausführung des Modelles Stunden bis Tage beanspruchen kann.

Gleichzeitig sind Unsicherheiten in diesen Modellen und Prozessen allgegenwärtig. In vielen Fällen sind die Modelleingangsgrössen nicht bekannt, werden von Datensätzen abgeleitet und werden anschliessend mit einem probabilistischen Ansatz modelliert. Jedoch kommt es in der Praxis häufig vor, dass das Kontingent für Datensammlungsaktionen begrenzt ist und die Datensätze darum klein sind. In diesem Fall treten neben aleatorischen Unsicherheiten (natürliche Variabilität) auch epistemische Unsicherheiten (Wissenslücken) auf. Die Mischung dieser beiden Unsicherheitsquellen wird in der Literatur mit dem Term unpräzise Wahrscheinlichkeiten (englisch: imprecise probabilities) beschrieben.

Unter den zahlreichen Konzepten zur Beschreibung von unpräzisen Wahrscheinlichkeiten sind die Wahrscheinlichkeitsboxen (Englisch: p-boxes) zu finden. Wahrscheinlichkeitsboxen beschreiben die Unsicherheit mit einer intervallwertigen kumulativen Verteilungsfunktion (CDF). Darum sind die Wahrscheinlichkeitsboxen eine Verallgemeinerung der beliebten Wahrscheinlichkeitstheorie. Wahrscheinlichkeitsboxen bieten einen intuitiven Ansatz, welcher allgemeiner als die klassische Wahrscheinlichkeitstheorie ist und welcher eine klare Unterscheidung zwischen aleatorischen und epistemischen Unsicherheiten erlaubt. Wegen des intuitiven Ansatzes sind Wahrscheinlichkeitsboxen einfach zu interpretieren und darum relevant für praktische Problemstellungen.

In diesem Kontext befassen sich Ingenieure mit der Quantifizierung von Unsicherheiten. Ingenieure analysieren, wie die Unsicherheit in den Eingangsgrössen die Unsicherheit in den Ausgangsgrössen beeinflusst. Die Quantifizierung von Unsicherheiten umfasst unter anderem die Übertragungen von Unsicherheiten, Verlässlichkeitsanalysen, Sensitivitätsanalysen und Entwurfsoptimierungen.

Wegen der Komplexität der Computersimulationen und der Verwendung von Wahrscheinlichkeitsboxen sind diese Analysen allerdings oft schwierig zu bewältigen. Metamodelle werden in dieser Doktorarbeit verwendet um die hohen Computeraufwände zu reduzieren und um die Analysen zur Quantifizierung von Unsicherheiten schliesslich zu ermöglichen. Kriging (im Englischen auch bekannt als "Gaussian process modelling") und "Polynomial Chaos Expansions" (PCE) sind zwei Metamodellierungsansätze auf dem neusten Stand der Technik, welche die komplexen Computersimulationen mit einer simplen Funktion annähern. Ausserdem ermöglicht die vorgeschlagene Polynomial-Chaos-Kriging-Methode die Genauigkeit solcher Annäherungen im

Vergleich zu Kriging und PCE weiter zu steigern.

Diese Metamodellmethoden werden so angepasst, dass sie den spezifischen Charakteristiken der verschiedenen Analysen zur Quantifizierung von Unsicherheiten entsprechen. Für die Übertragung von Unsicherheiten und die Verlässlichkeitsanalyse werden zweistufige Metamodellierungsalgorithmen vorgeschlagen. Die Verwendung von Metamodellen an verschiedenen Stellen der Analyse erlaubt eine effiziente, *i.e.* präzise und günstige, Abschätzung der Ausgangsgrösse. Der Fokus steht dabei auf der Begrenzung von Computerressourcen, insbesondere der Anzahl Computersimulationen.

Die vorgeschlagenen Algorithmen sind an einer Vielzahl von Benchmarkbeispielen validiert, einschliesslich rein analytischen Funktionen bis Finite-Element-Modelle. Ausserdem werden Fallbeispiele aus der Aeronautik und der Geotechnik diskutiert um die Anwendbarkeit der vorgeschlagenen Algorithmen in realistischen und komplexen Ingenieursproblemstellungen zu erörtern. Die Vielfalt der Anwendungen zeigt die Flexibilität und Vielseitigkeit der vorgeschlagenen Algorithmen. Darum sind die vorgeschlagenen Algorithmen relevant für die Ingenieurspraxis, insbesondere wenn der wahrscheinlichkeitsbasierte Standardansatz versagt.

Acknowledgement

My sincere gratitude goes to Prof. Bruno Sudret, which hired me in 2013 after my graduation as a Civil Engineer and a one-year position in Prof. Eleni Chatzi's group at ETH Zurich. His expertise on uncertainty quantification allowed me to quickly learn the essentials of this research field and gave me ideal guidelines for my research and employment at the Chair of Risk, Safety and Uncertainty Quantification. Furthermore, Prof. Bruno Sudret encouraged me to discover new fields of research and helped me to broaden my horizon by attending a number of events such as international conferences. Being assistant of the lecture "Structural safety and risk analysis" allowed me to finally disseminate my knowledge to the next generation of engineers and researchers.

Dr. Stefano Marelli was my officemate for the first part of my phd which I am grateful for. His technical advice helped me to evolve my approach to research, scientific writing as well as coding. He spent countless hours on educating me in LaTeX and Matlab. Through the software framework UQLab, I was able to work on a software development project that included my personal research results and that involved the subsequent publication as open-source software. I am thankful for this experience that involved developing pieces of code in collaboration with other group members.

My thank includes my two co-examiners Prof. Michael Beer and Prof. Edoardo Patelli. Their questions and comments helped me to improve my phd manuscript and finally to graduate successfully my phd program. Furthermore, the scientific discussions during past conferences and workshops extended my knowledge on uncertainty quantification, especially in the field of imprecise probabilities.

Throughout my phd at the Chair of Risk, Safety and Uncertainty Quantification, I had the pleasure to work with a very friendly group of fellow students and employees. In particular, I thank Joseph for the many discussions on Bayesian versus non-Bayesian approaches. Moreover, I am thankful for the numerous scientific and non-scientific discussions with Christos, Chu, and Katerina during the last four years. Additionally, I always appreciated the discussions with other members of Prof. Bruno Sudret's Chair as well as members of the Chair of Structural Mechanics (Prof. Eleni Chatzi) and the Chair of Structural Dynamics and Earthquake Engineering (Prof. Bozidar Stojadinovic).

Last but not least, I thank my family and friends for the countless support that I obtained throughout my studies and particularly during my phd. Without them, this thesis would not have been possible.

Contents

CHAPTER 1

Introduction

1.1 Context

In modern engineering, mathematical models are used to represent physical and engineering processes and systems. The application fields are various, including the modelling of subsurface flow of contaminated water through different layers of soil (Deman et al. 2016), slope stability under spatially varying soil parameters (Sudret & Schöbi 2017), and the deflection of a frame structure subjected to horizontal wind loads (Schöbi & Sudret 2016b). The mathematical model aims at reproducing the behaviour of the physical system for a given set of parameters. As an example, the soil model may predict the stability of a newly built slope. Moreover, the set of parameters may include the physical and mechanical properties of soil layers (*e.g.* Young's modulus or internal friction angle), the loading at the initial state, and the degradation behaviour over time.

Due to the diversity of the applications and the sophistication of the analyses, the complexity of the mathematical representation may range from simple analytical functions to complex models (*e.g.* partial differential equations, continuum mechanics, or dynamics). Often, complex mathematical models have no closed form solution and rely on solvers such as finite element or finite difference schemes. A large number of commercial and non-commercial software packages have been developed in the past decade intending to solve these finite element models.

Despite the differences in application field and complexity, mathematical models have one feature in common: they represent simplifications of the reality. The simplification can be characterized by two aspects: (i) limited applicability and (ii) limited accuracy. The mathematical model typically represents a limited system with boundary conditions rather than the entire universe. In other words, the mathematical models represent a small aspect of reality and, hence, have limited applicability. As an example, the model for slope stability is not suitable for analysing the deflection of the frame structure. The limited accuracy of the mathematical model is coupled to the definition of the target system. Due to the limiting boundaries of the system under consideration, some influence factors are neglected in the mathematical model. This naturally introduces discrepancies between the behaviour of the physical system and the mathematical model. Hence, the mathematical model generally approximates the real behaviour at a certain level of accuracy. As an examples, the soil model could use a linear elastic constitutive model whereas reality is different.

In the context of this discussion, the famous quote ”essentially all models are wrong but some are useful” by Box & Draper (1987) becomes meaningful. Among all possible models available to the modeller, the choice of an appropriate model is crucial for the success of the analysis. The modeller aims at describing the physical system by defining an appropriate subsystem, boundary conditions, set of parameters, and (potentially) a solver.

When having identified a suitable mathematical model, different sources of inaccuracies may occur. As the mathematical model represents a simplification of reality, it entails an intrinsic model error. Furthermore, modelling complex systems requires numerical solvers, as mentioned above, which may introduce numerical and discretization errors. Last but not least, the parameters of the mathematical model might not be known perfectly. The different types of inaccuracies might or might not be present in a given problem setting. However, they may interact with each other. Hence, an important task is to quantify the errors in order to be able to interpret the results.

The inaccuracies and errors are generally modelled by uncertainty models/concepts. In order to do so, two main sources of uncertainty are identified in the literature: aleatory and epistemic uncertainty. *Aleatory* (”alea” is Latin for ”dice”) uncertainty describes the natural/intrinsic variability of a quantity of interest and is hence non-reducible. An example of pure aleatory uncertainty is a fair dice, where the outcome of a dice throw varies naturally between one to six and cannot be predicted. *Epistemic* (”$\epsilon\pi\iota\sigma\tau\eta\mu\eta$” is Greek for ”knowledge/science”) uncertainty, on the other hand, describes the lack of knowledge and is potentially reducible by acquiring more knowledge. An example of pure epistemic uncertainty is the measurement of the height of an existing building. Due to the fact that the building has been constructed, there exists a real but unknown value for the building height. The height might be measured by *e.g.* eyesight, measuring tape, GPS, or theodolite. However, using different measurement devices will lead to different values. The theodolite will likely give a more precise value than measuring with your own eyes. In practical applications, the two sources of uncertainty can appear in the same problem setting and therefore should be considered as such by the modeller.

Having defined a proper mathematical model and having quantified the corresponding uncertainties, the modeller becomes a decision maker. Ultimately, the mathematical model is used to understand the physical process and make some related decisions, which can be manifold. In the slope stability example, the model might be used to estimate whether the constructed slope is reliable in the long term. In the case of the frame structure, the model might indicate whether the building would be safe in case of a hurricane and can hence be built or not.

1.2 General framework for uncertainty quantification

Uncertainty quantification (UQ) is a general term to summarize the various tasks and challenges described previously. Following the developments by de Rocquigny (2006a,b), Sudret (2007), and the mindset of the Chair of Risk, Safety and Uncertainty Quantification at ETH Zurich, a general framework for uncertainty quantification is presented here and frames the developments in this thesis.

A summary of the main elements is sketched in Figure 1.1. Each box represents a different task (grey area) as well as a number of examples (white area). The basic steps are presented in depth in the following.

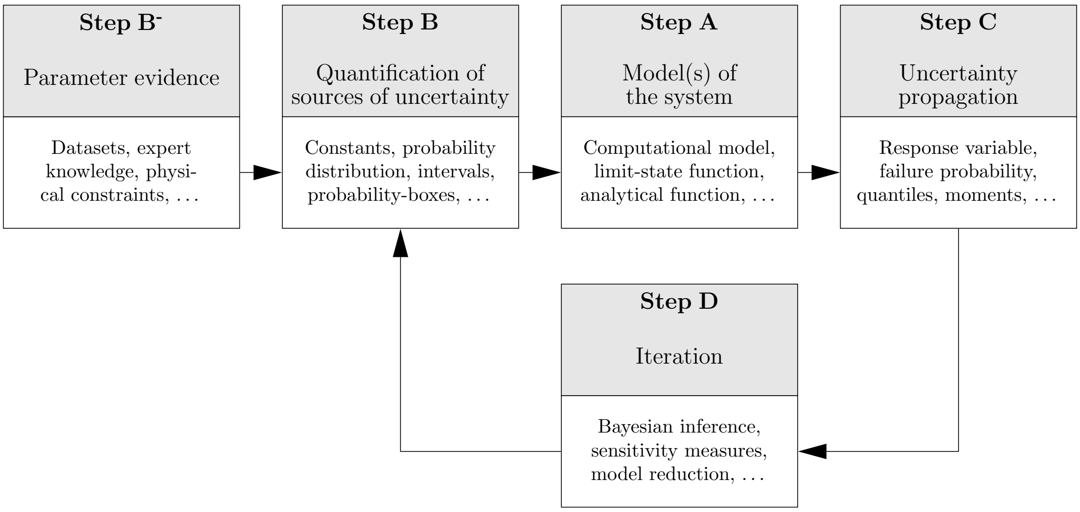


Figure 1.1: Uncertainty quantification framework.

- **Step A**: in a first step, the computational model representing the physical system or process is defined. It may consist of an analytical function in the simplest case. However, the computational model may also consist of an entire workflow containing different pieces of software or even physical experiments (as in *e.g.* hybrid modelling, which combines physical experiments and finite element models (Abbiati et al. 2017)). As a summary, the computational model maps a set of input parameters to a quantity of interest (QoI) that is further used for decision making.

- **Step B**: the quantification of sources of uncertainty aims at identifying input parameters and modelling them appropriately. A variety of modelling choices are available, including constants (*i.e.* no uncertainty), probability distributions, intervals, and imprecise probabilities. Hence, the task of the modeller is to choose a suitable model and calibrate it for each input parameter. The suitability of an uncertainty model may be defined through various factors, such as the available information on the input parameters, general knowledge on the system, as well as the purpose of the analysis.

- **Step B$^-$**: the definition of the uncertainty model in Step B is based on pieces of knowledge, which can be manifold. In general, however, the knowledge on a given parameter consists of two types of information: (i) datasets and (ii) expert knowledge. Datasets are collections of data obtained by field investigations. The term expert knowledge/judgement includes experiences made in similar previous projects as well as logical constraints, among other components.

- **Step C**: uncertainty propagation analyses how the uncertainty in the input parameters is transformed through the computational model towards the QoI. When the input is modelled by uncertain parameters, the QoI is likely to be uncertain, too. Hence, Step C analyses different statistics of interest, depending on the type of question at hand. Statistics of interest include mean values, higher-order moments, failure probabilities, quantiles, and distributions.

- **Step D**: the iteration step connects the response QoI to the input parameters. There are two main frameworks mentioned here to do so: (i) sensitivity analysis and (ii) Bayesian inference. Sensitivity analysis measures the impact of each uncertain input parameter onto the uncertainty in the QoI. This can be seen as a post-processing technique to Step C. Bayesian inference refines the uncertainty model in the input parameters as a function of the uncertainty in the QoI and potentially additional data.

As it is indicated in Figure 1.1, the uncertainty quantification analysis is an iterative process. In fact, a typical engineering workflow includes a continuous refinement of the uncertainty model based on additional knowledge (*i.e.* datasets, observations). This iterative learning process allows for an improved prediction accuracy with time.

1.3 Problem statement

In the general framework for uncertainty quantification, a number of challenges appear. Starting in Step A, consider a computational model that represents a physical system well enough for the purpose of uncertainty quantification (*i.e.* numerical and discretization error are neglected). Note that, as a simplification in this thesis, the computational model is considered to be deterministic, *i.e.* the repeated evaluation with the same set of input parameters leads to the same model response. In practice this assumption is not always fulfilled. However, the computational model may be complex and expensive-to-evaluate. Considering the subsurface flow example, a finite element model simulation may take minutes to days to converge to a stable solution. Furthermore, the computational model is often a black box, *i.e.* only the input parameters and the response value are observable. The internal machinery of the computational model is not necessarily accessible, as it is typically found when handling legacy codes. Hence, only non-intrusive analysis methods are applicable. The input values and the corresponding value of the QoI are the basis of a non-intrusive analysis. Any potential knowledge on the workings of the computational model is neglected.

Regarding Step B, probability theory is the most popular technique used in the literature for modelling uncertainties. However, in many practical application examples, the knowledge on input parameters is limited, due to time, money and labour constraints. Then, epistemic uncertainty is present in combination with aleatory uncertainty, and probability theory may not be any more an appropriate tool to characterize the uncertainty. Hence, in this thesis, the more general class of *imprecise probabilities* is used to describe the uncertainty in the input parameters. Imprecise probabilities allow for modelling separately aleatory and epistemic sources of uncertainty and hence can represent both types of uncertainty more adequately than probability theory alone.

The use of imprecise probability in uncertainty quantification analyses has become more popular in recent years but is still in its infancy compared to probabilistic modelling. The reason for this phenomenon is the added complexity and the increased computational costs when using imprecise probabilities. When the model is expensive-to-evaluate, the total computational costs is dominated by the number of runs of the computational model. Therefore, standard methods, such as Monte Carlo simulation, should be avoided, despite their robustness. As a countermeasure, meta-modelling (also called *response surface modelling* and *surrogate modelling*) techniques are widely established in probabilistic problem settings (Sudret 2007, 2012). Meta-models approximate the computational model by a simple and inexpensive-to-evaluate function. Subsequently, the meta-models can be used for uncertainty propagation analyses allowing for Monte Carlo simulation-based algorithms. These meta-modelling methods, however, have been applied to imprecise probabilities in rare occasions only. Hence, there is a need to further develop the well-established meta-modelling techniques to the context of imprecise probabilities.

1.4 Objectives and outline of the thesis

Taking into considerations the previous problem statement, this thesis is focused on the following objectives:

(i) Develop an intuitive framework to model imprecise probabilities, *i.e.* the mixture of aleatory and epistemic uncertainty, for the context of engineering problems. The determined framework shall be more general than probability theory and, at the same time, shall be handy for practising engineers.

(ii) Use state-of-the-art meta-modelling techniques to conduct uncertainty quantification analyses in the presence of imprecise probabilities.

(iii) Apply the proposed framework to a variety of classical uncertainty quantification analyses, such as structural reliability analysis and sensitivity analysis.

(iv) Show the relevance of the proposed framework for realistic engineering problems.

In order to address the objectives, the document is organized in nine chapters beyond the introduction and conclusions chapters.

Chapter 2 contains a review of common uncertainty concepts, including the conventional probability theory, probability-boxes, and fuzzy distributions. The concepts are discussed and compared based on the application of a small dataset. Finally, a suitable imprecise probabilities concept is identified and used for the remaining chapters.

In the context of probability theory, Chapter 3 introduces two state-of-the-art meta-modelling techniques, namely Kriging and sparse Polynomial Chaos Expansions (PCE). Kriging assumes that the computational model is a realization of a Gaussian process. PCE represents the response of the computational model by a series expansion of orthonormal polynomials. Using smart selections algorithms such as least-angle regression, sparse PCE models are calibrated. The two techniques are compared based on a one-dimensional example and their advantages and disadvantages are highlighted.

An original combination of Kriging and PCE is presented in Chapter 4. The so-called PC-Kriging meta-modelling technique was proposed by the author and combines the local interpolation property of Kriging with the global regression property of PCE. The resulting meta-model is shown to be more accurate than the conventional techniques taken separately. Apart from uncertainty propagation, the PC-Kriging technique is applied to the estimation of failure probabilities and quantiles in Chapter 5. Making use of active learning algorithms, the efficiency of PC-Kriging in structural reliability analysis is demonstrated.

Starting from Chapter 6, meta-modelling is applied to imprecise probabilities and the algorithms proposed thereof were developed by the author. In particular, Chapter 6 describes the uncertainty propagation of probability-boxes through a computational model. Making extensive use of meta-modelling techniques allows for an efficient estimation of the quantity of interest. In particular, the novel two-level meta-modelling algorithm approximates the computational model at different levels of the uncertainty propagation.

Chapter 7 discusses structural reliability analysis in the presence of imprecise probabilities. Similar to Chapter 6, the use of meta-models at different levels of the imprecise structural reliability analysis yields an efficient estimation of the statistics of interest. The efficiency can be improved further by using the idea of active learning algorithms discussed in Chapter 5.

Sensitivity analysis for imprecise probabilities is discussed in Chapter 8. Two sensitivity measures that are standard tools in the probabilistic literature are discussed in the context of imprecise probabilities. Again, using meta-models allows for an efficient estimation of the distinct sensitivity measures, resulting in a novel sensitivity measure as well as further development of existing sensitivity measures.

Apart from the various benchmark applications presented alongside the theory chapters of this thesis, Chapter 9 presents the NASA uncertainty quantification challenge whereas Chapter 10 illustrates a set of geotechnical engineering problems. The NASA challenge describes a problem set addressing the flight behaviour of an aircraft in adverse conditions and shows results obtained with the previously proposed frameworks. The geotechnics chapter discusses a number of commonly encountered problem settings in the field of geotechnics. Due to the comparably large uncertainties and expensive computational models (typically finite element models in combination with random field theory), geotechnical engineering is an ideal case study for the proposed frameworks.

A short summary of the topics presented in this thesis is shown in Table 1.1. Based on the general uncertainty quantification framework presented in Figure 1.1, its main elements are discussed in this thesis. Moreover, the discussion includes the case of probability theory, which is the conventional approach, as well as the proposed imprecise probabilities, which are the main contributions of the author. Note that for each chapter, "SOA" refers to state-of-the-art methodologies whereas "NEW" refers to original contributions by the author.

Table 1.1: Summary of the content of the thesis.

	Probability theory	Imprecise probabilities	Objectives
Uncertainty modelling	Chapter 2 (SOA)	Chapter 2 (SOA)	(i)
Uncertainty propagation	Chapter 3 (SOA), Chapter 4 (NEW)	Chapter 6 (NEW)	(ii)
Structural reliability	Chapter 5 (NEW)	Chapter 7 (NEW)	(iii)
Sensitivity analysis	Chapter 8 (SOA)	Chapter 8 (NEW)	(iii)
Case studies	Chapter 10 (NEW)	Chapter 9 (NEW), Chapter 10 (NEW)	(iv)

CHAPTER 2

Imprecise probabilities

Uncertainty in parameters can be modelled by a variety of concepts. In this chapter, a number of such concepts are introduced, including conventional probability theory, Bayesian hierarchical models, evidence theory, probability-boxes and fuzzy distributions. Each concept is defined by its characteristic terminology and equations. Further, a case study illustrates its application to a small dataset. At the end of the chapter, the concepts are connected, compared and discussed. Probability-boxes are chosen as the uncertainty modelling approach for targeting the problem statement in this thesis.

2.1 Probability theory

2.1.1 Theory

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω denotes an event space (also called sample space, universal set, or outcome space) equipped with the σ-algebra $\mathcal{F}$ and a probability measure $\mathbb{P} \in [0,1]$. Consider an event $\mathcal{E} \in \Omega$ and its complementary event $\mathcal{E}^c$. By definition, their union reads $\mathcal{E} \cup \mathcal{E}^c = \Omega$ and their intersection results in the zero event $\mathcal{E} \cap \mathcal{E}^c = \emptyset$. Then, the probability of $\mathcal{E}$ and $\mathcal{E}^c$ add up to one: $\mathbb{P}(\mathcal{E}) + \mathbb{P}(\mathcal{E}^c) = 1$. The probability of the empty set $\emptyset$ and the complete set are $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(\Omega) = 1$, respectively.

In this context, a random variable X is defined by the mapping $X(\omega) : \Omega \mapsto \mathcal{D}_X \subset \mathbb{R}$, where $\omega \in \Omega$ is an elementary event and $\mathcal{D}_X$ is the support domain of X. A realization of the random variable X is denoted by the corresponding lower case letter x. A random variable X is typically characterized by a *cumulative distribution function* (CDF) F_X which assigns a probability to the event $\{X \leq x\}$, *i.e.* $F_X(x) = \mathbb{P}(X \leq x)$. It follows from this definition that any CDF is monotonically non-decreasing, tends to zero for low values of x, and tends to one for large values of x. For continuous random variables, the first derivative of the CDF is called *probability density function* (PDF) and is defined by $f_X(x) = \mathrm{d}F_X(x)/\mathrm{d}x$. The PDF describes the likelihood of X being in the neighbourhood of x. Due to the monotonicity property of CDFs, $f_X(x) \geq 0$ for all $x \in \mathcal{D}_X$.

2.1.2 Data to CDF

Empirical CDF

Consider a set of realizations $\mathcal{X} = \left\{\chi^{(1)}, \dots, \chi^{(N)}\right\}$ of a random variable X, whose probability distribution is unknown a-priori. In order to describe X properly, the data $\mathcal{X}$ is used to infer a probability distribution.

A basic method of statistics is then to compute the empirical CDF, which is defined as:

$$F_X^{\text{emp}}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{x \geq \chi^{(i)}}(x), \tag{2.1}$$

where $\mathbb{I}$ is the indicator function, which reads $\mathbb{I} = 1$ for a true subscript statement and $\mathbb{I} = 0$ otherwise. Due to the limited number of samples N, the empirical CDF is a stair-shaped curve with constant CDF values between samples and steps at $x = \chi^{(i)},\ i = 1, \dots, N$. Assuming that the underlying probability distribution X may be continuous, the empirical CDF provides a poor but simple estimate of F_X. Hence, more sophisticated methods, such as the method of moments or the maximum likelihood method, are commonly used in practice.

Method of moments

Let us consider a distribution family $F_X(x|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes a vector of parameters defining the shape of the CDF. Further, we denote $n_{\boldsymbol{\theta}} = |\boldsymbol{\theta}|$ the number of such parameters. Note that for distributions commonly used in the literature and practice, $n_{\boldsymbol{\theta}} = 2$. Then, the method of moments determines the optimal distribution by matching the first moments of $F_X(x|\boldsymbol{\theta})$ with the sample-based estimations of the first moments based on $\mathcal{X}$. Precisely, denote the mean value and variance of X by $\mu(\boldsymbol{\theta})$ and $\sigma^2(\boldsymbol{\theta})$, which depend on the yet-to-be-determined parameters $\boldsymbol{\theta}$. Further, denote the sample mean and variance by $\mathbb{E}\left[\mathcal{X}\right]$ and $\mathrm{Var}\left[\mathcal{X}\right]$:

$$\mathbb{E}\left[\mathcal{X}\right] = \frac{1}{N} \sum_{i=1}^{N} \chi^{(i)}, \qquad \mathrm{Var}\left[\mathcal{X}\right] = \frac{1}{N-1} \sum_{i=1}^{N} \left(\chi^{(i)} - \mathbb{E}\left[\mathcal{X}\right]\right)^2. \tag{2.2}$$

Then, the parameters $\boldsymbol{\theta}$ are obtained by solving: $\mu(\boldsymbol{\theta}) = \mathbb{E}\left[\mathcal{X}\right], \sigma^2(\boldsymbol{\theta}) = \mathrm{Var}\left[\mathcal{X}\right]$.

The method of moments relies on the knowledge of the underlying distribution family, which is an additional assumption compared to the previous empirical CDF. However, this assumption avoids the stair-shaped CDF curves of the empirical CDF and allows for a smooth CDF curve.

Maximum likelihood method

Similarly to the method of moments, the maximum likelihood method requires the knowledge of a distribution family $F_X(x|\boldsymbol{\theta})$ with unknown distribution parameters $\boldsymbol{\theta}$. Then, the likelihood of observing $\mathcal{X}$ depending on the parameter value $\boldsymbol{\theta}$ can be computed by:

$$\mathcal{L}(\boldsymbol{\theta}|\mathcal{X}) = \prod_{i=1}^{N} f_X(\chi^{(i)}|\boldsymbol{\theta}), \tag{2.3}$$

where $f_X(x|\boldsymbol{\theta})$ is the PDF conditional on $\boldsymbol{\theta}$. The optimal parameter values are inferred by maximizing the likelihood function:

$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}|\mathcal{X}). \tag{2.4}$$

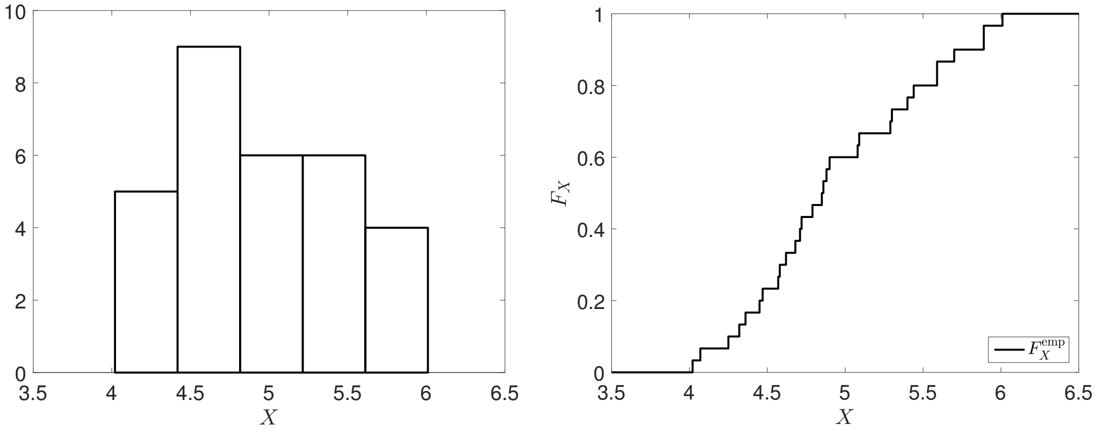
The maximum likelihood method suffers from the same pitfall as the method of moments. The feasibility of the method depends on the initial assumption on the distribution family. In fact, any suitable and non-suitable distribution family can be fitted with these methods. Moreover, the optimization in Eq. (2.4) is not always trivial to solve.

Dataset $\mathcal{X}_1$

In order to compare the various methods, a small dataset $\mathcal{X}_1$ of $N = 30$ samples is used. The dataset contains the following ordered values:

$$\mathcal{X}_1 = \left\{ \begin{matrix} 4.02 & 4.07 & 4.25 & 4.32 & 4.36 & 4.45 \\ 4.47 & 4.57 & 4.58 & 4.62 & 4.68 & 4.71 \\ 4.72 & 4.79 & 4.85 & 4.86 & 4.88 & 4.90 \\ 5.08 & 5.09 & 5.29 & 5.30 & 5.40 & 5.44 \\ 5.59 & 5.59 & 5.70 & 5.89 & 5.89 & 6.01 \end{matrix} \right\}, \tag{2.5}$$

which is sampled from a Gaussian distribution with mean value $\mu_X = 5$ and standard deviation $\sigma_X = 0.5$. Figure 2.1 shows a histogram and the empirical CDF of $\mathcal{X}_1$.



(a) Histogram (b) Empirical CDF

Figure 2.1: Dataset $\mathcal{X}_1$ – simple visualizations.

In order to conduct an analysis by the method of moments, the Gaussian distribution family is chosen. This distribution family has $n_{\boldsymbol{\theta}} = 2$ parameters (*i.e.* mean μ and standard deviation σ). In order to match these parameters, the mean value and variance of $\mathcal{X}_1$ are estimated:

$$\mathbb{E}\left[\mathcal{X}_1\right] = \frac{1}{N} \sum_{i=1}^{N} \chi^{(i)} = 4.95, \tag{2.6}$$

$$\mathrm{Var}\left[\mathcal{X}_1\right] = \frac{1}{N-1} \sum_{i=1}^{N} \left[\chi^{(i)} - \mathbb{E}\left[\mathcal{X}_1\right]\right]^2 = 0.31. \tag{2.7}$$

The resulting Gaussian distribution is defined by $\mu_X = \mathbb{E}\left[\mathcal{X}_1\right] = 4.95$ and $\sigma_X = \sqrt{\mathrm{Var}\left[\mathcal{X}_1\right]} = 0.56$. The CDF of the fitted distribution is shown in Figure 2.2 where it is compared to the empirical CDF. The fitted distribution follows nicely the empirical CDF.

At last, the parameters of the Gaussian distribution family are estimated by maximum likelihood method. For Gaussian distributions, the likelihood reads:

$$\mathcal{L}(\mu, \sigma | \mathcal{X}_1) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\chi^{(i)} - \mu}{\sigma}\right)^2\right]. \tag{2.8}$$

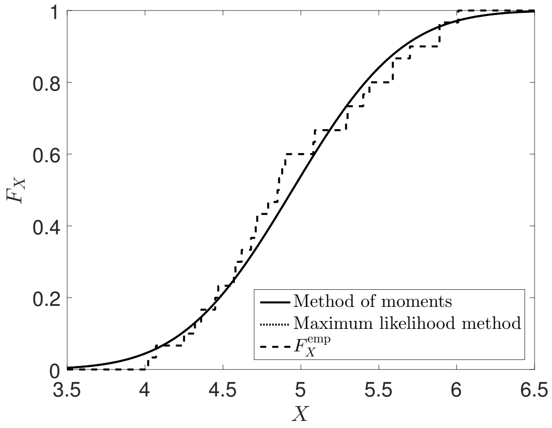


Figure 2.2: Dataset $\mathcal{X}_1$ – fitted distribution versus empirical CDF of the dataset.

The maximum likelihood estimate reads $\max_{\boldsymbol{\theta}} \mathcal{L}(\mu, \sigma|\mathcal{X}_1) = 2.37 \cdot 10^{-11}$ when using $\mu = 4.95$ and $\sigma = 0.56$. These parameters are identical to the method of moments at two-decimal precision. Hence, the resulting distribution function is equivalent to the one obtained by the method of moments as seen by the overlapping curves in Figure 2.2. This, however, might not be the case for other classical distributions.

2.1.3 Limitations

Probability theory provides a single measure for uncertainty. In other words, the uncertainty in X must be known and quantifiable by the probability measure $\mathbb{P}$. The distinction of aleatory and epistemic uncertainty, as described in Chapter 1.1, is not possible. Considering *e.g.* the method of moments, uncertainty arises from the variability of X, the choice of distribution family, and also the estimation of the moments based on a finite sample set N. This is why in many situations, the uncertainty in X consists of a mix of aleatory and epistemic uncertainty. Probability theory is not able to treat each of these sources of uncertainty separately. Therefore, a more general framework, such as Bayesian hierarchical modelling, p-box, or fuzzy variable, is more adequate to characterize imprecise probabilities.

2.2 Bayesian hierarchical modelling

2.2.1 Theory

Bayesian hierarchical modelling belongs to the group of subjective probabilities (Beer, Ferson, & Kreinovich 2016) and is a natural extension to the traditional probability theory. Considering a distribution family $F_X(x|\boldsymbol{\theta})$, a Bayesian hierarchical model describes the epistemic uncertainty with a probability distribution function on $\boldsymbol{\theta}$ and the aleatory uncertainty with the conditional probability distribution $F_X(x|\boldsymbol{\theta})$ (Gelman et al. 2009):

$$F_X(x) = F_X(x|\boldsymbol{\theta}), \qquad \text{where} \qquad F_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = F_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|\boldsymbol{\vartheta}), \tag{2.9}$$

where $F_{\boldsymbol{\Theta}}$ is the CDF of the parameters of F_X, and $\boldsymbol{\vartheta}$ are the corresponding distribution hyperparameters with $n_{\boldsymbol{\vartheta}} = |\boldsymbol{\vartheta}|$. Note that both types of uncertainty are modelled by probabilistic distributions. Hence, the properties and axioms of the previous section are valid on both levels of the hierarchical uncertainty model.

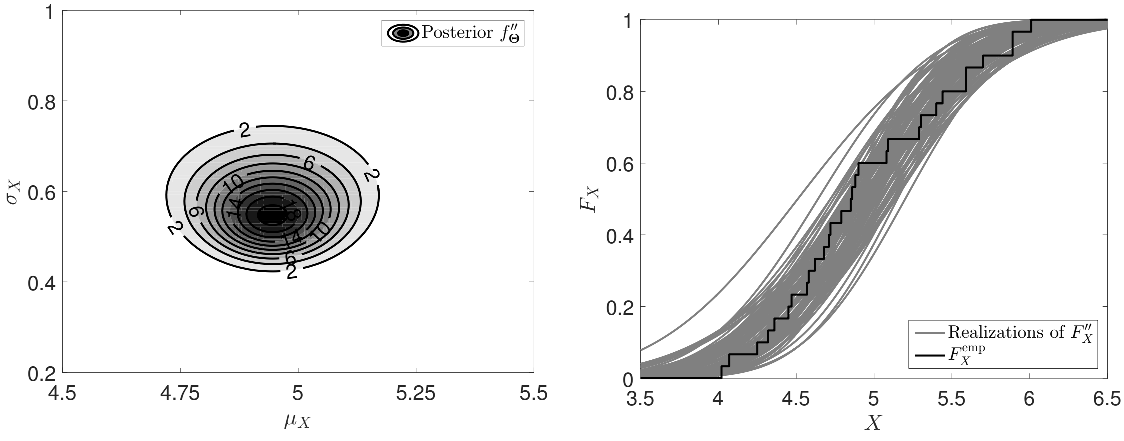
2.2.2 Data to Bayesian hierarchical model

Given a Bayesian hierarchical model and a dataset $\mathcal{X}$, Bayesian inference can be applied to estimate the posterior distribution of the epistemic parameters $\boldsymbol{\theta}$:

$$f''_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|\mathcal{X}, \boldsymbol{\vartheta}) = \frac{\mathcal{L}\left(\boldsymbol{\theta}|\mathcal{X}\right)}{\int_{\mathcal{D}_{\boldsymbol{\Theta}}} \mathcal{L}\left(\boldsymbol{\theta}|\mathcal{X}\right)\ f'_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|\boldsymbol{\vartheta})\ \mathrm{d}\boldsymbol{\theta}} f'_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|\boldsymbol{\vartheta}), \tag{2.10}$$

where $f'_{\boldsymbol{\Theta}}$ and $f''_{\boldsymbol{\Theta}}$ denote the prior and posterior probability density function of $\boldsymbol{\Theta}$, respectively, and $\mathcal{L}$ is the likelihood defined in Eq. (2.3). Note that the denominator has a constant value, independent of $\boldsymbol{\theta}$ and is sometimes also called *evidence*. This evidence has, however, nothing in common with the later presented evidence theory.

Consider again the example case of dataset $\mathcal{X}_1$. As a prior hierarchical model, a Gaussian distribution is assumed where the parameters are uniformly distributed as $\mu'_X \sim \mathcal{U}\left(0, 10\right)$ and $\sigma'_X \sim \mathcal{U}\left(0.2, 1\right)$. The corresponding posterior distribution of $\left(\mu_X, \sigma_X\right)^{\mathsf{T}}$ is shown in Figure 2.3(a). The large values of the posterior PDF group around the true, but unknown value of $\{\mu_X = 5, \sigma_X = 0.5\}$. The maximum posterior probability density is found at hyperparmeter values $\{\mu_X = 4.95,\ \sigma_X = 0.56\}$, which is equal to the point found in the method of maximum likelihood. When assuming a uniform prior for the Bayesian hierarchical model, the two methods lead to identical parameters with maximum likelihood. Note that the posterior distribution function $F''_{\boldsymbol{\Theta}}$ has a non-analytical form here and in the general case. Hence, the sampling of such a distribution can be a non-trivial task.



(a) Posterior PDF $f''_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = f''_{\boldsymbol{\Theta}}(\mu_X, \sigma_X)$

(b) Posterior hierarchical model

Figure 2.3: Dataset $\mathcal{X}_1$ – Bayesian hierarchical model.

Figure 2.3(b) compares the empirical CDF of $\mathcal{X}_1$ and 100 realizations of conditional CDFs $F_X(x|\boldsymbol{\theta})$ based on the posterior distribution $F''_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$. The posterior CDF realizations group nicely around the empirical CDF. It is, however, still possible to generate realizations far away from the empirical CDF, such as the most left curve with $\{\mu_X = 4.51, \sigma_X = 0.71\}$.

2.2.3 Limitations

The Bayesian hierarchical model greatly depends on the assumption that epistemic uncertainty can be quantified by a probability distribution (*i.e.* $F_{\boldsymbol{\Theta}}$). If there are good reasons (*e.g.* expert knowledge) to assume such a distribution, then Bayesian hierarchical models can be adequate. However, when the lack of knowledge does not allow to quantify a subjective distribution, Bayesian hierarchical models may be inadequate.

In many practical cases, a uniform prior $f'_{\mathbf{\Theta}}$ is assumed when prior knowledge is so vague that no distribution can be specified. The uniform distribution is a viable choice respecting the principle of maximum entropy when only the bounds on $\boldsymbol{\theta}$ are known. However, a uniform distribution implies a weighing between different values of $\mathbf{\Theta}$, which in turn might not correspond to any physical meaning in the system. This point adds to the discussion and criticism whether treating epistemic uncertainty with probability theory is appropriate.

A third point of discussion is the following: by using a probability distribution for $\mathbf{\Theta}$, the modeller focuses on the analysis of the average behaviour of a variable. In other words, extreme realizations of $\mathbf{\Theta}$ are weighted against all other realizations and are thus likely to be under-represented. When the modeller puts importance on the worst/best case scenario, Bayesian hierarchical models appear to be suboptimal.

2.3 Evidence theory

2.3.1 Theory

Evidence theory is based on the developments by Dempster (1967) and Shafer (1976) and is hence often called *Dempster-Shafer's theory of evidence.* Evidence theory loosens the strict assumption of a single probability measure $\mathbb{P}$ in probability theory. In particular, evidence theory considers two measures, *i.e.* belief and plausibility, for each event $\mathcal{E}$ in the event space Ω.

Consider the triple $(\Omega, \mathcal{F}, m)$, where Ω is again the event space, $\mathcal{F}$ is a countable collection of subsets of Ω, and m is the *basic probability assignment* (BPA) defined as:

$$m(\mathcal{J}) = \begin{cases} > 0 & \text{if } \mathcal{J} \in \mathcal{F} \\ 0 & \text{if } \mathcal{J} \subset \Omega \text{ and } \mathcal{J} \notin \mathcal{F} \end{cases} \tag{2.11}$$

$$\sum_{\mathcal{J} \in \mathcal{F}} m(\mathcal{J}) = 1. \tag{2.12}$$

In this concept, $m(\mathcal{J})$ is interpreted as the amount of likelihood that is associated with event $\mathcal{J}$ but without any specifications of how this likelihood might be appointed in any sub-events $\mathcal{J}_* \subset \mathcal{J}$.

Then, the *belief* measure is defined as the minimum amount of likelihood that *must* be associated to an event $\mathcal{E}$, whereas the *plausibility* is defined as the maximum amount of likelihood that *could* be associated to the same event $\mathcal{E}$. In mathematical terms, the belief and plausibility are defined as follows in terms of the event $\mathcal{E}$ and the event $\mathcal{J}$ (or possibly a set of events $\mathcal{J}$):

$$\mathrm{Bel}\,(\mathcal{E}) = \sum_{\mathcal{J} \subset \mathcal{E}} m(\mathcal{J}), \tag{2.13}$$

$$\mathrm{Pl}\,(\mathcal{E}) = \sum_{\mathcal{J} \cap \mathcal{E} \neq \emptyset} m(\mathcal{J}). \tag{2.14}$$

It follows from Eqs. (2.13) and (2.14) that $\mathrm{Bel}\,(\mathcal{E}) + \mathrm{Bel}\,(\mathcal{E}^c) \leq 1$, $\mathrm{Pl}\,(\mathcal{E}) + \mathrm{Pl}\,(\mathcal{E}^c) \geq 1$, and $\mathrm{Bel}\,(\mathcal{E}) + \mathrm{Pl}\,(\mathcal{E}^c) = 1$. Note that probability theory can be seen as a special case of evidence theory, when $\mathrm{Bel}\,(\mathcal{E}) = \mathrm{Pl}\,(\mathcal{E})$. The construction of $\mathrm{Bel}\,(\cdot)$, $\mathrm{Pl}\,(\cdot)$, $\mathcal{J}$, and $m(\mathcal{J})$ is also referred to as *Dempster-Shafer structure* in the literature.

A special case of the Dempster-Shafer structure emerges when considering the event $\{X \leq x\}$, used earlier in the section on probability theory. Let $\mathcal{J}_X = \left\{\mathcal{J}_Y^{(1)}, \ldots, \mathcal{J}_Y^{(N)}\right\}$ denote a set of events whose components can be represented by an interval in the domain $\mathcal{D}_X \in \mathbb{R}$. Then, the belief and plausibility measures can be interpreted as the minimum and maximum values for the CDF of variable X. Eqs. (2.13) and (2.14) can be rewritten as:

$$\mathrm{Bel}\left(X \leq x\right) = \mathrm{Bel}\left(F_X(x)\right) = \sum_{i=1}^{N} \mathbb{I}_{\left(\max \mathcal{J}_X^{(i)} \leq x\right)}\left(\mathcal{J}_X^{(i)}\right)\, m\left(\mathcal{J}_X^{(i)}\right), \tag{2.15}$$

$$\mathrm{Pl}\left(X \leq x\right) = \mathrm{Pl}\left(F_X(x)\right) = \sum_{i=1}^{N} \mathbb{I}_{\left(\min \mathcal{J}_X^{(i)} \leq x\right)}\left(\mathcal{J}_X^{(i)}\right)\, m\left(\mathcal{J}_X^{(i)}\right), \tag{2.16}$$

where $\min \mathcal{J}_X^{(i)}$ and $\max \mathcal{J}_X^{(i)}$ denote the minimum and maximum x-value in event $\mathcal{J}_X^{(i)}$, respectively.

2.3.2 Data to Dempster-Shafer structure

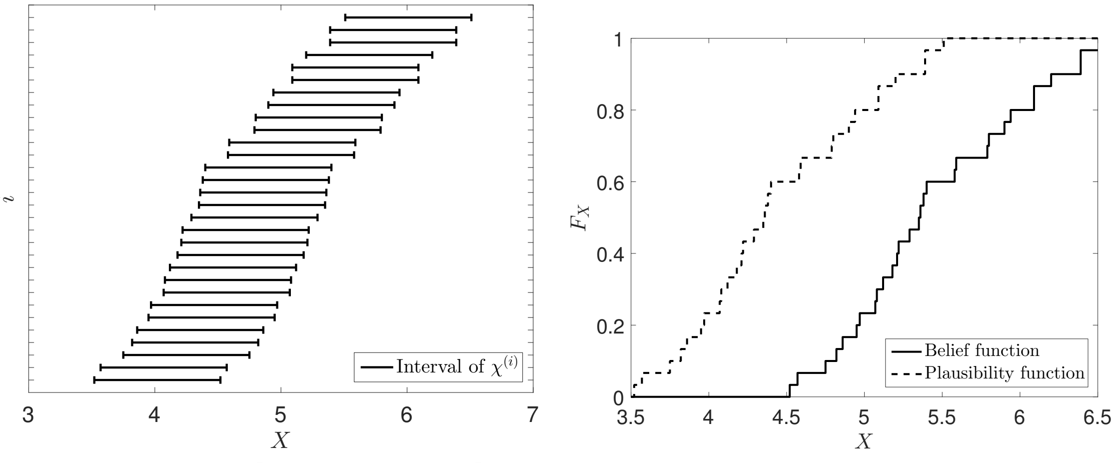
Consider again the dataset $\mathcal{X}_1$. In order to construct a Dempster-Shafer structure, it is assumed that the sampled data points $\chi^{(i)}$ are imprecise due to inaccuracies in the measurement devices. The true value of each sample point is found in the interval $\chi_*^{(i)} \in \left[\chi^{(i)} - 0.5,\ \chi^{(i)} + 0.5\right]$. The intervals are shown in Figure 2.4(a). Further, the probability mass assignment is chosen as equal for all samples, *i.e.* $m\left(\chi^{(i)}\right) = 1/N = 1/30$.

Then, the belief and plausibility measure for the event $\{X \leq x\}$ reads:

$$\mathrm{Bel}\left(X \leq x\right) = \sum_{i=1}^{N} \mathbb{I}_{\left(\chi^{(i)}+0.5 \leq x\right)}(x)\ m\left(\chi^{(i)}\right), \tag{2.17}$$

$$\mathrm{Pl}\left(X \leq x\right) = \sum_{i=1}^{N} \mathbb{I}_{\left(\chi^{(i)}-0.5 \leq x\right)}(x)\ m\left(\chi^{(i)}\right). \tag{2.18}$$

The resulting belief and plausibility values are shown in Figure 2.4(b). The resulting CDFs have the same shape as the empirical CDF translated along the horizontal direction by +0.5 and −0.5, respectively.



(a) Intervals $\chi_*^{(i)} \in \left[\chi^{(i)} - 0.5, \chi^{(i)} + 0.5\right]$

(b) Belief and plausibility measures

Figure 2.4: Dataset $\mathcal{X}_1$ – Dempster-Shafer structure.

2.3.3 Limitations

The concept of Dempster-Shafer's theory of evidence depends on the BPA and hence on the ability to determine bounds for specific events. The definition of the bounds is, however, often a non-trivial task in practical applications. In case it is not possible, the reader is referred to the subsequent sections of this thesis.

2.4 Probability-boxes

2.4.1 Theory

Probability-boxes (p-boxes) define the CDF of a variable X by lower and upper bounds denoted by $\underline{F}_X$ and $\overline{F}_X$, respectively (Ferson & Ginzburg 1996; Ferson & Hajagos 2004). For any value $x \in \mathcal{D}_X$, the true-but-unknown CDF lies within these bounds such that $\underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x)$. The two boundary curves form an intermediate area, hence the name probability-*box*. In the literature, two types of p-boxes are distinguished, namely *free* and *parametric* p-boxes, which are discussed in the following.

Free p-boxes

Free p-boxes are defined as introduced above, *i.e.* by lower and upper bounds on the CDF only. This implies that the true CDF can have an arbitrary shape as long as it fulfils the characteristics of a generic CDF and lies within the bounds of the p-box. Figure 2.5 shows the boundary curves of a free p-box and a number of possible realizations of the true CDF. Because the shape of the true CDF is not specified, different types of curves are possible, including non-smooth ones (see realization #3 in Figure 2.5).

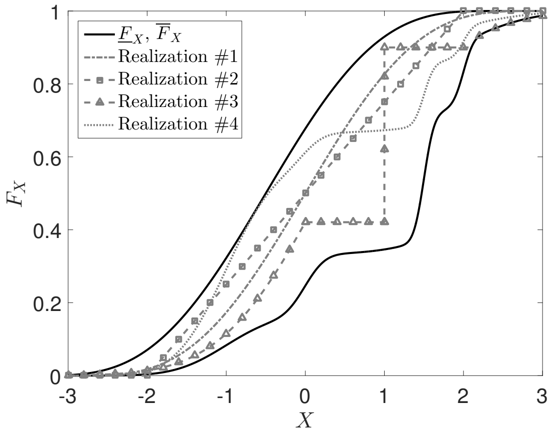


Figure 2.5: Free p-box – bounds and realizations of possible CDFs.

Free p-boxes are a special case of Dempster-Shafer structures, when considering the event $\{X \leq x\}$. Then, the belief function $\mathrm{Bel}\,(X \leq x)$ is equivalent to $\underline{F}_X(x)$ and the plausibility function $\mathrm{Pl}\,(X \leq x)$ is equivalent to $\overline{F}_X(x)$, for all $x \in \mathcal{D}_X$. This implies a duality between free p-boxes and Dempster-Shafer structures: a free p-box can be represented by a Dempster-Shafer structure and vice versa (Walley 2000; Ferson, Kreinovich, et al. 2003; Möller & Beer 2008).

Parametric p-boxes

Parametric p-boxes (a.k.a. distributional p-boxes) are defined as distribution function families the parameters of which are known in intervals:

$$F_X(x) = F_X(x|\boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}} \subset \mathbb{R}^{n_{\boldsymbol{\theta}}}, \tag{2.19}$$

where $\mathcal{D}_{\boldsymbol{\Theta}}$ is the interval domain of the distribution parameters of dimension $n_{\boldsymbol{\theta}} = |\boldsymbol{\theta}|$. Precisely, $\mathcal{D}_{\boldsymbol{\Theta}} = \left[\underline{\theta}_1, \overline{\theta}_1\right] \times \ldots \times \left[\underline{\theta}_{n_{\boldsymbol{\theta}}}, \overline{\theta}_{n_{\boldsymbol{\theta}}}\right]$ denotes a hyper-rectangular domain in this thesis for the sake of simplicity. Note that $\underline{\theta}$ and $\overline{\theta}$ denote the lower and upper bound value of the interval $\left[\underline{\theta}, \overline{\theta}\right]$.

Parametric p-boxes allow for a clear separation of aleatory and epistemic uncertainty: aleatory uncertainty is represented by the distribution function family, whereas epistemic uncertainty is represented by the intervals in the distribution parameters. However, parametric p-boxes are more restrictive than free p-boxes because they require knowledge on the distribution family. In other words, free p-boxes can be interpreted as a generalization of parametric p-boxes where the distribution family has $n_{\boldsymbol{\theta}} \to \infty$ parameters.

Figure 2.6 illustrates a parametric p-box consisting of a Gaussian distribution family with mean value $\mu \in [-0.5, 0.5]$ and standard deviation $\sigma \in [0.7, 1.0]$. The lower and upper boundary curves of the parametric p-box drawn in Figure 2.6 are obtained by:

$$\underline{F}_X(x) = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} F_X(x|\boldsymbol{\theta}), \qquad \overline{F}_X(x) = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} F_X(x|\boldsymbol{\theta}), \tag{2.20}$$

where $\boldsymbol{\theta} = (\mu, \sigma)^{\mathsf{T}}$ in this example. Hence, the boundary curves have the characteristics of a CDF but are not necessarily a realization with a specific parameter vector $\boldsymbol{\theta}^{(0)}$. They generally consist of sections of different realizations. As an example, the lower boundary CDF $\underline{F}_X$ is a combination of realization #2 (defined by $\mu = 0.5$, $\sigma = 0.7$) and realization #3 ($\mu = 0.5$, $\sigma = 1.0$).

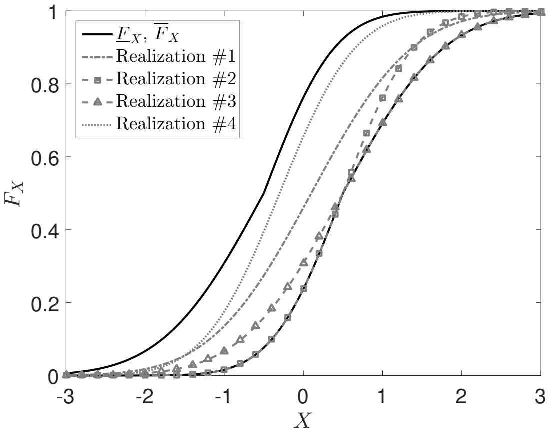


Figure 2.6: Parametric p-box – bounds and realizations of possible CDFs.

Parametric p-boxes are related to the previously discussed Bayesian hierarchical models (Section 2.2). The two-level construction of a parametric p-box in Eq. (2.9) resembles a Bayesian hierarchical model in Eq. (2.19) in which the distribution of the hyper-parameters $\boldsymbol{\theta} \sim \boldsymbol{\Theta}$ is replaced by an interval domain $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$. Note that in the p-box approach, no weight function is attached to the various values of $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$.

2.4.2 Data to free p-box

A number of methods have been proposed to construct free p-boxes and the equivalent Dempster-Shafer structures, some of which are presented here in brief. For further details and other methods, it is referred to Ferson, Kreinovich, et al. (2003). In this thesis, the Kolmogorov-Smirnov confidence bound, Chebyshev's inequalities, and robust Bayes' method are reviewed briefly.

Kolmogorov-Smirnoff confidence bounds

Consider a dataset with N samples. Based on the Kolmogorov-Smirnov (KS) statistics (Smirnov 1939; Kolmogoroff 1941; Feller 1948; Miller 1956), Grosof (1986) and Ferson, Kreinovich, et al. (2003) define the bounds of a free p-box by KS confidence limits as follows:

$$\underline{F}_X(x) = \min\left[1,\, \max\left(0, F_n(x) - D_N^\alpha\right)\right], \tag{2.21}$$

$$\overline{F}_X(x) = \min\left[1,\, \max\left(0, F_n(x) + D_N^\alpha\right)\right], \tag{2.22}$$

where $F_N(x)$ is the empirical CDF of the dataset and D_N^α is the KS critical value at significance level α and sample size N. The values for D_N^α can be found in Kolmogoroff (1941). Note that the obtained p-boxes are based on confidence bounds rather than absolute bounds. Hence strictly speaking, this approach violates the definition of a p-box. However, using confidence bounds is often the only possible and reasonable choice in practice.

Chebychev's inequalities

In some situations it is not possible to have a dataset to characterize a free p-box, but it is possible to obtain certain statistics, such as mean value, variance, or minimum/maximum value. In this context, a number of methods have been published to bound the CDF of interest. In particular, Rowe (1988) discusses the case of known {minimum, maximum, mean}, Smith (1995) deals with the combination of {minimum, maximum, mode}, Chebyshev (1874) and Feller (1968) discuss {mean, variance}, and Mallows (1956) and Smith (1995) handle PDF limits. These methods can be used in the context of free p-boxes with little modification.

As an example, Oberguggenberger & Fellin (2008) and H. Zhang, Dai, et al. (2013) define the bounds of a free p-box by the two-sided Chebyshev's inequality:

$$\underline{F}_X(x) = \begin{cases} 0, & \text{for } x < \mu + \sigma \\ 1 - \frac{\sigma^2}{(x-\mu)^2}, & \text{for } x \geq \mu + \sigma \end{cases}, \tag{2.23}$$

$$\overline{F}_X(x) = \begin{cases} \frac{\sigma^2}{(x-\mu)^2}, & \text{for } x < \mu - \sigma \\ 1, & \text{for } x \geq \mu - \sigma, \end{cases}, \tag{2.24}$$

where μ and σ are the known mean value and standard deviation of X. Tighter bounds can be obtained by using one-sided Chebyshev's inequalities and knowledge about the minimum and maximum value of X (Ferson, Kreinovich, et al. 2003; H. Zhang, Dai, et al. 2013):

$$\underline{F}_X(x) = \begin{cases} 0, & \text{for } x \leq \mu + \sigma^2/(\mu - \overline{x}) \\ 1 - [b(1+a) - c - b^2]/a, & \text{for } \mu + \sigma^2/(\mu - \overline{x}) < x < \mu + \sigma^2/(\mu - \underline{x}) \\ 1/[1 + \sigma^2/(x-\mu)^2], & \text{for } \mu + \sigma^2/(\mu - \underline{x}) \leq x < \overline{x} \\ 1, & \text{for } x \geq \overline{x} \end{cases}, \tag{2.25}$$

$$\overline{F}_X(x) = \begin{cases} 0, & \text{for } x \leq \underline{x} \\ 1/[1+(\mu-x)^2/\sigma^2] & \text{for } \underline{x} < x \leq \mu + \sigma^2/(\mu - \overline{x}) \\ 1-(b^2-ab+c)/(1-a), & \text{for } \mu + \sigma^2/(\mu - \overline{x}) < x \leq \mu + \sigma^2/(\mu - \underline{x}) \\ 1, & \text{for } x \geq \mu + \sigma^2/(\mu - \underline{x}) \end{cases}, \tag{2.26}$$

where $x \in [\underline{x}, \overline{x}]$, $a = (x - \underline{x})/(\overline{x} - \underline{x})$, $b = (\mu - \underline{x})/(\overline{x} - \underline{x})$, and $c = \sigma^2/(\overline{x} - \underline{x})^2$.

Robust Bayes' method

The *robust Bayes' method* (Berger 1985) is based on the basic Bayes' theorem. Consider again the hierarchical model presented in Section 2.2 and denote $\boldsymbol{\theta}^*$ the hyper-parameter values which maximize the posterior PDF $f''_{\boldsymbol{\Theta}}$. $\boldsymbol{\theta}^*$ and the corresponding CDF model $F''_X(x|\boldsymbol{\theta}^*)$ depend on the distribution family F_X, the formulation of the likelihood $\mathcal{L}$, and the prior distribution $f'_{\boldsymbol{\Theta}}$. In the context of robust Bayes' method, the insistence on having a single, precise prior distribution and a single, specific likelihood function are relaxed. Instead, an entire class of distribution families and prior distributions $f'_{\boldsymbol{\Theta}}$ are used. Different combinations of prior and likelihood lead to different posterior distributions F''_X. Hence, a p-box can be obtained by forming the envelope or convex hull of this set of posterior CDFs. Consider a set of $k = 1, \ldots, n_b$ pairs of distribution families and prior distributions, *i.e.* $\left\{F_X^{(k)}, F_{\boldsymbol{\Theta}}^{\prime(k)}\right\}$, the bounds of the free p-box are obtained by:

$$\underline{F}_X(x) = \min_{k=1,\ldots,n_b} F_X^{\prime\prime(k)}\left(x|\boldsymbol{\theta}^{*(k)}\right), \tag{2.27}$$

$$\overline{F}_X(x) = \max_{k=1,\ldots,n_b} F_X^{\prime\prime(k)}\left(x|\boldsymbol{\theta}^{*(k)}\right), \tag{2.28}$$

where $\boldsymbol{\theta}^{*(k)}$ denotes the optimal hyper-parameter value of combination k. Note that n_b can be arbitrarily large when considering the prior distribution to be a beta distribution with unknown-but-bound hyper-parameters.

Compound of expert opinions

In the previous methods, a single source of information has been considered. However, a common situation in practice is when multiple sources are available, *e.g.* when a number of experts is asked for their opinion. There exists an extensive literature on the aggregation of information (Dubois & Prade 1992; Ayyub 2001; Sentz & Ferson 2002; Ferson, Kreinovich, et al. 2003; Ayyub & Klir 2006). Methods include null aggregation, intersection, enveloping, Dempster's rule and its modifications, Bayes' rule, mixture rules, logarithmic pooling and averaging. For further details, it is referred to the corresponding literature.

Dataset $\mathcal{X}_1$

The KS-confidence bounds, Chebyshev's inequalities, and robust Bayes' method are applied to the dataset $\mathcal{X}_1$. For the KS-confidence bounds, the level of significance is set to $\alpha = 0.10$. The resulting free p-box is shown in Figure 2.7 alongside the empirical CDF of $\mathcal{X}_1$. The bounds of the resulting p-box are a vertical translation of the empirical CDF by a distance of $D_N^\alpha = 0.22$. Note that the support of the p-box is $[-\infty, \infty]$, when assuming no knowledge on the real support of X. This shows the amount of conservatism in this approach.

Chebyshev's inequalities require knowledge on the mean value and standard deviation of X. Therefore, it is assumed that the empirical mean value (Eq. (2.6)) and variance (Eq. (2.7)) correspond to the true mean value and variance of X. Figure 2.8 shows the resulting p-boxes

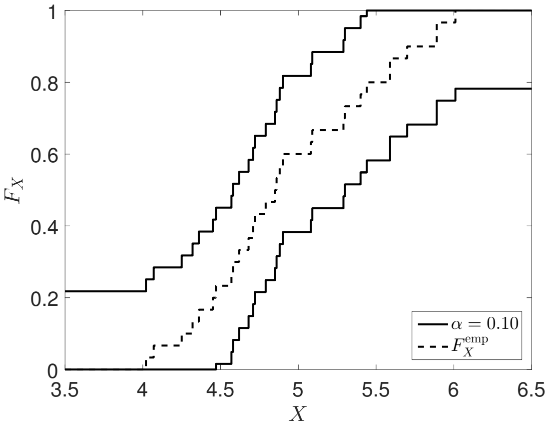
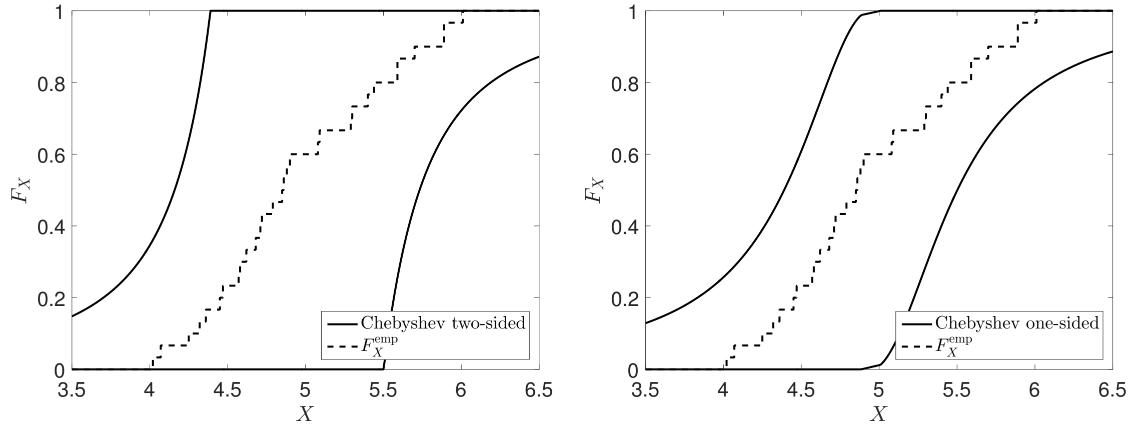


Figure 2.7: Dataset $\mathcal{X}_1$ – free p-boxes – KS-confidence bounds.

based on the Chebyshev's inequalities compared to the empirical CDF of the dataset. Note that for the one-sided results, it is further assumed that $x \in [0, 10]$. Figure 2.8 illustrates nicely how the free p-box changes, when the information on the support of X is added to the analysis: the p-box becomes narrower, *i.e.* the epistemic uncertainty becomes smaller.



(a) Two-sided Chebyshev's inequalities

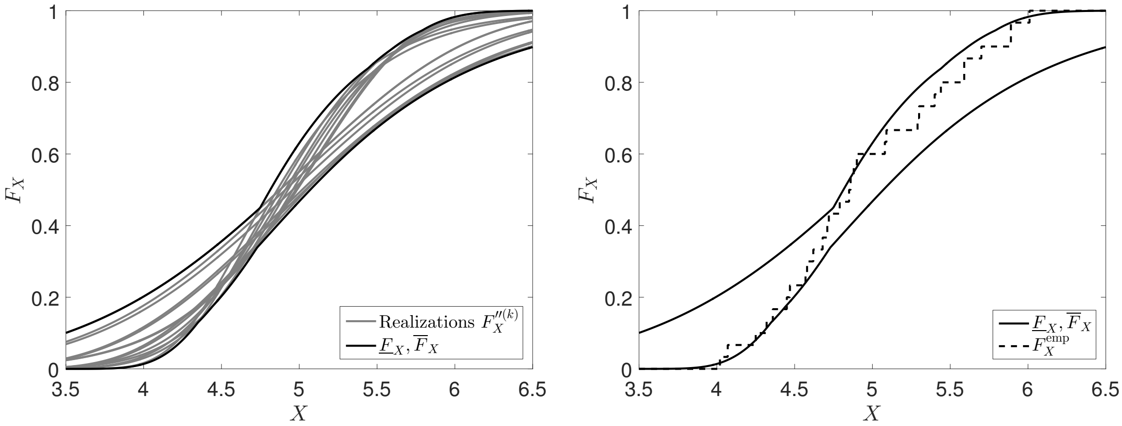
(b) One-sided Chebyshev's inequalities

Figure 2.8: Dataset $\mathcal{X}_1$ – free p-boxes – Chebyshev's inequalities.

Chebyshev's inequalities lead to wider, but smoother p-box bounds than the KS-confidence method around the mean value of the unknown distribution. In the tails, however, the behaviour is the contrary: the KS-confidence method leads to wider bounds in the tail compared to the Chebyshev's inequalities. In fact, the KS-confidence bounds do not converge in the tails of the distributions, as seen by the constant value of $\underline{F}_X(x)$ for $x > 6$ in Figure 2.7.

For the robust Bayes' method, the distribution families of interest are the Gaussian, lognormal, Weibull, and Gumbel distribution. The hyper-parameters of interest are the mean value μ_X and the standard deviation σ_X. The set of prior distributions for the mean value consists of a uniform distribution within the range $\mu_X \in [0, 10]$ and a Gaussian distribution with $\{\mu_{\mu_X} = 4, \sigma_{\mu_X} = 1\}$. The set of prior distributions for the standard deviation consists of a uniform distribution with $\sigma_X \in [0.1, 1]$ and a lognormal distribution with $\{\mu_{\sigma_X} = 0.6, \sigma_{\sigma_X} = 0.2\}$. For each combination of distribution function and prior distribution, the Bayesian inference is used to compute the (posterior-based) most likely value of $\{\mu_X, \sigma_X\}$. This results in a set of $n_b = 4 \cdot 2 \cdot 2 = 16$ values for $\{\mu_X, \sigma_X\}$ and hence to 16 distributions for X. Their CDFs and the resulting p-box boundary curves are illustrated in Figure 2.9.

The final boundary curves of the p-box envelope nicely the empirical CDF with tight bounds. Due to the small sample size of $\mathcal{X}_1$, the empirical CDF exceeds the bounds of the p-box locally. Note, however, that the tightness of the bounds depends on the number of candidate distribution families and prior distributions. When more candidates are available (*i.e.* $k \uparrow$), it is likely that the bounds are wider than shown in this example.



(a) Optimal CDFs and p-box boundaries

(b) P-box versus empirical CDF

Figure 2.9: Dataset $\mathcal{X}_1$ – free p-boxes – robust Bayes' method.

2.4.3 Data to parametric p-boxes

Confidence bounds

Considering the likelihood method in Section 2.1 and its generalized Bayesian inference algorithm in Section 2.2, a parametric p-box can be generated by determining confidence bounds onto the distribution parameters $\boldsymbol{\Theta}$, *i.e.* on $F''_{\boldsymbol{\Theta}}$. The interval-domain $\mathcal{D}_{\boldsymbol{\Theta}}$ (see Eq. (2.19)) is then constructed by cutting off the areas with low posterior density $f''_{\boldsymbol{\Theta}}$. Note that in the general case of an arbitrary distribution in $\boldsymbol{\Theta}$, $\mathcal{D}_{\boldsymbol{\Theta}}$ does not necessarily form a rectangular domain (see also Figure 2.3(a)).

In the special case of a Gaussian distribution with uniform prior on its mean value μ_X and variance σ_X^2, their confidence bounds are often assumed to be independent, forming a rectangular domain $\mathcal{D}_{\boldsymbol{\Theta}}$. The confidence intervals for μ_X and σ_X^2 are then calculated as follows (McClave & Sincish 2013):

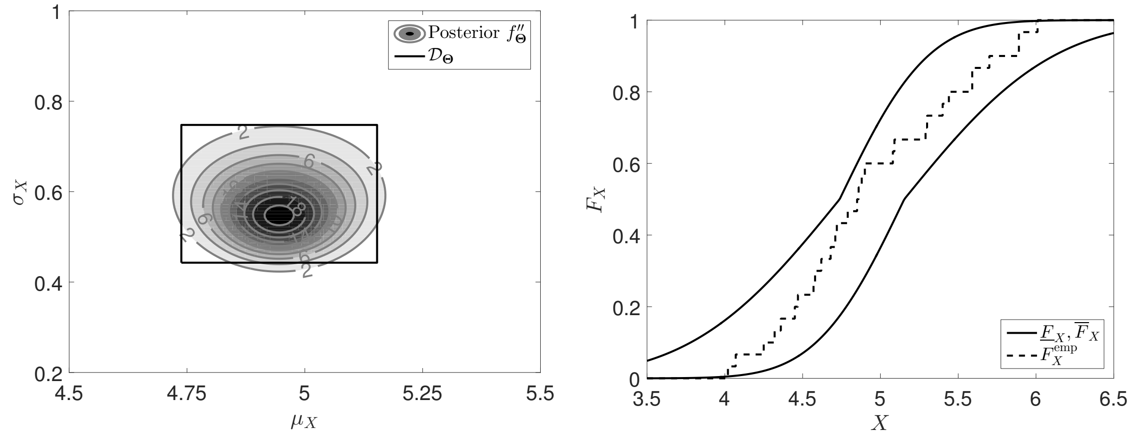
$$\mu_X \in \left[\mathbb{E}\left[\mathcal{X}\right] + \frac{t_{N-1,1-\gamma/2}}{\sqrt{N}}\mathrm{Std}\left[\mathcal{X}\right],\ \mathbb{E}\left[\mathcal{X}\right] - \frac{t_{N-1,1-\gamma/2}}{\sqrt{N}}\mathrm{Std}\left[\mathcal{X}\right]\right], \tag{2.29}$$

$$\sigma_X^2 \in \left[\frac{N-1}{\chi^2_{N-1,1-\gamma/2}}\mathrm{Var}\left[\mathcal{X}\right],\ \frac{N-1}{\chi^2_{N-1,\gamma/2}}\mathrm{Var}\left[\mathcal{X}\right]\right], \tag{2.30}$$

where $\mathrm{Std}\left[\mathcal{X}\right] = \sqrt{\mathrm{Var}\left[\mathcal{X}\right]}$ is the standard deviation of the dataset $\mathcal{X}$, $t_{N-1,1-\gamma/2}$ is the $(1-\gamma/2)$-th quantile of the t-distribution with $N-1$ degrees of freedom, $\chi^2_{N-1,\gamma/2}$ is the $(\gamma/2)$-th quantile of the χ^2 (chi-square) distribution with $N-1$ degrees of freedom, and $\gamma \in [0,1]$ is the confidence level.

Dataset $\mathcal{X}_1$

Considering $\mathcal{X}_1$ and the analysis in Section 2.2, the posterior likelihood $f''_{\mathbf{\Theta}}$ is shown in Figure 2.3(a) based on a uniform prior distribution $f'_{\mathbf{\Theta}}$. Using Eqs. (2.29) and (2.30), the confidence intervals for $\gamma = 0.05$ are $\mu_X \in [4.74, 5.15]$ and $\sigma_X = [0.44, 0.75]$. The corresponding $\mathcal{D}_{\mathbf{\Theta}}$ is illustrated in Figure 2.10(a) in comparison to the posterior distribution $f''_{\mathbf{\Theta}}$ of the Bayesian hierarchical model. The interval domain $\mathcal{D}_{\mathbf{\Theta}}$ incorporates the major part of the posterior probability density located around the true mean value/standard deviation.



(a) Interval domain $\mathcal{D}_{\mathbf{\Theta}}$

(b) P-box boundaries

Figure 2.10: Dataset $\mathcal{X}_1$ – parametric p-box – confidence-bounds-based p-box.

Figure 2.10(b) compares the parametric p-box boundaries with the empirical CDF. Also from this point of view, the obtained parametric p-box provides a conservative estimate for the true, but unknown distribution function F_X. Note that the boundary curves are obtained through Eq. (2.20).

2.4.4 Limitations

P-boxes provide an intuitive concept to distinguish between aleatory and epistemic uncertainty as seen by the definition of free and parametric p-boxes. However, the crucial point lies in the definition of the (absolute) bounds of the epistemic uncertainty. A particular case is illustrated by the study of the dataset $\mathcal{X}_1$, where no physical reasons or expert knowledge regarding bounds is available. Then, a practical solution is to define the bounds in terms of an appropriate confidence level, as seen in Section 2.4.3. The level is, however, a choice made by the modeller and should be considered carefully in the analysis and post-processing of such p-boxes.

2.5 Fuzzy variables

2.5.1 Definitions

Fuzzy variables are strongly related to the set theory. Hence, the basic definitions of sets are presented here for the sake of clarity and in coherency with Möller & Beer (2004). Considering the event space (*i.e.* universal set) Ω, a *membership function* $\xi_{\mathcal{E}}(\omega) \in [0, 1]$ describes whether an elementary event ω belong to a set $\mathcal{E}$. When the membership function is binary (*i.e.* $\{0, 1\}$), the set $\mathcal{E}$ is called a *crisp set*. When the membership function is not binary but continuous, $\mathcal{E}$ describes a *fuzzy set*.

In practical applications, continuous random variables are common. Hence, the fuzzy sets are adapted to the theory of continuous numbers. In particular, an interval is defined by a binary membership function analogously to a crisp set. Moreover, a *fuzzy number* is a convex fuzzy set whose membership function is at least segmentally continuous and $\xi_{\mathcal{E}}(x) = 1$ at a single value $x = x^*$. This point x^* is also called the mean value of the fuzzy number. When $\xi_{\mathcal{E}}(x) = 1$ for more than a single point $x \in \mathcal{D}_X$, then $\mathcal{E}$ describes a *fuzzy interval.* In both cases, the support is defined as the interval where $\xi_{\mathcal{E}}(x) > 0$, $x \in \mathcal{D}_X$.

The membership function of an interval, a fuzzy number, and a fuzzy interval in the domain $\mathcal{D}_X$ are shown in Figure 2.11. The interval has a membership value of $\xi_{\mathcal{E}} = 1$ for $x \in [2, 4]$ and 0 otherwise. The fuzzy number is illustrated here by a triangular membership function whose mean value is $x = 2.7$. The support of the fuzzy number is $[0.3, 5.2]$. The third curve in Figure 2.11 is a fuzzy interval where the membership function reads $\xi_{\mathcal{E}} = 1$ for the interval $x \in [1.5, 3.0]$ and has a support of $[0.5, 5.0]$.

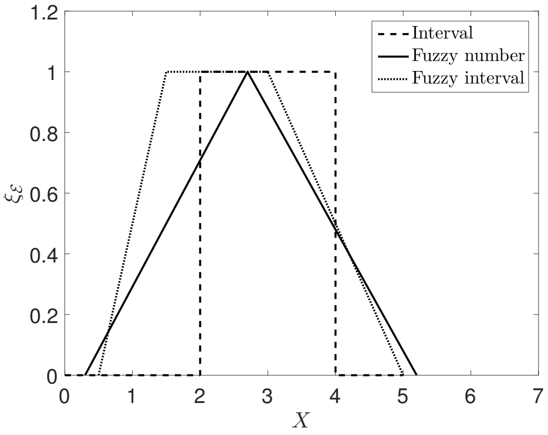


Figure 2.11: Fuzzy numbers for event $\mathcal{E}$ – membership functions of interval, fuzzy number and fuzzy interval.

For the further analysis of fuzzy numbers, the α-level interval is defined as:

$$\mathcal{E}^{(\alpha)} = \{x \in \mathcal{D}_X : \ \xi_{\mathcal{E}}(x) \geq \alpha\}\,, \tag{2.31}$$

where $\alpha \in [0, 1]$ is the α-cut level. Note that the α-cuts provide nested intervals, as $\mathcal{E}^{(\alpha_i)} \subseteq \mathcal{E}^{(\alpha_j)}$ when $\alpha_i \geq \alpha_j$ for all $\alpha_i, \alpha_j \in [0, 1]$. The case of $\alpha = 0^+$ describes the support of the fuzzy number/interval.

2.5.2 Fuzzy distributions

General idea

Fuzziness can be interpreted as the imprecision in measurements, which is a generalization of the description of epistemic uncertainty modelled by intervals (*i.e.* crisp sets). Then, the membership function quantifies the precision of the measurements. When $\alpha = 1$, the measurement is as precise as possible (however, $\mathcal{E}^{(\alpha=1)}$ can still be an interval due to epistemic uncertainty). When $\alpha \approx 0^+$, the measurement is very uncertain and $\mathcal{E}^{(\alpha\approx 0^+)}$ denotes the widest possible interval for event $\mathcal{E}$.

Fuzzy probability theory is a combination of fuzzy set theory and probability theory (Beer, Ferson, & Kreinovich 2013). In fact, a *fuzzy distribution* is defined as a probability distribution whose CDF $F_X(x)$ is a fuzzy interval for every $x \in \mathcal{D}_X$:

$$F_X(x) \in \Xi_X(x,c), \tag{2.32}$$

where $\Xi_X(x,c)$ denotes the membership function for variable X at $x \in \mathcal{D}_X$ and CDF value $c \in [0,1]$. $\Xi_X(x,c)$ can be defined in different ways, including as a generalization of free and parametric p-boxes.

Fuzzy distribution: a generalization of a free p-box

Considering the free p-box defined in Section 2.4, $\underline{F}_X(x)$ and $\overline{F}_X(x)$ describe an interval for $F_X(x)$, which corresponds to a membership function with $\Xi_X(x,c) = 1$ if $c \in \left[\underline{F}_X(x), \overline{F}_X(x)\right]$, and $\Xi_X(x,c) = 0$ otherwise, for all $x \in \mathcal{D}_X$. In analogy with Eq. (2.31), every α-cut of $\Xi_X(x,c)$ leads to the following free p-box bounds:

$$\underline{F}_X^{(\alpha)}(x) = c^*, \qquad c^* = \arg\min_{c\in[0,1]} \Xi_X(x,c) \geq \alpha, \tag{2.33}$$

$$\overline{F}_X^{(\alpha)}(x) = c^*, \qquad c^* = \arg\max_{c\in[0,1]} \Xi_X(x,c) \geq \alpha. \tag{2.34}$$

When generalizing the membership function to describe a fuzzy interval, a fuzzy distribution is generated. Note that Eqs. (2.33) and (2.34) hold for this general case too: each α-cut leads to a free p-box with bounds $\underline{F}_X^{(\alpha)}$ and $\overline{F}_X^{(\alpha)}$. In fact, Eqs. (2.33) and (2.34) lead to nested free p-boxes with $\underline{F}_X^{(\alpha_i)}(x) \leq \underline{F}_X^{(\alpha_j)}(x)$ and $\overline{F}_X^{(\alpha_i)}(x) \geq \overline{F}_X^{(\alpha_j)}(x)$ for $\alpha_i < \alpha_j$ and any $x \in \mathcal{D}_X$.

Figure 2.12 illustrates the construction of a fuzzy distribution as a generalization of a free p-box. For a given value $x^{(0)}$, the associated membership function is shown in red. Further, for a given level $\alpha^{(0)}$, the lower and upper bounds of the CDF are obtained analytically by Eqs. (2.33) and (2.34). This procedure can be repeated for every value $x \in \mathcal{D}_X$.

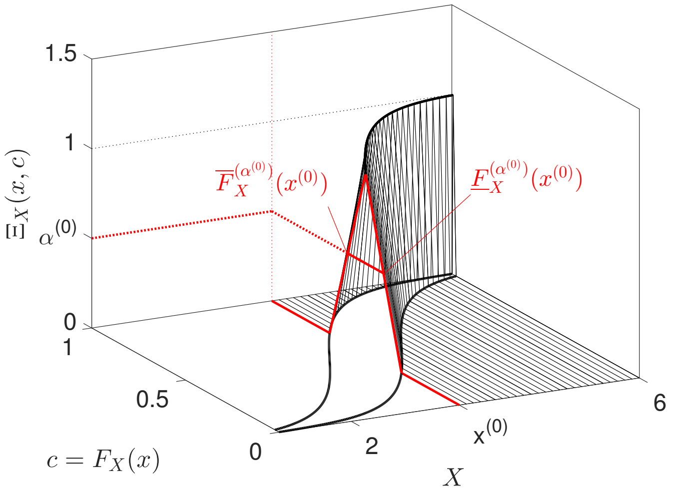


Figure 2.12: Fuzzy distribution – illustration of Eqs. (2.33) and (2.34).

Fuzzy hierarchical model: a generalization of a parametric p-box

Considering again the parametric p-box defined in Eq. (2.19), the interval domain $\mathcal{D}_{\boldsymbol{\Theta}}$ can be interpreted as a special case of a membership function for the parameters $\boldsymbol{\Theta}$. Then, $\xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$

describes the fuzziness of the hyper-parameters $\boldsymbol{\Theta}$, where $\xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = 1$ for $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$ and $\xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = 0$ otherwise.

The generalization of the membership function to describing a fuzzy interval leads to a second definition of a fuzzy variable (analogous to Eq. (2.19)):

$$F_X(x) = F_X(x|\boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}} : \ \xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) > \alpha\}\,. \tag{2.35}$$

Again, each α-cut leads to a parametric p-box with interval-valued hyper-parameters in $\mathcal{D}_{\boldsymbol{\Theta}}^{(\alpha)} = \{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}} : \ \xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) > \alpha\}$. Note that $\xi_{\boldsymbol{\Theta}}$ describes a $n_{\boldsymbol{\Theta}}$-dimensional fuzzy interval.

2.5.3 Data to fuzzy distribution

A variety of methods are available in the literature to define fuzzy numbers and fuzzy distributions from data. The optimal strategy greatly depends on the amount and type of knowledge available to the modeller. Hence, in the following, only two approaches are presented which relate to the dataset $\mathcal{X}_1$. In the first approach fuzziness is used to imprecise measurements and in the second approach it is used to model different confidence levels. For further methods, the reader is referred to *e.g.* Möller & Beer (2004), Klir (2006), Beer (2009), Beer, Mingqiang, et al. (2011), Beer, Y. Zhang, et al. (2012), and Beer, Ferson, & Kreinovich (2013).

Fuzzy measurements

Let us assume that the measurement device has an unknown but limited precision, which can be modelled by a membership function. In order to determine the corresponding fuzzy distribution, free p-boxes are generated at different values of α and subsequently assembled.

For a given α value, Eq. (2.31) is evaluated for each sample $\chi^{(i)}, \ i = 1, \ldots, N$, leading to a set of intervals:

$$\chi^{(i)(\alpha)} \in \left[\underline{\chi}^{(i)(\alpha)}, \overline{\chi}^{(i)(\alpha)}\right], \tag{2.36}$$

$$\underline{\chi}^{(i)(\alpha)} = \arg \min_{x \in \mathcal{D}_X} \xi_{\chi^{(i)}}(x) \geq \alpha, \tag{2.37}$$

$$\overline{\chi}^{(i)(\alpha)} = \arg \max_{x \in \mathcal{D}_X} \xi_{\chi^{(i)}}(x) \geq \alpha, \tag{2.38}$$

where $\xi_{\chi^{(i)}}$ denotes the membership function of sample $\chi^{(i)}$. Then, the free p-box for α is obtained by using Eqs. (2.15) and (2.16):

$$\underline{F}_X^{(\alpha)}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\overline{\chi}^{(i)(\alpha)} \leq x}(x)\,, \tag{2.39}$$

$$\overline{F}_X^{(\alpha)}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\underline{\chi}^{(i)(\alpha)} \leq x}(x)\,, \tag{2.40}$$

where N is the number of samples in the dataset.

This simple approach follows the same methodology as in Section 2.3.2, which only uses the information contained in the dataset to model a CDF. At each α-cut, the belief and plausibility measures of the empirical CDF are constructed. However, a more sophisticated approach could be applied to determine the bounds of the free p-box.

Variable confidence levels

Consider again Section 2.4.3, where a parametric p-box is determined through fixed confidence intervals (CI) on the posterior distribution of the hyper-parameters $\boldsymbol{\Theta}$. The choice of the confidence level is crucial for the size of the interval domain $\mathcal{D}_{\boldsymbol{\Theta}}$ and hence the width of the parametric p-box. In this case, fuzzy distributions allow for more flexibility by varying the CI and hence by defining a membership function for $\boldsymbol{\Theta}$ dependent on this CI.

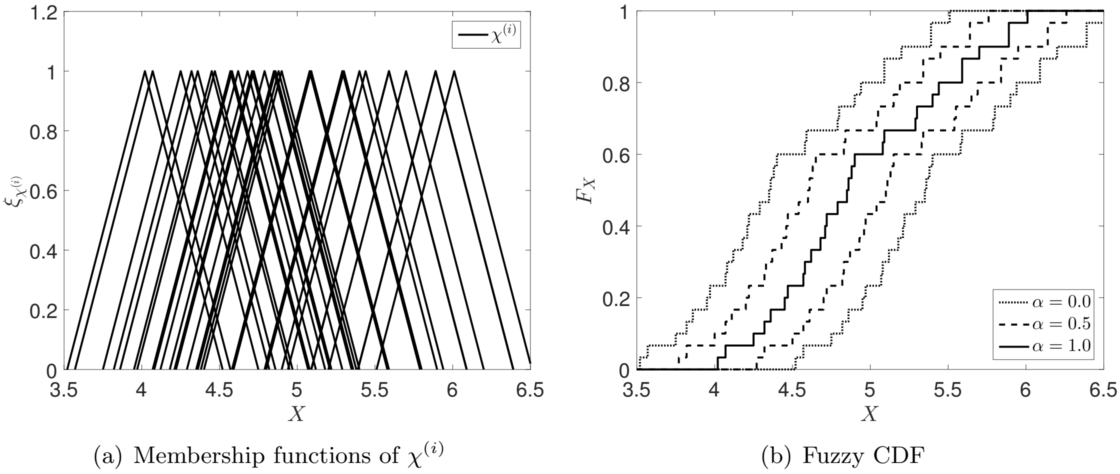
The membership function $\xi_{\boldsymbol{\Theta}}$ can be defined indirectly through its α-cuts:

$$\{\boldsymbol{\theta}:\ \xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \geq \alpha\} = \left\{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}^{(\gamma)}\right\}, \tag{2.41}$$

where $\mathcal{D}_{\boldsymbol{\Theta}}^{(\gamma)}$ denotes the interval domain corresponding to a confidence level γ. Then, a function $f_{\alpha\gamma}$ is required to map α to γ to define the membership function $\xi_{\boldsymbol{\Theta}}$. A number of such functions can be proposed, *i.e.* there is no unique choice.

Dataset $\mathcal{X}_1$

In a first approach, it is assumed that the samples of the dataset $\mathcal{X}_1$ can be modelled by a triangular membership function $\xi_{\chi^{(i)}}$ with mean value $\chi^{(i)}$ and support interval $\left[\chi^{(i)} - 0.5, \chi^{(i)} + 0.5\right]$. The membership functions are shown graphically in Figure 2.13(a), where $\xi_{\chi^{(i)}}(x)$ are plotted as a function of x for $i = 1, \ldots, N$. Each triangular shape corresponding to a sample point $\chi^{(i)}$.



(a) Membership functions of $\chi^{(i)}$ (b) Fuzzy CDF

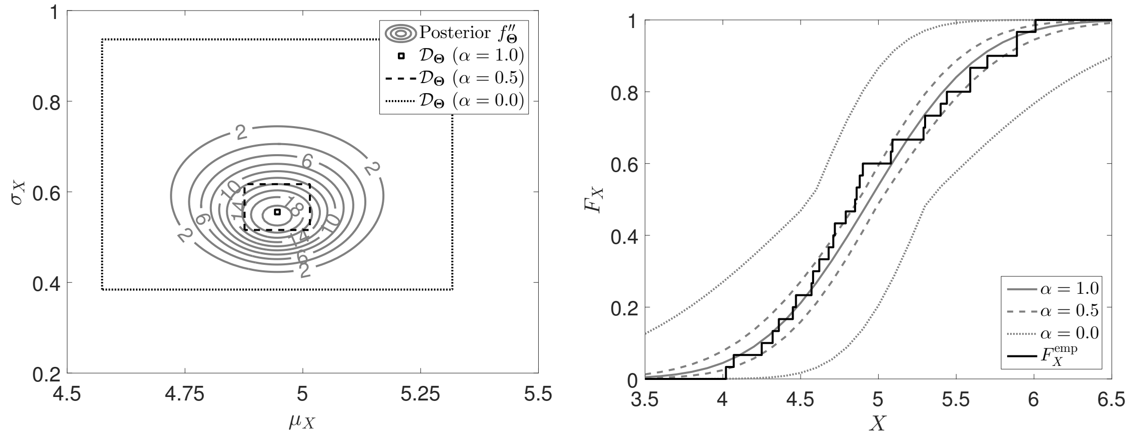
Figure 2.13: Dataset $\mathcal{X}_1$ – fuzzy distributions – measurement fuzziness.

Figure 2.13(b) illustrates the corresponding fuzzy distribution by showing the boundary curves of the free p-boxes at $\alpha = \{0.0, 0.5, 1.0\}$. When $\alpha = 1.0$, the free p-box reduces to the empirical CDF found in Figure 2.1(b). At this level, the fuzzy intervals reduce to the actual measurements $\chi^{(i)}$. On the other side, $\alpha = 0.0$ leads to the same boundary curves as in Figure 2.4, because of the identical support in both application examples.

In order to apply the second approach, which deals with variable confidence levels (Eq. (2.41)) and no measurement imprecision, the function $f_{\alpha\gamma}$ is defined:

$$\gamma = f_{\alpha\gamma}(\alpha) = \alpha, \tag{2.42}$$

which maps the α value to the confidence level γ. Figure 2.14(a) shows the interval domain $\mathcal{D}_{\boldsymbol{\Theta}}^{(\gamma)}$ corresponding to $\alpha = \{0.0, 0.5, 1.0\}$ and in comparison to the posterior likelihood distribution



(a) Two-dimensional membership functions and interval domains $\mathcal{D}_{\Theta}$

(b) Fuzzy CDF

Figure 2.14: Dataset $\mathcal{X}_1$ – fuzzy distributions – fuzziness due to confidence level.

obtained in Figure 2.10(a). Naturally, the lower α, the larger is $\mathcal{D}_{\Theta}$. Note that the minimum value of α is chosen as 0.001 (instead of 0.0) in order to avoid infinitely wide domains.

Figure 2.14(b) shows the corresponding fuzzy CDF. In particular, the boundary curves of the same α-cuts are compared to the empirical CDF of $\mathcal{X}_1$. Again, the lower α, the wider are the parametric p-box bounds. Hence, the lower α-cuts lead to more conservative estimates of the CDF. Note that despite the linear function between α and γ, the width of the α-cut parametric p-boxes is highly nonlinear. $\alpha = 0.5$ results in bounds that are close to the bounds corresponding to $\alpha = 1.0$, whereas $\alpha = 0.0$ results in much wider bounds.

2.6 Further concepts

Apart from the concepts presented for uncertainty modelling, further methodologies are discussed in the literature, including possibility theory, clouds, credal sets, info-gap theory and lack-of-knowledge theory. These methods are related to the presented uncertainty concepts but are not discussed here in detail for the sake of brevity.

Possibility theory defines two specifications of the likelihood of events, namely *necessity* (lower bound) and *possibility* (Dubois & Prade 1988, 1998; De Cooman, Ruan, & Kerre 1995; De Cooman 1997). The possibility theory is closely related to Dempster-Shafer's theory of evidence, which also provides two measures of likelihood (Guan & Bell 1991).

Neumaier (2004) introduced *clouds*, which is somewhere between fuzzy sets and probability distributions and which is related to the interval-valued fuzzy sets (Dubois & Prade 2005). In the context of clouds, a fuzzy interval is modelled by a pair of membership functions, which lead to a lower an upper α-cut. Further details can be found in *e.g.* Kozine & Utkin (2005), Destercke, Dubois, & Chojnacki (2008a,b), and Fuchs & Neumaier (2009).

Info-gap theory (Ben-Haim 2006) can be interpreted as the concept of fuzzy intervals, where the membership function is defined in a different way (Ferson & Tucker 2008). Further details and applications can be found in *e.g.* Kanno & Takewaki (2006a,b) and Pierce, Worden, & Manson (2006). A closely related concept is also the *lack-of-knowledge theory* developed by Barthe et al. (2003) and Ladevèze, Puel, & Romeuf (2006).

2.7 Comparison

The notations of the various concepts presented in this thesis are summarized in Table 2.1. For each concept, the table presents the uncertainty model as well as its aleatory and epistemic portion of uncertainty. Note that the epistemic uncertainty is split into interval-valued and fuzzy set-valued methods to distinguish the complexity of modelling. For completeness, simple constants, intervals, and a fuzzy interval/number are added. The interval is an example of a purely interval-valued epistemic quantity. In this compilation, the equivalence of Dempster-Shafer structure and free p-boxes is clearly visible since $\underline{F}_X(x) = \mathrm{Bel}\,(X \leq x)$ and $\overline{F}_X(x) = \mathrm{Pl}\,(X \leq x)$.

A distinction is made between hierarchical and non-hierarchical models, as indicated by the last column in Table 2.1. The three hierarchical concepts allocate epistemic uncertainty on the definition of the hyper-parameters $\boldsymbol{\theta}$. This allows for a clear separation of the three aspects of uncertainty: aleatory uncertainty models the variability in X, epistemic uncertainty models the non-determinacy of $\boldsymbol{\Theta}$, and fuzziness models the variation of epistemic uncertainty. In contrast, the non-hierarchical models mix the different aspects of uncertainty in a single modelling level. The distinction is not as straightforward in this case.

The connections of the different concepts in visualized in Figure 2.15. Each concept is represented by a rectangular box. Note that free p-boxes share a box with Dempster-Shafer (DS) structures due to their equivalence. The arrows indicate the simplifications in the uncertainty model. The figure is split into four levels. The highest level shows the most general uncertainty representations considered in this thesis, *i.e.* models with all three aspects of uncertainty. When heading to a lower level, an aspect of uncertainty modelling is removed. As an example, a fuzzy distribution reduces to a free p-box when removing fuzziness from the uncertainty model. The free p-box further reduces to an interval when removing aleatory uncertainty. Finally, the interval reduces to a constant when neglecting any epistemic uncertainty.

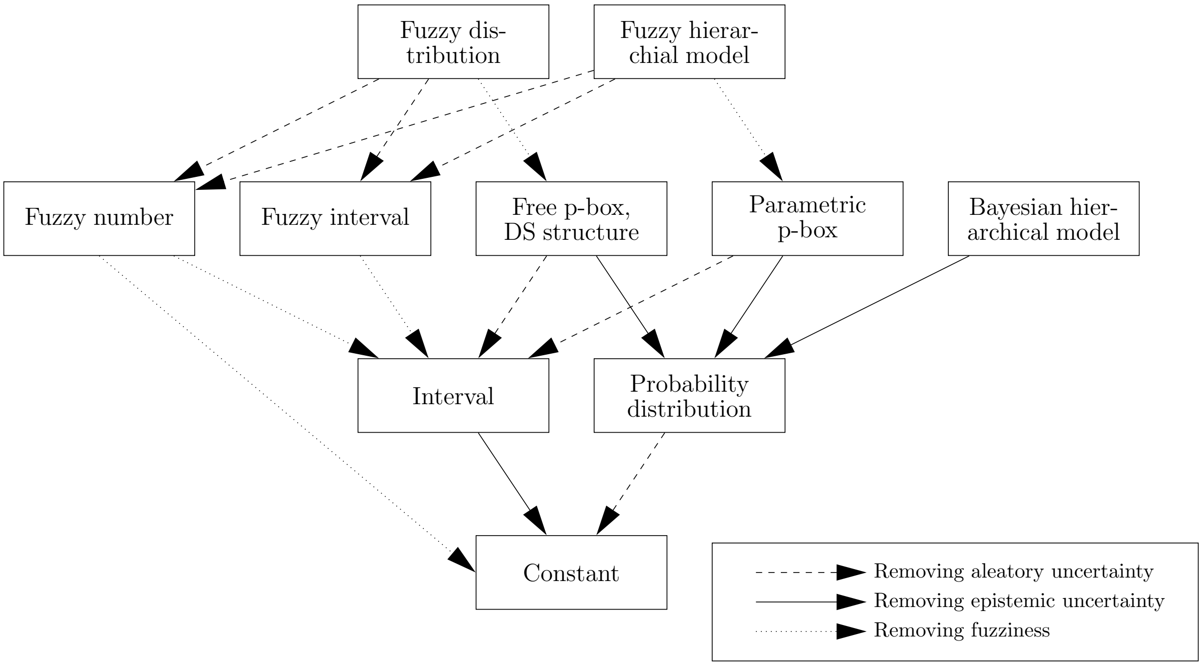


Figure 2.15: Relations between different concepts of uncertainty modelling.

Going upwards in Figure 2.15 leads to more general and more complex uncertainty represen-

Table 2.1: Characteristics of different concepts of uncertainty modelling for a variable X.

Concept	Uncertainty model	Aleatory	Epistemic uncertainty		Hierarchical
			Interval-valued (crisp sets)	Fuzzy sets	
Constant	x	-	-	-	
Probability theory	$F_X(x)$	$X \sim F_X(x)$	-	-	
Bayesian hierarchical model	$F_X(x\|\boldsymbol{\theta})$, $F_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$	$X \sim F_X\left(x\|\boldsymbol{\theta}\right)$	$\boldsymbol{\Theta} \sim F_{\boldsymbol{\Theta}}\left(\boldsymbol{\theta}\right)$	-	✓
Interval	$x \in [\underline{x}, \overline{x}]$	-	$x \in [\underline{x}, \overline{x}]$	-	
Dempster-Shafer structure	$\mathrm{Bel}\left(X \leq x\right)$, $\mathrm{Pl}\left(X \leq x\right)$	$X \sim F_X\left(x\right)$	$\mathrm{Bel}\left(X \leq x\right) \leq F_X(x) \leq \mathrm{Pl}\left(X \leq x\right)$	-	
Free p-box	$\underline{F}_X(x)$, $\overline{F}_X(x)$	$X \sim F_X\left(x\right)$	$\underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x)$	-	
Parametric p-box	$F_X(x\|\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$	$X \sim F_X\left(x\|\boldsymbol{\theta}\right)$	$\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$	-	✓
Fuzzy number/interval	$\xi_X(x)$	-	$x \in \mathcal{D}_X^{(\alpha)}$	$\xi_X(x)$	
Fuzzy distribution	$\Xi_X(x, c)$	$X \sim F_X\left(x\right)$	$\underline{F}_X^{(\alpha)}(x) \leq F_X\left(x\right) \leq \overline{F}_X^{(\alpha)}(x)$	$\Xi_X(x, c)$	
Fuzzy hierarchical model	$F_X(x\|\boldsymbol{\theta})$, $\xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$	$X \sim F_X\left(x\|\boldsymbol{\theta}\right)$	$\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}^{(\alpha)}$	$\xi_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$	✓

tations. As an example, allowing for epistemic uncertainty transforms a probability distribution into either a free p-box, a parametric p-box or a Bayesian hierarchical model. Generalizing the parametric p-box leads to a fuzzy hierarchical model.

2.8 Conclusion

The concepts presented in this chapter give an overview of uncertainty models with varying complexity and with varying information content. In practical applications, however, a choice must be made for a single uncertainty model or a combination of suitable models. This choice should be driven by the available information on the one side and by the purpose of the analysis on the other side (Beer, Ferson, & Kreinovich 2016). The type and amount of information can be manifold as discussed in Chapter 1: from pure datasets to pure expert knowledge or a combination of multiple sources of information. The same diversity can be imagined for the purpose of the analysis. As an example, the purpose of the analysis could be the estimation of sensitivities in early design stage to understand the impact of input uncertainties to the output. Another case would be to estimate the failure probability with respect to prescribed code requirements.

Conventionally, probability theory is used to cope with uncertainty. After a comprehensive review, the p-box model is retained in this thesis, particularly the two concepts of free and parametric p-boxes, due to the following arguments:

- The p-box concept allows for a natural extension of the conventional probability theory. Apart from aleatory uncertainty, which is related to probability theory, p-boxes allow to include epistemic uncertainty in the uncertainty model.
- A common situation in practice is to have limited information due to resource constraints (*e.g.* scarce datasets, imprecise measurements, distribution bounds). Then, p-boxes allow one to model the lack of knowledge by considering the set-theoretical approach: epistemic uncertainty is represented by interval-valued quantities.
- The p-box concept allows for a clear distinction of aleatory and epistemic uncertainty, which makes it intuitive to understand and interpret. This is crucial when dealing with engineers and decision makers in practice, where the communication of uncertainty concepts is challenging.
- The concepts of free and parametric p-boxes cover the two typical scenarios of non-hierarchical and hierarchical models, respectively (see also Table 2.1). This allows for a flexible modelling of uncertainty and the applicability to diverse problems.
- Last but not least, p-boxes are a special case of fuzzy distributions and fuzzy hierarchical models. Hence, the approaches presented in the following chapters can also be applied to more general cases including fuzziness with only little modifications.

Hence on the one side, p-boxes provide a versatile concept of uncertainty quantification that goes beyond the conventional probability theory but does not touch the realm of fuzzy sets. On the other side, p-boxes allow for modelling a mix of aleatory and epistemic uncertainty. Considering the aspect of modelling complexity and modelling generality, p-boxes maintain a good trade-off.

In the remainder of this thesis, p-boxes are used for uncertainty quantification analyses. In particular, specific algorithms are developed for the two types of p-boxes, namely free and parametric p-boxes.

CHAPTER 3

Classical meta-models

This chapter provides an overview of state-of-the-art meta-modelling techniques, in particular polynomial chaos expansions and Kriging. The main parts of this chapter have been published in Schöbi, Sudret, & Wiart (2015), Schöbi & Sudret (2014a,d).

3.1 State of the art

Propagating uncertainties usually requires a large number of repeated calls to the model for different values of the input parameters, for instance through a Monte Carlo simulation procedure. Such an approach requires thousands to millions of runs which is not affordable in many practical cases even with high-performance computing architectures. To circumvent this problem, *surrogate models* may be used, which replace the original computational model by an easy-to-evaluate function (Hastie, Tibshirani, & Friedman 2001; Forrester, Sóbester, & Keane 2008; Storlie et al. 2009). These surrogate models, also known as *response surfaces* or *meta-models*, are capable of quickly predicting responses to new input realizations. This allows for conducting analyses which require a large number of model evaluations, such as simple uncertainty propagation, structural reliability analysis and design optimization, in a reasonable time.

Among the various options for constructing meta-models, this thesis focuses on *non-intrusive approaches*, meaning that only the input vector (*i.e.* a realization of the random input in the process of uncertainty propagation) is selected, the model is run and the output vector of quantities of interest (QoI) is collected. No additional knowledge on the inner structure of the computer code is accessible and/or considered. Popular types of meta-models include polynomial chaos expansions (PCE) (Ghanem & Spanos 2003), Kriging (also known as Gaussian process modelling) (Sacks et al. 1989; Santner, B. Williams, & Notz 2003; Rasmussen & C. Williams 2006), support vector machines (Gunn 1998; Vazquez & Walter 2003; Clarke, Griebsch, & Simpson 2003; Smola & Schölkopf 2006), and artificial neural networks (Hurtado & Alvarez 2001; Schueremans & Van Gemert 2005a; Cartwright 2015), which have been extensively investigated in the last decade. In this thesis, the focus lies on PCE and Kriging.

PCE, also known as *spectral expansions*, approximate the computational model by a series of multivariate polynomials which are orthogonal with respect to the distributions of the input random variables. Traditionally, spectral expansions have been used to solve partial differential equations in an *intrusive* manner (Ghanem & Spanos 2003). In this setup, truncated expansions are inserted into the governing equations and the expansion coefficients are obtained using a Galerkin scheme. The pioneering approach called *spectral stochastic finite element method* (SSFEM) was later developed by Xiu & Karniadakis (2002), Sudret, Defaux, & Pendola (2004), Wan & Karniadakis (2005, 2006), and Berveiller, Sudret, & Lemaire (2006a), among others. These intrusive methods require specific, problem-dependent algorithmic developments though. Because it is not always possible and/or feasible to treat a computational model intrusively, especially when legacy codes are at hand in *e.g.* an industrial context, non-intrusive PCE were developed. So-called *projection methods* were developed by Ghiocel & Ghanem (2002), Le Maître et al. (2002), Keese & Matthies (2005), and Xiu & Hesthaven (2005), see Xiu (2009) for a comprehensive review. *Least-squares minimization* techniques have been introduced by Choi et al. (2004) and Berveiller, Sudret, & Lemaire (2006b) based on the pioneering work of Tatang et al. (1997) and Isukapalli (1999). Further developments which combine spectral expansions and compressive sensing ideas have lead to so-called *sparse PCE* (Blatman & Sudret 2008, 2010a,b, 2011; Doostan & Owhadi 2011; Doostan, Validi, & Iaccarino 2013; Jakeman, Eldred, & Sargsyan 2015). Sparse PCE is the approach discussed in this thesis (see also Section 3.3). Recent applications of sparse PCE in the context of structural reliability analysis and design optimization can be found in Eldred, Webster, & Constantine (2008), Eldred (2009), and Sarangi et al. (2014). A state-of-the-art implementation of sparse PCE is available in the MATLAB-based framework UQLAB (Marelli & Sudret 2014, 2015).

The second meta-modelling technique of interest in this thesis is Kriging, which originates from geographical data in mining (Krige 1951) and is today also known as Gaussian process modelling (Santner, B. Williams, & Notz 2003; Rasmussen & C. Williams 2006). The Kriging meta-model is interpreted as a realization of a Gaussian process. The latest developments of Kriging are contributed in the aspects of optimal estimation of model parameters (called hyper-parameters in Section 3.4) (Bachoc 2013; Bachoc et al. 2014; Bachoc 2014), the use of adaptive kernels (Duvenaud, Nickisch, & Rasmussen 2012; Ginsbourger, Roustant, & Durrande 2013), and the use of additive kernels (Durrande, Ginsbourger, & Roustant 2012, 2013; Ginsbourger, Durrande, & Roustant 2013). Practical applications can be found in many fields, such as structural reliability analysis (Kaymaz 2005; Bichon, Eldred, et al. 2008; Echard, Gayton, & Lemaire 2011; Bect et al. 2012; Dubourg, Sudret, & Deheeger 2013; Dubourg & Sudret 2014; Schöbi, Sudret, & Marelli 2016), and design optimization (Jones, Schonlau, & Welch 1998; Dubourg, Sudret, & Bourinet 2011; Dubourg 2011; Moustapha et al. 2016). The implementation of the Kriging meta-modelling techniques can be found in a variety of software including the MATLAB implementation (The MathWorks Inc. 2016), the MATLAB-based DACE toolbox (Lophaven, Nielsen, & Sondergaard 2002), the MATLAB-based UQLAB framework (Lataniotis, Marelli, & Sudret 2015; Marelli, Lamas-Fernandes, et al. 2015), and the R-based DICEKRIGING package (Roustant, Ginsbourger, & Deville 2012, 2013).

3.2 Computational model

Consider a system whose behaviour is represented by a computational model $\mathcal{M}$ which maps the M-dimensional input space to the one-dimensional output space:

$$\mathcal{M}: \ \boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} \subset \mathbb{R}^M \ \mapsto \ y \in \mathbb{R}, \tag{3.1}$$

where $\boldsymbol{x} = (x_1, \ldots, x_M)^{\mathsf{T}}$. In other words, the computational model $\mathcal{M}$ is a function from $\boldsymbol{x}$ to y:

$$y = \mathcal{M}(\boldsymbol{x}). \tag{3.2}$$

The computational model considered here provides a *deterministic* mapping from the input to the output space, *i.e.* repeated evaluations with the same $\boldsymbol{x}^{(0)} \in \mathcal{D}_{\boldsymbol{X}}$ lead to the same output value $y^{(0)} = \mathcal{M}\left(\boldsymbol{x}^{(0)}\right)$. Further, it is assumed that $\mathcal{M}$ is a *black box*, *i.e.* for each input vector $\boldsymbol{x}$, only the corresponding response $y = \mathcal{M}(\boldsymbol{x})$ is accessible. In other words, the inner structure and mechanisms of the computational model are not observable. This is typically the case for *e.g.* a finite element model (FEM) in which, generally speaking, the governing equations cannot be solved analytically.

As the input vector $\boldsymbol{x}$ is assumed to be affected by uncertainty, an uncertainty model is introduced. In the context of most classical meta-modelling techniques, probability theory is used. Hence, due to uncertainties in the input vector, it is represented by a random vector $\boldsymbol{X}$ with given joint probability density function (PDF) $f_{\boldsymbol{X}}$ and corresponding joint cumulative distribution function (CDF) $F_{\boldsymbol{X}}$. For the sake of simplicity, the components of $\boldsymbol{X}$ are assumed independent throughout this thesis, so that the joint PDF may be written as the product of the marginal PDFs denoted by $f_{X_i},\ i = 1, \ldots, M$. Note that the case of dependent input variables can be addressed by using an isoprobabilistic transform first, such as the Nataf or Rosenblatt transform (Lemaire 2009; Lebrun & Dutfoy 2009a,b). The output of the model is a random variable Y obtained by propagating the input uncertainty in $\boldsymbol{X}$ through the computational model $\mathcal{M}$:

$$Y = \mathcal{M}(\boldsymbol{X}). \tag{3.3}$$

3.3 Polynomial Chaos Expansions

3.3.1 Definition

Provided that the output random variable Y is a second-order variable (*i.e.* $\mathbb{E}\left[Y^2\right] < +\infty$), the computational model $\mathcal{M}$ can be cast as the following polynomial chaos expansion (PCE) (Ghanem & Spanos 2003; Soize & Ghanem 2004):

$$Y = \mathcal{M}(\boldsymbol{X}) \equiv \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} a_{\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}), \tag{3.4}$$

where $\left\{a_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \mathbb{N}^M\right\}$ are coefficients of the multivariate orthonormal polynomials $\psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$ which are built in coherency with the distribution of the input random vector $\boldsymbol{X}$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M)$ is the multi-index and M is the number of input variables (dimensions). Since the components of $\boldsymbol{X}$ are independent, the joint PDF is the product of its marginals. Then, a *functional inner product* for each marginal PDF f_{X_i} is defined by:

$$\langle \phi_1, \phi_2 \rangle_i = \int_{\mathcal{D}_{X_i}} \phi_1(x)\ \phi_2(x)\ f_{X_i}(x)\ \mathrm{d}x, \tag{3.5}$$

for any two functions $\{\phi_1, \phi_2\}$ such that the integral exists. For each variable X_i, an orthonormal polynomial basis can be constructed which satisfies (Xiu & Karniadakis 2002):

$$\langle P_j^{(i)}, P_k^{(i)} \rangle = \int_{\mathcal{D}_{X_i}} P_j^{(i)}(x)\ P_k^{(i)}(x)\ f_{X_i}(x)\ \mathrm{d}x = \delta_{jk}, \tag{3.6}$$

where $\left\{P_j^{(i)}, P_k^{(i)}\right\}$ are two candidate univariate polynomials in the i-th variable, $\mathcal{D}_{X_i}$ is the support of the random variable X_i, and δ_{jk} is the Kronecker symbol with is equal to 1 for $j = k$ and equal to 0 otherwise. Xiu & Karniadakis (2002) summarize bases for some classical PDFs.

The multivariate polynomials in Eq. (3.4) are then composed of univariate polynomials by tensor product, *i.e.* by multiplying the various polynomials in each input variable:

$$\psi_{\boldsymbol{\alpha}}\left(\boldsymbol{X}\right) = \prod_{i=1}^{M} \psi_{\alpha_i}^{(i)}\left(X_i\right), \tag{3.7}$$

where $\psi_{\alpha_i}^{(i)}$ is the polynomial of degree α_i in the i-th variable.

3.3.2 Truncation schemes

The main idea of PCE is to surrogate the computational model by an infinite series of polynomials as seen in Eq. (3.4). In practice however, it is not feasible to handle infinite series, thus the need for a truncation scheme. Such a truncation scheme corresponds to a set of multi-indices $\boldsymbol{\alpha} \in \mathcal{A} \subset \mathbb{N}^M$ such that the system response is accurately approximated with respect to some error measure (Blatman & Sudret 2010a, 2011):

$$Y \approx Y^{(\mathrm{PCE})} \overset{\mathrm{def}}{=} \mathcal{M}^{(\mathrm{PCE})}\left(\boldsymbol{X}\right) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}\left(\boldsymbol{X}\right). \tag{3.8}$$

There are several ways to select a priori a truncation set $\mathcal{A}$. A simple and commonly applied scheme consists in upper-bounding the total degree of polynomials to a maximal value p. The total degree of polynomials is defined by:

$$|\boldsymbol{\alpha}| = \sum_{i=1}^{M} \alpha_i. \tag{3.9}$$

In this case, the set of multi-indices is denoted by $\mathcal{A}^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : \ |\boldsymbol{\alpha}| \leq p\}$ where p is the maximal total polynomial degree. The number of elements in the set $\mathcal{A}^{M,p}$ is:

$$\left|\mathcal{A}^{M,p}\right| = \frac{(M+p)!}{M!\,p!}. \tag{3.10}$$

This cardinality grows polynomially with both M and p. Such a truncation scheme thus leads to non-tractable problems if the response is highly nonlinear in its input parameters (need for a large p) and/or if the size of the input vector $\boldsymbol{X}$ is large ($M > 10$). This problem is referred to as the *curse of dimensionality*.

Blatman (2009) and Blatman & Sudret (2010a) proposed a more restrictive truncation scheme called *hyperbolic index set*. The authors observed that many systems tend to have only low-degree interaction polynomials and thus it is not necessary to compute all interaction terms of higher polynomial degree. The hyperbolic index set is based on the following q-norm:

$$\mathcal{A}_q^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M : \ \|\boldsymbol{\alpha}\|_q \leq p\}, \tag{3.11}$$

where

$$\|\boldsymbol{\alpha}\|_q = \left(\sum_{i=1}^{M} \alpha_i^q\right)^{\frac{1}{q}}, \tag{3.12}$$

where $0 < q \leq 1$ is a parameter and p is the maximal total degree of the polynomials retained. A decreasing q leads to a smaller number of interactive polynomials, *i.e.* a smaller set of polynomials. When $q \to 0$, only univariate polynomials are left in the set of polynomials which is called an additive model (Sudret 2015). For the sake of illustration, the retained polynomial indices $\boldsymbol{\alpha} \in \mathcal{A}_q^{M,p}$ of a two-dimensional input space ($M = 2$) and varying p and q are illustrated in Figure 3.1. The indices denoted by $\bullet$ are elements of $\mathcal{A}_q^{M,p}$ and the solid black line represents $\|\boldsymbol{\alpha}\|_q = p$. Note that for $q = 1$, the hyperbolic index sets are equivalent to the total degree index set (see Eq. (3.9)).

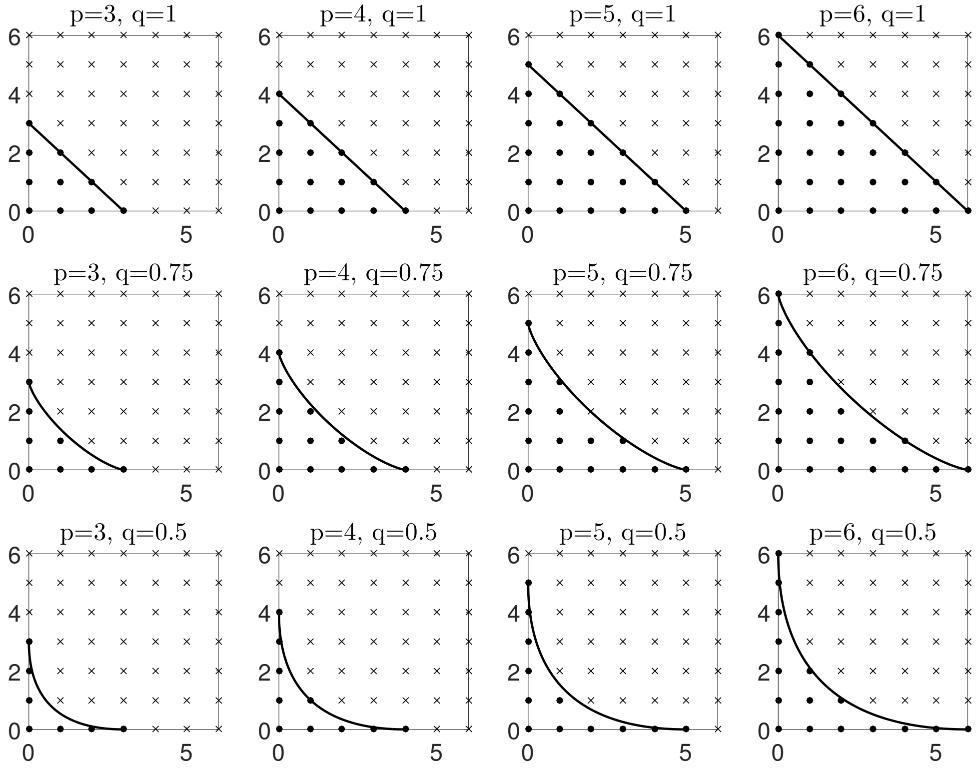


Figure 3.1: Representation of a hyperbolic index set $\mathcal{A}_q^{M,p}$ for various p and q ($M = 2$).

3.3.3 Computation of the coefficients

After defining the set of candidate polynomials, the next step is to determine the expansion coefficients $a_{\boldsymbol{\alpha}}$ of each multivariate polynomial $\psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$. In this thesis, only *non-intrusive* methods are considered, which are based on repeatedly evaluating the model $\mathcal{M}$ over a set of input realizations $\boldsymbol{\mathcal{X}} = \{\boldsymbol{\chi}^{(1)}, \dots, \boldsymbol{\chi}^{(N)}\}$, the so-called *experimental design*. A variety of methods has been proposed to sample the input variables $\boldsymbol{X}$ to ensure high accuracy for the meta-model. Apart from random sampling, such as Monte Carlo simulation and Latin-hypercube sampling (LHS) (McKay, Beckman, & Conover 1979), quasi-random sequences may be used to generate an experimental design. Popular quasi-random sequences include Sobol' sequences (Sobol' 1967), Faure sequences (Faure 1982), Halton series (Halton 1960), Niederreiter series (Niederreiter 1988), and orthogonal arrays (Suen & Kuhfeld 2005; Y. Zhang 2007). For a comparison of the various methods, the reader is referred to Schöbi & Sudret (2014a).

After generating the experimental design, the coefficients $a_{\boldsymbol{\alpha}}$ are computed. Different non-intrusive methods have been proposed in the last decade to calibrate PC meta-models, namely projection (Ghiocel & Ghanem 2002; Le Maître et al. 2002; Keese & Matthies 2005), stochastic collocation (Xiu & Hesthaven 2005; Xiu 2009) and least-square minimization methods (Berveiller, Sudret, & Lemaire 2006b; Blatman & Sudret 2010a, 2011; Chkifa et al. 2013; Migliorati et al. 2014). In this thesis, least-square minimization methods are used.

The expansion coefficients $\boldsymbol{a} = \{a_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \mathcal{A} \subset \mathbb{N}^M\}$ are calculated by minimizing the expectation of the least-squares residual:

$$\boldsymbol{a} = \arg \min_{\boldsymbol{a} \in \mathbb{R}^{n_{\mathcal{A}}}} \mathbb{E}\left[\left(Y - \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\right)^2\right], \tag{3.13}$$

where $n_{\mathcal{A}} = |\mathcal{A}|$ are the number of polynomials in PCE. In practice, the expectation in Eq. (3.13) is evaluated by an empirical sample-based estimator. Denoting by $\mathcal{Y} = \left\{\mathcal{Y}^{(1)}, \dots, \mathcal{Y}^{(N)}\right\}$ the set of outputs of the exact model $\mathcal{M}$ for each point in the experimental design $\boldsymbol{\mathcal{X}}$ (*i.e.* $\mathcal{Y}^{(i)} = \mathcal{M}(\boldsymbol{\chi}^{(i)}),\ i = 1, \dots, N$), the discretized least-squares error minimization problem derived from Eq. (3.13) reads:

$$\widehat{\boldsymbol{a}} = \arg \min_{\boldsymbol{a} \in \mathbb{R}^{n_{\mathcal{A}}}} \frac{1}{N} \sum_{i=1}^{N} \left(\mathcal{Y}^{(i)} - \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}\left(\boldsymbol{\chi}^{(i)}\right)\right)^2. \tag{3.14}$$

The optimal expansion coefficients $\widehat{\boldsymbol{a}}$ can be computed by solving the linear system

$$\widehat{\boldsymbol{a}} = (\mathbf{F}^{\mathsf{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathsf{T}}\mathcal{Y}, \tag{3.15}$$

where $\mathbf{F}$ is the information matrix of size $N \times n_{\mathcal{A}}$ whose generic term reads:

$$\mathrm{F}_{ij} = \psi_j\left(\boldsymbol{\chi}^{(i)}\right), \quad i = 1, \dots, N, \quad j = 1, \dots, n_{\mathcal{A}}. \tag{3.16}$$

3.3.4 Sparse PCE

Typically for smooth functions, a small number of polynomials is able to represent accurately the output of the computational model. Thus, a further reduction of the set of predictors is possible. Various types of generalized regression algorithms have been proposed in the literature, namely the *least absolute shrinkage operator* (LASSO) (Tibshirani 1996), the *Least Angle Regression* (LAR) (Efron, Hastie, et al. 2004), the *Orthogonal Matching Pursuit* (OMP) (Pati, Rezaiifar, & Krishnaprasad 1993; Mallat & Z. Zhang 1993), *low-rank approximations* (Doostan, Validi, & Iaccarino 2013; Peng, Hampton, & Doostan 2014) and compressive sensing (Sargsyan et al. 2014).

Based on LAR, sparse PCE models defined by a hyperbolic index set $\mathcal{A}_q^{M,p}$ in combination with the LARS algorithm (Blatman & Sudret 2011) are used in this thesis. LARS allows for an efficient selection of a small number of polynomial predictors out of the candidate set $\mathcal{A}_q^{M,p}$. For a thorough discussion of the LARS algorithm, the reader is referred to Blatman (2009) and Blatman & Sudret (2011).

3.3.5 Error measures

Validation set-based error

As seen in Eq. (3.8), PCE are approximations of the exact computational model and thus the prediction at new input samples leads to some residual error. Hence, error measures are developed to quantify the deviation between exact output Y and the meta-model output $Y^{(\mathrm{PCE})}$. The *generalization error* (also called L^2-error) is the expectation of the squared output residuals (Vapnik 1995):

$$Err_{gen}^{(\mathrm{P})} = \mathbb{E}\left[\left(Y - Y^{(\mathrm{PCE})}\right)^2\right], \tag{3.17}$$

where the expectation is defined with respect to the PDF of the input variables $\boldsymbol{X}$. If the computational model $\mathcal{M}$ is inexpensive to evaluate, the generalization error can be estimated accurately using an auxiliary *validation set* $\mathcal{S} = \left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n)}\right\}$, which is sampled from the input distribution $f_{\boldsymbol{X}}$. The estimate of the generalization error then reads:

$$\widehat{Err}_{gen}^{(\mathrm{P})} = \frac{1}{n} \sum_{i=1}^{n} \left[\mathcal{M}\left(\boldsymbol{x}^{(i)}\right) - \mathcal{M}^{(\mathrm{PCE})}\left(\boldsymbol{x}^{(i)}\right)\right]^2. \tag{3.18}$$

Normalizing the generalization error by the variance of the output Y leads to the *relative generalization error*:

$$\widehat{err}_{gen}^{(\mathrm{P})} = \frac{\widehat{Err}_{gen}}{\mathrm{Var}\left[Y\right]} = \frac{\sum_{i=1}^{n} \left[\mathcal{M}\left(\boldsymbol{x}^{(i)}\right) - \mathcal{M}^{(\mathrm{PCE})}\left(\boldsymbol{x}^{(i)}\right)\right]^2}{\sum_{i=1}^{n} \left[\mathcal{M}\left(\boldsymbol{x}^{(i)}\right) - \widehat{\mu}_Y\right]^2}, \tag{3.19}$$

where $\widehat{\mu}_Y = \frac{1}{n}\sum_{i=1}^{n} \mathcal{M}\left(\boldsymbol{x}^{(i)}\right)$ is the estimated mean value of Y. By normalizing with respect to the variance, the generalization error provides a dimensionless measure relative to the scatter/variability of the response quantity Y.

Experimental design-based error

However, a validation set is rarely available in real applications since the very purpose of building a meta-model is to avoid evaluating $\mathcal{M}$ on a large sample set $\mathcal{S}$. When the use of a large validation set is not affordable, the *empirical error* based on the available experimental design $\boldsymbol{\mathcal{X}}$ may be defined:

$$Err_{emp} = \frac{1}{N} \sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \mathcal{M}^{(\mathrm{PCE})}\left(\boldsymbol{\chi}^{(i)}\right)\right]^2. \tag{3.20}$$

Normalizing the empirical error by the variance of the output values leads to the *relative empirical error*, which is defined as:

$$err_{emp} = \frac{Err_{emp}}{\mathrm{Var}\left[\mathcal{Y}\right]} = \frac{\sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \mathcal{M}^{(\mathrm{PCE})}\left(\boldsymbol{\chi}^{(i)}\right)\right]^2}{\sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \widehat{\mu}_{\mathcal{Y}}\right]^2}, \tag{3.21}$$

where $\widehat{\mu}_{\mathcal{Y}} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{M}\left(\boldsymbol{\chi}^{(i)}\right)$ is the estimated mean value of the output values $\mathcal{Y}$. The empirical error, which is based on the experimental design, generally underestimates the generalization error. In particular, if the number of polynomials $|\mathcal{A}|$ is close to the number of samples N in the experimental design, the empirical error tends to zero whereas the true (generalization) error may not. In the extreme case where $N = |\mathcal{A}|$ the predictors may interpolate the experimental design points and thus the empirical error vanishes. This phenomenon is called *overfitting.*

In order to avoid overfitting, the *leave-one-out* (LOO) has been proposed as a better estimate of the generalization error (Stone 1974; Geisser 1975). The general formulation of the LOO error reads:

$$Err_{\mathrm{LOO}}^{(\mathrm{PCE})} = \frac{1}{N} \sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \mathcal{M}_{(-i)}^{(\mathrm{PCE})}\left(\boldsymbol{\chi}^{(i)}\right)\right]^2, \tag{3.22}$$

where $\mathcal{M}_{(-i)}^{(\mathrm{PCE})}(\cdot)$ is a PCE model built from the reduced experimental design set $\boldsymbol{\mathcal{X}}^{(-i)} = \boldsymbol{\mathcal{X}} \backslash \boldsymbol{\chi}^{(i)} = \left\{\boldsymbol{\chi}^{(j)},\ j = 1, \dots, i-1, i+1, \dots, N\right\}$ and $\mathcal{Y} = \left\{\mathcal{Y}^{(i)},\ i = 1, \dots, N\right\}$ are the corresponding response values of the exact computational model. The LOO error is a special case of the *leave-k-out cross validation* error (Allen 1971), which discards k samples from the initial experimental design to build up a model and predict the error at the k samples left aside.

In theory, the computational cost of the LOO error is proportional to the number of samples N, since N PCE meta-models corresponding to each experimental design $\boldsymbol{\mathcal{X}}^{(-i)}$ seem to be required in its estimation in Eq. (3.22). In the special case of linearly parameterized regression, which is the case for PCE, it is possible to calculate the LOO error analytically *without* building N separate models. Then, the LOO error reads (see *e.g.* Saporta (2006) and Blatman (2009) for the proof):

$$Err_{LOO}^{(\mathrm{PCE})} = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{\mathcal{Y}^{(i)} - \mathcal{M}^{(\mathrm{PCE})}(\boldsymbol{\chi}^{(i)})}{1 - h_i} \right]^2 , \tag{3.23}$$

where h_i is the i-th diagonal term of the matrix $\mathbf{F}\left(\mathbf{F}^{\mathsf{T}}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathsf{T}}$ and the information matrix $\mathbf{F}$ is defined in Eq. (3.16). Note that the PCE used in Eq. (3.23) is built only once from the full experimental design $\boldsymbol{\mathcal{X}}$.

Similar to the generalization error, the relative leave-one-out error is defined as the following ratio:

$$err_{LOO}^{(\mathrm{PCE})} = \frac{Err_{LOO}^{(\mathrm{PCE})}}{\mathrm{Var}\left[\mathcal{Y}\right]} = \frac{\sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \mathcal{M}_{(-i)}^{(\mathrm{PCE})}\left(\boldsymbol{\chi}^{(i)}\right)\right]^2}{\sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \widehat{\mu}_{\mathcal{Y}}\right]^2}, \tag{3.24}$$

where $\widehat{\mu}_{\mathcal{Y}} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{M}\left(\boldsymbol{\chi}^{(i)}\right)$ is the estimated mean value of Y.

3.4 Kriging

3.4.1 Definition

Kriging, also known as *Gaussian process modeling*, assumes that the computational model is a realization of a Gaussian random process (Santner, B. Williams, & Notz 2003):

$$\mathcal{M}(\boldsymbol{x}) \approx \mathcal{M}^{(\mathrm{K})}(\boldsymbol{x}) = \boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{f}(\boldsymbol{x}) + \sigma^2 \, Z(\boldsymbol{x}, \omega), \tag{3.25}$$

where $\boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{f}(\boldsymbol{x}) = \sum_{j=1}^{n_T} \beta_j f_j(\boldsymbol{x})$ is the mean value (a.k.a *trend*) of the Gaussian process, σ^2 is the process variance and $Z(\boldsymbol{x}, \omega)$ is a zero-mean, unit-variance stationary Gaussian process. The zero-mean Gaussian process $Z(\boldsymbol{x}, \omega)$ is characterized by an auto-correlation function $R(\boldsymbol{x}, \boldsymbol{x}') = R(|\boldsymbol{x} - \boldsymbol{x}'|; \boldsymbol{\rho})$ and its hyper-parameters $\boldsymbol{\rho}$.

Various correlation functions can be found in the literature (Santner, B. Williams, & Notz 2003; Rasmussen & C. Williams 2006), including *linear*, *exponential*, *Gaussian* (also called *squared exponential*), and *Matérn* autocorrelation functions. A summary of the commonly used autocorrelation functions can be found in Appendix C. Depending on the characteristics of the computational model $\mathcal{M}$, one or the other represents the function's behaviour better. Note that for stochastic computational models, *i.e.* models where the repeated evaluation of the same input vector $\boldsymbol{x}$ leads to different values of y, a *nugget* is added to the autocorrelation function.

Apart from the correlation part in Eq. (3.25), there is also the trend $\boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{f}(\boldsymbol{x})$. Three different flavours of Kriging are defined in the literature (M. Stein 1999; Santner, B. Williams, & Notz 2003; Rasmussen & C. Williams 2006), namely simple, ordinary and universal Kriging according to the complexity of the trend. *Simple Kriging* assumes that the trend has a known constant value, *i.e.* $\boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{f}(\boldsymbol{x}) = \beta_0$. In *ordinary Kriging*, the trend has a constant but unknown value, *i.e.* $n_T = 1$, $f_1(\boldsymbol{x}) = 1$ and β_1 is unknown. The most general and flexible formulation is *universal Kriging*, which assumes that the trend is composed of a sum of n_T pre-selected functions $f_k(\boldsymbol{x})$:

$$\boldsymbol{\beta}^{\mathsf{T}} \boldsymbol{f}(\boldsymbol{x}) = \sum_{k=1}^{n_T} \beta_k f_k(\boldsymbol{x}), \tag{3.26}$$

where β_k is the trend coefficient corresponding to the k-th function f_k. Note that simple and ordinary Kriging are special cases of universal Kriging. Hence, the following derivations are elaborated for the case of universal Kriging only.

3.4.2 Calibration

Given a value for the auto-correlation hyper-parameters $\widehat{\boldsymbol{\rho}}$, the calibration of the Kriging model parameters $\left\{\boldsymbol{\beta}(\widehat{\boldsymbol{\rho}}), \sigma_y^2(\widehat{\boldsymbol{\rho}})\right\}$ may be computed using an *empirical best linear unbiased estimator* (BLUE). The optimization yields an analytical expression as a function of $\widehat{\boldsymbol{\rho}}$:

$$\boldsymbol{\beta}(\widehat{\boldsymbol{\rho}}) = \left(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F}\right)^{-1} \mathbf{F}\,\mathbf{R}^{-1}\mathcal{Y}, \tag{3.27}$$

$$\sigma_y^2(\widehat{\boldsymbol{\rho}}) = \frac{1}{N}\left(\mathcal{Y} - \mathbf{F}\,\boldsymbol{\beta}\right)^\mathsf{T} \mathbf{R}^{-1}\left(\mathcal{Y} - \mathbf{F}\,\boldsymbol{\beta}\right), \tag{3.28}$$

where $\mathcal{Y} = \{\mathcal{Y}^{(i)},\ i = 1, \dots, N\}$ are model responses of the exact computational model on the experimental design $\boldsymbol{\mathcal{X}} = \{\boldsymbol{\chi}^{(i)},\ i = 1, \dots, N\}$, $\mathbf{R}_{ij} = R(|\boldsymbol{\chi}^{(i)} - \boldsymbol{\chi}^{(j)}|; \widehat{\boldsymbol{\rho}})$ is the correlation matrix and $\mathbf{F}_{ij} = f_j\left(\boldsymbol{\chi}^{(i)}\right)$ is the information matrix.

Following recent developments, the optimal correlation parameters $\widehat{\boldsymbol{\rho}}$ may be determined by either a maximum-likelihood-estimate (denoted by *ML*) (Marrel et al. 2008; Dubourg 2011) or by leave-one-out cross-validation (*CV*) (Bachoc 2013). The optimal parameters are determined through a minimization:

$$\widehat{\boldsymbol{\rho}}_{ML} = \arg\min_{\boldsymbol{\rho}} \left[\frac{1}{N}\left(\mathcal{Y} - \mathbf{F}\,\boldsymbol{\beta}\right)^\mathsf{T}\ \mathbf{R}^{-1}(\boldsymbol{\rho})\ \left(\mathcal{Y} - \mathbf{F}\,\boldsymbol{\beta}\right)\left(\det \mathbf{R}(\boldsymbol{\rho})\right)^{1/N}\right], \tag{3.29}$$

$$\widehat{\boldsymbol{\rho}}_{CV} = \arg\min_{\boldsymbol{\rho}} \left[\mathcal{Y}^\mathsf{T}\ \mathbf{R}^{-1}(\boldsymbol{\rho})\ \mathrm{diag}\left(\mathbf{R}^{-1}(\boldsymbol{\rho})\right)^{-2}\ \mathbf{R}^{-1}(\boldsymbol{\rho})\ \mathcal{Y}\right]. \tag{3.30}$$

The comparison of both approaches shows that ML is preferable to CV in well-specified cases, *i.e.* when the meta-model autocorrelation function family is identical to the autocorrelation function of the computational model. For practical problems, *i.e.* assuming a black-box model, the autocorrelation function family is not known with certainty. In this case, CV shall lead to more robust results than ML, as discussed in Bachoc (2013).

3.4.3 Prediction

Based on the assumption of a Gaussian process, the prediction of a Kriging model of a new point $\boldsymbol{x}$ is a Gaussian random variable with mean $\mu_{\widehat{Y}}(\boldsymbol{x})$ and variance $\sigma^2_{\widehat{Y}}(\boldsymbol{x})$:

$$\mu_{\widehat{Y}}(\boldsymbol{x}) = \boldsymbol{f}^\mathsf{T}(\boldsymbol{x})\boldsymbol{\beta}\ +\ \boldsymbol{r}^\mathsf{T}(\boldsymbol{x})\mathbf{R}^{-1}\left(\mathcal{Y} - \mathbf{F}\boldsymbol{\beta}\right), \tag{3.31}$$

$$\sigma^2_{\widehat{Y}}(\boldsymbol{x}) = \sigma_y^2\left(1\ -\ \boldsymbol{r}^\mathsf{T}(\boldsymbol{x})\,\mathbf{R}\,\boldsymbol{r}(\boldsymbol{x})\ +\ \boldsymbol{u}^\mathsf{T}(\boldsymbol{x})\left(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\boldsymbol{u}(\boldsymbol{x})\right), \tag{3.32}$$

where $r_i(\boldsymbol{x}) = R\left(|\boldsymbol{x} - \boldsymbol{\chi}^{(i)}|; \boldsymbol{\rho}\right)$ is the correlation between the new sample $\boldsymbol{x}$ and the sample $\boldsymbol{\chi}^{(i)}$ of the experimental design and $\boldsymbol{u}(\boldsymbol{x}) = \mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})$. The prediction mean is used as the surrogate to the original model $\mathcal{M}$, whereas the variance gives a *local* error indicator about its prediction precision. It is important to note that the Kriging model interpolates the data, namely:

$$\mu_{\widehat{Y}}\left(\boldsymbol{\chi}^{(i)}\right) = \mathcal{M}\left(\boldsymbol{\chi}^{(i)}\right), \qquad \sigma^2_{\widehat{Y}}\left(\boldsymbol{\chi}^{(i)}\right) = 0, \qquad \forall \boldsymbol{\chi}^{(i)} \in \boldsymbol{\mathcal{X}}. \tag{3.33}$$

3.4.4 Error measures

A local error measure for any sample $\boldsymbol{x}$ is given by the prediction variance $\sigma^2_{\widehat{Y}}(\boldsymbol{x})$ defined in Eq. (3.32). As an example, a confidence interval of the prediction can be defined as $\mu_{\widehat{Y}}(\boldsymbol{x}) \pm 1.96 \cdot \sigma_{\widehat{Y}}(\boldsymbol{x})$ for every point $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}}$. This information is useful to detect regions where the prediction accuracy is low. Adding new samples to the experimental design $\boldsymbol{\mathcal{X}}$ in the regions with high prediction variance may lead to an overall increase in the accuracy of the meta-model in that region. This characteristics is exploited when devising adaptive experimental designs in structural reliability analysis (Bichon, Eldred, et al. 2008; Bichon, McFarland, & Mahadevan 2011; Dubourg, Sudret, & Bourinet 2011; Echard, Gayton, Lemaire, & Relun 2013; Schöbi, Sudret, & Marelli 2016). For further details, the reader is referred to Chapter 5.

A simple global error measure of the accuracy of the meta-model (such as the empirical error in Eq. (3.21) for PCE) is not available for Kriging due to its interpolating properties (see Eq. (3.33)), which results in $Err_{emp} = 0$ (considering no nugget effect in its auto-correlation function). However, the LOO error is available and is defined as:

$$Err^{(\mathrm{K})}_{LOO} = \frac{1}{N} \sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \mu_{\widehat{Y}(-i)} \left(\boldsymbol{\chi}^{(i)} \right) \right]^2 , \tag{3.34}$$

where $\mu_{\widehat{Y}(-i)} \left(\boldsymbol{\chi}^{(i)} \right)$ is the prediction mean $\mu_{\widehat{Y}}$ at sample $\boldsymbol{\chi}^{(i)}$ by a Kriging meta-model based on the experimental design $\boldsymbol{\mathcal{X}}^{(-i)} = \boldsymbol{\mathcal{X}} \backslash \boldsymbol{\chi}^{(i)}$ and $\mathcal{Y} = \left\{ \mathcal{Y}^{(i)}, i = 1, \ldots, N \right\}$ is the exact model response. Dubrule (1983) derived an analytical solution for the LOO error for universal Kriging without computing the N meta-models explicitly, in the same spirit as Eq. (3.23) for PCE. The prediction mean and variance are given by:

$$\mu_{\widehat{Y}(-i)} = - \sum_{j=1, j \neq i}^{N} \frac{\mathbf{B}_{ij}}{\mathbf{B}_{ii}} \mathcal{Y}^{(j)} = - \sum_{j=1}^{N} \frac{\mathbf{B}_{ij}}{\mathbf{B}_{ii}} \mathcal{Y}^{(j)} + \mathcal{Y}^{(i)}, \tag{3.35}$$

$$\sigma^2_{\widehat{Y}(-i)} = \frac{1}{\mathbf{B}_{ii}}, \tag{3.36}$$

where $\mathbf{B}$ is a square matrix of size $(N + n_T)$ with N and n_T denoting the number of samples in the experimental design and the number of functions f_k in the trend part, respectively:

$$\mathbf{B} = \begin{bmatrix} \sigma^2_y \mathbf{R} & \mathbf{F} \\ \mathbf{F}^{\mathsf{T}} & \mathbf{0} \end{bmatrix}^{-1} , \tag{3.37}$$

where σ^2_y is the Kriging variance for the full experimental design $\boldsymbol{\mathcal{X}}$ estimated by Eq. (3.28). A generalized version of this algorithm called *v-fold cross-correlation error* can be found in Dubrule (1983). Finally, the relative leave-one-out error is obtained by normalizing the leave-one-out error by the variance of Y:

$$err^{(\mathrm{K})}_{LOO} = \frac{Err^{(\mathrm{K})}_{LOO}}{\mathrm{Var}\left[Y\right]} = \frac{\sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \mu_{\widehat{Y}(-i)} \left(\boldsymbol{\chi}^{(i)} \right) \right]^2}{\sum_{i=1}^{N} \left[\mathcal{Y}^{(i)} - \widehat{\mu}_{\mathcal{Y}} \right]^2}, \tag{3.38}$$

where $\widehat{\mu}_{\mathcal{Y}} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{M} \left(\boldsymbol{\chi}^{(i)} \right)$ is the estimated mean value of Y.

When a validation set is available, the generalization error and relative generalization error defined in Eqs. (3.18) and (3.19) for PCE, respectively, can be adopted to the Kriging prediction as follows:

$$\widehat{Err}^{(\mathrm{K})}_{gen} = \frac{1}{n} \sum_{i=1}^{n} \left[\mathcal{M} \left(\boldsymbol{x}^{(i)} \right) - \mu_{\widehat{Y}} \left(\boldsymbol{x}^{(i)} \right) \right]^2 , \tag{3.39}$$

$$\widehat{err}^{(\mathrm{K})}_{gen} = \frac{\sum_{i=1}^{n} \left[\mathcal{M}\left(\boldsymbol{x}^{(i)}\right) - \mu_{\widehat{Y}}\left(\boldsymbol{x}^{(i)}\right)\right]^2}{\sum_{i=1}^{n} \left[\mathcal{M}\left(\boldsymbol{x}^{(i)}\right) - \widehat{\mu}_Y\right]^2}, \tag{3.40}$$

where $\widehat{\mu}_Y = \frac{1}{n}\sum_{i=1}^{n} \mathcal{M}\left(\boldsymbol{x}^{(i)}\right)$ is the estimated mean value of Y.

3.5 PCE versus Kriging

3.5.1 Model behaviour

As mentioned before, PCE approximates the global behaviour of the computational model well, due to the regression-type approach that is used in this thesis and the related least-squares minimization of the residuals. Kriging, however, focuses on the local behaviour of the computational model, resulting in high prediction accuracy close to sample points of the experimental design. On the other side, the global behaviour can be poor. In fact, far away from the experimental design points $\boldsymbol{\chi}^{(i)}$, Kriging tends to the trend. Hence, the goodness-of-fit of the Kriging meta-model does not only depend on the experimental design points but also on the definition of the trend.

The dependency on the experimental design algorithm is a characteristic of Kriging. In fact, the experimental design $\boldsymbol{\mathcal{X}}$ and the corresponding responses $\mathcal{Y}$ are required to predict a new point $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}}$, as seen in Eqs. (3.31) and (3.32). This is due to the correlation vector $\boldsymbol{r}$ which estimates the correlation between the new sample $\boldsymbol{x}$ and the samples of the experimental design $\boldsymbol{\mathcal{X}}$. This results in large matrices and the need for inverse of large matrices when the experimental design is large. PCE, however, is independent of the experimental design $\boldsymbol{\mathcal{X}}$ once the meta-model is calibrated. The prediction of a new sample $\boldsymbol{x}$ is achieved by simply solving Eq. (3.8), which consists of evaluating a sum of multivariate polynomials.

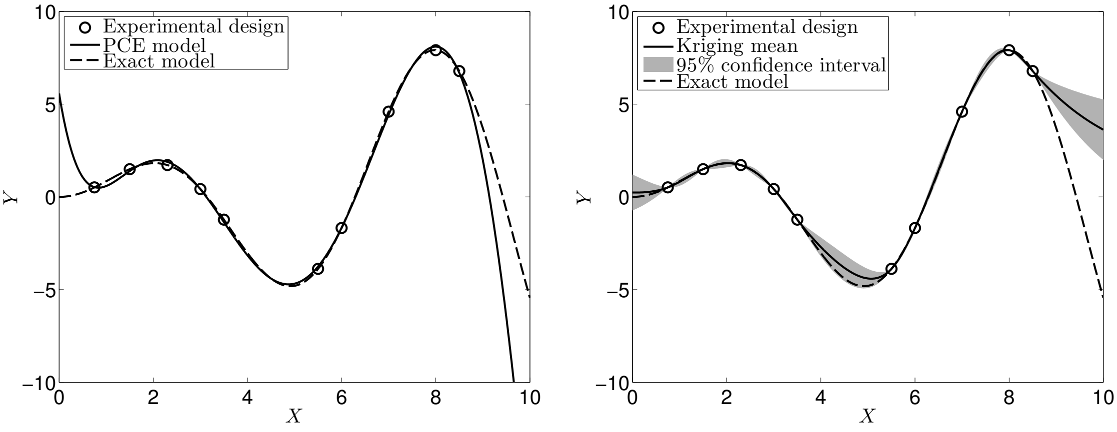
Due to the independence on $\boldsymbol{\mathcal{X}}$, PCE does not interpolate samples of the experimental design in the general case ($\mathcal{Y}^{(i)} \neq \mathcal{M}^{(\mathrm{PCE})}\left(\boldsymbol{\chi}^{(i)}\right)$). This downside originates in the regression-types approach used for PCE. Note that in contrast, stochastic collocation approaches lead to multivariate interpolating polynomials. Kriging, however, interpolates the experimental design points as shown in Eq. (3.33) whatever the structure of the trend and the autocorrelation function. Considering that the computational model is deterministic, *i.e.* $\mathcal{Y}^{(i)}$ is the true and precise value of $\mathcal{M}\left(\boldsymbol{\chi}^{(i)}\right)$, Kriging may locally lead to better prediction results than PCE.

3.5.2 Visualization

In order to visualize the general behaviour of PCE and Kriging, consider the following experimental design $\mathcal{X}_2$, where $\chi^{(i)},\ i = 1, \ldots, 10$ are sampled from a uniform distribution $X \sim \mathcal{U}(0, 10)$. The computational model is defined as $\mathcal{M}(x) = x \cdot \sin(x)$, which defines the vector of response values $\mathcal{Y}_2$ corresponding to $\mathcal{X}_2$.

PCE and Kriging meta-models are calibrated using the Matlab-based framework UQLab (Marelli & Sudret 2014). In this example, PCE is calibrated by a truncation set $\mathcal{A}^{M,p}$ where the maximal total polynomial degree is $p = 10$ and the degree-adaptive LARS algorithm is used. Note that the use of hyperbolic index sets is redundant here due to $M = 1$. For the Kriging meta-model, an ordinary Kriging model is used together with a Matérn autocorrelation function ($\nu = 3/2$). The corresponding hyper-parameter ρ is determined by cross-validation.

Figure 3.2 compares the true computational model to PCE and Kriging. The circles mark the samples of the experimental design whereas the dashed line represents the exact computational model $\mathcal{M}$. In Figure 3.2(a), the solid line marks the prediction of the PCE model. The prediction line follows the line of the exact model nicely within the range of the experimental design $x \in \left[\min_i \chi^{(i)}, \max_i \chi^{(i)}\right]$. At the boundaries of the domain $\mathcal{D}_X$, however, the PCE model tends to be inaccurate: close to $x = 0$ the exact model response in overestimated whereas close to $x = 10$ it is underestimated.



(a) PCE model (b) Kriging model

Figure 3.2: Comparison of PCE and Kriging – meta-models based on $\mathcal{X}_2$ and corresponding $\mathcal{Y}_2$.

The Kriging model is shown in Figure 3.2(b), where the prediction mean value $\mu_{\widehat{Y}}(\boldsymbol{x})$ are shown by the solid line and where the grey area marks the 95% confidence interval of the prediction. The confidence interval $\left[\mu_{\widehat{Y}}(\boldsymbol{x}) - 1.96 \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}),\, \mu_{\widehat{Y}}(\boldsymbol{x}) + 1.96 \cdot \sigma_{\widehat{Y}}(\boldsymbol{x})\right]$ is obtained by accounting for the fact that the prediction of each $x \in \mathcal{D}_X$ results in a Gaussian random variable characterised by $\left\{\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma_{\widehat{Y}}(\boldsymbol{x})\right\}$. The grey confidence intervals point out areas of low accuracy, which correspond to the areas where the prediction mean value $\mu_{\widehat{Y}}(\boldsymbol{x})$ does not overlap with the true value $y = \mathcal{M}(\boldsymbol{x})$. Similar to PCE, the Kriging model predicts well the exact model in the centre of the input domain and not so well towards the boundaries of the input domain.

As discussed in Section 3.5.1, the behaviour of the two meta-models is different when extrapolating outside of the range of experimental design samples. The polynomial format of PCE results in extreme response values at the boundaries of $\mathcal{D}_X$, whereas the ordinary Kriging model results in predictions towards its trend (here estimated $\beta_1 = 1.45$). Hence in this case, Kriging results in more confined response values than PCE.

The accuracy of both meta-models is compared in Table 3.1 by means of error measures introduced in Sections 3.3.5 and 3.4.4. The empirical and LOO error are based on the experimental design. Note that the empirical error of Kriging is zero due to the interpolation feature. The low values of both error measures indicate that the meta-models accurately predict the computational model response Y, which confirms the visual impression gained from Figure 3.2.

The (relative) generalization error measures the accuracy with respect to a validation set of size $n_{\mathrm{MCS}} = 10^5$ sampled uniformly in the input domain. In order to see the effect of extrapolation, *i.e.* prediction outside of the range of experimental design samples, the generalization error is computed for $x \in [0, 10]$ and $x \in [0.75, 8.5]$. In this example, the difference in the error estimates is significant. Comparing the domain $x \in [0.75, 8.5]$ to $x \in [0, 10]$, the generalization error varies by two orders of magnitude. This is due to the previously discussed behaviour when extrapolating beyond the range of experimental design samples and confirms the visual impression gained in Figure 3.2.

Finally, the relative generalization error normalizes the generalization error with respect to $\mathrm{Var}\left[Y\right] = 13.61$ (see also Eqs. (3.19) and (3.40)). Then, an error of $\widehat{err}_{gen} = 1$ would correspond

Table 3.1: Comparison of PCE and Kriging – error measures for $\mathcal{X}_2$ and corresponding $\mathcal{Y}_2$ (validation set with $n = 10^5$ samples).

Error measure	Domain	PCE	Kriging
Err_{emp}	-	$1.47 \cdot 10^{-2}$	0
Err_{LOO}	-	1.11	$3.90 \cdot 10^{-1}$
err_{LOO}	-	$7.84 \cdot 10^{-2}$	$2.76 \cdot 10^{-2}$
$\widehat{Err}_{gen}$	$x \in [0.0, 10.0]$	5.66	3.26
	$x \in [0.75, 8.5]$	$1.61 \cdot 10^{-2}$	$3.90 \cdot 10^{-2}$
$\widehat{err}_{gen}$	$x \in [0.0, 10.0]$	$4.16 \cdot 10^{-1}$	$2.39 \cdot 10^{-1}$
	$x \in [0.75, 8.5]$	$1.19 \cdot 10^{-3}$	$2.86 \cdot 10^{-3}$

to a model where the meta-model residuals scatter by a similar amount as the variable Y. Hence, $\widehat{err}_{gen} \ll 1$ would correspond to an accurate model. In the present example, the relative generalization errors are small for $x \in [0.75, 8.5]$ and between 0.2 and 0.4 for $x \in [0, 10]$. This confirms the discussion on extrapolation in the previous paragraph on the generalization error.

3.5.3 PCE as a particular case of Kriging

Despite the many differences in the definition of PCE and Kriging, the two meta-models are related. In fact, PCE can be interpreted as universal Kriging models when the autocorrelation function consists of the Kronecker delta function:

$$R(|\boldsymbol{x} - \boldsymbol{x}'|) = \delta(\boldsymbol{x} - \boldsymbol{x}'), \tag{3.41}$$

where $\delta = 1$ for $\boldsymbol{x} = \boldsymbol{x}'$ and $\delta = 0$ otherwise. The correlation matrix is then the identity matrix of size N, *i.e.* $\mathbf{R} = \mathbf{I}_N$. This reduces Eqs. (3.27) to (3.15) and hence Eqs. (3.31) to (3.8) (indeed $r(\boldsymbol{x}, \boldsymbol{x}') = 0$ when $\boldsymbol{x} \notin \boldsymbol{\mathcal{X}}$ in this particular case).

Further, it can be shown that the LOO error in Eqs. (3.34) to (3.36) reduces to Eq. (3.23) by the following derivations. Consider the symmetric partitioned matrix $\mathbf{C}$ and the corresponding inverse $\mathbf{D}$, *i.e.* $\mathbf{D} = \mathbf{C}^{-1}$ which are defined as:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}^{\mathsf{T}} & \mathbf{C}_{22} \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{12}^{\mathsf{T}} & \mathbf{D}_{22} \end{bmatrix}, \tag{3.42}$$

where $\mathbf{C}_{11}$ and $\mathbf{D}_{11}$ (respectively $\mathbf{C}_{22}$ and $\mathbf{D}_{22}$) are square matrices with dimension N (respectively n_T). Using block matrix inversion one can derive:

$$\mathbf{D}_{11} = \mathbf{C}_{11}^{-1} + \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\left(\mathbf{C}_{22} - \mathbf{C}_{12}^{\mathsf{T}}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\right)^{-1}\mathbf{C}_{12}^{\mathsf{T}}\mathbf{C}_{11}^{-1}. \tag{3.43}$$

In the context of the LOO error, taking $\mathbf{C} = \mathbf{B}$, $\mathbf{C}_{11} = \sigma_y^2 \mathbf{I}_N$, $\mathbf{C}_{12} = \mathbf{F}$, and $\mathbf{C}_{22} = \mathbf{0}_{n_T}$ in Eq. (3.37) leads to:

$$\begin{aligned} \mathbf{D}_{11} &= \frac{1}{\sigma_y^2}\mathbf{I}_N + \frac{1}{\sigma_y^2}\mathbf{I}_N\mathbf{F}\left(\mathbf{0}_P - \mathbf{F}^{\mathsf{T}}\frac{1}{\sigma_y^2}\mathbf{I}_N\mathbf{F}\right)^{-1}\mathbf{F}^{\mathsf{T}}\frac{1}{\sigma_y^2}\mathbf{I}_N \\ &= \frac{1}{\sigma_y^2}\mathbf{I} + \frac{1}{\sigma_y^2}\mathbf{F}\left(-\mathbf{F}^{\mathsf{T}}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathsf{T}} \\ &= \frac{1}{\sigma_y^2}\left(\mathbf{I} - \mathbf{F}\left(\mathbf{F}^{\mathsf{T}}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathsf{T}}\right). \end{aligned} \tag{3.44}$$

Then, the LOO error in Eq. (3.34) combined with Eq. (3.35) and the above inverse formulation of the $\mathbf{B}$ matrix reads (Schöbi, Sudret, & Wiart 2015):

$$\begin{aligned}
Err_{LOO}^{(\mathrm{K})} &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathcal{Y}^{(i)} + \sum_{j=1}^{N}\frac{B_{ij}}{B_{ii}}\mathcal{Y}^{(j)} - \mathcal{Y}^{(i)}\right)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{B_{ii}}\sum_{j=1}^{N}B_{ij}\mathcal{Y}^{(j)}\right)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{B_{ii}\sigma_y^2}\sum_{j=1}^{N}\left(\mathbf{I}-\mathbf{F}\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\right)_{ij}\mathcal{Y}^{(j)}\right)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{1-\left(\operatorname{diag}\left[\mathbf{F}\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\right]\right)_i}\cdot\left[\mathcal{Y}^{(i)}-\sum_{j=1}^{N}\left[\mathbf{F}\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\right]_{ij}\mathcal{Y}^{(j)}\right]\right)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathcal{Y}^{(i)}-\left[\mathbf{F}\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\mathcal{Y}\right]_i}{1-\left(\operatorname{diag}\left[\mathbf{F}\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\right]\right)_i}\right)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathcal{Y}^{(i)}-\left[\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\mathcal{Y}\right]^\mathsf{T}f(\boldsymbol{\chi}^{(i)})}{1-\left(\operatorname{diag}\left[\mathbf{F}\left(\mathbf{F}^\mathsf{T}\mathbf{F}\right)^{-1}\mathbf{F}^\mathsf{T}\right]\right)_i}\right)^2,
\end{aligned} \tag{3.45}$$

which is equivalent to the formulation of the LOO error in Eq. (3.23) for the case of PCE and $f(\boldsymbol{\chi}^{(i)}) \equiv \psi(\boldsymbol{\chi}^{(i)})$. Thus, the LOO error in PCE can be seen as a special case of the LOO error in the Kriging framework.

3.6 Conclusion

PCE and Kriging are two popular meta-modelling techniques which approximate the computational model by an inexpensive-to-evaluate function. These techniques have been developed by different communities in the past, hence the different formulations, notations, and usages. When looking at Section 3.5, however, the two meta-modelling techniques are more similar than seen at first sight. The question then naturally arises, whether it is possible to combine the two techniques to form an even more accurate and efficient meta-modelling technique. To the author's knowledge, no attempt has been taken in combining them up till now. Hence, a new meta-modelling technique is introduced in the next chapter taking advantage of the properties of PCE and Kriging.

CHAPTER 4

PC-Kriging

This chapter introduces the novel PC-Kriging approach, which is a combination of the classical PCE and Kriging meta-modelling techniques, and compares the three methods. The content of this chapter has been published in Schöbi & Sudret (2014a,d) and Schöbi, Sudret, & Wiart (2015).

4.1 Theory

The characteristics of Kriging is to interpolate *local* variations of the output of the computational model Y as a function of the neighbouring experimental design points. In contrast, PCE are used for approximating the *global* behaviour of $\mathcal{M}$ using a set of orthogonal polynomials. By combining the two techniques, the goal is to capture the global behaviour of the computational model with a set of orthogonal polynomials in the trend of a universal Kriging model and the local variability with the Gaussian process. The new approach called *Polynomial-Chaos-Kriging* (PC-Kriging) combines these two distinct meta-modelling techniques and their characteristics.

Using the same notations as in the previous chapter for truncated PCE (see Eq. (3.8)), the PC-Kriging meta-model is cast as follows (Schöbi & Sudret 2014d):

$$\mathcal{M}(\boldsymbol{x}) \approx \mathcal{M}^{(\mathrm{PCK})}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) + \sigma^2 Z(\boldsymbol{x}, \omega), \tag{4.1}$$

where $\sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$ is a weighted sum of orthonormal polynomials describing the mean value of the Gaussian process and $\mathcal{A}$ is the index set of the polynomials. $Z(\boldsymbol{x}, \omega)$ is a zero-mean, unit-variance stationary Gaussian process defined by an autocorrelation function $R(|\boldsymbol{x} - \boldsymbol{x}'|; \boldsymbol{\rho})$ and is parametrized by a set of hyper-parameters $\boldsymbol{\rho}$.

Building a PC-Kriging meta-model consists of two parts: (i) the determination of the optimal set of polynomials contained in the regression part (*i.e.* in the truncation set $\mathcal{A}$) and (ii) the calibration of the correlation hyper-parameters $\boldsymbol{\rho}$ as well as the Kriging parameters $\left\{\sigma_y^2, \boldsymbol{a}\right\}$. The set of polynomials is determined using the LARS algorithm (Blatman & Sudret 2011) together with hyperbolic index sets to obtain sparse sets of polynomials. After the set is fixed, the trend and correlation parameters are evaluated using the universal Kriging equations (see Eqs. (3.27)-(3.30)).

4.2 Calibration and prediction

The two distinct frameworks for PCE and Kriging can be combined in various ways. In this chapter, two approaches are explained and discussed in details, *i.e.* *Sequential* PC-Kriging and *Optimal* PC-Kriging. Both approaches are based on the same input information: (i) an experimental design $\boldsymbol{\mathcal{X}}$, (ii) the corresponding response values $\mathcal{Y}$ obtained from the computational model evaluations $\mathcal{M}\left(\boldsymbol{\mathcal{X}}\right)$, (iii) the description of the random input variables by their marginal PDFs (to define orthonormal polynomials), and (iv) the parametric expression of the auto-correlation function $R\left(|\boldsymbol{x}-\boldsymbol{x}'|;\boldsymbol{\rho}\right)$.

4.2.1 Sequential PC-Kriging

In sequential PC-Kriging (SPC-Kriging), the set of polynomials and the Kriging meta-models are determined *sequentially*. The assumption behind this procedure is that the optimal set of polynomials found by the LARS algorithm in the context of pure PCE can be used directly as a suitable trend for the universal Kriging model. Hence in a first step, the optimal set of polynomials is determined using the PCE framework: $\mathcal{A}$ is found by applying the LARS procedure as in Blatman & Sudret (2011). Then, the set of multivariate orthonormal polynomials $\mathcal{A}$ is embedded into a universal Kriging model as its trend. The universal Kriging model is finally calibrated using Eqs. (3.27)-(3.30).

At the end of the algorithm, the accuracy of the calibrated meta-model can be measured by *e.g.* the LOO error given in Eq. (3.34) or, when using a validation set, by the sample-based generalization error in Eq. (3.39).

The SPC-Kriging algorithm is illustrated in Figure 4.1, where the white boxes mark the required input information and the grey boxes represent the computational tasks. Given a calibrated SPC-Kriging model, the response of a new input realization (*i.e.* prediction) is computed by Eqs. (3.31) and (3.32).

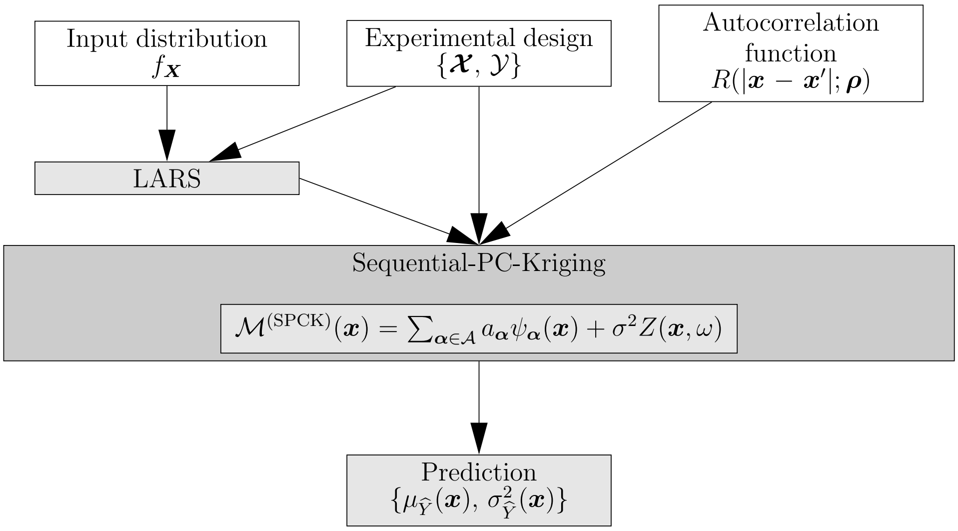


Figure 4.1: Sequential PC-Kriging – flowchart for calibration of the meta-model.

4.2.2 Optimal PC-Kriging

In optimal PC-Kriging (OPC-Kriging), the PC-Kriging meta-model is obtained iteratively. The set of orthonormal multivariate polynomials is determined by the LARS algorithm in the same way as in SPC-Kriging. Yet the LARS algorithm results in a list of ranked polynomials which are chosen depending on their correlation to the current residual at each iteration, in decreasing order. OPC-Kriging consists of an iterative algorithm where each polynomial is added one-by-one (starting with the one ranked highest) to the trend part of the universal Kriging model. In each iteration, a Kriging model is calibrated, *i.e.* the coefficients of the trend and the parameters of the auto-correlation function are computed. In the end, a number of $n_{\mathcal{A}} = |\mathcal{A}|$ different PC-Kriging models are available. The $n_{\mathcal{A}}$ meta-models are then compared in terms of the LOO error in Eq. (3.34). Finally, the optimal PC-Kriging model is chosen as the one with minimal LOO error.

Figure 4.2 illustrates the OPC-Kriging algorithm in a flowchart. Similar to Figure 4.1, the white boxes mark the required input information and the grey boxes represent the computational tasks. The model $\mathcal{M}^{(\mathrm{K})(i)}(\boldsymbol{x})$, $i = 1, \dots, n_{\mathcal{A}}$ denotes a universal Kriging meta-model where the trend is modelled by the first i polynomials selected in $\mathcal{A}$ according to the ranking obtained by LARS. The LOO error corresponding to the Kriging model $\mathcal{M}^{(\mathrm{K})(i)}$ is denoted by $Err_{LOO}^{(\mathrm{K})(i)}$. The final meta-model is chosen as the one minimizing the LOO error, *i.e.* $i^* = \arg\min_{i=1,\dots,n_{\mathcal{A}}} Err_{LOO}^{(\mathrm{K})(i)}$ and hence $\mathcal{M}^{(\mathrm{OPCK})}(\boldsymbol{x}) = \mathcal{M}^{(\mathrm{K})(i^*)}(\boldsymbol{x})$. The final box marks the prediction model responses to new realizations of the input vector $\boldsymbol{X}$. As in SPC-Kriging and any universal Kriging model, the prediction values are obtained by Eqs. (3.31) and (3.32).

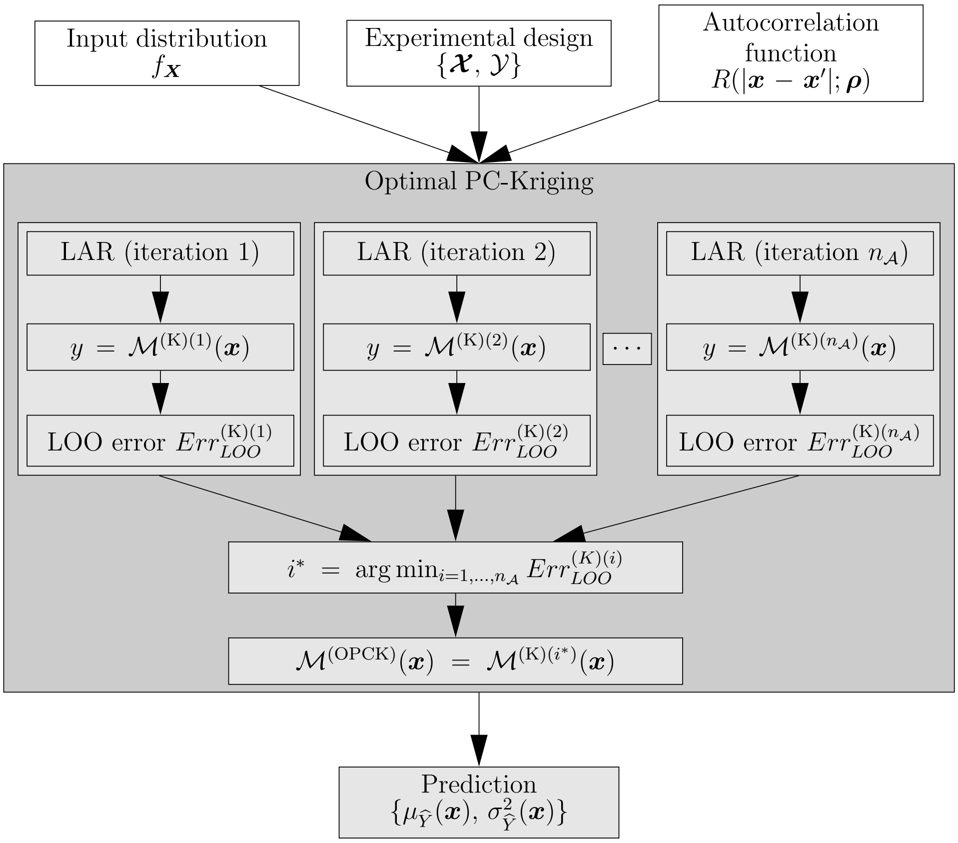


Figure 4.2: Optimal PC-Kriging – flowchart for calibration of the meta-model.

4.3 Error measures

PC-Kriging is nothing but a universal Kriging model with an advanced trend. Thus, the error measures in Section 3.4.4 are valid. In particular, the LOO error in Eq. (3.34) is used to compare different PC-Kriging models and to compute the optimal meta-model in OPC-Kriging.

4.4 Numerical examples

4.4.1 Setup

A number of benchmark analytical functions are analysed in this section to compare the performance of the PC-Kriging approach and the traditional PCE and Kriging. Six analytical functions of different dimensionality are illustrated, namely four with uniformly distributed input variables (*i.e.* Ishigami, Sobol', Rosenbrock and Morris function) and two with Gaussian input variables (*i.e.* Rastrigin and O'Hagan function).

- The *Ishigami* function is a smooth function with three independent input parameters commonly used for benchmarking methods in meta-modelling and global sensitivity analysis:

  $$f_1(\boldsymbol{x}) = \sin x_1 + 7\,\sin^2 x_2 + 0.1\,x_3^4\,\sin x_1, \tag{4.2}$$

  where $X_i \sim \mathcal{U}(-\pi, \pi),\ i = 1, 2, 3$.

- The *Sobol'* function is a well-known function in sensitivity analysis because of the analytical derivation of the Sobol' indices (Sobol' 1993):

  $$f_2(\boldsymbol{x}) = \prod_{i=i}^{8} \frac{|4\,x_i - 2| + c_i}{1 + c_i}, \tag{4.3}$$

  where $X_i \sim \mathcal{U}(0, 1),\ i = 1, \dots, 8$ and $c = (1, 2, 5, 10, 20, 50, 100, 500)^\mathsf{T}$ as in Sudret (2008). Due to the absolute value operator in the enumerator the function behaves non-smoothly at the point $x_i = 0.5$.

- The *Rosenbrock* function is a polynomial function with a two-dimensional input space (Rosenbrock 1960):

  $$f_3(\boldsymbol{x}) = 100\,\left(x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2, \tag{4.4}$$

  where $X_i \sim \mathcal{U}(-2, 2),\ i = 1, 2$.

- The *Morris* function is a function of a high-dimensional input space ($M = 20$) and is defined by (Morris 1991):

  $$f_4(\boldsymbol{x}) = \sum_{i=1}^{20} \beta_i\, w_i + \sum_{i<j}^{20} \beta_{ij}\, w_i w_j + \sum_{i<j<l}^{20} \beta_{ijl}\, w_i w_j w_l + 5\, w_1 w_2 w_3 w_4, \tag{4.5}$$

  where $X_i \sim \mathcal{U}(0, 1),\ i = 1, \dots, 20,\ w_i = 2\,(x_i - 1/2)$ for all i except for $i = 3, 5, 7$ where $w_i = 2\,(\dfrac{1.1\,x_i}{x_i + 0.1} - 1/2)$. The coefficients are defined as: $\beta_i = 20$ for $i = 1, \dots, 10$, $\beta_{ij} = -15$ for $i, j = 1, \dots, 6$, and $\beta_{ijl} = -10$ for $i, j, l = 1, \dots, 5$. The remaining coefficients are set equal to $\beta_i = (-1)^i$ and $\beta_{ij} = (-1)^{i+j}$ as in Blatman (2009).

- The *Rastrigin* function has a two-dimensional input space and is defined by (Rastrigin 1974):
$$f_5(\boldsymbol{x}) = 10 - \sum_{i=1}^{2} (x_i^2 - 5\,\cos(2\pi\, x_i)), \tag{4.6}$$
where $X_i \sim \mathcal{N}(0,1),\ i = 1, 2$ is modelled as a standard normal variable.

- The last function is the *O'Hagan* function which is defined by (Oakley & O'Hagan 2004):
$$f_6(\boldsymbol{x}) = \boldsymbol{a}_1^{\mathsf{T}}\boldsymbol{x} + \boldsymbol{a}_2^{\mathsf{T}}\sin(\boldsymbol{x}) + \boldsymbol{a}_3^{\mathsf{T}}\cos(\boldsymbol{x}) + \boldsymbol{x}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{x}, \tag{4.7}$$
where $X_i \sim \mathcal{N}(0,1),\ i = 1, \ldots, 15$. The vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2.\boldsymbol{a}_3\}$ and matrix $\boldsymbol{Q}$ are defined in Oakley & O'Hagan (2004) and given in Appendix D for the sake of completeness.

Note that the functions f_1 to f_4 have uniform input random variables. Accordingly, the PC trend in PC-Kriging is built up from multivariate Legendre polynomials. In contrast f_5 and f_6 have Gaussian input random variables. Thus, the PC trend is modelled by multivariate Hermite polynomials (see also Sudret (2015)).

4.4.2 Analysis

At the beginning of each algorithm, the experimental design is generated with the Latin-hypercube sampling technique (McKay, Beckman, & Conover 1979) in the unit-hypercube $[0,1]^M$ and a subsequent isoprobabilistic transform to the input domain $\mathcal{D}_{\boldsymbol{X}}$. Then, meta-models are calibrated applying the four previously discussed meta-modelling techniques, *i.e.* ordinary Kriging, PCE, SPC-Kriging, and OPC-Kriging. Their performance is compared by means of relative generalization error (see Eqs. (3.19) and (3.40)) based on a validation set of $n = 10^5$ samples. Note that the validation set is sampled from the input vector $\boldsymbol{X}$ with the Monte Carlo method.

For each experimental setup, the analysis is replicated 50 times to account for the statistical uncertainties in the experimental design. The results of the 50 independent runs are represented by box plots. In a box plot, the central mark visualizes the median value of the 50 runs, the edges are the 25-th and 75-th percentile denoted by q_{25} and q_{75}, respectively. The whiskers describe the boundary to the outliers, which are defined as the values smaller than $q_{25} - 1.5\,(q_{75} - q_{25})$ or larger than $q_{75} + (q_{75} - q_{25})$.

4.4.3 Results

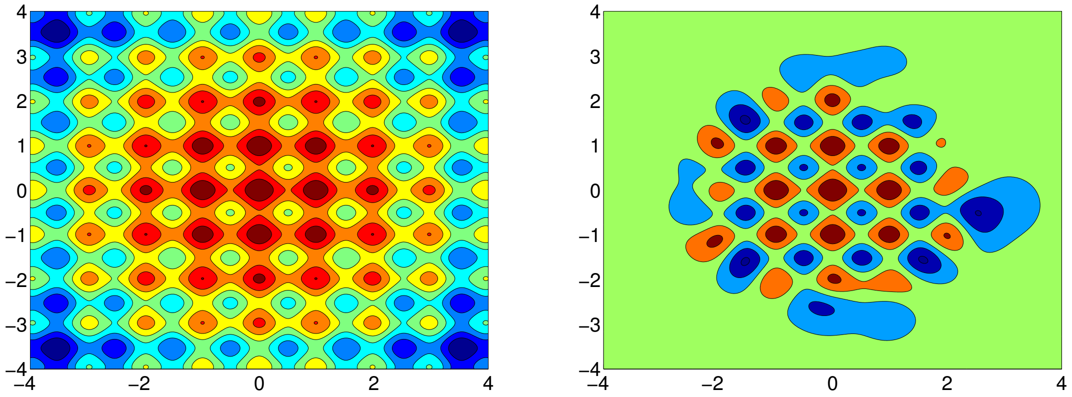
Visualization of PC-Kriging's behaviour

The general behaviour of the different types of meta-models are illustrated in Figure 4.3. Figure 4.3(a) shows a two-dimensional contour plot of the model response of the Rastrigin function as a function of x_1 and x_2. The model response is compared to PCE, ordinary Kriging and PC-Kriging meta-models (Figures 4.3(b)-4.3(d)), which are based on an experimental design of size $N = 128$. The Rastrigin function has a highly oscillatory behaviour over the entire input space as seen in Figure 4.3(a). This behaviour is difficult to meta-model with a small number of samples because many local minima/maxima are missed out.

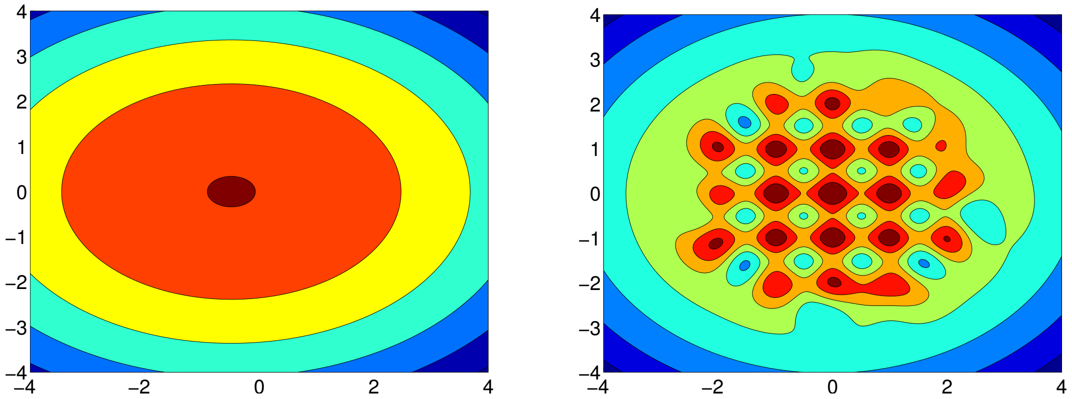
The analytical formulation of the Rastrigin function is a combination of a quadratic component and a high-frequency trigonometric component. The PCE model in Figure 4.3(c) captures the global characteristic of the function, *i.e.* the quadratic component, whereas the ordinary

Kriging model in Figure 4.3(b) approximates the local characteristics, *i.e.* the high-frequency trigonometric component. Finally, the combination of PCE and Kriging leads to a better meta-model as shown in Figure 4.3(d).

Note that the meta-models in Figure 4.3 have a high accuracy around the origin of the coordinate system due to the definition of the input vector PDF as standard normal distributions ($X_i \sim \mathcal{N}(0,1)$).



(a) Rastrigin function (b) Ordinary Kriging



(c) PCE (d) PC-Kriging

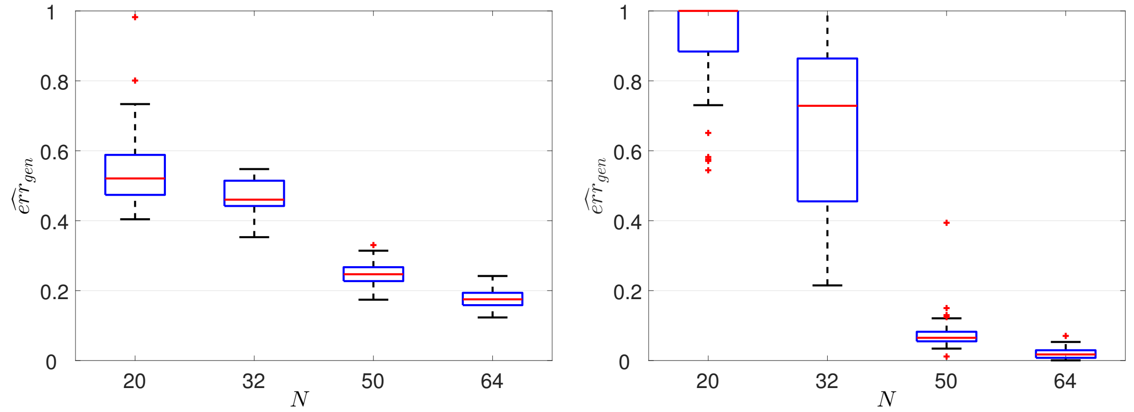
Figure 4.3: Rastrigin function – visual composition of PC-Kriging.

Small experimental design

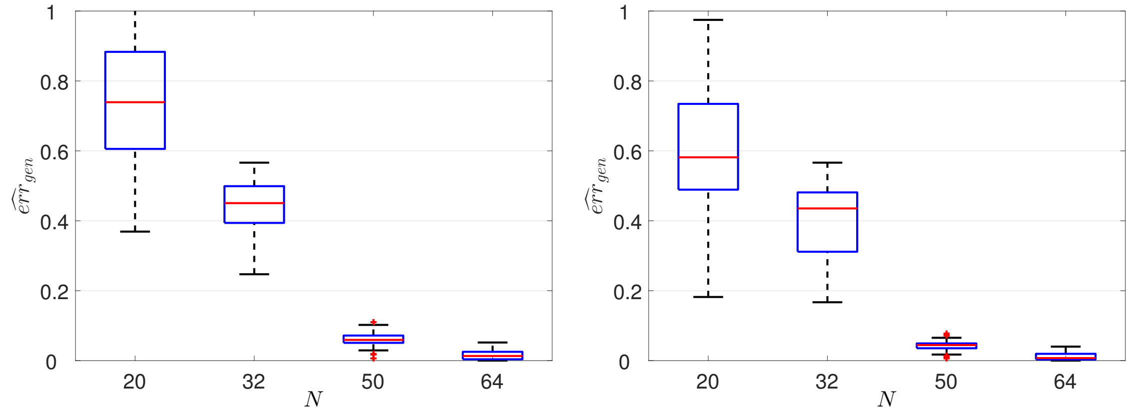
The four meta-modelling techniques are compared for the six analytical functions f_1 to f_6 using experimental designs of increasing size N. The number of samples is chosen so that the analysis yields a large range of relative generalization errors. The results are illustrated in Figures 4.4 to 4.9. In each figure, (a) shows the ordinary Kriging model and (b) shows the PCE model. The new approaches SPC-Kriging and OPC-Kriging are shown in (c) and (d), respectively.

Ishigami function Figure 4.4 shows the relative generalization error for the Ishigami function. For a small sample size of $N = 20$ samples, ordinary Kriging performs best with respect to the median value of the box plots. When the number of samples is increased, the two PC-Kriging approaches perform better than the traditional approaches because their median value and the range of the error is lower. For large sample sizes ($N \geq 50$), PC-Kriging performs similarly to

PCE, though slightly better. OPC-Kriging is slightly more accurate than SPC-Kriging over the whole range of sample sizes.



(a) Ordinary Kriging (b) PCE



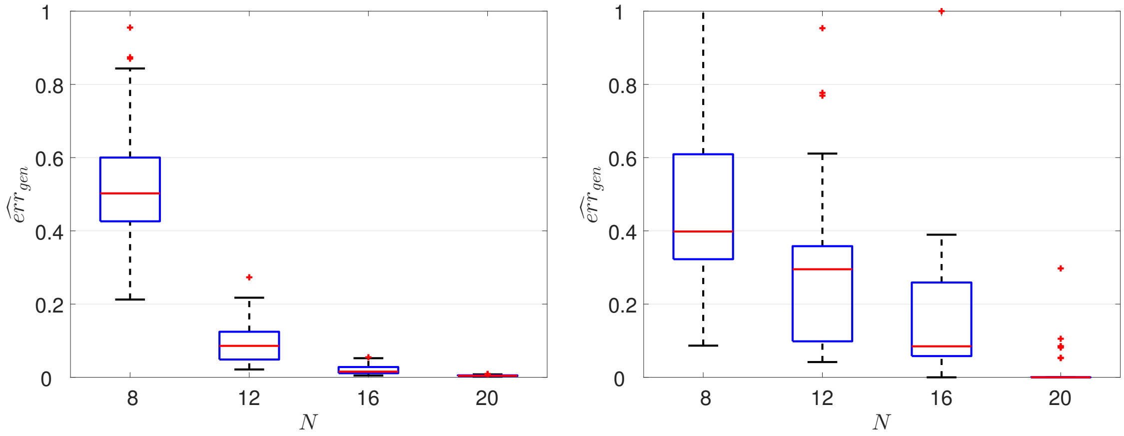
(c) SPC-Kriging (d) OPC-Kriging

Figure 4.4: Ishigami function – relative generalization error for four meta-modelling techniques.

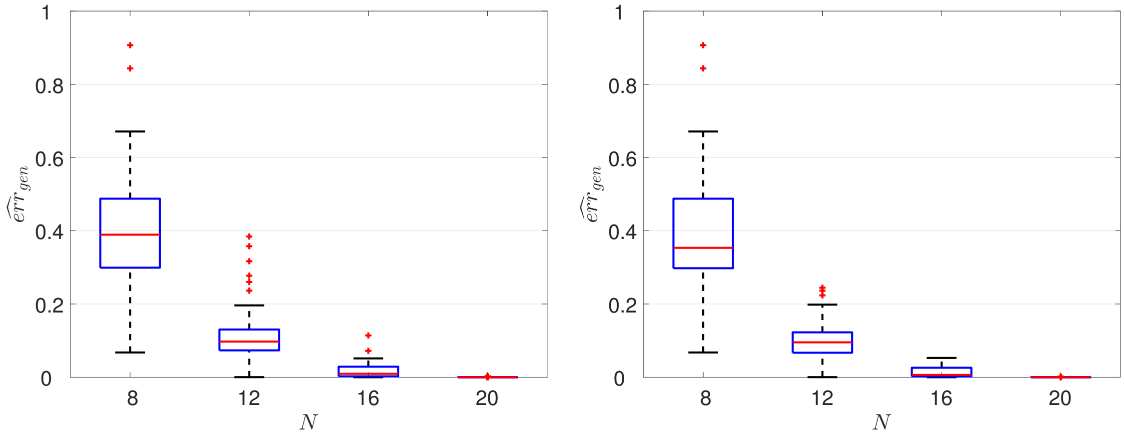
Rosenbrock function Figure 4.5 presents the results of the Rosenbrock function, which is a purely polynomial function and can be modelled accordingly with a small number of polynomials based on a small number of points in the experimental design (*e.g.* in the case of PCE). This is the reason why the number of points lies within $N = 8, \ldots, 20$. Highly accurate surrogate models are obtained with only 20 samples. For small sample sizes OPC-Kriging performs best among the four techniques in terms of the relative generalization error.

Sobol' function The Sobol' function is more complex than the two previous functions because of the dimensionality ($M = 8$) and the non-smooth behaviour at $x_i = 0.5,\ i = 1, \ldots, 8$. Thus, more samples are needed to obtain a similar range of relative generalization errors compared to the previous functions, as seen in Figure 4.6. Behaviours for the larger sample sizes ($N = 64,\ 128$) are very similar among the meta-modelling approaches, although PC-Kriging performs slightly better than the traditional PCE and ordinary Kriging approaches. For very small sample sizes ($N = 16,\ 32$), OPC-Kriging performs significantly better than the other meta-models.

Morris function Figure 4.7 shows the results for the Morris function. A large experimental design is required to properly surrogate the computational model because of the high dimensionality of the input vector $\boldsymbol{X}$ and the amount of interactive terms of different input variables



(a) Ordinary Kriging (b) PCE



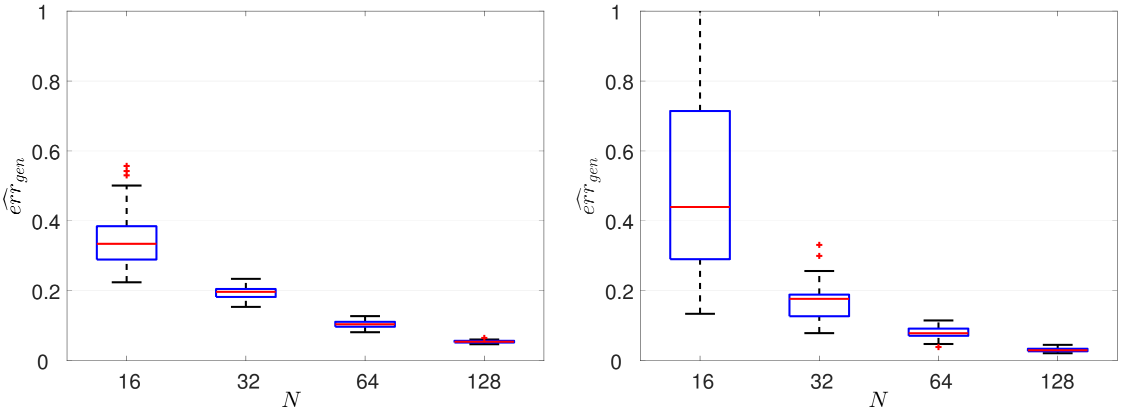
(c) SPC-Kriging (d) OPC-Kriging

Figure 4.5: Rosenbrock function – relative generalization error for four meta-modelling techniques.

X_i in the analytical formulation (see Eq. (4.5)). The relative generalization error of the two PC-Kriging approaches resembles more the one of ordinary Kriging than the one of PCE in this case. Sparse PCE is not capable of modelling this analytical function with a small number of samples, due to the large number of interactive terms in the formulation of the function at hand.

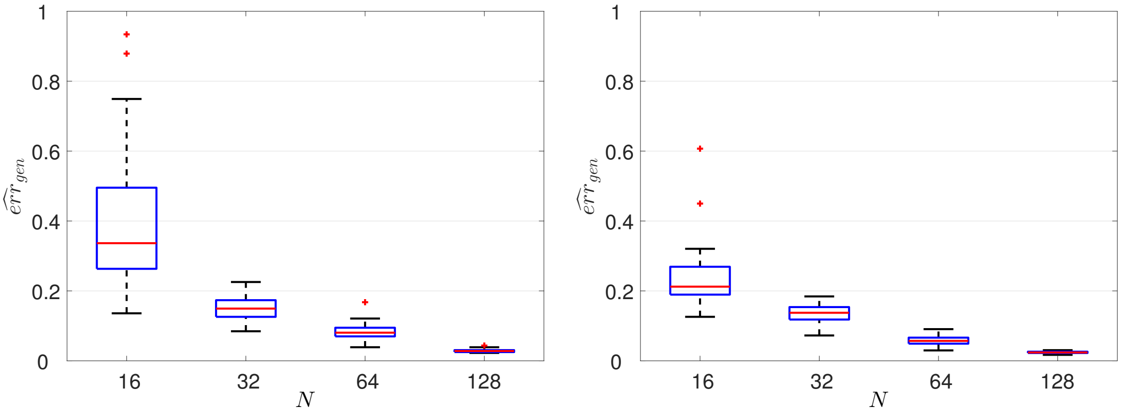
Rastrigin function The results associated with the Rastrigin function are shown in Figure 4.8. Despite the low dimensionality of the input ($M = 2$), many samples are needed to obtain small error estimates. This is due to the fact that the function output is highly oscillatory over the entire input space as previously illustrated in Figure 4.3. In comparison to Figure 4.3, which describes the qualitative performance of PC-Kriging on the Rastrigin function, the quantitative benefit of combining PCE and Kriging becomes visible in Figure 4.8: PC-Kriging performs better than the traditional approaches. Ordinary Kriging performs the worst followed by PCE. OPC-Kriging has statistically the lowest relative generalization errors over the whole range of experimental design sizes.

O'Hagan function Figure 4.9 displays the results associated with the O'Hagan function. Similarly to the Morris function, the performance of PC-Kriging in the case of the O'Hagan function resembles that of ordinary Kriging whereas PCE performs worse than the other three



(a) Ordinary Kriging

(b) PCE



(c) SPC-Kriging

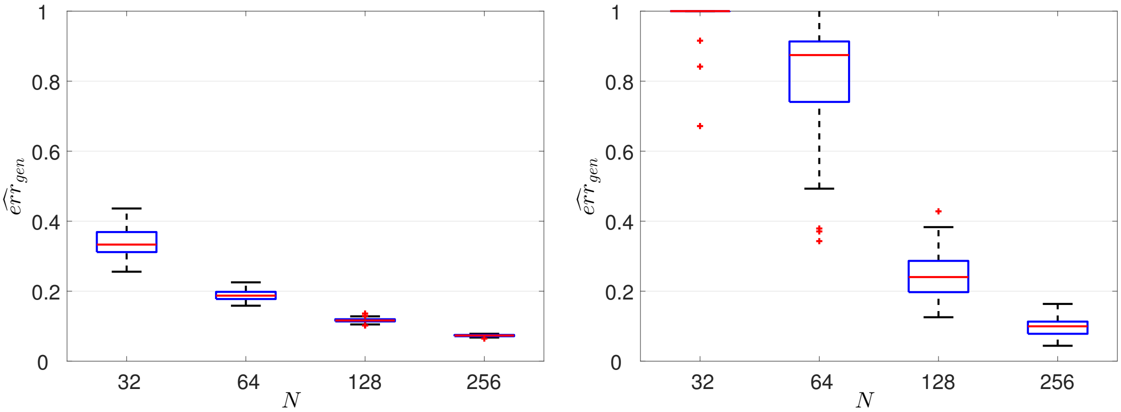
(d) OPC-Kriging

Figure 4.6: Sobol' function – relative generalization error for four meta-modelling techniques.

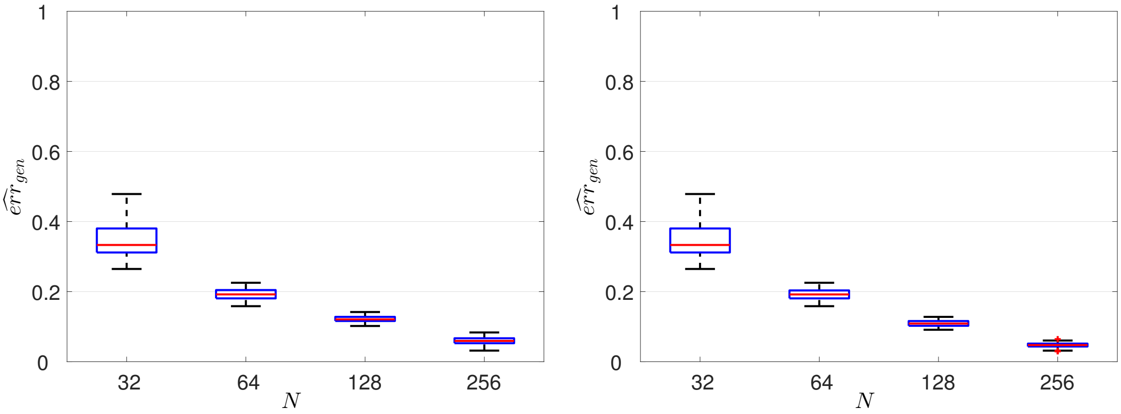
approaches. Over the entire displayed range of experimental designs in Figure 4.9, the performance of OPC-Kriging is slightly better than the performance of SPC-Kriging and ordinary Kriging. Note that meta-modelling the O'Hagan function requires less samples in the experimental design to obtain the same accuracy as in the case of the Rastrigin function despite the fact that the O'Hagan function has a 15-dimensional input space and that both functions are very smooth.

Accuracy comparison Summarizing the results of all six analytical functions in Figures 4.4-4.9, the proposed PC-Kriging approaches perform better than or at least as good as the traditional PCE and Kriging approaches. Note that for functions like Morris (Figure 4.7) and O'Hagan (Figure 4.9), the performance of PC-Kriging is more similar to Kriging than PCE, whereas for the other functions the performance of PC-Kriging resembles more that of PCE. As one could expect, there is no general rule so as to decide whether PCE or Kriging provides the most accurate meta-models for a given experimental design. The advantage of PC-Kriging is to perform as least as well as the best of the two.

Computational costs The combination of PCE and Kriging and its increased accuracy comes with a higher computational cost. The traditional ordinary Kriging and PCE approaches have the lowest computational cost, SPC-Kriging has an intermediate and OPC-Kriging has the highest cost. The high cost of OPC-Kriging originates from the iterative character of the algorithm and the accompanying repetitive calibration of Kriging models. For a single calibration of



(a) Ordinary Kriging (b) PCE



(c) SPC-Kriging (d) OPC-Kriging

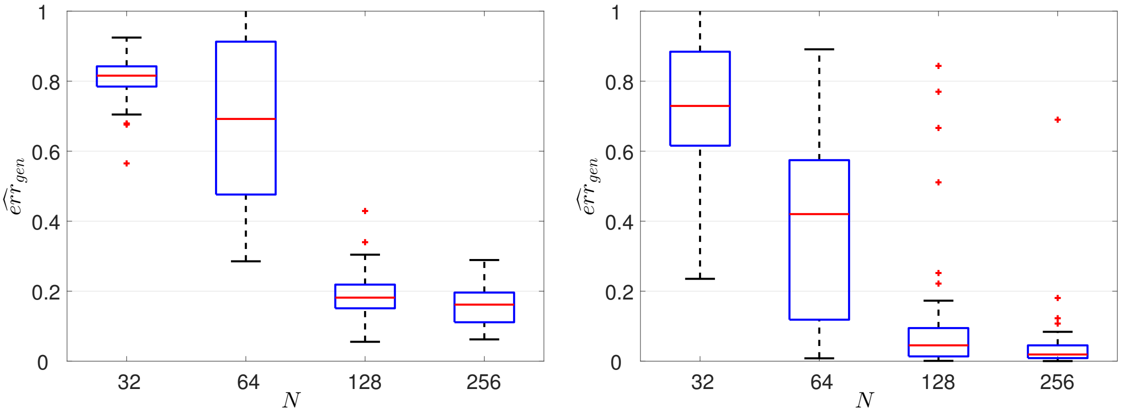
Figure 4.7: Morris function – relative generalization error for four meta-modelling techniques.

a surrogate of the Ishigami function (experimental design of size $N = 128$ samples) the ratio of computational times when comparing PCE to ordinary Kriging, SPC-Kriging and OPC-Kriging is approximately $1 : 5$, $1 : 20$ and $1 : 200$, respectively.

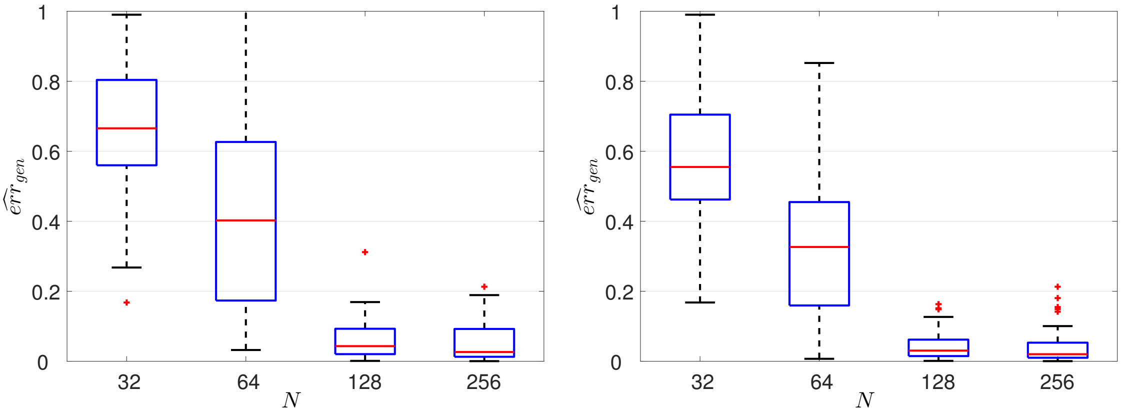
Note that it is intended to apply these techniques to realistic problems where the evaluation of the exact computational model response lasts much longer than the computation of a meta-model. Then, the apparent computational overload of OPC-Kriging will not be anymore an issue in many practical applications.

Large experimental design

When the resources for experiments are limited, the focus lies on doing as few computational model runs as possible, as discussed in the previous section. However, in order to describe the entire behaviour of the meta-modelling approaches, the Ishigami function is also studied for larger experimental designs in order to assess the convergence of the various schemes. Figure 4.10 illustrates the evolution of the relative generalization error from small to large sample sizes on the *logarithmic* (base 10) scale. When using PC-based approaches, the relative generalization error decreases fast when enlarging the sample set. As the Ishigami function is composed of sine and cosine functions, an approximation with series of polynomials works well (see also Taylor series expansions). Hence, PCE and PC-Kriging result in highly accurate meta-models. In case of PC-Kriging, the polynomial trend is the dominating effect whereas the Gaussian random process is almost negligible.



(a) Ordinary Kriging (b) PCE



(c) SPC-Kriging (d) OPC-Kriging

Figure 4.8: Rastrigin function – relative generalization error for four meta-modelling techniques.

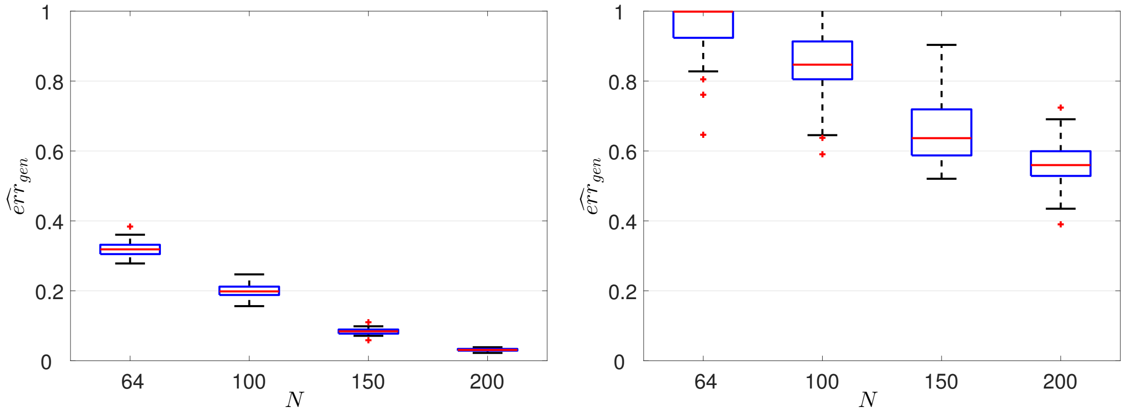
For large sample sizes, ordinary Kriging is outperformed by the other three approaches. In general, Kriging works well with small sample sizes. If too many samples are used, the interpolation algorithm becomes unstable due to singularities and bad conditioning in the correlation matrix $\mathbf{R}$. If a large sample size is available, regional Kriging models on a subset of samples, *e.g.* the neighbouring samples, are more suitable (Dubrule 1983).

The performance of PC-Kriging and also PCE for a large number of samples (here 128 and 256 samples in Figure 4.10) is in the order of magnitude of the machine computation precision. Errors around $\widehat{err}_{gen} \approx 10^{-12}$ originate from numerical round-off errors which are not reducible by adding more samples. It is questionable though, whether in reality such a high accuracy in the meta-model prediction is needed.

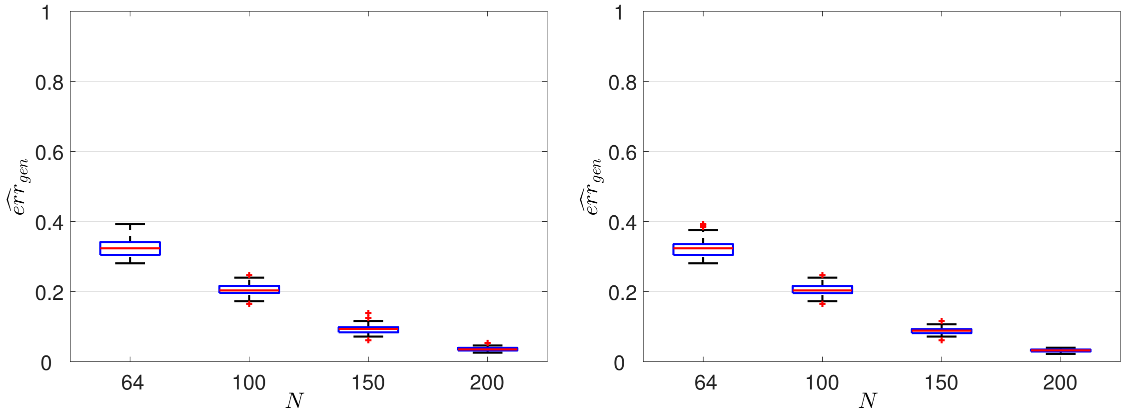
Evolution of the error measures

The OPC-Kriging algorithm includes the tracking of the LOO error to optimally choose the sparse set of orthonormal polynomials. Hence, the evolution of the LOO error for the Ishigami function and a sample size of $N = 128$ samples is presented in Figure 4.11. The experimental-design-based LOO error (dashed line) is compared to the relative generalization error which is shown as the solid black line.

The first point to notice is that the LOO error slightly under-predicts the true value of the relative generalization error for all sizes of polynomial sets. This is due to the fact that the LOO error is based solely on the information contained in the experimental design samples,



(a) Ordinary Kriging (b) PCE



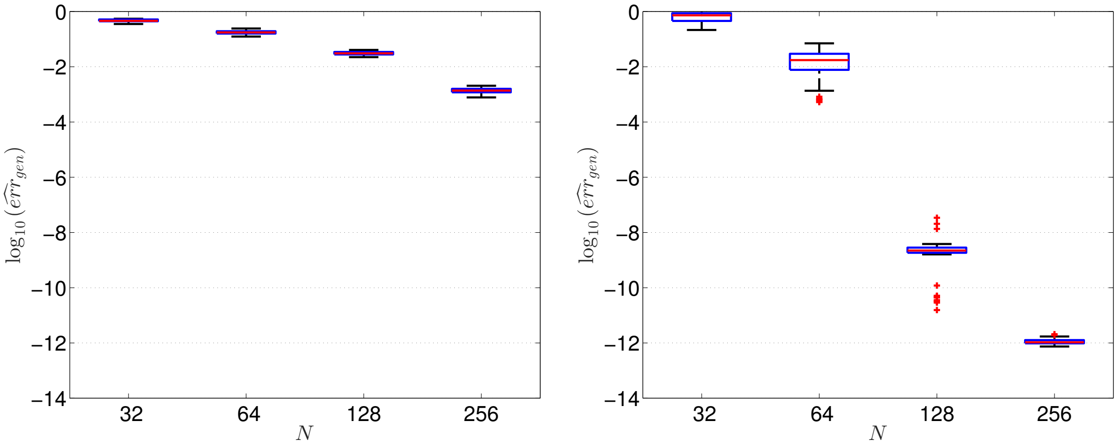
(c) SPC-Kriging (d) OPC-Kriging

Figure 4.9: O'Hagan function – relative generalization error for four meta-modelling techniques.

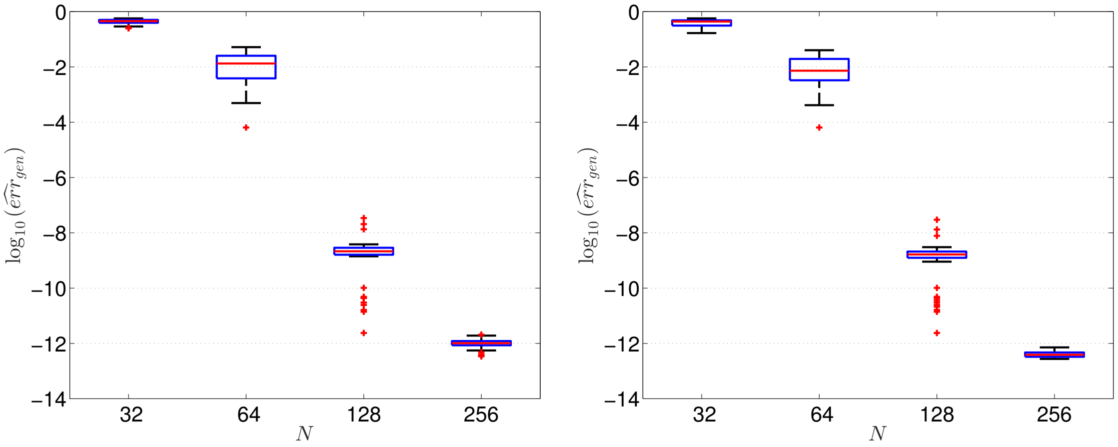
whereas the relative generalization error is based on a large validation set ($n = 10^5$). Although there is an inherent under-prediction, the overall behaviour of the two error measures is similar. Thus, the choice of the optimal set of polynomials for the OPC-Kriging can be based on the LOO error, which is obtained as a by-product of the procedure used to fit the parameters of the PC-Kriging model. In the example case of Figure 4.11, choosing only half of the polynomials, *i.e.* $i^* = 27$, leads to a meta-model which is almost as accurate as using all 56 polynomials. The optimal set of polynomials can be chosen at the point where the decrease in LOO error becomes insignificant. This reduces the number of polynomials needed and thus also reduces the complexity of the OPC-Kriging meta-model.

4.5 Conclusion

PC-Kriging is a combination of the traditional non-intrusive PCE and Kriging methods. In other words, PC-Kriging is a generalization of PCE and Kriging. In fact in some applications, PC-Kriging tends to behave like an ordinary Kriging model, whereas in other applications it tends to behave like PCE. In any examined cases, however, PC-Kriging performs better than the two traditional meta-modelling techniques taken separately. Hence, PC-Kriging is an efficient meta-modelling technique which allows for an efficient propagation of uncertainty in the context of probability distributions.



(a) Ordinary Kriging (b) PCE



(c) SPC-Kriging (d) OPC-Kriging

Figure 4.10: Ishigami function – relative generalization error for four meta-modelling techniques and large experimental designs.

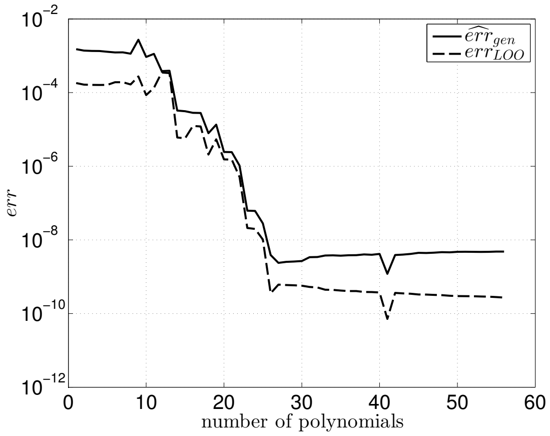


Figure 4.11: Evolution of the LOO error and relative generalization error inside the OPC-Kriging algorithm as a function of the number of polynomials in the trend.

CHAPTER 5

PC-Kriging in rare event estimation

PC-Kriging is used in structural reliability analysis, where adaptive experimental design algorithms are used to increase the efficiency of estimating failure probabilities and quantiles. Different aspects of this topic have been published in Schöbi & Sudret (2014b,c,e), and Schöbi, Sudret, & Marelli (2016).

5.1 State of the art

Apart from pure uncertainty propagation (as discussed in Chapter 3 and 4), analysts are often interested in the estimation of rare events. Two cases are generally distinguished in the context of rare event estimation, namely the estimation of failure probabilities (so-called structural reliability analysis) and that of extreme quantiles. Both quantities describe important statistics of the response of the computational models. In the context of engineering and high reliability targets, failure probabilities are generally low (in the order of 10^{-3} to 10^{-6}) and the quantiles are extreme (*e.g.* 99 % quantiles), therefore structural reliability analysis falls into the broad class of rare events estimation. Note that traditionally "structural reliability analysis" was developed in the context of civil engineering structures. However, the same methods can be applied in other fields, hence the more general name of "rare event estimation" and simply "reliability analysis".

A large number of tools have been proposed and are widely available for solving such reliability analyses. Methods include the traditional Monte Carlo simulation (MCS) (Metropolis & Ulam 1949; Hammersley & Handscomb 1964), first order reliability method (FORM) (Hasofer & Lind 1974; Rackwitz & Fiessler 1978), second order reliability analysis (SORM) (Hohenbichler et al. 1987; Breitung 1989), subset simulation (Au & Beck 2001; Au 2016), line sampling (Koutsourelakis, Pradlwarter, & Schuëller 2004; Koutsourelakis 2004), importance sampling (Melchers 1989; Engelund & Rackwitz 1993), and meta-model-based approaches. For a more detailed overview and applications of structural reliability analyses, see *e.g.* Ditlevsen & Madsen (1996), Melchers (1999), Lemaire (2009), and Morio, Balesdent, et al. (2014).

In the context of structural reliability analysis and meta-modelling, Kriging is often connected to the idea of adaptive design enrichment strategies. The reason is that Kriging provides a local error estimate of the meta-model accuracy by the prediction variance $\sigma^2_{\widehat{Y}}(\boldsymbol{x})$. Kaymaz (2005) and Schueremans & Van Gemert (2005b) introduced Kriging to structural reliability analysis. The use of adaptive experimental design algorithms has been investigated by Bichon, Eldred, et al. (2008), Echard, Gayton, & Lemaire (2011), Dubourg, Sudret, & Bourinet (2011), and Ranjan, Bingham, & Michailidis (2012) who combine an adaptive Kriging meta-model with Monte Carlo simulation, and Echard, Gayton, Lemaire, & Relun (2013) and Cadini, Santos, & Zio (2014), who modify it for the case of importance sampling. Under the name "stepwise uncertainty reduction" Arnaud et al. (2010), Bect et al. (2012), and Jala (2013) discuss an adaptive Kriging approach in a Bayesian context. Dani, Hayes, & Kakade (2008) and Srinivas et al. (2012) use upper confidence bounds. Finally, the use of adaptive Kriging to build up a quasi-optimal importance sampling density (thus leading to unbiased estimates of the failure probability) has been recently proposed in Dubourg, Sudret, & Deheeger (2013) and Dubourg & Sudret (2014).

Other meta-modelling techniques have been applied in fewer occasions. Bourinet, Deheeger, & Lemaire (2011) use support vector machines in the context of subset simulations. Dai et al. (2014) discuss the use of wavelet-density-based adaptive importance sampling. Marelli & Sudret (2016) use bootstrapping to enrich the experimental design of a PCE meta-model.

Due to their versatility and popularity, Kriging meta-models with experimental design strategies are discussed in the following in the context of structural reliability methods. In particular, existing approaches are expanded to the new PC-Kriging algorithms.

5.2 Structural reliability analysis

5.2.1 Limit-state function

In the context of structural reliability analyses, a *limit-state function* describes the performance of a process or system as a function of a set of input parameters. The deterministic mapping is defined as:

$$\mathcal{G}: \ \boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} \subset \mathbb{R}^M \ \mapsto \ y = \mathcal{G}(\boldsymbol{x}) \in \mathbb{R}, \tag{5.1}$$

where $\boldsymbol{x}$ is a M-dimensional vector defined in $\mathcal{D}_{\boldsymbol{X}}$ and y is the output scalar indicating the performance. The sign of y, and hence $\mathcal{G}(\boldsymbol{x})$, determines whether an input vector $\boldsymbol{x}$ corresponds to a safe system ($\mathcal{G}(\boldsymbol{x}) > 0$) or a failed system ($\mathcal{G}(\boldsymbol{x}) \leq 0$).

The limit-state function $\mathcal{G}$ and the computational model $\mathcal{M}$ are closely related. For simplicity, it is assumed that the input vector $\boldsymbol{x}$ is identical for $\mathcal{G}$ and $\mathcal{M}$. Moreover the limit-state function can be interpreted as a function of the computational model, *i.e.* $\mathcal{G}(\boldsymbol{x}) = f(\mathcal{M}(\boldsymbol{x}))$. In practice, this function is often of simple form, such as a translation. As an example, let us consider that the computational model $\mathcal{M}$ is a finite element model of a building. The response value $\mathcal{M}(\boldsymbol{x})$ may be the displacement at the top floor. Then, a possible limit-state function could describe whether the top floor displacement exceeds a threshold value of y_{adm}. This can be modelled by the following limit-state function:

$$\mathcal{G}(\boldsymbol{x}) = y_{\text{adm}} - \mathcal{M}(\boldsymbol{x}). \tag{5.2}$$

In other examples, it can be that the computational model is modelling the performance itself, hence $\mathcal{G}(\boldsymbol{x}) = \mathcal{M}(\boldsymbol{x})$. Due to this close relationship between $\mathcal{G}$ and $\mathcal{M}$, only $\mathcal{G}$ is used in the following section in the context of structural reliability analysis. The use of $\mathcal{M}$ instead would not change the spirit of the methods proposed in the following.

5.2.2 Estimation of a failure probability

In structural reliability analysis, the *failure probability* P_f is defined as the probability that the limit-state function takes negative values:

$$P_f = \mathbb{P}\left(\mathcal{G}\left(\boldsymbol{X}\right) \leq 0\right). \tag{5.3}$$

The failure probability may be recast as the following integral:

$$P_f = \int_{\mathcal{D}_f} f_{\boldsymbol{X}}(\boldsymbol{x})\ \mathrm{d}\boldsymbol{x}, \tag{5.4}$$

where $\mathcal{D}_f = \{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} :\ \mathcal{G}\left(\boldsymbol{x}\right) \leq 0\}$ is the failure domain and $f_{\boldsymbol{X}}$ is the joint PDF of the input vector $\boldsymbol{X}$. The integration operation in Eq. (5.4) cannot be performed analytically in the general case where the failure domain $\mathcal{D}_f$ has a complex shape. Hence, numerical estimates for the failure probability were developed such as Monte Carlo simulation. Considering a large sample of $\boldsymbol{X}$ denoted by $\mathcal{S} = \left\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}\right\}$, the failure probability can be estimated by:

$$\widehat{P}_f = \frac{n_f}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\mathcal{G}(\boldsymbol{x})\leq 0}\left(\boldsymbol{x}^{(i)}\right), \tag{5.5}$$

where n_f is the number of failure samples $\boldsymbol{x}^{(i)} \in \mathcal{D}_f$, $n = |\mathcal{S}|$ is the total number of samples and $\mathbb{I}$ is the indicator function with $\mathbb{I} = 1$ for failure and $\mathbb{I} = 0$ otherwise. Note that the use of $\mathbb{I}$ transforms the structural reliability problem into a classification problem where only the sign of $\mathcal{G}(\boldsymbol{x})$ is relevant.

An advantage of the Monte Carlo simulation is that the accuracy of the failure probability estimate can be computed easily. The theoretical coefficient of variation of the failure portability estimate reads:

$$\mathrm{CoV}\left[\widehat{P}_f\right] = \sqrt{\frac{1 - \widehat{P}_f}{n \cdot \widehat{P}_f}}. \tag{5.6}$$

The main limitation of pure Monte Carlo approaches, namely the loss of efficiency for small failure probabilities, is apparent in Eq. (5.6). The number of samples required to obtain a target coefficient of variation is rapidly increasing with decreasing failure probability. For instance, to achieve a target coefficient of variation of $\mathrm{CoV}\left[\widehat{P}_f\right] = 0.1$ given a failure probability $P_f = 10^{-k}$, $n \approx \left[\widehat{P}_f \cdot \mathrm{CoV}\left[\widehat{P}_f\right]^2\right]^{-1} = 10^{k+2}$ samples are needed.

5.2.3 Estimation of a quantile

Quantile estimation consists in determining a quantile q_α so that the probability of $\mathcal{G}\left(\boldsymbol{X}\right)$ being smaller than q_α is equal to a preset value α:

$$\mathbb{P}\left(\mathcal{G}\left(\boldsymbol{X}\right) \leq q_\alpha\right) = \alpha, \tag{5.7}$$

where $\alpha \in [0, 1]$. This constitutes an inverse problem compared to classical failure probability estimation shown in Eq. (5.3).

The quantile estimation problem can also be solved by Monte Carlo simulation. Consider again a reasonably large sample set $\mathcal{S} = \left\{\boldsymbol{x}^{(1)}, ..., \boldsymbol{x}^{(n)}\right\}$ and the corresponding response values $y = \left\{y^{(1)}, ..., y^{(n)}\right\} = \left\{\mathcal{G}\left(\boldsymbol{x}^{(1)}\right), \ldots, \mathcal{G}\left(\boldsymbol{x}^{(n)}\right)\right\}$. Assume that the response values are ranked in ascending order and denote them by $y_{(1)} \leq y_{(2)} \leq \ldots \leq y_{(n)}$. Then the estimator of the quantile q_α reads:

$$\widehat{q}_\alpha = y_{\lfloor n\alpha \rfloor}, \tag{5.8}$$

where $\lfloor n\alpha \rfloor$ is the largest integer smaller than $n\alpha$. Analogously to failure probability estimation, the quantile estimation problem can be cast as a classification problem:

$$\text{find} \quad \widehat{q}_\alpha \quad \text{s.t.} \quad \alpha \approx \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\mathcal{G}(\boldsymbol{x}) \leq \widehat{q}_\alpha} \left(\boldsymbol{x}^{(i)} \right). \tag{5.9}$$

5.2.4 Limit-state surface

Both the estimation of failure probabilities in Eq. (5.5) and of quantiles in Eq. (5.9) may be treated as classification problems dividing the input domain into failure and safety regions. The *limit-state surface* marks the boundary between the failure and safety domain. Therefore, the *limit-state parameter* y_0 is introduced, which determines whether a response y belongs to the failure ($y \leq y_0$) or to the safety domain ($y > y_0$). For the estimation of failure probabilities $y_0 = 0$, whereas for the estimation of quantiles $y_0 = \widehat{q}_\alpha$. Due to these similarities, failure probability and quantiles are denoted as statistics of interest in the remaining of the chapter.

5.2.5 Model accuracy

In the context of engineering and high reliability targets, failure probabilities are generally low (in the order of 10^{-3} to 10^{-6}) and quantiles are extreme (*e.g.* 99% quantile). Therefore, structural reliability falls into the broad class of rare event estimation. Hence, the failure domain $\mathcal{D}_f$ is either small and/or far away from the mean value of the input distribution $\boldsymbol{X}$ in the input space $\mathcal{D}_{\boldsymbol{X}}$. The meta-modelling techniques introduced so far focus, however, on the area around the mean of the input distribution rather than on the distribution's tail. Moreover, they use a static experimental design which is independent on the limit-state surface that characterizes the classification problems introduced in Eqs. (5.5) and (5.9). Thus, the accuracy of the estimation of the statistics of interest is generally low using these techniques.

In order to improve the accuracy of this techniques, rare event estimation has been often related to the concept of experimental design enrichment (a.k.a. adaptive experimental design algorithms). The core idea is to start calibrating a meta-model with a limited set of limit-state function runs and then to add new evaluations $\boldsymbol{x}$ iteratively to increase the accuracy of the statistics of interest.

5.3 Adaptive experimental design algorithm

5.3.1 Main algorithm

A slightly modified version of the *Adaptive leaning reliability methods combining Kriging and Monte Carlo Simulation* (AK-MCS) (Echard, Gayton, & Lemaire 2011) algorithm is introduced here making use of the PC-Kriging meta-models presented in the previous chapter. The main steps can be summarized as follows:

(i) *Initial experimental design*: The initial experimental design $\boldsymbol{\mathcal{X}}$ (composed of N_0 samples) is generated by a space-filling design (*e.g.* Latin-hypercube sampling) and the corresponding response values $\mathcal{Y}^{(i)} = \mathcal{G}\left(\chi^{(i)}\right)$ are computed.

(ii) *Calibration of meta-model* $\widehat{\mathcal{G}}$: A meta-model $\widehat{\mathcal{G}}$ is calibrated based on $\{\boldsymbol{\mathcal{X}}, \mathcal{Y}\}$. In the following, PC-Kriging and ordinary Kriging models are used for comparison.

(iii) *MC population* $\mathcal{S}$: A large set of candidate samples $\mathcal{S} = \left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n)}\right\}$ is generated from $\boldsymbol{X}$ by Monte Carlo simulation. This candidate set is the basis for estimating the statistics of interest later on.

(iv) *Computation of meta-model responses*: Compute the prediction mean $\mu_{\widehat{Y}}(\boldsymbol{x})$ and variance $\sigma^2_{\widehat{Y}}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \mathcal{S}$.

(v) *Limit-state parameter* y_0: Estimation of the limit-state parameter y_0 based on the current meta-model, *i.e.* based on $\left\{\mu_{\widehat{Y}}(\boldsymbol{x}),\ \boldsymbol{x} \in \mathcal{S}\right\}$. Note that in the case of failure probability estimation, the parameter is automatically set to $y_0 = 0$.

(vi) *Probability of misclassification*: All candidate samples $\boldsymbol{x} \in \mathcal{S}$ are ranked according to an enrichment criterion.

(vii) *Candidate selection*: In the selection step, the optimal sample(s) $\boldsymbol{\chi}^*$ which are to be added to the experimental design are determined, so that $\boldsymbol{\mathcal{X}} \leftarrow \{\boldsymbol{\mathcal{X}}, \boldsymbol{\chi}^*\}$.

(viii) *Addition of* $\boldsymbol{\chi}^*$: The response values corresponding to $\boldsymbol{\chi}^*$ is/are evaluated and added to $\mathcal{Y} \leftarrow \{\mathcal{Y}, \mathcal{Y}^*\}$. Then, the iterative algorithm goes back to step (ii), where the meta-model $\widehat{\mathcal{G}}$ is recalibrated with the enriched experimental design.

(ix) *Stopping criterion*: The iterative part of the algorithm is terminated through a convergence measure (stopping criterion) with respect to the statistics of interest.

(x) *End of algorithm*: After the termination of the iterative algorithm, the statistics of interest is estimated based on the last meta-model $\widehat{\mathcal{G}}$ and the candidate set $\mathcal{S}$.

An overview of the algorithm is shown in the flowchart in Figure 5.1. The white boxes mark the information required to start the algorithm, whereas the grey boxes represent the computational tasks.

5.3.2 Initial experimental design

In order to have a well-behaved initial experimental design, a space-filling sampling method is required. In this thesis, the Latin-hypercube sampling algorithm is used (McKay, Beckman, & Conover 1979). In particular, the unit hyper-cube $[0, 1]^M$ is sampled and the samples are mapped to the input space $\mathcal{D}_{\boldsymbol{X}}$ by the inverse CDF of the marginals. Other sampling strategies include quasi-random sequences, such as Sobol' sequences (Sobol' 1967) and Halton series (Halton 1960).

5.3.3 Candidate selection

Single sample selection

A *learning function* (LF) estimates the expected value of information gained with respect to the statistics of interest when point $\boldsymbol{x}$ is added to the experimental design $\boldsymbol{\mathcal{X}}$ of the meta-model. A variety of learning functions has been proposed in the literature including the *Expected Feasibility Function* (EFF) (Bichon, Eldred, et al. 2008), the *Expected Improvement* (EI) functions (Jones, Schonlau, & Welch 1998; Ginsbourger, Rosspopoff, et al. 2013), the *Gaussian process upper confidence bounds* (GP-UCB) (Srinivas et al. 2012) and the *U-function* (Echard, Gayton, &

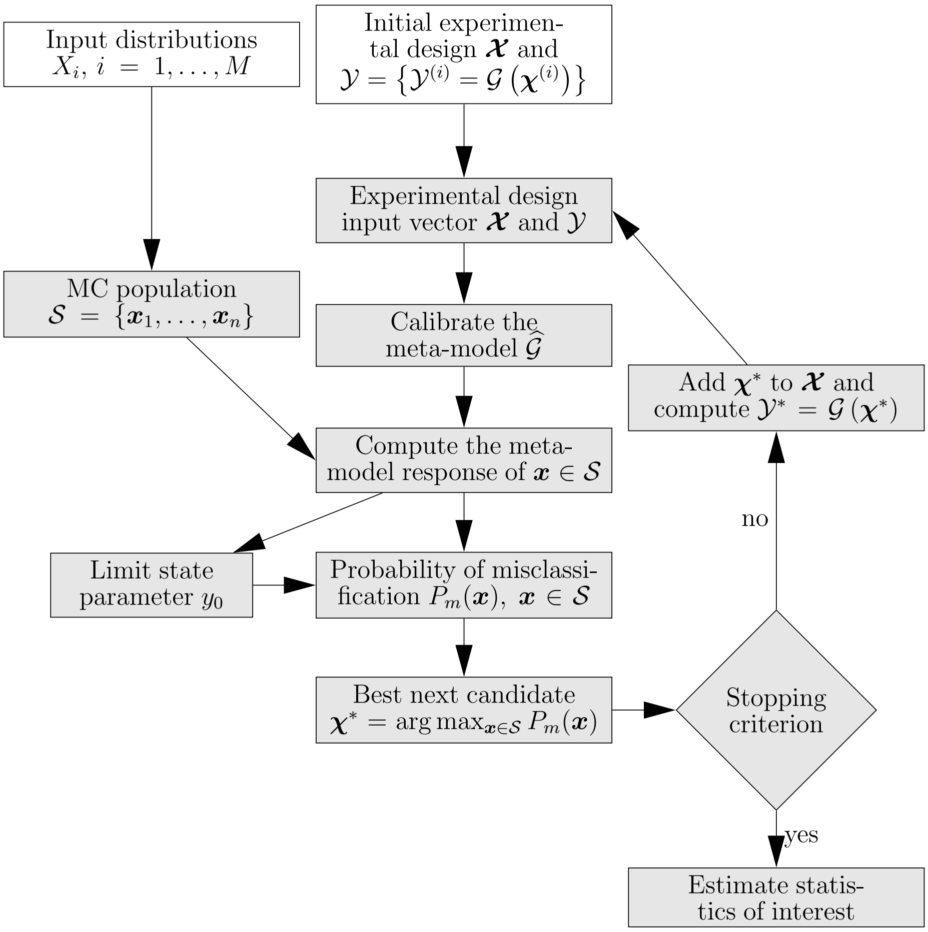


Figure 5.1: Flowchart of the adaptive algorithm for rare event estimation.

Lemaire 2011). For further details regarding the learning function, the reader is referred to Appendix E. In the following, the U-function is discussed more in depth.

The U-function is based on the concept of misclassification and the very nature of the Gaussian process (GP) meta-model. Since the GP prediction $\widehat{Y}(\boldsymbol{x})$ is a Gaussian variable for each sample point $\boldsymbol{x} \in \mathcal{S}$, there is a non-zero probability that the prediction mean $\mu_{\widehat{Y}}(\boldsymbol{x}) > y_0$ (safety domain) whereas the true value satisfies $\mathcal{G}(\boldsymbol{x}) \leq y_0$ (failure domain) or vice versa. Due to the fact that the prediction $\widehat{Y}(\boldsymbol{x})$ is a Gaussian random variable described by the first two moments $\left\{\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma_{\widehat{Y}}(\boldsymbol{x})\right\}$, the probability of misclassification P_m can be written as (Bect et al. 2012):

$$P_m(\boldsymbol{x}) = \min\left[\Phi\left(\frac{\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \Phi\left(\frac{y_0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)\right] = \Phi\left(-\frac{|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right). \tag{5.10}$$

The probability of misclassification is maximized when the fraction tends to zero, *i.e.* $-|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|/\sigma_{\widehat{Y}}(\boldsymbol{x}) \sim 0 \rightarrow P_m(\boldsymbol{x}) = 0.5$, and is small when either the prediction mean is far away from the limit-state parameter y_0 and/or the prediction variance is small. Note that the probability measure used to define misclassification is the one associated to the Gaussian nature of the predictor $\mathcal{N}\left(\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma_{\widehat{Y}}(\boldsymbol{x})\right)$ in each point $\boldsymbol{x}$. It shall not be confused with the probability $\mathbb{P}(\cdot)$ in Eq. (5.3) which corresponds to the input random vector $\boldsymbol{X}$.

In this formalism, the U-function is defined as the “reliability index” corresponding to the probability of misclassification (Echard, Gayton, & Lemaire 2011):

$$U(\boldsymbol{x}) = \frac{|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}. \tag{5.11}$$

The optimal sample to enrich the experimental design is the one minimizing the U-function among $\boldsymbol{x} \in \mathcal{S}$, thus maximizing the probability of misclassification:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathcal{S}} U(\boldsymbol{x}) \equiv \arg\max_{\boldsymbol{x}\in\mathcal{S}} P_m(\boldsymbol{x}). \tag{5.12}$$

Due to this similarity, both the probability of misclassification and the U-function can be used as learning functions.

Multiple sample selection

In cases where parallel computing is available, it can be beneficial to add *multiple points* in each iteration to speed up the overall computation even if the procedure may be slightly suboptimal. Assuming that K samples can be determined at the beginning of each iteration, the corresponding model responses could be computed simultaneously on K independent CPUs.

There are several ways to sample multiple points from the set of candidates $\mathcal{S}$. Generally speaking though, attractive candidates are found close to the limit-state surface. Thus, a lower and upper boundary of the limit-state surface is defined taking into account the prediction uncertainty in the Kriging model. Then, two bounds can be defined as:

$$\mu_{\widehat{Y}}(\boldsymbol{x}) - k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) = y_0, \tag{5.13}$$

$$\mu_{\widehat{Y}}(\boldsymbol{x}) + k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) = y_0, \tag{5.14}$$

where k sets the confidence level typically equal to $1.96 = \Phi^{-1}(97.5\%)$. In other words, the lower bound estimates the limit-state surface assuming that the real value of every sample $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}}$ is $\mu_{\widehat{Y}}(\boldsymbol{x}) - k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x})$ instead of the mean value $\mu_{\widehat{Y}}(\boldsymbol{x})$.

Analogously, we define the "mean" failure domain in terms of the prediction mean value:

$$\mathcal{D}_f^0 \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} :\ \mu_{\widehat{Y}}(\boldsymbol{x}) \leq y_0\right\}, \tag{5.15}$$

and the corresponding lower and upper bounds to the failure domain:

$$\mathcal{D}_f^- \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} :\ \mu_{\widehat{Y}}(\boldsymbol{x}) + k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) \leq y_0\right\}, \tag{5.16}$$

$$\mathcal{D}_f^+ \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} :\ \mu_{\widehat{Y}}(\boldsymbol{x}) - k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) \leq y_0\right\}, \tag{5.17}$$

so that $\mathcal{D}_f^- \subset \mathcal{D}_f^0 \subset \mathcal{D}_f^+$ are nested failure domains. The lower and upper bounds can be interpreted as the least and most conservative estimate of the failure domain, respectively.

The *limit-state margin* $\mathbb{M}_f$ is defined as the difference between the lower and the upper boundaries of the failure domain (Dubourg 2011):

$$\mathbb{M}_f \stackrel{\text{def}}{=} \mathcal{D}_f^+ \backslash \mathcal{D}_f^-. \tag{5.18}$$

The limit-state margin is a natural region where to focus for candidate samples for the design enrichment. Considering a large set of samples of the input vector $\mathcal{S} = \left\{\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n)}\right\}$, we define the following sets corresponding to the domains in Eqs. (5.15) to (5.18):

$$\mathcal{S}_f^0 \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \in \mathcal{S} :\ \mu_{\widehat{Y}}(\boldsymbol{x}) \leq y_0\right\}, \tag{5.19}$$

$$\mathcal{S}_f^+ \overset{\text{def}}{=} \left\{ \boldsymbol{x} \in \mathcal{S} : \; \mu_{\widehat{Y}}(\boldsymbol{x}) - k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) \leq y_0 \right\}, \tag{5.20}$$

$$\mathcal{S}_f^- \overset{\text{def}}{=} \left\{ \boldsymbol{x} \in \mathcal{S} : \; \mu_{\widehat{Y}}(\boldsymbol{x}) + k \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) \leq y_0 \right\}, \tag{5.21}$$

$$\mathcal{S}_{\mathbb{M}} \overset{\text{def}}{=} \mathcal{S}_f^+ \backslash \mathcal{S}_f^-, \tag{5.22}$$

where $\mathcal{S}_f^- \subset \mathcal{S}_f^0 \subset \mathcal{S}_f^+$ are nested sets. These sets are the discrete counterparts corresponding to the domains defined in Eqs. (5.15) to (5.18). The margin set $\mathcal{S}_{\mathbb{M}}$ contains points of interest for enriching the experimental design since they lie close to the true limit-state surface. Virtually any point in $\mathcal{S}_{\mathbb{M}}$ could be added to $\boldsymbol{\mathcal{X}}$ and a simple approach would be to sample the K different points randomly. However, a better coverage of $\mathbb{M}_f$ can be obtained using clustering techniques.

To account for the relative importance of the samples in $\mathcal{S}_{\mathbb{M}}$, a weighted K-means clustering algorithm is used (Zaki & Meira 2014). In this algorithm, samples with high information value according to the learning function have larger weights, whereas samples with low information value have low weights. The weights are set equal to the probability of misclassification (see Eq. (5.10)) of each sample $\boldsymbol{x} \in \mathcal{S}_{\mathbb{M}}$. By definition of the probability of misclassification, the weights are bounded on $[0, 0.5]$. The additional K samples $\boldsymbol{\chi}^*$ are then determined as the samples $\boldsymbol{x} \in \mathcal{S}_{\mathbb{M}}$ which are closest to the nuclei of each i-th cluster.

To showcase the process of selecting additional samples, let us consider the following limit-state function with two-dimensional input vector:

$$y = \mathcal{G}(\boldsymbol{x}) = 20 - (x_1 - x_2)^2 - 8 \cdot (x_1 + x_2 - 4)^3, \tag{5.23}$$

where the input variables have uniform distributions $X_i \sim \mathcal{U}(-5, 5), \; i = 1, 2$.

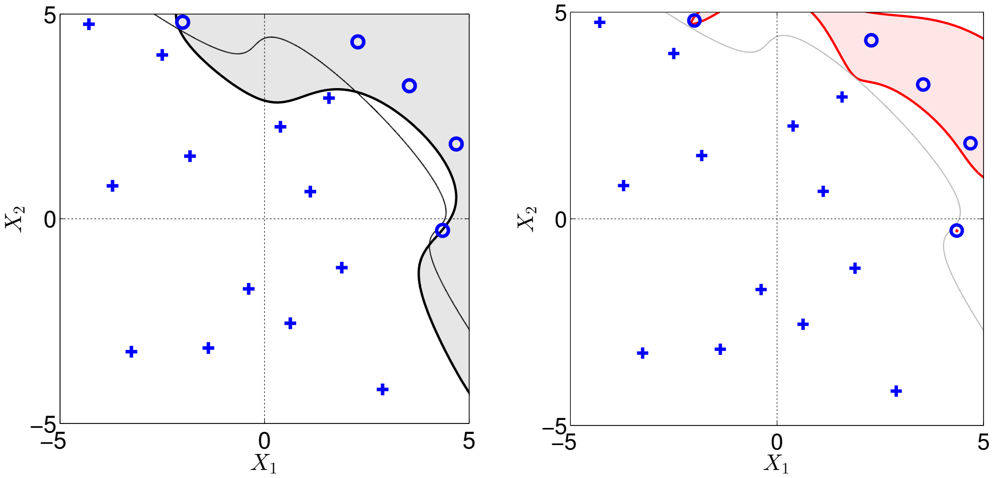
The function is meta-modelled with an ordinary Kriging model. Figure 5.2(a) displays the initial experimental design ("$\circ$" for failure samples and "+" for safety samples), the exact failure domain (thin line) and the failure region estimated from the prediction mean values $\mu_{\widehat{Y}}(\boldsymbol{x})$ (thick black line). Figures 5.2(b)-5.2(d) show the lower and upper failure domains as well as the domain of the limit-state margin, respectively. Finally, Figure 5.2(e) shows the probability of misclassification in the limit-state margin and Figure 5.2(f) shows the selected candidates. The grey area in Figure 5.2(f) represents the samples in the limit-state margin, the large black diamond represents the optimal sample (with highest probability of misclassification) and the hollow black diamonds represent the $K = 5$ samples obtained by weighted K-means clustering.

It can be observed that the $K = 5$ samples in Figure 5.2(f) cover the domain of the limit-state margin well and are close to the limit-state surface determined by the prediction mean value. This leads to an efficient selection of additional samples for enriching the experimental design. Interestingly, however, the optimal sample in terms of the probability of misclassification is not part of the $K = 5$ sample set. This implies that each of the $K = 5$ samples is suboptimal in terms of misclassification.

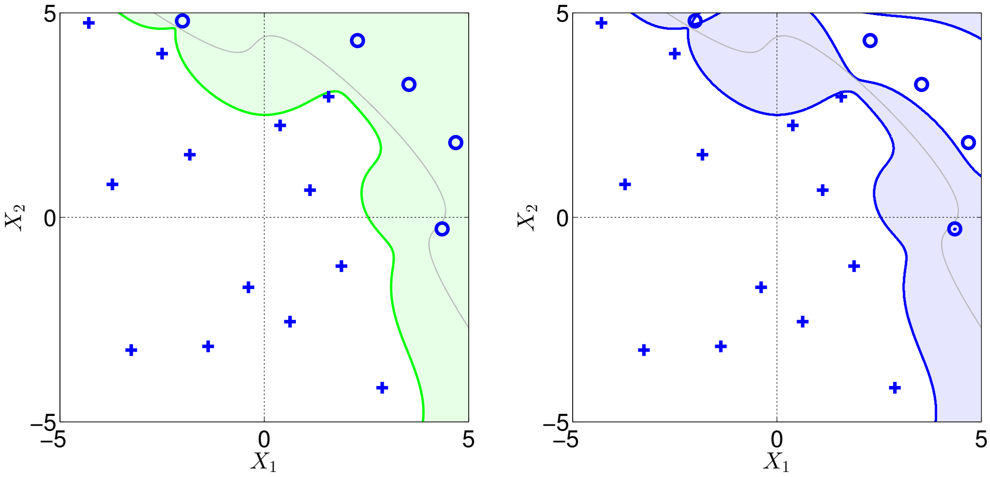
The optimal sample (filled diamond marker in Figure 5.2(f)) depends highly on the sampling of the limit-state margin. In principle any sample with $\mu_{\widehat{Y}}(\boldsymbol{x}) = y_0 = 0$ could be optimal in terms of the probability of misclassification. This implies that the choice of a single sample with respect to the maximal probability of misclassification is non-unique in theory. However, it is unique and well-defined in the discrete optimization problem Eq. (5.12).

5.3.4 Stopping criterion

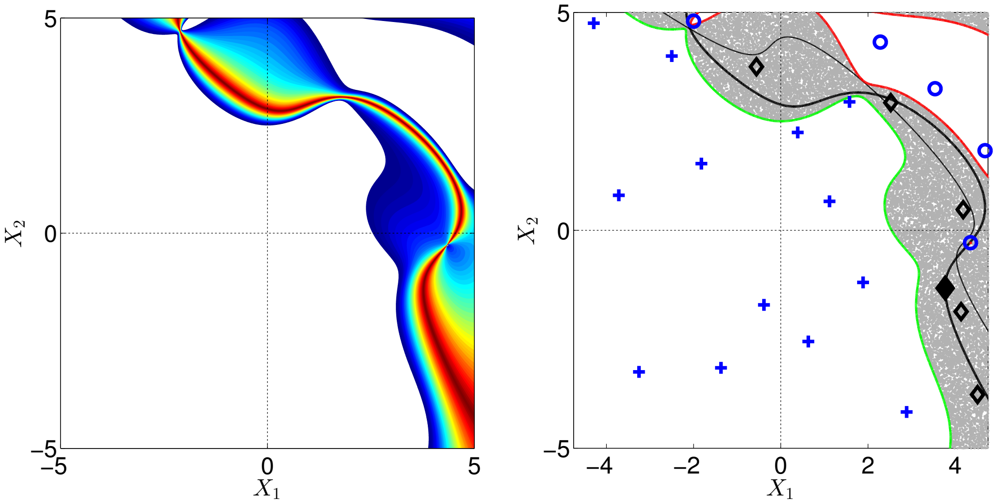
In Echard, Gayton, & Lemaire (2011), the convergence measure (a.k.a. stopping criterion) is based on the accuracy of the meta-model around the limit-state surface rather than on the



(a) Mean $\mathcal{D}_f^0$ (b) Lower bound $\mathcal{D}_f^-$



(c) Upper bound $\mathcal{D}_f^+$ (d) Limit-state margin $\mathbb{M}_f$



(e) Probabiliy of misclassification P_m in $\mathbb{M}_f$ (f) Additional samples

Figure 5.2: Design enrichment when adding $K = 5$ samples – definition of failure domains and illustration of the candidate selection algorithm.

estimation of the statistics of interest. Thus, two stopping criteria are proposed for the estimation of failure probabilities and quantiles designed to maximize the accuracy of the statistics of interest while minimizing computational costs.

The stability of the estimate of the statistics of interest can be measured by the size of the limit-state margin $\mathbb{M}_f$ and consequently by the associated values of the upper and lower boundaries of the limit-state surface. When the boundaries are close to each other (*i.e.* a small limit-state margin), then, the estimate of the statistics of interest is accurate. Therefore, the stopping criterion for estimating failure probabilities is defined as:

$$\frac{\widehat{P}_f^+ - \widehat{P}_f^-}{\widehat{P}_f^0} \leq \epsilon_{\widehat{P}_f}, \tag{5.24}$$

for two consecutive iteration steps where $\epsilon_{\widehat{P}_f} = 5\%$ is used in applications. The upper and lower bound failure probabilities are defined as:

$$\widehat{P}_f^+ \stackrel{\text{def}}{=} \mathbb{P}\left(\mu_{\widehat{Y}}(\boldsymbol{X}) - k\,\sigma_{\widehat{Y}}(\boldsymbol{X}) \leq y_0\right), \tag{5.25}$$

$$\widehat{P}_f^- \stackrel{\text{def}}{=} \mathbb{P}\left(\mu_{\widehat{Y}}(\boldsymbol{X}) + k\,\sigma_{\widehat{Y}}(\boldsymbol{X}) \leq y_0\right), \tag{5.26}$$

and the best estimate for the failure probability reads:

$$\widehat{P}_f^0 \stackrel{\text{def}}{=} \mathbb{P}\left(\mu_{\widehat{Y}}(\boldsymbol{X}) \leq y_0\right). \tag{5.27}$$

These failure probabilities are estimated based on $\mathcal{S}$. As an example, $\widehat{P}_f^+$ is estimated by an equation analogue to Eq. (5.5):

$$\widehat{P}_f^+ \approx \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\left(\mu_{\widehat{Y}}(\boldsymbol{x}) - k\sigma_{\widehat{Y}}(\boldsymbol{x}) \leq y_0\right)}\left(\boldsymbol{x}^{(i)}\right). \tag{5.28}$$

The stopping criterion defined in Eq. (5.24) can be adapted to the context of quantile estimation as follows:

$$\frac{\widehat{q}_\alpha^+ - \widehat{q}_\alpha^-}{\operatorname{Std}\left[\widehat{Y}\right]} \leq \epsilon_{\widehat{q}_\alpha}, \tag{5.29}$$

for two consecutive iteration steps where $\epsilon_{\widehat{q}_\alpha} = 5\%$ is used in practical applications. The quantiles $\widehat{q}_\alpha^\pm$ are computed from a large Monte Carlo sampling of the bounds $\mu_{\widehat{Y}}(\boldsymbol{x}) \pm k\,\sigma_{\widehat{Y}}(\boldsymbol{x})$, where $\operatorname{Std}\left[\widehat{Y}\right]$ is the empirical standard deviation of the sample $\left\{\mu_{\widehat{Y}}\left(\boldsymbol{x}^{(i)}\right),\, i = 1, \ldots, n\right\}$. The normalization in Eq. (5.29) makes use of the standard deviation rather than $\widehat{q}_\alpha^0$ in analogy with Eq. (5.24). Indeed, unlike $\widehat{P}_f^0$ in Eq. (5.24), $\widehat{q}_\alpha^0$ can be any real number in $\mathbb{R}$, which makes it inappropriate for normalization.

5.4 Numerical examples

5.4.1 Four-branch function

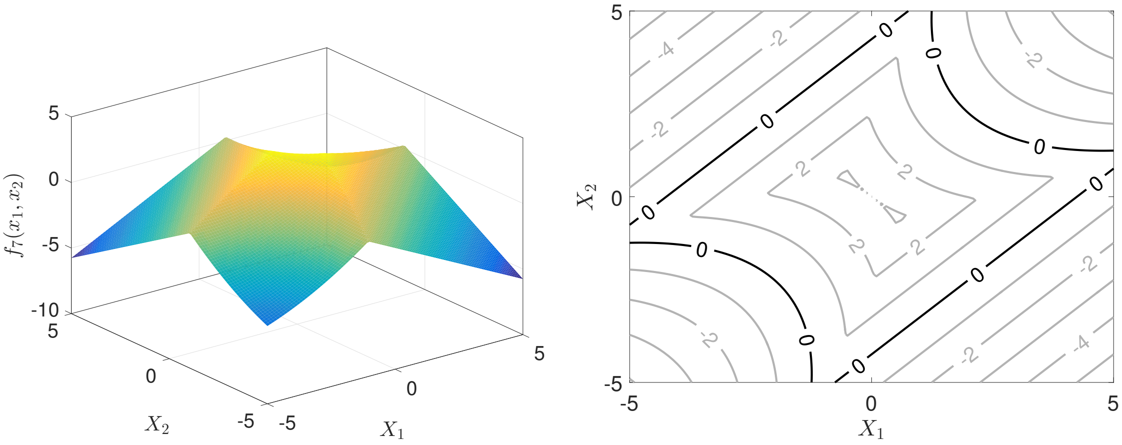
Problem statement

The four-branch function is a common benchmark in structural reliability analysis (SRA) that describes the failure of a series system with four distinct component limit states. Its mathematical formulation reads (Waarts 2000; Schueremans & Van Gemert 2005a,b):

$$f_7(\boldsymbol{x}) = \min \left\{ \begin{array}{c} 3 + 0.1\,(x_1 - x_2)^2 - \frac{x_1+x_2}{\sqrt{2}} \\ 3 + 0.1\,(x_1 - x_2)^2 + \frac{x_1+x_2}{\sqrt{2}} \\ (x_1 - x_2) + \frac{6}{\sqrt{2}} \\ (x_2 - x_1) + \frac{6}{\sqrt{2}} \end{array} \right\}, \tag{5.30}$$

where the input variables are modelled by two independent Gaussian random variables $X_i \sim \mathcal{N}(0, 1)$. The failure event is defined as $f_7(\boldsymbol{x}) \leq 0$, *i.e.* the failure probability is $P_f = \mathbb{P}\left(f_7(\boldsymbol{X}) \leq 0\right)$. Hence, in this example the limit-state function $\mathcal{G}\left(\boldsymbol{x}\right)$ is equivalent to the computational model $\mathcal{M}(\boldsymbol{x}) = f_7(\boldsymbol{x})$.

Figure 5.3 illustrates the response surface of the four-branch function and the contour lines including a highlighted limit-state surface. The response surface nicely shows the four parts of the limit-state function. Moreover, the failure domain has a complex shape accounting for the four elements of the series system.



(a) Response surface (b) Contour lines

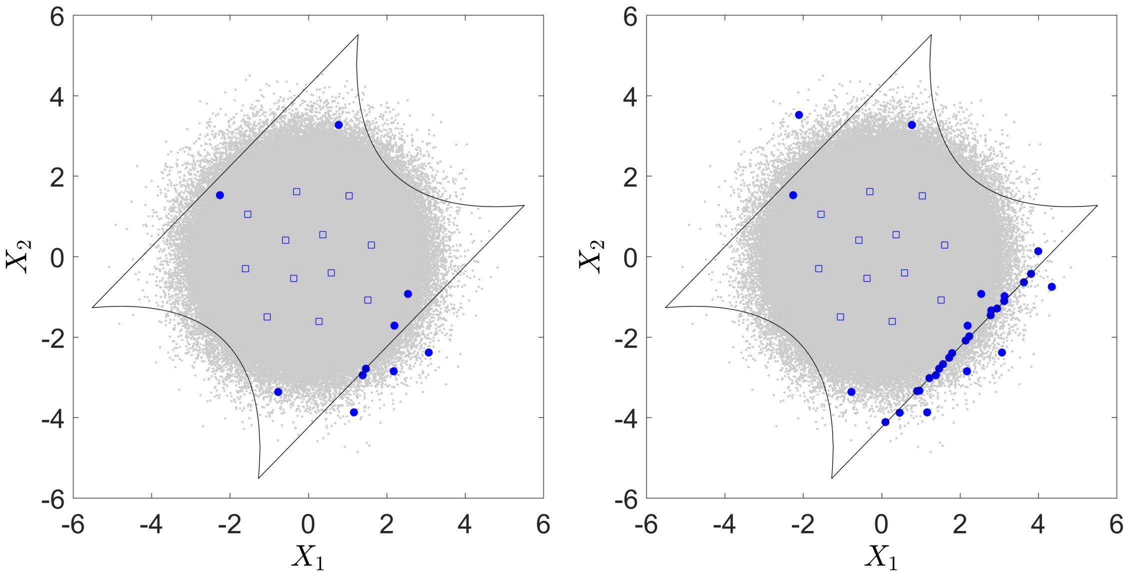
Figure 5.3: Four-branch function – exact model response.

Failure probability estimation

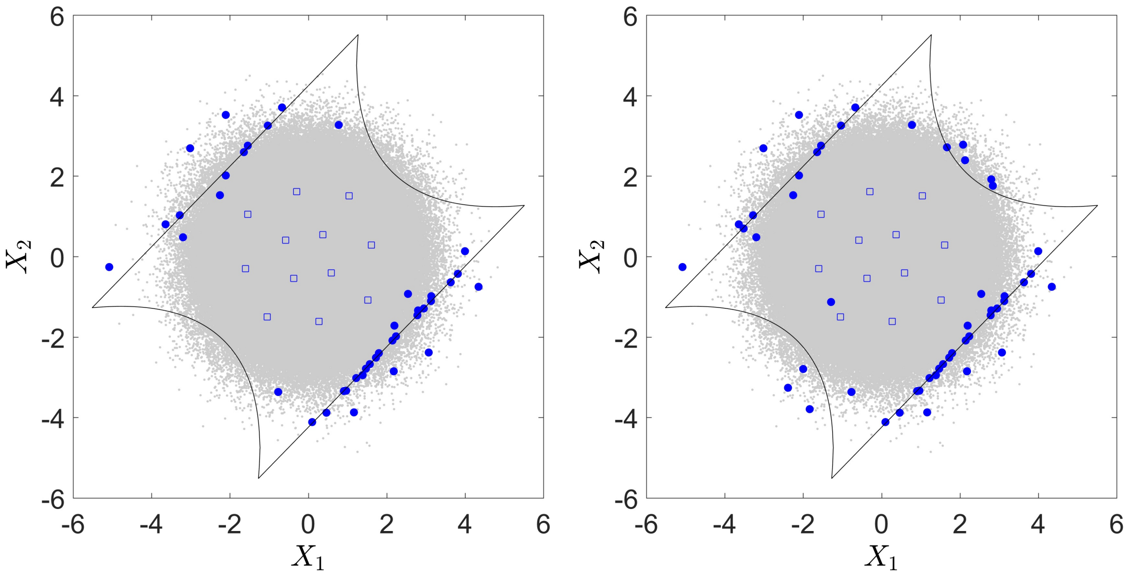
Settings The adaptive experimental design algorithm is initiated with $N_0 = 12$ Latin-hypercube samples (LHS) and a candidate Monte Carlo population $\mathcal{S}$ of $n = 10^6$ samples. The performances of PC-Kriging (PCK) are compared to those of ordinary Kriging (OK). The trend in PC-Kriging consists of Hermite polynomials of maximal degree of eight and a Gaussian autocorrelation function is chosen for the Gaussian process. Both meta-modelling techniques are calibrated using the MATLAB-based toolbox UQLAB (Marelli & Sudret 2014). The adaptive experimental design algorithm is tested in both single and multiple ($K = 6$) samples mode.

Visualization Figures 5.4 and 5.5 visualize several iterations of the algorithm for OK and PCK, respectively, for single sample enrichment. The grey dots represent the candidate MC population $\mathcal{S}$. The empty squares mark the initial experimental design, whereas the blue filled circles mark the additional samples. The solid black line represents the exact limit-state surface.

Both variants discover the four distinct failure modes in the limit-state function within the first 50 iterations. The main difference lies in the limit-state exploration pattern. When using

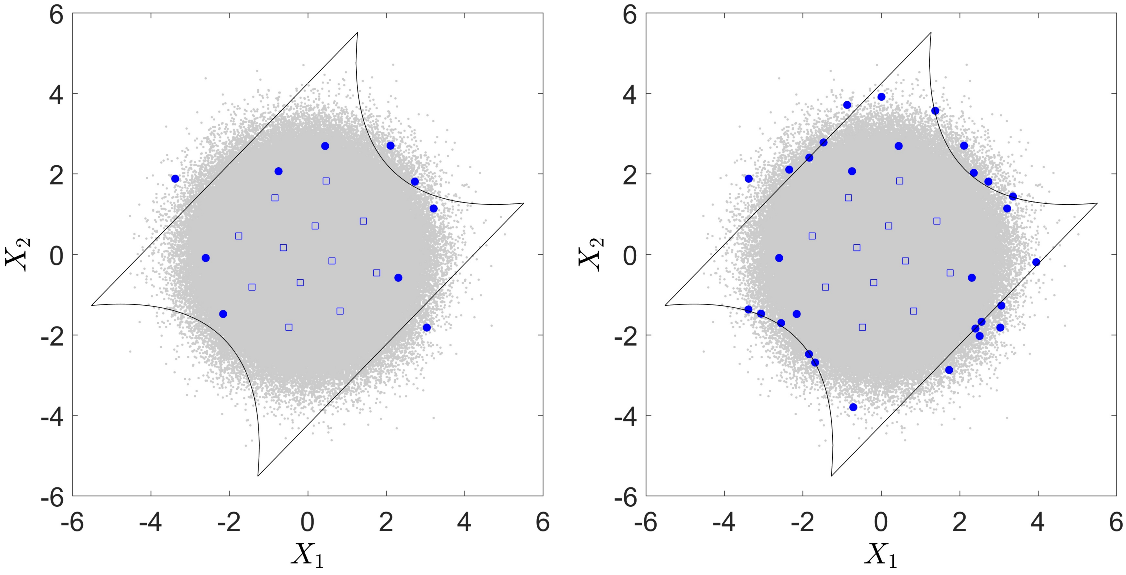


(a) Iteration 10 (b) Iteration 30

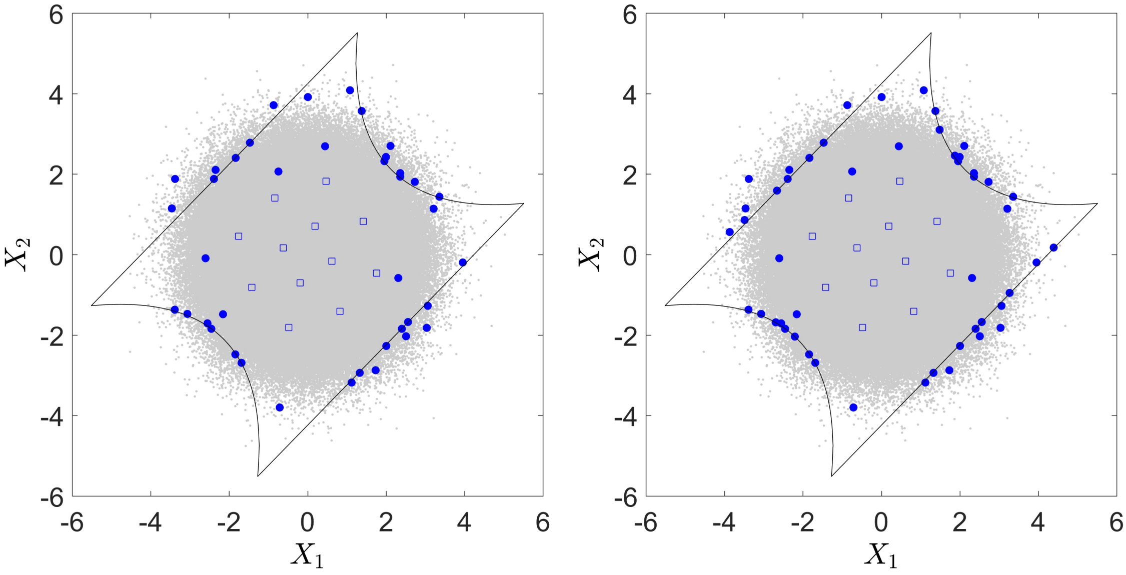


(c) Iteration 40 (d) Iteration 50

Figure 5.4: Four-branch function – iterations of the adaptive experimental design algorithm with ordinary Kriging meta-models ($K = 1$).



(a) Iteration 10 (b) Iteration 30

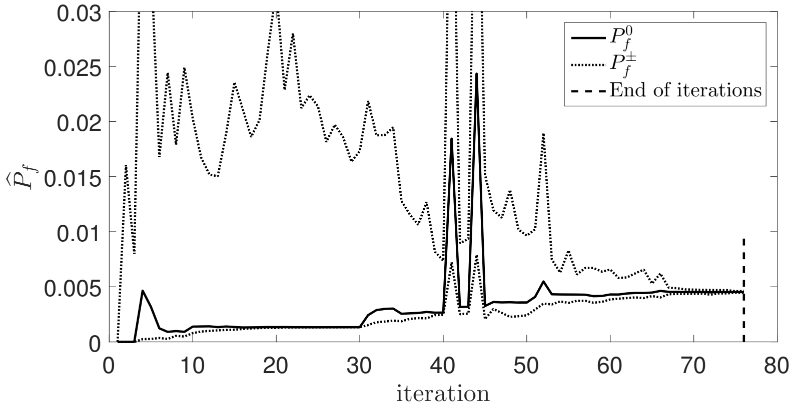


(c) Iteration 40 (d) Iteration 50

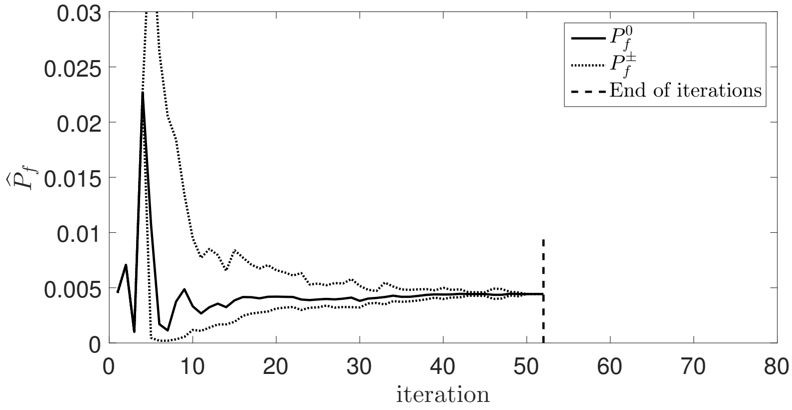
Figure 5.5: Four-branch function – iterations of the adaptive experimental design algorithm with PC-Kriging meta-models ($K = 1$).

ordinary Kriging in Figure 5.4, the distinct failure mechanisms are explored one-by-one. As soon as a failure mechanism is characterized, the adaptive sampling algorithm moves to the next. At iteration 50 all four failure mechanisms are discovered. In the case of PC-Kriging, all four failure mechanisms are already discovered after 10 iterations (Figure 5.5(a)). Thus, PC-Kriging converges on average faster to an accurate estimate of the failure probability as compared to ordinary Kriging. This is to be expected due to the more accurate global trend in PC-Kriging.

Figure 5.6 illustrates the convergence of the estimate of the failure probability by showing $\widehat{P}_f^{\pm}$ ($k = 2$) and $\widehat{P}_f^0$. Further, the final iteration is marked where the iterative algorithm is stopped due to the stopping criterion in Eq. (5.24). Figure 5.6 confirms the faster convergence of the PC-Kriging meta-models. Note that for the Kriging meta-model, Figure 5.6 shows four plateaus in $\widehat{P}_f^0$ which correspond to the exploration of the four branches of the limit-state function.



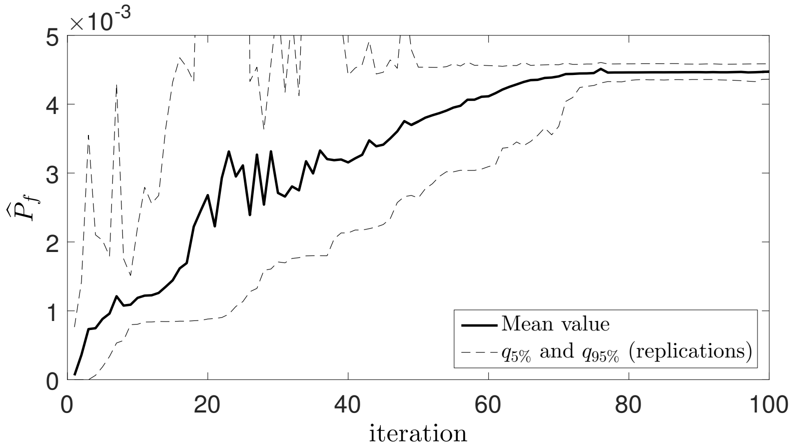
(a) Ordinary Kriging



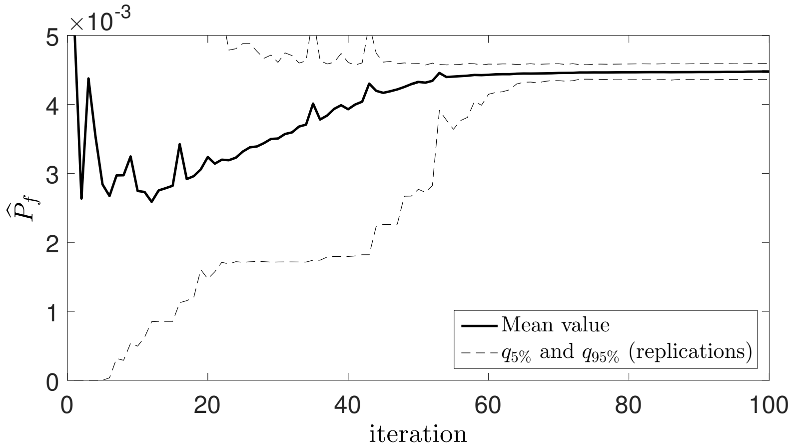
(b) PC-Kriging

Figure 5.6: Four-branch function – convergence curves of the adaptive experimental design algorithm for the single run in Figures 5.4 and 5.5.

In order to test the statistical significance of the convergence, the analysis is replicated 50 times with different initial experimental designs. Figure 5.7 illustrates the convergence of the failure probability estimates $\widehat{P}_f^0$ based on ordinary Kriging and PC-Kriging up to iteration 100, respectively. The solid line represents the mean value of 50 independent replications of the same analysis with different initial Latin-hypercube experimental designs. The dashed lines represent the 5% and 95% quantiles of the 50 independent replications, *i.e.* the 90% confidence interval (CI). The results show that PC-Kriging converges faster than ordinary Kriging. Note that the variation in $\widehat{P}_f$ after convergence (iterations ≥ 80) originates from the finite size of $\mathcal{S}$ and the corresponding variance in Monte Carlo simulations (see also Eq. (5.6)). Hence, the remaining variance is identical for ordinary Kriging and PC-Kriging.



(a) Ordinary Kriging



(b) PC-Kriging

Figure 5.7: Four-branch function – convergence curves of the adaptive experimental design algorithms (statistics of 50 replications of the analysis with different initial Latin-hypercube experimental design).

Parametric study In this paragraph, the candidate selection algorithm and the choice of stopping criterion are discussed. The candidate selection algorithm compares the selection of the single optimal candidate to the selection of multiple candidates through weighted K-means clustering. In terms of stopping criterion, Eq. (5.24) ($k = 2$) is compared to the existing stopping criterion defined in Echard, Gayton, & Lemaire (2011):

$$\min[U(\boldsymbol{x})] \geq 2 \quad \forall \boldsymbol{x} \in \mathcal{S}, \tag{5.31}$$

which indicates that the probability of misclassification must be smaller than $\Phi(-2) \approx 2$ % for all candidate samples in order to stop the iterations.

The full failure probability analysis is replicated 50 times with different initial LHS experimental design in order to assess the statistical significance (*e.g.* standard deviation of the resulting $\widehat{P}_f$). The analysis is focused on the total number of computational model evaluations and on the accuracy of the estimate of the failure probability. The total number of computational model runs is $N_{tot} = N_0 + N_1$, where N_1 is the number of samples added by the adaptive algorithm.

Table 5.1: Four-branch function – results using different meta-models, candidate selection algorithms and stopping criteria. Coefficients of variation (CoV) are computed based on the 50 replications of the analysis with different initial experimental designs.

Method	Enrich.	Stop. crit.	$\mathbb{E}\left[\widehat{P}_f\right]$	$\mathrm{CoV}\left[\widehat{P}_f\right]$	$\mathbb{E}\left[N_{tot}\right]$	$\mathrm{CoV}\left[N_1\right]$
MCS			$4.460 \cdot 10^{-3}$	0.15 %	10^8	-
OK	single	Eq. (5.31)	$4.464 \cdot 10^{-3}$	1.4 %	$12 + 96.4 = 108.4$	6.7 %
PCK	single	Eq. (5.31)	$4.471 \cdot 10^{-3}$	1.4 %	$12 + 115.8 = 127.8$	24.5 %
OK (*)	single	Eq. (5.31)	$4.416 \cdot 10^{-3}$	-	126	-
OK	single	Eq. (5.24)	$4.440 \cdot 10^{-3}$	1.6 %	$12 + 66.3 = 78.3$	11.0 %
PCK	single	Eq. (5.24)	$4.457 \cdot 10^{-3}$	1.5 %	$12 + 61.2 = 73.2$	28.5 %
PCK	$K = 6$	Eq. (5.24)	$4.458 \cdot 10^{-3}$	1.5 %	$12 + 14.4 \cdot 6 = 98.4$	8.9 %

(*) single run results from Echard, Gayton, & Lemaire (2011)

Table 5.1 summarizes the results for the four-branch function. The results of the adaptive experimental design algorithm are presented for both stopping criteria (Eqs. (5.24) and (5.31)) and meta-models (ordinary Kriging and PC-Kriging). The reference value for the failure probability is computed by Monte Carlo simulation with 10^8 samples.

Considering the cases of adding a single candidate, all configurations estimate accurately the failure probability. The coefficient of variation of the estimate of the failure probability $\mathrm{CoV}\left[\widehat{P}_f\right]$ is comparable with the theoretical coefficient of variation of a Monte Carlo sampling (see Eq. (5.6)) with $n_{\mathrm{MCS}} = 10^6$ samples, *i.e.* $\mathrm{CoV}\left[\widehat{P}_f = 4.46 \cdot 10^{-3}, n_{\mathrm{MCS}} = 10^6\right] \approx 1.5$ %.

The results in Table 5.1 show that the stopping criterion in Eq. (5.24) leads to accurate results despite the smaller number of computational model runs N_{tot} when compared to the stopping criterion in Eq. (5.31). This indicates that Eq. (5.24) is more suitable to the estimation of failure probabilities than Eq. (5.31). It is interesting though that PC-Kriging converges faster than ordinary Kriging with the stopping criterion in Eq. (5.24), as opposed to the one in Eq. (5.31). The explanation lies in the fact that the stopping criterion in Eq. (5.24) measures the convergence of the statistics of interest, whereas Eq. (5.31) assesses the overall accuracy of the approximation of the Kriging surrogate.

For the case of PCK and the stopping criterion in Eq. (5.24), Table 5.1 includes also the results obtained by adding $K = 6$ candidates at each iteration. The comparison of the total number of computational model runs N_{tot} to the case of adding a single point in each iteration shows that the addition of multiple points at each iteration requires a larger N_{tot}: indeed in the case of single point enrichment, the total number of model runs is $N_{tot} = 73$, whereas adding $K = 6$ points in each step leads to $N_{tot} = 98$. Considering however that the runs of the computational model can be performed in parallel ($K = 6$ on six independent CPUs), the total physical time for the K-point enrichment algorithm is approximately four times smaller.

Quantile estimation

Consider now the quantile estimation problem related to the four-branch function. The quantiles are estimated for $\alpha = \{0.01, 0.001, 0.0001\}$. PC-Kriging is used, a single sample is added at each iteration out of a candidate Monte Carlo population of $n = 10^6$ samples, and the algorithm is stopped with the stopping criterion in Eq. (5.29).

Table 5.2 summarizes the results and compares the estimate of the quantiles $\widehat{q}_\alpha$ with a reference solution from a Monte Carlo simulation ($n_{\mathrm{MCS}} = 10^8$). The three α values are accurately estimated with the adaptive algorithm despite the small initial experimental design of $N_0 = 12$ samples. The coefficient of variation of the results increases with decreasing α due to the decreasing size of the failure domain. The results obtained by PC-Kriging are more efficient due to the relatively low number of model evaluations compared to the ordinary Kriging results.

Table 5.2: Four-branch function – quantile estimation results. MCS is based $n_{\mathrm{MCS}} = 10^8$ samples. Coefficients of variation (CoV) are computed based on 50 replications of the analysis with different initial experimental designs.

α	Method	$\mathbb{E}[\widehat{q}_\alpha]$	CoV $[\widehat{q}_\alpha]$	$\mathbb{E}[N_{tot}]$	CoV $[N_1]$
0.01	MCS	0.303			
	OK	0.303	1.7 %	$12 + 139.6 = 151.6$	32.5 %
	PCK	0.305	1.4 %	$12 + 62.4 = 74.4$	21.0 %
0.001	MCS	−0.528			
	OK	−0.530	2.2 %	$12 + 138.9 = 150.9$	34.8 %
	PCK	−0.531	2.1 %	$12 + 57.7 = 69.7$	23.3 %
0.0001	MCS	−1.299			
	OK	−1.280	5.0%	$12 + 165 = 177$	29.4 %
	PCK	−1.292	2.4 %	$12 + 47.9 = 59.9$	13.6 %

It is interesting to notice that in the case of PC-Kriging the smaller the α value is, the fewer samples are required to estimate the quantile. This phenomenon can be explained by considering the failure domain defined by Eq. (5.30). For small values of α, indeed the failure domain is also small. Hence to meta-model the limit-state surface, a smaller number of experimental design points is required.

5.4.2 Borehole model

Problem statement

The first realistic engineering problem considered is the so-called *borehole-function*, which describes the water flow through a borehole. This benchmark function has been discussed in papers such as Harper & Gupta (1983), Morris (1993), An & Owen (2001), and Kersaudy et al. (2015). It is a fast-to-evaluate function depending on an eight-dimensional input vector $\boldsymbol{x} = (r_w, r, T_u, H_u, T_l, H_l, L, K_w)^\mathsf{T}$ (Harper & Gupta 1983):

$$v(\boldsymbol{x}) = \frac{2\pi T_u (H_u - H_l)}{\ln(r/r_w)\left(1 + \frac{2LT_u}{\ln(r/r_w) r_w^2 K_w} + \frac{T_u}{T_l}\right)}, \tag{5.32}$$

where $v(\boldsymbol{x})$ is the fluid water flow measured in m^3/year, r_w is the radius of the borehole, r the radius of influence, T_u the transmissivity of the upper aquifer, H_u the potentiometric head of the upper aquifer, T_l the transmissivity of the lower aquifer, H_l the potentiometric head of the lower aquifer, L the length of the borehole, and K_w the hydraulic conductivity of the soil.

The uncertainties in the input vector are modelled as independent random variables whose properties are summarized in Table 5.3. For the lognormal distribution, the parameters are the mean and standard deviation of the natural logarithm of the variable. For the other variables, they describe the range of uniform distributions.

Table 5.3: Borehole model – definition of the probabilistic model of the input variables. For the uniform distributions the parameters denote the range, whereas for the lognormal distribution, the parameters denote the mean and standard deviation of the natural logarithm of the variable.

Variable	Units	Distribution	Parameters
r_w	[m]	Uniform	$[0.05, 0.15]$
r	[m]	Lognormal	$[7.71, 1.0056]$
T_u	[m^2/year]	Uniform	$[63070, 115600]$
H_u	[m]	Uniform	$[990, 1110]$
T_l	[m^2/year]	Uniform	$[63.1, 116]$
H_l	[m]	Uniform	$[700, 820]$
L	[m]	Uniform	$[1120, 1680]$
K_w	[m/year]	Uniform	$[9855, 12045]$

Quantile estimation

The statistics of interest are the quantiles of the water flow $v(\boldsymbol{x})$ corresponding to $\alpha = \{0.99, 0.999, 0.9999\}$. An initial experimental design of $N_0 = 12$ Latin-hypercube samples in used. The number of samples in the candidate Monte Carlo population is set to $n = 10^6$ and the remaining settings are the same as in the four-branch function application example.

The results of the adaptive algorithm and the reference values are reported in Table 5.4 for the various values of α, the meta-modelling method (PCK and OK) and enrichment strategies (single or multiple) used.

The prediction of the quantiles is accurate for both meta-modelling techniques (OK and PCK) compared to the Monte Carlo ($n_{\text{MCS}} = 10^6$) solution when adding a single sample and multiple samples to the experimental design in each iteration. There is, however, a noticeable difference in the total number of evaluations of the computational model. Ordinary Kriging (OK) requires more model evaluations than PC-Kriging (PCK) for all values of α. The addition of $K = 6$ samples at a time induces slightly more model evaluations N_{tot} but at the same time reduces the total computational time due to the parallel evaluation of $K = 6$ samples.

5.4.3 Two-dimensional truss structure

Problem statement

For the purpose of illustrating a realistic structural engineering application including FEM, a two-dimensional truss structure is analysed. This specific truss example structure has been presented and discussed previously in Lee & Kwak (2006), Sudret, Blatman, & Berveiller (2007), and Blatman & Sudret (2008, 2010a).

Consider the simply supported two-dimensional truss structure sketched in Figure 5.8, which consists of 23 bars and 13 nodes. The geometry is known deterministically, whereas the material properties and the loadings are modelled stochastically. Ten stochastic input variables form the input vector $\boldsymbol{X}$:

$$\boldsymbol{X} = (E_1, E_2, A_1, A_2, P_1, P_2, P_3, P_4, P_5, P_6)^{\mathsf{T}}, \tag{5.33}$$

Table 5.4: Borehole model – results of quantile estimation based on a single run of the adaptive experimental design algorithm with stopping criterion defined in Eq. (5.29).

α	Method	Enrichment	$\widehat{q}_\alpha$	N_{tot}
0.99	MCS	-	157.5 m^3/year	10^6
	OK	single	157.3 m^3/year	$12 + 199 = 211$
	PCK	single	157.4 m^3/year	$12 + 34 = 46$
	OK	$K = 6$	157.5 m^3/year	$12 + 41 \cdot 6 = 258$
	PCK	$K = 6$	157.7 m^3/year	$12 + 10 \cdot 6 = 72$
0.999	MCS	-	197.8 m^3/year	10^6
	OK	single	197.8 m^3/year	$12 + 137 = 149$
	PCK	single	198.2 m^3/year	$12 + 30 = 42$
	OK	$K = 6$	198.0 m^3/year	$12 + 32 \cdot 6 = 204$
	PCK	$K = 6$	198.2 m^3/year	$12 + 4 \cdot 6 = 32$
0.9999	MCS	-	235.0 m^3/year	10^6
	OK	single	235.2 m^3/year	$12 + 143 = 155$
	PCK	single	235.2 m^3/year	$12 + 26 = 38$
	OK	$K = 6$	235.3 m^3/year	$12 + 33 \cdot 6 = 210$
	PCK	$K = 6$	235.3 m^3/year	$12 + 10 \cdot 6 = 72$

where E_1, E_2 are the Young's moduli of the linear elastic material, A_1, A_2 are the cross-sections of the bars and $P_1, \ldots, P_6$ are the vertical loads acting on the nodes of the upper part of the structure. The horizontal bars have the properties $\{E_1, A_1\}$, whereas the diagonal bars have the properties $\{E_2, A_2\}$. The input variables are modelled by the probability distributions summarized in Table 5.5. It is assumed that the input variables are statistically independent. A finite element model of the structure is used to calculate the mid-span deflection, denoted by $u(\boldsymbol{x})$, as a function of the ten variables in the input vector $\boldsymbol{x}$, which is defined as positive in the direction indicated in Figure 5.8. This defines the computational model $\mathcal{M}(\boldsymbol{x}) = u(\boldsymbol{x})$.

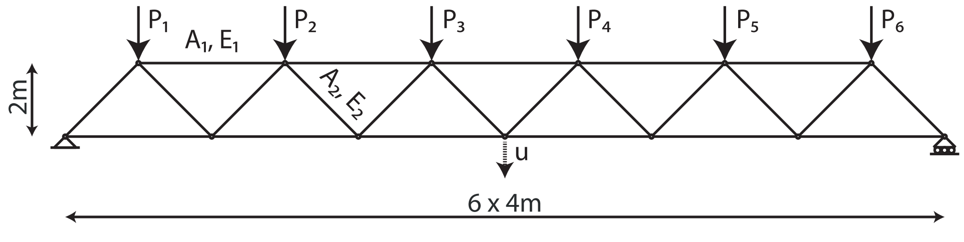


Figure 5.8: Truss structure – sketch of the geometry, material parameters $\{A_i, E_i\}$ and loads P_i.

Failure probability estimation

The failure probability is defined as the probability that the mid-span deflection u is larger than an admissible value u_{adm}, *i.e.* $P_f = \mathbb{P}\left(u\left(\boldsymbol{X}\right) \geq u_{\text{adm}}\right)$. Hence, the limit-state function is defined as $\mathcal{G}\left(\boldsymbol{x}\right) = u_{\text{adm}} - u\left(\boldsymbol{x}\right)$, which is equivalent to $\mathcal{G}(\boldsymbol{x}) = u_{\text{adm}} - \mathcal{M}\left(\boldsymbol{x}\right)$.

The adaptive experimental design algorithm is started with an initial experimental design

Table 5.5: Truss structure – definition of the probabilistic model of the input variables.

Variable	Units	Distribution	Mean	Standard deviation
$E_1,\ E_2$	[Pa]	Lognormal	$2.1 \cdot 10^{11}$	$2.1 \cdot 10^{10}$
A_1	[m^2]	Lognormal	$2.0 \cdot 10^{-3}$	$2 \cdot 10^{-4}$
A_2	[m^2]	Lognormal	$1.0 \cdot 10^{-3}$	$1 \cdot 10^{-4}$
$P_1, \dots, P_6$	[N]	Gumbel	$5.0 \cdot 10^{4}$	$7.5 \cdot 10^{3}$

of $N_0 = 12$ Latin-hypercube samples and the candidate Monte Carlo population $\mathcal{S}$ has a size of $n = 10^6$ samples. A single sample or $K = 6$ samples are added to the experimental design at each iteration.

Table 5.6 summarizes the results of the failure probability estimation. The results are compared to a reference plain Monte Carlo simulation with $n_{\text{MCS}} = 10^6$ samples and to the *first order reliability method* (FORM) in terms of failure probability and Hasofer-Lind reliability index (Y. Zhang & Der Kiureghian 1995). For comparison purposes, the generalized reliability index $\widehat{\beta} = \Phi^{-1}(\widehat{P}_f)$ is also given for the various sampling-based approaches.

Table 5.6: Truss structure – results of the failure probability estimate based on a single run of the adaptive experimental design algorithm for different meta-modelling techniques and enrichment modes.

u_{adm}	Method	Enrichment	$\widehat{P}_f$ (CoV $[P_f]$)	$\widehat{\beta}$	N_{tot}
10 cm	MCS	-	$4.29 \cdot 10^{-2}$ (0.5 %)	1.72	10^6
	FORM	-	$2.81 \cdot 10^{-2}$	1.91	251
	OK	single	$4.32 \cdot 10^{-2}$	1.71	$12 + 135 = 147$
	PCK	single	$4.32 \cdot 10^{-2}$	1.71	$12 + 158 = 170$
	OK	$K = 6$	$4.31 \cdot 10^{-2}$	1.72	$12 + 26 \cdot 6 = 168$
	PCK	$K = 6$	$4.32 \cdot 10^{-2}$	1.71	$12 + 31 \cdot 6 = 198$
12 cm	MCS	-	$1.55 \cdot 10^{-3}$ (2.5 %)	2.96	10^6
	FORM	-	$7.57 \cdot 10^{-4}$	3.17	236
	OK	single	$1.53 \cdot 10^{-3}$	2.96	$12 + 164 = 176$
	PCK	single	$1.52 \cdot 10^{-3}$	2.96	$12 + 157 = 169$
	OK	$K = 6$	$1.53 \cdot 10^{-3}$	2.96	$12 + 27 \cdot 6 = 174$
	PCK	$K = 6$	$1.53 \cdot 10^{-3}$	2.96	$12 + 25 \cdot 6 = 162$
14 cm	MCS	-	$3.6 \cdot 10^{-5}$ (16.7 %)	3.97	10^6
	FORM	-	$1.29 \cdot 10^{-5}$	4.21	231
	OK	single	$3.7 \cdot 10^{-5}$	3.96	$12 + 110 = 122$
	PCK	single	$3.7 \cdot 10^{-5}$	3.96	$12 + 63 = 75$
	OK	$K = 6$	$3.4 \cdot 10^{-5}$	3.99	$12 + 27 \cdot 6 = 174$
	PCK	$K = 6$	$3.2 \cdot 10^{-5}$	4.00	$12 + 11 \cdot 6 = 78$

The values of $\widehat{P}_f$ shown in Table 5.6 are consistent with the reference values for all the adaptive experimental design configurations. The number of runs of the computational model, however, differ considerably. For the cases of $u_{\text{adm}} = 10$ cm and 12 cm the number of exact model runs are similar for both ordinary Kriging (OK) and PC-Kriging (PCK). In the case of $u_{\text{adm}} = 14$ cm, PCK requires significantly fewer runs than OK, for both single and multiple sample selection modes. Similarly to the borehole example, multiple sample selection results in more computational model runs, but overall lower computational times (up to 5-times smaller) due to high performance computing.

The well-established FORM gives an estimate that is worse than the adaptive experimental design algorithm despite the higher cost. It is thus preferable in this application to conduct an adaptive experimental design search rather than using FORM algorithm.

5.5 Discussion

Apart from simple uncertainty propagation, PC-Kriging can be applied in structural reliability analysis. A unified framework is available to estimate both failure probabilities and quantiles. The combination of PC-Kriging and design enrichment based on the probability of misclassification increases the accuracy in the estimation of the statistics of interest. An efficient strategy to add multiple samples at each iteration of the adaptive experimental design algorithm is introduced to enable the use of high-performance computing resources. Further, the adaptive design algorithm is equipped with a new stopping criterion which monitors the convergence of the statistics of interest better than existing stopping criteria. This further reduces the total computational resources needed to accurately estimate the statistics of interest.

The applications in Sections 4.4 and 5.4 confirm the efficiency of the PC-Kriging model on benchmark analytical functions as well as realistic engineering applications. These applications shows the versatility of PC-Kriging and its significance for the engineering practice.

CHAPTER 6

Propagation of p-boxes with meta-models

In this chapter, a number of algorithms are proposed to propagate p-boxes efficiently using meta-models at different stages of the analysis. The algorithms have been published for free p-boxes in Schöbi & Sudret (2017b) and for parametric p-boxes in Schöbi & Sudret (2015c).

6.1 State of the art

As previously seen in the context of probabilistic input modelling, uncertainty propagation methods have been widely studied in the last decades through Monte Carlo simulation. A much lower number of methods have been developed for propagating p-boxes, amongst which are nested Monte Carlo algorithms (Eldred 2009; Chowdhary & Dupuis 2013; He, Mirzargar, & Kirby 2015; Schöbi & Sudret 2015c) and interval-analysis-based algorithms (Helton & Oberkampf 2004; Helton, Johnson, & Oberkampf 2004). Nested Monte Carlo simulations are generally associated to parametric p-boxes, whereas interval analysis is related to free p-boxes. These algorithms require a large number of model evaluations to ensure an accurate estimate of the uncertainty in the quantity of interest (QoI). Hence, in the general case of expensive-to-evaluate models, these types of algorithms become intractable.

Moreover, the use of meta-models to facilitate the uncertainty propagation analysis has been addressed rarely in the literature. Chen et al. (2015) propagate fuzzy sets using generalized polynomial chaos meta-models. Hu & Du (2015) use Kriging meta-models to estimate failure probabilities. Shah, Hosder, & Winter (2015) use collocation methods to propagate Dempster-Shafer structures. Jakeman, Eldred, & Xiu (2010) use PCE to propagate different types of imprecise probabilities. Eldred (2009) and Eldred & Swiler (2009) propagate imprecise probabilities using generalized PCE and stochastic collocation methods.

Making use of the sparse PCE models discussed in Section 3.3, two approaches are proposed in this chapter to propagate the two types of p-boxes efficiently.

6.2 Free p-boxes

6.2.1 Scenarios

In Chapter 2, a variety of methods are presented to define a free p-box based on a dataset and more generally based on information. In the context of uncertainty propagation however, two cases can be distinguished for free p-boxes: bounded and unbounded free p-boxes. Hence, two scenarios are introduced here and discussed in the remainder of this chapter.

Case #1 – interval-valued measurements or expert opinions

Consider the case of interval-valued measurements described in Section 2.3. This scenario can be interpreted as twofold: the intervals are imprecise measurements or the intervals represent the opinion of different experts. The first case is discussed in the latter section and the corresponding free p-box can be defined by Eqs. (2.15) and (2.16).

The second case assumes that a number of experts are asked to name an interval for describing the possible values of a variable. Expert $i = 1, \ldots, n_E$ provides an interval $x^{(i)} \in \left[\underline{x}^{(i)}, \overline{x}^{(i)}\right]$. Additionally, a mass of credibility $w^{(i)}$ is assigned to each expert accounting for the expert's knowledge. Note that the credibility is defined here as a relative value so that $\sum_{i=1}^{n_E} w^{(i)} = 1$. Modifying Eqs. (2.15) and (2.16) and assuming that the disagreement between the various estimates of the variable represents actual variability, the bounds of the free p-box are obtained by the mixture method (Ferson, Kreinovich, et al. 2003):

$$\underline{F}_X(x) = \sum_{i=1}^{n_E} w^{(i)} \cdot \mathbb{I}_{x \geq \overline{x}^{(i)}}(x), \qquad \overline{F}_X(x) = \sum_{i=1}^{n_E} w^{(i)} \cdot \mathbb{I}_{x \geq \underline{x}^{(i)}}(x). \tag{6.1}$$

Case #2 – CDF-shaped expert opinions

Case #2(a) Assume that experts are asked to give their opinion on the behaviour of a variable in the form of a CDF. Each expert provides a formulation for the CDF denoted by $F_X^{(i)}, i = 1, \ldots, n_E$, where the support of X is possibly unbounded. Assuming that these CDFs describe the uncertainty (epistemic and aleatory) in the system, the free p-box can be generated by the envelope of the experts' CDFs (Ferson, Kreinovich, et al. 2003; Fu et al. 2011):

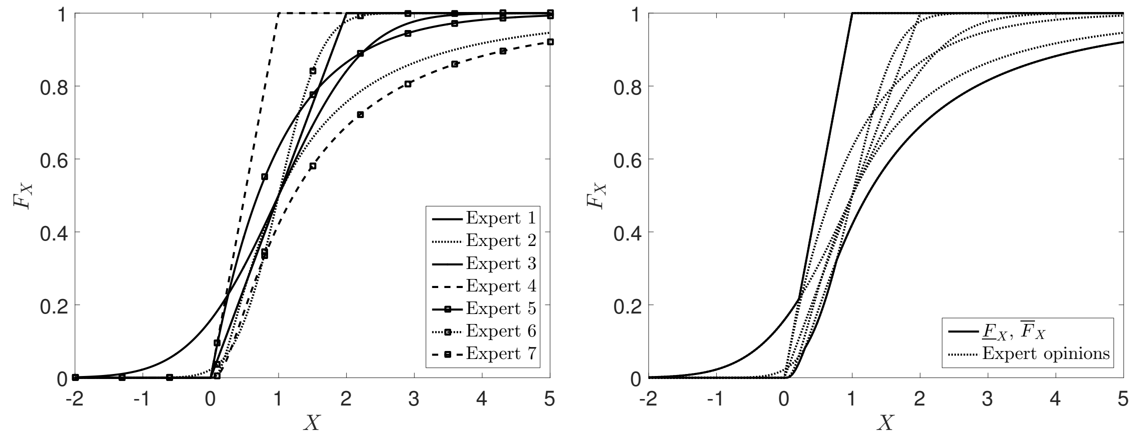
$$\underline{F}_X(x) = \min_{i=1,\ldots,n_E} F_X^{(i)}(x), \qquad \overline{F}_X(x) = \max_{i=1,\ldots,n_E} F_X^{(i)}(x), \qquad \forall x \in \mathcal{D}_X. \tag{6.2}$$

Note that contrary to Case #1, the credibility of the experts is not considered in Case #2. Assuming that the true CDF lies between the experts' CDF, the envelope of all opinions includes the true CDF, thus forming a valid p-box.

An example of seven expert CDFs is shown in Figure 6.1(a). The resulting free p-box for the CDFs in Figure 6.1(a) can be found in Figure 6.1(b). The boundary curves of the free p-box consist of sections of different expert CDFs.

Case #2(b) An alternative way to define a p-box is to choose a distribution family and set interval-valued distribution parameters, *i.e.* a parametric p-box. The corresponding free p-box can be defined by taking the envelope of the set of distributions as in Case #2(a). The free p-box bounds are then found by (see also Eq. (2.20)):

$$\underline{F}_X(x) = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} F_X(x|\boldsymbol{\theta}), \quad \overline{F}_X(x) = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} F_X(x|\boldsymbol{\theta}), \quad \forall x \in \mathcal{D}_X, \tag{6.3}$$



(a) Expert CDFs (b) Aggregated p-box

Figure 6.1: Case #2 – free p-box as a combination of expert CDFs.

where $\boldsymbol{\theta}$ is a vector of distribution parameters which are defined in a domain $\mathcal{D}_{\boldsymbol{\Theta}}$. Note that in this case, only the bounds are used for further analyses. The information about the distribution family is ignored (as opposed to parametric p-boxes) once $\underline{F}_X$ and $\overline{F}_X$ are defined by Eq. (6.3).

6.2.2 Slicing algorithm

The uncertainty propagation of free p-boxes uses the same formalism as in Eq. (3.2):

$$Y = \mathcal{M}\left(\boldsymbol{X}\right) \tag{6.4}$$

where each X_i is modelled by a free p-box and hence Y is a free p-box, too. The traditional approach to estimate the p-box of the response variable Y is the slicing algorithm in combination with interval analysis. The *slicing algorithm* transforms the propagation of p-boxes into the propagation of a large number of intervals, the propagation of which is a well-established field of research related to constraint optimization algorithms (Moore 1966; Dong & Shah 1987; Stolfi & De Figueiredo 2003). The main steps are described in the following:

(i) *Input definition*: The input variables X_i, $i = 1, \ldots, M$ of the computational model $\mathcal{M}$ are modelled by free p-boxes.

(ii) *Discretization*: Discretization methods approximate the p-box by a set of intervals and corresponding probability masses in order to facilitate the uncertainty propagation task. A number of discretization schemes are available in the literature (see *e.g.* Tonon (2004) and H. Zhang, Mullen, & Muhanna (2010)). The *outer discretization method* is now briefly reviewed here. Each input p-box is discretized into a number of intervals and associated probability masses. For variable X_i, the interval $[0, 1]$ is divided into n_{X_i} subintervals with corresponding thickness $m_i^{(j)}$ where $j = 1, \ldots, n_{X_i}$ and $\sum_j m_i^{(j)} = 1$. Let us denote the lower and upper boundary of these intervals by $\underline{c}_i^{(j)}$ and $\overline{c}_i^{(j)}$, respectively. Given the bounds of the free p-box $[\underline{F}_{X_i}, \overline{F}_{X_i}]$, the corresponding intervals in $\mathcal{D}_{X_i}$ are:

$$\underline{x}_i^{(j)} = \overline{F}_{X_i}^{-1}\left(\underline{c}_i^{(j)}\right), \qquad \overline{x}_i^{(j)} = \underline{F}_{X_i}^{-1}\left(\overline{c}_i^{(j)}\right), \tag{6.5}$$

for $j = 1, \ldots, n_{X_i}$ and $i = 1, \ldots, M$. The intervals of interest are then $\left[\underline{x}_i^{(j)}, \overline{x}_i^{(j)}\right]$ and the associated probability masses are $m_{X_i}^{(j)}$, which together characterize the free p-box of the input variable X_i.

(iii) *Interval propagation*: Let $\mathcal{K}$ be a set of multi-indices defining a combination of intervals of each input parameter X_i:

$$\mathcal{K} = \{\boldsymbol{k} = (k_1, \dots, k_M),\ k_i \in \{1, \dots, n_{X_i}\},\ i = 1, \dots, M\}. \tag{6.6}$$

Let $\mathcal{D}_{\boldsymbol{k}}$ be the hyper-rectangle defined by:

$$\mathcal{D}_{\boldsymbol{k}} = \left[\underline{x}_1^{(k_1)}, \overline{x}_1^{(k_1)}\right] \times \dots \times \left[\underline{x}_M^{(k_M)}, \overline{x}_M^{(k_M)}\right]. \tag{6.7}$$

For each hyper-rectangle $\mathcal{D}_{\boldsymbol{k}}$, two optimization problems are solved to define the associated bounds of Y:

$$\underline{y}^{(\boldsymbol{k})} = \min_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{k}}} \mathcal{M}(\boldsymbol{x}), \qquad \overline{y}^{(\boldsymbol{k})} = \max_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{k}}} \mathcal{M}(\boldsymbol{x}). \tag{6.8}$$

The probability mass associated to $\mathcal{D}_{\boldsymbol{k}}$ can be computed by:

$$m_Y^{(\boldsymbol{k})} = m_{X_1}^{(k_1)} \cdot m_{X_2}^{(k_2)} \cdot \ldots \cdot m_{X_M}^{(k_M)}. \tag{6.9}$$

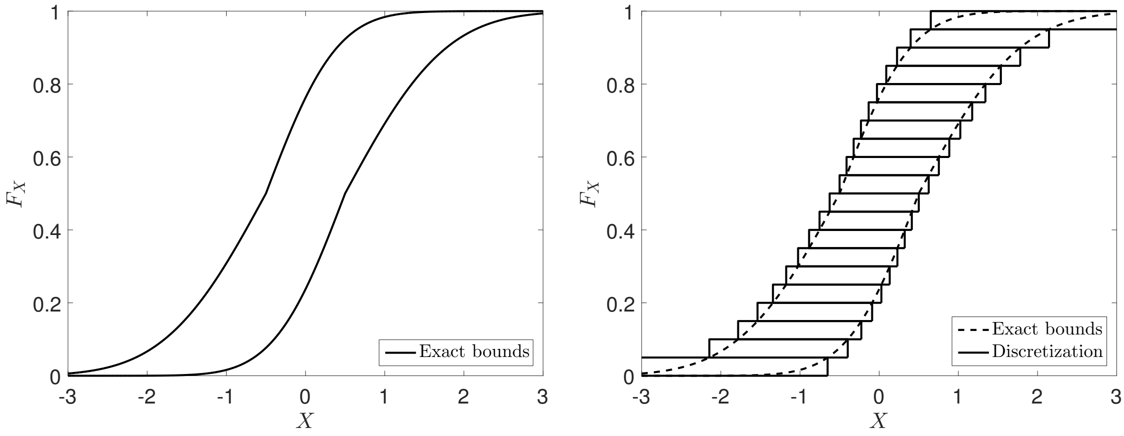
Correspondingly, the free p-box of Y is eventually characterized by $n_Y = n_{X_1} \cdot n_{X_2} \cdot \ldots \cdot n_{X_M}$ intervals $\left[\underline{y}^{(\boldsymbol{k})}, \overline{y}^{(\boldsymbol{k})}\right]$ given in Eq. (6.8) with associated probability masses given in Eq. (6.9). Hence, $2 \cdot n_Y$ optimization algorithms (see Eq. (6.8)) are required in the M-dimensional optimization domain in order to propagate the input p-boxes. When M and n_{X_i} become large, this quickly becomes intractable due to the large number of optimizations. This problem is often referred to as the *curse of dimensionality.*

A number of methodologies can be found in the literature to simplify the optimizations, amongst which are the classical interval analysis (Moore 1966), affine arithmetic (Stolfi & De Figueiredo 2003) and the vertex method (Dong & Shah 1987; Dubois, Fargier, & Fortin 2004). However, these simplifications require restrictive assumptions, such as monotonicity in the computational model, to ensure accuracy. Other optimization algorithms can be applied such as local, derivative-based methods (*e.g.* BFGS algorithm (Byrd, Hribar, & Nocedal 1999)), global methods (*e.g.* genetic algorithms (Goldberg 1989) and differential evolution algorithms (Storn & K. Price 1997; Deng et al. 2013)) and hybrid methods (*i.e.* with a global and local component) for solving directly Eq. (6.8). They are more accurate in the general case but require extensive computational resources.

(iv) *Merging*: The result of the previous step is a set of $n_Y = |\mathcal{K}|$ intervals $\left[\underline{y}^{(\boldsymbol{k})}, \overline{y}^{(\boldsymbol{k})}\right]$ and the corresponding probability masses $m_y^{(\boldsymbol{k})}$. The response free p-box is then obtained by converting $\left\{\overline{y}^{(\boldsymbol{k})}, m_y^{(\boldsymbol{k})}\right\}$ and $\left\{\underline{y}^{(\boldsymbol{k})}, m_y^{(\boldsymbol{k})}\right\}$ to weighted empirical CDFs $\underline{F}_Y$ and $\overline{F}_Y$ as explained in Eq. (6.1).

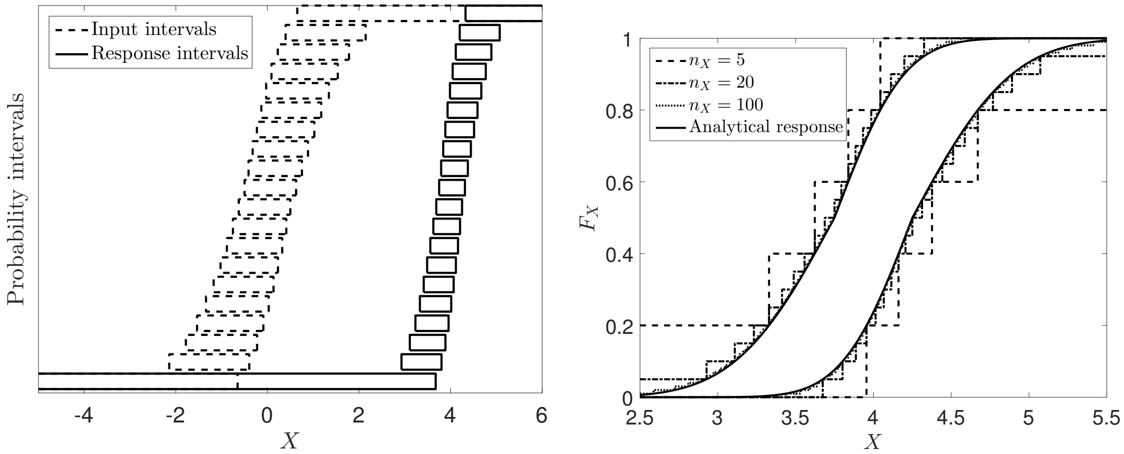
Figure 6.2 illustrates the main steps of the slicing algorithm on a one-dimensional problem. The bounds of the input p-box are defined by two Gaussian distributions with $\mu_X = [1.5,\ 2]$ and $\sigma_X = [0.7,\ 1.0]$ (Case #2(b)) (Figure 6.2(a)). The p-box is discretized with $n_X = 20$ equally spaced subintervals, *i.e.* $m^{(j)} = 1/20$ (Figure 6.2(b)). The computational model is $y = x/2 + 4$ (Figure 6.2(c)). Finally, the resulting intervals are merged to a free p-box in Figure 6.2(d).

Figure 6.2(d) illustrates the influence of n_X on the accuracy of the response p-box. It can be seen from this simple example already that the number of discretization points n_{X_i} is crucial to the accuracy of the response p-box: the approximated response p-box is conservative in the sense that it is wider than the exact response p-box obtained analytically. The larger n_{X_i}, the more accurate is the approximated p-box, and at the same time the larger the computational



(a) Input p-box

(b) Discretization ($n_X = 20$)



(c) Interval propagation ($y = \mathcal{M}(x) = x/2 + 4$)

(d) Merging

Figure 6.2: Illustration of the slicing algorithm. The bounds of the input p-box are defined from Gaussian CDFs with interval-valued mean value and standard deviation. $n_X = 20$ intervals are used.

costs (due to the number of intervals and the related number of optimization operations). This effect is more pronounced when the input vector is multi-dimensional, as discussed later in the application examples in Section 6.3.

Note that in Case #1, the intervals might be chosen to represent each expert's opinion. Then, the expert's credibility is equal to the probability mass in uncertainty propagation, *i.e.* $m^{(j)} = w^{(j)}$, $j = 1, \dots, n_{X_i}$. This is feasible particularly when M and n_{X_i} are small.

6.2.3 Problem conversion

A disadvantage of the slicing algorithm is the full-factorial design, which leads to a potentially large number of optimizations for high-dimensional problems, namely $n_Y = |\mathcal{K}| = n_{X_1} \cdot n_{X_2} \cdot \ldots \cdot n_{X_M}$. In order to circumvent the effects of the full factorial approach, the imprecise problem setting is reformulated in this section, as originally proposed by H. Zhang, Mullen, & Muhanna (2010) and Schöbi & Sudret (2015b). They replace the full factorial design of the slicing algorithm with a random sampling-based approach as follows. Consider the slicing algorithm with a large number of subintervals $n_{X_i} \to \infty$ (and $|\mathcal{K}| \to \infty$) in order to accurately estimate the response p-box Y. The corresponding number of response intervals will be $n_Y \to \infty$, too. Then,

Y can be approximated by propagating a random subset $\mathcal{K}' \subset \mathcal{K}$ to reduce the computational costs. This approximation allows to estimate Y more efficiently than the full-factorial design.

Consider the random vector $\boldsymbol{C}$ made of independent, uniformly distributed variables in the unit-hypercube domain $\mathcal{D}_{\boldsymbol{C}} = [0,1]^M$. The random variable C_i shall describe the CDF value of input variable X_i, *i.e.* $C_i = F_{X_i}(X_i)$. In other words, C_i describes the index of one of the $n_{X_i} \to \infty$ subintervals in $\mathcal{K}$. Given a free p-box in X_i, each $c_i \in [0,1]$ corresponds then to an interval $[\underline{x}_i, \overline{x}_i]$ in $\mathcal{D}_{X_i}$ through the inverse CDF of the bounds of the p-box:

$$\underline{x}_i(c_i) = \overline{F}_{X_i}^{-1}(c_i), \qquad \overline{x}_i(c_i) = \underline{F}_{X_i}^{-1}(c_i). \tag{6.10}$$

This equation is a pseudo inverse transformation due to the potentially stairs-shaped CDF boundary curves (*e.g.* Case 1). In fact, when $\underline{F}_{X_i}$ and $\overline{F}_{X_i}$ are a stairs-shaped functions, then:

$$\underline{x}_i(c_i) = \arg \min_{x_i \in \mathcal{D}_{X_i}} \overline{F}_{X_i}(x) \geq c_i, \qquad \overline{x}_i(c_i) = \arg \min_{x_i \in \mathcal{D}_{X_i}} \underline{F}_{X_i}(x) \geq c_i. \tag{6.11}$$

Note that compared to the previous definition in Eq. (6.5) where the interval $[\underline{c}_i, \overline{c}_i]$ is the input argument, Eq. (6.10) uses a single value c_i. Eq. (6.10) can be interpreted as a special case of Eq. (6.5) when setting $n_{X_i} \to \infty$ and thus $m_i^{(j)} \to 0$ and $\underline{c}_i \approx \overline{c}_i$. For a given realization $\boldsymbol{c} \in \mathcal{D}_{\boldsymbol{C}}$, let us denote a hyper-rectangular domain by $\mathcal{D}_{\boldsymbol{c}} = [\underline{x}_1(c_1), \overline{x}_1(c_1)] \times \ldots \times [\underline{x}_M(c_M), \overline{x}_M(c_M)]$. As a consequence, two computational models can be formulated as a function of $\boldsymbol{c}$, equivalent to Eq. (6.8) (Alvarez, Hurtado, & Uribe 2014; Schöbi & Sudret 2015b):

$$\underline{y} = \underline{\mathcal{M}}(\boldsymbol{c}) = \min_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{c}}} \mathcal{M}(\boldsymbol{x}), \qquad \overline{y} = \overline{\mathcal{M}}(\boldsymbol{c}) = \max_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{c}}} \mathcal{M}(\boldsymbol{x}). \tag{6.12}$$

These equations lead to intervals $\left[\underline{y}, \overline{y}\right]$ of Y. In analogy with the slicing algorithm (see also Figure 6.2), the lower bounds of the intervals model the upper boundary curve of Y (and vice versa). Hence, the lower bound model $\underline{\mathcal{M}}$ maps $\boldsymbol{C}$ to the upper bound $\overline{Y}$ and the upper bound model $\overline{\mathcal{M}}$ maps $\boldsymbol{C}$ to the lower bound response $\underline{Y}$:

$$\underline{Y} = \overline{\mathcal{M}}\left(\boldsymbol{C}\right), \qquad \overline{Y} = \underline{\mathcal{M}}\left(\boldsymbol{C}\right), \tag{6.13}$$

where $\underline{Y}$ and $\overline{Y}$ are characterized by the CDF $\underline{F}_Y$ and $\overline{F}_Y$ respectively. Interestingly, Eq. (6.13) splits the propagation of free p-boxes into two standard uncertainty propagation problems associated with input random vector $\boldsymbol{C}$.

The probabilistic description of the auxiliary input vector $\boldsymbol{C}$ allows for conventional methods of uncertainty propagation such as random sampling (Monte Carlo simulation, Latin-hypercube sampling (McKay, Beckman, & Conover 1979)) and low-discrepancy sequences (Sobol' sequence (Sobol' 1967), Halton sequence (Halton 1960)). These methods are more efficient than the full factorial design approach described in Section 6.2.2, but they are not conservative with respect to the p-box due to the nature of sampling methods, as opposed to the previously discussed slicing algorithm. Sufficiently large sample sets must then be used to ensure proper convergence and high accuracy.

6.2.4 Two-level meta-modelling

Basic idea

Considering the main steps of the uncertainty propagation of p-boxes, there are three main factors contributing to the total computational effort. Firstly, the computational model is evaluated a large number of times due to the optimization methods, in particular when using global optimization methods in the general case when model monotonicity does not hold. Secondly, the number of optimization operations is large considering the sampling-based approach for estimating the bounds of the response p-box, *i.e.* $\underline{Y}$ and $\overline{Y}$. Last but not least, the cost of a single evaluation of the computational model may affect the total costs considerably.

In order to address these three factors, it is proposed to surrogate the computational model at two levels by sparse PCE models. More specifically, the first-level meta-model approximates the response of $\mathcal{M}$, whereas the second-level one approximates the response of the lower and upper model denoted by $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$.

Meta-modelling $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$

Eq. (6.13) describes the bounds of the response p-box as a function of the probabilistic vector $\boldsymbol{C}$. Hence, classical meta-modelling techniques can be applied in a straightforward manner. The corresponding meta-models using the sparse PCE technique read:

$$\underline{Y} \approx \overline{\mathcal{M}}^{(\mathrm{P})}(\boldsymbol{C}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_l} \overline{a}_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{C}), \tag{6.14}$$

$$\overline{Y} \approx \underline{\mathcal{M}}^{(\mathrm{P})}(\boldsymbol{C}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_u} \underline{a}_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{C}), \tag{6.15}$$

where $\overline{\boldsymbol{a}}$ and $\underline{\boldsymbol{a}}$ are the corresponding PC coefficients. Note that because of the LARS algorithm in the sparse PCE models, the two index sets $\mathcal{A}_l$ and $\mathcal{A}_u$ are likely to be not equal. The meta-models $\overline{\mathcal{M}}^{(\mathrm{P})}$ and $\underline{\mathcal{M}}^{(\mathrm{P})}$ are based on the experimental design $\mathfrak{C} = \left\{\mathfrak{c}^{(1)}, \ldots, \mathfrak{c}^{(N)}\right\}$ and the corresponding response values

$$\underline{\mathcal{Y}} = \left\{\underline{\mathcal{Y}}^{(1)} = \overline{\mathcal{M}}\left(\mathfrak{c}^{(1)}\right), \ldots, \underline{\mathcal{Y}}^{(N)} = \overline{\mathcal{M}}\left(\mathfrak{c}^{(N)}\right)\right\},$$

$$\overline{\mathcal{Y}} = \left\{\overline{\mathcal{Y}}^{(1)} = \underline{\mathcal{M}}\left(\mathfrak{c}^{(1)}\right), \ldots, \overline{\mathcal{Y}}^{(N)} = \underline{\mathcal{M}}\left(\mathfrak{c}^{(N)}\right)\right\},$$

respectively.

Meta-modelling $\mathcal{M}$

Condensation of free p-boxes The expensive-to-evaluate computational model $\mathcal{M}$ may itself be approximated with a sparse PCE model in view of solving the global optimization problems defining $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$. The input of $\mathcal{M}$ is a free p-box in M dimensions in the problems on consideration. However, in order to apply a sparse PCE model, a probabilistic input vector is required. Hence, auxiliary input variables $\widetilde{X}_i$ are defined for the sole purpose of meta-modelling $\mathcal{M}$. The auxiliary input variables should represent the probability mass in the input free p-boxes in an appropriate manner. In other words, the auxiliary distributions aim at "summarizing" the free p-box, such as fulfilling $\underline{F}_X(x) < F_{\widetilde{X}}(x) < \overline{F}_X(x), \forall x \in \mathcal{D}_X$. This is the so-called *condensation* phase. As free p-boxes are defined on interval-valued CDFs, a number of distributions can be proposed, *i.e.* there is no unique choice.

Generally speaking, for Case #1 (Section 6.2.1) and bounded free p-boxes, it is proposed to use a uniform distribution between the minimum and maximum value of the free p-boxes:

$$\widetilde{X}_i \sim \mathcal{U}\left(\min_{j=1\ldots,n_E}\left(\underline{x}_i^{(j)}\right), \max_{j=1,\ldots,n_E}\left(\overline{x}_i^{(j)}\right)\right), \tag{6.16}$$

where $\left[\underline{x}_i^{(j)}, \overline{x}_i^{(j)}\right]$ is the interval on variable i provided by the j-th expert. For Case #2 (Section 6.2.1) and generally for unbounded free p-boxes, it is proposed to define the CDF of $\widetilde{X}_i$ as an average curve of its input free p-boxes:

$$F_{\widetilde{X}_i}(x_i) = \frac{1}{2}\left(\underline{F}_{X_i}(x_i) + \overline{F}_{X_i}(x_i)\right). \tag{6.17}$$

In both cases, the proposed auxiliary input distribution covers the free p-box in the areas where most of the probability mass is located and is therefore suitable to represent the free p-box in a sparse PCE model. However, in terms of accuracy of the meta-model, different distributions might be more suitable than the one in Eq. (6.17).

Arbitrary input PCE Having defined an auxiliary input vector $\widetilde{\boldsymbol{X}}$, it can be propagated through the computational model. The corresponding uncertainty propagation problem then reads:

$$\widetilde{Y} = \mathcal{M}\left(\widetilde{\boldsymbol{X}}\right), \tag{6.18}$$

and the corresponding approximation with PCE (see also Eq. (3.8)):

$$\widetilde{Y} \approx \mathcal{M}^{(\mathrm{P})}\left(\widetilde{\boldsymbol{X}}\right) = \sum_{\boldsymbol{\alpha}\in\widetilde{\mathcal{A}}} \widetilde{a}_{\boldsymbol{\alpha}}\ \psi_{\boldsymbol{\alpha}}\left(\widetilde{\boldsymbol{X}}\right). \tag{6.19}$$

This PCE model can be trained by pure vanilla least-square analysis, least-angle regression or any other non-intrusive technique (see Section 3.3).

An important aspect here is when $\widetilde{X}_i$ has an arbitrary CDF shape for which it might not be trivial to define a set of orthogonal polynomials (Gautschi 2004). Wan & Karniadakis (2006), Witteveen, Sarkar, & Bijl (2007), Ahlfeld, Belkouchi, & Montomoli (2016) and Dey et al. (2016) discuss the use of so-called *arbitrary PCE*, which are based on arbitrarily shaped input distributions. An alternative way is to formalize an isoprobabilistic transform from variables $\widetilde{X}_i$ to Z_i for which a suitable set of orthogonal polynomials is known already. The mapping from one distribution to the other is denoted by $T : Z_i = T\left(\widetilde{X}_i\right)$. A point $x_i \sim \widetilde{X}_i$ can be mapped to Z_i as $z_i = F_{Z_i}^{-1}\left(F_{\widetilde{X}_i}(x_i)\right)$. When $\widetilde{X}_i$ and Z_i are chosen in a smart way, the transform T shall be nearly linear. The sparse PCE model can then be written as:

$$\widetilde{Y} \approx \mathcal{M}^{(\mathrm{P})}(\widetilde{\boldsymbol{X}}) = \sum_{\boldsymbol{\alpha}\in\widetilde{\mathcal{A}}} \widetilde{a}_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}\left(T\left(\widetilde{\boldsymbol{X}}\right)\right) = \sum_{\boldsymbol{\alpha}\in\widetilde{\mathcal{A}}} \widetilde{a}_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}\left(\boldsymbol{Z}\right), \tag{6.20}$$

where the set of $\psi_{\boldsymbol{\alpha}}$'s are orthogonal with respect to the vector $\boldsymbol{Z} = T\left(\widetilde{\boldsymbol{X}}\right)$.

Aggregation of the two levels

Framework of meta-models Figure 6.3 summarizes the two levels of meta-modelling presented previously. A first level of sparse PCE approximates the computational model $\mathcal{M}$ on the basis of the auxiliary input vector $\widetilde{\boldsymbol{X}}$, which is itself defined according to the type of input free p-box considered. This results in the meta-model $\mathcal{M}^{(\mathrm{P})}$. Then, the problem is divided into the estimation of $\underline{Y}$ and $\overline{Y}$ by Eq. (6.13). They are obtained by approximating $\overline{\mathcal{M}}$ and $\underline{\mathcal{M}}$, respectively, via sparse PCE and the probabilistic input vector $\boldsymbol{C}$, which defines the second level of the two-level approach.

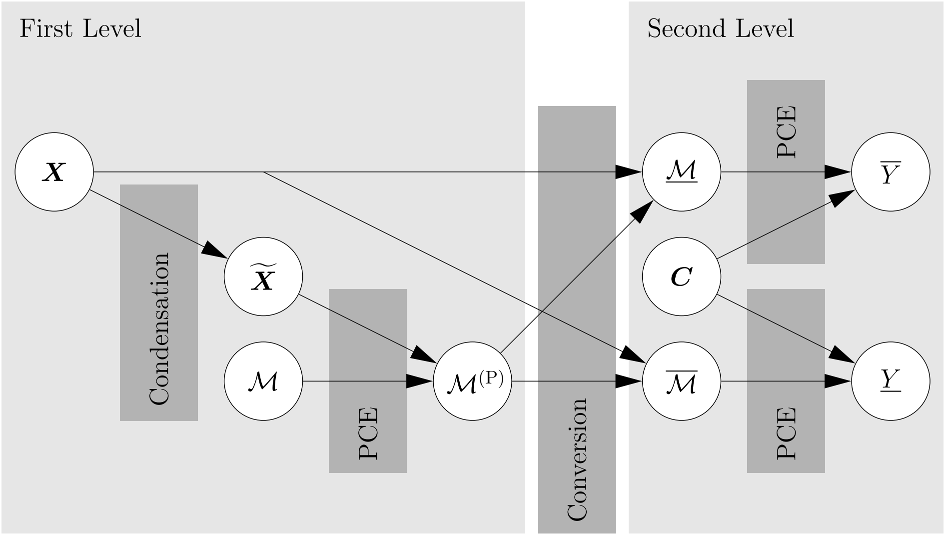


Figure 6.3: Two-level meta-modelling for the propagation of p-boxes – definition and connection of the two levels.

Case #1 In the case of bounded free p-boxes, *i.e.* when the support of the boundary CDFs $\underline{F}_{X_i}$ and $\overline{F}_{X_i}$ are compact, the methods can be applied as described in the previous sections. For the first level of meta-modelling, bounded variables $\widetilde{\boldsymbol{X}}$ are used so that the response $\widetilde{Y}$ may be approximated efficiently. On the second-level meta-model however, the response $\underline{Y} = \overline{\mathcal{M}}(\boldsymbol{C})$ and $\overline{Y} = \underline{\mathcal{M}}(\boldsymbol{C})$ are non-smooth due to the stair-like input free p-boxes X_i originating in a finite set of expert intervals. The effect of this non-smoothness is investigated further in the example in Section 6.3.2. However, when the number of expert intervals and the number of variables are small, the second-level meta-model may be redundant, as the full factorial design approach is tractable (*i.e.* n_Y small). In the other extreme case of large number of variables and intervals, the response boundary curves loose the stair-shaped nature of the CDF curves due to the large number of response intervals (*i.e.* n_Y large).

Case #2 In the case of unbounded free p-boxes, the uncertainty propagation analysis, as proposed above, may become inefficient due to the usage of a bounded $\boldsymbol{C}$ on the second-level meta-model. In general, it is not advised to model an unbounded variable X by a bounded variable C and the corresponding isoprobabilistic transform, due to the highly non-linear functional form of this transform (Schöbi & Sudret 2015b, 2016b).

In order to reduce the effect of the non-linear transform and to ensure fast convergence, the second-level meta-model is trained in the auxiliary input domain $\widetilde{\boldsymbol{X}}$ as previously discussed in Schöbi & Sudret (2015b). An isoprobabilistic transform can be formulated such that $\boldsymbol{C} = T\left(\widetilde{\boldsymbol{X}}\right)$. The corresponding models are:

$$\underline{Y} = \overline{\mathcal{M}}\left(T\left(\widetilde{\boldsymbol{X}}\right)\right), \qquad \overline{Y} = \underline{\mathcal{M}}\left(T\left(\widetilde{\boldsymbol{X}}\right)\right). \tag{6.21}$$

These two computational models can be approximated using sparse PCE meta-models. The effect of this additional isoprobabilistic transform is discussed further in the application example in Section 6.3.2.

Special case: monotone computational models When the computational model $\mathcal{M}$ is known to be monotonic with respect to all variables X_i, then the constraint optimization in

Eq. (6.12) reduces to the analysis of the corners of the search domain. In the context of interval analysis, this approach is called *vertex method* (Dong & Shah 1987; Dubois, Fargier, & Fortin 2004). Then, uncertainty propagation of p-boxes simplifies to the propagation of the bounds of the input p-boxes. It follows that the meta-modelling of $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$ becomes out of scope. $\underline{Y}$ and $\overline{Y}$ can be directly estimated based on $\mathcal{M}^{(\mathrm{P})}$ and the bounds of each X_i.

6.3 Numerical examples for free p-boxes

6.3.1 Reference solution

In the following application examples, the proposed two-level meta-modelling algorithm is compared to a reference solution. The reference solution is obtained by (i) using the original computational model $\mathcal{M}$ (no first-level meta-model $\mathcal{M}^{(\mathrm{P})}\left(\widetilde{\boldsymbol{X}}\right)$), (ii) using a large number of points ($n = 10^6$) in the prediction of $\underline{Y}$ and $\overline{Y}$, and by (iii) using either a hybrid optimization algorithm or the exact optimization value to find $\underline{y}$ and $\overline{y}$.

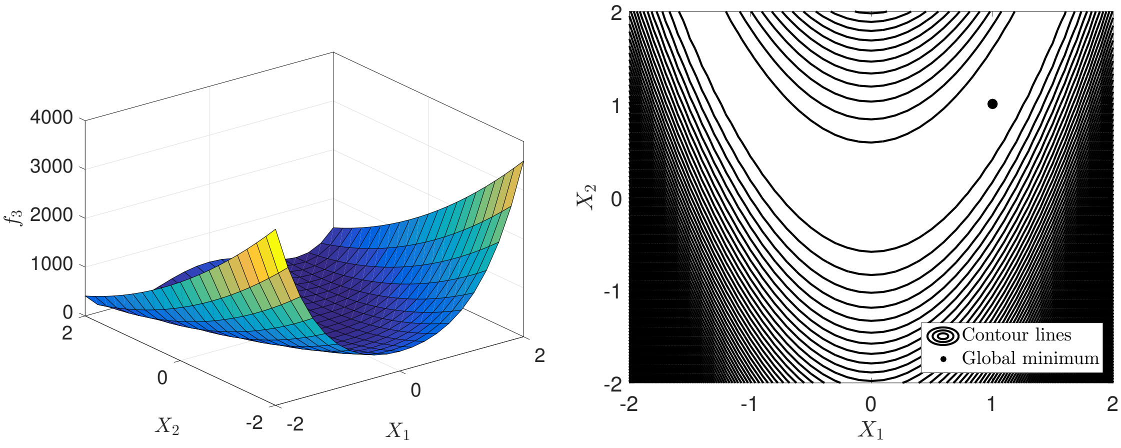
6.3.2 Rosenbrock function

Problem statement

The first example is the Rosenbrock function, defined previously in Eq. (4.4), which is recalled here:

$$f_3(\boldsymbol{x}) = 100\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2 . \tag{6.22}$$

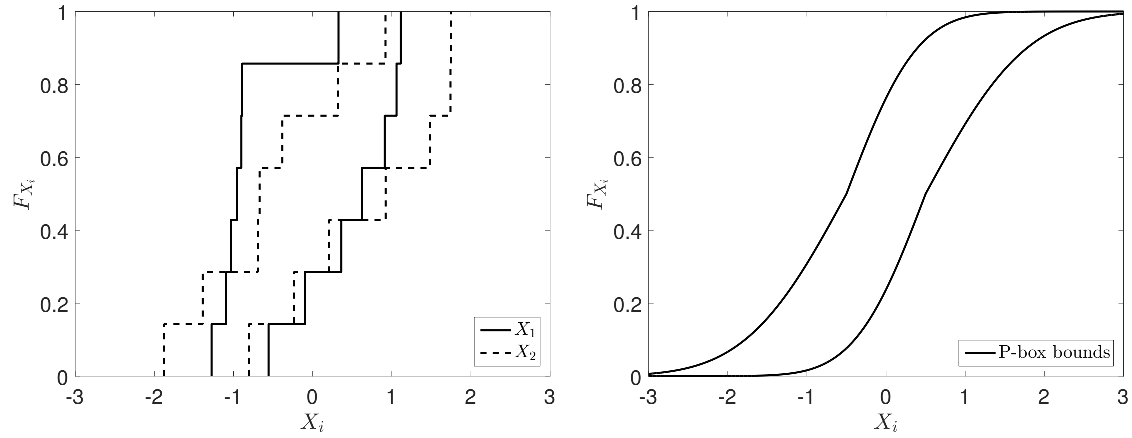
Here, the two input parameters X_1 and X_2 are modelled by independent free p-boxes. Figure 6.4(a) shows the response surface of f_3 as a function of x_1 and x_2. The response is non-monotone around the origin of the input space and has a global minimum of $y = 0$ at $\{x_1 = 1,\, x_2 = 1\}$. Thus, Y is bounded on one side to the domain $\mathcal{D}_Y = [0, \infty]$. Figure 6.4(b) shows a contour plot of f_3 and its global minimum marked by the large black dot. The response surface is flat around the minimum response value and steep at the boundaries of the figure domain.



(a) 3D visualization (b) Contour lines

Figure 6.4: Rosenbrock function – response surface $y = f_3(\boldsymbol{x})$.

Case #1 The free p-boxes based on the opinions of seven experts are illustrated in Figure 6.5(a). The solid lines mark the p-box for X_1, whereas the dashed lines mark the p-box for X_2. It can be seen that the credibility of the experts is uniform due to the constant vertical step size between the vertical plateaus of the boundary CDFs, and that the free p-boxes are bounded on both sides.



(a) Case #1

(b) Case #2

Figure 6.5: Rosenbrock function – input free p-boxes – boundary curves of X_i.

Case #2 The bounds of the two input variables X_i are defined as follows (corresponding to Case #2(b)):

$$\underline{F}_{X_i}(x_i) = \min_{\mu \in [\underline{\mu}, \overline{\mu}], \sigma \in [\underline{\sigma}, \overline{\sigma}]} F_{\mathcal{N}}(x_i | \mu, \sigma),$$

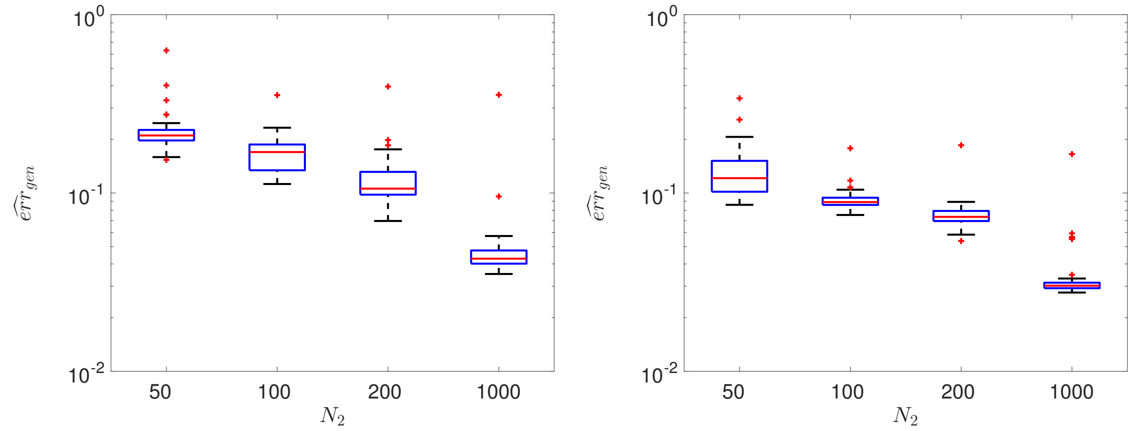
$$\overline{F}_{X_i}(x_i) = \max_{\mu \in [\underline{\mu}, \overline{\mu}], \sigma \in [\underline{\sigma}, \overline{\sigma}]} F_{\mathcal{N}}(x_i | \mu, \sigma),$$

where $F_{\mathcal{N}}(x|\mu, \sigma)$ is a Gaussian CDF with mean value μ and standard deviation σ. In order to generate a free p-box, the distribution parameters are given in intervals: $\mu \in [\underline{\mu}, \overline{\mu}] = [-0.5, 0.5]$ and $\sigma \in [\underline{\sigma}, \overline{\sigma}] = [0.7, 1.0]$. Figure 6.5(b) shows the boundary curves of the free p-boxes for X_i, $i = 1, 2$.

Analysis

The two-level meta-modelling approach has been implemented taking advantage of the Matlab-based uncertainty quantification framework UQLab (Marelli & Sudret 2014, 2015). The sparse PCE meta-models were calibrated with Latin-hypercube experimental designs and appropriate sets of orthonormal polynomials (Legendre polynomials for bounded variables and Hermite polynomials for unbounded variables). The set of polynomials is determined by a degree-adaptive LARS of which the maximum total polynomial degree is set to 30 and the parameter for the hyperbolic truncation set $\mathcal{A}_q^{M,p}$ is set to $q = 0.75$. The number of samples in the experimental design is denoted by N_1 and N_2 for constructing the meta-models of $\mathcal{M}$ and $\{\underline{\mathcal{M}}, \overline{\mathcal{M}}\}$, respectively. In order to achieve a statistical significance, the uncertainty propagation analyses are replicated 50 times with different LHS experimental designs.

For Case #1, the auxiliary input variables are defined as in Eq. (6.16), whereas for Case #2, the auxiliary input variables are defined as $F_{\widetilde{X}_i} = F_{\mathcal{N}}(x|0, 1)$. Note that for Case #2, both levels of meta-models are trained in the same auxiliary domain, *i.e.* $\mathcal{D}_{\widetilde{\boldsymbol{X}}}$.



(a) Response $\underline{Y}$ (*i.e.* based on $\overline{\mathcal{M}}$) (b) Response $\overline{Y}$ (*i.e.* based on $\underline{\mathcal{M}}$)

Figure 6.6: Rosenbrock function – Case #1 – relative generalization error as a function of N_2 and based on 50 replications of the same analysis ($N_1 = 30$).

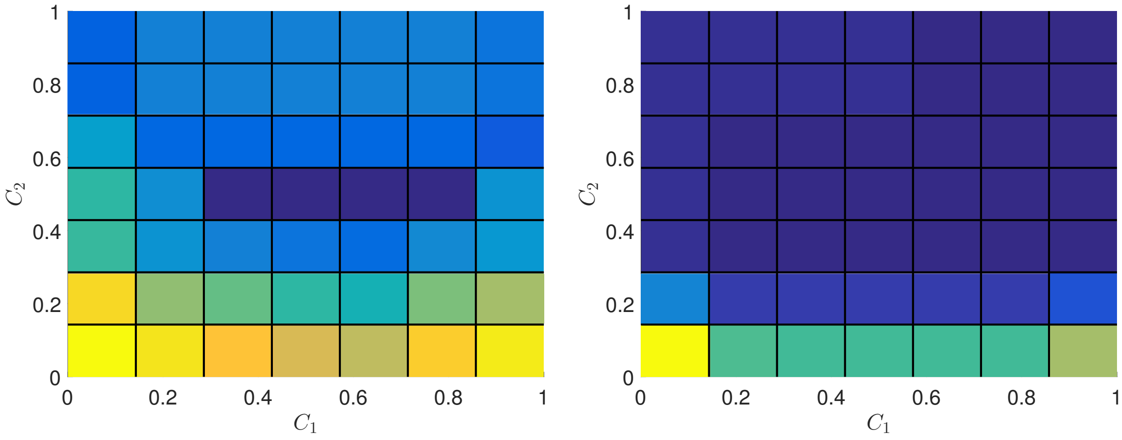
Results

Case #1 The polynomial form of the Rosenbrock function can be modelled exactly with the polynomial-based PCE models. Hence, the first-level meta-model does not introduce any approximation in this application example. Therefore, the first-level meta-model is not discussed thereafter.

The reference solution for the second-level models $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$ is obtained by propagating all possible combinations of input intervals, *i.e.* $n_Y = n_{E,X_1} \cdot n_{E,X_2} = 7 \cdot 7 = 49$ intervals. This results in stair-shaped boundary curves for the response p-box. These shapes are difficult to model with polynomials, which are smooth functions. Hence, the accuracy of the second-level meta-models is low, as seen in Figure 6.6, which summarizes the relative generalization error for the two second-level meta-models, *i.e.* $\overline{\mathcal{M}}^{(P)}$ and $\underline{\mathcal{M}}^{(P)}$. In fact, it requires $N_2 = 1000$ samples to achieve an accuracy of $\widehat{err}_{gen} \approx 0.05$ due to the shape of the input p-boxes. The continuous input vector $\boldsymbol{C}$ is mapped onto 49 interval domains of the input space $\mathcal{D}_{\boldsymbol{X}}$ and hence also in the response variables $\underline{Y}$ and $\overline{Y}$.

The explanation for this convergence behaviour is illustrated in Figure 6.7, which shows the response surface of the computational models as a function of the input random variables. The computational model $\mathcal{M}$, which is the basis for the first-level meta-model, shows smooth contour lines (see Figure 6.4(b)), whereas $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$ show plateau-shaped response surfaces (see Figures 6.7(a) and 6.7(b)), respectively, with constant response values within each plateau. A total of $7 \times 7 = 49$ distinct plateaus can be identified corresponding to the 49 possible combinations of input intervals defined by the expert opinions. As an alternative to training second-level meta-models, the original (small) number of input intervals, *i.e.* the intervals given by the experts, may be propagated directly to estimate the bounds of the response p-box. In other words, the slicing algorithm may be applied on the second-level of the uncertainty analysis, due to the low number of response intervals, *i.e.* $|\mathcal{K}| = 49$.

When comparing the boundary curves of the response p-box, however, results show that the proposed approach provides accurate (although smooth) approximations of the response p-box boundaries. The exact boundary curves of the p-box of Y are shown in Figure 6.8, where the distinct values are clearly visible in the stair-shaped CDF curves. For comparison, the boundary curves of the two-level approach are presented for a single run with $N_1 = 50$ and $N_2 = 200$. For



(a) $\overline{\mathcal{M}}$ (b) $\underline{\mathcal{M}}$

Figure 6.7: Rosenbrock function – Case #1 – shape of the response surfaces.

this run, the relative generalization error are $\widehat{err}_{gen}\left[\overline{Y}\right] = 1.14 \cdot 10^{-1}$ and $\widehat{err}_{gen}\left[\underline{Y}\right] = 8.76 \cdot 10^{-2}$. The PCE-based CDFs are not capable of reproducing the stair-like functions but still follow the exact boundary curves sufficiently accurately for visual comparison. This confirms the large relative generalization errors in the second-level meta-models.

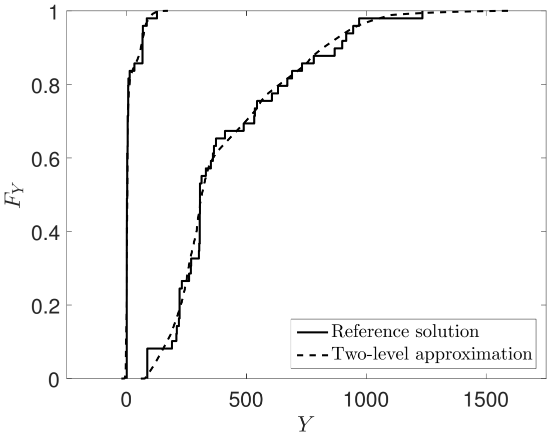
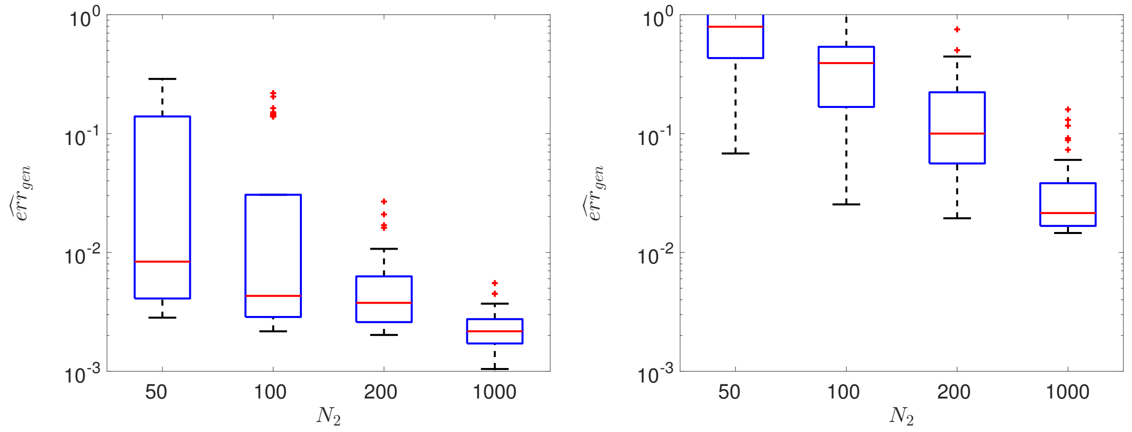


Figure 6.8: Rosenbrock function – Case #1 – response p-boxes $Y \in \left[\underline{Y}, \overline{Y}\right]$: reference solution ($n = 7 \cdot 7 = 49$) versus two-level approach ($N_1 = 50$, $N_2 = 200$).

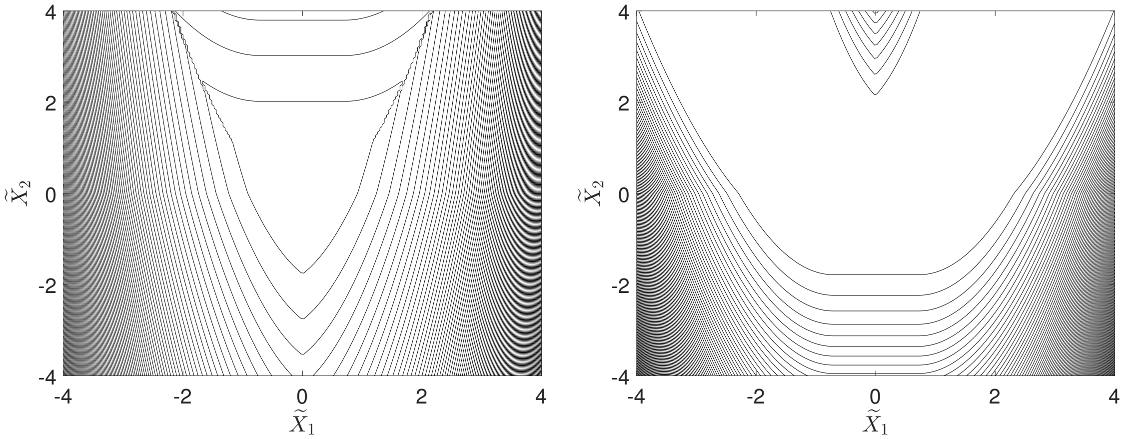
Case #2 Analogously to Case #1, the performance of the second-level meta-models is shown in the two Figures 6.9(a) and 6.9(b) for $\underline{Y}$ and $\overline{Y}$, respectively, using $N_1 = 100$ and $N_2 = \{50, 100, 200, 1000\}$. As expected, the relative generalization error becomes smaller with larger experimental design, *i.e.* larger N_2. Interestingly however, the values are considerably different for $\underline{Y}$ and $\overline{Y}$, due to the shape of the response function $\overline{\mathcal{M}}$ and $\underline{\mathcal{M}}$, respectively.

Figure 6.10 shows the response surface for $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$, respectively, as a function of the auxiliary input variables $\widetilde{\boldsymbol{X}}$. Comparing the two sets of contour lines in Figure 6.10 and the contour lines in Figure 6.4(b), the response surface of $\underline{\mathcal{M}}$ shows a large plateau around the origin and steep value increases around the plateau. This behaviour is difficult to model with polynomials, which results in a lower accuracy seen in Figure 6.9(b). Note that in case the PCE meta-model is calibrated in $\mathcal{D}_{\boldsymbol{C}}$ directly, then the response surfaces are even more complex, as shown in Figure 6.11. Hence, an even lower accuracy of the meta-models can be observed.



(a) Response $\underline{Y}$ (*i.e.* based on $\overline{\mathcal{M}}$) (b) Response $\overline{Y}$ (*i.e.* based on $\underline{\mathcal{M}}$)

Figure 6.9: Rosenbrock function – Case #2 – relative generalization error as a function of N_2 and based on 50 replications of the same analysis ($N_1 = 100$).



(a) $\overline{\mathcal{M}}$ (b) $\underline{\mathcal{M}}$

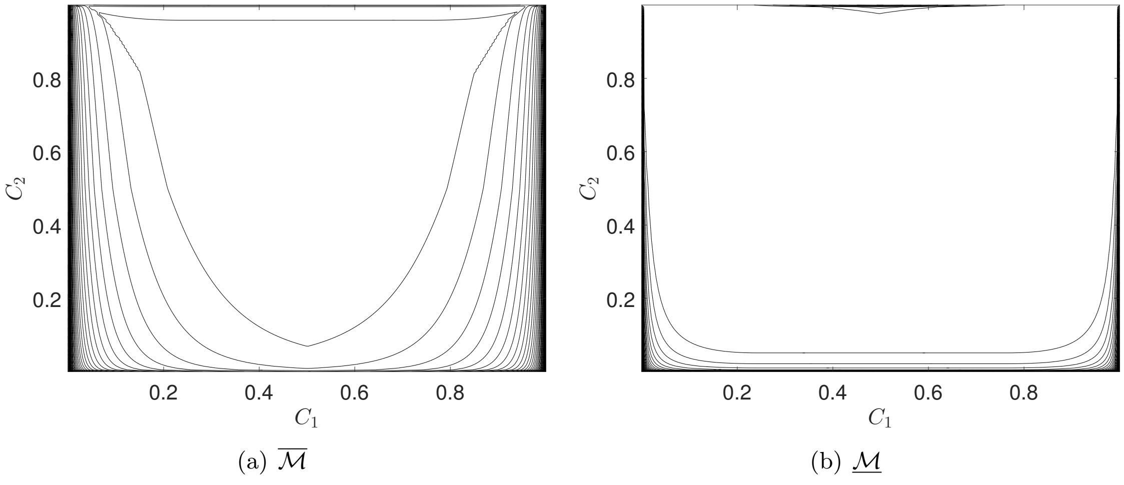
Figure 6.10: Rosenbrock function – shape of the response surfaces of $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$ as a function of $\widetilde{\boldsymbol{X}}$.

Finally, the resulting boundary curves of the p-box of Y are shown in Figure 6.12. The reference solution was computed with the exact computational model on the first level of the algorithm and $n = 10^6$ samples on the second level. The effect of the identified plateau in $\underline{\mathcal{M}}$ in Figure 6.10(b) is that a large quantile of $\overline{Y}$ is close to $y = 0$ (see $\overline{F}_Y$ in Figure 6.12). Additionally, a single realization of those boundary curves is drawn in Figure 6.12 for the case of $N_1 = 100$ and $N_2 = 200$. The corresponding relative generalization error for $\underline{Y}$ and $\overline{Y}$ are $\widehat{err}_{gen}\left[\overline{Y}\right] = 3.43 \cdot 10^{-1}$ and $\widehat{err}_{gen}\left[\underline{Y}\right] = 5.66 \cdot 10^{-3}$, respectively. Despite the relatively large values of the relative generalization error of the second-level meta-model, the bounds of the response p-box are remarkably accurate.

6.3.3 Two-degree-of-freedom damped oscillator

Problem statement

Consider the two-degree-of-freedom (2-dof) damped oscillator subjected to white noise excitation $S(t)$ sketched in Figure 6.13. The subscripts p and s refer to the primary and secondary mass,



(a) $\overline{\mathcal{M}}$ (b) $\underline{\mathcal{M}}$

Figure 6.11: Rosenbrock function – shape of the response surfaces of $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$ as a function of $\boldsymbol{C}$.

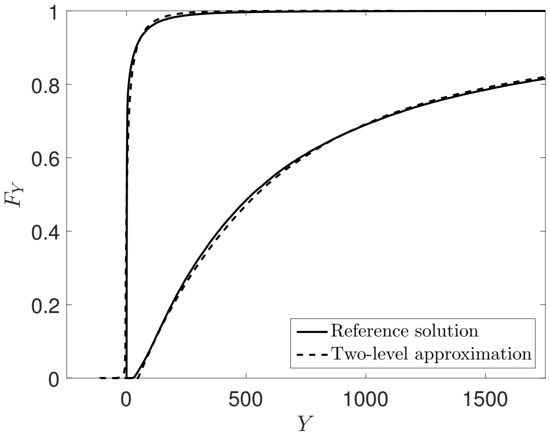


Figure 6.12: Rosenbrock function – Case #2 – response p-boxes $Y \in \left[\underline{Y}, \overline{Y}\right]$: reference solution ($n = 10^6$) versus two-level approach ($N_1 = 100$, $N_2 = 200$).

respectively. The QoI is the force acting on the secondary spring. According to Der Kiureghian & de Stefano (1990) and Dubourg, Sudret, & Deheeger (2013), the peak force P_s in the secondary spring can be computed by:

$$P_s = 3\, k_s \sqrt{\mathbb{E}_S\left[x_s^2\right]}, \tag{6.23}$$

$$\mathbb{E}_S\left[x_s^2\right] = \pi \frac{S_0}{4\zeta_s\omega_s^3} \cdot \frac{\zeta_a\zeta_s}{\zeta_p\zeta_s\left(4\zeta_a^2 + \xi^2\right) + \gamma\zeta_a^2} \cdot \frac{\left(\zeta_p\omega_p^3 + \zeta_s\omega_s^3\right)\omega_p}{4\zeta_a\omega_a^4}, \tag{6.24}$$

where $\omega_p = \sqrt{k_p/m_p}$ and $\omega_s = \sqrt{k_s/m_s}$ are the natural frequencies, $\gamma = m_s/m_p$ is the relative mass, $\omega_a = \left(\omega_p + \omega_s\right)/2$ is the average natural frequency, $\zeta_a = \left(\zeta_p + \zeta_s\right)/2$ is the average damping ratio, $\xi = \left(\omega_p - \omega_s\right)/\omega_a$, m is the mass, k_p and k_s are the spring stiffnesses, and ζ_p and ζ_S are the damping ratios. The input parameters in $\boldsymbol{x} = (m_p, m_s, k_p, k_s, \zeta_p, \zeta_s, S_0)$ are considered statistically independent. The parameters $\{m_p, m_s, k_p, k_s, S_0\}$ are assumed to be well-known (*i.e.* negligible epistemic uncertainty) and defined as independent lognormal variables, with properties summarized in Table 6.1. Each lognormal variable is characterized by a mean value and a coefficient of variation (CoV). Assume that the properties of the damping ratios $\{\zeta_p, \zeta_s\}$ are investigated through a survey among experts because of the highly uncertain nature of damping in dynamic systems. Analogously to Case #1, each expert provides an interval of values for the two damping ratios. Eleven and ten intervals have been collected for ζ_p and ζ_s, respectively. The aggregated p-boxes are shown in Figure 6.14 assuming equal credibility among the experts.

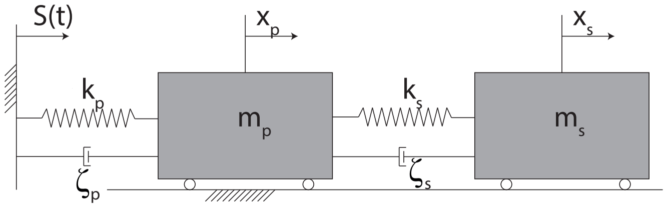


Figure 6.13: Sketch of the 2-dof damped oscillator

Table 6.1: 2-dof damped oscillator – probabilistic input variables.

Variable	Distribution	Mean	CoV
m_p	Lognormal	1.50	10%
m_s	Lognormal	0.01	10%
k_p	Lognormal	1.00	20%
k_s	Lognormal	0.05	20%
S_0	Lognormal	100	10%

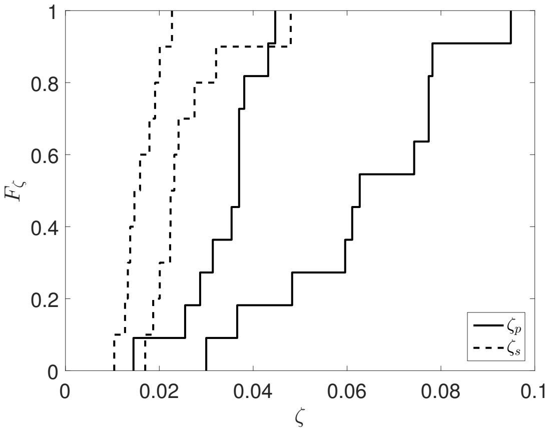


Figure 6.14: 2-dof damped oscillator – aggregated p-boxes of the damping ratio (10 expert opinions of equal credibility).

Analysis

The settings for the two-level meta-modelling approach are kept the same as in the Rosenbrock function example in Section 6.3.2. Sparse PCE is trained with a candidate basis of maximum total polynomial degree equal to 20 and hyperbolic truncation sets with $q = 0.75$. The auxiliary variables $\widetilde{X}_i$ for ζ_p and ζ_s are defined as in Eq. (6.16). In order to analyse the influence of the sample size, N_1 and N_2 are varied in $\{100, 200, 300, \to \infty\}$ and $\{100, 200, 300, 1000\}$, respectively. Note that $N_1 \to \infty$ denotes a case where the first-level meta-model is not applied; instead, the true model $\mathcal{M}$ is used. 50 independent analyses with different experimental designs are conducted to assess the statistical significance of the QoI.

Results

The results for the relative generalization error are summarized in Table 6.2. As expected, increasing N_1 reduces the relative generalization error of the first-level meta-model (see $\widehat{err}_{gen}\left[\widetilde{P}_s\right]$). However, the accuracy of the second-level meta-model is lower than the first-level meta-model, and does not depend significantly on the accuracy of the first-level meta-model. Thus, the error is dominated by the second-level meta-modelling operation. The explanation lies in the free p-boxes of the damping coefficients ζ_p and ζ_s, which have stair-shaped boundary CDF curves. Analogously to Case #1 in the Rosenbrock function (Section 6.3.2), this shape reduces the accuracy of the meta-model on the second level.

Table 6.2: 2-dof damped oscillator – resulting relative generalization error $\widehat{err}_{gen}$ based on a Monte Carlo simulation with $n = 10^5$ samples – mean value of 50 repetitions.

N_1	$\widehat{err}_{gen}\left[\widetilde{P}_s\right]$	N_2	$\widehat{err}_{gen}\left[\underline{P}_s\right]$	$\widehat{err}_{gen}\left[\overline{P}_s\right]$
100	$1.59 \cdot 10^{-2}$	100	$5.85 \cdot 10^{-2}$	$9.46 \cdot 10^{-2}$
		200	$4.59 \cdot 10^{-2}$	$7.32 \cdot 10^{-2}$
		300	$4.18 \cdot 10^{-2}$	$6.76 \cdot 10^{-2}$
		1000	$3.19 \cdot 10^{-2}$	$5.17 \cdot 10^{-2}$
200	$5.24 \cdot 10^{-3}$	100	$5.20 \cdot 10^{-2}$	$7.99 \cdot 10^{-2}$
		200	$3.41 \cdot 10^{-2}$	$6.08 \cdot 10^{-2}$
		300	$2.97 \cdot 10^{-2}$	$5.21 \cdot 10^{-2}$
		1000	$1.87 \cdot 10^{-2}$	$3.57 \cdot 10^{-2}$
300	$2.85 \cdot 10^{-3}$	100	$5.12 \cdot 10^{-2}$	$8.09 \cdot 10^{-2}$
		200	$3.13 \cdot 10^{-2}$	$5.93 \cdot 10^{-2}$
		300	$2.75 \cdot 10^{-2}$	$4.99 \cdot 10^{-2}$
		1000	$1.71 \cdot 10^{-2}$	$3.30 \cdot 10^{-2}$
$\to \infty$	0	100	$5.44 \cdot 10^{-2}$	$8.49 \cdot 10^{-2}$
		200	$3.07 \cdot 10^{-2}$	$6.11 \cdot 10^{-2}$
		300	$2.71 \cdot 10^{-2}$	$4.96 \cdot 10^{-2}$
		1000	$1.52 \cdot 10^{-2}$	$3.07 \cdot 10^{-2}$

Although the two-level meta-models are only moderately accurate, the bounds of the resulting free p-box are computed accurately, as seen in Figure 6.15. A reference solution based on $n = 10^5$ Monte Carlo samples is compared to the two-level meta-modelling approach with $N_1 = 300$ and $N_2 = 100$. The boundary curves ($\underline{F}_{P_s}$ and $\overline{F}_{P_s}$) obtained by using the proposed approach are almost superimposed with the reference ones.

Comparison to slicing algorithm

In order to illustrate the curse of dimensionality, the slicing algorithm is applied on the second level of the two-level approach replacing the calibration of the meta-models $\underline{\mathcal{M}}^{(\mathrm{P})}$ and $\overline{\mathcal{M}}^{(\mathrm{P})}$. Note that in order to apply the slicing algorithm consistently, each probabilistic input variable is interpreted as free p-box too. These continuous random variables are bounded within their 1% and 99% quantiles in order to obtain finite intervals for the interval analysis. The number of intervals is chosen as $n_{X_i} = 3$ in order to obtain a similar number of optimization operations as

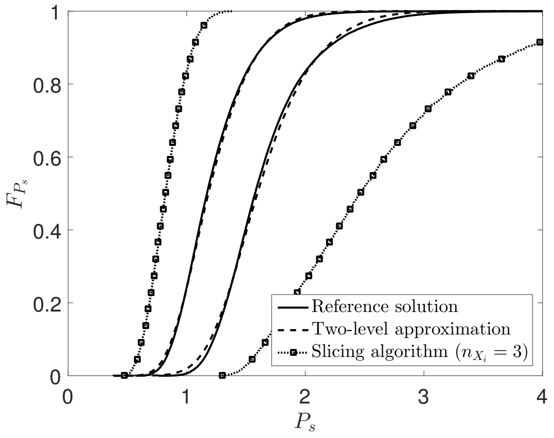


Figure 6.15: 2-dof damped oscillator – reference solution ($n = 10^5$) versus two-level approximation approach ($N_1 = 300$, $N_2 = 100$, $\widehat{err}_{gen}\left[\widetilde{P}_s\right] = 2.57 \cdot 10^{-3}$, $\widehat{err}_{gen}\left[\overline{P}_s\right] = 7.50 \cdot 10^{-2}$, $\widehat{err}_{gen}\left[\underline{P}_s\right] = 3.89 \cdot 10^{-2}$) and slicing algorithm with $n_{X_i} = 3$.

the proposed two-level approach, *i.e.* $|\mathcal{K}| = 3^7 = 2187$. Figure 6.16 illustrates the discretization for m_p and ζ_p. The boundary curves of the discretized free p-boxes are wider than the original input p-boxes. This effect is pronounced for the probabilistic input variables, as seen in the left part of Figure 6.16.

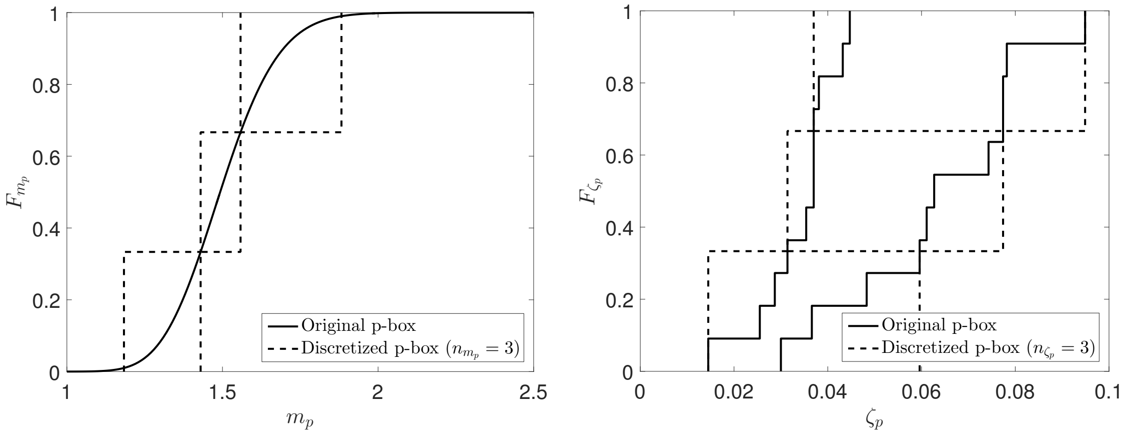


Figure 6.16: 2-dof damped oscillator – discretization of m_p and ζ_p with $n_{X_i} = 3$.

The discretization algorithm leads to conservative approximations of the true input free p-boxes. The propagation of the discretized intervals results in a free p-box that is much wider (*i.e.* more conservative) compared to the proposed two-level approach, as indicated in Figure 6.15, despite the fact that more model evaluations were computed. This would be hardly interpretable in practice. The conservative estimate of the response free p-box originates mainly in the coarse approximation of the five probabilistic input variables.

6.3.4 Two-dimensional truss structure

Problem statement

The third application example is the two-dimensional truss structure described in Section 5.4.3. Analogously to Case #2(b), the input variables are modelled by free p-boxes defined by the envelope of a set of curves (see Eq. (6.3)). The distribution function and its interval-valued parameters are summarized in Table 6.3. The variables are assumed to be statistically independent.

Table 6.3: Two-dimensional truss – input p-boxes.

Variable		Distribution	Mean	CoV
A_1	[m^2]	Lognormal	$[1.9, 2.1] \cdot 10^{-3}$	[8, 12]%
A_2	[m^2]	Lognormal	$[0.9, 1.1] \cdot 10^{-3}$	[8, 12]%
E_1, E_2	[Pa]	Lognormal	$[2.0, 2.2] \cdot 10^{11}$	[9, 11]%
$P_1, \ldots, P_6$	[N]	Gumbel	$[4.0, 6.0] \cdot 10^{4}$	[10, 15]%

The QoI is the deflection u at mid-span as a function of the loading and the material parameters. The arrow in Figure 5.8 indicates positive values for u.

Analysis

For the truss structure, the auxiliary variable $\widetilde{X}_i$ is defined as follows. The mean value $\mu_{\widetilde{X}_i}$ is the mid-range of the mean value of X_i, and the coefficient of variation $CoV_{\widetilde{X}_i}$ is set to the maximum value of the coefficient of variation of X_i. The auxiliary variables are used on both levels of meta-models to ensure a good convergence behaviour.

Similar to the previous examples, the number of samples is varied in both levels of meta-models, *i.e.* $N_1 = \{100, 300, \rightarrow \infty\}$ and $N_2 = \{100, 300\}$. The sparse PCE meta-models are built with Hermite polynomials with a candidate basis of maximal total degree of 20 and hyperbolic truncation scheme with $q = 0.75$. To ensure statistical significance of the results, 50 independent runs of the analysis are performed with different LHS experimental designs.

Results

A summary of the relative generalization error is given in Table 6.4. As expected, increasing the number of samples in the experimental design decreases the error of the meta-model on both meta-modelling levels. In particular, the influence of a larger experimental design on the first-level meta-model is visible. For the second-level meta-model, the absolute value of the relative generalization error is larger than for the first level due to the larger complexity of the analysis. In fact, the accuracy of $\underline{u}$ and $\overline{u}$ depends on the quality of (i) the first-level meta-model, (ii) the optimization algorithm, and (iii) the second-level meta-model.

Table 6.4: Two-dimensional truss – resulting relative generalization error based on a Monte Carlo simulation with $n = 10^5$ samples – mean value of 50 repetitions.

N_1	$\widehat{err}_{gen}\,[\widetilde{u}]$	N_2	$\widehat{err}_{gen}\,[\underline{u}]$	$\widehat{err}_{gen}\,[\overline{u}]$
100	$1.96 \cdot 10^{-3}$	100	$8.30 \cdot 10^{-2}$	$4.85 \cdot 10^{-2}$
		300	$2.39 \cdot 10^{-2}$	$2.50 \cdot 10^{-2}$
300	$1.44 \cdot 10^{-4}$	100	$9.10 \cdot 10^{-2}$	$4.66 \cdot 10^{-2}$
		300	$2.09 \cdot 10^{-2}$	$2.01 \cdot 10^{-2}$
$\rightarrow \infty$	0	100	$1.24 \cdot 10^{-2}$	$2.18 \cdot 10^{-2}$
		300	$7.05 \cdot 10^{-3}$	$1.65 \cdot 10^{-2}$

Another aspect of the modelling is the shape of the response p-box, which is shown in Figure 6.17. The reference solution is obtained by a Monte Carlo simulation ($n = 10^5$) and by taking advantage of the monotonicity of the truss model. A two-level approximation with $N_1 = 300$ is able to reproduce the boundary curves of the response p-box accurately: the two curves for $\overline{u}$ coincide, whereas the two curves for $\underline{u}$ are remarkably close to each other, too.

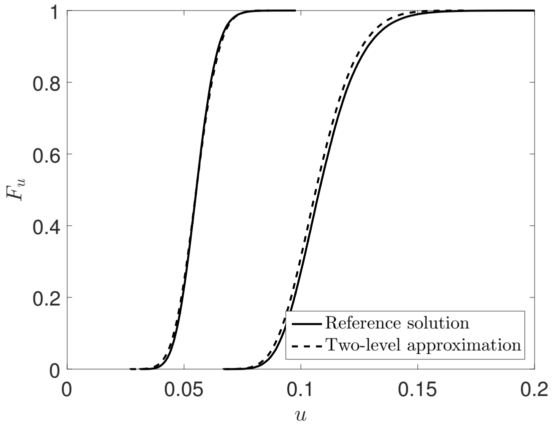


Figure 6.17: Two-dimensional truss – free p-box of the deflection u: reference solution ($n = 10^5$) versus two-level meta-modelling approach ($N_1 = 300$, $N_2 = 300$, $\widehat{err}_{gen}\left[\widetilde{u}\right] = 9.82 \cdot 10^{-5}$, $\widehat{err}_{gen}\left[\overline{u}\right] = 3.70 \cdot 10^{-2}$, and $\widehat{err}_{gen}\left[\underline{u}\right] = 2.64 \cdot 10^{-2}$).

6.4 Parametric p-boxes

6.4.1 Nested Monte Carlo simulation

The hierarchical setup of parametric p-boxes allows for a nested uncertainty propagation approach. A straightforward sampling-based algorithm is the *nested Monte Carlo simulation* (nMCS) (Eldred & Swiler 2009; Chowdhary & Dupuis 2013; Schöbi & Sudret 2015c), which is illustrated in Figure 6.18. Consider the input vector $\boldsymbol{X}$ described by M independent parametric p-boxes, which are characterized by a conditional CDF $F_{\boldsymbol{X}} = F_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$ and $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$. Note that $\boldsymbol{\theta}_i$ denotes the vector of epistemic parameters (*i.e.* intervals) describing variable X_i, whereas $\boldsymbol{\Theta} \overset{\text{def}}{=} \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M\}$ gathers those parameters for all variables in $\boldsymbol{X}$. Then, in the outer loop of nMCS, realizations of the distribution parameters $\boldsymbol{\Theta}$ are generated. In the inner loop, a Monte Carlo simulation is conducted based on the conditional CDF $F_{\boldsymbol{X}}\left(\boldsymbol{x}\middle|\boldsymbol{\theta}^{(k)}\right)$ and the computational model $\mathcal{M}$. Each realization $\boldsymbol{\theta}^{(k)}$, hence, results in a realization of the response CDF $F_Y\left(y\middle|\boldsymbol{\theta}^{(k)}\right)$. Finally, the boundary curves of the response p-box are obtained by:

$$\underline{F}_Y(y) = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} F_Y\left(y|\boldsymbol{\theta}\right), \qquad \overline{F}_Y(y) = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} F_Y\left(y|\boldsymbol{\theta}\right). \tag{6.25}$$

The nMCS provides a robust but inefficient tool to estimate the boundaries of the response p-box. The inefficiency arises from the large number of model evaluations due to the nested Monte Carlo simulations. In case of an expensive-to-evaluate computational model $\mathcal{M}$, nMCS becomes intractable. Therefore, it is proposed to replace the computational model by a PCE model as described in the following sections.

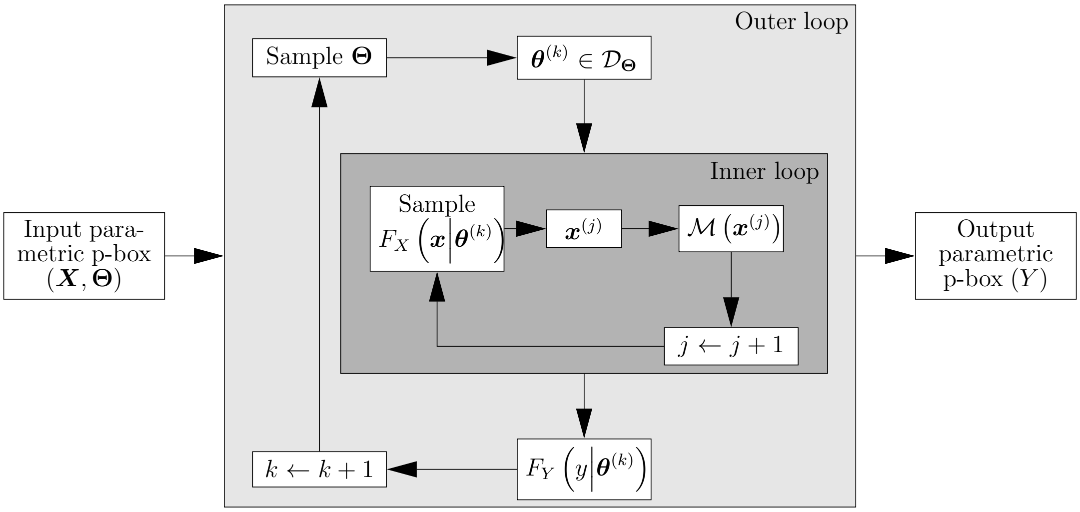


Figure 6.18: Nested Monte Carlo simulation.

6.4.2 Augmented PCE

Augmented input space

Due to the clear separation of aleatory and epistemic uncertainty in the formulation of parametric p-boxes, those sources of uncertainty can be treated as separate entities, *i.e.* $\boldsymbol{X}$ and $\boldsymbol{\Theta}$, as seen previously in Section 2.4 for a single variable X. Hence, the QoI Y can be interpreted as a function of the augmented input vector $(\boldsymbol{X}, \boldsymbol{\Theta})$. The corresponding map then reads:

$$W = \mathcal{M}^{(\text{aug})}\left(\boldsymbol{X}, \boldsymbol{\Theta}\right), \tag{6.26}$$

which is based on the original computational model as $\mathcal{M}^{(\text{aug})}\left(\boldsymbol{x}, \boldsymbol{\theta}\right) \equiv \mathcal{M}\left(\boldsymbol{x}\right)$.

The components of the augmented input vector are dependent on each other. In particular, $\boldsymbol{X}$ depends on $\boldsymbol{\Theta}$ as a result of the hierarchical formulation of the parametric p-box. The dependencies are visualized in Figure 6.19(a). In this thesis, however, the input variables to a computational model are assumed to be independent. In order to obtain independent variables in the augmented input vector, an isoprobabilistic transform is used which maps $(\boldsymbol{X}, \boldsymbol{\Theta})$ to $\boldsymbol{V}$, which shall be a vector with independent components. Assuming that the components X_i are independent and that the parameter intervals in $\boldsymbol{\Theta}$ are independent too (*i.e.* $\mathcal{D}_{\boldsymbol{\Theta}}$ is a hyper-rectangular domain), the parametric p-box $\boldsymbol{X}$ can be written as:

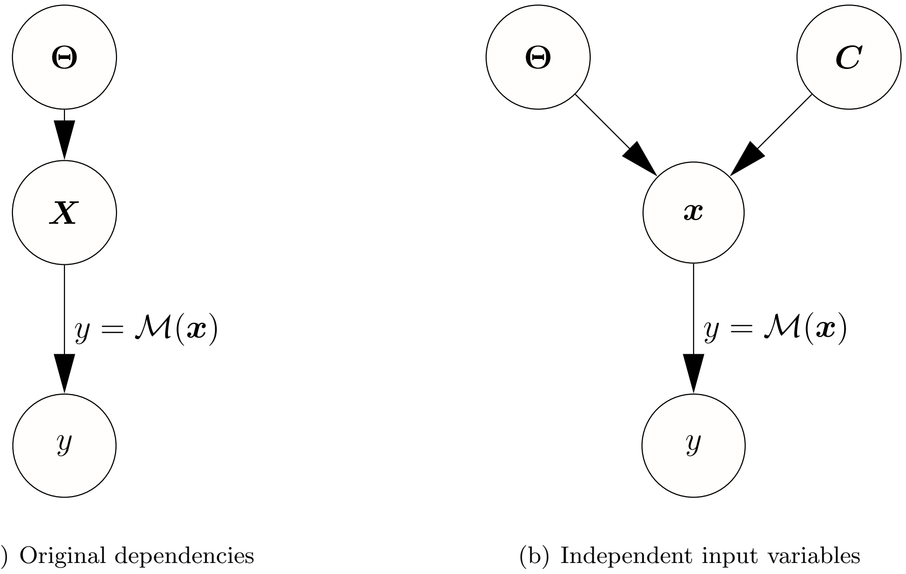
$$X_i = F_{X_i}^{-1}\left(C_i | \boldsymbol{\theta}_i\right), \quad i = 1, \ldots, M. \tag{6.27}$$

where $\boldsymbol{C} = (C_1, \ldots, C_M)$ is a vector of independent uniform distributions and where $C_i \sim \mathcal{U}(0,1)$ describes the CDF value of X_i. Then, the augmented vector $\boldsymbol{V} = (\boldsymbol{C}, \boldsymbol{\Theta})$ is a vector with independent components.

The Bayesian network corresponding to this model is visualized in Figure 6.19(b). Note that in the terminology of the Bayesian network, $\boldsymbol{x}$ is used in lower case due to the deterministic transform of $\boldsymbol{V} = (\boldsymbol{C}, \boldsymbol{\Theta}) \mapsto \boldsymbol{X}$. In Figure 6.19(a), the transform of $\boldsymbol{\Theta}$ to $\boldsymbol{X}$ is not deterministic, hence, the upper case $\boldsymbol{X}$.

Special cases

A special case is when the distribution family is Gaussian, *i.e.* $X \sim \mathcal{N}\left(\mu_X, \sigma_X\right)$ and $\boldsymbol{\theta} = (\mu_X, \sigma_X)$, and the isoprobabilistic transform in Eq. (6.27) may be replaced by a simpler con-



(a) Original dependencies (b) Independent input variables

Figure 6.19: Parametric p-box – dependency structures.

struction. Consider a parametric p-box where the mean value μ_X and standard deviation σ_X are given in intervals. Instead of using a uniform distribution C, a standard Gaussian variable $\zeta \sim \mathcal{N}(0, 1)$ can be used in the independent input vector $\boldsymbol{V} = (\zeta, \mu_X, \sigma_X)$. Then, the isoprobabilistic transform reads:

$$X = \mu_X + \sigma_X \cdot \zeta. \tag{6.28}$$

This transform is simpler than the inverse CDF formulation in Eq. (6.27) because the unbounded variable X is modelled by the unbounded ζ. A similar construction is possible for the lognormal distribution, due to the fact that $\log X \sim \mathcal{N}$ follows a normal distribution if $X \sim \mathcal{LN}$ follows a lognormal distribution.

Another special case of interest is the Gumbel distribution whose CDF reads:

$$F_{\mathcal{GU}}(x|\alpha, \beta) = \exp\left[-\exp\left(-(x-\alpha)/\beta\right)\right],$$

where $\alpha = \mu_X - \beta\gamma_e$ and $\beta = \sigma_X \sqrt{6}/\pi$ are its distribution hyperparameters and $\gamma_e = 0.5772\ldots$ is the Euler constant. Then, denote by $\zeta \sim \mathcal{GU}(\alpha = 0, \beta = 1)$ a standard Gumbel variable. It follows that a Gumbel distribution with arbitrary hyperparameters $\{\alpha, \beta\}$ can be formulated in terms of a standard Gumbel variable:

$$\begin{aligned} F_{\mathcal{GU}}(x|\alpha, \beta) &= F_{\mathcal{GU}}(\zeta|0, 1) \\ \exp\left[-\exp\left(-(x-\alpha)/\beta\right)\right] &= \exp\left[-\exp\left(-\zeta\right)\right] \\ (x-\alpha)/\beta &= \zeta. \end{aligned}$$

Hence, a standard Gumbel distribution can be transformed into an arbitrary Gumbel distribution by:

$$X = \alpha + \beta \cdot \zeta, \tag{6.29}$$

which resembles the construction for Gaussian variables in Eq. (6.28) with different transformation parameters. Note that there are potentially other distribution families where such a simplification might be feasible. However, they are not discussed in this thesis.

Meta-model

The augmented computational model $\mathcal{M}^{(\text{aug})}$ can be formulated as a function of vector $\boldsymbol{V}$:

$$W = \mathcal{M}^{(\text{aug})}\left(T\left(\boldsymbol{V}\right)\right), \tag{6.30}$$

where T denotes the isoprobabilistic transform from $\boldsymbol{V}$ to $(\boldsymbol{X}, \boldsymbol{\Theta})$. When interpreting the epistemic intervals in $\boldsymbol{\Theta}$ as uniform distributions, a sparse PCE model can be calibrated to meta-model the augmented computational model:

$$W \approx \mathcal{M}^{(\text{PCE})}\left(\boldsymbol{V}\right) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \; \psi_{\boldsymbol{\alpha}}\left(\boldsymbol{V}\right), \tag{6.31}$$

where the dimensionality is $M = |\boldsymbol{V}| = |\boldsymbol{C}| + |\boldsymbol{\Theta}| \equiv |\boldsymbol{X}| + |\boldsymbol{\Theta}|$. The meta-model is trained with an experimental design $\mathfrak{V} = \left\{\mathfrak{v}^{(1)}, \dots, \mathfrak{v}^{(N)}\right\}$ and the corresponding responses:

$$\mathfrak{W} = \left\{\mathfrak{W}^{(1)} = \mathcal{M}^{(\text{aug})}\left(T\left(\mathfrak{v}^{(1)}\right)\right), \dots, \mathfrak{W}^{(N)} = \mathcal{M}^{(\text{aug})}\left(T\left(\mathfrak{v}^{(N)}\right)\right)\right\},$$

by using Eq. (6.30). In order to efficiently calibrate a sparse PCE meta-model, the reader is referred to Section 3.3 for further details. Note that in this setup, the input vector $\boldsymbol{V}$ is sampled in the same way for the $\boldsymbol{C}$- and the $\boldsymbol{\Theta}$-components.

6.4.3 Phantom points

Dependencies in the experimental design

The number of dimensions in the augmented input space $\mathcal{D}_{\boldsymbol{V}}$ is larger than the one of the original input space $\mathcal{D}_{\boldsymbol{X}}$ due to the epistemic uncertainty. Hence, Eq. (6.27) leads to an interesting feature of the augmented PCE. For a given realization of $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}}$, there is an infinite number of realizations $\boldsymbol{v}$ that satisfy Eq. (6.27). In other words, the inverse operation of Eq. (6.27) is non-unique.

For illustrative purpose, consider a parametric p-box defined by a Gaussian distribution family with interval-valued mean and standard deviation. As an example, it is assumed that $\mu_X \in [-1, 1]$ and $\sigma_X \in [0.5, 1.0]$. Then for a given x_0, the manifold defined by the triplet

$$\left\{\mu_X \in [-1,1],\, \sigma_X \in [0.5,\, 1.0],\, c \in [0,1] : \, F_X^{-1}\left(x_0 | \mu_X, \sigma_X\right) = c\right\}$$

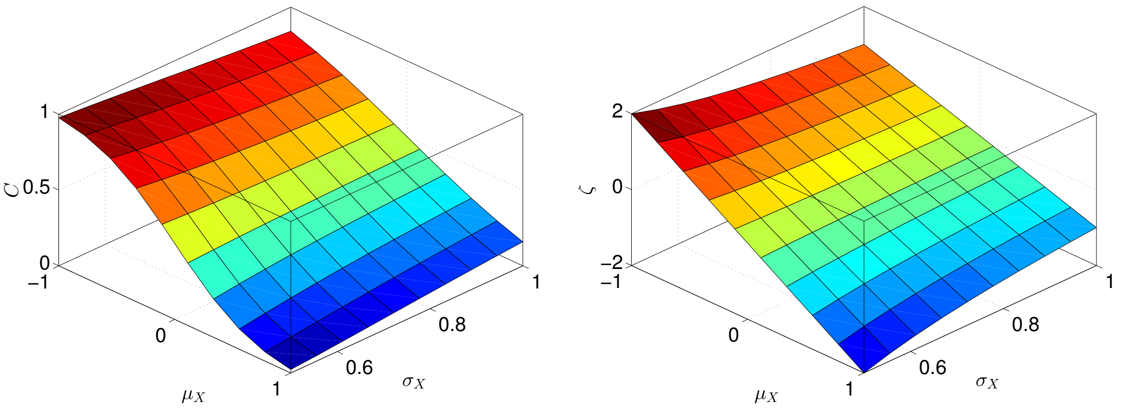
leads to the same x_0 through the isoprobabilistic transform. Similarly for the special case of a Gaussian variable, the manifold is defined as

$$\left\{\mu_X \in [-1,1],\, \sigma_X \in [0.5,\, 1.0],\, \zeta \in \mathbb{R} : \, \mu_X + \sigma_X \zeta = x_0\right\}.$$

Figure 6.20 illustrates the resulting manifolds in the μ_X-σ_X-C and μ_X-σ_X-ζ domain for $x_0 = 0$. Note that they are limited in μ_X and σ_X by the corresponding interval values, wheres the extent in the ζ dimension is infinite. In terms of C, the manifold is naturally bounded in $[0, 1]$.

Enrichment of the experimental design

Consider again the experimental design $\mathfrak{V}$. Each sample $\mathfrak{v}^{(j)}$ can be transformed into an equivalent $\boldsymbol{\chi}^{(j)}$ using Eq. (6.27). A number of realizations $\mathfrak{v} = (\mathfrak{c}, \boldsymbol{\tau})$ can be generated by sampling $\boldsymbol{\tau} \in \mathcal{D}_{\boldsymbol{\Theta}}$ and inserting them into the inverse isoprobabilistic transform to obtain realizations $\mathfrak{c}$.



(a) Inverse CDF formulation in Eq. (6.27)

(b) Special case formulation in Eq. (6.28)

Figure 6.20: Gaussian parametric p-box – manifolds corresponding to a particular x_0.

In particular, for each dimension $i = 1, \dots, M$ and experimental design sample $j = 1, \dots, N$, n_{ph} samples can be generated from $\boldsymbol{\theta}_i \in \mathcal{D}_{\boldsymbol{\Theta}_i}$ and collected in $\left\{\boldsymbol{\tau}_i^{(j)(k)}, k = 1, \dots, n_{ph}\right\}$. Then, the inverse transform consists of evaluating simple CDFs and reads:

$$\mathfrak{c}_i^{(j)(k)} = F_{X_i}\left(\chi_i^{(j)} \middle| \boldsymbol{\tau}_i^{(j)(k)}\right), \tag{6.32}$$

where each experimental design sample point is $\boldsymbol{\chi}^{(j)} = \left(\chi_1^{(j)}, \dots, \chi_M^{(j)}\right)$. The new components in dimension i of vector $\mathfrak{v}^{(j)}$ reads then:

$$\left\{\mathfrak{v}_i^{(j)(k)} = \left(\mathfrak{c}_i^{(j)(k)}, \boldsymbol{\tau}_i^{(j)(k)}\right),\ k = 1, \dots, n_{ph}\right\}. \tag{6.33}$$

When pursuing this procedure for all $i = 1, \dots, M$ dimensions of the input vector, new samples of the experimental design are generated. These samples are called *phantom points* (Schöbi & Sudret 2015c) because they contribute to the experimental design $\mathfrak{V}$ without increasing the number of model evaluations $\mathcal{M}$. Indeed, the n_{ph} points in Eq. (6.33) all correspond to a *single* $\boldsymbol{\chi}^{(j)}$ in the original experimental design, meaning a *single run* $\mathcal{Y}^{(j)} = \mathcal{M}\left(\boldsymbol{\chi}^{(j)}\right)$. This has a large effect on the global efficiency when considering that model evaluations are dominating the total computational costs. The experimental design with phantom points and the corresponding responses then read:

$$\begin{cases} \mathfrak{V} = \left\{\mathfrak{v}^{(j)(k)}, \quad j = 1, \dots, N, \quad k = 1, \dots, n_{ph}\right\} \\ \\ \mathfrak{W} = \left\{\mathfrak{W}^{(j)(1)} = \ldots = \mathfrak{W}^{(j)(n_{ph})}, \quad j = 1, \dots, N\right\} \end{cases}, \tag{6.34}$$

which is obtained at a cost of exactly N evaluations of the computational model $\mathcal{M}$.

Consider again the Gaussian distribution with interval-valued distribution parameters and the isoprobabilistic transform in Eq. (6.28). Further, consider a single realization of the augmented input vector, denoted by $\left\{\mu_X^{(1)}, \sigma_X^{(1)}, \zeta^{(1)}\right\}$, and the corresponding $\chi^{(1)} = \mu_X^{(1)} + \sigma_X^{(1)} \cdot \zeta^{(1)}$. A set of realizations of the distribution parameters is generated and denoted by $\left\{\mu_X^{(1)(k)}, \sigma_X^{(1)(k)}\right\}$, $k = 1, \dots, n_{ph}$. Then, the equation corresponding to Eq. (6.32) reads:

$$\zeta^{(1)(k)} = \frac{\chi^{(1)} - \mu_X^{(1)(k)}}{\sigma_X^{(1)(k)}}, \tag{6.35}$$

which leads to the following experimental design enriched by phantom points and the corresponding responses:

$$\begin{cases} \mathfrak{V} = \left\{ \left\{ \mu_X^{(1)(k)}, \sigma_X^{(1)(k)}, \zeta^{(1)(k)} \right\},\ k = 1, \dots, n_{ph} \right\} \\ \\ \mathfrak{W} = \left\{ \mathfrak{W}^{(1)(1)} = \dots = \mathfrak{W}^{(1)(n_{ph})} = \mathcal{M}\left(\chi^{(1)}\right) = \mathcal{Y}^{(1)} \right\} \end{cases} .$$

6.5 Numerical examples for parametric p-boxes

6.5.1 Rosenbrock function

Problem statement

Consider the Rosenbrock function defined in Eq. (6.22). The input vector is modelled by parametric p-boxes. Each input variable is defined by a Gaussian distribution function with interval-valued mean value and standard deviation: $\mu_{X_i} \in [-0.5, 0.5]$ and $\sigma_{X_i} \in [0.7, 1.0]$. Its boundary curves are plotted in Figure 6.5(b).

Analysis

Two cases are considered for parametric p-boxes in combination with augmented PCE models. The first case consists of the general formulation for the isoprobabilistic transform in Eq. (6.27) and the corresponding input vector $\boldsymbol{V} = (\boldsymbol{C}, \boldsymbol{\Theta})$. The second case consists of the special formulation in Eq. (6.28) and the corresponding input vector $\boldsymbol{V} = (\boldsymbol{\zeta}, \boldsymbol{\Theta})$.

The sparse PCE models are calibrated with a maximal polynomial degree of $p = 10$ and the hyperbolic index set characterized by $q = 1$, due to the expected large number of interaction polynomials. The phantom points are obtained by Eq. (6.35) and by Latin-hypercube sampling. The same analysis is replicated 50 times with different experimental designs to achieve statistical significance in the results.

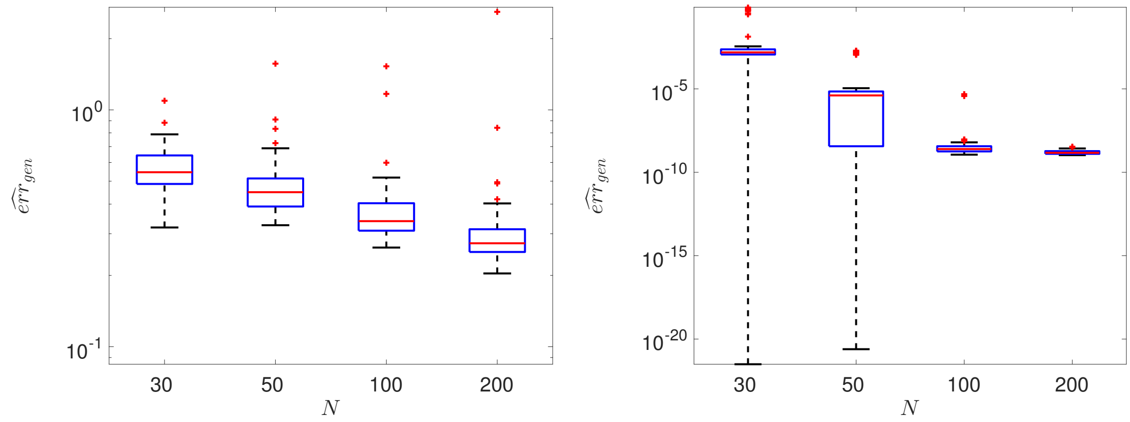
Results

Figure 6.21 shows the boxplots for a varying experimental design size N and a constant value of $n_{ph} = 50$ phantom points. As expected, the larger the experimental design, the smaller the relative generalization error, which is estimated based on $n = 10^5$ Monte Carlo samples in this example.

The two formulations of the augmented input vector $\boldsymbol{V}$, however, behave differently. The general formulation in Eq. (6.27) leads to a slower convergence behaviour than the formulation in Eq. (6.28). The high accuracy of the meta-models in the case of $\boldsymbol{V} = (\boldsymbol{\zeta}, \boldsymbol{\Theta})$ can be explained by reformulating the computational model. By substituting the input vector $\boldsymbol{X}$ with $(\boldsymbol{\zeta}, \boldsymbol{\Theta})$ and the corresponding isoprobabilistic transform, the augmented computational model read:

$$w = f_3^{(\mathrm{aug})}(\boldsymbol{v}) = 100 \left((\mu_{X_2} + \sigma_{X_2} \cdot \zeta_2) - (\mu_{X_1} + \sigma_{X_1} \cdot \zeta_1)^2 \right)^2 + (1 - (\mu_{X_1} + \sigma_{X_1} \cdot \zeta_1))^2 , \quad (6.36)$$

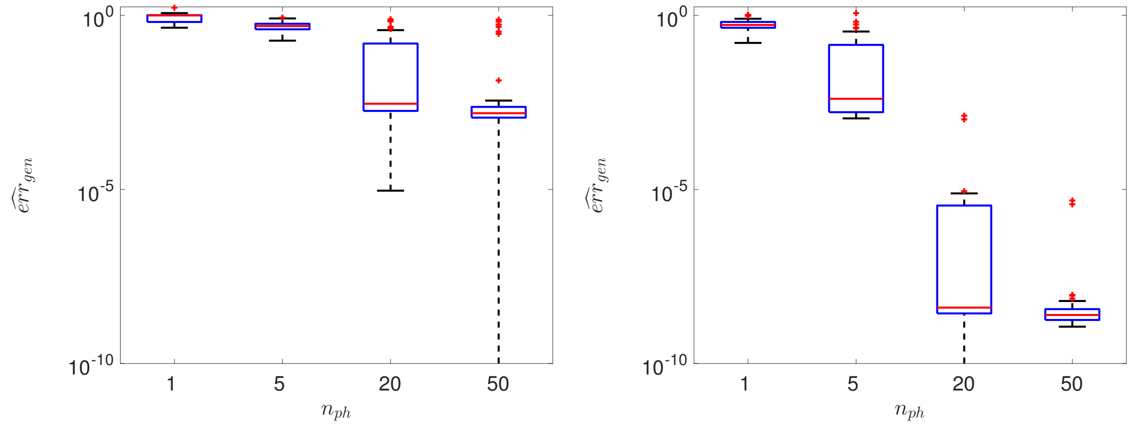
where $\zeta_1 \sim \mathcal{N}(0,1)$ and $\zeta_2 \sim \mathcal{N}(0,1)$. This function is a smooth polynomial function with a total degree of $p = 8$. Note that the number of interaction terms is large. However, this model can be approximated well by PCE as seen in Figure 6.21(b). With an experimental design of $N = 100$ samples, the average relative generalization error is in the order of $\widehat{err}_{gen} \approx 10^{-9}$.



(a) Augmented input vector $\boldsymbol{V} = (\boldsymbol{C}, \boldsymbol{\Theta})$ (b) Augmented input vector $\boldsymbol{V} = (\boldsymbol{\zeta}, \boldsymbol{\Theta})$

Figure 6.21: Rosenbrock function – relative generalization error as a function of N and the definition of the augmented space ($n_{ph} = 50$).

Focusing on the case of $\boldsymbol{V} = (\boldsymbol{\zeta}, \boldsymbol{\Theta})$, the influence of the number of phantom points n_{ph} is illustrated in Figure 6.22. The two figures show nicely that the larger the number of phantom points, the more accurate is the meta-model in the augmented input space. When no phantom points are used (*i.e.* $n_{ph} = 1$), the relative generalization error is of order $\widehat{err}_{gen} \approx \left[10^{-1}, 10^{0}\right]$, which depicts a poor meta-model. On the right hand side of Figure 6.22(b), the relative generalization error is of order $\widehat{err}_{gen} \approx 10^{-9}$ for $n_{ph} = 50$, which depicts an accurate meta-model at the same number of computational model runs N.



(a) Small $N = 30$ (b) Large $N = 100$

Figure 6.22: Rosenbrock function – relative generalization error as a function of n_{ph} ($\boldsymbol{V} = (\boldsymbol{\zeta}, \boldsymbol{\Theta})$).

Comparison to free p-boxes

In Section 6.3.2, the Rosenbrock function is discussed in the context of free p-boxes. The p-box boundaries of the input variables X_i are the same in the two examples, except a free (resp.a parametric) structure is assumed in the latter (resp. present) example.

Both sections result in p-boxes for the response variable Y, the boundary curves of which are shown in Figure 6.23. The free p-box formulation leads to wider boundary curves than the parametric p-box formulation. This result is confirmed by the intuition that the free p-box approach is more general than the parametric p-box approach and models a larger amount of epistemic uncertainty (basically the absence of knowledge on the shape of the CDF).

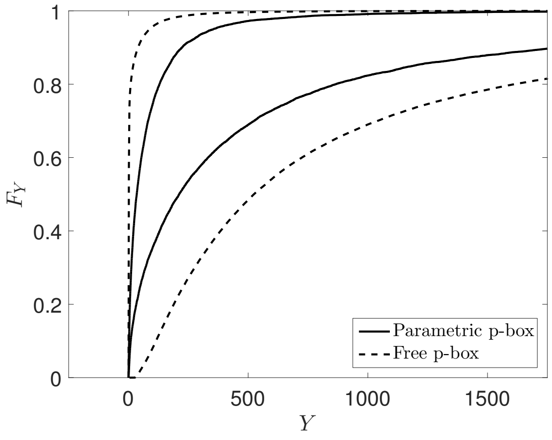


Figure 6.23: Rosenbrock function – response parametric versus free p-box.

6.5.2 Two-dimensional truss structure

Problem statement

Consider again the two-dimensional truss structure described in Sections 5.4.3 and 6.3.4. The parametric p-boxes in the input vector $\boldsymbol{X}$ are defined as distribution families with interval-valued mean value and coefficient of variation, as described in Table 6.3.

Analysis

The augmented input vector in this application example is chosen to be a mix of interval-valued epistemic parameters and standard Gaussian distributions, *i.e.* $\boldsymbol{V} = (\boldsymbol{\zeta}, \boldsymbol{\Theta})$. Note that the vector $\boldsymbol{\theta}_i$ consists of a mean value μ_X and coefficient of variation η_X, as seen in Table 6.3. The standard Gaussian distributions $\boldsymbol{\zeta}$ represent the aleatory uncertainty of the input vector. The Gaussian distribution is chosen because the input random variables are unbounded and hence the resulting isoprobabilistic transform (see Eq. (6.27)) is nicely behaving.

The sparse PCE models are calibrated with a maximal polynomial degree of $p = 10$ and a hyperbolic index set with $q = 1$, due to the expected large number of interaction polynomials. The phantom points are generated by Latin-hypercube sampling technique. In particular, the samples are generated in the domain of $\boldsymbol{\mu_X}$ and $\boldsymbol{\eta_X}$.

For lognormal variables (denoted by $\mathcal{LN}$), the phantom points are generated as follows. Analogously to Eq. (6.28), the isoprobabilistic transform reads for the logarithm of x:

$$\ln x = \lambda_X + \varpi_X \cdot \zeta_X, \tag{6.37}$$

where the mean value and standard deviation of $\ln x$ are defined as:

$$\lambda_X\left(\mu_X, \eta\right) = \ln \mu_X - \varpi_X^2/2 = \ln \frac{\mu_X}{\sqrt{1+\eta_X^2}},$$

$$\varpi_X\left(\mu_X, \eta_X\right) = \sqrt{\ln\left(1+\eta_X^2\right)},$$

and where ζ_X is a standard Gaussian variable. Then, in the context of phantom points, Eq. (6.35) transforms to:

$$\zeta_X^{(j)(k)} = \frac{\ln \chi^{(j)} - \lambda_X\left(\mu_X^{(j)(k)}, \eta_X^{(j)(k)}\right)}{\varpi_X\left(\mu_X^{(j)(k)}, \eta_X^{(j)(k)}\right)}, \tag{6.38}$$

where $k = 1, \ldots, n_{ph}$. Note that the input parametric p-boxes are defined in terms of interval-valued coefficient of variation rather than standard deviation. Hence, the experimental design consist of $\tau^{(j)(k)} = \left\{\mu_X^{(j)(k)}, \eta_X^{(j)(k)}\right\}$ and the corresponding standard deviation reads $\sigma_X^{(j)(k)} = \eta_X^{(j)(k)} \cdot \mu_X^{(j)(k)}$.

For the Gumbel distributions (denoted by $\mathcal{GU}$) with interval-valued mean value and coefficient of variation, the phantom points are computed directly with CDF and inverse CDF:

$$\zeta_X^{(j)(k)} = F_{\mathcal{N}}^{-1}\left(F_{\mathcal{GU}}\left(\chi^{(j)}\middle|\mu_X^{(j)(k)}, \eta_X^{(j)(k)}\right)\right), \tag{6.39}$$

where $F_{\mathcal{GU}}$ denotes the CDF of a Gumbel distribution, defined by:

$$F_{\mathcal{GU}}(x) = \exp\left[-\exp\left(-(x-\alpha)/\beta\right)\right], \tag{6.40}$$

where $\alpha = \mu_X - \beta\gamma_e$ and $\beta = \sigma_X\sqrt{6}/\pi$ are the distribution parameters and $\gamma_e = 0.577216\ldots$ is the Euler constant. Note that Eq. (6.29) could also be used in this example to simplify the isoprobabilistic transform.

Results

Table 6.5 summarizes the relative generalization error of the augmented PCE model of a single run of the analysis. In particular, the dependency of the relative generalization error on the number of runs of the computational model N and the number of phantom points n_{ph}. As expected, the relative generalization error decreases, when N increases as well as when n_{ph} increases.

Table 6.5: Two-dimensional truss structure – relative generalization error $\widehat{err}_{gen}[W]$ for varying N and n_{ph} and relative generalization error for a conventional PCE model on the auxiliary input vector $\widetilde{\boldsymbol{X}}$ for comparison.

	$N = 100$	$N = 200$	$N = 500$	$N = 1000$
$n_{ph} = 1$	$4.07 \cdot 10^{-2}$	$1.12 \cdot 10^{-2}$	$2.86 \cdot 10^{-3}$	$1.14 \cdot 10^{-3}$
$n_{ph} = 2$	$8.26 \cdot 10^{-2}$	$2.39 \cdot 10^{-2}$	$1.32 \cdot 10^{-2}$	$5.21 \cdot 10^{-3}$
$n_{ph} = 10$	$1.27 \cdot 10^{-2}$	$7.95 \cdot 10^{-3}$	$2.73 \cdot 10^{-3}$	$3.68 \cdot 10^{-4}$
$n_{ph} = 20$	$8.73 \cdot 10^{-3}$	$3.64 \cdot 10^{-3}$	$3.42 \cdot 10^{-4}$	$2.25 \cdot 10^{-4}$
$\widetilde{\boldsymbol{X}}$	$7.63 \cdot 10^{-3}$	$5.53 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$	$1.30 \cdot 10^{-6}$

The relative generalization error of the augmented PCE is also compared to a PCE model in the original input domain $\mathcal{D}_{\boldsymbol{X}}$. For this purpose an auxiliary, probabilistic input vector $\widetilde{\boldsymbol{X}}$ is used as a basis for the PCE meta-model. Each component $\widetilde{X}_i$ is defined as the p-box distribution family with central value mean value (*i.e.* $\mu_{\widetilde{X}_i} = \left(\underline{\mu}_{X_i} + \overline{\mu}_{X_i}\right)/2$) and maximal coefficient of variation (*i.e.* $\eta_{\widetilde{X}_i} = \overline{\eta}_{X_i}$). Despite the higher-dimensional input domain $\mathcal{D}_{\boldsymbol{V}}$ (30 dimensions instead of 10 for $\mathcal{D}_{\widetilde{\boldsymbol{X}}}$), the augmented PCE performs comparable to the PCE in the auxiliary domain $\mathcal{D}_{\widetilde{\boldsymbol{X}}}$ when using a reasonably large number of phantom points.

The accuracy of the meta-model is also measured in terms of the shape of response p-box. Figure 6.24 compares the boundary curves of the response p-box of the augmented PCE meta-model ($N = 200$, $n_{ph} = 20$, $\widehat{err}_{gen} = 3.64 \cdot 10^{-3}$) to a reference solution. The reference solution is obtained by using the exact computational model $\mathcal{M}$ instead of the meta-model. The boundary curves of the reference solution and the augmented PCE-based solution are close, hence, indicating a high accuracy of the augmented PCE meta-model. The values of the relative generalization error in Table 6.5 confirm this impression.

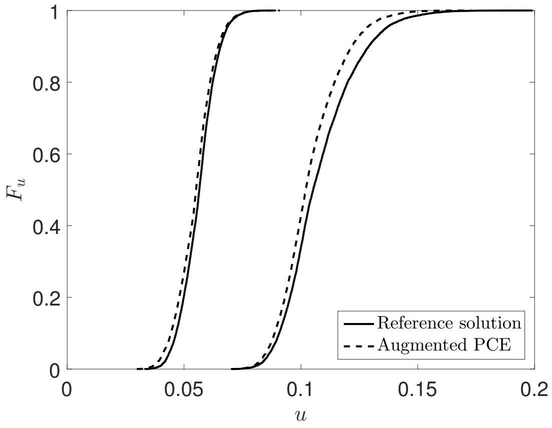


Figure 6.24: Two-dimensional truss structure – boundary curves of the response p-box – exact model versus augmented PCE model.

Comparison to free p-boxes

In Section 6.3.4, the same structural truss is discussed in the context of free p-boxes. The p-box boundaries of the input variables X_i of the latter and present section coincide by construction.

Both sections result in p-boxes for the midspan deflection u, the boundary curves of which are shown in Figure 6.25. Despite the different formulation of the input p-boxes, boundary curves are almost identical. In fact, the free p-box formulation leads to wider boundary curves as expected, but the difference to the parametric p-box formulation is negligibly small in contrast to the Rosenbrock example.

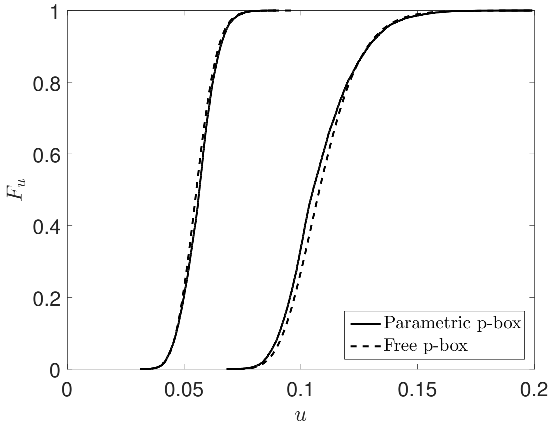


Figure 6.25: Two-dimensional truss structure – boundary curves of the response p-box – free versus parametric p-box approach.

In terms of the number of model evaluations N, the two p-box approaches are comparable, too. Considering an experimental design size of $N \approx 200$, the meta-modelling approaches are capable of representing the underlying computational model and hence to propagate the p-boxes: the two-level approach for free p-boxes and the augmented PCE model for parametric p-boxes. As a result, the boundary curves are obtained efficiently in both p-box approaches.

6.6 Comparison free and parametric p-box approaches

In the case of parametric p-boxes, PCE in the augmented space approximate the computational model $\mathcal{M}$ and at the same time account for the distinction between epistemic and aleatory uncertainty. The number of computational model runs remains small because of the introduction of phantom points. However, there is still a nMCS required to estimate the response p-box after the calibration of the meta-model. In order to construct the bounds efficiently (see also Eq. (6.3)), an optimization algorithm in combination with simple MCS can be applied instead of a plain nMCS.

In the context of free p-boxes (see Section 6.2), the nature of the problem solving scheme is different. In fact, the uncertainty propagation problem is solved based on the idea of interval analysis. Hence, the two-level meta-modelling approach follows a different path than the augmented PCE approach for parametric p-boxes. The two levels of meta-models are required to reduce the computational costs introduced mainly by the large number of optimizations in the converted problem described in Section 6.2.3 and Eq. (6.12). However, the second level meta-models, approximating $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$, directly represent the bounds of the response p-box. In the end, each bound can be estimated by simple MCS.

Table 6.6 summarizes the characteristics of the two approaches corresponding to the two p-box types. Consider that the complexity of a meta-model is given by its dimensionality. From this compilation, it is apparent that the free p-box approach requires three different meta-models of complexity M. The parametric p-box approach requires just a single meta-model but of increased complexity. Considering that most common distributions have two hyper-parameters, the augmented PCE model has likely a complexity of $3 \cdot M$. Although the complexity can be handled by phantom points, the amount of storage required to calibrate such meta-models might be infeasible. Hence, it is important to use sparse PCE algorithms to limit the required computational resources.

Table 6.6: Imprecise uncertainty propagation – comparison of free and parametric p-box approaches.

	Free p-box	Parametric p-box
P-box	$\underline{F}_{\boldsymbol{X}},\ \overline{F}_{\boldsymbol{X}}$	$F_{\boldsymbol{X}}(\boldsymbol{x}\|\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$
Main solver	Interval analysis	Nested MCS
PCE models	$\mathcal{M}(\boldsymbol{X}), \underline{\mathcal{M}}(\boldsymbol{C}), \overline{\mathcal{M}}(\boldsymbol{C})$	$\mathcal{M}^{(\text{aug})}(\boldsymbol{V})$
PCE dimensionality	$\|\boldsymbol{X}\| = \|\boldsymbol{C}\| = M$	$\|\boldsymbol{V}\| = \|\boldsymbol{X}\| + \|\boldsymbol{\Theta}\| = M + \|\boldsymbol{\Theta}\|$
Optimizations	Eq. (6.12)	Eq. (6.3)

CHAPTER 7

Imprecise structural reliability analysis

In this chapter, a number of algorithms are proposed to conduct structural reliability analysis in the presence of p-boxes. The algorithms are based on the developments discussed previously for the case of imprecise uncertainty propagation. Moreover, the use of meta-models at various stages of the analysis allows for an efficient estimation of failure probabilities. The meta-model-based algorithm for free p-boxes has been published in Schöbi & Sudret (2015b, 2016a,b).

7.1 State of the art

As seen in Chapter 5, there are many methods available for conducting structural reliability analysis (SRA) in the presence of probabilistic variables. In the presence of imprecise probabilities, however, fewer methods and publications are available, some of which are mentioned here. Eldred (2009) and Hurtado (2013) use nested Monte Carlo simulations to cope with imprecise probabilities. Alvarez & Hurtado (2014) combine subset simulation with random sets. Z. Zhang, Jiang, Wang, et al. (2015) and Alibrandi & Koh (2015) generalize FORM/SORM for the context of Dempster-Shafer's evidence theory. A combination of line sampling and imprecise probabilities is discussed in De Angelis, Patelli, & Beer (2014) and De Angelis, Patelli, & Beer (2015). Balu & Rao (2013) discuss imprecise structural reliability analyses (ISRA) in the context of fuzzy variables. Muscolino & Sofi (2012) and Muscolino, Santoro, & Sofi (2015) discuss the interval perturbation method for uncertain-but-bound parameters. These publications and their methods typically require a large number of model evaluations and hence rely on an inexpensive-to-evaluate performance function to become a tractable analysis (Beer, Ferson, & Kreinovich 2013).

A strategy to reduce the computational costs and to make the analysis tractable is the use of meta-models, as seen also in Chapter 5 for probabilistic input vectors. As seen already, using meta-models allows for an efficient estimation of both failure probabilities and quantiles. Again, only few authors combine meta-modelling techniques with imprecise probabilities and

SRA. Balesdent, Morio, & Brevault (2016) and Morio & Balesdent (2016) use an adaptive-Kriging importance sampling algorithm for distributions with interval-valued parameters. A combination of Karhunen-Loève expansions, evidence theory and structural reliability analysis is presented in Oberguggenberger (2014). Z. Zhang, Jiang, Han, et al. (2014) use radial basis expansions and quasi-Monte Carlo simulations in the case of evidence theory.

Based on these developments, a set of algorithms is proposed in this chapter to further increase the efficiency of SRA methods in the presence of epistemic uncertainties. In particular, free and parametric p-box approaches are presented separately and compared later in Section 7.6. The application of Kriging meta-models at several stages of the imprecise structural reliability analysis promises a reduction of the overall computational costs in analogy to Chapter 5.

7.2 Imprecise failure probability

In the context of probability theory, the definition of the failure probability is defined and discussed in Section 5.2. In the context of interval-valued epistemic uncertainty and hence p-boxes, Eq. (5.3) leads to a range of failure probabilities $P_f \in \left[\underline{P}_f, \overline{P}_f\right]$, the bounds of which are defined in the general case as:

$$\underline{P}_f = {}''\min_{f_{\boldsymbol{X}}}{}'' \int_{\mathcal{D}_f} f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \qquad \overline{P}_f = {}''\max_{f_{\boldsymbol{X}}}{}'' \int_{\mathcal{D}_f} f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \tag{7.1}$$

where ″min″ (resp. ″max″) means that the optimization would be carried out over all PDF $f_{\boldsymbol{X}}$ which satisfy some constraints related to the definition of the underlying p-box. This *imprecise structural reliability analysis*(ISRA) is not straightforward as it involves an optimization with a multi-dimensional PDF $f_{\boldsymbol{X}}$ as the argument of the objective function. Moreover, depending on the type of p-boxes, the shape of the PDF $f_{\boldsymbol{X}}$ is more or less restricted. In the following sections, solution algorithms for the two types of p-boxes are discussed based on the developments in the previous sections, in particular based on the proposed imprecise uncertainty quantification algorithms (Chapter 6) and adaptive meta-modelling techniques (Chapter 5).

7.3 Free p-boxes

7.3.1 General idea

The uncertainty propagation of free p-boxes has been discussed in Section 6.2. The resulting two-level meta-modelling approach, displayed in Figure 6.3, allows for an efficient propagation of free p-boxes. The resulting uncertainty model in Y is again a free p-box, characterized by CDFs $F_Y(y) \in \left[\underline{F}_Y(y), \overline{F}_Y(y)\right]$. Based on the definition of the limit-state function, *i.e.* $P_f = \mathbb{P}(Y \leq 0)$ in the usual case, the failure probability can be found as $P_f = F_Y(0) \in \left[\underline{F}_Y(0), \overline{F}_Y(0)\right]$. In other words, the belief and plausibility values of the failure probability are found as $\underline{P}_f = \underline{F}_Y(0)$ and $\overline{P}_f = \overline{F}_Y(0)$, respectively. Figure 7.1 shows the response p-box of Y by the corresponding boundary curves. Further, the two bounds for the failure probability are illustrated.

Considering a limit-state function $\mathcal{G}$ instead of the computational model $\mathcal{M}$, the failure probability can be formulated as a function of the response variable Y. Then, the bounds to the failure probability can be formulated in terms of the bounds of the response free p-box:

$$\underline{P}_f = \mathbb{P}\left(\underline{Y} \leq 0\right) = \mathbb{P}\left(\overline{\mathcal{G}}\left(\boldsymbol{C}\right) \leq 0\right), \tag{7.2}$$

$$\overline{P}_f = \mathbb{P}\left(\overline{Y} \leq 0\right) = \mathbb{P}\left(\underline{\mathcal{G}}\left(\boldsymbol{C}\right) \leq 0\right), \tag{7.3}$$

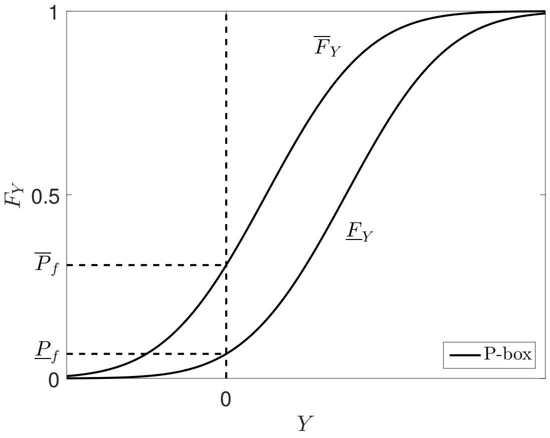


Figure 7.1: Imprecise structural reliability analysis (ISRA) using free p-boxes – definition of the bounds of the failure probability.

where $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$ are defined analogously to Eq. (6.12) for $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$, respectively:

$$\underline{\mathcal{G}}\left(\boldsymbol{c}\right) = \min_{\boldsymbol{x}\in\mathcal{D}_{\boldsymbol{c}}} \mathcal{G}\left(\boldsymbol{x}\right), \qquad \overline{\mathcal{G}}\left(\boldsymbol{c}\right) = \max_{\boldsymbol{x}\in\mathcal{D}_{\boldsymbol{c}}} \mathcal{G}\left(\boldsymbol{x}\right). \tag{7.4}$$

Hence, for the case of free p-boxes, the imprecise structural reliability problem can be recast as two structural reliability problems in the probabilistic sense (H. Zhang, Mullen, & Muhanna 2010; H. Zhang 2012; Schöbi & Sudret 2015b).

A major drawback in this approach is, however, the high computational costs caused by the same factors discussed in Chapter 6.2: expensive-to-evaluate limit-state functions, large number of model evaluations, and small failure probabilities. Using the same idea as in the latter section, the limit-state function can be replaced by a meta-model to reduce these computational costs. In contrast to Section 6.2 where PCE models are used, however, Kriging models with adaptive experimental designs are used here. The adaptive experimental design algorithm allows for an efficient modelling of the limit-state surface and hence estimation of failure probabilities, as discussed in Chapter 5.

7.3.2 First-level meta-model

A first meta-model is applied to the limit-state function $\mathcal{G}$ and, in particular, to model the limit-state surface $\mathcal{G}\left(\boldsymbol{x}\right) = 0$. However, in order to conduct an AK-MCS analysis as introduced in Section 5.3, a probabilistic input vector is required. Analogously to Section 6.2, an auxiliary input vector $\widetilde{\boldsymbol{X}}$ is created by condensation. The auxiliary distribution shall cover the shape of the p-box such that the resulting meta-model $\mathcal{G}^{(\mathrm{K})}$ is accurate in the very neighbourhood of the limit-state surface that contributes to the failure probability in the p-box setting.

7.3.3 Second-level meta-models

The second meta-model is applied to the limit-state functions $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$ and the estimation of the bounds of the failure probability. By using the approximation of $\mathcal{G}$, Eq. (7.4) may be replaced by:

$$\underline{\mathcal{G}}\left(\boldsymbol{c}\right) \approx \min_{\boldsymbol{x}\in\mathcal{D}_{\boldsymbol{c}}} \mathcal{G}^{(\mathrm{K})}\left(\boldsymbol{x}\right), \qquad \overline{\mathcal{G}}\left(\boldsymbol{c}\right) \approx \max_{\boldsymbol{x}\in\mathcal{D}_{\boldsymbol{c}}} \mathcal{G}^{(\mathrm{K})}\left(\boldsymbol{x}\right), \tag{7.5}$$

where $\mathcal{G}^{(\mathrm{K})}$ is the meta-model resulting from the first-level approximation. The optimizations can be solved by available methods, such as genetic algorithms, due to the small computational costs of evaluating the meta-model $\mathcal{G}^{(\mathrm{K})}$. Therefore, the bounds of the failure probability $\underline{P}_f$ and $\overline{P}_f$ can be estimated by two independent AK-MCS analyses of $\overline{\mathcal{G}}$ and $\underline{\mathcal{G}}$, respectively.

The bounds to the failure probability can be estimated as a function of the probabilistic input vector $\boldsymbol{C}$. However, to improve the convergence of the AK-MCS algorithm, an auxiliary distribution may be beneficial. In fact, the same auxiliary distribution $\widetilde{\boldsymbol{X}}$ is suitable when the free p-boxes are unbounded, as shown Section 6.2.4 and in Schöbi & Sudret (2015b). Then, an isoprobabilistic transform maps $\boldsymbol{C}$ to $\widetilde{\boldsymbol{X}}$ and vice versa.

7.3.4 Compilation

Figure 7.2 summarizes the procedure, which consists of two sequential meta-modelling levels connected by the problem conversion step. Analogously to Figure 6.3, the first-level meta-model surrogates the original limit-state function $\mathcal{G}$, whereas the second-level meta-model surrogates the limit-state function for estimating the lower and upper bounds of the failure probability. Both levels use auxiliary random variables, *i.e.* $\widetilde{\boldsymbol{X}}$ on the first level and $\boldsymbol{C}$ on the second level.

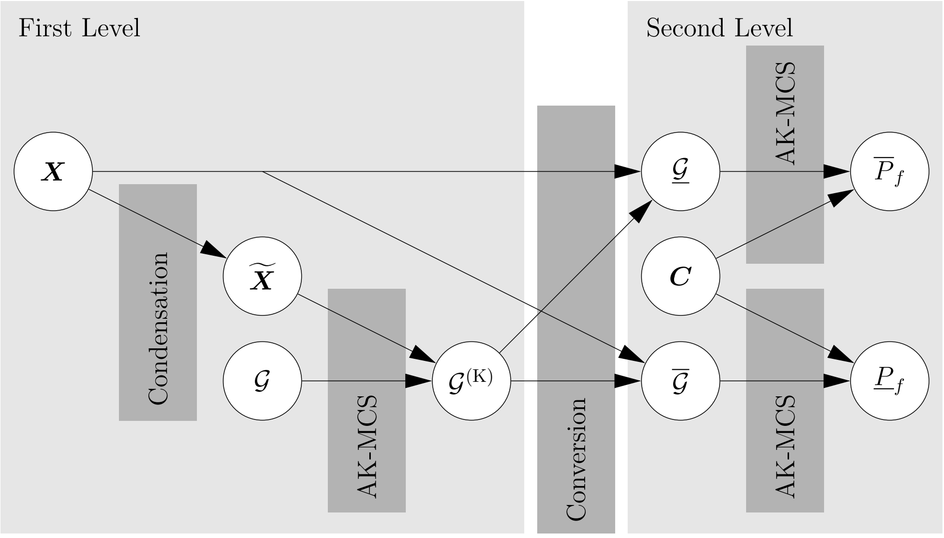


Figure 7.2: ISRA for free p-boxes – flowchart of the two-level meta-modelling algorithm.

7.4 Parametric p-boxes

7.4.1 General idea

The definition of parametric p-boxes indicates a hierarchical model where the distribution of X is defined conditionally on its distribution parameters. Hence, nested simulation algorithms can be applied as previously discussed in Section 6.4 in the context of uncertainty propagation. In the context of structural reliability analysis, the computational model $\mathcal{M}$ is replaced by the limit-state function $\mathcal{G}$ and the quantity of interest is the failure probability instead of the entire response p-box.

Given parametric p-boxes, the bounds for the failure probability can be found by optimizing the value of the distribution parameters $\boldsymbol{\theta}$:

$$\underline{P}_f = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} P_{f|\boldsymbol{\theta}}\left(\boldsymbol{\theta}\right), \qquad \overline{P}_f = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} P_{f|\boldsymbol{\theta}}\left(\boldsymbol{\theta}\right), \tag{7.6}$$

where $P_{f|\boldsymbol{\theta}} = \mathbb{P}\left(\mathcal{G}\left(\boldsymbol{X}_{\boldsymbol{\theta}}\right) \leq 0\right)$ is the failure probability conditional on $\boldsymbol{\theta}$, and $\boldsymbol{X}_{\boldsymbol{\theta}}$ is the conditional distribution characterized by the CDF $F_{\boldsymbol{X}_{\boldsymbol{\theta}}}(\boldsymbol{x}) = F_{\boldsymbol{X}}\left(\boldsymbol{x}|\boldsymbol{\theta}\right)$. Considering a Monte Carlo simulation, the conditional failure probability is estimated by (see also Eq. (5.5)):

$$\widehat{P}_{f|\boldsymbol{\theta}} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\mathcal{G}(\boldsymbol{x}) \leq 0} \left(\boldsymbol{x}^{(i)}\right), \tag{7.7}$$

where $\mathcal{S}_{\boldsymbol{\theta}} = \left\{\boldsymbol{x}^{(i)},\ i = 1, \ldots, n\right\}$ is a MC sample of the conditional input vector $\boldsymbol{X}_{\boldsymbol{\theta}}$. Note that each $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$ will lead to a different failure probability.

Considering this setup, it is important (i) to estimate the individual conditional failure probability efficiently and (ii) to choose $\boldsymbol{\theta}$ in a way as to reach the bounds of the failure probability efficiently. One way to do so is to use the meta-modelling techniques with adaptive experimental designs as introduced in Chapter 5.

7.4.2 Conditional failure probabilities

In the context of parametric p-boxes, the core task consists of estimating efficiently the conditional failure probability $P_{f|\boldsymbol{\theta}}$ by AK-MCS. Given a vector $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$ and consequently conditional variables $\boldsymbol{X}_{\boldsymbol{\theta}}$, AK-MCS can be applied to estimate the conditional failure probability analogously as described in Chapter 5 and Eq. (7.7).

7.4.3 Optimization on distribution parameters

Efficient global optimization (EGO) is a global optimization algorithm which is based on Kriging meta-models with a design enrichment strategy similar to AK-MCS (Jones, Schonlau, & Welch 1998; Echard, Gayton, & Lemaire 2011). In the context of ISRA for parametric p-boxes, EGO can be used to find the extreme values of the failure probability by optimizing on the distribution parameters $\boldsymbol{\theta}$. The main steps are discussed here:

(i) Given the distribution parameter $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$, generate a small initial experimental design $\boldsymbol{\mathcal{T}} = \left\{\boldsymbol{\tau}^{(1)}, \ldots, \boldsymbol{\tau}^{(N)}\right\}$.

(ii) Compute the corresponding responses $\mathcal{P}^{(i)} = \mathbb{P}\left(\mathcal{G}\left(\boldsymbol{X}_{\boldsymbol{\tau}^{(i)}}\right) \leq 0\right)$ for $i = 1, \ldots, N$ by AK-MCS. Note that for the first evaluation of such a conditional failure probability, the procedure of Section 5.3 is applicable one-to-one with the distribution $\boldsymbol{X}_{\boldsymbol{\tau}^{(1)}}$. However, for $i = 2, \ldots, n$, the limit-state function evaluations of the previous AK-MCS analyses are kept as initial experimental design. The difference lies then solely in a modified $\boldsymbol{X}_{\boldsymbol{\tau}^{(i)}}$ and hence a modified $\mathcal{S}_{\boldsymbol{\tau}^{(i)}}$. Due to the similarity of the MC populations $\mathcal{S}_{\boldsymbol{\tau}^{(i)}}$, the number of limit-state function evaluations can be kept small.

(iii) Train a Kriging meta-model $P_f^{(\mathrm{K})}\left(\boldsymbol{\theta}\right)$ based on $\{\boldsymbol{\mathcal{T}}, \mathcal{P}\}$. This meta-model approximates the estimation of the conditional failure probability $P_{f|\boldsymbol{\theta}}(\boldsymbol{\theta})$.

(iv) Search for the optimal samples to be added to the experimental design for the minimization and maximization of the failure probability:

$$\boldsymbol{\tau}^{*}_{min} = \arg \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} \left[EI_{min}\left(\boldsymbol{\theta}\right)\right], \tag{7.8}$$

$$\boldsymbol{\tau}^{*}_{max} = \arg \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} \left[EI_{max}\left(\boldsymbol{\theta}\right)\right]. \tag{7.9}$$

where the expected improvement is defined as (Mockus, Tiesis, & Zilinskas 1978; Jones, Schonlau, & Welch 1998):

$$EI_{min}(\boldsymbol{\theta}) = \left(\mathcal{P}_{min} - \mu_{\widehat{P}_f}(\boldsymbol{\theta})\right) \Phi\left(\frac{\mathcal{P}_{min} - \mu_{\widehat{P}_f}(\boldsymbol{\theta})}{\sigma_{\widehat{P}_f}(\boldsymbol{\theta})}\right) + \sigma_{\widehat{P}_f}(\boldsymbol{\theta})\ \varphi\left(\frac{\mathcal{P}_{min} - \mu_{\widehat{P}_f}(\boldsymbol{\theta})}{\sigma_{\widehat{P}_f}(\boldsymbol{\theta})}\right), \tag{7.10}$$

$$EI_{max}(\boldsymbol{\theta}) = \left(\mu_{\widehat{P}_f}(\boldsymbol{\theta}) - \mathcal{P}_{max}\right) \Phi\left(\frac{\mu_{\widehat{P}_f}(\boldsymbol{\theta}) - \mathcal{P}_{max}}{\sigma_{\widehat{P}_f}(\boldsymbol{\theta})}\right) + \sigma_{\widehat{P}_f}(\boldsymbol{\theta})\ \varphi\left(\frac{\mu_{\widehat{P}_f}(\boldsymbol{\theta}) - \mathcal{P}_{max}}{\sigma_{\widehat{P}_f}(\boldsymbol{\theta})}\right), \tag{7.11}$$

where EI_{min} and EI_{max} are the expected improvement for minimizing and maximizing P_f, respectively, $\mathcal{P}_{min} = \min_i \left[\mathcal{P}^{(i)}\right]$, $\mathcal{P}_{max} = \max_i \left[\mathcal{P}^{(i)}\right]$, and $\Phi(\cdot)$ and $\varphi(\cdot)$ are the CDF and PDF of a standard normal variable, respectively. The expected improvement allows for a balance between local and global search (Jones, Schonlau, & Welch 1998).

(v) Check the convergence criterion proposed by Jones, Schonlau, & Welch (1998), applied to the maximum and minimum estimates of the failure probability:

$$EI_{min}\left(\boldsymbol{\tau}^*_{min}\right) \leq \epsilon_{EI}, \tag{7.12}$$

$$EI_{max}\left(\boldsymbol{\tau}^*_{max}\right) \leq \epsilon_{EI}. \tag{7.13}$$

If both criteria are fulfilled, terminate the optimization algorithm and return the last estimates of the boundary value for the failure probability: $\underline{P}_f = \min_i \left[\mathcal{P}^{(i)}\right]$, $\overline{P}_f = \max_i \left[\mathcal{P}^{(i)}\right]$. Otherwise, continue with Step (vi). Note that a value of $\epsilon_{EI} = 10^{-5}$ was identified as a reliable choice for the applications in Section 7.6.

(vi) Add the best next sample(s) to the experimental design $\boldsymbol{\mathcal{T}}$. The Steps (iv) and (v) allow for a simultaneous or separate optimization for the bounds of the failure probability. For simultaneous optimization, both boundary values of the failure probability are estimated with the same meta-model and $\{\boldsymbol{\tau}^*_{min}, \boldsymbol{\tau}^*_{max}\}$ are added in each iteration. For separate optimization, the lower and upper bounds of the failure probability are estimated one after another. The effect of this choice is discussed in Section 7.6. Then, depending on Step (v) and the optimization mode, $\{\boldsymbol{\tau}^*_{min}, \boldsymbol{\tau}^*_{max}\}$ is added if Eqs. (7.12) and (7.13) are fulfilled. $\{\boldsymbol{\tau}^*_{min}\}$ is added if only Eq. (7.12) is fulfilled and $\{\boldsymbol{\tau}^*_{min}\}$ is added if only Eq. (7.13) is fulfilled.

(vii) Evaluate the failure probability corresponding to $\boldsymbol{\tau}^*$ and add it to $\mathcal{P}$. Note that the limit-state function evaluations computed in AK-MCS can be recycled. They can be used as the initial experimental design in AK-MCS for estimating $\mathcal{P}^*$ corresponding to $\boldsymbol{\tau}^*$ (see also Step (ii) for further details). Finally, continue with Step (iii).

Figure 7.3 summarizes the EGO algorithm for finding the minimum failure probability $\underline{P}_f$ in analogy to Figure 5.1. The white boxes represent the starting point of the algorithm, whereas the grey boxes mark the computational tasks described previously. Note that for obtaining the maximum failure probability, the same flowchart is valid when replacing EI_{min} by EI_{max} and $\underline{P}_f$ by $\overline{P}_f$.

When comparing Figures 7.3 to 5.1, some differences between EGO and AK-MCS are apparent. The main difference is that AK-MCS optimizes on a sample set $\mathcal{S}$, whereas EGO optimizes on the continuous, $n_{\boldsymbol{\Theta}}$-dimensional parameter domain $\mathcal{D}_{\boldsymbol{\Theta}}$. The reason for searching in $\mathcal{D}_{\boldsymbol{\Theta}}$ is threefold. Firstly, the input domain is given in intervals without a probabilistic description (*i.e.*

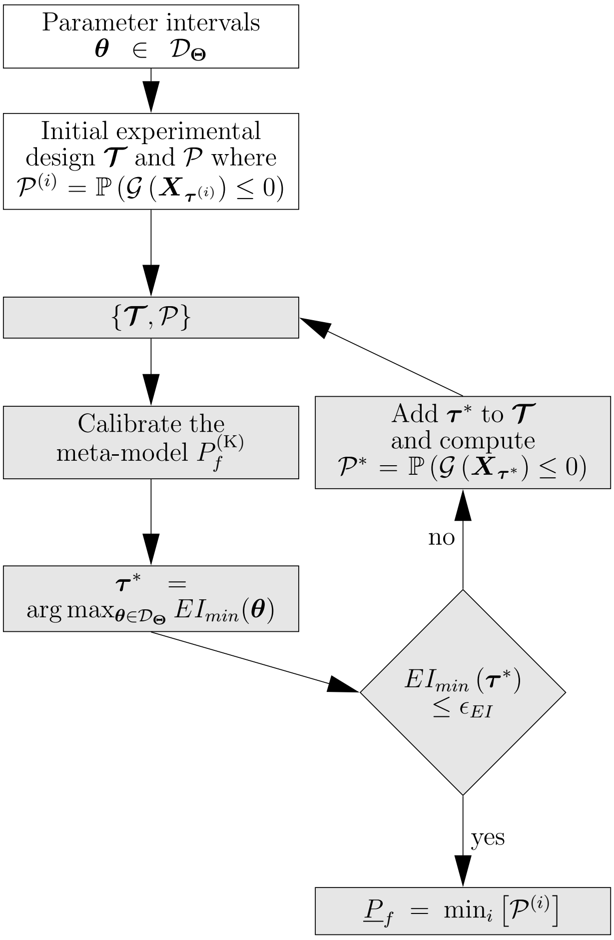


Figure 7.3: Flowchart of the EGO algorithm to find extreme failure probabilities.

PDF). Hence, an auxiliary distribution is required to sample from in order to obtain a candidate set $\mathcal{S}$. The use of $\mathcal{D}_{\boldsymbol{\Theta}}$ avoids this task. Secondly, $\mathcal{D}_{\boldsymbol{\Theta}}$ can be interpreted as an infinite set of candidate samples. Hence, the chance of hitting the global optimum is larger than in the case of $\mathcal{S}$. Thirdly, in AK-MCS, the quantity of interest (*i.e.* P_f) depends on the entire set $\mathcal{S}$ rather than a single sample $\boldsymbol{x}^{(i)} \in \mathcal{S}$. On the other side, EGO searches for the optimal single sample $\boldsymbol{\tau}$ that optimizes the failure probability. Hence, the target result of EGO is point-dependent, whereas the result of AK-MCS is ensemble-dependent.

7.4.4 Compilation

Figure 7.4 illustrates the main components of the proposed algorithm. Similar to the imprecise uncertainty quantification algorithm in Figure 6.18, the nested algorithm is visible. The outer loop consists of the EGO method on the distribution parameters $\boldsymbol{\theta}$, whereas the inner loop consists of the AK-MCS method to estimate failure probabilities. The main difference to Figure 6.18 is, however, that the outer loop consists of an optimization algorithm rather than a pure MCS.

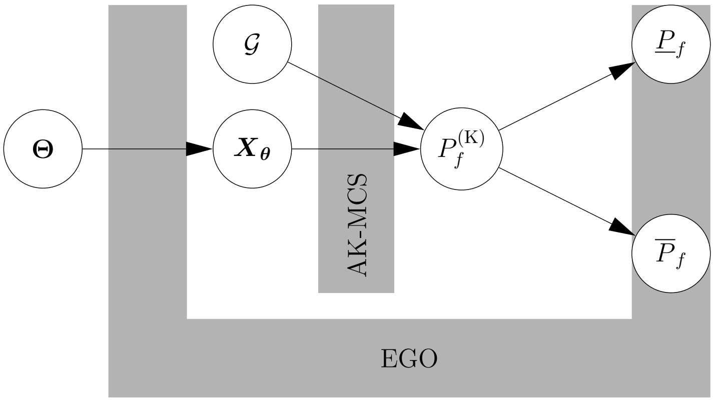


Figure 7.4: ISRA for parametric p-boxes – overview of the nested algorithm.

7.5 Multiple p-box types

The previous two sections dealt with the case where either free or parametric p-boxes are present. However, if both types of p-boxes appear in a problem setting for different variables, these concepts may be combined to one.

Denote the variables modelled by free p-boxes as $\boldsymbol{X}^{(\mathrm{f})}$ and the ones modelled by parametric p-boxes as $\boldsymbol{X}^{(\mathrm{p})}$, such that $\boldsymbol{X} = \left(\boldsymbol{X}^{(\mathrm{f})}, \boldsymbol{X}^{(\mathrm{p})}\right)$ and $\left|\boldsymbol{X}^{(\mathrm{f})}\right| + \left|\boldsymbol{X}^{(\mathrm{p})}\right| = M$. Then, for a given $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$, the bounds of the failure probability can be estimated as described in Section 7.3 in the context of free p-boxes. Note that the conditional variables $X^{(\mathrm{p})}_{\boldsymbol{\theta}_i,i}$ are interpreted as free p-boxes of which the two boundary CDFs coincide. As a result, boundary values of the failure probability are obtained conditional on the parameter values $\boldsymbol{\theta}$, *i.e.* $\underline{P}_{f|\boldsymbol{\theta}}$ and $\overline{P}_{f|\boldsymbol{\theta}}$. In a second stage, the minimum value of $\underline{P}_{f|\boldsymbol{\theta}}$ and the maximum value of $\overline{P}_{f|\boldsymbol{\theta}}$ are obtained by applying EGO twice in the parameter space $\mathcal{D}_{\boldsymbol{\Theta}}$, in a similar fashion as described in Section 7.4.

The procedure is summarized in Figure 7.5. It consists of three levels of meta-modelling: (i) a Kriging meta-model of $\mathcal{G}$, (ii) Kriging models of $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$, and (iii) two Kriging models for optimization (*i.e.* for $\underline{P}_{f|\boldsymbol{\theta}}$ and $\overline{P}_{f|\boldsymbol{\theta}}$). Note that evaluations of the true limit-state function $\mathcal{G}$ are only computed in the AK-MCS procedure resulting in $\mathcal{G}^{(\mathrm{K})}$. The evaluations of $\mathcal{G}$ can be recycled in the different iterations of EGO as previously discussed in Section 7.4.

7.6 Numerical examples

In order to visualize how the proposed two-level meta-modelling approaches work, four application examples are presented in this section. The toy function visualizes the adaptive Kriging approach on a simple analytical function. Thereafter, three examples discuss realistic problem settings involving more complex computational models. The analyses focus on the comparison of free and parametric p-box modelling.

7.6.1 Toy function

Problem statement

The first example is a simple two-dimensional problem to illustrate the machinery of the proposed approaches. The limit-state function is defined as:

$$g_1(\boldsymbol{x}) = x_1 - x_2^2. \tag{7.14}$$

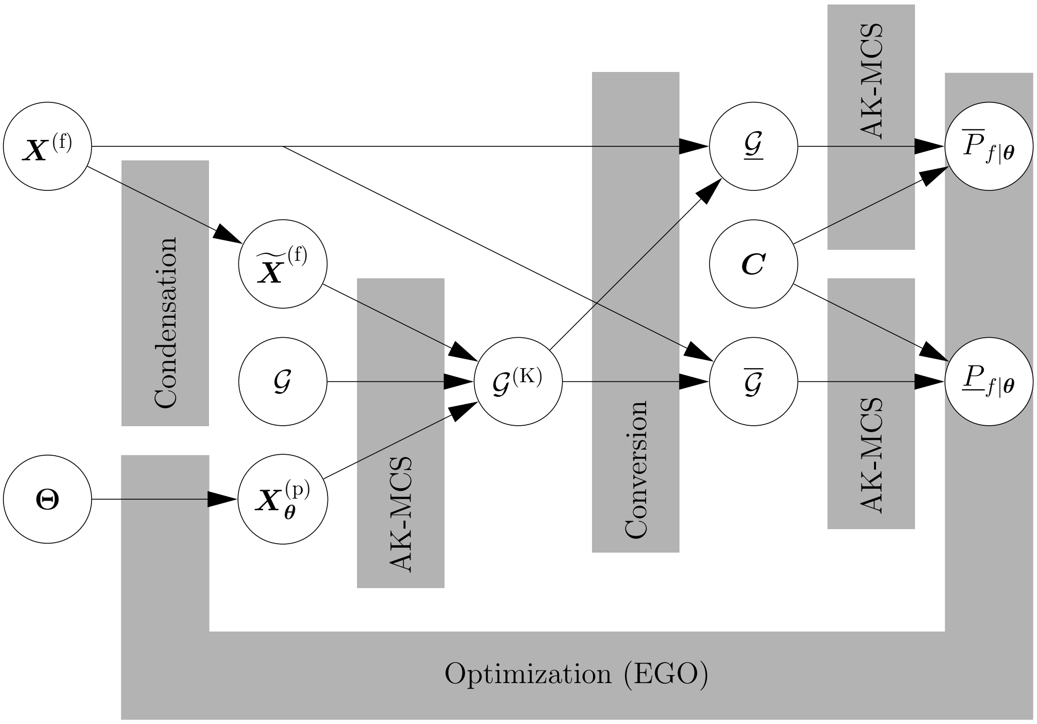


Figure 7.5: ISRA for multiple p-box types (*i.e.* free and parametric) – overview of the proposed algorithm.

Failure is defined as $g_1(\boldsymbol{x}) \leq 0$ and hence the failure probability reads $P_f = \mathbb{P}\left(g_1\left(\boldsymbol{X}\right) \leq 0\right)$. The two variables X_i are modelled by p-boxes. In the following, two cases are distinguished: (i) X_i is modelled by free p-boxes and (ii) X_i is modelled by parametric p-boxes. The bounds of the free p-boxes are defined as:

$$\underline{F}_{X_i}(x_i) = F_{\mathcal{N}}\left(x_i|\mu = 2.5, \sigma = 1\right),$$

$$\overline{F}_{X_i}(x_i) = F_{\mathcal{N}}\left(x_i|\mu = 1.5, \sigma = 1\right),$$

where $F_{\mathcal{N}}(x|\mu, \sigma)$ denotes the CDF of a Gaussian distribution with mean value μ and standard deviation σ. For the parametric p-box case, the distribution parameters $\boldsymbol{\theta}$ are chosen so that the boundary curves are the same as in the case of free p-boxes:

$$F_{X_i}(x_i) = F_{\mathcal{N}}\left(x_i|\mu, \sigma = 1\right), \qquad \mu \in [1.5,\ 2.5].$$

Analysis

The settings for the ISRA are the following. The meta-models and MCS are computed using the Matlab-based uncertainty quantification framework UQLab (Marelli & Sudret 2014), including its implementation of AK-MCS (Marelli, Lamas-Fernandes, et al. 2015). All Kriging models are defined by a Gaussian autocorrelation function and a constant trend ($f(\boldsymbol{x}) = 1$, a.k.a. ordinary Kriging). The number of samples in the initial experimental design are $N_0^{(1)} = 12$ and $N_0^{(2)} = 12$ for the first-level and second-level AK-MCS algorithms, respectively. Furthermore, the initial experimental design consists of $N_0^{(\mathrm{EGO})} = 4$ samples for EGO. The initial experimental designs are generated by the LHS method. For free p-boxes, the auxiliary input distribution is defined as $\widetilde{X}_i \sim \mathcal{N}(x_i|2, 1)$. For parametric p-boxes and EGO, the threshold value of the stopping criterion is set to $\epsilon_{EI} = 10^{-5}$. The failure probabilities are estimated with a set of $n_{\text{MCS}} = 10^6$ MC samples. The analysis is repeated 50 times with different initial experimental designs in order to obtain statistically significant results.

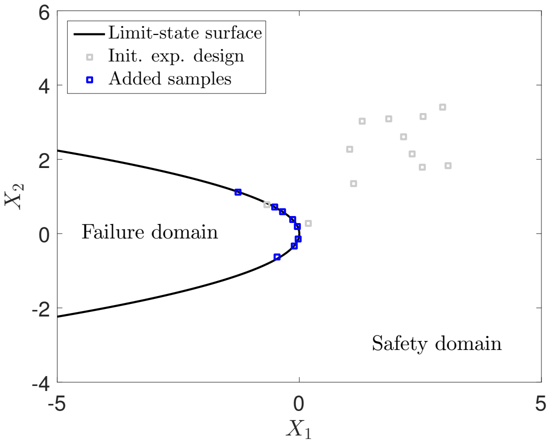
Results of free p-box case

Figure 7.6 visualizes a typical realization of the ISRA for free p-boxes. Note that the stopping criterion in AK-MCS was deactivated in this simulation run to visualize the general behaviour of the method. Figure 7.6(a) illustrates the experimental design of the final first-level meta-model. The initial experimental design is marked by the grey squares. The additional samples (blue squares) group around the limit-state surface (black solid line) in the physical domain, which implies an efficient selection of additional samples. Further, Figures 7.6(b) and 7.6(c) show the constraint optimization domains $\mathcal{D}_{\boldsymbol{c}}$ in Eq. (7.4) in the physical space $\mathcal{D}_{\boldsymbol{X}}$ for the experimental design on the second-level meta-model. The optimization domains $\mathcal{D}_{\boldsymbol{c}}$ mark rectangular areas due to the independence of the input components X_i. Again, the optimization domains are selected in an efficient manner seen in the positioning of the blue rectangles. For $\underline{\mathcal{G}}$ in Figure 7.6(b), the minimum value of the limit-state function within each blue rectangle is achieved at a point which lies close to the limit-state surface. Thus, the rectangular are contained in the safe domain. Analogously in Figure 7.6(c) for $\overline{\mathcal{G}}$, the maximum value of the limit-state function within each blue rectangle lies close to the limit-state surface. Hence, the rectangles are mostly contained in the failure domain.

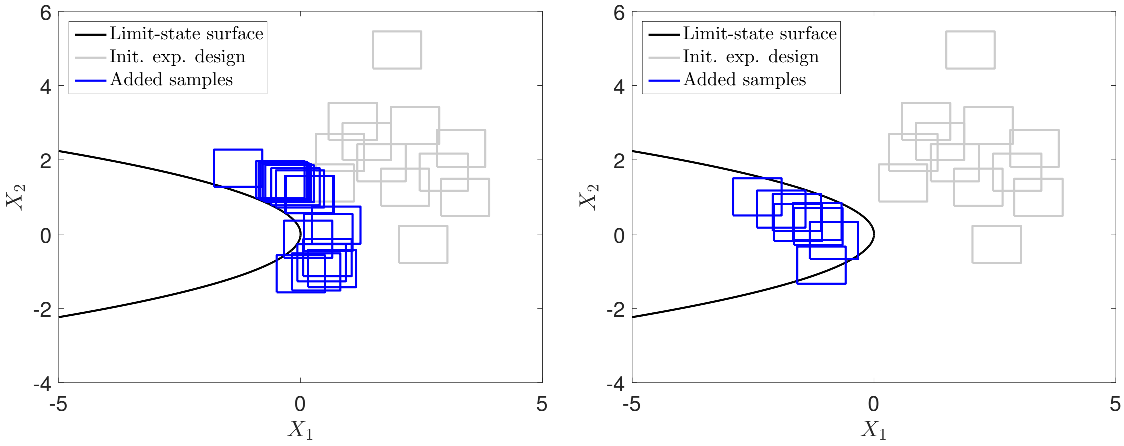
Accounting for the 50 replications of the ISRA, the estimates of the failure probability are summarized in Table 7.1. The reference value of the failure probability (denoted by $P_{f,ref}$) is obtained by an Importance Sampling analysis of $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$ with a sample size of $n = 10^6$. The table shows that all three meta-models accurately estimate the corresponding failure probabilities in terms of mean estimate. However, in terms of coefficient of variation, the second-level meta-models are less accurate than the expected Monte Carlo sampling-related values: for $\underline{P}_f$ the expected value is $\text{CoV}\left[\underline{P}_f\right] = \sqrt{(1-\underline{P}_f)/(n_{\text{MCS}} \cdot \underline{P}_f)} = 11.3\% < 13.8\%$ and for $\overline{P}_f$ the expected value is $\text{CoV}\left[\overline{P}_f\right] = 0.9\% < 7.4\%$. The large variation in $\overline{P}_f$ is caused by the larger number of samples in the failure domain and hence the larger length of the limit-state surface relevant for the failure probability estimation, when comparing to the auxiliary space $\mathcal{D}_{\widetilde{\boldsymbol{X}}}$. Selecting a different auxiliary distribution $\widetilde{\boldsymbol{X}}$ might reduce this phenomenon.

Table 7.1: Toy function – free p-boxes – failure probability estimation (results of 50 repetitions of the same analysis with different sample sets and initial experimental designs).

	$\widetilde{P}_f$	$\underline{P}_f$	$\overline{P}_f$
$P_{f,ref}$	$1.68 \cdot 10^{-3}$	$7.78 \cdot 10^{-5}$	$1.27 \cdot 10^{-2}$
$\mathbb{E}\left[\widehat{P}_f\right]$	$1.67 \cdot 10^{-3}$	$7.62 \cdot 10^{-5}$	$1.25 \cdot 10^{-2}$
$\text{CoV}\left[\widehat{P}_f\right]$	2.5%	13.8%	7.4%
$\mathbb{E}\left[N^{(1)}\right]$	$12 + 2.7 = 14.7$	-	-
$\text{Std}\left[N^{(1)}\right]$	0.6	-	-
$\mathbb{E}\left[N^{(2)}\right]$	-	$4 + 36.9 = 40.9$	$4 + 19.4 = 23.4$
$\text{Std}\left[N^{(2)}\right]$	-	13.3	17.7



(a) Experimental desgin for $\mathcal{G}$



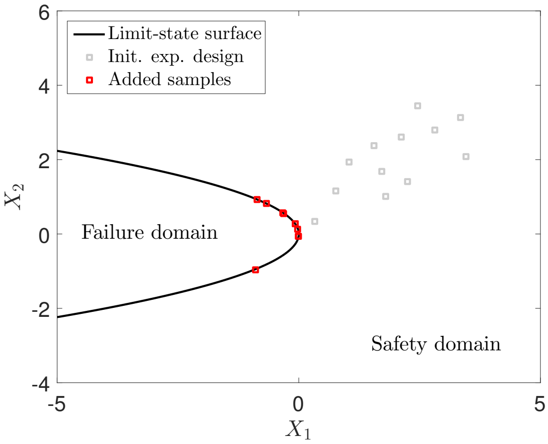
(b) Optimization domains $\mathcal{D}_c$ for $\underline{\mathcal{G}}$ (c) Optimization domains $\mathcal{D}_c$ for $\overline{\mathcal{G}}$

Figure 7.6: Toy function – free p-boxes – adaptive experimental designs of a single run of the analysis.

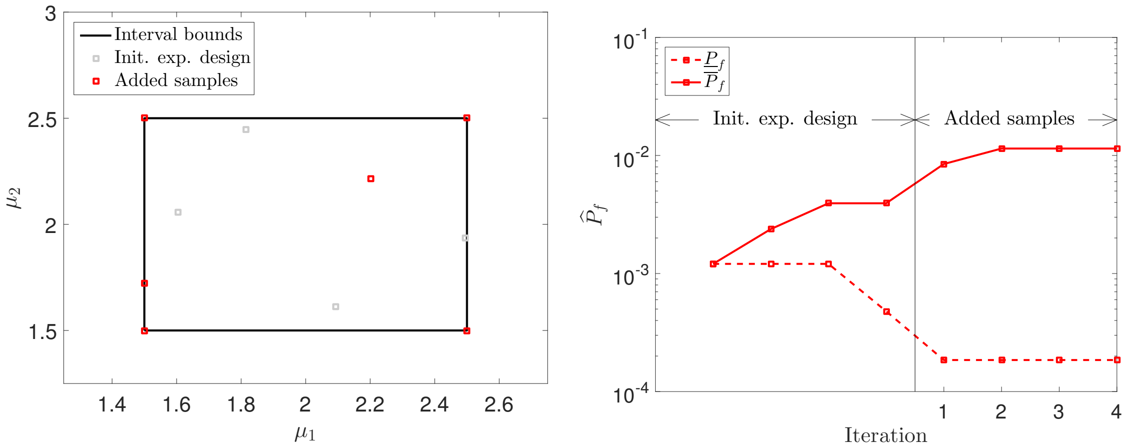
The lower part of Table 7.1 shows the number of samples in the experimental design of the final meta-models. $N^{(1)}$ and $N^{(2)}$ denote the number of model evaluations on the first-level and second-level meta-model, respectively. $N^{(1)}$ is low due to the smooth shape of the limit-state surface. In fact, an addition of $\Delta N^{(1)} = 2.7$ points on average is sufficient to model it. On the second-level meta-models, the number of samples $N^{(2)}$ is larger due to the shape of $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$ and the corresponding limit-state surfaces.

Results of parametric p-box case

A typical realization of the ISRA for parametric p-boxes in visualized in Figure 7.7. Figure 7.7(a) shows the experimental design of the final AK-MCS meta-model. Similar to Figure 7.6(a), the additional samples (red squares) group efficiently around the limit-state surface which is represented by the solid black line. The experimental design of the EGO algorithm is shown in Figure 7.7(b), which illustrates the exploratory behaviour of the EGO algorithm. Despite the small number of added samples, the boundary values of the failure probability are estimated accurately, as seen in Figure 7.7(c). The left side of Figure 7.7(c) shows the evolution of the failure probability boundary values in the initial experimental design ($N_0^{(\mathrm{EGO})} = 4$) which are evaluated according to their a priori defined design of experiments. The right side of the latter figure shows the current boundary values of the failure probability during each iteration of EGO. After iteration 3, the optimization algorithm converges to a stable value of the failure probability estimates.



(a) Experimental design $\boldsymbol{\mathcal{X}}$ for $\mathcal{G}$



(b) Experimental design $\boldsymbol{\mathcal{T}}$ in $\mathcal{D}_{\boldsymbol{\Theta}}$

(c) Evolution of the estimate of the failure probability bounds

Figure 7.7: Toy function – parametric p-boxes – single realization of the ISRA.

The results of the 50 replications of the analysis are summarized in Table 7.2 in terms of failure probability estimates and experimental design sizes. Further, it is distinguished whether the EGO algorithm is used to simultaneously optimize for $\underline{P}_f$ and $\overline{P}_f$ (left block of values) or whether EGO is performed separately for $\underline{P}_f$ and $\overline{P}_f$ (right block of values). Note that for the separate optimization case, samples from one optimization are not re-used in the other one. In terms of the failure probability, both settings result in accurate estimates. Moreover, the coefficient of variation is close to the expected value of $\text{CoV}\left[P_f\right] = 8.0\%$ and 0.9% for $\underline{P}_f$ and $\overline{P}_f$, respectively, for a Monte Carlo simulation with $n_{\text{MCS}} = 10^6$ samples.

Further, the number of evaluations of the true limit-state function $N^{(1)}$ is similar for the simultaneous and separate optimization cases. The number of samples added during AK-MCS is generally low. At the level of EGO, however, the simultaneous optimization requires more samples ($N^{(\text{EGO})}$) as the separate optimization. Even when adding up the samples for the separate optimization case, $N_{tot}^{(\text{EGO})} = 4 + 4.7 + 2.0 = 10.7$ is less than $N^{(\text{EGO})} = 14.3$. Hence, the separate optimization case is more efficient than the simultaneous optimization. Based on this observation, the remaining application examples are analysed using separate optimizations.

Table 7.2: Toy function – parametric p-boxes – failure probability estimation (results of 50 repetitions of the same analysis with different sample sets and initial experimental designs).

	Simultaneous optimization		Separate optimization	
	$\underline{P}_f$	$\overline{P}_f$	$\underline{P}_f$	$\overline{P}_f$
$P_{f,ref}$	$1.57 \cdot 10^{-4}$	$1.14 \cdot 10^{-2}$	$1.57 \cdot 10^{-4}$	$1.14 \cdot 10^{-2}$
$\mathbb{E}\left[\widehat{P}_f\right]$	$1.58 \cdot 10^{-4}$	$1.14 \cdot 10^{-2}$	$1.58 \cdot 10^{-4}$	$1.14 \cdot 10^{-2}$
$\mathrm{CoV}\left[\widehat{P}_f\right]$	8.0%	1.2%	8.1%	1.2%
$\mathbb{E}\left[N^{(1)}\right]$	$12 + 3.4 = 15.4$		$12 + 3.4 = 15.4$	$12 + 3.2 = 15.2$
$\mathrm{Std}\left[N^{(1)}\right]$	0.9		1.0	0.9
$\mathbb{E}\left[N^{(\mathrm{EGO})}\right]$	$4 + 10.4 = 14.3$		$4 + 4.7 = 8.7$	$4 + 2.0 = 6.0$
$\mathrm{Std}\left[N^{(\mathrm{EGO})}\right]$	1.4		1.2	1.1

Results comparison

The algorithms for free and parametric p-boxes originate from two different paradigms: interval analysis and nested Monte Carlo simulations, respectively. However, the statistics of the two proposed two-level meta-modelling algorithms perform similar in terms of model evaluations. Free p-boxes result in $N^{(1)} = 14.7$ evaluations of the limit-state function whereas parametric p-boxes result in $N^{(1)} = 15.4$ on average. This is intuitively understood due to the identical limit-state surface $\mathcal{G}(\boldsymbol{x}) = 0$ in both settings.

7.6.2 Single-degree-of-freedom oscillator

Problem statement

In order to visualize the effects of non-monotone functions on the failure probability, the following example is analysed. Consider the non-linear undamped single-degree-of-freedom (SDOF) oscillator sketched in Figure 7.8 (Schueremans & Van Gemert 2005a; Echard, Gayton, & Lemaire 2011; Echard, Gayton, Lemaire, & Relun 2013). The corresponding limit-state function reads:

$$g_{\mathrm{SDOF}}\left(r, F_1, t_1, c_1, c_2, m\right) = 3r - \left|\frac{2F_1}{m\omega_0^2}\sin\left(\frac{\omega_0 T_1}{2}\right)\right|, \tag{7.15}$$

where m is the mass, c_1, c_2 are the spring constants of the primary and secondary springs, r is the displacement at which the secondary spring yields, t_1 is the duration of the loading, F_1 is the amplitude of the force and $\omega_0 = \sqrt{\frac{c_1+c_2}{m}}$ is the natural frequency of the oscillator. Failure is defined as $g_{\mathrm{SDOF}} \leq 0$, hence the failure probability reads $P_f = \mathbb{P}\left(g_{\mathrm{SDOF}} \leq 0\right)$.

The input vector is modelled by a mix of probabilistic variables and p-boxes accounting for the different levels of knowledge. The description of the input variables is provided in Table 7.3. It is assumed that the spring stiffnesses and the mass are well-known. Hence $\{c_1, c_2, m\}$ are modelled by precise CDFs. On the other side, knowledge on $\{r, F_1, t_1\}$ is scarce. Hence, these variables are modelled by p-boxes. The two cases of free and parametric p-boxes are distinguished and compared in the following. As seen in Table 7.3, the parametric p-box is characterized by a distribution function with interval-valued mean value. For the case of free p-boxes, the same

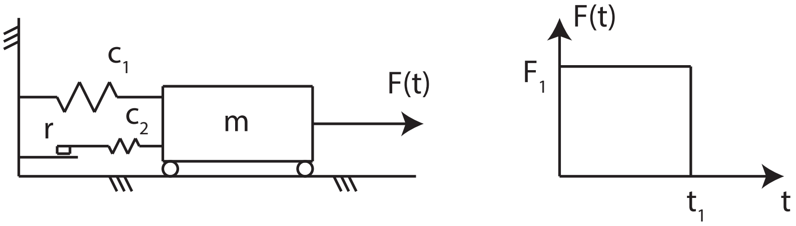


Figure 7.8: SDOF oscillator – geometry sketch and definition of the variables

p-box boundary curves are used as for parametric p-boxes. Note that $F_{\mathcal{N}}(x|\mu,\sigma)$ refers to a Gaussian distribution with mean value μ and standard deviation σ.

Table 7.3: SDOF oscillator – definition of the input vector for the cases of parametric and free p-boxes.

		Parametric p-box		Free p-box	
X_i	Distribution	Mean	Std.	$\underline{F}_{X_i}$	$\overline{F}_{X_i}$
r	Gaussian	$[0.49, 0.51]$	0.05	$F_{\mathcal{N}}(r\|0.51, 0.05)$	$F_{\mathcal{N}}(r\|0.49, 0.05)$
F_1	Gaussian	$[-0.2, 0.2]$	0.5	$F_{\mathcal{N}}(F_1\|0.2, 0.5)$	$F_{\mathcal{N}}(F_1\|-0.2, 0.5)$
t_1	Gaussian	$[0.95, 1.05]$	0.2	$F_{\mathcal{N}}(t_1\|1.05, 0.2)$	$F_{\mathcal{N}}(t_1\|0.95, 0.2)$
c_1	Gaussian	1	0.1	$F_{\mathcal{N}}(c_1\|1, 0.1)$	$= F_{\mathcal{N}}(c_1\|1, 0.1)$
c_2	Gaussian	0.1	0.01	$F_{\mathcal{N}}(c_2\|0.1, 0.01)$	$= F_{\mathcal{N}}(c_2\|0.1, 0.01)$
m	Gaussian	1	0.05	$F_{\mathcal{N}}(m\|1, 0.05)$	$= F_{\mathcal{N}}(m\|1, 0.05)$

Analysis

The settings for the imprecise structural reliability analyses are kept the same as in Section 7.6.1, except for the size of the initial experimental designs. Here, this size is set to $N_0 = 12$ for all meta-models (including EGO). Further, the threshold value in EI is set to $\epsilon_{EI} = 10^{-5}$. For free p-boxes, the auxiliary distributions for $\{r, F_1, t_1\}$ are defined by $F_{\mathcal{N}}(r|0.5, 0.05)$, $F_{\mathcal{N}}(F_1|0, 0.5)$, and $F_{\mathcal{N}}(t_1|1.00, 0.2)$, respectively.

Efficiency of the proposed approaches

Table 7.4 summarizes the results of a single run of the ISRA. The estimates of the failure probability $\widehat{P}_f$ are compared to a reference Monte Carlo simulation with $n_{\mathrm{MCS}} = 10^7$ samples. All four failure probabilities are estimated accurately. Further, the number of evaluations of the limit-state function $N^{(1)}$ is low. The case of parametric p-boxes results in a larger $N^{(1)}$ compared to the free p-box. The reason for this is the added samples during the iterations of the EGO. In comparison, the two levels of the free p-box analysis are independent and hence the total $N^{(1)}$ is lower.

The number of samples in the second-level meta-models are varying more than the first-level meta-models. In the case of parametric p-boxes, a few samples $N^{(\mathrm{EGO})}$ are sufficient to find the optimal distribution parameters. In the case of free p-boxes, $N^{(2)}$ has the same order of magnitude as $N^{(1)}$ because the analysis is of the same type, *i.e.* AK-MCS. However, the difference in $N^{(2)}$ of estimating $\underline{P}_f$ and $\overline{P}_f$ is large, which indicates the different complexity of the limit-state surfaces $\overline{\mathcal{G}} = 0$ and $\underline{\mathcal{G}} = 0$, respectively.

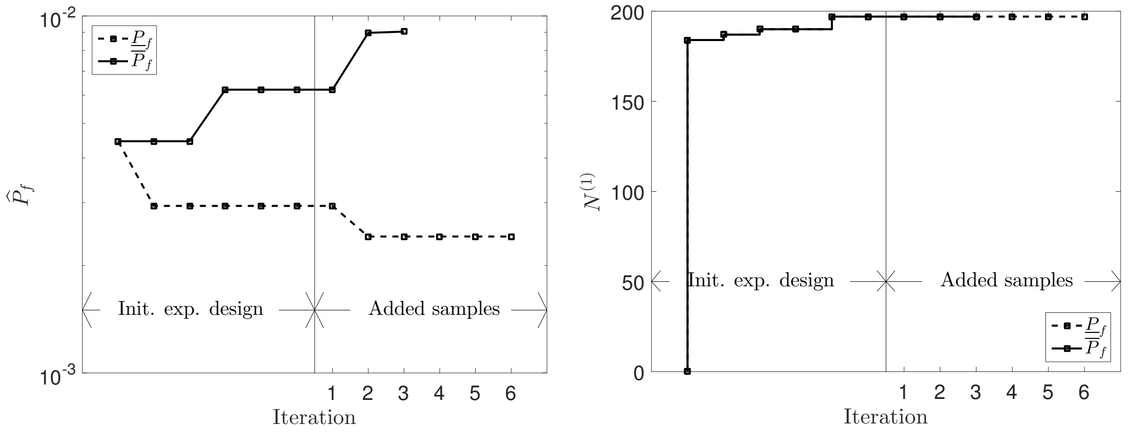
Table 7.4: SDOF oscillator – results of the ISRA (reference solution obtained by Monte Carlo simulation with $n_{\mathrm{MCS}} = 10^7$ samples).

	Parametric p-box		Free p-box	
	$\underline{P}_f$	$\overline{P}_f$	$\underline{P}_f$	$\overline{P}_f$
$P_{f,ref}$	$2.42 \cdot 10^{-3}$	$9.04 \cdot 10^{-3}$	$7.08 \cdot 10^{-4}$	$1.63 \cdot 10^{-2}$
$\widehat{P}_f$	$2.41 \cdot 10^{-3}$	$9.04 \cdot 10^{-3}$	$6.69 \cdot 10^{-4}$	$1.61 \cdot 10^{-2}$
$N^{(1)}$	$12 + 185 = 197$	$12 + 185 = 197$	$12 + 120 = 132$	
$N^{(\mathrm{EGO})}$	$6 + 3 = 9$	$6 + 6 = 12$	-	-
$N^{(2)}$	-	-	$12 + 60 = 82$	$12 + 153 = 165$
μ_r^*	1.02	0.98	-	-
$\mu_{F_1}^*$	-0.01	0.20	-	-
$\mu_{t_1}^*$	0.95	1.05	-	-

For the case of parametric p-boxes, Table 7.4 shows also the optimal mean values for $\{r, F_1, t_1\}$, denoted by $\{\mu_r^*, \mu_{F_1}^*, \mu_{t_1}^*\}$, which result from the EGO optimization. The results confirm the intuition: the largest maximal failure probability is obtained by the smallest yield displacement in combination with a long activation time and an eccentric force ($\mu_{F_1} \approx 0.2$ or $\mu_{F_1} \approx -0.2$).

Evolution of experimental design

As mentioned above, the evolution of the experimental design is different for the two types of p-boxes. For parametric p-boxes, the details are shown in Figure 7.9. The size of $N^{(1)}$ is plotted as a function of the iteration in EGO in Figure 7.9(b). Interestingly, the major part of evaluations are required to estimate the first sample of EGO's initial experimental design. Only a few additional evaluations are needed for the later iterations in EGO in order to refine the estimates of the conditional failure probabilities. Hence, the recycling of limit-state function evaluations is an efficient tool to keep the computational costs at a low level.



(a) Failure probability estimates

(b) Experimental design size $N^{(1)}$

Figure 7.9: Oscillator – parametric p-boxes – convergence of the ISRA.

In the same figure, the evolution of the extreme failure probability estimates is also given (see

Fig. 7.9(a)). Interestingly, the failure probabilities evolve with each iteration in EGO where at the same time $N^{(1)}$ remains constant. In other words, the limit-state surface $\mathcal{G} = 0$ is modelled accurately enough to estimate the extreme failure probabilities.

Effect of non-monotonicity on failure probability

In this example, the two types of p-boxes have the same boundary curves, as seen in Table 7.3. In case the limit-state function was a monotone function, the failure probabilities would be identical. The oscillator function, however, results in non-identical failure probabilities. In fact, the imprecise failure probability of the free p-box case encapsulates the one of the parametric p-box case. As free p-boxes are more general than parametric p-boxes by definition, more extreme failure probabilities are possible. In this example, free p-boxes result in a 3.4-times lower $\underline{P}_f$ and in a 1.8-times larger $\overline{P}_f$ compared to the parametric p-boxes.

7.6.3 Simply supported truss structure

Problem statement

Hurtado (2013) introduced a two-dimensional truss structure, the geometry of which is shown in Figure 7.10. The truss is subjected to seven loads P_i which are modelled with p-boxes. The parametric p-boxes are defined by lognormal distributions with mean value $\mu_{P_i} \in [95, 105]$ kN and standard deviation $\sigma_{P_i} \in [13, 17]$ kN. The geometry of the structure and the material properties are assumed constant. The modulus of elasticity is $E = 200 \cdot 10^9$ Pa, whereas the cross section area varies along the different bars: $A = 0.00535$ m^2 for bars marked by $\bullet$, $A = 0.0068$ m^2 for bars marked by $\circ$, and $A = 0.004$ m^2 for the remaining bars.

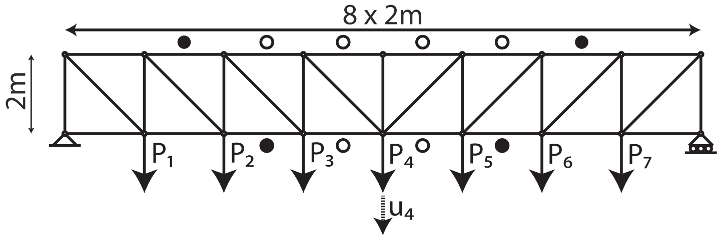


Figure 7.10: Simply-supported truss structure – sketch of the geometry and definition of the input variables.

In the following analysis, a second scenario is considered, where the former parametric p-boxes are modelled by free p-boxes. As in the other examples, the boundary curves of the p-boxes should coincide, hence:

$$\underline{F}_{P_i}(p_i) = \min_{\mu_{P_i} \in [95,105] \text{ kN}, \sigma_{P_i} \in [13,17] \text{ kN}} F_{P_i}(p_i | \mu_{P_i}, \sigma_{P_i}),$$

$$\overline{F}_{P_i}(p_i) = \max_{\mu_{P_i} \in [95,105] \text{ kN}, \sigma_{P_i} \in [13,17] \text{ kN}} F_{P_i}(p_i | \mu_{P_i}, \sigma_{P_i}).$$

In the context of structural reliability analysis, the limit-state function is defined as:

$$g_{\text{truss}}(\boldsymbol{p}) = 0.029 \text{ m} - u_4(\boldsymbol{p}), \tag{7.16}$$

where u_4 is the deflection of the truss at midspan, as indicated in Figure 7.10. Then, the failure probability describes the probability that the deflection of the truss exceeds 0.029 m, *i.e.* $P_f = \mathbb{P}(u_4(\boldsymbol{P}) \geq 0.029 \text{ m})$.

Analysis

The deflection of the truss is computed by a finite element model (FEM), which is implemented in the software framework UQLab (Marelli & Sudret 2014). The FEM interprets each bar as a bar element, whereas the loads are modelled as point loads at the intersections of the corresponding bars as indicated in Figure 7.10.

The ISRA settings are kept the same as in the previous examples. However, the initial experimental design is set to $N_0^{(1)} = N_0^{(2)} = N_0^{(\mathrm{EGO})} = 12$ Latin-hypercube samples and $\epsilon_{EI} = 10^{-5}$. For free p-boxes, the auxiliary distributions $\widetilde{X}_i$ are chosen as lognormal distributions with mean value $\mu_{\widetilde{P}_i} = 100$ kN and standard deviation $\mu_{\widetilde{P}_i} = 15$ kN.

Results

Table 7.5 summarizes the results in terms of failure probability estimates and number of model evaluations. The failure probabilities are estimated accurately for the cases of parametric and free p-boxes. The numbers of FEM evaluations $N^{(1)}$ are low for both cases, too. Hence, the proposed multi-level meta-modelling techniques yield efficient algorithms for ISRA in this application example.

Table 7.5: Simply-supported truss structure – results of the ISRA (reference solution obtained by MCS with $n_{MC} = 10^7$ samples).

	Parametric p-box		Free p-box	
	$\underline{P}_f$	$\overline{P}_f$	$\underline{P}_f$	$\overline{P}_f$
$P_{f,ref}$	$2.49 \cdot 10^{-4}$	$7.62 \cdot 10^{-2}$	$2.21 \cdot 10^{-4}$	$9.22 \cdot 10^{-2}$
$\widehat{P}_f$	$2.64 \cdot 10^{-4}$	$7.62 \cdot 10^{-2}$	$2.36 \cdot 10^{-4}$	$9.16 \cdot 10^{-2}$
$N^{(1)}$	$12 + 130 = 142$	$12 + 129 = 141$	$12 + 106 = 118$	$12 + 106 = 118$
$N^{(\mathrm{EGO})}$	$12 + 174 = 186$	$12 + 40 = 52$	-	-
$N_*^{(\mathrm{EGO})}$	$12 + 26 = 38$	$12 + 4 = 16$	-	-
$N^{(2)}$	-	-	$12 + 116 = 128$	$12 + 192 = 204$
$\mu_{P_i}^*$	95 kN	105 kN	-	-
$\sigma_{P_i}^*$	13 kN	17 kN	-	-

In the case of parametric p-boxes, the minimum failure probability is obtained by $\mu_{P_i}^* = 95$ kN and $\sigma_{P_i}^* = 13$ kN, whereas the maximum failure probability is obtained by $\mu_{P_i}^* = 105$ kN and $\sigma_{P_i}^* = 17$ kN, $i = 1, \ldots, 7$, as one would expect. EGO is capable of identifying these two extreme cases in this application. The number of iterations required to do so are indicated by $N_*^{(\mathrm{EGO})}$ in Table 7.5. The values are lower than the final number of samples $N^{(\mathrm{EGO})}$, which represent the number of samples used to terminate the EGO algorithm. The additional samples beyond $N_*^{(\mathrm{EGO})}$ show the exploratory behaviour of the EGO algorithm in this application example.

Hurtado (2013) analysed the same truss structure for the parametric p-box case. However, due to its use of a Monte Carlo simulation on the second level of the algorithm, the published failure probability range is narrower, *i.e.* $P_f \in \left[3.2 \cdot 10^{-3}, 2.37 \cdot 10^{-2}\right]$, than reported in Table 7.5. This indicates the importance of using a proper optimization algorithm in order to find the extreme failure probabilities.

7.6.4 Two-dimensional truss

Problem statement

Consider again the two-dimensional truss structure described in Section 5.4.3 and whose geometry is shown again in Figure 7.11. The deflection at midspan u is modelled as a function of the deterministic geometry and a random ten-dimensional input vector consisting of Young's moduli, cross section areas and six loads. The limit-state function is defined as $\mathcal{G}(\boldsymbol{x}) = u_{\text{adm}} - \mathcal{M}(\boldsymbol{x})$, where the admissible deflection is $u_{\text{adm}} = 12$ cm in this example. In order to obtain a reasonable range of the imprecise failure probability, the ten input parameters are defined in Table 7.6, which is different from the definition in Table 6.3.

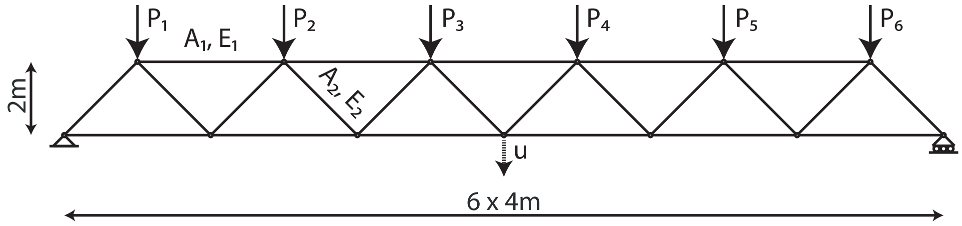


Figure 7.11: Two-dimensional truss – sketch of the geometry, material parameters $\{A_i, E_i\}$ and loads P_i (see also Figure 5.8).

Table 7.6: Two-dimensional truss – input p-boxes for ISRA.

Variable		Distribution	Mean	CoV
A_1	[m^2]	Lognormal	$[0.98,\ 1.02] \cdot 2 \cdot 10^{-3}$	$[0.98,\ 1.02] \cdot 10\%$
A_2	[m^2]	Lognormal	$[0.98,\ 1.02] \cdot 10^{-3}$	$[0.98,\ 1.02] \cdot 10\%$
E_1, E_2	[Pa]	Lognormal	$[0.98,\ 1.02] \cdot 2 \cdot 10^{11}$	$[0.98,\ 1.02] \cdot 10\%$
$P_1, \ldots, P_6$	[N]	Gumbel	$[0.95,\ 1.05] \cdot 5 \cdot 10^{4}$	$[0.95,\ 1.05] \cdot 15\%$

Three scenarios are considered: (i) all variables are modelled as free p-boxes, (ii) all variables are modelled as parametric p-boxes, and (iii) $P_1, \ldots, P_6$ are modelled as free p-boxes whereas E_1, E_2, A_1, A_2 are modelled by parametric p-boxes. As seen in previous examples, the bounds of the free p-boxes are generated by the envelope of the parametric p-boxes (see also Eq. (6.3)). Note that the range for the mean value and coefficient of variation of all p-boxes are chosen so that the structural reliability analysis results in manageable failure probabilities when using Monte Carlo simulation. In practical applications, the ranges could be considerably wider.

Analysis

The same settings for ISRA are used as described in Section 7.6.1, except the initial experimental design for AK-MCS and EGO are chosen to have $N_0^{(1)} = N_0^{(2)} = N_0^{(\text{EGO})} = 20$ samples. The three cases of different input p-boxes are analysed using the three algorithms proposed in (i) Section 7.3, (ii) Section 7.4, and (iii) Section 7.5, respectively. For any free p-box, the auxiliary distribution $\widetilde{X}_i$ is defined by the given distribution with mid-range mean value and maximum coefficient of variation. The auxiliary distributions are used on both levels of the two-level meta-modelling approach to improve convergence.

Results

The results of a single run analysis are summarized in Table 7.7. The ranges of failure probabilities are similar for all three cases. In other words, the decision on free or parametric p-boxes does not affect the failure probability considerably in this example application. The reason for this behaviour lies in the monotonicity of the limit-state function as discussed previously in Section 6.5.2. As a result the input CDFs leading to the extreme failure probabilities are very similar for all three cases.

Table 7.7: Two-dimensional truss – results of the ISRA (reference solution obtained by Importance Sampling with $n = 10^6$ samples).

Case		$\underline{P}_f$	$\overline{P}_f$
(i) free p-boxes	$P_{f,ref}$	$1.11 \cdot 10^{-4}$	$1.19 \cdot 10^{-2}$
	$\widehat{P}_f$	$1.35 \cdot 10^{-4}$	$1.06 \cdot 10^{-2}$
	$N^{(1)}$	$20 + 114 = 134$	$20 + 114 = 134$
	$N^{(2)}$	$20 + 81 = 101$	$20 + 292 = 312$
(ii) parametric p-boxes	$P_{f,ref}$	$1.18 \cdot 10^{-4}$	$1.12 \cdot 10^{-2}$
	$\widehat{P}_f$	$1.34 \cdot 10^{-4}$	$9.91 \cdot 10^{-3}$
	$N^{(1)}$	$20 + 269 = 289$	$20 + 169 = 189$
	$N^{(\mathrm{EGO})}$	$20 + 31 = 51$	$20 + 8 = 28$
(iii) mixed p-boxes	$P_{f,ref}$	$1.12 \cdot 10^{-4}$	$1.18 \cdot 10^{-2}$
	$\widehat{P}_f$	$1.19 \cdot 10^{-4}$	$1.18 \cdot 10^{-2}$
	$N^{(1)}$	$20 + 311 = 331$	$20 + 286 = 306$
	$N^{(2)}$	$20 + 83 = 103$	$20 + 181 = 201$
	$N^{(\mathrm{EGO})}$	$20 + 19 = 39$	$20 + 5 = 25$

The estimated bounds for the failure probability are accurate in all three cases. The differences to the reference failure probabilities $P_{f,ref}$ originate from the finite size of the MC sample, *i.e.* $n_{\mathrm{MCS}} = 10^6$. However, in terms of efficiency, the differences between the three cases is larger. Case (i) results in fewer evaluations of the exact limit-state function $\mathcal{G}$ compared to the other two cases, as seen by comparing the values of $N^{(1)}$. The reason for the larger $N^{(1)}$ in Cases (ii) and (iii) lies in the nested algorithm and its iterative refinement of the first-level meta-model in the presence of parametric p-boxes. The largest number of limit-state function evaluations are found in Case (iii) due to the fact that this case poses the most complex analysis among the three ones discussed.

In terms of the second-level meta-model for free p-boxes, $N^{(2)}$ lies in the same order of magnitude as $N^{(1)}$ but varies more. For EGO in the context of parametric p-boxes, the number of iterations in the optimization is in the range $N^{(\mathrm{EGO})} \in [25, 51]$. Considering that the optimization domain has a dimensionality of $M = 20$ and $M = 8$ for Cases (ii) and (iii), respectively, $N^{(\mathrm{EGO})}$ is remarkably low meaning that EGO is efficient to find the extreme failure probabilities.

7.7 Comparison of free and parametric p-box approaches

The main features of the two proposed algorithms are shown in Table 7.8. In the case of free p-boxes, the main solving algorithm is AK-MCS, whereas in the case of parametric p-boxes, it is AK-MCS in combination with EGO. Both types of p-boxes require two levels on meta-models. The dimensionality of these meta-models is M except for $\mathcal{M}^{(\mathrm{K})}$ in EGO, where the dimensionality is $|\boldsymbol{\Theta}|$. EGO handles the optimization algorithms for the two bounds of the failure probability. In the case of free p-boxes, EGO could also be applied in the optimizations in Eq. (7.4). The number of EGO algorithms would, however, be much higher than for parametric p-boxes, *i.e.* $2 \cdot N^{(2)} \gg 2$.

Table 7.8: ISRA – comparison of free and parametric p-box approaches.

	Free p-box	Parametric p-box
P-box	$\underline{F}_{\boldsymbol{X}},\ \overline{F}_{\boldsymbol{X}}$	$F_{\boldsymbol{X}}(\boldsymbol{x}\vert\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$
Main solver	AK-MCS	AK-MCS + EGO
Kriging models	$\mathcal{G}(\boldsymbol{x}),\ \underline{\mathcal{G}}(\boldsymbol{c}),\ \overline{\mathcal{G}}(\boldsymbol{c})$	$\mathcal{M}^{(\mathrm{K})}(\boldsymbol{\theta}),\ \mathcal{G}^{(\mathrm{K})}(\boldsymbol{x})$
Kriging dimensionality	$\vert\boldsymbol{X}\vert = \vert\boldsymbol{C}\vert = M$	$\vert\boldsymbol{\Theta}\vert,\ \vert\boldsymbol{X}\vert = M$
Optimization equation	Eq. (7.4)	Eq. (7.6)
Number of optimizations	$2 \cdot N^{(2)}$	2

The ISRA approaches are similar to the imprecise uncertainty propagation (IUP) algorithms in Chapter 6. In fact, when replacing $\mathcal{M}$ and MCS by $\mathcal{G}$ and AK-MCS, respectively, the IUP approach for free p-boxes transforms to the ISRA approach presented in this chapter. The meta-models calibrated in the analyses have the same complexity (*i.e.* dimensionality) for first- and second-level meta-models. The difference lies in the fact that PCE models are used in the case of IUP. These, however, might be replaced by Kriging meta-models to increase the similarities to ISRA.

For parametric p-boxes, the ISRA approach is different from the IUP approach. In ISRA, the statistics of interest is obtained by two different meta-models (used in AK-MCS and EGO), whereas in uncertainty propagation, the response p-box is obtained through a single meta-model (*i.e.* augmented PCE). In other words, the complexity of the augmented PCE model in IUP is split into two meta-models in ISRA. The reason for this change is the efficiency of Kriging models in the context of adaptive experimental designs and structural reliability analysis.

CHAPTER 8

Imprecise sensitivity analysis

In this chapter, sensitivity analysis is discussed in the presence of p-boxes. In particular, two approaches existing in the context of probability theory are extended to imprecise probabilities. Pinching analysis provides a sensitivity measure for free p-boxes as well as parametric p-boxes. Imprecise Sobol' indices is an extension of the traditional Sobol' indices for the case of parametric p-boxes and has been published in Schöbi & Sudret (2017a).

8.1 State of the art

8.1.1 Local versus global sensitivity

Sensitivity analysis (SA) examines the impact of the uncertainty in $X_i \in \boldsymbol{X}$ onto the uncertainty in the response quantity Y. This is of importance in practice where the relation between input variables and quantity of interest Y is often implicitly defined by a complex computer code. A large number of SA methods can be found in the literature, including reviews of several methods in Saltelli, Chan, & Scott (2000), Helton, Johnson, Sallaberry, et al. (2006), and Xu & Gertner (2008). The methods can be separated into two broad classes, namely *local* and *global* sensitivity analyses.

Local sensitivity analysis examines how small variations in the input variables X_i in the vicinity of a nominal value influence the response variable Y. As a result, many methods are based on the computation of the gradient of the computational model around the nominal value. Efficient computation of the gradient is achieved by applying finite-difference schemes, direct differentiations, and adjoint differentiation (Cacuci 2003). Popular local sensitivity measures include importance factors in FORM, the Cotter method (Cotter 1979), the perturbation method, and one-at-a-time methods. Note that these methods generally-speaking do not take into account the interactions between input variables and hence are of limited use in practical examples.

Opposed to local sensitivity analysis examining the small-scale variation around a nominal value, global sensitivity analysis focuses on the entire variation of an input variable X_i or a combination of several input variables. Hence, the uncertainty in Y is analysed as a function of the variability in the input variable X_i. Saltelli, Chan, & Scott (2000) group the various methods into *regression-based* and *variance-based* methods:

(i) Regression-based methods exploit the linear regression of the model response on the input vector. Hence, these methods are useful when the computational model is a linear or nearly linear function of the input variables $\boldsymbol{X}$. On the other side, these methods are inappropriate for non-monotone and highly non-linear computational models (Saltelli, Chan, & Scott 2000). Popular methods in this group include the standard regression coefficient (SRC), the input/output Pearson correlation coefficients, partial correlation coefficients (PCC), standard rank regression coefficients (SRRC), and partial rank regression coefficients (PRCC).

(ii) Variance-based methods aim at decomposing the variance in Y as a sum of the variances of the input contributions. Therefore, these methods are known as ANOVA (*i.e.* ANalysis Of VAriance) techniques in statistics (Efron & C. Stein 1981). Popular methods include the correlation ratios (McKay 1995), Fourier amplitude sensitivity test (FAST) (Cukier 1973; Cukier, Levine, & Shuler 1978; Saltelli, Tarantola, & Chan 1999; Mara 2009), and Sobol' indices (Sobol' 1993, 2001; Archer, Saltelli, & Sobol' 1997; Sobol' & Kucherenko 2009).

Other global SA methods that do not fit in either of the groups above include the Morris method (Morris 1991) and derivative-based global sensitivity measures (DGSM) (Sobol' & Kucherenko 2009, 2010; Kucherenko et al. 2009; Lamboni et al. 2013). Further, there are a number of methods based on specific experimental designs, such as fractional factorial designs (Saltelli & Sobol' 1995) and Plackett-Burman designs (Beres & Hawkins 2001).

8.1.2 Meta-models

The SA methods usually require a large number of computational model runs for different realizations of the input vector $\boldsymbol{X}$. Hence, such analyses can become intractable when the computational model is expensive-to-evaluate. Then, the use of meta-models is a popular solution to lower the total computational costs, as also seen previously in this thesis in the context of uncertainty propagation and structural reliability analysis.

The trivial approach of using meta-models is by the implementation of a sequential algorithm: (i) calibrate a meta-model as previously done in an uncertainty propagation problem and (ii) conduct a SA using Monte Carlo-based estimates on the meta-model. There are, however, more elaborate ways to use meta-models in this context. Sudret (2006, 2008) computes Sobol' indices as a post-processing algorithm on an existing PCE model. Furthermore, Sudret & Mai (2015) compute DGSM based on a sparse PCE model.

8.1.3 Imprecise probabilities

In the context of imprecise probabilities, the aim of sensitivity analysis is not straightforward. Sensitivity can be measured in terms of (i) aleatory uncertainty, (ii) epistemic uncertainty, or (iii) a combination of both (such as in the present case of p-boxes). In this thesis, the contributions of aleatory and epistemic uncertainty are not mixed due to the different modelling with probabilistic variables and intervals, respectively.

The SA methods presented above are typically defined as a measure of variability in the sense of aleatory uncertainty. In the context of epistemic uncertainty and hence imprecise probabilities, these methods can be applied/extended under some rigorous assumptions. However, the number of publications for *imprecise sensitivity analysis* (ISA) is small. The following publications include sensitivity analysis and epistemic uncertainty:

- Krzykacz-Hausmann (2006) examines Bayesian hierarchical models. The author computes sensitivity indices for each source of epistemic uncertainty in the overall joint epistemic and aleatory uncertainty domain. In other words, sensitivity indices are computed based on the expected response value over aleatory uncertainties.

- Helton, Johnson, Sallaberry, et al. (2006) provide a survey of available methods for uncertainty quantification and sensitivity analysis. They point out that epistemic uncertainties in ISA have not been considered adequately yet in 2006.

- Helton, Johnson, Oberkampf, & Sallaberry (2006) describe three methods for sensitivity analysis in the presence of evidence theory (belief and plausibility measures). The three methods range from an initial exploratory analysis over an incremental effects analysis to a variance-based SA.

- Guo & Du (2012) describe a sensitivity analysis method to identify the most important epistemic variables with respect to the design performance (in imprecise structural reliability analysis). They make use of the Kolmogorov-Smirnov distance and one-at-a-time strategy as importance measure.

- Sankararaman & Mahadevan (2013) discuss global sensitivity analysis for Bayesian hierarchical models. They use auxiliary variables to describe the aleatory variability in each variable.

- C. Li & Mahadevan (2016a,b) describe Sobol' indices in the presence of input uncertainty (aleatory and epistemic; modelled as Bayesian hierarchical model) and model uncertainty (epistemic). Additionally, they propose techniques to analyse time series efficiently.

- Oberguggenberger & Fellin (2005) and Oberguggenberger, King, & Schmelzer (2009) discuss SA for different concepts of imprecise probabilities, including random sets, fuzzy sets, p-boxes and CDFs. The focus lies on pinching strategies (keeping one variable constant and see the effect on the response structure) and Sobol' indices.

- Ferson & Troy Tucker (2006) use pinching to estimate the sensitivity of variables. In particular, the sensitivity is measured by the ratio of the areas between the lower and upper boundaries of a p-box.

- Hall (2006) considers that the input variables are modelled by intervals, free, or parametric p-boxes. Then, the author examines variance-based sensitivity analysis (VBSA), partial expected value of perfect information (partial EVPI) and aggregate uncertainty entropy-based sensitivity indices (AUEBSI).

As previously seen in the context of IUQ and ISRA, the use of imprecise probability concepts increases the computational cost, possibly making ISA intractable. Hence, the ISA methods presented in the literature are only applicable to inexpensive-to-evaluate computational models, which are often analytical toy functions. Therefore, in this chapter, the meta-modelling tools introduced previously are extended and applied to ISA. In the following, the focus lies on the pinching algorithm in the context of free and parametric p-boxes. Additionally, the Sobol' indices are extended to the new concept of imprecise Sobol' indices for the case of parametric p-boxes.

8.2 Free p-box

8.2.1 Pinching

Idea

In this section, sensitivity analysis is used to estimate the value of additional information when reducing the epistemic uncertainty in the response variable Y of a computational model $\mathcal{M}$. In other words, it is examined how much the imprecision of the input variables X_i affects the imprecision of the response Y.

A popular strategy herein is to compare the uncertainty in Y before and after *pinching* an input variable (Ferson & Troy Tucker 2006; Oberguggenberger, King, & Schmelzer 2009). Pinching in this context means that the epistemic uncertainty in an input variable X_i (or a group of input variables) is replaced by a deterministic value. The pinched variables might, however, still inherit aleatory uncertainty. In other words, the pinching transforms a set of p-box-modelled input variables into probabilistic input variables.

In this context, the value of information (VoI) can be manifold, depending on the goals of the sensitivity analysis and generally the statistics of interest. As a general concept, however, the value of information can be defined as (Ferson & Troy Tucker 2006):

$$\mathrm{VoI}\left[B, T\right] = 1 - \frac{\mathrm{unc}\left(T\right)}{\mathrm{unc}\left(B\right)}, \tag{8.1}$$

where $\mathrm{unc}\left(\cdot\right)$ is a function measuring uncertainty, B is the base value and T is the pinched value of the statistics of interest. The larger the VoI, the larger is the influence of epistemic uncertainty in the considered variable. The VoI is limited to $\mathrm{VoI} \in [0, 1]$ because the pinched input vector consists of less epistemic uncertainty, by definition, compared to the original input vector.

Pinching response p-boxes

In the context of imprecise uncertainty propagation, the epistemic uncertainty can be measured by the area between the bounds of the free p-box. Then, the VoI can be defined as the following ratio of p-box areas:

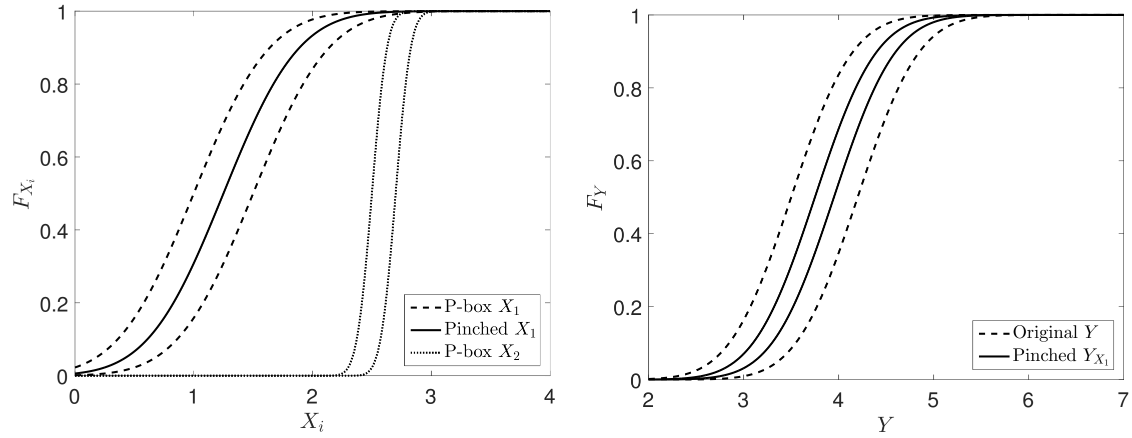
$$\mathrm{VoI}\left[Y, Y_{\check{X}_i}\right] = 1 - \frac{A_{Y_{\check{X}_i}}}{A_Y}, \tag{8.2}$$

where Y is the original response variable, $Y_{\check{X}_i}$ is the corresponding response variable when $\check{X}_i$ is pinched, and A_Y and $A_{Y_{\check{X}_i}}$ are the corresponding p-box areas, respectively. For a given p-box, this p-box area is defined as:

$$A_Y = \int_{\mathcal{D}_Y} \left(\overline{F}_Y(y) - \underline{F}_Y(y)\right) \mathrm{d}y. \tag{8.3}$$

The integral can be computed numerically by using *e.g.* standard quadrature method or a large number of MC samples of each boundary curve.

Figure 8.1 sketches the idea of pinching. Figure 8.1(a) shows two input variables modelled by p-boxes, *i.e.* X_1 (dashed lines) and X_2 (dotted lines). Additionally, the solid line marks the pinched variable $\check{X}_1$, which consists a p-box with no epistemic uncertainty, *i.e.* of a probabilistic variable. The original and pinched response p-box are compared in Figure 8.1(b). Note that the computational model is $y = \mathcal{M}\left(\boldsymbol{x}\right) = x_1 + x_2$. As seen in Figure 8.1(b), the pinching of X_1 reduces the epistemic uncertainty in the response p-box considerably. The corresponding value of information results in $\mathrm{VoI}\left[Y, Y_{\check{X}_1}\right] = 1 - 0.2/0.7 = 0.7143 \approx 71\ \%$.



(a) Input variables

(b) Response p-box

Figure 8.1: Pinching free p-boxes – comparison of p-box areas.

Pinching failure probabilities

In the context of imprecise structural reliability analysis, the statistics of interest is the failure probability. Hence, a VoI equivalent to Eq. (8.2) can be formulated as follows:

$$\text{VoI}\left[Y, Y_{\check{X}_i}\right] = 1 - \frac{\Delta\beta_{\text{HL}}\left(Y_{\check{X}_i}\right)}{\Delta\beta_{\text{HL}}\left(Y\right)}, \tag{8.4}$$

where β_{HL} is the Hasofer-Lind reliability index defined as $\beta_{\text{HL}} = -\Phi^{-1}\left(P_f\right)$ and $\Delta\beta_{\text{HL}}\left(Y\right) = \Phi^{-1}\left(\overline{P}_f\right) - \Phi^{-1}\left(\underline{P}_f\right)$ describes the range of the reliability index for a p-box in Y. Note that this is not a unique solution for the case of imprecise structural reliability analysis. However, the Hasofer-Lind reliability index describes a value that is widely accepted as a measure of the reliability level with typical values of $\beta_{\text{HL}} = \{1,\, 3,\, 5\}$ for $P_f = \left\{1.587 \cdot 10^{-1},\, 1.350 \cdot 10^{-3},\, 2.867 \cdot 10^{-7}\right\}$, respectively.

Consider again the example in Figure 8.1 and assume that the failure probability of interest is $P_f = \mathbb{P}\left(Y \leq 3\right)$. Then according to Eq. (8.4), the value of information when pinching X_1 reads $\text{VoI}\left[Y, Y_{\check{X}_1}\right] = 1 - \Delta\beta_{\text{HL}}\left(Y_{\check{X}_1}\right) / \Delta\beta_{\text{HL}}\left(Y\right) = 1 - 0.3922/1.3728 = 0.7143 \approx 71$ %. In this special case, the two measures of Eqs. (8.2) and (8.4) lead to identical VoI values. However, in the general case where $\underline{F}_Y$ and $\overline{F}_Y$ have arbitrary shapes, this conclusion may not hold. For further details, see also the application example in Section 8.4.2.

Definition of $\check{X}_i$

The straightforward definition of a pinched variable consists of using the average CDF curve:

$$F_{\check{X}_i} = \frac{1}{2}\left(\underline{F}_{X_i}(x_i) + \overline{F}_{X_i}(x_i)\right). \tag{8.5}$$

The definition of the pinched distribution is, however, not unique. Moreover, the results of a pinching analysis can depend on the definition of the CDF of the pinched variable. In particular, the VoI is a function of $F_{\check{X}_i}(x_i) \in \left[\underline{F}_{X_i}(x_i), \overline{F}_{X_i}(x_i)\right]$, $\forall x_i \in \mathcal{D}_{X_i}$. This dependency is pronounced when the computational model is non-linear and non-monotone. Then, different definitions of $\check{X}_i$ may result in different VoI. In extreme cases where the dependency is strong, the examination of a bunch of different definitions of $F_{\check{X}_i}\left(x_i\right)$ might be more suitable to measure the VoI than the examination of a single one, which is by default the average of the bounds $\underline{F}_{X_i}$ and $\overline{F}_{X_i}$ defined in Eq. (8.5).

8.2.2 Efficient pinching

The estimation of the VoI requires the computation of a series of response p-boxes as a function of slightly varying input variables X_i. Hence, the approaches presented in Sections 6.2 and 7.3 can be adopted to solve Eqs. (8.2) and (8.4), respectively.

VoI with pinched p-boxes

In order to estimate the pinched response p-boxes efficiently, the two-level meta-modelling approach shown in Figure 6.3 can be adopted. The first-level meta-model can be kept the same as in imprecise uncertainty propagation. The second-level meta-model, however, is modified to account for the pinching. In fact when pinching X_i, Eq. (6.12) transforms into:

$$\underline{y}_{\check{X}_i} = \underline{\mathcal{M}}_{\check{X}_i}(\boldsymbol{c}) = \min_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{c}_{\check{X}_i}}} \mathcal{M}(\boldsymbol{x}), \qquad \overline{y}_{\check{X}_i} = \overline{\mathcal{M}}_{\check{X}_i}(\boldsymbol{c}) = \max_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{c}_{\check{X}_i}}} \mathcal{M}(\boldsymbol{x}), \tag{8.6}$$

where the optimization domain is defined as:

$$\mathcal{D}_{\boldsymbol{c}_{\check{X}_i}} = [\underline{x}_1(c_1), \overline{x}_1(c_1)] \times \ldots \times [\check{x}_i\,(c_i)] \times \ldots \times [\underline{x}_M(c_M), \overline{x}_M(c_M)]\,, \tag{8.7}$$

where the p-box of variable X_i degenerates to a probabilistic variable, *i.e.* $\check{x}_i\,(c_i) = \underline{x}_i\,(c_i) = \overline{x}_i\,(c_i)$ and $\underline{F}_{\check{X}_i}\,(\check{x}_i) = \overline{F}_{\check{X}_i}\,(\check{x}_i)$ for all $\check{x}_i \in \mathcal{D}_{\check{X}_i}$.

Considering the computational models in Eq. (8.6), Figure 8.2 shows the flowchart of the two-level meta-modelling algorithm for pinching variables. The pinched input vector is denoted by $\boldsymbol{X}_{\check{X}_i}$, which is located between the first and the second level in the flowchart. After the computation of the response free p-box $Y_{\check{X}_i}$ (characterized by $\overline{Y}_{\check{X}_i}$ and $\underline{Y}_{\check{X}_i}$), the area $A_{Y_{\check{X}_i}}$ is computed and used in Eq. (8.2).

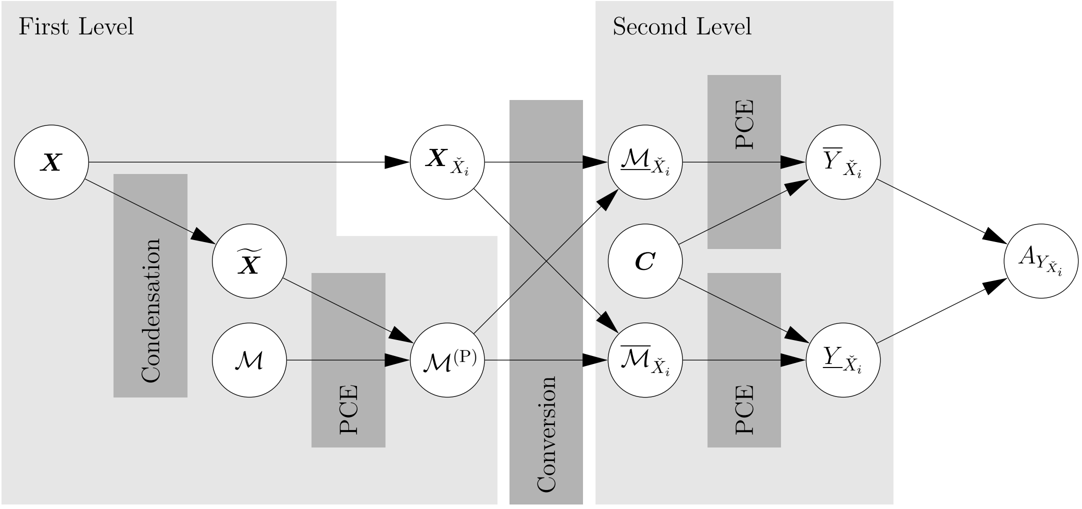


Figure 8.2: Pinching of free p-boxes – two-level meta-modelling approach for computation of $A_{Y_{\check{X}_i}}$.

Analysing the global meta-modelling approach, one can see that the two levels of meta-modelling are conducted in series. In other words, the second-level meta-model does not interact with the first-level meta-model. Hence, the second-level meta-model can be calibrated for each pinched variable $\check{X}_i$, $i = 1, \ldots, M$ without recomputing the first-level meta-model $\mathcal{M}^{(\mathrm{P})}$. In this setting, the first-level meta-model is recycled during the estimation of $A_{Y_{\check{X}_i}}$, $i = 1, \ldots, M$, which limits the total computational costs to the ones in the (unpinched) imprecise uncertainty propagation analysis.

VoI with pinched failure probabilities

When the value of information is connected to the range of the reliability index β_{HL}, the two-level meta-modelling framework presented in Figure 7.2 can be adopted straightforwardly with little modifications. In particular, Eq. (7.4) is transformed to:

$$\underline{\mathcal{G}}_{\check{X}_i}(\boldsymbol{c}) = \min_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{c}_{\check{X}_i}}} \mathcal{G}(\boldsymbol{x}), \qquad \overline{\mathcal{G}}_{\check{X}_i}(\boldsymbol{c}) = \max_{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{c}_{\check{X}_i}}} \mathcal{G}(\boldsymbol{x}), \tag{8.8}$$

where the domain $\mathcal{D}_{\boldsymbol{c}_{\check{X}_i}}$ is defined in Eq. (8.7).

The modified flowchart is shown in Figure 8.3. Similar to the computation of $A_{Y_{\check{X}_i}}$, the second-level meta-model is computed after the first-level meta-model. Hence, the same first-level meta-model $\mathcal{G}^{(\mathrm{K})}$ can be used to compute all $\Delta\beta_{\mathrm{HL}}\left(Y_{\check{X}_i}\right)$, $i = 1, \ldots, M$ on the second level. This reduces the total computational costs as it limits the number of runs of the limit-state function $\mathcal{G}$.

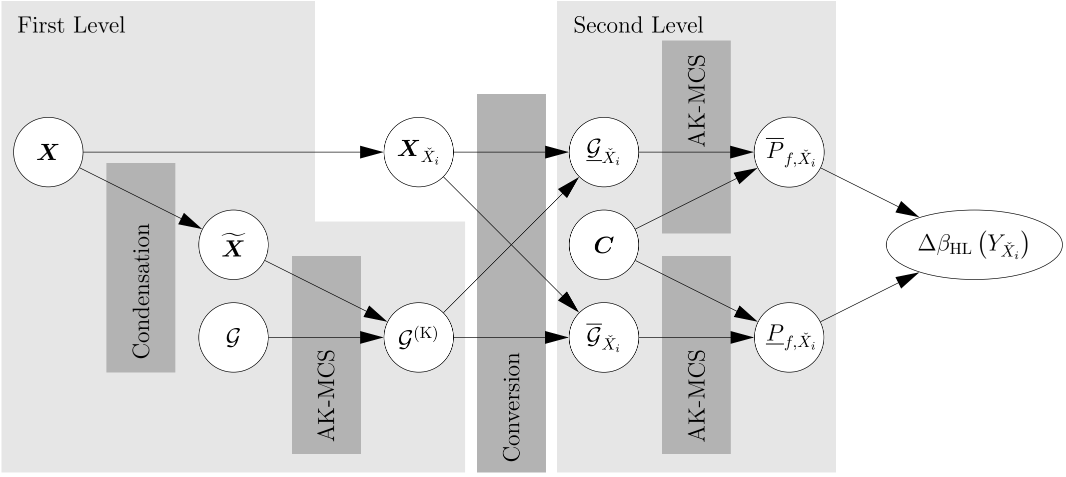


Figure 8.3: Pinching of free p-boxes – two-level meta-modelling approach for computation of $\Delta\beta_{\mathrm{HL}}\left(Y_{\check{X}_i}\right)$.

8.3 Parametric p-boxes

8.3.1 Pinching

In the previous section, pinching is introduced for the case of free p-boxes. The concept of pinching can be applied to parametric p-boxes in an analogous manner, when considering the boundary curves of the response parametric p-box (see also Eq. (2.20)):

$$\underline{F}_Y(y) = \min_{\boldsymbol{\theta}_Y \in \mathcal{D}_{\boldsymbol{\Theta}_Y}} F_Y(y|\boldsymbol{\theta}_Y), \qquad \overline{F}_Y(y) = \max_{\boldsymbol{\theta}_Y \in \mathcal{D}_{\boldsymbol{\Theta}_Y}} F_Y(y|\boldsymbol{\theta}_Y), \tag{8.9}$$

where F_Y denotes the distribution family of Y with distribution parameters $\boldsymbol{\theta}_Y \in \mathcal{D}_{\boldsymbol{\Theta}_Y}$. Note that in the general case, $\underline{F}_Y$ and $\overline{F}_Y$ have non-analytical shapes and values.

8.3.2 Efficient pinching

VoI with pinched p-boxes

Using the developments of Section 6.4, an augmented PCE model can be calibrated for the full response parametric p-box Y. In order to apply Eq. (8.2), a particular realization of Y is computed by setting the distribution parameters $\boldsymbol{\Theta}$ to a fixed value $\boldsymbol{\theta}$ and by conducting a Monte Carlo simulation for the (aleatory) variables $\boldsymbol{X}(\boldsymbol{\theta})$, or their transformed counterpart $\boldsymbol{C}$. Let us denote the resulting variable $Y_{\boldsymbol{\theta}}$. Then, the response variable $Y_{\boldsymbol{\theta}}$ can be obtained by using the augmented PCE model:

$$Y_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \mathcal{M}^{(\mathrm{P})}_{(\mathrm{aug})}(\boldsymbol{C}, \boldsymbol{\theta}). \tag{8.10}$$

In this context, the boundaries of a pinched response p-box are computed with a pinched input vector $\boldsymbol{X}_{\check{X}_i}$:

$$\underline{F}_{Y_{\check{X}_i}}(y) = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}_{\check{X}_i}}} F_{Y_{\boldsymbol{\theta}}}(y, \boldsymbol{\theta}), \qquad \overline{F}_{Y_{\check{X}_i}}(y) = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}_{\check{X}_i}}} F_{Y_{\boldsymbol{\theta}}}(y, \boldsymbol{\theta}), \tag{8.11}$$

where the pinched optimization domain $\mathcal{D}_{\boldsymbol{\Theta}_{\check{X}_i}}$ is equal to the original optimization domain $\mathcal{D}_{\boldsymbol{\Theta}}$ with the components $\boldsymbol{\theta}_{\check{X}_i}$ being constant instead of interval-valued.

Figure 8.4 summarizes the main steps of the algorithm in a flowchart. Following the paths towards the augmented PCE model $\mathcal{M}^{(\mathrm{P})}_{(\mathrm{aug})}$ corresponds to the procedure presented in Section 6.4. The pinching algorithm introduces the paths from $\boldsymbol{V}_{\check{X}_i}$ to $A_{Y_{\check{X}_i}}$. Similar to the previous section, the augmented meta-model $\mathcal{M}^{(\mathrm{P})}_{(\mathrm{aug})}$ is computed only once using the un-pinched input vector $\boldsymbol{V}$. This model is reused for every pinched variable $\check{X}_i$, $i = 1, \ldots, M$. Hence, the main additions in terms of computational effort are the optimizations to estimate $\underline{Y}_{\check{X}_i}$ and $\overline{Y}_{\check{X}_i}$, which is, in turn, small compared to the computation of the experimental design of the augmented PCE model.

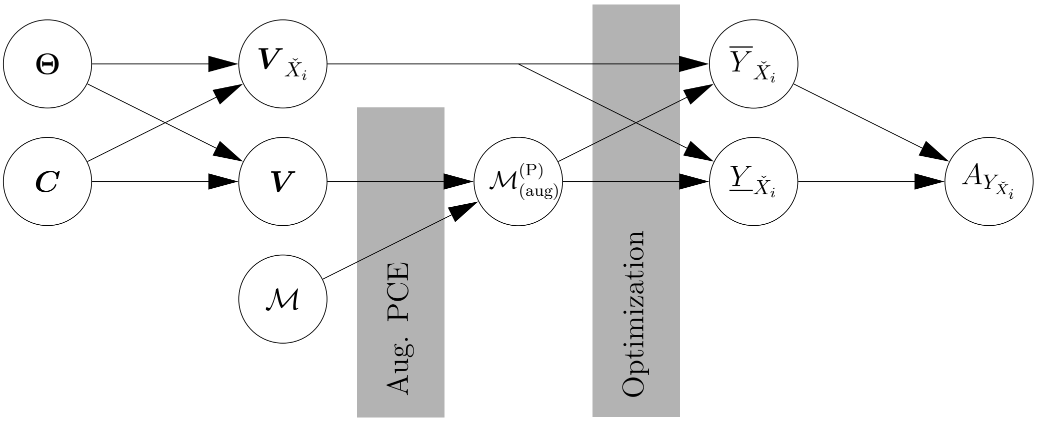


Figure 8.4: Pinching of parametric p-boxes – augmented PCE-based computation of $A_{Y_{\check{X}_i}}$.

VoI with pinched failure probabilities

In the case of Eq. (8.4), the two-level meta-modelling approach presented in Figure 7.4 can be adjusted to the pinching algorithm. Instead of using the domain $\mathcal{D}_{\boldsymbol{\Theta}}$, the domain $\mathcal{D}_{\boldsymbol{\Theta}_{\check{X}_i}}$ is used to estimate the bounds of the failure probability by EGO and AK-MCS. After estimating the bounds $\underline{P}_{f,\check{X}_i}$ and $\overline{P}_{f,\check{X}_i}$, the VoI is computed by Eq. (8.4).

The limit-state function evaluations in AK-MCS can be recycled for different input vectors $\boldsymbol{\Theta}_{\check{X}_i}$, $i = 1, \ldots, M$. This limits the number of limit-state function evaluations and hence also the total computational costs in the case of an expensive-to-evaluate $\mathcal{G}$.

The resulting two-level meta-modelling approach is summarized in Figure 8.5. Compared to Figure 7.4, the variables $\boldsymbol{\Theta}_{\check{X}_i}$ replaces the corresponding $\boldsymbol{\Theta}$ and the box for computing the range of reliability index is added. The core of the two-level meta-modelling approach stays the same.

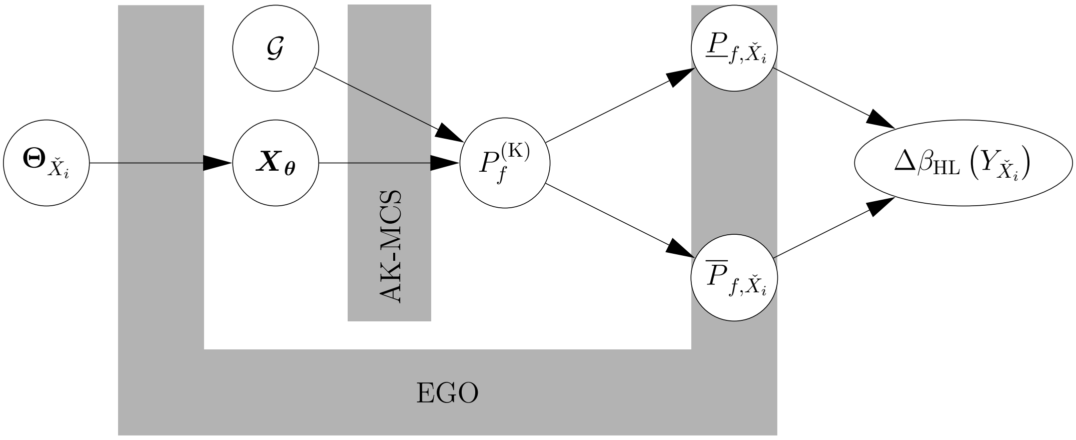


Figure 8.5: Pinching for parametric p-boxes – two-level meta-modelling approach for computation of $\Delta\beta_{\mathrm{HL}}\left(Y_{\check{X}_i}\right)$.

8.3.3 Sobol' indices

Theory

Consider a probabilistic input vector $\boldsymbol{X}$, whose components X_i, $i = 1, \ldots, M$ are modelled by independent random variables. The output scalar of a computational model is then $Y = \mathcal{M}(\boldsymbol{X})$ as seen in Eq. (3.2). The so-called Sobol' decomposition represents the computational model by a series of summands of increasing dimension (Sobol' 1993):

$$\begin{aligned}\mathcal{M}(\boldsymbol{x}) = \mathcal{M}_0 + \sum_{i=1}^{M} \mathcal{M}_i(x_i) + \sum_{1 \leq i < j \leq M} \mathcal{M}_{ij}(x_i, x_j) + \ldots \\ + \sum_{1 \leq i_1 < \ldots < i_s \leq M} \mathcal{M}_{i_1 \ldots i_s}\left(x_{i_1}, \ldots, x_{i_s}\right) + \ldots + \mathcal{M}_{1,2,\ldots,M}(\boldsymbol{x}),\end{aligned} \tag{8.12}$$

where $\mathcal{M}_0$ is a constant (mean value of the function) and where it is imposed that the integral of each summand $\mathcal{M}_{i_1,\ldots,i_s}\left(x_{i_1}, \ldots, x_{i_s}\right)$ over any of its arguments is zero:

$$\int_{\mathcal{D}_{X_k}} \mathcal{M}_{i_1 \ldots i_s}(x_{i_1}, \ldots, x_{i_s})\, f_{X_k}(x_k)\, \mathrm{d}x_k = 0, \qquad 1 \leq i_1 < \ldots < i_s \leq M, \qquad k \in \{i_1, \ldots, i_s\}, \tag{8.13}$$

where $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}}$. Then, the mean value of the function can be computed as:

$$\mathcal{M}_0 = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}. \tag{8.14}$$

Due to the imposed constraints in Eq. (8.13), the summands are orthogonal to each other so that (Homma & Saltelli 1996):

$$\int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s})\, \mathcal{M}_{j_1 \dots j_t}(x_{j_1, \dots, j_t})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = 0, \qquad \{i_1, \dots, i_s\} \neq \{j_1, \dots, j_t\}. \tag{8.15}$$

Considering now that the input parameters are modelled by independent random variables, the *total variance* of the computational model is defined as:

$$D = \mathrm{Var}\left[\mathcal{M}(\boldsymbol{X})\right] = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}^2(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} - \mathcal{M}_0^2. \tag{8.16}$$

By using the properties of Eq. (8.15) in Eq. (8.12), one obtains a decomposition of the variance:

$$D = \sum_{i=1}^{M} D_i + \sum_{1 \leq i < j \leq M} D_{ij} + \ldots + D_{1,2,\dots,M}, \tag{8.17}$$

where the *partial variances* are computed as:

$$D_{i_1, \dots, i_s} = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}^2_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \qquad \text{for} \quad 1 \leq i_1 < \ldots < i_s \leq M, \qquad s = 1, \dots, M. \tag{8.18}$$

This equation can be simplified due to the orthogonality property:

$$D_{i_1, \dots, i_s} = \int_{\mathcal{D}_{\boldsymbol{X}}} \mathcal{M}^2_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s})\, f_{x_{i_1}}(x_{i_1}) \dots f_{x_{i_s}}(x_{i_s})\, \mathrm{d}x_{i_1} \dots \mathrm{d}x_{i_s} \qquad \text{for} \quad 1 \leq i_1 < \ldots < i_s \leq M, \qquad s = 1, \dots, M. \tag{8.19}$$

Making use of this decomposition, the Sobol' indices are defined as the relative partial variances:

$$S_{i_1 \dots i_s} = \frac{D_{i_1 \dots i_s}}{D}, \tag{8.20}$$

and by definition in Eq. (8.17), it holds that:

$$\sum_{i=1}^{M} S_i + \sum_{1 \leq i < j \leq M} S_{ij} + \ldots + S_{1,2,\dots,M} = 1. \tag{8.21}$$

The Sobol' indices measure the amount of the total variance due to the uncertainties in a set of input parameters. In practice, it is common to measure the first order indices S_i, which measure the influence of each parameter taken separately. Higher order indices account for the interactive contributions to the total variance. In this sense, the *total sensitivity indices* $S_i^{(T)}$ are defined as:

$$S_i^{(T)} = \sum_{\{i_1, \dots, i_s\} \supset \{i\}} \frac{D_{i_1 \dots i_s}}{D}. \tag{8.22}$$

In other words, $S_i^{(T)} = 1 - S_{-i}$, where S_{-i} is the sum of all $S_{i_1 \dots i_s}$ that do not include index i. Note that the sum of all total Sobol' indices is generally larger than one due to the interactive terms $D_{i_1 \dots i_s}$, which are considered multiple times.

$$\sum_{i=1}^{M} S_i^{(T)} \geq 1, \tag{8.23}$$

The equality only holds when there are no interactive terms, *i.e.* the computational model $\mathcal{M}$ is a sum of univariate functions $\mathcal{M}_i\left(x_i\right)$, $i=1,\ldots,M$. Likewise, the sum of all first order Sobol' indices generally sums up to less than one due to the interactive terms $D_{i_1\ldots i_s}$, which are not accounted for:

$$\sum_{i=1}^{M} S_i \leq 1, \tag{8.24}$$

The equality only holds when there are no interactive terms. Sobol' indices of higher order can be computed in the same manner. Often in practice, however, only the first and the total order Sobol' indices are of interest.

PCE-based Sobol' indices

Let us consider now that an appropriate approximation of the computational model $\mathcal{M}$ exists in the form of a truncated PCE, as seen in Eq. (3.8) and recalled here:

$$Y \approx \mathcal{M}^{(\mathrm{P})}\left(\boldsymbol{X}\right) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}\left(\boldsymbol{X}\right). \tag{8.25}$$

Then, a set of multi-indices $\mathcal{I}_{i_1,\ldots,i_s}$ can be defined so that only the indices $i_1,\ldots,i_s$ are non-zero, *i.e.* :

$$\mathcal{I}_{i_i,\ldots,i_s} = \left\{\boldsymbol{\alpha}\in\mathcal{A} : \begin{array}{lll} \alpha_k > 0 & \text{for all} & k \in \{i_1,\ldots,i_s\} \\ \alpha_k = 0 & \text{for all} & k \notin \{i_1,\ldots,i_s\} \end{array}\right\}. \tag{8.26}$$

Then, the elements of Eq. (3.8) can be reordered according to the decomposition in Eq. (8.12) (Sudret 2008):

$$\begin{aligned}\mathcal{M}^{(\mathrm{P})}(\boldsymbol{x}) = a_0 &+ \sum_{i=1}^{N}\left(\sum_{\boldsymbol{\alpha}\in\mathcal{I}_i} a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(x_i)\right) + \sum_{1\leq i_1<i_2\leq M}\left(\sum_{\boldsymbol{\alpha}\in\mathcal{I}_{i_1,i_2}} a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(x_{i_1},x_{i_2})\right) + \ldots \\ &+ \sum_{1\leq i_1<\ldots<i_s\leq M}\left(\sum_{\boldsymbol{\alpha}\in\mathcal{I}_{i_1,\ldots,i_s}} a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(x_{i_1},\ldots,x_{i_s})\right) + \ldots + \sum_{\boldsymbol{\alpha}\in\mathcal{I}_{1,2,\ldots,M}} a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(x_1,\ldots,x_M).\end{aligned} \tag{8.27}$$

The summands fo Eq. (8.27) can be identified as summands of the Sobol' decomposition:

$$\mathcal{M}_{i_1\ldots i_s}(x_{i_1},\ldots,x_{i_s}) = \sum_{\boldsymbol{\alpha}\in\mathcal{I}_{i_1,\ldots,i_s}} a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(x_{i_1},\ldots,x_{i_s}). \tag{8.28}$$

Due to the property of uniqueness, it can be concluded that Eq. (8.27) is indeed the Sobol' decomposition of the PCE model.

The corresponding *polynomial-chaos-based Sobol' indices*, denoted by $S^{(\mathrm{P})}_{i_1\ldots i_s}$, are then defined as:

$$S^{(\mathrm{P})}_{i_1\ldots i_s} = \frac{1}{D^{(\mathrm{P})}} \sum_{\boldsymbol{\alpha}\in\mathcal{I}_{i_1,\ldots,i_s}} a^2_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha}\in\mathcal{I}_{i_1,\ldots,i_s}} a^2_{\boldsymbol{\alpha}} \Bigg/ \sum_{\boldsymbol{\alpha}\in\mathcal{A},\ \boldsymbol{\alpha}\neq\boldsymbol{0}} a^2_{\boldsymbol{\alpha}}\ , \tag{8.29}$$

where $D^{(\mathrm{P})}$ is the variance of the response variable Y obtained by the PCE model $\mathcal{M}^{(\mathrm{P})}$. The *total PC-based sensitivity indices* are then computed analogously to Eq. (8.22):

$$S_i^{(T)(\mathrm{P})} = \sum_{\{i_1,\ldots,i_s\}\supset\{i\}} S^{(\mathrm{P})}_{i_1,\ldots,i_s}. \tag{8.30}$$

As seen in this section, the computation of the Sobol' indices for a PCE model boils down to working with the coefficients of the PC-expansion, which is a simple operation at almost no additional computational cost.

8.3.4 Imprecise Sobol' indices

Idea

In the context of parametric p-boxes, uncertainty propagation is typically conducted in nested Monte Carlo simulations (nMCS) algorithms where an outer loop samples the parameters $\boldsymbol{\theta}$ and an inner loop samples the conditional probability distribution $F_{\boldsymbol{X}}\left(\boldsymbol{x}|\boldsymbol{\theta}\right)$. Taking this idea into the context of sensitivity analysis, the Sobol' indices can be computed for each realization $\boldsymbol{\theta}^{(j)}$ and the corresponding $F_{\boldsymbol{X}}\left(\boldsymbol{x}|\boldsymbol{\theta}^{(j)}\right)$. Each realization can result in a different set of Sobol' indices. Finally, the bounds of the Sobol' indices can be obtained by analysing the set of realizations.

In order to determine the bounds of the imprecise Sobol' indices, an optimization problem can be defined. As an example, the bounds of the total Sobol' indices are obtained by:

$$\underline{S}_j^{(T)} = \min_{\boldsymbol{\theta}\in\mathcal{D}_{\boldsymbol{\Theta}}} S_j^{(T)}\left(\boldsymbol{\theta}\right), \qquad \overline{S}_j^{(T)} = \max_{\boldsymbol{\theta}\in\mathcal{D}_{\boldsymbol{\Theta}}} S_j^{(T)}\left(\boldsymbol{\theta}\right), \tag{8.31}$$

where $S_j^{(T)}\left(\boldsymbol{\theta}\right)$ is the j-th total Sobol' index for the conditional input distribution $F_{\boldsymbol{X}}\left(\boldsymbol{x}|\boldsymbol{\theta}\right)$, and $j = 1, \ldots, M$.

Eq. (8.31) requires a repeated evaluation of the Sobol' indices, which includes a repeated computation of the variance decomposition in Eq. (8.12). In order to speed up the process, PCE models can be used:

$$\underline{S}_j^{(T)(\mathrm{P})} = \min_{\boldsymbol{\theta}\in\mathcal{D}_{\boldsymbol{\Theta}}} S_j^{(T)(\mathrm{P})}\left(\boldsymbol{\theta}\right), \qquad \overline{S}_j^{(T)(\mathrm{P})} = \max_{\boldsymbol{\theta}\in\mathcal{D}_{\boldsymbol{\Theta}}} S_j^{(T)(\mathrm{P})}\left(\boldsymbol{\theta}\right), \tag{8.32}$$

where $S_j^{(T)(\mathrm{P})}$ is the PCE-based total Sobol' index equivalent to $S_j^{(T)}$. The use of the PCE models reduces the computational effort of computing Sobol' indices but does not remove the repeated computation of the variance decomposition for every $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$.

Augmented PCE-based indices

When an augmented PCE model is available as in Eq. (6.31) in Section 6.4.2, the imprecise Sobol' indices can be obtained by post-processing this augmented PCE model. Recall that:

$$W \approx \mathcal{M}^{(\mathrm{P})}\left(\boldsymbol{V}\right) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} a_{\boldsymbol{\alpha}}\ \psi_{\boldsymbol{\alpha}}\left(\boldsymbol{V}\right), \tag{8.33}$$

where $\boldsymbol{V} = \left(\boldsymbol{C}, \boldsymbol{\Theta}\right)$ is a composite of uniform random variables $\boldsymbol{C}$ accounting for the aleatory uncertainty in the input vector and the vector of distribution parameters $\boldsymbol{\Theta}$ accounting for epistemic uncertainty. Hence, the Sobol' indices shall be depending on $\boldsymbol{C}$ and conditioned on $\boldsymbol{\Theta}$. In order to do so, the augmented PCE model is reformulated:

$$W \approx \mathcal{M}^{(\mathrm{P})}\left(\boldsymbol{C}, \boldsymbol{\Theta}\right) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} a_{\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}_{\boldsymbol{C}}}\left(\boldsymbol{C}\right)\ \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}}\left(\boldsymbol{\Theta}\right), \tag{8.34}$$

where $\boldsymbol{\alpha} = \left(\boldsymbol{\alpha}_{\boldsymbol{C}}, \boldsymbol{\alpha}_{\boldsymbol{\Theta}}\right)$ forms the same index set as in Eq. (8.33) and $\psi_{\boldsymbol{\alpha}}\left(\boldsymbol{V}\right) = \psi_{\boldsymbol{\alpha}_{\boldsymbol{C}}}\left(\boldsymbol{C}\right) \cdot \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}}\left(\boldsymbol{\Theta}\right)$ are the corresponding multivariate orthonormal polynomials. Then, the computational model can be further rearranged for a given value $\boldsymbol{\theta}$ as:

$$W\left(\boldsymbol{\theta}\right) \approx \mathcal{M}^{(\mathrm{P})}\left(\boldsymbol{C}, \boldsymbol{\theta}\right) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} \left(a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}}\left(\boldsymbol{\theta}\right)\right)\ \psi_{\boldsymbol{\alpha}_{\boldsymbol{C}}}\left(\boldsymbol{C}\right) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} a_{\boldsymbol{\alpha},\boldsymbol{\theta}}(\boldsymbol{\theta})\ \psi_{\boldsymbol{\alpha}_{\boldsymbol{C}}}\left(\boldsymbol{C}\right), \tag{8.35}$$

where $a_{\boldsymbol{\alpha},\boldsymbol{\theta}}(\boldsymbol{\theta}) = a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}}(\boldsymbol{\theta})$ is a new coefficient dependent on the value of $\boldsymbol{\theta}$. Note that several $\boldsymbol{\alpha}_{\boldsymbol{\Theta}}$ may correspond to the same $\boldsymbol{\alpha}_{\boldsymbol{C}}$ in $\boldsymbol{\alpha}$. Hence, Eq. (8.35) can be rewritten to a model with proper variance decomposition as in Eq. (8.27):

$$W(\boldsymbol{\theta}) \approx \mathcal{M}^{(\mathrm{P})}(\boldsymbol{C},\boldsymbol{\theta}) = \sum_{\boldsymbol{\alpha}^*_{\boldsymbol{C}} \in \mathcal{A}^*_{\boldsymbol{C}}} \left(\sum_{i=1,\ldots,n_{\mathcal{A}}} \mathbb{I}_{\left\{\boldsymbol{\alpha}^{(i)}_{\boldsymbol{C}} = \boldsymbol{\alpha}^*_{\boldsymbol{C}}\right\}} \left(\boldsymbol{\alpha}^{(i)}_{\boldsymbol{C}}\right) \, a_{\boldsymbol{\alpha}^{(i)}} \, \psi_{\boldsymbol{\alpha}^{(i)}_{\boldsymbol{\Theta}}}(\boldsymbol{\theta}) \right) \psi_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{C}), \tag{8.36}$$

where i marks the $1, \ldots, n_{\mathcal{A}}$ multi-indices in $\mathcal{A}$, $\mathbb{I}$ is the indicator function with $\mathbb{I} = 1$ for $\boldsymbol{\alpha}^{(i)}_{\boldsymbol{C}} = \boldsymbol{\alpha}^*_{\boldsymbol{C}}$ and $\mathbb{I} = 0$ otherwise. $\mathcal{A}^*_{\boldsymbol{C}}$ is the set of unique multi-indices $\boldsymbol{\alpha}_{\boldsymbol{C}}$ in $\mathcal{A}$ and $\boldsymbol{\alpha}^{(i)} = \left(\boldsymbol{\alpha}^{(i)}_{\boldsymbol{C}}, \boldsymbol{\alpha}^{(i)}_{\boldsymbol{\Theta}}\right)$. Rearranging the terms to obtain the same structure as in Eq. (8.27) leads to the following coefficients depending on $\boldsymbol{\theta}$:

$$a_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\theta}) = \sum_{i=1,\ldots,n_{\mathcal{A}}} \mathbb{I}_{\left\{\boldsymbol{\alpha}^{(i)}_{\boldsymbol{C}} = \boldsymbol{\alpha}^*_{\boldsymbol{C}}\right\}} \left(\boldsymbol{\alpha}^{(i)}_{\boldsymbol{C}}\right) \, a_{\boldsymbol{\alpha}^{(i)}} \, \psi_{\boldsymbol{\alpha}^{(i)}_{\boldsymbol{\Theta}}}(\boldsymbol{\theta}). \tag{8.37}$$

The simplified decomposition then reads:

$$W_{\boldsymbol{\theta}} \approx \mathcal{M}^{(\mathrm{P})}(\boldsymbol{C}, \boldsymbol{\Theta} = \boldsymbol{\theta}) = \sum_{\boldsymbol{\alpha}^*_{\boldsymbol{C}} \in \mathcal{A}^*_{\boldsymbol{C}}} a_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\Theta} = \boldsymbol{\theta}) \, \psi_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{C}). \tag{8.38}$$

Finally, in order to obtain the extreme values of the Sobol' indices, the optimizations in Eq. (8.31) use the reformulated model in Eq. (8.38). Then, the bounds of imprecise Sobol' indices can be computed by:

$$\underline{S}^{(\mathrm{P})}_{i_1 \ldots i_s} = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} \frac{D_{i_1 \ldots i_s}(\boldsymbol{\theta})}{D(\boldsymbol{\theta})} = \min_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} \sum_{\boldsymbol{\alpha}^*_{\boldsymbol{C}} \in \mathcal{I}_{i_1 \ldots i_s}} a^2_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\theta}) \Bigg/ \sum_{\boldsymbol{\alpha}^*_{\boldsymbol{C}} \in \mathcal{A}^*_{\boldsymbol{C}},\ \boldsymbol{\alpha}^*_{\boldsymbol{C}} \neq \boldsymbol{0}} a^2_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\theta}), \tag{8.39}$$

$$\overline{S}^{(\mathrm{P})}_{i_1 \ldots i_s} = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} \frac{D_{i_1 \ldots i_s}(\boldsymbol{\theta})}{D(\boldsymbol{\theta})} = \max_{\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}} \sum_{\boldsymbol{\alpha}^*_{\boldsymbol{C}} \in \mathcal{I}_{i_1 \ldots i_s}} a^2_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\theta}) \Bigg/ \sum_{\boldsymbol{\alpha}^*_{\boldsymbol{C}} \in \mathcal{A}^*_{\boldsymbol{C}},\ \boldsymbol{\alpha}^*_{\boldsymbol{C}} \neq \boldsymbol{0}} a^2_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\theta}), \tag{8.40}$$

Using the augmented PCE, the extreme values are obtained by optimizing the PCE coefficients $a_{\boldsymbol{\alpha}^*_{\boldsymbol{C}}}(\boldsymbol{\theta})$ in $\boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$ without recomputing the meta-model.

8.4 Numerical examples

The derived concepts are discussed on three application examples. The toy function example examines the machinery of the imprecise Sobol' indices in the case of parametric p-boxes. The second example discusses the two-degree-of-freedom oscillator in the context of a pinching analysis. Last but not least, the third example consists of the two-dimensional truss structure presented in previous chapters. A full analysis, including parametric p-boxes and imprecise Sobol' indices as well as pinching analyses, is discussed for this example.

8.4.1 Toy function

Problem statement

A simple two-dimensional function is considered here to visualize the algorithm for computing the imprecise Sobol' indices presented in Section 8.3.3. The analytical function is defined as:

$$f_8(\boldsymbol{x}) = x_1 \cdot x_2, \tag{8.41}$$

where x_1 and x_2 are modelled by parametric p-boxes. The input variables are modelled by Gaussian distributions with interval-valued mean value and standard deviation. In particular, $\mu_i \in [-1, 1]$ and $\sigma_i \in [0.5, 1.0]$.

Imprecise Sobol' indices

Each input parameter may be cast as $x_i = \mu_i + \sigma_i \cdot \zeta_i$ where ζ_i will be a standard normal variable. Then, the augmented computational model can be written as:

$$\begin{aligned} f_8^{(\mathrm{aug})}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\zeta}) &= (\mu_1 + \sigma_1 \cdot \zeta_1) \cdot (\mu_2 + \sigma_2 \cdot \zeta_2) \\ &= (\mu_1 \cdot \mu_2) + (\mu_2 \cdot \sigma_1) \cdot \zeta_1 + (\mu_1 \cdot \sigma_2) \cdot \zeta_2 + (\sigma_1 \cdot \sigma_2) \cdot \zeta_1 \cdot \zeta_2. \end{aligned} \tag{8.42}$$

The variance of the response variable conditional on the values of $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ then reads:

$$D(\boldsymbol{\mu}, \boldsymbol{\sigma}) \equiv \mathrm{Var}\left[f_8 | \boldsymbol{\mu}, \boldsymbol{\sigma}\right] = (\mu_2\sigma_1)^2 + (\mu_1\sigma_2)^2 + (\sigma_1\sigma_2)^2. \tag{8.43}$$

Hence, the first order Sobol' indices can be computed as a function of $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\sigma})$:

$$S_1(\boldsymbol{\theta}) = \frac{(\mu_2\sigma_1)^2}{(\mu_2\sigma_1)^2 + (\mu_1\sigma_2)^2 + (\sigma_1\sigma_2)^2}, \qquad S_2(\boldsymbol{\theta}) = \frac{(\mu_1\sigma_2)^2}{(\mu_2\sigma_1)^2 + (\mu_1\sigma_2)^2 + (\sigma_1\sigma_2)^2}, \tag{8.44}$$

The analytical derivation of the extreme Sobol' indices results in the following bounds for the first order Sobol' indices:

$$\underline{S}_1 = \underline{S}_2 = 0.0, \qquad \overline{S}_1 = \overline{S}_2 = 0.8. \tag{8.45}$$

As an example, the minimal first order Sobol' index for X_1 is obtained by setting $\mu_2 = 0$.

Similarly to the first order indices, the total order Sobol' indices can be written as:

$$S_1^{(T)}(\boldsymbol{\theta}) = \frac{(\mu_2\sigma_1)^2 + (\sigma_1\sigma_2)^2}{(\mu_2\sigma_1)^2 + (\mu_1\sigma_2)^2 + (\sigma_1\sigma_2)^2}, \qquad S_2^{(T)}(\boldsymbol{\theta}) = \frac{(\mu_1\sigma_2)^2 + (\sigma_1\sigma_2)^2}{(\mu_2\sigma_1)^2 + (\mu_1\sigma_2)^2 + (\sigma_1\sigma_2)^2}, \tag{8.46}$$

The bounds of the total order Sobol' indices read:

$$\underline{S}_1^{(T)} = \underline{S}_2^{(T)} = 0.2, \qquad \overline{S}_1^{(T)} = \overline{S}_2^{(T)} = 1.0. \tag{8.47}$$

As an example, the minimal total order Sobol' index for X_1 is obtained by $\mu_1 \pm 1$, $\sigma_1 = 0.5$, $\mu_2 = 0$, and $\sigma_2 = 1$. Due to the symmetry of the problem definition, *i.e.* computational model $f_8^{(\mathrm{aug})}$ and input p-boxes X_i, the extreme Sobol' indices are equivalent for X_1 and X_2. Note that in this example, the first and total order Sobol' indices are not identical due to the interactive term $(\sigma_1 \cdot \sigma_2) \cdot \zeta_1 \cdot \zeta_2$ in Eq. (8.42).

Augmented PCE-based Sobol' indices

Implemented in UQLab (Marelli & Sudret 2015), the augmented PCE of $f_8^{(\mathrm{aug})}$ and the parametric p-boxes results in the following set of multi-indices $\mathcal{A}$:

$$\mathcal{A} = \begin{pmatrix} \boldsymbol{\alpha}^{(1)} \\ \vdots \\ \boldsymbol{\alpha}^{(10)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}, \tag{8.48}$$

which contains ten vectors $\boldsymbol{\alpha}$ and where the indices in each row correspond to the vector $(\mu_1, \sigma_1, \zeta_1, \mu_2, \sigma_2, \zeta_2)$. Splitting the multi-index set $\mathcal{A}$ to the contributions of the epistemic $(\mu_1,\, \sigma_1,\, \mu_2,\, \sigma_2)$ and aleatory $(\zeta_1,\, \zeta_2)$ variables leads to:

$$\mathcal{A}_{\boldsymbol{\Theta}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathcal{A}_{\boldsymbol{C}} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}. \tag{8.49}$$

Note that according to Eq. (8.42), only four elements should be part of $\mathcal{A}$. However, due to the normalization to a standard space (uniform distributions to $[-1, 1]$ and Gaussian to standard normal space), an additional six elements are present.

$\mathcal{A}_{\boldsymbol{C}}^{*}$ summarizes the set of unique row vectors $\boldsymbol{\alpha}_{\boldsymbol{C}}$ in $\mathcal{A}_{\boldsymbol{C}}$ and contains the following four elements:

$$\mathcal{A}_{\boldsymbol{C}}^{*} = \begin{pmatrix} \boldsymbol{\alpha}_{\boldsymbol{C}}^{*(1)} \\ \boldsymbol{\alpha}_{\boldsymbol{C}}^{*(2)} \\ \boldsymbol{\alpha}_{\boldsymbol{C}}^{*(3)} \\ \boldsymbol{\alpha}_{\boldsymbol{C}}^{*(4)} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{8.50}$$

The final coefficients are composed as follows:

$$\begin{aligned} a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(1)}} &= a_{\boldsymbol{\alpha}^{(1)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(1)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(5)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(5)}}(\boldsymbol{\theta}), && (8.51) \\ a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(2)}} &= a_{\boldsymbol{\alpha}^{(2)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(2)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(6)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(6)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(7)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(7)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(10)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(10)}}(\boldsymbol{\theta}), && (8.52) \\ a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(3)}} &= a_{\boldsymbol{\alpha}^{(3)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(3)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(8)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(8)}}(\boldsymbol{\theta}), && (8.53) \\ a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(4)}} &= a_{\boldsymbol{\alpha}^{(4)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(4)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(9)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(9)}}(\boldsymbol{\theta}), && (8.54) \end{aligned}$$

where Eqs. (8.51) and (8.52) can be simplified to:

$$\begin{aligned} a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(1)}} &= a_{\boldsymbol{\alpha}^{(1)}} + a_{\boldsymbol{\alpha}^{(5)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(5)}}(\boldsymbol{\theta}), && (8.55) \\ a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(2)}} &= a_{\boldsymbol{\alpha}^{(2)}} + a_{\boldsymbol{\alpha}^{(6)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(6)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(7)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(7)}}(\boldsymbol{\theta}) + a_{\boldsymbol{\alpha}^{(10)}} \psi_{\boldsymbol{\alpha}_{\boldsymbol{\Theta}}^{(10)}}(\boldsymbol{\theta}). && (8.56) \end{aligned}$$

The variance decomposition of the PCE model then reads (see also Eq. (8.38)):

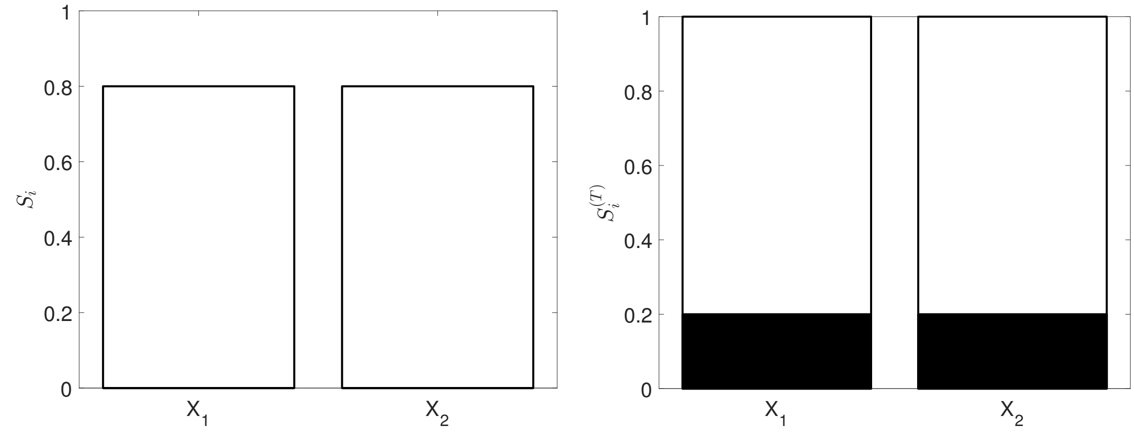
$$\mathcal{M}^{(\mathrm{P})}\left(\boldsymbol{C}, \boldsymbol{\Theta}\right) = \sum_{i=1,\ldots,4} a_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(i)}}\left(\boldsymbol{\Theta}\right)\ \psi_{\boldsymbol{\alpha}_{\boldsymbol{C}}^{(i)}}\left(\boldsymbol{C}\right), \tag{8.57}$$

where $\boldsymbol{C} = (\zeta_1, \zeta_2)$. By optimization on $\boldsymbol{\theta}$, the Sobol' indices of Eqs. (8.45) and (8.47) are obtained.

Visualization of the results

The resulting imprecise Sobol' indices are illustrated in Figure 8.6. In particular, the first order indices are shown in Figure 8.6(a), whereas the total order indices are shown in Figure 8.6(b).

Each sensitivity index is visualized by two overlapping bars. The solid bar represent the lower bound of the Sobol' index, whereas the transparent bar represent the maximum bound of the same Sobol' index. Hence, the imprecision in the sensitivity measure is shown by the difference of these two bars.



(a) First order indices (b) Total order indices

Figure 8.6: Toy function – imprecise Sobol' indices.

The results nicely visualize the symmetry of the model f_8 and the p-boxes describing X_1 and X_2. Further, the influence of the interactive terms (see $\boldsymbol{\alpha}_{\boldsymbol{C}}^{*(2)}$) can be seen when comparing the first order to the total order indices, *i.e.* Figures 8.6(a) to 8.6(b).

8.4.2 Single-degree-of-freedom oscillator

Problem statement

Consider again the SDOF oscillator described in Section 7.6.2. The limit-state function is defined in Eq. (7.15) and recalled here:

$$g_{\text{SDOF}}\left(r, F_1, t_1, c_1, c_2, m\right) = 3r - \left| \frac{2F_1}{m\omega_0^2} \sin\left(\frac{\omega_0 T_1}{2}\right) \right|, \tag{8.58}$$

where the six input parameters are modelled statistically independent. The two cases of (i) free and (ii) parametric p-boxes are considered here (see also Table 7.3). Note that only the first three variables are p-boxes whereas the remaining three variables are purely probabilistic variables.

Analysis

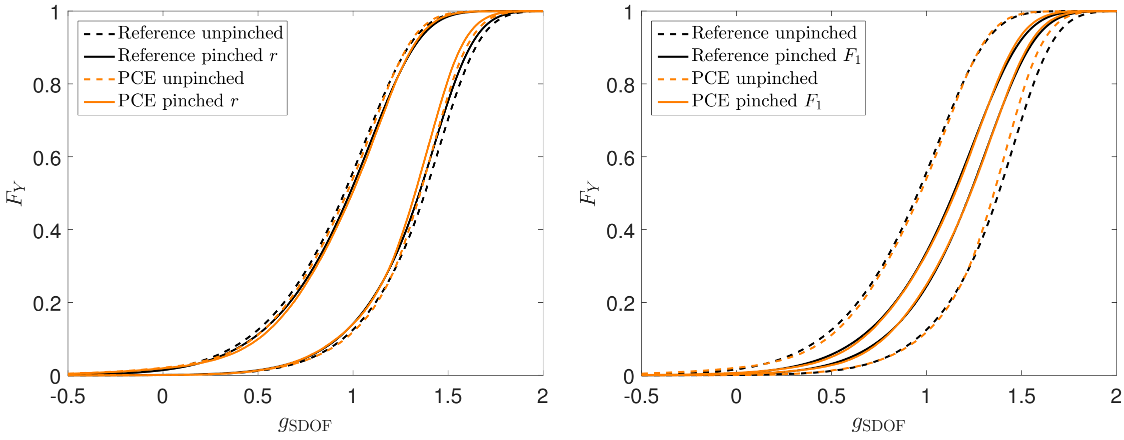
The pinching algorithms are used in this example to measure the influence of the epistemic uncertainty in each input variable onto the response p-box and the failure probability. Both VoI of Eqs. (8.2) and (8.4) are estimated as sensitivity measures. Note that for Eq. (8.4), failure is defined as $g_{\text{SDOF}} \leq 0$. As the algorithms presented in Section 8.2 and 8.3.1 are extensions of the previously discussed meta-modelling algorithms, their accuracy and efficiency is not discussed extensively in this section.

For the case of free p-boxes, the reference solutions are obtained by brute force Monte Carlo simulation with a large sample size $n_{\text{MCS}} = 10^7$. For the case of parametric p-boxes, the reference solution is obtained by (i) a brute force Monte Carlo simulation with $n_{\text{MCS}} = 10^7$ and (ii) a standard genetic algorithm for the optimal value of $\boldsymbol{\theta}$.

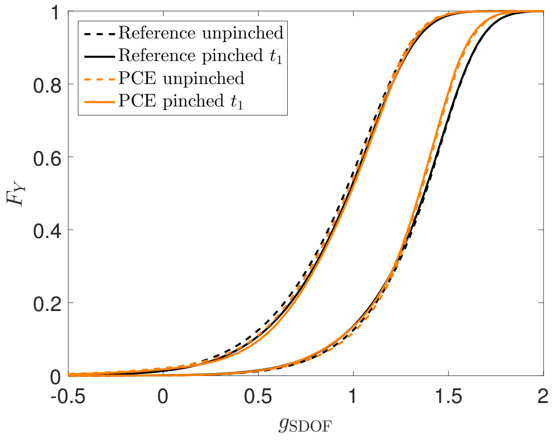
Free p-boxes

For free p-boxes, Eq. (8.2) is estimated by using PCE models with an experimental design of $N = 200$ LHS samples. Moreover, the maximal polynomial degree is set of $p = 15$ and a hyperbolic truncation set with $q = 0.75$ is used as the candidate basis. For estimating Eq. (8.4), all AK-MCS algorithms are initiated with an initial experimental design of $N_0 = 12$ LHS samples. The first-level algorithm resulted in a total number of computational model runs of $N^{(1)} = 180$. Hence, in terms of number of model runs, the analyses are comparable.

The resulting pinched p-boxes are shown in Figure 8.7. The dashed lines represent the unpinched response p-boxes, whereas the solid lines represent the pinched response p-boxes. Moreover, the orange curves mark the meta-model-based results whereas the black lines mark the reference results. The meta-model-based results approximate the reference solution accurately, as seen from the overlapping lines.



(a) Pinched r (b) Pinched F_1



(c) Pinched t_1

Figure 8.7: SDOF oscillator – pinched response free p-boxes.

The same figure can be used to infer the sensitivity to the epistemic uncertainty in this problem. Qualitatively, the epistemic uncertainty of variable F_1 has the largest impact on the imprecision in the response p-box since there is a great difference between the "pinched" and "unpinched" curves. The two variables $\{r, t_1\}$ have a minor impact on the width of the response p-box. Note that the three probabilistic variables $\{c_1, c_2, m\}$ are not shown here, because they are not affected by epistemic uncertainty, hence the width of the response p-box is independent of them.

The numerical values for the VoI of Eqs. (8.2) and (8.4) are summarized in Table 8.1. The meta-model-based results approximate well the reference values. Furthermore, the results confirm the impression gained in Figure 8.7: variable F_1 has the largest impact on the imprecision of the response p-box.

Table 8.1: SDOF oscillator – VoI of pinched free p-boxes.

		Eq. (8.2)			Eq. (8.4)		
$\check{X}_i$	$F_{\check{X}_i}$	VoI_{ref}	$\widehat{\mathrm{VoI}}$	Ranking	VoI_{ref}	$\widehat{\mathrm{VoI}}$	Ranking
r	$F_{\mathcal{N}}(r\|0.5, 0.05)$	0.140	0.146	#2	0.095	0.099	#3
F_1	$F_{\mathcal{N}}(F_1\|0.0, 0.5)$	0.784	0.771	#1	0.724	0.735	#1
t_1	$F_{\mathcal{N}}(t_1\|1.00, 0.2)$	0.079	0.086	#3	0.171	0.182	#2

The variables are ranked according to their VoI for both criteria. For both criteria, F_1 leads to the largest value of VoI. However, the variables r and t_1 lead to different rankings. r is more influential for the imprecision in the response p-box (*i.e.* Eq. (8.2)), whereas t_1 is more influential for the imprecision in the failure probability (*i.e.* Eq. (8.4)). This phenomenon originates from the definition of the two VoIs: Eq. (8.2) compares the volume of the p-box, whereas Eq. (8.4) compares the tail of the distributions.

Parametric p-boxes

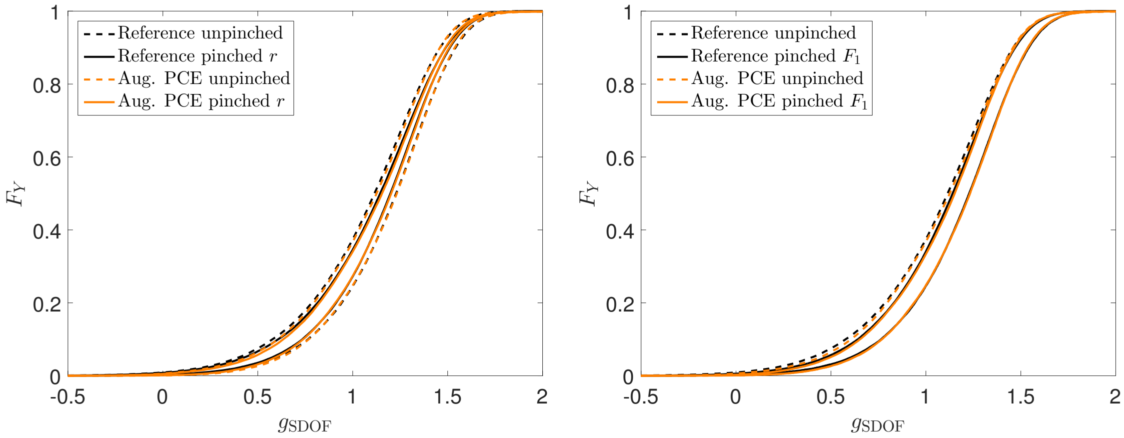
For parametric p-boxes, the same meta-model specifications are used to obtain comparable results to the case of free p-boxes. In the augmented PCE model, the number of phantom points is set to $n_{ph} = 10$ to ensure an accurate meta-model. For estimating Eq. (8.4) and the corresponding conditional failure probabilities by AK-MCS, the total number of evaluations of the limit-state function is $N^{(1)} = 223$, which is comparable to the $N^{(1)} = 200$ in the augmented PCE model.

Figure 8.8 shows the boundaries of the response p-boxes. The figure design is the same as in Figure 8.7. In this case, the largest reduction of the width of the response p-box is obtained by pinching variable r. Variables F_1 and t_1 have a smaller but non-negligible effect on the width of the response p-box.

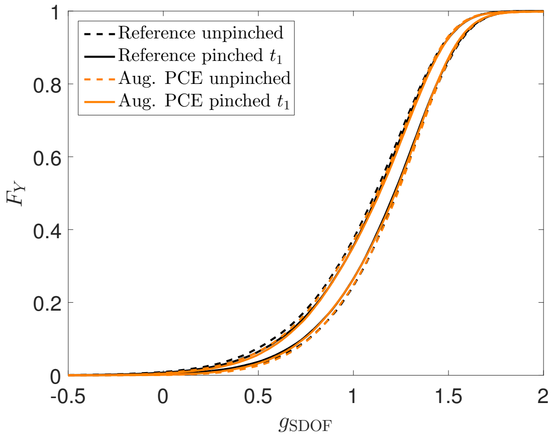
The quantitative values of the VoI are summarized in Table 8.2. The values confirm the impression gained from Figure 8.8. The VoI is approximated well by using the meta-model-based approaches. All three variables r, F_1, and t_1 have a non-negligible influence onto the imprecision in the response p-box. However, the relative ranking of their importance depends on the definition of the VoI.

Table 8.2: SDOF oscillator – VoI of pinched parametric p-boxes.

		Eq. (8.2)			Eq. (8.4)		
$\check{X}_i$	$F_{\check{X}_i}$	VoI_{ref}	$\widehat{\mathrm{VoI}}$	Ranking	VoI_{ref}	$\widehat{\mathrm{VoI}}$	Ranking
r	$F_{\mathcal{N}}(r\|0.5, 0.05)$	0.483	0.494	#1	0.237	0.255	#3
F_1	$F_{\mathcal{N}}(F_1\|0.0, 0.5)$	0.252	0.254	#2	0.343	0.319	#2
t_1	$F_{\mathcal{N}}(t_1\|1.00, 0.2)$	0.275	0.257	#2	0.392	0.421	#1



(a) Pinched r (b) Pinched F_1



(c) Pinched t_1

Figure 8.8: SDOF oscillator – pinched response parametric p-boxes.

Comparison

The width of the pinched response p-boxes are obtained with a comparable number of runs of the limit-state function, *i.e.* $N^{(1)} \approx 200$. Hence, the meta-model-based algorithms provide an efficient tool to estimate the pinched response p-boxes. For both types of p-boxes, the boundary curves of the response p-boxes are estimated accurately.

In terms of pinching and the corresponding VoI, the results differ for the two types of p-boxes. The relative importance of the variables changes when comparing Table 8.1 to Table 8.2. In other words, the definition of the input p-boxes affects the results of the sensitivity analysis. The reason for this behaviour is the definition of the limit-state function, which is a non-monotone and non-smooth function (see also Eq. (8.58)). The non-monotonicity can also be seen by the different shapes of the response p-boxes in Figures 8.7 and 8.8. The case of free p-boxes leads to much wider response p-boxes than the case of parametric p-boxes. Hence, the resulting sensitivities are likely to be different too.

8.4.3 Simply supported truss structure

Problem statement

Consider again the simply supported truss structure presented in Section 7.6.3. The midspan deflection is computed as a function of the seven loads, which are described here by parametric p-boxes. In order to estimate the influence of the loads onto the deflection u_4, imprecise Sobol' indices and pinching-based VoI are computed.

Imprecise Sobol' indices

Reference results The finite element model corresponds to a monotone function of the load parameters. Hence, the extreme cases of the first order Sobol' indices can be obtained by setting one variable to $\{\mu_{P_i} = 105 \text{ kN},\ \sigma_{P_i} = 17 \text{ kN}\}$ and all other ones to $\{\mu_{P_i} = 95 \text{ kN},\ \sigma_{P_i} = 13 \text{ kN}\}$ or vice versa. The resulting extreme Sobol' indices are summarized in Table 8.3. The indices are obtained by PCE and a large experimental design. As expected, the values group symmetrically around P_4. As an example, P_1 and P_7 have the same values. Note that the largest imprecision (*i.e.* largest $\Delta S_i^{(\mathrm{P})} = \overline{S}_i^{(\mathrm{P})} - \underline{S}_i^{(\mathrm{P})}$) is found where the largest upper bound Sobol' indices is found, *i.e.* P_4.

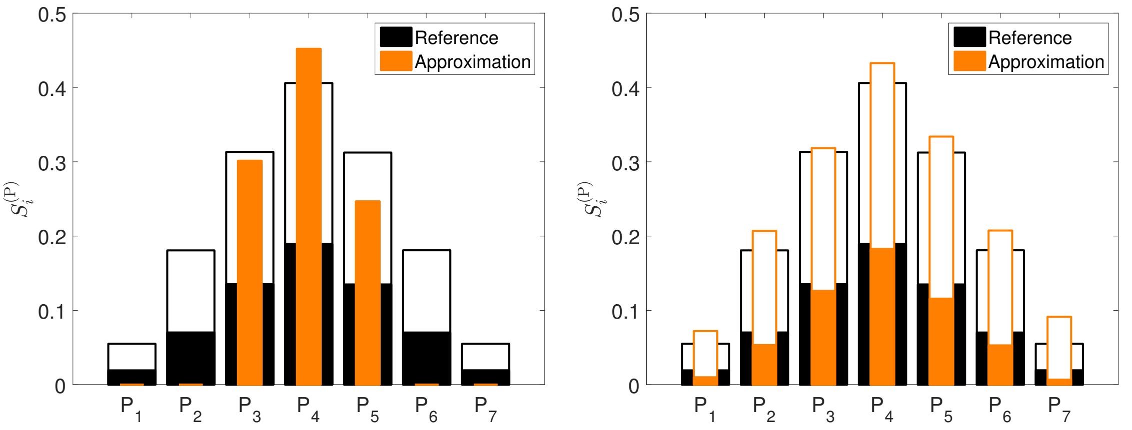
Table 8.3: Simply supported truss structure – imprecise Sobol' indices (first order) – reference results.

Variable	$\underline{S}_i^{(\mathrm{P})}$	$\overline{S}_i^{(\mathrm{P})}$	$\Delta S_i^{(\mathrm{P})}$
P_1	0.0193	0.0550	0.0357
P_2	0.0703	0.1808	0.1105
P_3	0.1353	0.3133	0.1780
P_4	0.1895	0.4060	0.2165
P_5	0.1349	0.3124	0.1775
P_6	0.0702	0.1809	0.1107
P_7	0.0194	0.0550	0.0356

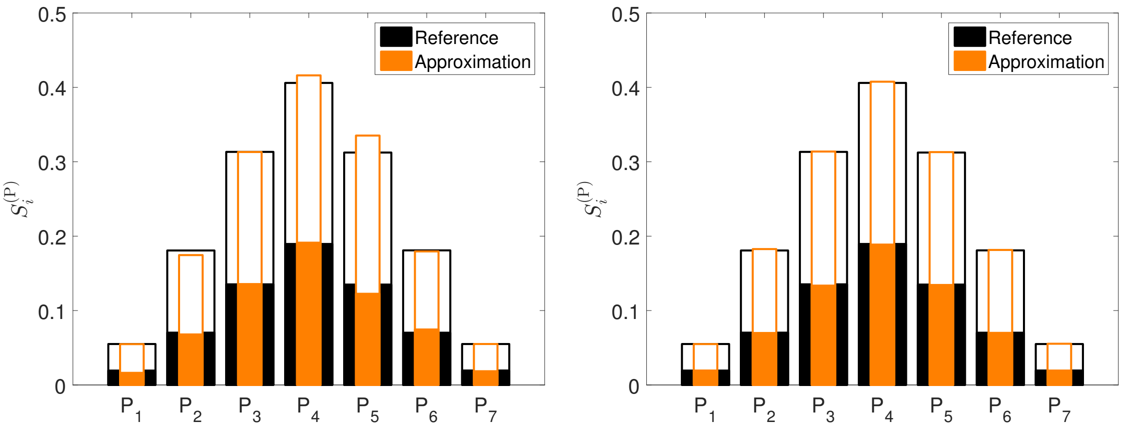
Augmented PCE analysis In order to estimate the imprecise Sobol' indices, an augmented PCE model is computed based on an experimental design of $N = 100$ LHS samples and the corresponding response values of the computational model. The original computational model requires a seven-dimensional input vector describing the loading forces. Due to the definition of the parametric p-boxes by interval-valued mean value and standard deviation, the augmented PCE model requires a 21-dimensional input vector. Phantom points are used to construct the meta-model in order to cope with the increased dimensionality of the augmented input space. The number of phantom points n_{ph} is varied to examine the influence of these points onto the accuracy of the Sobol' indices.

Influence of phantom points The imprecise Sobol' indices are visualized in Figure 8.9 as a function of n_{ph}. The estimated values (plotted with orange bars) are compared to the reference solution (black bars). As n_{ph} increases, the extreme values of the Sobol' indices converge to the true values without adding additional FEM runs. In this example, $n_{ph} = 5$ is sufficient to obtain reliable results. This impression is confirmed by the relative generalization error which is provided in the caption of the figure. In the case of $n_{ph} = 2$ and a relative generalization error of $\widehat{err}_{gen} = 4.9 \cdot 10^{-3}$, the imprecise PCE-based analysis provides usable estimates of the imprecise Sobol' indices.

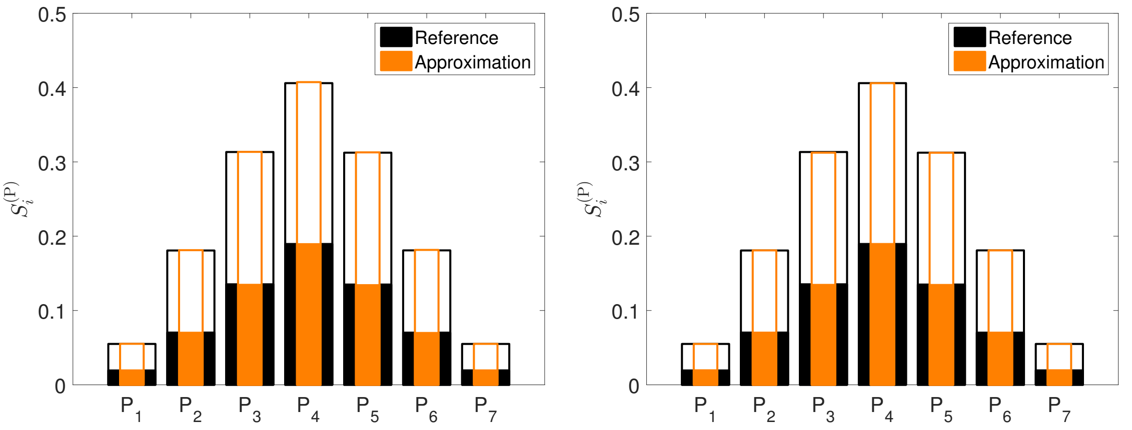
Table 8.4 summarizes the first order Sobol' indices obtained by the augmented PCE model with $N = 100$ samples and $n_{ph} = 20$ phantom points. The meta-model-based results are identical to the reference values given in Table 8.3. This confirms the impression gained in the graphical illustration in Figure 8.9(f).



(a) $n_{ph} = 1$ ($\widehat{err}_{gen} = 3.4 \cdot 10^{-1}$) (b) $n_{ph} = 2$ ($\widehat{err}_{gen} = 4.9 \cdot 10^{-3}$)



(c) $n_{ph} = 3$ ($\widehat{err}_{gen} = 8.4 \cdot 10^{-4}$) (d) $n_{ph} = 5$ ($\widehat{err}_{gen} = 1.1 \cdot 10^{-5}$)



(e) $n_{ph} = 10$ ($\widehat{err}_{gen} = 3.5 \cdot 10^{-6}$) (f) $n_{ph} = 20$ ($\widehat{err}_{gen} = 5.7 \cdot 10^{-7}$)

Figure 8.9: Simply supported truss structure – imprecise Sobol' indices (first order) – augmented-PCE estimates (orange bars) versus the reference values (black bars).

Table 8.4: Simply supported truss structure – imprecise Sobol' indices (first order) – augmented PCE results ($N = 100$, $n_{ph} = 20$).

Variable	$\underline{S}_i^{(\mathrm{P})}$	$\overline{S}_i^{(\mathrm{P})}$
P_1	0.0195	0.0550
P_2	0.0702	0.1809
P_3	0.1344	0.3124
P_4	0.1895	0.4060
P_5	0.1344	0.3123
P_6	0.0702	0.1809
P_7	0.0196	0.0550

Pinching analysis

Augmented PCE analysis The augmented PCE model calibrated for the imprecise Sobol' indices analysis is used in the pinching analysis. In order to ensure an accurate approximation of the augmented computational model $\mathcal{M}^{(\mathrm{aug})}$, $n_{ph} = 20$ is used in the following presentation of results.

Results summary Reusing the same augmented PCE model allows for an efficient estimation of the pinched response p-boxes and the corresponding $A_{Y_{\check{X}_i}}$. The results are summarized in Table 8.5, which shows the VoI of Eq. (8.2) alongside a ranking of the variables. The ranking sorts the variables in ascending order from large to small VoI. As expected, the closer a load to the symmetry axis of the truss structure (*i.e.* closer to u_4), the larger its influence on the epistemic uncertainty in u_4 and hence the larger the VoI.

Table 8.5: Simply supported truss structure – pinching of parametric p-boxes – augmented PCE results ($N = 100$, $n_{ph} = 20$) – VoI according to Eq. (8.2).

$\check{P}_i$	$\widehat{A}_{Y_{\check{X}_i}}$	$\widehat{\mathrm{VoI}}$	VoI	Ranking
P_1	$2.65 \cdot 10^{-3}$	$0.068 \approx 7$ %	0.069	#4
P_2	$2.46 \cdot 10^{-3}$	$0.133 \approx 13$ %	0.134	#3
P_3	$2.31 \cdot 10^{-3}$	$0.187 \approx 19$ %	0.188	#2
P_4	$2.21 \cdot 10^{-3}$	$0.223 \approx 22$ %	0.224	#1
P_5	$2.31 \cdot 10^{-3}$	$0.187 \approx 19$ %	0.188	#2
P_6	$2.46 \cdot 10^{-3}$	$0.133 \approx 13$ %	0.134	#3
P_7	$2.65 \cdot 10^{-3}$	$0.068 \approx 7$ %	0.069	#4
none	$2.84 \cdot 10^{-3}$	0.000	0.000	-

Graphical comparison Figure 8.10 illustrates the response p-boxes when (i) the original input p-boxes is used, (ii) when P_4 is pinched and the other variables are parametric p-boxes. In particular, the reference results (black lines) are compared to the augmented PCE-based results (orange lines). The boundary curves of the original and pinched response p-boxes are overlapping, indicating a remarkably accurate pinching analysis. Despite the small experimental design of $N = 100$ LHS samples, the response p-box as well as the VoI are estimated accurately.

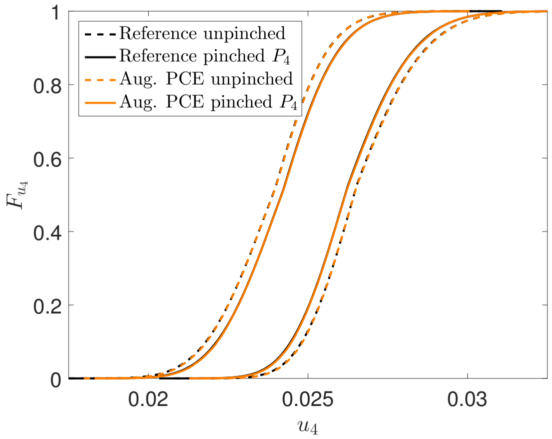


Figure 8.10: Simply supported truss structure – response p-boxes – original versus pinched p-boxes ($\check{X}_i = P_4$) – reference solution versus augmented PCE-based solution ($N = 100$, $n_{ph} = 20$).

Comparison to imprecise Sobol' indices Comparing Table 8.3 and Table 8.5, the imprecise Sobol' indices analysis and the pinching algorithm result in comparable results. The influential imprecision can be found in the midspan loading variable P_4. Moreover, the symmetry of the problem setting is confirmed by both analysis methods. As the augmented PCE methods coincide with the reference solution, the meta-modelling approach is capable of estimating the imprecise sensitivity measures accurately.

8.5 Discussion

In the context of p-boxes, a set of algorithms is proposed to compute the VoI in pinching analyses efficiently. The low computational costs are achieved by an extensive use of meta-models. The types and number of meta-models is summarized in Table 8.6. In all but one scenario, two levels of meta-models are used. Moreover, the ISA algorithms are closely related to previously discussed algorithms in the context of IUP and ISRA. The details of those algorithms are presented in the corresponding chapters of this thesis. Hence, the ISA algorithms presented here can be interpreted as an extension of the previously developed algorithms.

Table 8.6: ISA - pinching – comparison of free and parametric p-box approaches.

	Free p-box	Parametric p-box
P-box	$\underline{F}_{\boldsymbol{X}}, \overline{F}_{\boldsymbol{X}}$	$F_{\boldsymbol{X}}(\boldsymbol{x}\|\boldsymbol{\theta}), \boldsymbol{\theta} \in \mathcal{D}_{\boldsymbol{\Theta}}$
VoI of Eq. (8.2)	2 × PCE	Augmented PCE
Algorithm based on	Section 6.2	Section 6.4
VoI of Eq. (8.4)	2 × AK-MCS	AK-MCS + EGO
Algorithm based on	Section 7.3	Section 7.4

As a special case for parametric p-boxes, imprecise Sobol' indices are introduced as a second measure for sensitivity. This additional sensitivity measure compares the epistemic and aleatory uncertainty in a natural way. Furthermore, based on the augmented PCE model developed in Section 7.3, the VoI defined in Eq. (8.2) can be estimated efficiently. In fact, the same meta-model can be used as in the IUP analysis.

The proposed SA estimate the influence of aleatory and epistemic uncertainty in the input variables onto the output variable. However, considering that it is common practice to work solely with probabilistic input variables, the introduction of (interval-valued) epistemic uncertainty in the input variables and the consequent imprecise response variable can be interpreted as a sensitivity measure for probabilistic analyses. In particular, p-boxes (free and parametric) lead to bounds on the QoI. Then, the amount of imprecision in the QoI can be interpreted as a measure for the reliability of the analysis. In other words, when the QoI represent small imprecision compared to a given criterion, the results are insensitive to potential inaccuracies in the definition of the probabilistic variables.

CHAPTER 9

NASA uncertainty quantification challenge

The NASA uncertainy quantification challenge is a comprehensive challenge set for imprecise probabilities in aeronautical engineering. In this chapter, different aspects of the cheallenge set are solved with the frameowrk presented in this thesis and the results are discussed. In comparison to the application examples in the previous chapters, reference solutions are not available for this case study. Hence, it is important that the analyses lead to accurate and reliable results.

9.1 Context

The *NASA Langley Multidisciplinary Uncertainty Quantification Challenge* is a comprehensive challenge set for imprecise probabilities in aeronautical engineering proposed by Crespo, Kenny, & Giesy (2013). The details and software pieces can be found on the corresponding NASA website: https://uqtools.larc.nasa.gov/nda-uq-challenge-problem-2014/. The challenge set deals with key aspects of uncertainty quantification, namely uncertainty characterization, sensitivity analysis, uncertainty propagation, extreme-case analysis, and robust design. The various tasks include:

- Model updating: improvement of an initial uncertainty model by experimental data.
- Sensitivity analysis: determination of the key input variables and potential exclusion of unimportant variables with respect to different statistics of interest.
- Investment strategy: decisions are to be made as to which uncertainty models should be improved under constraint resources in order to reduce the uncertainty in the system efficiently.
- Uncertainty propagation: comparison of different input models and the effect on statistics of interest.

- Structural identification and structural reliability analysis: identification of the uncertainty models that yield to extreme response quantities.
- Robust design: determination of the best design point under epistemic (as well as aleatory) uncertainty.

9.2 Problem statement

9.2.1 Introduction

The computational model describes the dynamic of a *generic transport model*, *i.e.* a remotely operated twin-jet aircraft developed by NASA Langley Research Center. A typical test campaign consists of manually piloting the aircraft to save flying conditions controlled from the ground via radio frequency links. Of particular interest is the aircraft's behaviour when flying outside of the normal flying envelope. A set of requirements describes then the vehicles stability and performance characteristics in regard to pilot command tracking and handling qualities.

9.2.2 Computational model

The challenge set is based on a set of encrypted Matlab functions. Hence, the computational model is a true black box. The limit-state function is defined as:

$$\boldsymbol{g}_{\text{NASA}} = \boldsymbol{f}\left(\boldsymbol{x}, \boldsymbol{d}\right), \tag{9.1}$$

where $\boldsymbol{g}_{\text{NASA}} = \left(g_1, \ldots, g_8\right)^{\mathsf{T}}$ is the performance vector, $\boldsymbol{d} = \left(d_1, \ldots, d_{14}\right)^{\mathsf{T}}$ is a vector of design parameters, and $\boldsymbol{x} = \left(x_1, \ldots, x_5\right)^{\mathsf{T}}$ is a vector of intermediate variables. Note that the safety domain is defined as $\max_i \left[g_i\right] < 0$ and the failure domain is defined as $\max_i \left[g_i\right] \geq 0$. The intermediate variables are the result of five fixed discipline analyses and are defined as:

$$\boldsymbol{x} = \boldsymbol{h}\left(\boldsymbol{p}\right), \tag{9.2}$$

where $\boldsymbol{p} = \left(p_1, \ldots, p_{21}\right)^{\mathsf{T}}$ is a vector of uncertain parameters in $\mathbb{R}^{21}$. Due to the fixed discipline analyses, the following relationships between p_i and x_j are known:

$$\begin{aligned}
x_1 &= h_1\left(p_1, p_2, p_3, p_4, p_5\right), && (9.3)\\
x_2 &= h_2\left(p_6, p_7, p_8, p_9, p_{10}\right), && (9.4)\\
x_3 &= h_3\left(p_{11}, p_{12}, p_{13}, p_{14}, p_{15}\right), && (9.5)\\
x_4 &= h_4\left(p_{16}, p_{17}, p_{18}, p_{19}, p_{20}\right), && (9.6)\\
x_5 &= h_5\left(p_{21}\right). && (9.7)
\end{aligned}$$

Based on the limit-state function, two additional performance measures are defined:

$$\begin{aligned}
J_1 &= \mathbb{E}\left[w\left(\boldsymbol{p}, \boldsymbol{d}_{\text{baseline}}\right)\right], && (9.8)\\
J_2 &= 1 - \mathbb{P}\left(w\left(\boldsymbol{p}, \boldsymbol{d}_{\text{baseline}}\right) < 0\right), && (9.9)
\end{aligned}$$

where $\mathbb{E}\left[\cdot\right]$ denotes the expectation value with respect to probabilistic uncertainty and w describes the worst-case requirement metric:

$$w\left(\boldsymbol{p}, \boldsymbol{d}\right) = \max_{i=1,\ldots,8} g_i = \max_{i=1,\ldots,8} f_i\left(\boldsymbol{h}\left(\boldsymbol{p}\right), \boldsymbol{d}\right). \tag{9.10}$$

Figure 9.1 summarizes the computational models and the corresponding input and output quantities in a flowchart. The two white boxes mark the input parameters, whereas the grey boxes mark the computational models/limit-state functions. Note that the design parameters have fixed deterministic baseline values which are used throughout the following analyses.

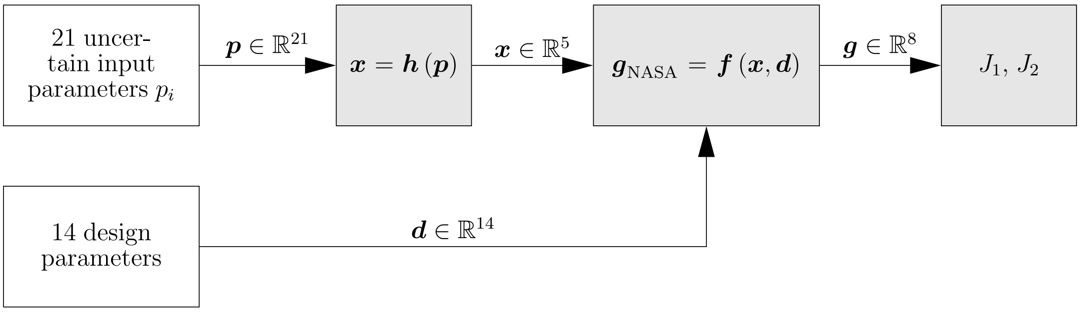


Figure 9.1: NASA challenge set – input/output variables and the connection of the computational models.

9.2.3 Input modelling

Each of the input variables p_i, $i = 1, \ldots, 21$ is modelled as either (i) a probability distribution (purely aleatory uncertainty, type I), (ii) an interval (purely epistemic uncertainty, type II), or (iii) a parametric p-box (type III).

The detailed description of the 21 input variables is listed in Table 9.1. Considering the three types of variables, the input vector $\boldsymbol{p}$ can be decomposed into 17 aleatory and 31 epistemic variables, which are denoted by a_i and e_i, respectively. The aleatory uncertainty is described by uniform, Gaussian and beta distributions. Note that the PDF of the beta distribution is parametrized by $f_{\mathcal{B}}(x) \propto (x-a)^{r-1}(b-x)^{s-1}$, where $x \in [a, b]$ describes the support of this distribution and r and s are the shape parameters.

9.3 Uncertainty characterization

9.3.1 Tasks

Consider the computational model defined in Eq. (9.3). The following challenge tasks read:

(i) Provided are 25 observations of x_1 obtained through the true uncertainty model, where the input variables have no epistemic uncertainty. Use this information to improve the uncertainty model of the category II and III parameters. In other words, reduce the interval widths of $e_1, \ldots, e_8$ where applicable.

(ii) Use an additional 25 observations to validate the models found in (i).

(iii) Improve the uncertainty model further by using all 50 samples.

9.3.2 Solution algorithm

As the input distributions consist of parametric p-boxes, a non-parametric statistic method based on the Kolmogorov-Smirnov (KS) test is used to improve the uncertainty model (see also Section 2.4.2). The workflow consists of the following main steps:

Table 9.1: NASA challenge set – input definition.

p_i	Type	Aleatory variable		Epistemic variable	
p_1	III	$a_1 \sim \mathcal{B}(0,1)$	Beta	$e_1 \in [3/5, 4/5]$	Mean value $\mathbb{E}[p_1]$
				$e_2 \in [1/50, 1/25]$	Variance $\mathrm{Var}[p_1]$
p_2	II			$e_3 \in [0,1]$	Interval
p_3	I	$a_2 \sim \mathcal{U}(0,1)$	Uniform		
p_4	III	$a_3 \sim \mathcal{N}$	Gaussian	$e_4 \in [-5, 5]$	Mean value $\mathbb{E}[p_4]$
				$e_5 \in [1/400, 4]$	Variance $\mathrm{Var}[p_4]$
p_5	III	$a_4 \sim \mathcal{N}$	Gaussian	$e_6 \in [-5, 5]$	Mean value $\mathbb{E}[p_5]$
				$e_7 \in [1/400, 4]$	Variance $\mathrm{Var}[p_5]$
p_4, p_5				$e_8 \in [-1, 1]$	Correlation $\rho(p_4, p_5)$
p_6	II			$e_9 \in [0,1]$	Interval
p_7	III	$a_5 \sim \mathcal{B}(0,1)$	Beta	$e_{10} \in [0.982, 3.537]$	Distr. parameter r_7
				$e_{11} \in [0.619, 1.080]$	Distr. parameter s_7
p_8	III	$a_6 \sim \mathcal{B}(0,1)$	Beta	$e_{12} \in [7.450, 14.093]$	Distr. parameter r_8
				$e_{13} \in [4.285, 7.864]$	Distr. parameter s_8
p_9	I	$a_7 \sim \mathcal{U}(0,1)$	Uniform		
p_{10}	III	$a_8 \sim \mathcal{B}(0,1)$	Beta	$e_{14} \in [1.520, 4.513]$	Distr. parameter r_{10}
				$e_{15} \in [1.536, 4.750]$	Distr. parameter s_{10}
p_{11}	I	$a_9 \sim \mathcal{U}(0,1)$	Uniform		
p_{12}	II			$e_{16} \in [0,1]$	Interval
p_{13}	III	$a_{10} \sim \mathcal{B}(0,1)$	Beta	$e_{17} \in [0.412, 0.737]$	Distr. parameter r_{13}
				$e_{18} \in [1.000, 2.068]$	Distr. parameter s_{13}
p_{14}	III	$a_{11} \sim \mathcal{B}(0,1)$	Beta	$e_{19} \in [0.931, 2.169]$	Distr. parameter r_{14}
				$e_{20} \in [1.000, 2.407]$	Distr. parameter s_{14}
p_{15}	III	$a_{12} \sim \mathcal{B}(0,1)$	Beta	$e_{21} \in [5.435, 7.095]$	Distr. parameter r_{15}
				$e_{22} \in [5.287, 6.945]$	Distr. parameter s_{15}
p_{16}	II			$e_{23} \in [0,1]$	Interval
p_{17}	III	$a_{13} \sim \mathcal{B}(0,1)$	Beta	$e_{24} \in [1.060, 1.662]$	Distr. parameter r_{17}
				$e_{25} \in [1.000, 1.488]$	Distr. parameter s_{17}
p_{18}	III	$a_{14} \sim \mathcal{B}(0,1)$	Beta	$e_{26} \in [1.000, 4.266]$	Distr. parameter r_{18}
				$e_{27} \in [0.553, 1.000]$	Distr. parameter s_{18}
p_{19}	I	$a_{15} \sim \mathcal{U}(0,1)$	Uniform		
p_{20}	III	$a_{16} \sim \mathcal{B}(0,1)$	Beta	$e_{28} \in [7.530, 13.492]$	Distr. parameter r_{20}
				$e_{29} \in [4.711, 8.148]$	Distr. parameter s_{20}
p_{21}	III	$a_{17} \sim \mathcal{B}(0,1)$	Beta	$e_{30} \in [0.421, 1.000]$	Distr. parameter r_{21}
				$e_{31} \in [7.772, 29.621]$	Distr. parameter s_{21}

1. Repeat for $i = 1, \dots, k$, where k is sufficiently large:
    (a) Generate a realization of the epistemic variables $\boldsymbol{e}$.
    (b) Propagate the resulting probability distributions through the computational model $\boldsymbol{h}$ and store the resulting probability distribution.
    (c) Determine the maximal (vertical) distance between the empirical CDF of the dataset at hand and the response probability distribution in a variable-CDF plot.
    (d) The threshold value for the distance is looked up in tables such as in Massey (1951), of which the values used here are listed in Table 9.2. If the maximum distance exceeds $D_{\alpha,n}$ (corresponding to a significance level α), reject this realization $\boldsymbol{e}$, otherwise accept it as a feasible vector of distribution parameters for generating the dataset.
2. Plot the rejected versus accepted vectors $\boldsymbol{e}$ in a parallel plot in order to discover patterns.

In order to obtain robust results, repeat the same analysis with different confidence levels in the KS-test.

Table 9.2: Kolmogorov-Smirnov test – maximal distances $D_{\alpha,n}$ for a level of significance α and a sample size of n (literature values by Massey (1951)).

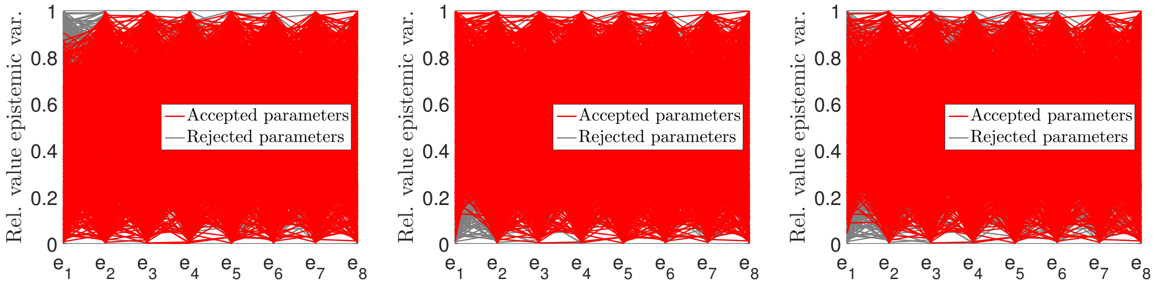
	Literature values		Matlab	
α	$n = 25$	$n = 50$	$n = 25$	$n = 50$
0.01	0.32	0.23	0.3166	0.2260
0.05	0.27	0.19	0.2641	0.1884
0.10	0.24	0.17	0.2377	0.1696
0.20	0.21	0.15	0.2080	0.1484

9.3.3 Results

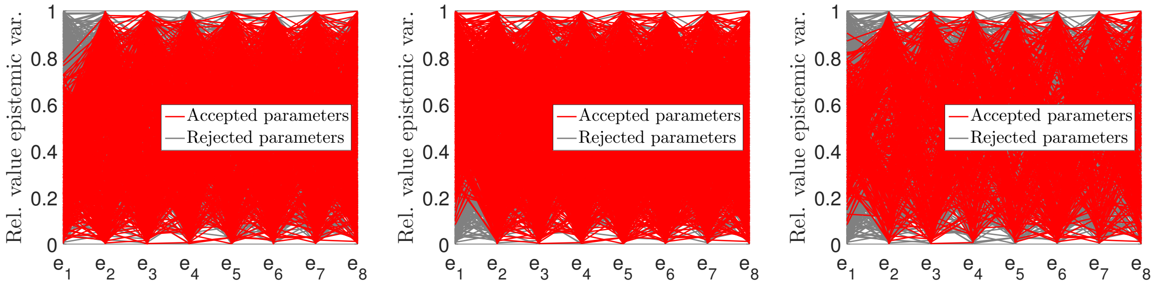
Figure 9.2 displays the parallel plots for the first, second, and combined datasets (for $k = 1000$ samples of $\boldsymbol{e}$). On the vertical axis, a relative value of the epistemic variables are displayed, *i.e.* 0 marks the lower boundary value and 1 marks the upper boundary value of the epistemic intervals. The red curves show the accepted parameter vectors $\boldsymbol{e}$ whereas the grey curves mark the rejected parameter vectors $\boldsymbol{e}$. As expected, the number of accepted parameter combinations reduces when the significance level is increased. For $\alpha \geq 0.1$, there is a significant reduction visible for two epistemic variables, namely e_1 and e_6, which correspond to $\mathbb{E}[p_1]$ and $\mathbb{E}[p_5]$, respectively. For the other epistemic variables, there is no large reduction visible.

Figure 9.3 shows the response p-box of the accepted CDF curves in comparison with the original response p-box. Again, the red area denotes the box of accepted CDF curves, whereas the grey area denoted the box of rejected CDF curves. Additionally, the empirical CDF of the corresponding dataset is shown by the solid black line. The accepted CDF curves envelope the empirical CDF in all cases well. Furthermore, the larger α, the narrower the band of accepted CDFs, which resembles the behaviour in Figure 9.2.

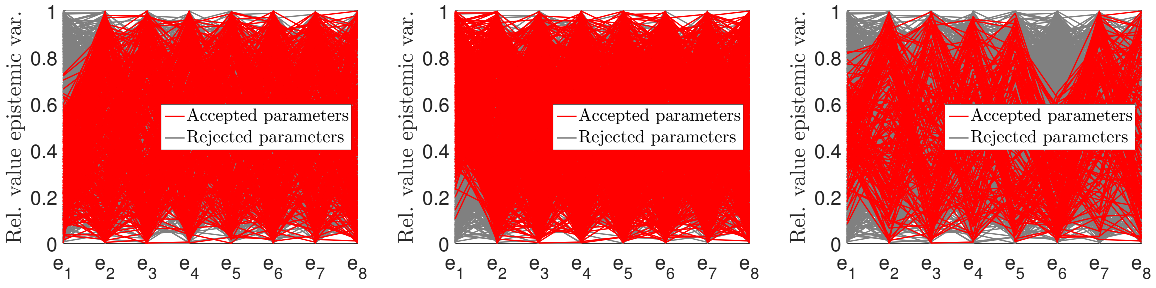
In order to improve the intervals of the epistemic variables $e_1, \dots, e_8$, new boundary values are chosen as the minimum and maximum values of the accepted parameter values in Figure 9.2. Using the results of $\alpha = 0.10$ and engineering judgement, the improved input model is shown in Table 9.3. Note that the solution to (i) can be found in the column $n = 25$, whereas the solution to (iii) can be found in the column $n = 50$. For the two epistemic uncertainties $\mathbb{E}[p_1]$



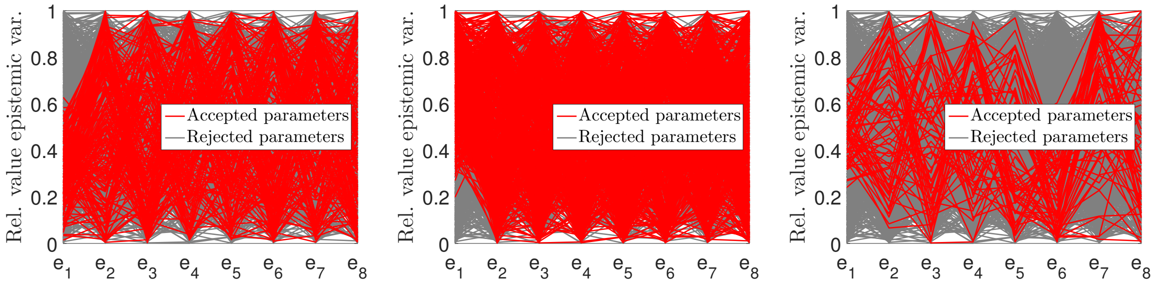
(a) First dataset ($\alpha = 0.01$) (b) Second dataset ($\alpha = 0.01$) (c) Both datasets ($\alpha = 0.01$)



(d) First dataset ($\alpha = 0.05$) (e) Second dataset ($\alpha = 0.05$) (f) Both datasets ($\alpha = 0.05$)



(g) First dataset ($\alpha = 0.10$) (h) Second dataset ($\alpha = 0.10$) (i) Both datasets ($\alpha = 0.10$)



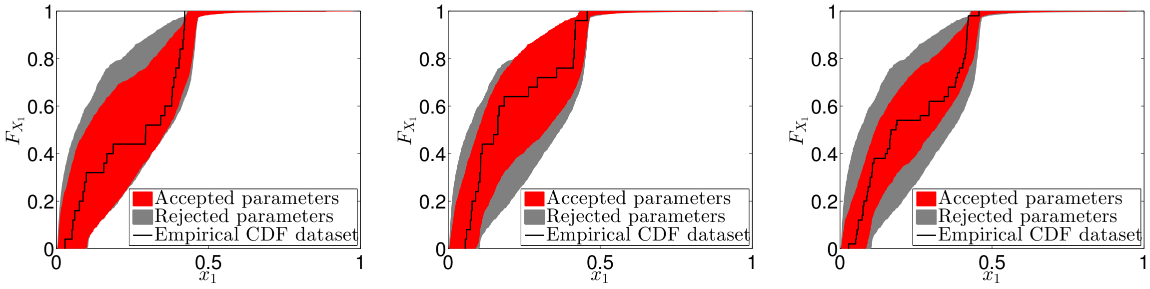
(j) First dataset ($\alpha = 0.20$) (k) Second dataset ($\alpha = 0.20$) (l) Both datasets ($\alpha = 0.20$)

Figure 9.2: NASA challenge set – uncertainty characterization – parallel plots.

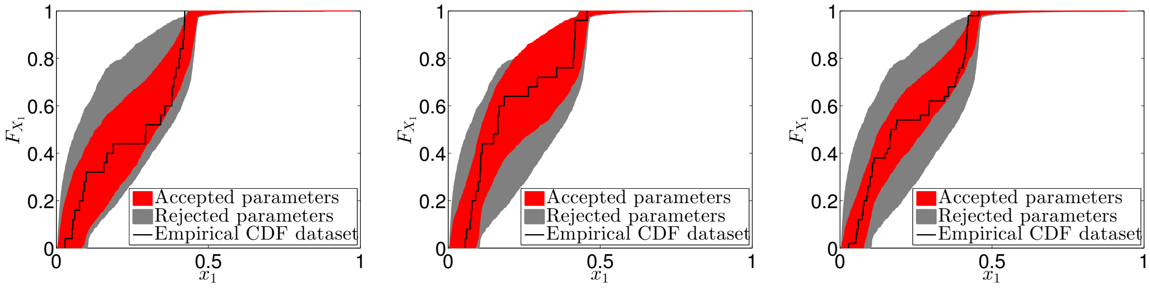
and $\mathbb{E}[p_5]$, a reduction in the interval width is proposed based on the case of $n = 50$ data points. The reduced values are used in the subsequent analysis where applicable.

Table 9.3: NASA challenge set – uncertainty characterization – reduced input intervals ($\alpha = 0.1$).

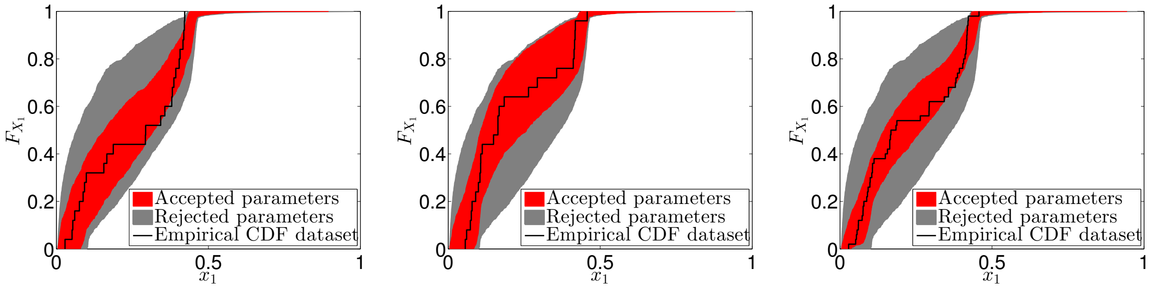
e_i	Description	Initial range	$n = 25$	$n = 50$
e_1	$\mathbb{E}[p_1]$	$[0.6, 0.8]$	$[0.6, 0.75]$	$[0.62, 0.76]$
e_6	$\mathbb{E}[p_5]$	$[-5, 5]$	$[-5, 5]$	$[-5, 1.5]$



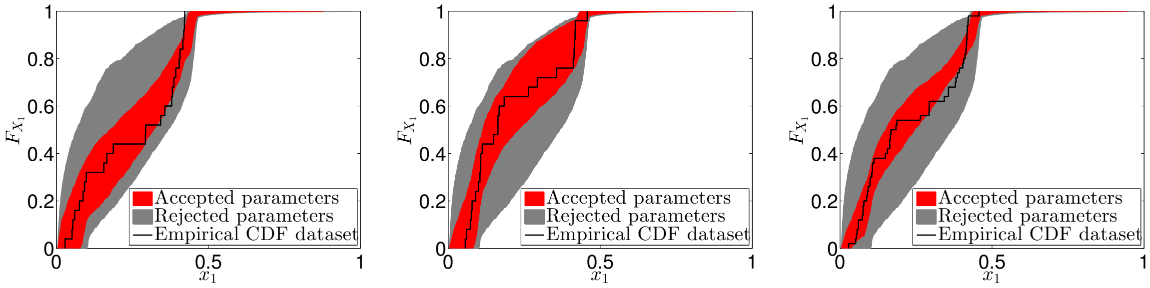
(a) First dataset ($\alpha = 0.01$) (b) Second dataset ($\alpha = 0.01$) (c) Both datasets ($\alpha = 0.01$)



(d) First dataset ($\alpha = 0.05$) (e) Second dataset ($\alpha = 0.05$) (f) Both datasets ($\alpha = 0.05$)



(g) First dataset ($\alpha = 0.10$) (h) Second dataset ($\alpha = 0.10$) (i) Both datasets ($\alpha = 0.10$)



(j) First dataset ($\alpha = 0.20$) (k) Second dataset ($\alpha = 0.20$) (l) Both datasets ($\alpha = 0.20$)

Figure 9.3: NASA challenge set – uncertainty characterization – response p-boxes compared to empirical CDF.

9.4 Sensitivity analysis

9.4.1 Tasks

For sensitivity analysis, consider the input vector $\boldsymbol{p} \in \mathbb{R}^{21}$ and the computational model defined in Eq. (9.2). Further, the results of Section 9.3 should be considered as a stating point for the following tasks:

(v) Rank the four imprecise variables in $x_1 = h_1(p_1, p_2, p_3, p_4, p_5)$ according to the degree of refinement in the p-box of x_1 which one could hope to obtain by refining the uncertainty models. Is it possible to set any parameters constant without introducing a significant error? If so, evaluate and bound the error. Do the same for x_2, x_3, and x_4 and the corresponding variables p_i.

(vi) Rank the 17 imprecise variables according to the reduction in the range of J_1 in Eq. (9.8). Are there any parameters that can take a constant value without introducing a significant error?

(vii) Do the same as in (vi) for J_2 in Eq. (9.9).

(viii) Compare the rankings.

9.4.2 Solution algorithm

The tasks in this section are solved by using the pinching method (see also Section 8.3.1). For subtask (v), the VoI defined in Eq. (8.2) is used. The epistemic variables e_i are pinched one-by-one at their central value of the interval and the area of the response p-box $A_{\check{e}_i}$ is measured. For subtask (vi), a new VoI measure is introduced, because Eq. (8.4) does not fit in this example. The epistemic variables are pinched one-by-one at their central value of the interval and the change in the response range of J_1 and J_2 is measured. Then, the VoI for J_1 is defined as:

$$\mathrm{VoI}^{(J_1)}\left(J_{1,0}, J_{1,\check{e}_i}\right) = 1 - \frac{\overline{J}_{1,\check{e}_i} - \underline{J}_{1,\check{e}_i}}{\overline{J}_{1,0} - \underline{J}_{1,0}}, \tag{9.11}$$

where $J_{1,\check{e}_i} \in \left[\underline{J}_{1,\check{e}_i}, \overline{J}_{1,\check{e}_i}\right]$ is the pinched interval and $J_{1,0} \in \left[\underline{J}_{1,0}, \overline{J}_{1,0}\right]$ is the unpinched interval of J_1. In a similar manner, the VoI for J_2 is defined as:

$$\mathrm{VoI}^{(J_2)}\left(J_{2,0}, J_{2,\check{e}_i}\right) = 1 - \frac{\overline{J}_{2,\check{e}_i} - \underline{J}_{2,\check{e}_i}}{\overline{J}_{2,0} - \underline{J}_{2,0}}, \tag{9.12}$$

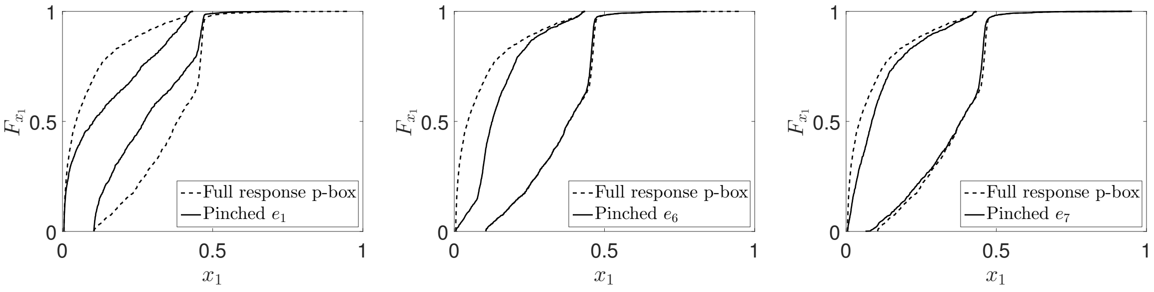
where $J_{2,\check{e}_i} \in \left[\underline{J}_{2,\check{e}_i}, \overline{J}_{2,\check{e}_i}\right]$ is the pinched interval and $J_{2,0} \in \left[\underline{J}_{2,0}, \overline{J}_{2,0}\right]$ is the unpinched interval of J_2. Note that for pinching J_2 representing a probability, the previously defined Eq. (8.4) is not used here due to the anticipated wide range for $J_2 \in [0, 1]$.

9.4.3 Pinching results of x_i

The computational model defined in Eq. (9.2) is an inexpensive-to-evaluate function. Hence, the pinching analysis is conducted without the use of a meta-model but rather with brute force Monte Carlo simulation. A further simplification is made by taking into account the dependencies of the output x_i on the input p_j as indicated in Eqs. (9.3) - (9.7).

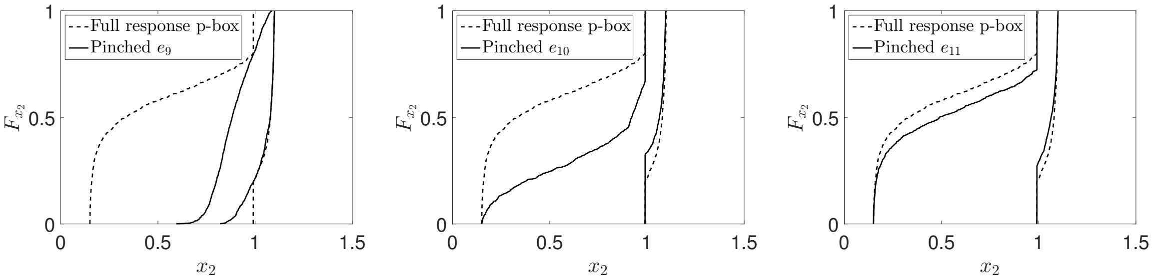
Figures 9.4 through 9.8 compare the pinched and the original response p-boxes in terms of x_i. The resulting p-boxes are plotted for the variables e_i that lead to the largest reduction of epistemic uncertainty according to the VoI defined in Eq. (8.2). Note that each epistemic uncertainty e_i is pinched separately. Based on the pictures, the largest reduction in imprecision in the response p-box can be observed for the variables $\{e_1, e_6, e_9, e_{10}, e_{11}, e_{23}, e_{31}\}$, which correspond to the physical input parameters $\{\mathbb{E}\left[p_1\right], \mathbb{E}\left[p_5\right], p_6, r_7, s_7, p_{16}, s_{21}\}$, respectively.

The quantitative values of Eq. (8.2) are summarized in Tables 9.4 through 9.8. Additionally, the epistemic variables are ranked according to VoI for each response variable x_i. The tables



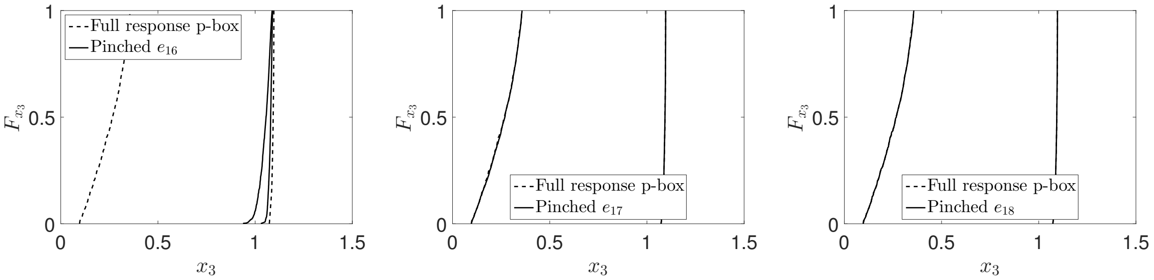
(a) $\mathbb{E}[p_1] = e_1 = 0.7$ (b) $\mathbb{E}[p_5] = e_6 = 0$ (c) $\mathrm{Var}[p_5] = e_7 = 2.0013$

Figure 9.4: NASA challenge – sensitivity analysis – pinching for x_1.



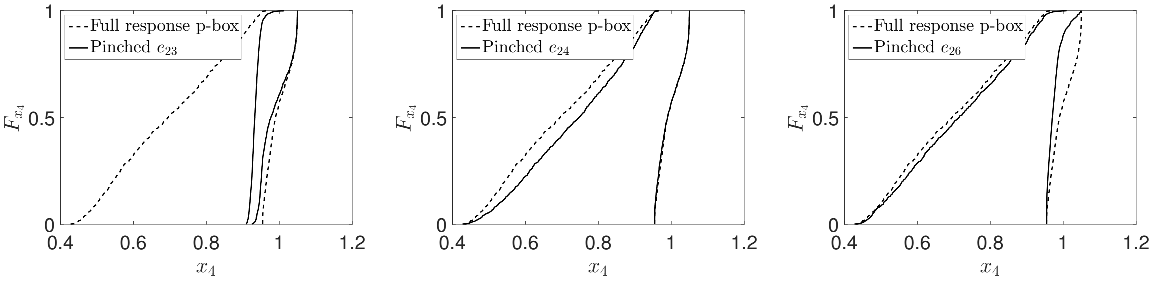
(a) $p_6 = e_9 = 0.5$ (b) $r_7 = e_{10} = 2.2595$ (c) $s_7 = e_{11} = 0.8495$

Figure 9.5: NASA challenge – sensitivity analysis – pinching for x_2.



(a) $p_{12} = e_{16} = 0.5$ (b) $r_{13} = e_{17} = 0.5745$ (c) $s_{13} = e_{18} = 1.534$

Figure 9.6: NASA challenge – sensitivity analysis – pinching for x_3.

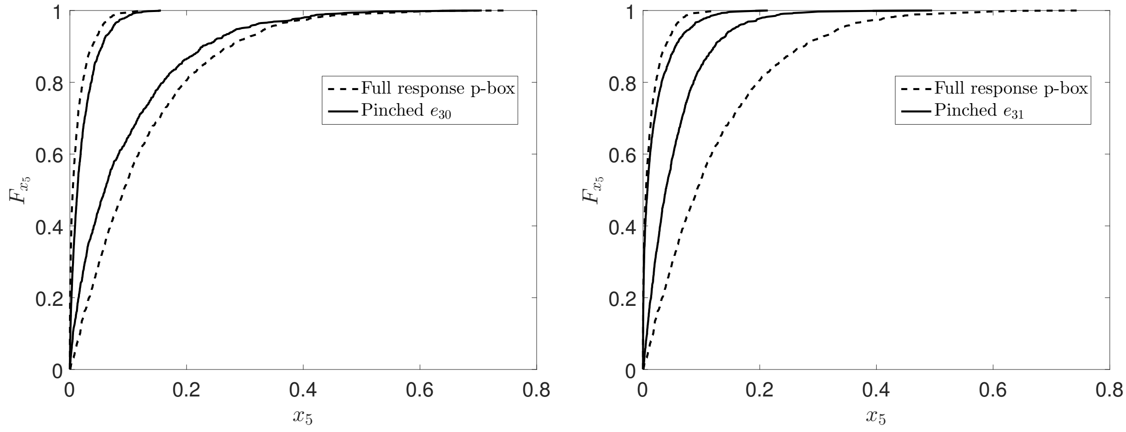


(a) $p_{16} = e_{23} = 0.5$ (b) $r_{17} = e_{24} = 1.3610$ (c) $r_{18} = e_{26} = 2.633$

Figure 9.7: NASA challenge – sensitivity analysis – pinching for x_4.

show that for each x_i, a reduced number of input parameters is responsible for the response epistemic uncertainty. According to the pinching algorithm, the variables

$$\{e_3,\, e_5,\, e_8,\, e_{12},\, e_{13},\, e_{14},\, e_{15},\, e_{17},\, e_{18},\, e_{19},\, e_{20}, e_{21},\, e_{22},\, e_{28},\, e_{29}\}$$



(a) $r_{21} = e_{30} = 0.7105$ (b) $s_{21} = e_{31} = 18.6965$

Figure 9.8: NASA challenge – sensitivity analysis – pinching for x_5.

do not contribute significantly to the imprecision in the corresponding response parametric p-boxes.

Table 9.4: NASA challenge – sensitivity analysis – pinching for x_1.

Variable	$\check{e}_i$	$A_{\check{e}_i}/A_0$	VoI	Ranking
$\mathbb{E}[p_1]$	e_1	0.49	0.51	#1
$\mathrm{Var}[p_1]$	e_2	0.91	0.09	#4
p_2	e_3	0.99	0.01	#6
$\mathbb{E}[p_4]$	e_4	0.94	0.06	#5
$\mathrm{Var}[p_4]$	e_5	0.99	0.01	#6
$\mathbb{E}[p_5]$	e_6	0.79	0.21	#2
$\mathrm{Var}[p_5]$	e_7	0.89	0.11	#3
$\rho(p_4, p_5)$	e_8	1.00	0.00	#8

Table 9.5: NASA challenge – sensitivity analysis – pinching for x_2.

Variable	$\check{e}_i$	$A_{\check{e}_i}/A_0$	VoI	Ranking
p_6	e_9	0.28	0.72	#1
r_7	e_{10}	0.56	0.44	#2
s_7	e_{11}	0.84	0.16	#3
r_8	e_{12}	0.95	0.05	#4
s_8	e_{13}	0.96	0.04	#5
r_{10}	e_{14}	1.00	0.00	#7
s_{10}	e_{15}	0.99	0.01	#6

9.4.4 Pinching results of J_1 and J_2

Compared to the computational model defined in Eq. (9.2), the computational model defined in Eqs. (9.8) and (9.9) are expensive-to-evaluate models. Hence, a brute force MC simulation is not applicable in this case. In order to make the pinching analysis tractable, sparse PCE models for the mapping $\boldsymbol{e} \mapsto J_1$ and $\boldsymbol{e} \mapsto J_2$ are calibrated.

Table 9.6: NASA challenge – sensitivity analysis – pinching for x_3.

Variable	$\check{e}_i$	$A_{\check{e}_i}/A_0$	VoI	Ranking
p_{12}	e_{16}	0.03	0.97	#1
r_{13}	e_{17}	1.00	0.00	#2
s_{13}	e_{18}	1.00	0.00	#2
r_{14}	e_{19}	1.00	0.00	#2
s_{14}	e_{20}	1.00	0.00	#2
r_{15}	e_{21}	1.00	0.00	#2
s_{15}	e_{22}	1.00	0.00	#2

Table 9.7: NASA challenge – sensitivity analysis – pinching for x_4.

Variable	$\check{e}_i$	$A_{\check{e}_i}/A_0$	VoI	Ranking
p_{16}	e_{23}	0.18	0.82	#1
r_{17}	e_{24}	0.85	0.15	#2
s_{17}	e_{25}	0.90	0.10	#4
r_{18}	e_{26}	0.86	0.14	#3
s_{18}	e_{27}	0.94	0.06	#5
r_{20}	e_{28}	0.96	0.04	#6
s_{20}	e_{29}	0.96	0.04	#6

Table 9.8: NASA challenge – sensitivity analysis – pinching for x_5.

Variable	$\check{e}_i$	$A_{\check{e}_i}/A_0$	VoI	Ranking
r_{21}	e_{30}	0.63	0.37	#2
s_{21}	e_{31}	0.32	0.68	#1

Due to the large dimensionality of $\boldsymbol{e}$ (*i.e.* $|\boldsymbol{e}| = 31$), the computational model is complex. In order to reduce the complexity, some input ranges are set constant. In particular, if e_i was identified an unimportant in subtask (v), it is set to its central value in the PCE model to reduce the number of random variables of the latter. Based on Tables 9.4 - 9.8, the twelve remaining epistemic variables are

$$\{e_1,\, e_6,\, e_7,\, e_9,\, e_{10},\, e_{11},\, e_{16},\, e_{23},\, e_{24},\, e_{26},\, e_{30},\, e_{31}\}\,,$$

when considering VoI > 0.1. The PCE model is then calibrated with an experimental design of $N = 10^5$ MC samples. Due to the non-linearity of the computational model, the sparse PCE model results in relatively large empirical errors of $\widehat{err}_{emp} = 0.65$ and $\widehat{err}_{emp} = 0.21$ for J_1 and J_2, respectively. Nevertheless, the pinching algorithm is applied onto the meta-models.

Table 9.9 summarizes the results of the pinching analysis for J_1 and J_2. For each p_i, the resulting range in J_i and the corresponding VoI are shown, as well as an overall ranking. Note that in contrast to the previous subtask (v), the parameters p_i are pinched to a single value/distribution whether they are modelled by one or two epistemic uncertainties, e_j. Hence, the pinched variables are denoted by $\check{p}_i$ rather than $\check{e}_i$.

As mentioned before, the accuracy of the meta-model is limited. This explains also the negative values for J_2, which are by definition of a probability measure not feasible. Despite the low accuracy of the meta-model, the influence of the input parameters p_i onto the measures J_i is estimated. For both measures, the most influential parameters are marked with a relative ranking. For J_1 and J_2, the influential factors are $\{p_{21},\, p_5,\, p_{12},\, p_1,\, p_{16},\, p_6,\, p_7\}$ and $\{p_{12},\, p_1,\, p_7,\, p_5,\, p_{21}\}$,

Table 9.9: NASA challenge – sensitivity analysis – pinching for J_i – PCE-based results.

	J_1			J_2		
$\check{p}_i$	Range	VoI	Ranking	Range	VoI	Ranking
None	$[-1.10, 3.05]$	0.00	-	$[0.00, 0.73]$	0.00	-
p_1	$[-0.99, 2.78]$	0.09	#4	$[0.06, 0.52]$	0.39	#2
p_5	$[-0.91, 1.33]$	0.46	#2	$[0.020.69]$	0.09	#4
p_6	$[-0.90, 3.05]$	0.05	#6	$[-0.01, 0.71]$	0.03	#6
p_7	$[-1.03, 2.89]$	0.05	#6	$[0.02, 0.68]$	0.11	#3
p_{12}	$[-0.40, 2.79]$	0.23	#3	$[0.01, 0.42]$	0.44	#1
p_{16}	$[-0.96, 2.91]$	0.07	#5	$[0.03, 0.74]$	0.03	#6
p_{21}	$[-0.32, 0.54]$	0.79	#1	$[0.01, 0.71]$	0.04	#5

respectively. Note that many of the influential input parameters p_i appear in both measures as influential.

9.5 Uncertainty propagation

9.5.1 Tasks

The uncertainty propagation tasks are targeted on the two measures J_1 and J_2, defined in Eqs. (9.8) and (9.9), respectively. The following subtasks are to be solved:

(ix) Determine the range of J_1 corresponding to the original uncertainty model and the improved uncertainty model found in subtask (iii).

(x) Do the same for J_2.

(xi) Select four imprecise random variables according to the rankings in subtasks (vi) and (vii). Request an improved model from the organizers of the challenge.

(xii) Use the reduced model obtained in (xi) to recalculate the ranges of J_1 and J_2.

9.5.2 Solution algorithms

Uncertainty quantification with parametric p-boxes can be solved as discussed in Section 7.4. Due to the complexity of the problem and the definitions of J_1 and J_2, however, a different approach is chosen here. Instead of the original limit-state function defined in Eq. (9.10), an alternative computational model is defined. Consider the direct mapping from the epistemic input vector $\boldsymbol{e}$ to the measures J_1 and J_2. These are stochastic mappings, which map realizations of $\boldsymbol{e}$ to J_i. The stochasticity originates from the sampling-based estimation of J_1 and J_2 and the corresponding expectation value and exceedance probability operation, respectively.

When interpreting the epistemic input parameters as independent uniformly distributed variables, a meta-model model can be used to approximate J_i. Finally, the extreme values of J_i can be obtained by optimization in $\mathcal{D}_{\boldsymbol{e}}$. In order to increase the efficiency of this analysis, the EGO algorithm is used to iteratively enrich the experimental design of the meta-models and finally to determine the bounds of J_i.

EGO is based on Kriging and is hence more efficient for low-dimensional objective functions. Based on the results of the sensitivity analysis, the uncertainty propagation analysis is conducted with a reduced epistemic input vector. The remaining interval-valued parameters are then:

$$\{e_1,\, e_6,\, e_7,\, e_9,\, e_{10},\, e_{11},\, e_{16},\, e_{23},\, e_{30},\, e_{31}\}\,,$$

which affect the input parameters:

$$\{p_1,\, p_5,\, p_6,\, p_7,\, p_{12},\, p_{16},\, p_{21}\}\,.$$

Note that in parameter p_{20}, the epistemic uncertainty is neglected due to the result of Table 9.7. Due to the simpler shape of the optimization domain $\mathcal{D}_{\boldsymbol{e}}$, *i.e.* smaller dimensionality of $\mathcal{D}_{\boldsymbol{e}}$, the number of function evaluations is expected to be lower than when using all epistemic parameters as intervals.

9.5.3 Results

The ranges for the measures J_1 and J_2 are summarized in Table 9.10. The two measures are estimated with the original uncertainty model as well as the improved model, where some input epistemic intervals are reduced based on the datasets (see values in Table 9.3). Note that the boundary values for J_1 and J_2 are obtained by a maximum of $N^{(\mathrm{EGO})} = 120$ samples in the experimental design. The iterative EGO algorithms are stopped manually.

The results show that the improved model of the input parameters does not change the intervals considerably compared to the original model. As expected, the original model encapsulates the improved model, due to the smaller epistemic uncertainty of the latter. The two models provide rather wide ranges, in particular for the probability J_2. Hence, it is impossible to make decisions based on these results. Further investigations are required to reduce the epistemic uncertainty in the input vector.

Table 9.10: NASA challenge set – uncertainty propagation – results of the adaptive Kriging algorithm.

QoI	Statistics	Original model	Improved model	NASA model
J_1	$\widehat{J}_1$	$[8.40 \cdot 10^{-3},\, 13.1]$	$[1.48 \cdot 10^{-2},\, 12.8]$	$[3.06 \cdot 10^{-2},\, 1.06]$
	$N^{(\mathrm{EGO})}$	$\{120,\, 120\}$	$\{120,\, 120\}$	$\{120,\, 120\}$
J_2	$\widehat{J}_2$	$[0.04,\, 0.89]$	$[0.04,\, 0.85]$	$[0.30,\, 0.40]$
	$N^{(\mathrm{EGO})}$	$\{120,\, 120\}$	$\{120,\, 120\}$	$\{120,\, 120\}$

Based on the results of pinching analysis summarized in Table 9.9, the parameters $\{p_1,\, p_7,\, p_{12},\, p_{21}\}$ are chosen for further refinement. Upon request, the corresponding refined epistemic parameters read: $e_1 \in [3.1/5,\, 3.2/5]$, $e_2 \in [0.8/25,\, 0.9/25]$, $e_{10} \in [2.7,\, 3.1]$, $e_{11} \in [0.9,\, 1.08]$, $e_{16} \in [0.96,\, 1.00]$, $e_{30} \in [0.87,\, 1.00]$, and $e_{31} \in [19,\, 23]$.

Using the reduced uncertainty model and the values obtained upon request, the corresponding ranges for J_1 and J_2 are shown in Table 9.10, denoted by "NASA model". The values are computed with the same algorithm as for the other uncertainty models. Compared to the previous results, the ranges for J_1 and J_2 are considerably reduced. In fact, the range of J_2 is reduced to a value of $\overline{J}_2 - \underline{J}_2 = 0.10$. Hence, the parameters chosen for refinement efficiently reduced the epistemic uncertainty in the statistics of interest, *i.e.* J_1 and J_2.

9.6 Conclusion

As seen in Figure 9.1, the challenge set consists of three different computational models/limit-state functions. In the flow of analyses, the computational model $\boldsymbol{h}$ is an inexpensive-to-evaluate computational model, whereas the computation of J_1 and J_2 is an expensive-to-evaluate operation, depending on the definition of the epistemic uncertainty in the input parameters. Due to the varying complexity, the choice of analysis is crucial to obtain results in an efficient manner. As an example, in the first analyses concerning $\boldsymbol{x} = \boldsymbol{h}(\boldsymbol{p})$, no meta-modelling algorithm was used due to the low computational costs. Moreover, a meta-model is a potential source of model errors. In the pinching of J_i and the final uncertainty propagation analyses, however, meta-models are extensively used to reduce the computational costs. As an example, a single evaluation of J_1 based on 1000 samples of the aleatory uncertainty-related parameters results in a computational time of 33 seconds on a quad-core computer.

In order to further reduce the computational efforts, the analyses of different subtasks are combined. In particular, the information gained from results of previous analyses are used in the definition of the solution algorithm of the next analysis. In doing so, the vector of 21 imprecise input parameters $\boldsymbol{p}$ is reduced in the final uncertainty propagation analysis to a vector of only 10 imprecise input parameters. This reduction allows for a more efficient search algorithm to find the extreme values of J_1 and J_2. In fact, the results confirm the ones found in Patelli, Broggi, & Angelis (2014), Patelli, Alvarez, et al. (2015), and Crespo & Kenny (2015), summarized in Table 9.11, which use similar approaches on the original, non-reduced input vector $\boldsymbol{p}$. Hence, the reduction of the input vector is feasible at this stage of the analysis.

Table 9.11: NASA challenge – comparison to results in Patelli, Broggi, & Angelis (2014).

Model	Quantity	Schöbi	Patelli, Broggi, & Angelis (2014)
Original model	J_1	$[8.40 \cdot 10^{-3}, 13.1]$	$[-1.55 \cdot 10^{-4}, 1.7608]$
	J_2	$[0.04, 0.89]$	$[0, 1]$
Improved model	e_i	$\left\{ \begin{array}{c} e_1 \in [0.62, 0.76] \\ e_6 \in [-5, 1.5] \end{array} \right\}$	$\left\{ \begin{array}{c} e_1 \in [0.63, 0.76] \\ e_6 \in [-4.9, 0.3] \end{array} \right\}$
	J_1	$[1.48 \cdot 10^{-2}, 12.8]$	$[1.35 \cdot 10^{-2}, 4.97]$
	J_2	$[0.04, 0.85]$	$[0.06, 0.82]$
NASA model	p_i	$\{p_1, p_7, p_{12}, p_{21}\}$	$\{p_1, p_4, p_{12}, p_{21}\}$
	J_1	$[3.06 \cdot 10^{-2}, 1.06]$	$[2.88 \cdot 10^{-2}, 1.11]$
	J_2	$[0.30, 0.40]$	$[0.24, 0.38]$

The presented results are obtained mainly with the algorithms developed in this thesis. In particular, the use of meta-models is enforced where applicable. In order to obtain reliable results, however, other methods shall be used for verification. As an example, Nagel & Sudret (2014) analysed the first uncertainty characterization tasks with Bayesian hierarchical models and Bayesian inference. Despite the different uncertainty concept as a starting point, similar conclusions are made about the intervals of the epistemic parameters e_i when comparing to the KS-test applied in this thesis.

CHAPTER 10

Geotechnical applications

Geotechnical engineering is a prime example of uncertainty quantification due to the generally speaking large uncertainties. Hence in this chapter, a number of common geotechnical engineering problems are setup and solved with the approaches developed in this thesis.

10.1 Overview

10.1.1 Context

Geotechnical engineering deals with construction projects in the subsurface and near-surface area of the earth. Hence, a geotechnical engineer deals with the interface between rock, soil, and anthropological constructions. Applications include tunnels, trenches, retaining walls, and foundations. A crucial task is to predict the behaviour of the rock and soil masses which influence the target construction site. However, the uncertainty in these materials is large, which makes uncertainty quantification necessary.

Soil masses consists of natural, often not-engineered materials. The material *soil* may range from silts to gravel and boulders. Moreover, soil is a composite material which can consist of rock and organic, as well as other intrusive components (*e.g.* in a sanitary landfill). Last but not least, soil is a multi-phase material, where solid (*e.g.* gravel), fluid (*e.g.* water and liquid gas), and gaseous (*e.g.* air) molecules are present at the same time. Long lasting processes, such as sedimentation and erosion, create soil masses over thousands of years. Additionally, short lasting processes such as a sudden raise of the groundwater table or drilling of a tunnel, can influence the soil behaviour in a limited geographical domain. Hence, the soil's mechanical/physical behaviour are generally-speaking inhomogeneous, anisotropic, non-linear, and time-dependent. The same natural processes are responsible for mechanical/physical properties that vary spatially, *i.e.* they are not constant on a large-scale physical domain.

Apart from the mechanical and physical behaviour of soil, the determination of the material model is another source of uncertainty. Soil masses consist of generally speaking opaque materials and large dimensions. Hence, accurate measurements of mechanical and physical properties are costly because large devices are required for data accumulation on site. Moreover, due to the spatial variability of soil, the measured value depends greatly on the location where it is taken.

Moving the measurement device a few meters can change data considerably. In order to describe a soil mass fully, *i.e.* including the spatial variability, a dense set of measurements are required, which in turn is not economically feasible. Hence, epistemic uncertainty is a large source of uncertainty in geotechnical engineering along with the inherent aleatory uncertainty.

As a result of the natural variability and limited observability properties, geotechnical engineering is a prime example for large uncertainties. Hence, there is an intrinsic need for uncertainty quantification in geotechnical applications. Additionally, the mix of natural variability and lack of knowledge fits well into the framework of imprecise probabilities discussed in this thesis.

10.1.2 Computational model

In geotechnical engineering, the computational models represent rock and soil masses as well as engineered structures (such as a retaining wall). The quantity of interest depends on the specific problem setting. However, there are three scenarios which compose the three most common situations in practice: (i) slope stability, (ii) foundation settlement and foundation bearing capacity, and (iii) stability of a retaining wall (see *e.g.* Lang et al. (2007)). Despite their popularity, analytical solutions are only available for a few special cases. The analytical solutions require rigorous assumptions on the problem setting, such as homogeneous soil masses, negligence of spatial variability, idealized soil behaviour, and a specific geometry. For such special case scenarios, a number of standard tools is available. As an example, a slope stability analysis of a homogeneous soil mass with no layering and a simple geometry can be solved by the *ordinary method of slices* (Fellenius 1936), the *Bishop's modified method* (Bishop 1955), the *force equilibrium method* (Lowe & Karafiath 1960), the *Janbu's generalized procedure for slices* (Janbu 1968), the *Morgenstein & Price's method* (Morgenstein & V. Price 1965), or the *Spencer's method* (Spencer 1967), amongst others. A survey of more methods can be found in Duncan (1996) and Duncan & Wright (2005).

For more general cases, finite element models (FEM) have become the standard tool to analyse the behaviour of geotechnical systems. FEMs require none of the above mentioned assumptions and are hence suitable for analysing geotechnical engineering problems (Sudret 2015). FE models can represent geotechnical problems with *e.g.* arbitrarily-shaped soil masses, loads, time-variant pore water pressures, and even spatial variability. The drawback is, however, the computational complexity of the analysis. The more complex the FE model, the larger are the computational resources required to solve the very model. As an example, the solution to a complex FEM can take minutes to hours using state-of-the-art software packages on nowadays computers. Hence, geotechnical FE models are an ideal example for the application of meta-models to speed up the uncertainty quantification analysis.

In this context, a large number of paper have been published, a few of which are mentioned here. Griffiths, Huang, & Fenton (2011) discuss the reliability of an infinite slope. The stability of slopes is analysed by using FE models in Griffiths & Fenton (2004). D.-Q. Li et al. (2016) use meta-models and subset simulation for efficient reliability-based design analyses in geotechnical engineering.

10.1.3 Modelling spatial variability

The spatial variability of mechanical properties in soil masses can be considered in FE models by assigning different property values to each node in the FE mesh. For this purpose, the spatial variability is commonly modelled by random field theory. The random field theory allows to model the mechanical properties of each geographical locations in a physical domain by a random variable. In addition, the variables at different locations are correlated by a correlation function (see Appendix F for further details).

In order to apply random fields in FE models, realizations of the random field are generated and its values are assigned to the corresponding nodes of the FE mesh. The simulation of such realizations is not straightforward. However, a number of methods have been proposed in the literature. A survey and comparison of popular methods can be found in *e.g.* Sudret & Der Kiureghian (2000). Among theses methods, the *expansion optimal linear estimation* (EOLE) method approximates the random field by a series expansion and a finite set of probabilistic random variables. For further details, the reader is referred to Appendix F, where the theory of EOLE is discussed and the method is illustrated on a simple example. Note that in the subsequent examples, the EOLE method is used to approximate all random fields involved.

In geotechnical engineering, a number of publications are available discussing the application of random fields. Al-Bittar & Soubra (2013, 2014) analyse strip footings on spatially random soils using PCE. Vorechovský (2008) simulates correlated random fields for FEM analyses. Papaioannou & Der Kiureghian (2010) conduct reliability-based design analyses for slopes where the soil is modelled by random fields. Griffiths, Huang, & Fenton (2009) discuss the influence of spatial variability on slope reliability. Random fields in combination with observational data and PCE meta-models are analysed in Schöbi & Sudret (2015a).

10.1.4 Examples

In the following sections, four application examples are discussed. The examples illustrate problem settings that appear frequently in the geotechnical engineering practice. A summary of the examples is given here:

(i) Infinite slope reliability (Section 10.2): the infinite slope reliability analysis is one of the few problems with analytical solution. In fact, the limit-state function is a simple equation. The input vector consists of six p-box-modelled input variables. Hence, the previously proposed two-level meta-modelling approaches can be applied straightforwardly. Note that there is no spatial variability in this example.

(ii) Foundation settlement with spatial variability (Section 10.3): the second application example aims at estimating the settlement of a foundation structure. In order to describe a realistic setting, the soil mass is composed of two soil layers which are connected by a random interface. A one-dimensional random field is used to describe the location of this interface. Due to the random interface geometry, the foundation settlement problem is analysed with FEM. Furthermore, meta-models are used to allow for uncertainty quantification analyses.

(iii) Foundation settlement with two random fields (Section 10.4): the third application example builds upon the two-layer soil model of the second application example. In order to increase the complexity of the analysis, though, the Young's modulus of the upper soil layer is also modelled by a random field, beside the layer interface.

(iv) Slope reliability (Section 10.5): another practical situation is the estimation of the stability of a slope. In this example, random fields are used to characterize the shear strength parameters of the soil mass (*i.e.* friction angle and cohesion). Furthermore, drained and undrained analyses are conducted to analyse the different failure mechanisms.

10.2 Infinite slope reliability

10.2.1 Problem setup

In geotechnical engineering, the stability of a infinitely long slope is a benchmark problem for structural reliability analysis, as seen in *e.g.* Lang et al. (2007). Phoon (2008) formulated the following limit-state function with six input random variables:

$$g_{slope}(\boldsymbol{x}) = \frac{[\gamma\,(H-h) + (\gamma_{sat}-\gamma_w)\,h]\cos\psi\,\tan\varphi}{[\gamma\,(H-h)+\gamma_{sat}h]\sin\psi} - 1, \tag{10.1}$$

where H is the depth of the soil above bedrock, h is the height of the groundwater table above bedrock, γ and γ_{sat} are the moist and saturated unit weight of the soil, respectively, $\gamma_w = 9.81$ kN/m^3 is the unit weight of water, φ is the effective stress friction angle, and ψ is the slope inclination. Figure 10.1 illustrates graphically the problem setting and the variables.

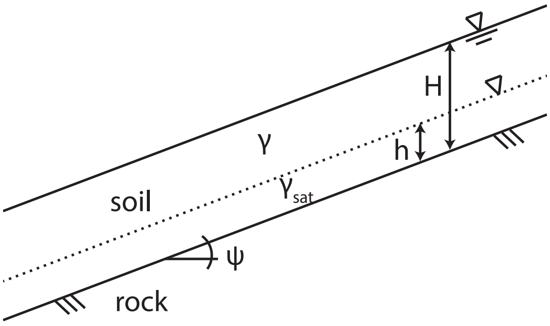


Figure 10.1: Infinite slope – geometry and definition of variables.

In order to model the input with a set of independent variables, the following auxiliary variables are introduced. The moist and saturated soil unit weights are related to each other by means of the specific gravity of soil solids G_s and the void ratio e:

$$\gamma = \gamma_w \frac{G_s + ke}{1+e}, \qquad \gamma_{sat} = \gamma_w \frac{G_s + e}{1+e}, \tag{10.2}$$

where $k = 0.2$ is the degree of saturation of moist soil. Further, h is modelled using an auxiliary variable U_h such that $h = U_h \cdot H$, because h is limited to the height of the soil. Finally, the input vector consists of six independent variables $\boldsymbol{x} = (H, U_h, \varphi, \psi, G_s, e)$.

Due to modelling assumption $h = U_h \cdot H$, the limit-state function can be reformulated:

$$\begin{aligned} g_{slope}(\boldsymbol{x}) &= \frac{[\gamma\,(H - U_h\cdot H) + (\gamma_{sat}-\gamma_w)\,U_h\cdot H]\tan\varphi}{[\gamma\,(H-U_h\cdot H)+\gamma_{sat}U_h\cdot H]\tan\psi} - 1 && (10.3)\\ &= \frac{[\gamma\,(1-U_h) + (\gamma_{sat}-\gamma_w)\,U_h]\tan\varphi}{[\gamma\,(1-U_h)+\gamma_{sat}U_h]\tan\psi} - 1, && (10.4)\end{aligned}$$

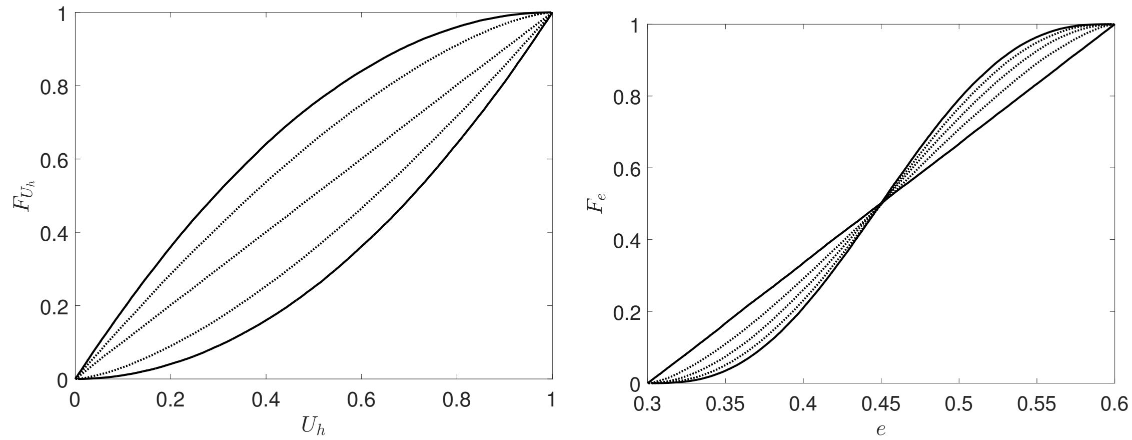
which is independent of H. Hence, the effective dimensionality of this problem is $M = 5$. H can be interpreted as a dummy variable. In the following computations, the variable H is considered although the effect on the results should be negligible.

In this example, the influence of the shape of the CDF onto the failure probability is analysed. Hence, the input parameters are modelled by free and parametric p-boxes. For the parametric p-boxes, each variable X_i is characterized by one interval-valued distribution parameter θ_i. The parametric p-boxes are summarized in Table 10.1. Note that the PDF of the beta distribution is parametrized by $f_{\mathcal{B}}(x) \propto (x-a)^{r-1}(b-x)^{s-1}$, where x is the variable of interest with $x \in$

$[a, b]$, r and s are the shape parameters. In order to visualize the effects of the parametric p-boxes, Figure 10.2 shows the boundary curves of the parametric p-boxes of U_h and e and some realizations of PDFs. Based on the definition of these two variables, U_h allows for skewed distribution functions, whereas e allows for uniform to bell-shaped distribution functions around the mean value.

Table 10.1: Infinite slope – definition of the parametric p-boxes

Variable	Distribution	Statistics	Interval parameter
φ	Lognormal	$\mu_\varphi = 35° \cdot \theta_\varphi$, $\sigma_\varphi = 2.8°$	$\theta_\varphi \in [0.95, 1.05]$
ψ	Lognormal	$\mu_\psi = 20° \cdot \theta_\psi$, $\sigma_\psi = 1°$	$\theta_\psi \in [0.95, 1.05]$
G_s	Lognormal	$\mu_{G_s} = 2.6 \cdot \theta_{G_s}$, $\sigma_{G_s} = 0.065$	$\theta_{G_s} \in [0.98, 1.02]$
H	Lognormal	$\mu_H = 4\text{ m} \cdot \theta_H$, $\sigma_H = 1.2\text{ m}$	$\theta_H \in [0.9, 1.1]$
U_h	Beta	$U_h \in [0, 1]$, $r = 1 + \max[0, \theta_{U_h}]$, $s = 1 + \max[0, -\theta_{U_h}]$	$\theta_{U_h} \in [-1, 1]$
e	Beta	$e \in [0.3, 0.6]$, $s = r = \theta_e$	$\theta_e \in [1, 3]$



(a) U_h

(b) e

Figure 10.2: Infinite slope – parametric p-boxes of the input.

A second scenario consists of all variables being modelled as free p-boxes. Then, the boundary curves are obtained by $\underline{F}_{X_i}(x_i) = \min_{\theta_i} F_{X_i}(x_i|\theta_i)$ and $\overline{F}_{X_i}(x_i) = \max_{\theta_i} F_{X_i}(x_i|\theta_i)$.

10.2.2 Analysis

The analysis settings are kept the same as in the application examples in Section 7.6. A two-level meta-modelling approach with adaptive Kriging models is applied for both cases of parametric and free p-boxes. However, the initial experimental designs for the first and second-level meta-models are set to $N_0^{(1)} = 12$ and $N_0^{(2)} = 12$, respectively. For parametric p-boxes, the threshold value for the stopping criterion in EGO is set to $\epsilon_{EI} = 10^{-5}$. For free p-boxes, the auxiliary distribution for the lognormal variables is defined by a lorgnormal variable the mean of which is the central value of the interval given in Table 10.1. For the beta-distributed variables U_h and e, the auxiliary distribution is defined as a beta distribution with $\{r = 1, s = 1\}$ (uniform distribution) and $\{r = 2, s = 2\}$, respectively.

10.2.3 Results

Table 10.2 summarizes the results of imprecise structural reliability analysis (ISRA). The proposed two-level methods approximate the boundary values of the imprecise failure probability efficiently when comparing to the reference values which are obtained by plain Monte Carlo simulation with $n_{\mathrm{MCS}} = 10^7$ samples.

As expected, the failure probability interval obtained by free p-boxes encapsulates the ones by parametric p-boxes. The reason for the slightly wider interval for free p-boxes is the shape of the input p-box in variable e. The optimal values of the distribution parameter r_e^* indicate that not the extreme values, *i.e.* $\{1, 3\}$, but rather an intermediate value leads to the extreme failure probabilities for the case of parametric p-boxes.

Table 10.2: Infinite slope – results of the ISRA (reference failure probability obtained with Monte Carlo simulation and $n_{\mathrm{MCS}} = 10^7$ samples).

	Parametric p-box		Free p-box	
	$\underline{P}_f$	$\overline{P}_f$	$\underline{P}_f$	$\overline{P}_f$
$P_{f,ref}$	$1.03 \cdot 10^{-3}$	$3.06 \cdot 10^{-1}$	$1.00 \cdot 10^{-3}$	$3.11 \cdot 10^{-1}$
$\widehat{P}_f$	$1.06 \cdot 10^{-3}$	$3.05 \cdot 10^{-1}$	$1.02 \cdot 10^{-3}$	$3.11 \cdot 10^{-1}$
$N^{(1)}$	$12 + 109 = 121$	$12 + 88 = 100$	$12 + 63 = 75$	
$N^{(2)}$	$12 + 72 = 84$	$12 + 11 = 23$	$12 + 116 = 128$	$12 + 80 = 92$
μ_φ^*	$1.05 \cdot 35°$	$0.95 \cdot 35°$	-	-
μ_ψ^*	$0.95 \cdot 20°$	$1.05 \cdot 20°$	-	-
$\mu_{G_s}^*$	$1.02 \cdot 2.6$	$0.98 \cdot 2.6$	-	-
μ_H^*	$1.1 \cdot 4$ m	$1.02 \cdot 4$ m	-	-
$\theta_{U_h}^*$	-1.00	1.00	-	-
r_e^*	2.68	3.00	-	-

Another interesting variable is H, which results in the optimal mean values $\mu_H^* = \{1.1, 1.02\} \cdot 4$ m. Due to modelling assumption $h = U_h \cdot H$, the limit-state function is independent of H as discussed in Eq. (10.4). It follows that the failure probability is independent of H, which explains the arbitrary values in μ_H^* in Table 10.2.

10.3 Foundation settlement with spatial variability

10.3.1 Problem statement

Consider the two-dimensional strip foundation sketched sketched in Figure 10.3 (Schöbi & Sudret 2015a). The soil mass is composed of two layers separated by an irregular horizontal interface $I(x)$ (where x is the horizontal coordinate) that can be modelled by a random field (RF) (see also Appendix F). It is assumed that the soil mass is weightless and lays on a rigid bedrock at a depth of 5 m below the soil surface.

The interface between the two soil layers is modelled by a one-dimensional Gaussian random field $I(x)$ with mean value $\mu_I = -1$ m measured from the soil surface, a standard deviation of $\sigma_I = 0.3$ m, and an exponential autocorrelation function with correlation length $\rho_I = 5$ m. Note that there is a probability of $\Phi\left(-\frac{1}{0.3}\right) = 4.3 \cdot 10^{-4}$ that the interface is *above* the soil surface. In this case, the lower layer would reach the soil surface at that location. Further, at

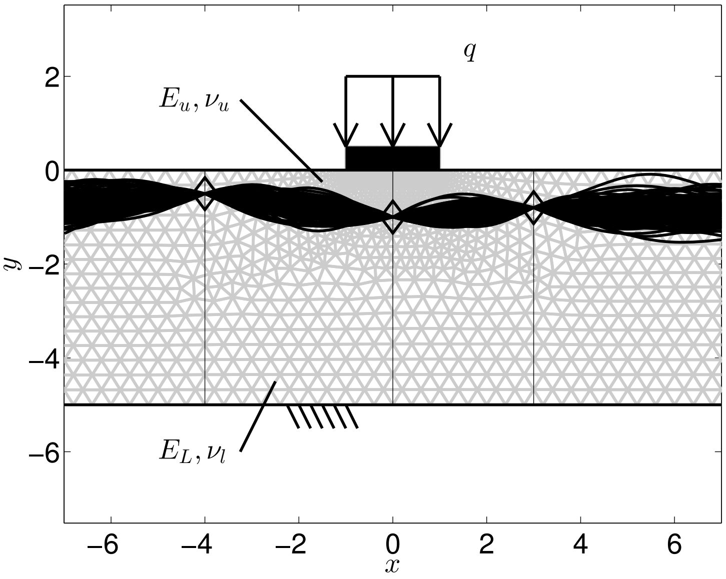


Figure 10.3: Foundation settlement – geometry and notations.

three locations on the horizontal axis indicated in Figure 10.3, the interface depth is measured: $I\,(x = -4 \text{ m}) = -0.5$ m, $I\,(0 \text{ m}) = 1$ m, and $I\,(3 \text{ m}) = -0.8$ m. Several realizations of the conditional random field are drawn in Figure 10.3 to showcase its variability. The measurements clearly reduce the spatial variability of the interface in the neighbourhood of the observations.

The mechanical properties of the soil mass are modelled by a linear elastic behaviour. In particular, Young's modulus E and Poisson's ratio ν describe the linear elastic constitutive model. In both soil layers, the mechanical properties are modelled by random variables without spatial variability. On top of the soil mass, a 2 m wide and 0.5 m thick foundation is located, which is modelled by a deterministic Young's modulus of $E_f = 25$ GPa and Poisson's ratio of $\nu_f = 0.2$. The foundation is subjected to a vertical pressure q.

The probabilistic model of all stochastic variables is summarized in Table 10.3. Each variable is defined by its distribution function, mean value and coefficient of variation (CoV). Note that the variables $\{\nu_u, \nu_l\}$ are modelled as constants due to their small influence on the results of the FEM analysis.

Table 10.3: Foundation settlement – input distributions (subscript u, and l stand for upper and lower layer of the soil model, respectively; $I(x)$ summarizes the random field interface).

Variable	Distribution	Mean	CoV
E_u	Lognormal	5 MPa	15 %
E_l	Lognormal	20 MPa	15 %
ν_u, ν_l	Deterministic	0.3	-
q	Gumbel	100 kPa	10 %
$I(x)$	Gaussian RF	−1 m	30 %

The quantity of interest is the settlement v of the midpoint of the foundation. The uncertainty in v, hence, depends on the stochastic variables $\{E_u, E_l, q\}$ as well as the spatial variability in the interface $I(x)$.

10.3.2 Analysis

The random field is discretized using EOLE (see Appendix F.3) and $M_{\mathrm{RF}} = 30$ Gaussian variables. The total dimensionality of the computational model is then $M = 30 + 3 = 33$.

The foundation settlement v is estimated by a finite element model implemented in the COMSOL Multiphysics software (COMSOL AB 2013). The resulting soil model has a total width of 40 m to avoid boundary conditions influencing the foundation settlement v. Its mesh is (partially) displayed in Figure 10.3 and consists of 17042 degrees-of-freedom. Note that the mesh is fine close to the foundation strip and coarse far away from the foundation strip. Depending on the realization of the random field, each node in the soil mass is assigned a value of $\{E_u^*, \nu_u\}$ or $\{E_l^*, \nu_l\}$ in the upper and lower layer, respectively.

In order to estimate the uncertainty in the foundation settlement v, a PCE model is calibrated. Its experimental design consists of $N = 300$ LHS samples. The sparse PCE model is calibrated using the Matlab-based framework UQLab (Marelli & Sudret 2015). Further, Sobol' indices are computed as a post-processing algorithm of the PCE model by using the same software framework (Marelli, Lamas, & Sudret 2015).

10.3.3 Results

The accuracy of the PCE model is estimated through a validation set of $n_{\mathrm{MCS}} = 1000$ MC samples. The relative generalization error between the prediction and the exact values of the validation set is $\widehat{err}_{gen} = 4.4~\%$. This indicates a sufficiently accurate approximation of the foundation settlement v to estimate the response PDF as well as sensitivity measures.

Due to the inexpensive-to-evaluate formulation of the sparse PCE model, the PDF $f_v(v)$ of the foundation settlement can be estimated. The solid line in Figure 10.4 illustrates the estimate of $f_v(v)$ obtained by a kernel smoothing of a MC sample ($n_{\mathrm{MCS}} = 10^6$) of the input vector and the PCE model.

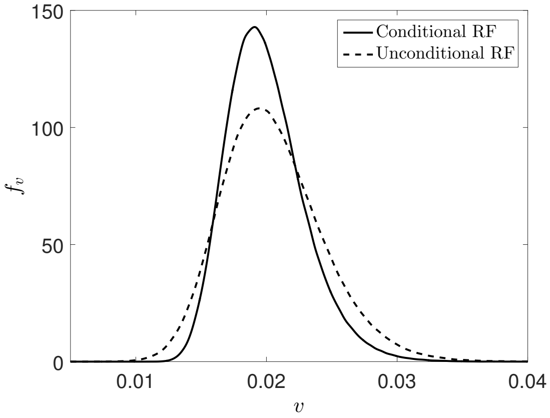
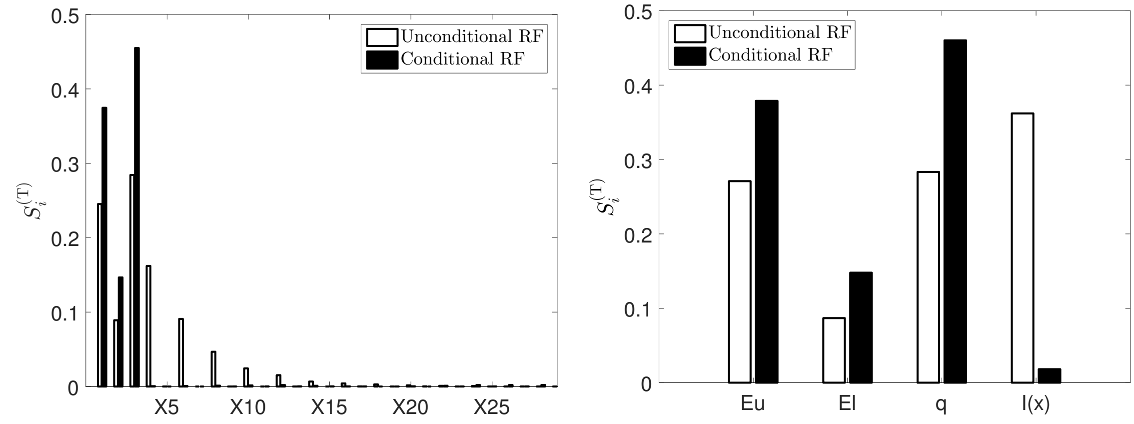


Figure 10.4: Foundation settlement – PDF of v.

The dashed line in Figure 10.4 represents the PDF of the settlement in case an unconditional random field (with the same parameters) is used, *i.e.* ignoring the borehole data. The difference between the two distributions shows the influence of the borehole data, namely a significant reduction of the uncertainty in the settlement estimation. The PDF of the conditional random field case is more peaked and narrower. The mean value and standard deviation of the foundation settlement are $\{\mu_{v,c} = 20.0 \text{ mm}, \sigma_{v,c} = 3.0 \text{ mm}\}$ and $\{\mu_v = 20.5 \text{ mm}, \sigma_v = 3.9 \text{ mm}\}$ for the conditional and unconditional random field, respectively. These statistics indicate that the observations reduce indeed the uncertainty in the settlement prediction.

In order to analyse the influence of the RF on the foundation settlement, Sobol' indices are computed from the coefficient of the PCE model. The resulting total order Sobol' indices, denoted by $S_i^{(\mathrm{T})}$, are illustrated in Figure 10.5 for the cases of conditional (black bars) and unconditional (white bars) random fields. In Figure 10.5(a), the variables $\{X_1, X_2, X_3\}$ correspond to $\{E_u, E_l, q\}$ whereas the variables $X_4, \ldots, X_{33}$ correspond to the $M_I = 30$ Gaussian variables modelling the random field $I(x)$. Note that in Figure 10.5(b) the variable $I(x)$ stands for the total order Sobol' index involving all 30 random variables describing the random field of the interface.



(a) Individual variables

(b) Grouped variables

Figure 10.5: Foundation settlement – total order Sobol' indices based on the PCE models.

In the case of an unconditional random field, the influence of the RF onto the variance of the settlement is largest among the input variables, whereas the influence is the smallest for the conditional random field. This indicates that the few observations (*i.e.* only three points) significantly reduce the uncertainty regarding the location of the interface. The remaining input variables have the same relative importance to each other in terms of Sobol' indices for both cases.

10.4 Foundation settlement with two random fields

10.4.1 Problem statement

Consider again the foundation settlement problem presented in the previous section. In order to increase the complexity of the computational model and hence the dimensionality of the calibrated meta-models, a second random field is introduced. Besides the interface of the two layers, the variability of the Young's modulus of the upper layer is modelled by a two-dimensional random field $E_u(x, y)$ where (x, y) are the coordinates in the soil mass.

Figure 10.6 illustrates the model setup. The coloured area represents a single realization of the two-dimensional RF $E_u(x, y)$. Additionally, the two layers of soil are connected by a realization of the RF of $I(x)$. The remaining random variables are then the load q and the Young's modulus E_l. The Poisson's ratios are kept constant as well as the Young's modulus of the foundation strip (E_f).

Table 10.4 summarizes the modelling of the random variables and constants in the input vector. Note that the value for the load and Young's modulus are different than in Section 10.3.

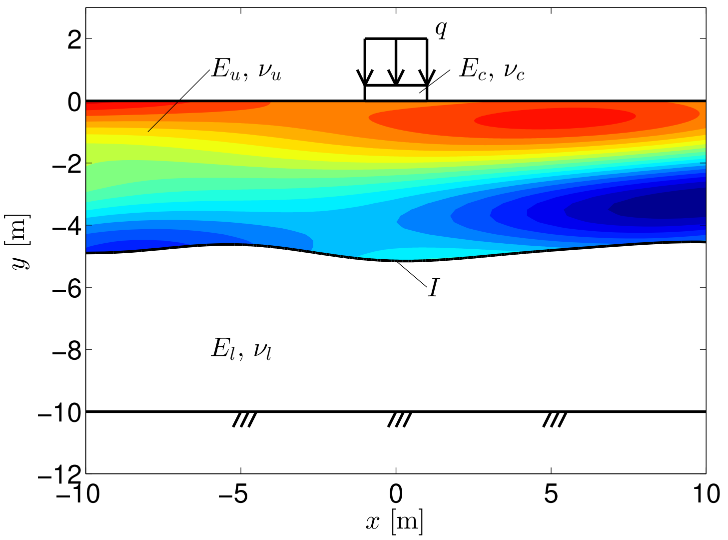


Figure 10.6: Two RF settlement – sketch of the geometry and the input variables.

The description of the random fields in summarized in Table 10.5. Note that the random field $E_u(x, y)$ is modelled by a lognormal random field. The logarithmic value of a lognormal random variable is a Gaussian random variable. Hence, realizations of the logarithmic random field are obtained by (i) generating realizations of a corresponding Gaussian random field and (ii) subsequently transforming the values by exponentiation.

Table 10.4: Two RF settlement – input random variables and constant parameters.

X_i	Name	Distribution	Mean	CoV
q	Load	Gumbel	200 kPa	15 %
E_l	E-modulus lower layer	Lognormal	50 MPa	15 %
E_f	E-modulus foundation	Constant	30 GPa	-
ν_f	Poisson's ratio foundation	Constant	0.20	-
ν_u	Poisson's ratio upper layer	Constant	0.45	-
ν_l	Poisson's ratio lower layer	Constant	0.30	-

The quantity of interest is again the settlement of the foundation strip at its mid-point, denoted by u in this example.

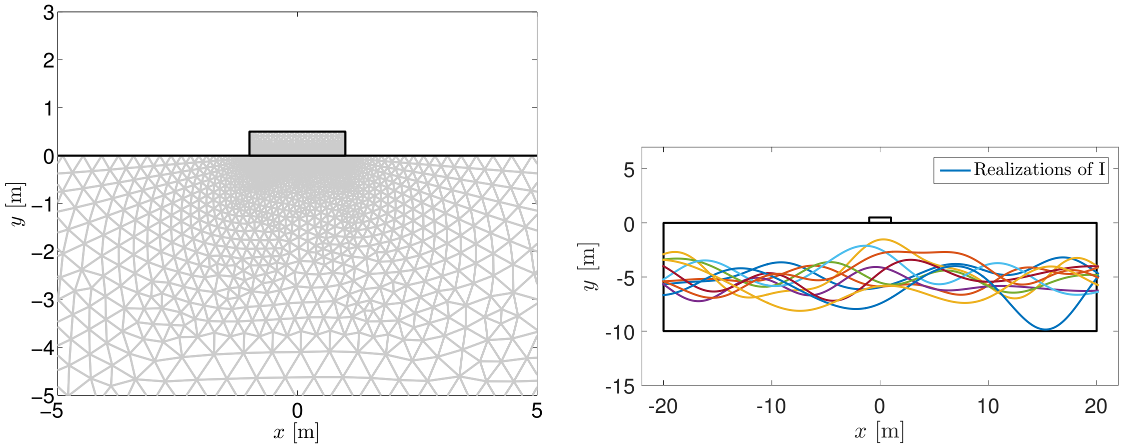
10.4.2 Analysis

The FEM is developed in COMSOL Multiphysics (COMSOL AB 2013). The mesh includes 8514 elements where the largest element has a dimension of 0.4 m which corresponds to a fifth of the vertical correlation length ρ_{y,E_u}. Figure 10.7(a) visualizes the mesh around the strip foundation. The mesh size is fine in the vicinity of the foundation strip, whereas it is coarse at the lateral boundaries. The solving time for a single evaluation is roughly six seconds on a quad-core processor.

The involved random field are discretized with the EOLE algorithm presented in Appendix F.3. In order to reproduce 99 % of the variance, $n_\zeta = 11$ and $n_\zeta = 45$ variables are used for modelling $I(x)$ and $E_u(x, y)$, respectively. Hence, the dimensionality of the computational model is $M = M_q + M_{E_l} + M_I + M_{E_u} = 1 + 1 + 11 + 45 = 58$.

Table 10.5: Two RF settlement – definition of the two random fields.

Name	Description	Value
$I(x)$	Distribution	Gaussian
	Mean	-5 m
	CoV	30 %
	Correlation family	squared exponential
	Correlation length	5 m
	Number of variables	11 ($\sim$ 99 % variance)
$E_u(x,y)$	Distribution	Lognormal
	Mean	10 MPa
	CoV	15 %
	Correlation family	squared exponential
	Corr. length horizontal	10 m
	Corr. length vertical	2 m
	Number of variables	45 ($\sim$ 99 % variance)



(a) FEM mesh (b) Realizations of $I(x)$

Figure 10.7: Two RF settlement – FEM model.

Figure 10.7(b) shows a number of realizations of the RF of the interface $I(x)$. The realizations are smooth curves due to the large correlation length and the squared exponential correlation family used in the application example. Due to the large coefficient of variation of 30 %, the depth of the interface from the surface varies a lot. In particular, one realization is shown which touches the rock surface (see the dark blue line).

The uncertainty quantification analysis in this application example is split into three consecutive tasks:

(i) Sensitivity analysis: A PCE model is calibrated based on an experimental design of $N = 1000$ LHS samples and the corresponding response values of the FEM. Based on the PCE model, the Sobol' indices for all variables are estimated.

(ii) Structural reliability analysis: Based on the PCE model, the failure probability $P_f = \mathbb{P}\left(u \geq u_{\text{adm}}\right)$ is estimated for different threshold values u_{adm}.

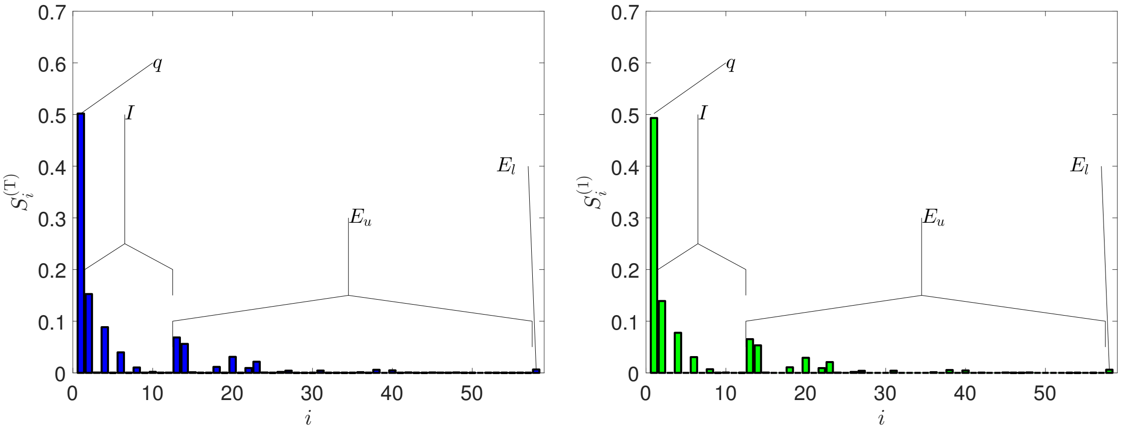
(iii) Improved structural reliability analysis: Based on the Sobol' indices, a reduced input vector is defined where each remaining variable has a non-negligible total order Sobol' index, *i.e.* $S_i^{(\mathrm{T})} \geq 0.01$. Subsequently, an improved structural reliability analysis is conduced using AK-MCS and the resulting failure probability estimate is compared to the PCE-based estimate.

10.4.3 Results

Sensitivity analysis

The PCE model is calibrated in UQLab (Marelli & Sudret 2015). The optimal maximum degree of polynomials is found as $p^* = 4$. The accuracy of the PCE model is estimated by a validation set of $n_{\mathrm{MCS}} = 10^4$ MC samples. The resulting relative generalization error is $\widehat{err}_{gen} = 1.85$ %, indicating an accurate meta-model despite the large dimensionality of the input vector.

The first and total order Sobol' indices are shown in Figure 10.8 for all $M = 58$ random variables. Note that the load q has the largest impact of the uncertainty in the settlement u. Moreover, some variables of the RF discretizations of $I(x)$ and $E_u(x, y)$ have a non-negligible influence too. As in the previous example application in Section 10.3, the variables with low index i in each RF have larger Sobol' indices than the ones with high index i.



(a) Total Sobol' indices

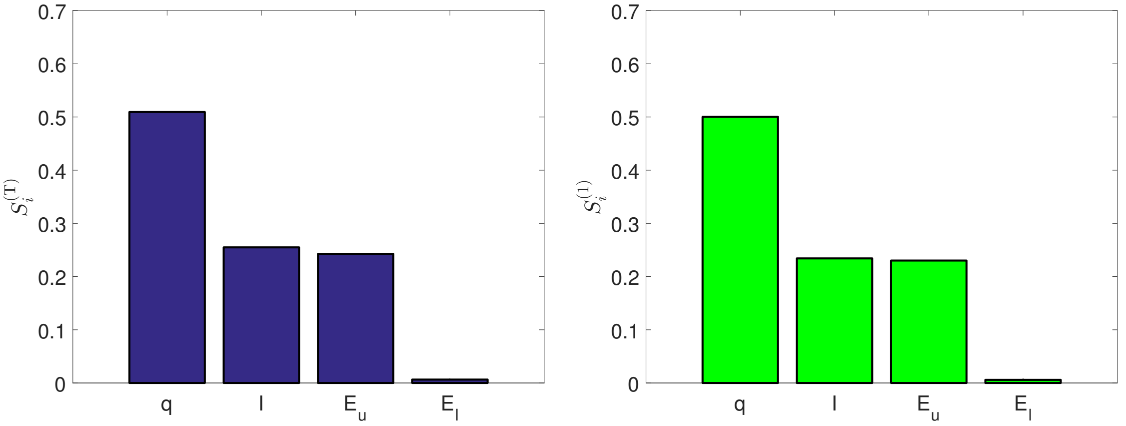
(b) First order Sobol' indices

Figure 10.8: Two RF settlement – Sobol' indices for individual input variables X_i.

The Sobol' indices in Figure 10.8 show the individual contributions of X_i to the uncertainty in the response variable u. In order to estimate the effect of the variability of a random field, the corresponding variables are grouped. The resulting Sobol' indices are summarized in Figure 10.9. The results show that the uncertainty introduced by the random fields have a considerable effect onto the uncertainty in u. The effects of $I(x)$ and $E_u(x, y)$ are comparable by means of first and total order Sobol' indices. The uncertainty in E_l, however, has a marginal effect on u.

Structural reliability analysis

Using the same PCE model as in the sensitivity analysis, the failure probabilities are estimated by MCS with a sample size of $n_{\mathrm{MCS}} = 10^7$ to ensure accurate estimates in terms of MCS analysis. The estimates of the failure probability and the corresponding reliability indices are summarized in Table 10.6. The failure probability decreases rapidly with increasing threshold value u_{adm}. Note that the accuracy for $u_{\mathrm{adm}} = \{6 \text{ cm}, 7 \text{ cm}\}$ is limited due to the use of PCE models which are generally speaking not always accurate describing the tail of the output distribution.



(a) Total order Sobol' indices

(b) First order Sobol' indices

Figure 10.9: Two RF settlement – Sobol' indices for groups of variables.

Table 10.6: Two RF settlement – failure probabilities for different threshold values (obtained via MCS on the PCE meta-model and $n_{\mathrm{MCS}} = 10^7$).

u_{adm}	$\widehat{P}_f$	$\widehat{\beta}_{\mathrm{HL}}$
4 cm	$1.97 \cdot 10^{-1}$	0.85
5 cm	$2.71 \cdot 10^{-2}$	1.93
6 cm	$2.85 \cdot 10^{-3}$	2.76
7 cm	$2.80 \cdot 10^{-4}$	3.45

Improved structural reliability analysis

In order to improve the accuracy of the failure probability estimates, a structural reliability analysis is conducted in this section taking into account the results of the sensitivity analysis and replacing PCE by AK-MCS. Selecting the variables with $S_i^{(T)} \geq 0.01$, a new input vector is formed consisting of only nine variables. Apart from the load q, three out of eleven variables in $I(x)$ and five out of 45 variables in $E_u(x, y)$ remain in the input vector.

The smaller dimensionality allows one to use Kriging meta-models and moreover AK-MCS to estimate the failure probabilities for different values of u_{adm}. The Kriging meta-models are trained with a Gaussian correlation function and an initial experimental design of $N_0 = 20$ Latin-hypercube samples. Further for AK-MCS, probability of misclassification is chosen as the learning function and a candidate set of $n_{\mathrm{MCS}} = 10^6$ Monte Carlo samples is generated.

Table 10.7 summarizes the results of the analysis, namely the estimated failure probability, corresponding reliability index, as well as the total number of runs of the FEM in AK-MCS N_{tot}. The failure probability estimates for $u_{\mathrm{adm}} = \{4\text{ cm}, 5\text{ cm}\}$ are close to the ones provided by PCE in Table 10.6. For $u_{\mathrm{adm}} = \{6\text{ cm}, 7\text{ cm}\}$, the deviation of the two estimates becomes larger. In terms of reliability index, however, the results of both methods are consistent. Note that N_{tot} denotes the number of FEM runs till a steady value of the failure probability is achieved. For all thresholds, the reliability analysis terminated with $N_{tot} \leq 150$, which is a low number considering failure probabilities as small as $\sim 10^{-4}$.

Table 10.7: Two RF settlement – AK-MCS results based on the reduced input vector.

u_{adm}	$\widehat{P}_f$	$\widehat{\beta}_{\text{HL}}$	N_{tot}
4 cm	$1.83 \cdot 10^{-1}$	0.90	$20 + 106 = 126$
5 cm	$2.65 \cdot 10^{-2}$	1.94	$20 + 122 = 142$
6 cm	$3.49 \cdot 10^{-3}$	2.70	$20 + 130 = 150$
7 cm	$4.10 \cdot 10^{-4}$	3.35	$20 + 129 = 149$

10.5 Slope stability with spatial variability

10.5.1 Problem statement

Consider the two-dimensional slope model sketched in Figure 10.10. The height of the embankment is 10 m and the slope inclination is 2 : 1. Below the embankment, there is a 10 m-thick layer of soil, which reaches the bedrock. The bedrock is assumed to be rigid, hence modelling the lower boundary of the slope model.

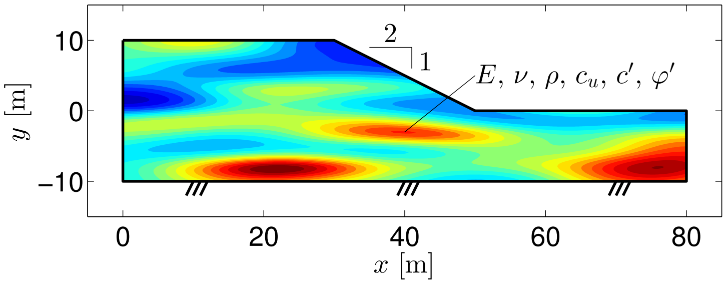


Figure 10.10: Slope stability – geometry and variables.

The analysis is split into two parts: (i) an undrained analysis and a (ii) drained analysis. The details are given here:

(i) Undrained analysis: Consider the case where the soil consists of silt and clay. The constitutive law for the soil mass consists of a linear elastic-ideal plastic behaviour. The soil mass is considered undrained, hence modelled by the Tresca failure criterion. The mechanical properties of the soil are thus described by the Young's modulus E, Poisson's ratio ν, and undrained shear strength c_u, which is modelled by a two-dimensional random field. The probabilistic modelling of the input parameters is summarized in Tables 10.8 and 10.9.

(ii) Drained analysis: Consider the case where the soil mass consists of gravel and sand with a small cohesion. Similar to the undrained analysis, the soil mass is modelled by a linear elastic-ideal plastic behaviour. However, the failure is modelled by the Mohr-Coulomb failure criterion, which is parametrized by the effective friction angle φ' and the effective cohesion c'. These two factors are assumed to be spatially varying and linearly correlated. Hence, they are modelled by two correlated random fields. The statistics of the random fields are summarized in Table 10.10. The linear elastic behaviour is assumed to be identical to the case of undrained soil behaviour, hence see Table 10.8 for the probabilistic descriptions.

Table 10.8: Slope stability – probabilistic input variables.

Name		Distribution	Mean	CoV
Young's modulus	E	Lognormal	50 MPa	5 %
Poisson's ratio	ν	Constant	0.3	-
Soil density	ρ	Constant	2000 kg/m^3	-

Table 10.9: Slope stability – undrained analysis – definition of the random field of c_u.

Description	Value
Distribution	Lognormal
Mean	50 kPa
CoV	20 %
Correlation family	Squared exponential
Correlation length horizontal	15 m
Correlation length vertical	2 m
Number of variables	65 ($\sim$ 95 % variance)

10.5.2 Analysis

The FEM is created using the free software `slope64` (Griffiths & Lane 1999; Griffiths 2015). The corresponding mesh is shown in Figure 10.11, which consists of 900 elements. The mechanical properties of each element is equal to the value of the RF realization taken at the central point of the element. Note that according to Sudret & Der Kiureghian (2000), the maximum element size in the vertical direction shall not be larger than a fourth of the vertical correlation length (*i.e.* 0.5 m), whereas in Figure 10.11 it is 1 m. The difference in terms of factor of safety between the present and a finer mesh is, however, small as it was checked from analyses not reported in this manuscript. Hence, this example is computed with the coarser mesh in order to speed up the computation.

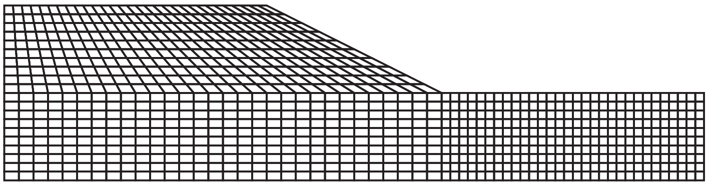


Figure 10.11: Slope stability – mesh of the finite element model.

As in the previous examples, the random fields are discretized by EOLE (see Appendix F.3). In the undrained case, the algorithm is applied straightforwardly. In the drained case, where the two random fields are correlated, the following transformations are used. Consider two independent Gaussian variables $\Xi^{(1)}, \Xi^{(2)} \sim \mathcal{N}(0,1)$. Then, φ' and c' can be modelled as:

$$\varphi' = T_{\varphi}\left(\Xi^{(1)}\right), \tag{10.5}$$

$$c' = T_{c'}\left(\rho_{\varphi c}\, \Xi^{(1)} + \sqrt{1-\rho_{\varphi c}^2}\, \Xi^{(2)}\right), \tag{10.6}$$

where $\rho_{\varphi c}$ is the correlation between φ' and c', and T_{φ} and $T_{c'}$ are isoprobabilistic transforms mapping for the Gaussian space to the original input space, *i.e.* lognormal space in this example.

Table 10.10: Slope stability – drained analysis – definition of the random fields for φ' and c'.

Description	φ'	c'
Distribution	Lognormal	Lognormal
Mean	25°	3 kPa
CoV	10 %	50 %
Correlation family	Squared exponential	Squared exponential
Correlation length horizontal	15 m	15 m
Correlation length vertical	2 m	2 m
Pseudo-correlation φ' and c'	-0.5	
Number of variables	65 ($\sim$ 95 % variance)	65 ($\sim$ 95 % variance)

Considering now two standard Gaussian fields $\Xi^{(1)}(\boldsymbol{x})$ and $\Xi^{(2)}(\boldsymbol{x})$, they can be approximated by the same series expansion due to the shared geometry and correlation structure of φ' and c'. Then, denote the EOLE approximation of $\Xi^{(1)}$ and $\Xi^{(2)}$ by $\widehat{H}_{\Xi^{(1)}}\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)}\right)$ and $\widehat{H}_{\Xi^{(2)}}\left(\boldsymbol{x},\boldsymbol{\zeta}^{(2)}\right)$, respectively. The random fields of φ' and c' can be modelled by:

$$\widehat{\varphi}'\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)}\right) = \exp\left(\lambda_\varphi + \varpi_\varphi \widehat{H}_{\Xi^{(1)}}\left(\boldsymbol{x}\right)\right), \tag{10.7}$$

$$\widehat{c}'\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)},\boldsymbol{\zeta}^{(2)}\right) = \exp\left(\lambda_c + \varpi_c\left[\rho_{\varphi c}\,\widehat{H}_{\Xi^{(1)}}\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)}\right) + \sqrt{1-\rho_{\varphi c}^2}\;\widehat{H}_{\Xi^{(2)}}\left(\boldsymbol{x},\boldsymbol{\zeta}^{(2)}\right)\right]\right), \tag{10.8}$$

where λ and ϖ are the mean and standard deviation of the logarithm of a lognormal random variable. This can be simplified taking into consideration the similar nature of the two random fields. Indeed, the eigenvalue problem associated to EOLE is solved once and for all for φ' and c':

$$\widehat{\varphi}'\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)}\right) = \exp\left[\lambda_\varphi + \varpi_\varphi \sum_{i=1}^{n_\zeta} \frac{\phi_i^{\mathsf{T}} \boldsymbol{\Sigma}_{H(\boldsymbol{x})\widehat{Y}}}{\sqrt{\lambda_i}} \zeta_i^{(1)}\right], \tag{10.9}$$

$$\widehat{c}'\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)},\boldsymbol{\zeta}^{(2)}\right) \quad = \quad \exp\left(\lambda_c + \varpi_c\left[\sum_{i=1}^{n_\zeta} \frac{\phi_i^{\mathsf{T}} \boldsymbol{\Sigma}_{H(\boldsymbol{x})\widehat{Y}}}{\sqrt{\lambda_i}}\left[\rho_{\varphi c}\,\zeta_i^{(1)} + \sqrt{1-\rho_{\varphi c}^2}\;\zeta_i^{(2)}\right]\right]\right), \tag{10.10}$$

where $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$ are the standard Gaussian variables of the RF discretizations $\widehat{H}_{\Xi^{(1)}}\left(\boldsymbol{x},\boldsymbol{\zeta}^{(1)}\right)$ and $\widehat{H}_{\Xi^{(2)}}\left(\boldsymbol{x},\boldsymbol{\zeta}^{(2)}\right)$, respectively.

Given the discretization of the random fields and the FEM, the performance of the system is assessed by the *factor of safety* (FOS). The FOS of each realization of the random fields is obtained by the *shear strength reduction method* (*e.g.* Matsui & San (1992)). Let us define a set of reduced shear strength parameters by:

$$c_{u,f} = \frac{c_u}{\mathrm{FOS}}, \tag{10.11}$$

for the undrained analysis and:

$$c'_f = \frac{c'}{\mathrm{FOS}}, \tag{10.12}$$

$$\varphi'_f = \arctan\left(\frac{\tan\left(\varphi'\right)}{\mathrm{FOS}}\right), \tag{10.13}$$

for the drained analysis. Assume that the gravity load is applied to the FEM. Then, the FOS is defined as the number which brings the slope to the point of failure. In other words, the FOS is the largest value where the FEM converges to a stable solution. The shear strength reduction method is implemented in `slope64`.

Finally, failure occurs when FOS ≤ 1. Hence, the corresponding limit-state function reads $g_{\mathrm{FOS}}(\boldsymbol{X}) = \mathrm{FOS}(\boldsymbol{X}) - 1$ and the corresponding failure probability is $P_f = \mathbb{P}(\mathrm{FOS}(\boldsymbol{X}) \leq 1) = \mathbb{P}(g_{\mathrm{FOS}}(\boldsymbol{X}) \leq 0)$. Note that the number of variables is $M = 1 + 65 = 66$ and $M = 1 + 65 + 65 = 131$ for the undrained and drained analyses, respectively.

Due to the large dimensionality of the limit-state function, PCE models are calibrated to estimate the distribution of the FOS. The experimental design consists of $N = 1000$ LHS samples. The sparse PCE model is calibrated with UQLab (Marelli & Sudret 2015). Then, the failure probability is obtained by MCS conducted on the PCE meta-model.

10.5.3 Results

Failure modes in undrained analysis

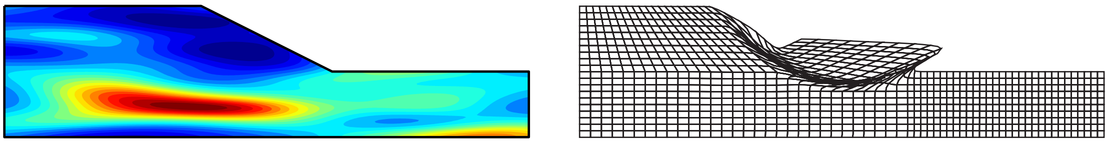
Figure 10.12 shows four realizations of the random field modelling the undrained cohesion c_u and the corresponding failure mode of the FEM model. In the left side figures, the blue areas correspond to low values of the effective cohesion whereas the red areas correspond to large values. The first row of figures shows a shallow failure mechanism, which takes place in the embankment part only. The second row visualizes a failure mechanism with medium depth, *i.e.* not touching the rock surface at the lower boundary of the FEM. The third and fourth row show RF realizations that lead to a deep failure mechanism. The fourth row leads to a combined shallow and deep failure mechanism. Comparing the failure mechanisms with the RF realization, it is apparent that the failure surface follows areas of low values in c_u. This behaviour is expected considering the shear strength reduction method.

Failure modes in drained analysis

Figure 10.13 shows visualizations of three random fields realizations and the corresponding failure mechanism for the case of a drained analysis. In the left side and central figures, dark areas correspond to low values of the mechanical properties whereas light areas correspond to high values. Again, the three rows correspond to three different failure mechanisms, as seen in Figures 10.13(c), 10.13(f), and 10.13(i), respectively. In the drained case, however, the failure surface can be found where the two parameters φ' and c_u are low valued. Note that all failure mechanisms occur close to the inclination of the surface. Hence, the failure mechanisms are more local than the ones observed in the undrained analysis in Figure 10.12.

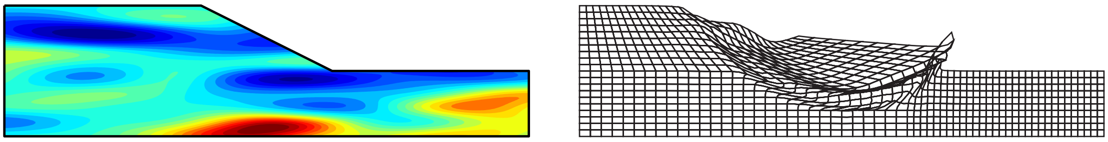
FOS in undrained analysis

The resulting PCE model, calibrated with $N = 1000$ samples, has a relative empirical error of $\widehat{err}_{emp} = 7.05 \cdot 10^{-2}$, which indicates an accurate meta-model despite the large dimensionality of the input domain. Based on the meta-model, Figure 10.14(a) illustrates the CDF of the FOS for the undrained analysis. Furthermore, the empirical CDF of the $N = 1000$ samples is added for comparison. The two lines are close to each other, which confirms the impression of an accurate meta-model. Using MCS with $n_{\mathrm{MCS}} = 10^7$ samples and the PCE meta-model, the failure probability is estimated as $\widehat{P}_f = 2.69 \cdot 10^{-5}$.



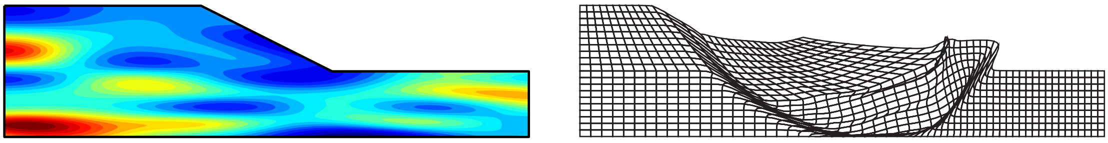
(a) Random field realization #1

(b) Failure mechanism #1



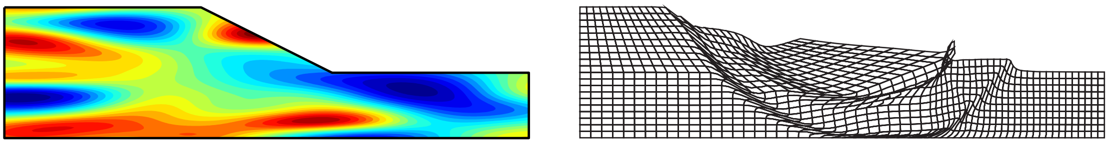
(c) Random field realization #2

(d) Failure mechanism #2



(e) Random field realization #3

(f) Failure mechanism #3



(g) Random field realization #4

(h) Failure mechanism #4

Figure 10.12: Slope stability – undrained analysis – RF realizations and associated failure mechanisms.

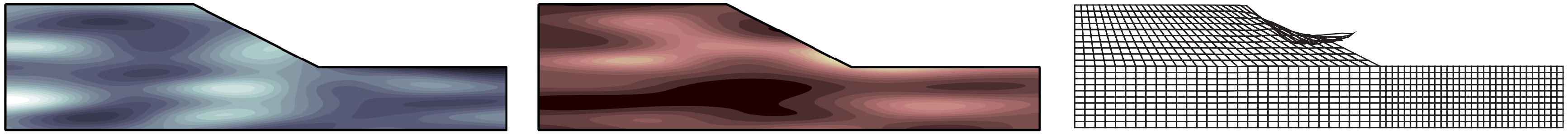
FOS in drained analysis

Based on an experimental design of $N = 1000$ samples, the calibrated PCE meta-model results in a relative empirical error of $\widehat{err}_{emp} = 7.84 \cdot 10^{-2}$. Despite the larger dimensionality M than in the undrained analysis, the meta-model is rather accurate. Figure 10.14(b) shows the CDF of the FOS for the drained analysis. Similar as in the undrained analysis, the meta-model reproduces the response CDF accurately when comparing to the empirical CDF of the $N = 1000$ experimental design samples. Using MCS with $n_{\mathrm{MCS}} = 10^6$ samples and the PCE meta-model, the failure probability is estimated as $\widehat{P}_f = 1.83 \cdot 10^{-3}$. It is important to note that using a PCE meta-model and Monte Carlo simulation may lead to gross results when $P_f < 10^{-3}$. However, the order of magnitude of P_f is deemed correctly evaluated.

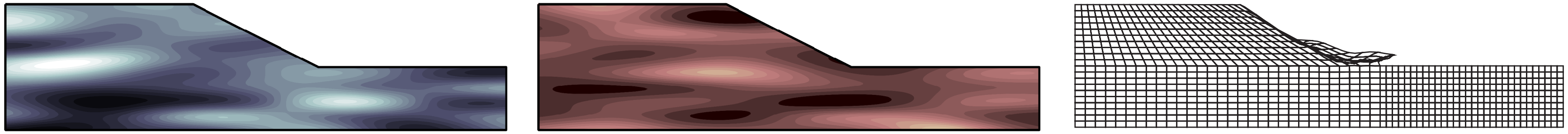
Imprecise failure probability (drained analysis)

The cohesion is a highly uncertain and at the same time a highly influential property of soil in drained analyses. Hence, the effect of an imprecise cohesion is considered here. The mean value of the cohesion is modelled by the interval $\mu_{c'} \in [1, 5]$ kPa and the standard deviation is kept constant, *i.e.* $\sigma_{c'} = 1.5$ kPa. Due to the use of random fields in the implementation of FEM, the uncertainty propagation is repeated for each realization $\mu_{c'} \in [1, 5]$ kPa. In other words, the previously used two-level approaches cannot be used for different realizations of $\mu_{c'}$. However as a first approximation of the influence of the epistemic uncertainty, only the two extreme cases of $\mu_{c'} = 1$ kPa and $\mu_{c'} = 5$ kPa are examined.

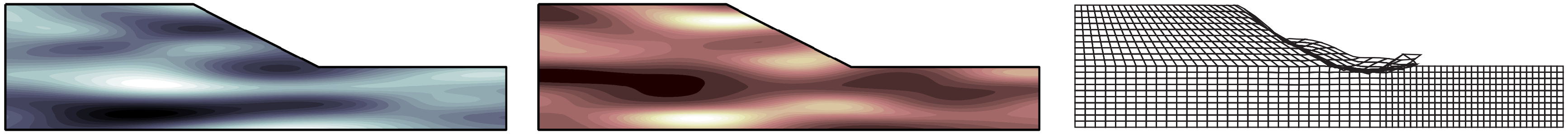
Using the same settings as in the previous paragraphs, two PCE meta-models are calibrated



(a) RF realization #1 of φ' (b) RF realization #1 of c' (c) Failure mechanism #1

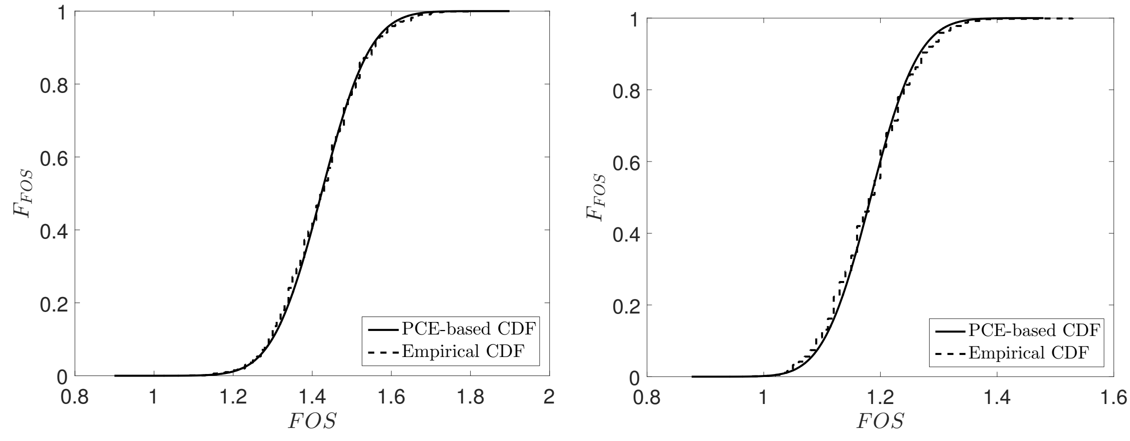


(d) RF realization #2 of φ' (e) RF realization #2 of c' (f) Failure mechanism #2



(g) RF realization #3 of φ' (h) RF realization #3 of c' (i) Failure mechanism #3

Figure 10.13: Slope stability – drained analysis – RF realizations and failure mechanisms.



(a) Undrained analysis (b) Drained analysis

Figure 10.14: Slope stability – CDF of the FOS.

based on an experimental design of size $N = 1000$. The relative empirical error reads $\widehat{err}_{emp} = 1.90 \cdot 10^{-1}$ and $\widehat{err}_{emp} = 4.55 \cdot 10^{-2}$ for $\mu_{c'} = 1$ kPa and $\mu_{c'} = 5$, respectively. The corresponding response CDFs, and hence the response p-box is illustrated in Figure 10.14(a). The width of the response p-box is large, indicating indeed a large influence of the cohesion on the behaviour of the slope. The same effect is observed when computing the imprecise FOSM (Cornell) reliability index defined as $\beta_C = \mu_{\text{FOS}}/\sigma_{\text{FOS}}$. In this case, the estimated imprecise reliability index reads $\beta_C \in [0.25, 4.85]$. The corresponding failure probability $P_f = \Phi(-\beta_{\text{HL}})$ ranges of roughly six orders of magnitude. This illustrates the large impact of the cohesion on the slope reliability in this example.

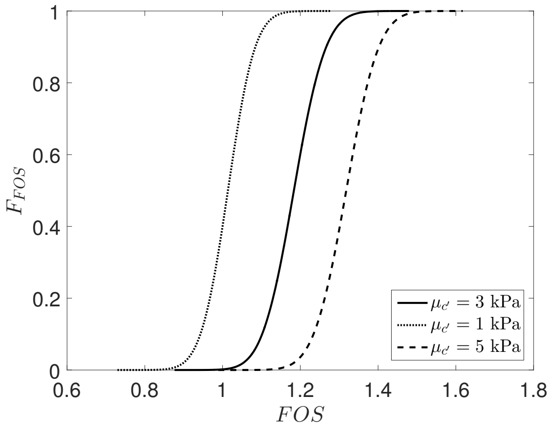


Figure 10.15: Slope stability – drained analysis – FOS as a function of an imprecise cohesion.

10.6 Conclusion

Four application examples are examined in this section of geotechnical engineering. The applications highlight typical geotechnical problem settings. The corresponding computational models have diverse dimensionality, complexity, and numerical solvers. In particular, third-party FEM software is used to model the behaviour of the soil mass. In order to allow for uncertainty quantification, the FEM software is coupled with the MATLAB-based software UQLAB. The modular setup of UQLAB allows for wrappers that consider the FEM software as a black box. Hence, the incorporation of FEM to the uncertainty quantification analyses conducted in MATLAB is simple.

In this thesis, it is assumed that the computational models themselves have no model error. In the geotechnical practice, however, the validation of the models is a crucial task in the analysis. A good practice is then to solve the same problem setting with several FEM softwares and compare the results. As an example, the results of the `slope64` model have been compared to the COMSOL MULTIPHYSICS model in this thesis.

Despite the complexity of the computational models, *i.e.* the FE models, the proposed meta-modelling approaches are applicable in this section. This shows the power of the non-intrusive approach that allows to analyse any computational model defined in Eq. (3.2). Moreover, the proposed two-level meta-modelling approaches allow for an efficient propagation of p-boxes, efficient estimation of failure probabilities, and efficient estimation of sensitivity measures.

CHAPTER 11

Conclusions

11.1 Summary

11.1.1 Imprecise probabilities

Imprecise probabilities describe a mix of aleatory (natural variability) and epistemic uncertainties (lack of knowledge). From the large number of possible concepts, the probability-box (p-box) provides a framework with clear distinction of aleatory and epistemic uncertainty. In this thesis, two cases are considered: free and parametric p-boxes. Parametric p-boxes describe a hierarchical model where the parameters of a probability distributions are modelled by intervals. Hence, the distribution function describes the aleatory uncertainty whereas the interval-valued parameters describe the epistemic uncertainty. Free p-boxes describe the cumulative distribution function of a variable by two bounds. In this case, the probability distribution describes the aleatory uncertainty, whereas the width of the two bounds describes the epistemic uncertainty. The behaviour of both types of p-boxes is intuitively clear when looking at the typical CDF plot (see *e.g.* Figure 2.5 for a free p-box). Due to the clear separation of the two types of uncertainty, p-boxes provide a convenient uncertainty concept for the practical modeller/engineer.

The proper assimilation of pieces of information is important to define a meaningful p-box. Given a dataset and/or expert knowledge, a number of methods have been presented to construct/define a p-box. This emphasizes the usability of p-boxes in practical applications where evidence is typically posed as a mix of sparse datasets and expert opinions.

In case the modeller is asking for an alternative uncertainty modelling concept, p-boxes provide a starting point for extension. On one side, p-boxes are a generalization of probability theory and interval analysis. On the other side, p-boxes can be seen as a special case of fuzzy distributions. Due to these similarities, switching from one to the next uncertainty concept is straightforward.

11.1.2 Meta-modelling

In the context of probability theory, a number of meta-modelling techniques are known and readily used in the literature and software packages (*e.g.* UQLab (Lataniotis, Marelli, & Sudret

2015; Marelli & Sudret 2015)). Among them, there are Polynomial Chaos Expansions (PCE) and Kriging (a.k.a. Gaussian process modelling). PCE assumes that the computational model can be modelled by a series of orthogonal polynomials, whereas Kriging assumes that the computational model is a realization of an underlying Gaussian process. As used in this thesis, PCE is a regression technique that models well the global behaviour of the computational model. On the other side, Kriging is known to model well local variations because of its interpolation properties.

In order to further increase the accuracy of meta-modelling, the author proposed to use a combination of the previous meta-modelling techniques. The new Polynomial-Chaos-Kriging (PC-Kriging) technique models the global behaviour with a set of orthogonal polynomials and the local behaviour with a Gaussian process. In fact, PC-Kriging is a special case of the universal Kriging model where the trend in modelled by orthogonal polynomials. A number of benchmark analytical functions show the increased accuracy when comparing to the traditional meta-modelling techniques.

PC-Kriging is applied to structural reliability analysis and compared to the conventional Kriging technique. The efficiency of computing rare events, such as small failure probabilities and extreme quantiles, is limited with both meta-modelling techniques. Then, a popular strategy is to iteratively enrich the experimental design of the meta-model to increase the accuracy of the estimation of the statistics of interest. Using ideas from machine learning, the adaptive-PC-Kriging-Monte Carlo simulation (APCK-MCS) algorithm is developed as a pendant to the well-established AK-MCS algorithm for Kriging. In order to show the efficiency of the proposed algorithm, APCK-MCS is compared to AK-MCS in a number of benchmark applications. In the shown applications, APCK-MCS allows for an accurate estimation of the statistics of interest for a reasonable amount of runs of the computational model.

UQLab is a state-of-the-art framework for uncertainty quantification developed at the Chair of Risk, Safety and Uncertainty Quantification at ETH Zurich. The conventional PCE and Kriging meta-modelling techniques are included in UQLab's meta-modelling module. As part of this thesis, PC-Kriging and also APCK-MCS have been implemented in UQLab by the author (Schöbi, Marelli, & Sudret 2016). This platform allows for a wide dissemination of these state-of-the-art PC-Kriging algorithms.

11.1.3 Efficient imprecise uncertainty quantification

The state-of-the-art meta-modelling techniques are used to solve uncertainty quantification analyses in the presence of p-boxes. Given the different nature of free and parametric p-boxes, a different set of algorithms is proposed for the two types of p-boxes. The basic idea for solving free p-box-related problems is interval analysis, whereas for parametric p-box-related problems it is nested Monte Carlo simulation.

For free p-boxes, the imprecise uncertainty quantification analysis is converted into two probabilistic uncertainty quantification analyses. Therefore, the meta-modelling techniques described previously can be applied. In particular, two-level meta-modelling algorithms are proposed. The two meta-models consists of PCE models in the case of uncertainty propagation, whereas they consists of adaptive Kriging models in the case of rare event estimations. The use of two meta-models allows for an efficient estimation of the statistics of interest by limiting the number of runs of the expensive-to-evaluate computational model. The efficiency is shown in a number of benchmark application examples. The proposed two-level meta-modelling algorithm can be post-processed for imprecise sensitivity analysis. As an example, pinching input variables modifies the p-box of the QoI. Due to the availability of the meta-model, this operation does not involve further runs of the computational model and is hence inexpensive.

For parametric p-boxes, an augmented input vector and the corresponding augmented PCE model is defined for uncertainty quantification analyses. The augmented input vector allows for a clear separation of the aleatory and epistemic uncertainty in the input to the augmented PCE model. The use of phantom points improves the prediction accuracy without the need for further runs of the expensive-to-evaluate computational model. Further, post-processing the augmented PCE model allows for the definition of imprecise Sobol' indices, an original contribution of the author. Similar as for free p-boxes, the same meta-model can be reused for a pinching analysis. For structural reliability analyses, a combination of AK-MCS and a Kriging-based optimization algorithm (which is similar to AK-MCS) allows for an efficient estimation of failure probabilities. The same benchmark applications are addressed using free and parametric p-boxes to compare the two types of p-boxes. Due to the same meta-modelling tools used in both p-box cases, the number of runs of the computational model are comparable.

11.1.4 Applications

The benchmark applications in the theory chapters are used to validate the proposed algorithms. They consist of mainly analytical functions or inexpensive-to-evaluate finite element models (*e.g.* truss structures). Hence, the reference values are generally available by using crude Monte Carlo simulation.

In the case studies, however, the application examples are more complex and often not solvable by crude Monte Carlo simulation. Reasons are the large computational costs for a single evaluation of the computational model as well as the large dimensionality of the input vector. Hence, a reference solution is not available. As a substitute, a validation set is used to estimate the accuracy of the proposed algorithms and, in particular, of the meta-models involved.

The NASA uncertainty quantification challenge provides a realistic, pure black-box computational model originating in aeronautics. All aspects of the general uncertainty quantification framework are addressed, including the definition of p-boxes, imprecise uncertainty propagation, imprecise structural reliability analysis, and imprecise sensitivity analysis. The main challenge lies in the iterative analysis of the problem, where the previous analysis influences the next analysis. Full advantage is taken from this dependence structure. The uncertainty model of the input is constantly refined by reducing the number of variables in the input vector. This allows for an efficient estimation of the statistics of interest in the final analyses.

Geotechnical engineering is a prime example for uncertainty quantification, including a mix of aleatory and epistemic uncertainties. Hence, four typical application examples of increasing complexity are analysed. Additional to the randomness of a parameter, spatial variability is added as a source of aleatory uncertainty. Using the theory of random fields, discretization algorithms are discussed to simulate realizations of the random fields. Another aspect of the complexity of geotechnical applications is the finite element models (FEM) typically involved. The efficient use of FEM requires a workflow including different pieces of software (*e.g.* UQLab and COMSOL Multiphysics). The proposed two-level meta-modelling approach, however, allow for an efficient estimation of various statistics of interest, such as sensitivity indices, failure probabilities and response p-boxes.

11.2 Outlook

The proposed algorithms offer efficient tools for uncertainty quantification in the context of p-boxes. A number of possible extensions and applications can be envisioned, a subset of which is described in the following.

11.2.1 Fuzzy distributions

In the first part of this thesis, the p-box is determined as the optimal uncertainty concept to deal with the target problem statement. In some cases however, an even more general uncertainty concept might be appropriate. As seen in Chapter 2, p-boxes are a special case for fuzzy distributions. When replacing the intervals by membership functions, p-boxes convert to fuzzy distributions. In particular, free p-boxes generalize to fuzzy distributions whereas parametric p-boxes generalize to fuzzy hierarchical models.

In the context of fuzzy distributions, uncertainty quantification is more cumbersome than in the case of p-boxes due to the more general uncertainty concept. However, due to similarities, the meta-modelling algorithms proposed in this thesis can be modified and extended to fit the needs of fuzzy distributions. Then, an efficient way to handle fuzzy distributions appears affordable.

11.2.2 Imprecise correlation structure

The computational model maps an input vector to a quantity of interest. In this thesis, it is assumed that the input variables are independent in order to be able to apply meta-models without the need of complex isoprobabilistic transforms. In practice, however, it often appears that certain variables are correlated. As an example, the friction angle and cohesion of a soil are normally negatively correlated, as seen in the geotechnical examples in Section 10.5.

In the context of probability theory, a number of correlation measures are popular, including Pearson's correlation coefficient, Spearman's correlation coefficient, Kendall's τ-coefficient, and copula theory (Nelsen 2006). The correlation coefficients consists of a single, bounded value that describes the correlation between two input variables. The copula theory describes the dependence between random variables by a function which is characterized by a set of copula parameters.

In the context of imprecise probabilities and in particular p-boxes, few publications are available in the literature that are concerned with dependence, see *e.g.* Ferson & Troy Tucker (2006). The authors consider the following cases for free p-boxes: (i) independence, (ii) comonotonicity (maximal dependence), (iii) countermonotonicity (minimal dependence), (iv) linear relationship and correlation within specified interval, (v) linear relationship with unknown correlation, (vi) signed but otherwise unknown dependence, and (vii) unknown dependency (including any non-linear relationship). In this thesis, a simple application of imprecise correlation structures is discussed in the NASA challenge set in Chapter 9. The linear correlation between two input parameters is modelled by introducing an epistemic variable.

Although there are a few publications and examples available, there is still a need to characterize correlation in the context of imprecise probabilities by a thorough framework. In this context, the meta-modelling-based algorithms proposed in this thesis could be extended to allow for efficient analyses in the context of correlated input variables.

11.2.3 Structural reliability methods

The estimation of failure probabilities and quantiles can be a non-trivial task. In this thesis, Monte Carlo simulation is used to estimate failure probabilities and quantiles, due to its robustness with respect to input distributions and shape of the limit-state function. The downside of Monte Carlo simulation is its low convergence rate and hence its low efficiency when dealing with engineering problems. The reason is that failure probabilities are generally low and quantiles of interest are extreme.

Monte Carlo simulation works well for failure probabilities in the range of 10^{-1} to 10^{-5}. When failure probabilities are smaller, the number of samples required for an accurate simulation result becomes too large for practical purposes. Hence, more sophisticated methods, such as importance sampling and subset simulation may be used. However, they still require a comparably large number of model evaluations. Hence, a number of publications discuss the combination of advanced structural reliability analyses in combination with meta-models, as described in Section 5.1.

In the context of imprecise probabilities, only a few publications are available. Especially, the consistent use of meta-models is rare in the case of imprecise structural reliability analysis. As an example, Z. Zhang, Jiang, Han, et al. (2014) uses radial-basis expansions in combination with evidence theory. Based on the proposed two-level meta-modelling frameworks for free and parametric p-boxes, a modified structural reliability analysis can be imagined when using importance sampling or subset simulation instead of crude Monte Carlo simulation.

11.2.4 Robust design optimization

Engineers are concerned with adequate design of systems and processes. However, as uncertainties are omnipresent at the design stage as well as at the operation stage, this is a non-trivial task. Reliability-based design optimization (RBDO) aims at designing a physical or engineering system under a number of constraints which include reliability targets. Attempts, such as by Moustapha (2016) and Moustapha et al. (2016), have been made to combine RBDO with meta-models. To the author's knowledge, however, the analyses are limited to probabilistic approaches. Hence, RBDO takes into account the natural variability of input parameters for searching the optimal design.

A different concept in design optimization is *robust* design optimization (RDO) (Mulvey, Vanderbei, & Zenios 1995; Beyer & Sendhoff 2007). A robust design is a design for which the sensitivity with respect to uncertainties in the definition of the system and its input variables is small. In the context of p-boxes, this can relate to designs which are insensitive to the epistemic uncertainty in the input parameters. Hence, RDO could account for the epistemic uncertainties in the problem definition as opposed to the aleatory uncertainty in RBDO.

The p-box framework and the related algorithm presented in this thesis could be extended to a combination of the traditionally disjoint RBDO and RDO. A *robust reliability-based design optimization* (RRBDO) could search for the optimal design given reliability-based constraints and imprecise input parameters. In particular, the modelling of input parameters as p-boxes would lead to imprecise failure probabilities in the assessment of the reliability constraints. The handling of such imprecise quantities is then related to the concepts of RDO.

11.2.5 Risk assessment

In this thesis, probability distributions are the main measure of importance when comparing different realizations of an input variable. In the case of structural reliability analysis, the measure of importance is also related to whether a sample is in the failure domain or in the safety domain. Hence, the measure is a binary value represented by the indicator function $\mathbb{I}$. In practical examples, however, decisions are made based on a utility function, such as the potential consequences in *e.g.* financial terms or number of fatalities. The utility function is a complex function in the general case. Hence, the associated risk can be defined as the expected utility as follows:

$$R = \mathbb{E}\left[u\right] = \int_{\mathcal{D}_{\boldsymbol{X}}} f_{\boldsymbol{X}}\left(\boldsymbol{x}\right)\, u\left(\boldsymbol{x}\right)\, \mathrm{d}\boldsymbol{x}, \tag{11.1}$$

where $u\left(\boldsymbol{x}\right)$ denotes the utility for a given realization $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}}$, $f_{\boldsymbol{X}}\left(\boldsymbol{x}\right)$ is the PDF of variable $\boldsymbol{X}$ and R denotes the risk. Hence, the measure of importance in this application is the product of utility and PDF value.

A modification of the proposed adaptive Kriging approach can be imagined in order to efficiently estimate the overall risk instead of the failure probability. The associated learning function could focus on exploring areas of large risk contribution rather than areas close to the limit-state surface. Moreover, the use of p-boxes rather than probability distributions could be used to analyse the sensitivity of the expected utility/risk, in particular, in situations where information is scarce.

11.2.6 Bayesian inference

In this thesis, the focus lies on so-called forward propagation problems, including crude uncertainty propagation, failure probability estimation, and quantile estimation. Sensitivity analysis is the only method discussed in this thesis that relates the uncertainty in the response quantity of interest backwards to the uncertainty in the input variables. In fact, sensitivity analysis decomposes the response uncertainty according to contributions of the input uncertainties.

A common situation in practice is, however, that additional data becomes available at a later stage of the analysis. Then, the initial uncertainty model can be updated to make an improved prediction of the quantity of interest. Bayesian inference is a popular tool to do so (Gelman et al. 2009; Nagel & Sudret 2016a). Chapter 2 shows a simple example of Bayesian inference and mentions the term *robust Bayesian inference*. In conventional Bayesian inference the prior is defined as a probability distribution. In robust Bayesian inference (Berger 1985), a bunch of prior distributions are considered and used in a conventional Bayesian inference. Each prior leads to a potentially different posterior distribution given the observed data.

In the context of this thesis, the bunch of prior distributions could be modelled by a p-box. Then, the associated bunch of posterior distributions could also be modelled by a p-box. For estimating the posterior p-box efficiently, the proposed two-level algorithms could be applied to meta-model the forward model used to compute the likelihood in the Bayesian inference algorithm. Moreover, the likelihood could be approximated by a meta-model itself as discussed in Nagel & Sudret (2016b).

APPENDIX A

Notations

The following lists summarize the notations in the theory sections of this thesis. For the applications section, some letters and symbols might denote a physical property required to define the computational model. These double usages are, however, not mentioned here for the sake of clarity.

Symbol	**Description**
A	Area of a p-box
$\boldsymbol{a}$	Vector of PC-coefficients
$\mathcal{A}$	Multi-index set in PCE
$\mathrm{Bel}\left(\cdot\right)$	Belief function
$\mathcal{B}$	Beta distribution
$\boldsymbol{C}$	Input vector with uniformly distributed components $C_i \in [0,1]$
$\boldsymbol{c}$	Realization of $\boldsymbol{C}$
$\boldsymbol{\mathfrak{C}}$	Experimental design of $\boldsymbol{C}$
$\boldsymbol{\mathfrak{c}}$	Experimental design sample in $\boldsymbol{\mathfrak{C}}$
$\mathrm{CoV}\left[\cdot\right]$	Coefficient of variation
D	Total variance of a computational model
D_N^α	Kolmogorov-Smirnov (KS) critical value for sample size N and confidence level α
d	Geographical dimensionality of a random field
$\mathcal{D}$	Domain
$\mathcal{E}$	Event, subset of Ω
$\mathcal{E}^c$	Complementary event to $\mathcal{E}$
$\mathbb{E}\left[\cdot\right]$	Expected value operator
$\mathrm{EI}(\cdot)$	Expected improvement function
Err	Absolute error measure
Err_{emp}	Empirical error
Err_{gen}	Generalization error
Err_{LOO}	Leave-one-out error
err	Relative error measure
err_{emp}	Relative empirical error
err_{gen}	Relative generalization error

err_{LOO}	Relative leave-one-out error
$F(\cdot)$	Cumulative distribution function (CDF)
$\underline{F}$	Lower bound of a p-box, belief
$\overline{F}$	Upper bound of a p-box, plausibility
$f(\cdot)$	Probability density function (PDF)
$f'(\cdot)$	Prior PDF
$f''(\cdot)$	Posterior PDF
$\mathbf{F}$	Information matrix
$\mathcal{F}$	Collection of events in Ω (σ-algebra)
$\boldsymbol{f}(\cdot)$	Trend functions in Kriging
$\mathcal{G}$	Limit-state function
$\mathcal{GU}$	Gumbel distribution
H	Random field
$\mathbb{I}(\cdot)$	Indicator function
$\mathbf{I}$	Identity matrix
$\mathcal{I}$	Multi-index set
$\mathcal{J}$	Event in Ω, subset of Ω
K	Number of nodal points
$\boldsymbol{k}$	Multi-index in the slicing algorithm
$\mathcal{K}$	Multi-index set in the slicing algorithm
$\mathcal{L}(\cdot)$	Likelihood function
$\mathcal{LN}$	Lognormal distribution
M	Dimensionality of the input vector to $\mathcal{M}$ or $\mathcal{G}$
m	Basic probability assignment, probability mass
$\mathcal{M}$	Computational model
$\mathbb{M}_f$	Limit-state margin domain
N	Experimental design size
N_0	Initial experimental design size
$N^{(1)}$	First-level meta-model experimental design size
$N^{(2)}$	Second-level meta-model experimental design size
$N^{(\mathrm{EGO})}$	EGO experimental design size
n	Number of samples in a validation set
$n_{\mathcal{A}}$	Number of multi-indices $\boldsymbol{\alpha} \in \mathcal{A}$
n_f	Number of failed samples in Monte Carlo simulation
n_{MCS}	Number of samples in a Monte Carlo simulation
n_{ph}	Number of phantom points
n_T	Number of functions in the trend of a universal Kriging model
n_X	Number of discretization intervals in X in the slicing algorithm
n_Y	Number of resulting intervals for Y in the slicing algorithm
n_ζ	Number of random variables in an EOLE discretization of a random field
$\mathcal{N}$	Normal (Gaussian) distribution
$\mathbb{N}$	Space of positive integer numbers
$P(\cdot)$	Candidate polynomial
P_f	Failure probability
P_m	Probability of misclassification
p	Maximal total polynomial degree
$\mathrm{Pl}(\cdot)$	Plausibility function

$\mathbb{P}\left(\cdot\right)$	Probability measure
$\mathcal{P}$	Response values to experimental design $\boldsymbol{\mathcal{T}}$
q	Parameter of the q-norm
q_α	α-quantile
$R(\cdot)$	Correlation function
r	Parameter of beta distribution
$\mathbf{R}$	Correlation matrix
$\mathbb{R}$	Space of real numbers
S_i	First order Sobol' indices of the i-th variable
$S_i^{(\mathrm{T})}$	Total order Sobol' indices of the i-th variable
s	Parameter of beta distribution
$\mathcal{S}$	Candidate set in AK-MCS and APCK-MCS
$\mathrm{Std}\left[\cdot\right]$	Standard deviation
$T(\cdot)$	Isoprobabilistic transform
t	Quantile value of the t-distribution
U	U-function in adaptive experimental design algorithm
$\mathcal{U}$	Uniform distribution
$\mathrm{unc}\left(\cdot\right)$	Uncertainty measure
$\boldsymbol{V}$	Independent input variables in the augmented input space
$\boldsymbol{\mathfrak{V}}$	Experimental design in the augmented space
$\boldsymbol{v}$	Realization of $\boldsymbol{V}$
$\boldsymbol{\mathfrak{v}}$	Experimental design sample point in the augmented space
$\mathrm{Var}\left[\cdot\right]$	Variance
$\mathrm{VoI}\left[\cdot\right]$	Value of information
W	Response variable in the augmented space
w	Realization of W
$\mathfrak{W}$	Vector of response values corresponding to experimental design $\boldsymbol{\mathfrak{V}}$
$\boldsymbol{X}$	Input vector
$\widetilde{\boldsymbol{X}}$	Auxiliary input vector
$\boldsymbol{x}$	Realization of $\boldsymbol{X}$
Y	Response variable
y	Realization of Y
y_{adm}	Admissible response value
y_0	Limit-state parameter
$\mathcal{Y}$	Vector of response values corresponding to experimental design $\boldsymbol{\mathcal{X}}$
$\mathbb{Y}$	Composed vector of response values in EOLE
$Z(\cdot)$	Gaussian process

Symbol	**Description**
α	Membership function value and basis of α-cuts
$\boldsymbol{\alpha}$	Multi-index in PCE
β_{HL}	Hasofer-Lind reliability index
β_C	Cornell's reliability index
$\boldsymbol{\beta}$	Vector of coefficients in Kriging
$\Gamma(\cdot)$	Gamma function
γ	Confidence level parameter

Δ	Difference/range operator
δ	Kronecker delta
ϵ	Threshold value
ϵ_{EI}	Threshold value for the EI function in EGO
ϵ_{P_f}	Threshold value for APCK-MCS when estimating failure probabilities
ϵ_{q_α}	Threshold value for APCK-MCS when estimating quantiles
ζ	Standard normal variable
η	Coefficient of variation
$\boldsymbol{\Theta}$	Vector of epistemic parameter intervals
$\boldsymbol{\theta}$	Vector of epistemic parameters
$\boldsymbol{\vartheta}$	Hyper-parameters in Bayesian hierarchical model
λ	Eigenvalue of the covariance matrix $\boldsymbol{\Sigma}$
μ	Mean value
$\mu_{\widehat{Y}}(\cdot)$	Prediction mean value in Kriging
$\Xi(\cdot)$	Membership function of a fuzzy variable
$\xi(\cdot)$	Membership function of a fuzzy integral
ρ	Correlation length
$\boldsymbol{\rho}$	Vector of correlation lengths
$\boldsymbol{\Sigma}$	Covariance matrix
σ	Standard deviation
$\sigma_{\widehat{Y}}(\cdot)$	Prediction standard deviation in Kriging
$\boldsymbol{\mathcal{T}}$	Experimental design of $\boldsymbol{\Theta}$
$\boldsymbol{\tau}$	Experimental design sample in $\boldsymbol{\mathcal{T}}$
$\Phi(\cdot)$	Cumulative distribution function (CDF) of a standard normal variable
$\phi(\cdot)$	Candidate function
$\varphi(\cdot)$	Probability density function (PDF) of a standard normal variable
$\mathcal{X}_1$	Dataset used in Chapter 2
$\boldsymbol{\mathcal{X}}$	Experimental design of $\boldsymbol{X}$
χ^2	Quantile value of the χ^2 (chi-square) distribution
$\boldsymbol{\chi}$	Experimental design sample point in $\boldsymbol{\mathcal{X}}$
$\psi(\cdot)$	Multivariate orthogonal polynomial
Ω	Event space
ω	Elementary event

APPENDIX B

Acronyms

The following list summarizes the acronyms used multiple times throughout the thesis. In the chapters, the individual acronyms are introduced properly where mentioned first.

Acronym	**Description**
AK-MCS	Active learning reliability method combining Kriging and Monte Carlo simulation
ANOVA	Analysis of variance
APCK-MCS	Adaptive-Polynomial-Chaos-Kriging Monte Carlo simulation
BPA	Basic probability assignment
CDF	Cumulative distribution function
CI	Confidence interval
CPU	Central processing unit
CoV	Coefficient of variation
CV	Cross validation
DGSM	Derivative-based global sensitivity measures
DS	Dempster-Shafer
EFF	Expected feasibility function
EGO	Efficient global optimization
EI	Expected improvement
EOLE	Expansion optimal linear estimation method
FE	Finite element
FEM	Finite element model
FORM	First order reliability method
FOSM	First order second moment method
FOS	Factor of safety
GP	Gaussian process
ISA	Imprecise sensitivity analysis
ISRA	Imprecise structural reliability analysis
IUP	Imprecise uncertainty propagation
IUQ	Imprecise uncertainty quantification

KL	Karhunen-Loève
KS	Kolmogorov-Smirnov
LAR	Least angle regression
LARS	Least angle regression selection
LF	Learning function
LHS	Latin-hypercube sampling
LOO	Leave-one-out
LSF	Limit-state function
LSS	Limit-state surface
MC	Monte Carlo
MCS	Monte Carlo simulation
ML	Maximum likelihood
nMCS	Nested Monte Carlo simulation
OK	Ordinary Kriging
OPCK	Optimal Polynomial-Chaos-Kriging
p-box	Probability-box
PC	Polynomial chaos
PCE	Polynomial chaos expansions
PCK	Polynomial-Chaos-Kriging
PC-Kriging	Polynomial-Chaos-Kriging
PDF	Probability density function
QoI	Quantity of interest
RBDO	Reliability-based design optimization
RDO	robust design optimization
RF	Random field
SA	Sensitivity analysis
SDOF	Single degree of freedom
SFEM	Stochastic finite element method
SRA	Structural reliability analysis
SSFEM	Spectral stochastic finite element method
SPCK	Sequential Polynomial-Chaos-Kriging
UCB	Upper confidence bound
UP	Uncertainty propagation
UQ	Uncertainty quantification
VoI	Value of information

APPENDIX C

Autocorrelation functions

An essential part of Kriging meta-models is the choice of autocorrelation function describing the Gaussian process. A number of common autocorrelation functions are presented here. Furthermore, different methodologies to assemble the one-dimensional autocorrelation functions in multi-dimensional examples are shown.

C.1 Popular autocorrelation functions

Consider a one-dimensional input space $\mathcal{D}_X$. The autocorrelation function describes the correlation between two points. Various formulations of autocorrelation functions can be found in the literature. Some of the widely used autocorrelation functions are listed here (see also Sacks et al. (1989) and Dubourg (2011)):

- *Linear autocorrelation function*:

  $$R(x, x'; \rho) = \max\left(0,\, 1 - \frac{|x - x'|}{\rho}\right), \tag{C.1}$$

  where ρ is the so-called correlation length. The correlation is limited to a defined range described by the parameter ρ. This corresponds to the assumption that there is no correlation beyond a distance ρ in each dimension.

- *Exponential autocorrelation function*:

  $$R(x, x'; \rho) = \exp\left(-\frac{|x - x'|}{\rho}\right). \tag{C.2}$$

- *Squared exponential autocorrelation function*:

  $$R(x, x'; \rho) = \exp\left(-\left(\frac{x_i - x_i'}{\rho}\right)^2\right). \tag{C.3}$$

  This autocorrelation function is also called *Gaussian autocorrelation function.*

- *Matérn autocorrelation function*:

$$R(x, x'; \rho, \nu) = \frac{1}{2^{\nu-1}\,\Gamma(\nu)} \left(\sqrt{2\nu}\, \frac{|x - x'|}{\rho} \right)^{\nu} \kappa_{\nu} \left(\sqrt{2\nu}\, \frac{|x - x'|}{\rho} \right), \tag{C.4}$$

  where $\nu \geq 1/2$ is the shape parameter, $\Gamma(\cdot)$ is the Euler Gamma function and $\kappa_{\nu}(\cdot)$ is the modified Bessel function of the second kind (also known as Bessel function of the third kind). In practice the values $\nu = 3/2$ and $\nu = 5/2$ are the most popular forms. The formulation then simplifies into (Roustant, Ginsbourger, & Deville 2012):

$$R(x, x'; \rho, \nu = 3/2) = \left(1 + \frac{\sqrt{3}\,|x - x'|}{\rho} \right) \exp\left(- \frac{\sqrt{3}\,|x - x'|}{\rho} \right), \tag{C.5}$$

$$R(x, x'; \rho, \nu \;=\; 5/2) \;=\; \left(1 + \frac{\sqrt{5}\,|x - x'|}{\rho} + \frac{5\,(x - x')^2}{3\,\rho^2} \right) \exp\left(- \frac{\sqrt{5}\,|x - x'|}{\rho} \right). \tag{C.6}$$

- *Dirac (a.k.a. nugget) autocorrelation function*:

$$R(x, x') = \delta(x - x'), \tag{C.7}$$

  where δ is the Dirac function which is $\delta = 1$ for $x = x'$ and 0 otherwise. Note that the Dirac autocorrelation function has no correlation hyper-parameter ρ.

Graphs of the behaviour of the various autocorrelation functions can be found in *e.g.* Dubourg (2011).

C.2 Multi-dimensional correlation models

Consider now that the input domain $\mathcal{D}_{\boldsymbol{X}}$ is M-dimensional. Then, the correlation of two points depends on their distance to each other. In the literature, there are a number of models available to model multi-dimensional correlation, including *separable*, *ellipsoidal*, and *additive* correlation models.

C.2.1 Separable autocorrelation functions

Separable autocorrelation functions are a popular choice and are characterized as the product of component-wise correlations:

$$R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\rho}) = \prod_{i=1}^{M} R(x_i, x_i'; \rho_i), \tag{C.8}$$

where $\boldsymbol{\rho} = (\rho_1, \dots, \rho_M)$ is the vector of correlation lengths corresponding to the input variables X_i, $i = 1, \dots, M$. As an example, the linear autocorrelation function reads then:

$$R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\rho}) = \prod_{i=1}^{M} \max\left(0,\; 1 - \frac{|x_i - x_i'|}{\rho_i} \right). \tag{C.9}$$

Separable autocorrelation functions work well for low-dimensional problems. Hence in this thesis, separable autocorrelation models are used in all examples when calibrating a Kriging model or a random field.

C.2.2 Ellipsoidal autocorrelation functions

Apart from the separable autocorrelation functions, ellipsoidal autocorrelation functions assume that the correlation depends on the distance between $\boldsymbol{x}$ and $\boldsymbol{x}'$ directly. The corresponding correlation functions then read (Rasmussen & C. Williams 2006):

$$R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\rho}) = R(h), \qquad h = \sqrt{\sum_{i=1}^{M} \left(\frac{x_i - x_i'}{\rho_i} \right)^2}, \tag{C.10}$$

where h is the weighted distance of $\boldsymbol{x}$ and $\boldsymbol{x}'$. As an example, the linear autocorrelation function reads then:

$$R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\rho}) = \max\left(0,\, 1 - h\right). \tag{C.11}$$

C.2.3 Additive autocorrelation functions

Recently, new *additive* autocorrelation functions were proposed (Ginsbourger, Durrande, & Roustant 2013; Durrande, Ginsbourger, & Roustant 2012; Duvenaud, Nickisch, & Rasmussen 2012; Durrande, Ginsbourger, & Roustant 2013) which seem to be suitable for high-dimensional problems. Additive kernels inherit a summation in the autocorrelation function instead of a multiplication:

$$R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\rho}) = \frac{1}{M} \sum_{i=1}^{M} R(x_i, x_i'; \rho_i). \tag{C.12}$$

As an example, the linear autocorrelation function reads then:

$$R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\rho}) = \frac{1}{M} \sum_{i=1}^{M} \max\left(0,\, 1 - \frac{|x_i - x_i'|}{\rho_i}\right). \tag{C.13}$$

By choosing adequate additive kernels, the behaviour of the Gaussian process can be modelled more accurately (Duvenaud, Nickisch, & Rasmussen 2012).

APPENDIX D

O'Hagan function

The O'Hagan function is defined as $f_6(\boldsymbol{x}) = \boldsymbol{a}_1^\mathsf{T}\boldsymbol{x} + \boldsymbol{a}_2^\mathsf{T}\sin(\boldsymbol{x}) + \boldsymbol{a}^\mathsf{T}\cos(\boldsymbol{x}) + \boldsymbol{x}^\mathsf{T}\boldsymbol{Q}\boldsymbol{x}$ (Oakley & O'Hagan 2004). This appendix provides the numerical values of $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{Q}\}$.

$$\boldsymbol{a}_1 = \begin{pmatrix} 0.0118 \\ 0.0456 \\ 0.2297 \\ 0.0393 \\ 0.1177 \\ 0.3865 \\ 0.3897 \\ 0.6061 \\ 0.6159 \\ 0.4005 \\ 1.0741 \\ 1.1474 \\ 0.7880 \\ 1.1242 \\ 1.1982 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 0.4341 \\ 0.0887 \\ 0.0512 \\ 0.3233 \\ 0.1489 \\ 1.0360 \\ 0.9892 \\ 0.9672 \\ 0.8977 \\ 0.8083 \\ 1.8426 \\ 2.4712 \\ 2.3946 \\ 2.0045 \\ 2.2621 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 0.1044 \\ 0.2057 \\ 0.0774 \\ 0.2730 \\ 0.1253 \\ 0.7526 \\ 0.8570 \\ 1.0331 \\ 0.8388 \\ 0.7970 \\ 2.2145 \\ 2.0382 \\ 2.4004 \\ 2.0541 \\ 1.9845 \end{pmatrix}$$

$$\boldsymbol{Q} = \begin{pmatrix}
-0.0225 & -0.1850 & 0.1342 & 0.3687 & 0.1717 & 0.1365 & -0.4403 & -0.0814 & 0.7132 & -0.4436 & 0.5038 & -0.0241 & -0.0459 & 0.2167 & 0.0559 \\
0.2566 & 0.0538 & 0.2580 & 0.2380 & -0.5913 & -0.0816 & -0.2875 & 0.4158 & 0.4975 & 0.0839 & -0.1106 & 0.0332 & -0.1398 & -0.0310 & -0.2232 \\
-0.0560 & 0.1954 & 0.0955 & -0.2863 & -0.1444 & 0.2237 & 0.1453 & 0.2900 & 0.2311 & -0.3193 & -0.2904 & -0.2096 & 0.4314 & 0.02440 & 0.0449 \\
0.6645 & 0.4307 & 0.2992 & -0.1620 & -0.3148 & -0.3903 & 0.1768 & 0.0580 & 0.1723 & 0.1347 & -0.3528 & 0.2515 & -0.0188 & 0.3648 & -0.3250 \\
-0.1213 & 0.1246 & 0.1066 & 0.0466 & -0.2168 & 0.1949 & -0.0655 & 0.02440 & -0.0968 & 0.1937 & 0.3335 & 0.3130 & -0.0836 & -0.2534 & 0.3733 \\
-0.2838 & -0.3282 & -0.1050 & -0.2207 & -0.1371 & -0.1443 & -0.1150 & 0.2242 & -0.0304 & -0.5151 & 0.0173 & 0.0390 & 0.3607 & 0.3090 & 0.0500 \\
-0.0779 & 0.0037 & 0.8869 & -0.2659 & -0.0793 & -0.0427 & -0.1865 & -0.3560 & -0.1750 & 0.0887 & 0.4003 & -0.0560 & 0.1372 & 0.2149 & -0.0113 \\
-0.0923 & 0.5921 & 0.0313 & -0.0331 & -0.2431 & -0.0998 & 0.0345 & 0.0951 & -0.3380 & 0.0064 & -0.6121 & 0.0813 & 0.8868 & 0.1425 & 0.1478 \\
-0.1319 & 0.5288 & 0.1265 & 0.0451 & 0.5837 & 0.3729 & 0.1140 & -0.2948 & -0.5701 & 0.4629 & -0.0941 & 0.1396 & -0.3861 & -0.449 & -0.1460 \\
0.0581 & -0.3229 & 0.0931 & 0.0724 & -0.5692 & 0.5255 & 0.2366 & -0.0118 & 0.0718 & 0.0783 & -0.1336 & 0.2272 & 0.1437 & -0.452 & -0.5557 \\
0.6615 & 0.3463 & 0.1410 & 0.5188 & -0.2802 & -0.1603 & -0.0684 & -0.2043 & 0.0697 & 0.2311 & -0.0444 & -0.1646 & 0.2162 & 0.0043 & -0.0874 \\
0.3160 & -0.0276 & 0.1343 & 0.1350 & 0.0540 & -0.1737 & 0.1753 & 0.0603 & -0.1791 & -0.3106 & -0.2536 & 0.0258 & -0.4301 & -0.6227 & -0.0340 \\
-0.2904 & 0.0341 & 0.0349 & -0.1212 & 0.0260 & -0.3355 & -0.4142 & 0.0532 & -0.271 & -0.0263 & 0.4102 & 0.2664 & 0.1558 & -0.1867 & 0.0199 \\
-0.2439 & -0.4410 & 0.0126 & 0.2495 & 0.0711 & 0.2462 & 0.1748 & 0.0085 & 0.2515 & -0.1466 & -0.0846 & 0.3693 & -0.2996 & 0.1104 & -0.7569 \\
0.0415 & -0.2598 & 0.4640 & -0.3611 & -0.9498 & -0.1650 & 0.0031 & 0.0528 & 0.2252 & 0.3839 & 0.4556 & -0.1863 & 0.0082 & 0.1667 & 0.1605
\end{pmatrix}$$

APPENDIX E

Learning functions

Learning functions are the core of Kriging meta-models with adaptive experimental design strategies. This appendix chapter illustrates four learning function found in the literature. Their comparison has been published previously in Schöbi & Sudret (2014e).

E.1 Overview

A *learning function* (LF) of a candidate state $\boldsymbol{x}$ estimates the expected value of information gained with respect to the objective function when point $\boldsymbol{x}$ is added to the experimental design $\boldsymbol{\mathcal{X}}$ of the (Kriging) meta-model. In the following, four learning functions are presented, namely the expected feasibility function, the U-function, the expected improvement function and the Gaussian process upper confidence bounds function. All of them are popular choices in adaptive experimental design algorithms and have been applied to optimization algorithms as well as to structural reliability problems.

The presentation of these four functions focuses on the task of estimating a failure probability, which is in principle a classification problem. The input domain is split into a safety and a failure domain. It is assumed that the failure domain is defined as $Y \leq y_0$ and the safety domain as $Y > y_0$.

E.2 Expected feasibility function

The *expected feasibility function* (EFF) (Bichon, Eldred, et al. 2008) originates from the *efficient global reliability analysis* algorithm (EGRA). In the context of structural reliability analysis, EFF provides an indication of how well the true value of the response is expected to satisfy the equality constraint $\mathcal{G}(\boldsymbol{x}) = y_0$. This expectation can be calculated by integrating over a region in the immediate vicinity of the limit-state parameter, say $y_0 \pm \epsilon$:

$$\mathrm{EFF}\left(\widehat{Y}(\boldsymbol{x})\right) = \int_{y_0-\epsilon}^{y_0+\epsilon} \left(\epsilon - |y_0 - y|\right) f_{\widehat{Y}}(y)\,\mathrm{d}y, \tag{E.1}$$

where $f_{\widehat{Y}}$ is the PDF of the Gaussian variable $\widehat{Y}(\boldsymbol{x})$, y_0 is the threshold value and ϵ is a tuning parameter with a small value. In the context of Kriging, $\widehat{Y}(\boldsymbol{x})$ is a Gaussian variable with mean value $\mu_{\widehat{Y}}(\boldsymbol{x})$ and standard deviation $\sigma_{\widehat{Y}}(\boldsymbol{x})$. Thus the integral can be computed analytically:

$$\begin{aligned}
&\mathrm{EFF}(\boldsymbol{x}) = \\
&\quad \left(\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0\right) \left[2\,\Phi\left(\frac{y_0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \Phi\left(\frac{(y_0 - \epsilon) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \Phi\left(\frac{(y_0 + \epsilon) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)\right] \\
&\qquad - \sigma_{\widehat{Y}}(\boldsymbol{x}) \left[2\,\varphi\left(\frac{y_0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \varphi\left(\frac{(y_0 - \epsilon) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \varphi\left(\frac{(y_0 + \epsilon) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)\right] \\
&\qquad\qquad + \epsilon \left[\Phi\left(\frac{(y_0 + \epsilon) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \Phi\left(\frac{(y_0 - \epsilon) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)\right], \qquad \text{(E.2)}
\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $\varphi(\cdot)$ is the standard normal PDF and y_0 describes the limit-state threshold value (which is equal to $y_0 = 0$ in the case of failure probability estimation). In EGRA, the expected feasibility function is built with a range ϵ that depends on the point $\boldsymbol{x}$. Usually, the value $\epsilon = 2\,\sigma_{\widehat{Y}}(\boldsymbol{x})$ is used, which describes the region or interest around y_0, *i.e.* $y_0 \pm 2\sigma_{\widehat{Y}}(\boldsymbol{x})$. Note that ϵ is a tuning parameter of the EFF algorithm and can thus be tuned towards points in the vicinity of the limit-state surface or towards a more global search.

The best next point of the input space is then determined by:

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}\in\mathcal{S}}(\mathrm{EFF}(\boldsymbol{x})), \qquad \text{(E.3)}$$

where $\mathcal{S}$ describes the set of candidate points in the enrichment.

Denote $t = \mu_{\widehat{Y}}(\boldsymbol{x}) - y_0$ and $s = \sigma_{\widehat{Y}}(\boldsymbol{x})$. It follows that $\epsilon = 2s$ and that EFF in Eq. (E.2) can be formulated in terms of $\{t, s\}$:

$$\begin{aligned}
\widetilde{\mathrm{EFF}}(t,s) = t &\left[2\,\Phi\left(\frac{-t}{s}\right) - \Phi\left(\frac{-t-2s}{s}\right) - \Phi\left(\frac{-t+2s}{s}\right)\right] \\
&- s\left[2\,\varphi\left(\frac{-t}{s}\right) - \varphi\left(\frac{-t-2s}{s}\right) - \varphi\left(\frac{-t+2s}{s}\right)\right] \\
&\qquad + 2s\left[\Phi\left(\frac{-t+2s}{s}\right) - \Phi\left(\frac{-t-2s}{s}\right)\right]. \qquad \text{(E.4)}
\end{aligned}$$

Figure E.1 displays the value of $\widetilde{\mathrm{EFF}}$ for an input variable $\boldsymbol{x}$ whose t ranges in the interval $[-10, 10]$ and whose s takes the values $\{1, 2, 3\}$. Note that the curves are symmetric around the value $t = 0$ (*i.e.* $\mu_{\widehat{Y}}(\boldsymbol{x}) = y_0$) and resemble Gaussian probability density functions.

E.3 *U*-function

The *U*-function and the related probability of misclassification P_m are introduced in Section 5.3. Note that the best next point of the input space is then determined by:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathcal{S}} U(\boldsymbol{x}), \qquad \text{(E.5)}$$

as opposed to other criteria in this chapter.

For illustration purposes, the variables $\{t, s\}$ are used to transform the *U*-function and the probability of misclassification into:

$$\widetilde{U}(t,s) = \frac{|t|}{s}, \qquad \widetilde{P}_M = \Phi\left(-\frac{|t|}{s}\right). \qquad \text{(E.6)}$$

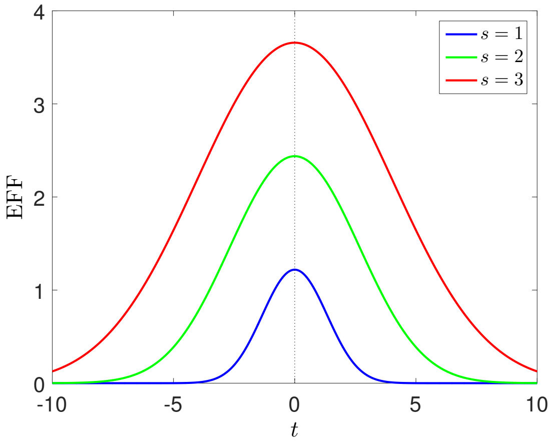
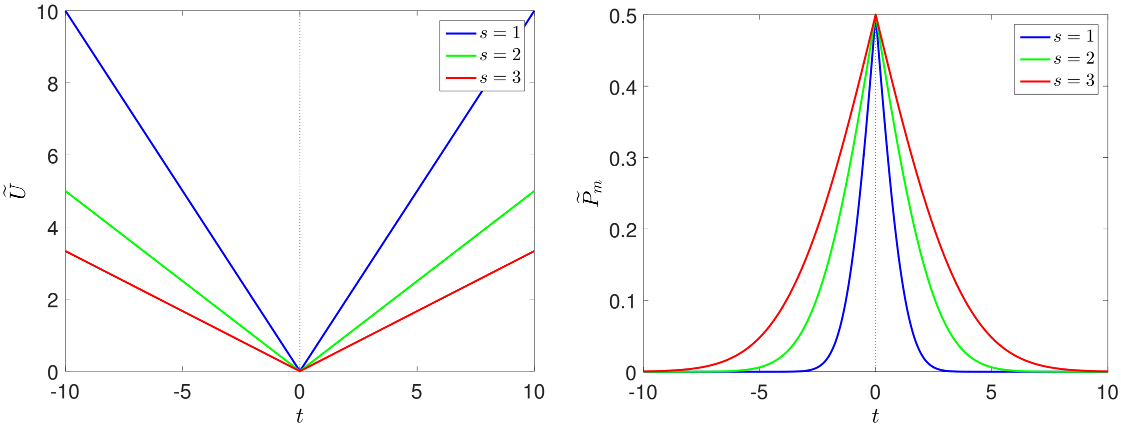


Figure E.1: Expected feasibility function $\widetilde{\text{EFF}}$.

Figure E.2 summarizes the $\widetilde{U}$-function and the corresponding probability of misclassification $\widetilde{P}_m$. Note that the maximum value of the probability of misclassification is $\max \widetilde{P}_m = 0.5$ for $t = 0$ (*i.e.* $\mu_{\widehat{Y}}(\boldsymbol{x}) = y_0$), which is independent of s. Thus the probability of misclassification is bounded within the interval $P_m \in [0, 0.5]$.



(a) $\widetilde{U}$-function

(b) Probability of misclassification $\widetilde{P}_m$

Figure E.2: $\widetilde{U}$-function and the corresponding probability of misclassification $\widetilde{P}_m$.

E.4 Expected improvement function

Ginsbourger, Rosspopoff, et al. (2013) proposed an expected improvement method (EI) using a Kriging meta-model in order to find the minimum of a computational model output when the input varies within a range. It is based on a formulation by Jones, Schonlau, & Welch (1998) for a minimization operation:

$$EI(\boldsymbol{x}) = \mathbb{E}\left[\max(\mu_{\min} - G, 0)\right], \tag{E.7}$$

which simplifies into:

$$EI(\boldsymbol{x}) = (\mu_{\min} - \mu_{\widehat{Y}}(\boldsymbol{x}))\, \Phi\left(\frac{\mu_{\min} - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) + \sigma_{\widehat{Y}}(\boldsymbol{x})\, \varphi\left(\frac{\mu_{\min} - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \tag{E.8}$$

where $\mu_{\min}$ is the minimum value of the prediction mean values, *i.e.* $\mu_{\min} = \min_{\boldsymbol{x}\in\mathcal{S}} \mu_{\widehat{Y}}(\boldsymbol{x})$. This minimization can be adapted to classification problems by taking the absolute Kriging prediction values, replacing the minimum value $\mu_{\min}$ with zero and replacing $\mu_{\widehat{Y}}(\boldsymbol{x})$ with $\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0$ in order to foster the closeness to the threshold y_0.

$$EI(\boldsymbol{x}) \quad = \quad \left(-|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|\right) \quad \Phi\left(\frac{-|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) \; + \; \sigma_{\widehat{Y}}(\boldsymbol{x}) \quad \varphi\left(\frac{-|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right). \tag{E.9}$$

The goal of adaptive experimental design algorithms is to accurately meta-model the limit-state function, but also to find a representative subset of input samples leading to a response curve close to the reference, *i.e.* to small misfits. Thus, Ginsbourger, Rosspopoff, et al. (2013) modified the original formulation in Eq. (E.8) by using the quantile value q_α instead of the minimum value $\mu_{\min}$ of the absolute prediction mean values. This lowers the repulsion effect of the current best point, *i.e.* the best next point shifts to the area of the quantile instead of the global maxima. The modified equation including the threshold y_0 reads then:

$$EI_\alpha(\boldsymbol{x}) \quad = \quad \left(q_\alpha - |\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|\right) \; \Phi\left(\frac{q_\alpha - |\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) \; + \; \sigma_{\widehat{Y}}(\boldsymbol{x})\, \varphi\left(\frac{q_\alpha - |\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \tag{E.10}$$

where q_α is the empirical α-quantile of the set of absolute shifted prediction mean values $|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|$. The tuning of α distinguishes between local and global search. When α is set to zero, then q_α coincides with the minimum of $|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0|$, leading back to Eq. (E.9).

The best next sample is then determined by choosing the maximum EI_α-value among the samples in the sample set $\mathcal{S}$:

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}\in\mathcal{S}} \left(EI_\alpha(\boldsymbol{x})\right).$$

Consider the expected improvement method as a function of $\{t, s\}$ which reads:

$$\widetilde{EI}_\alpha = \left(\widetilde{q}_\alpha - |t|\right) \, \Phi\left(\frac{\widetilde{q}_\alpha - |t|}{s}\right) + s\, \varphi\left(\frac{\widetilde{q}_\alpha - |t|}{s}\right), \tag{E.11}$$

where $\widetilde{q}_\alpha$ is the empirical α-quantile of the set of $\{|t_i|\}$. Figure E.3 illustrates the behaviour of the expected improvement function according to Eq. (E.11). Figures E.3(a) and E.3(b) show the cases of $\widetilde{q}_\alpha = 0$ and $\widetilde{q}_\alpha = 1$, respectively. The plots show that the expected improvement is positively correlated to s and negatively correlated to $|t|$.

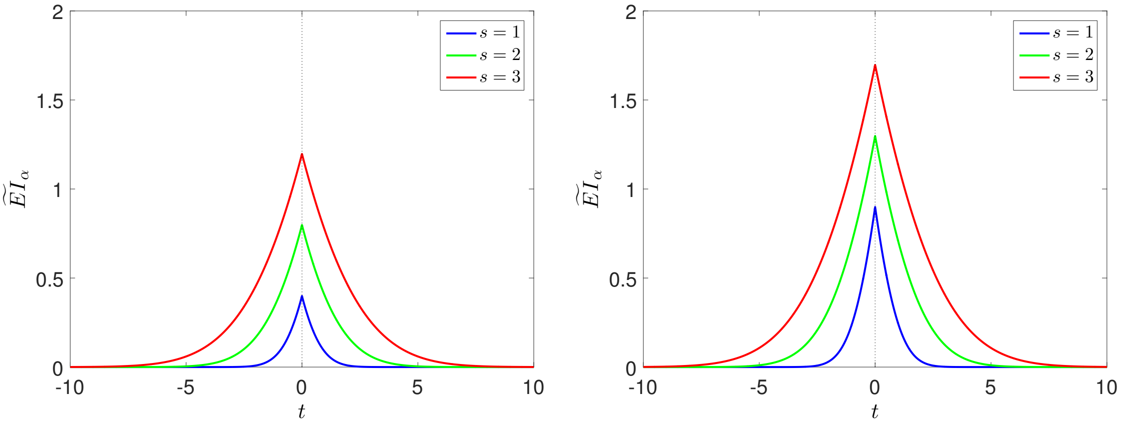
E.5 Gaussian process upper confidence bounds (GP-UCB)

Srinivas et al. (2012) proposed the *Gaussian process upper confidence bounds* approach (GP-UCB) primarily to find the maximum of a function. In the context of classification problems, the general formulation can be slightly modified to find points which are either close to the limit-state surface or whose prediction is highly uncertain, or both. The modified equation reads:

$$UCB(\boldsymbol{x}) = -|\mu_{\widehat{Y}}(\boldsymbol{x}) - y_0| + \sqrt{\beta_N}\, \sigma_{\widehat{Y}}(\boldsymbol{x}), \tag{E.12}$$

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}\in\mathcal{S}} UCB(\boldsymbol{x}), \tag{E.13}$$

where $\sqrt{\beta_N}$ is the parameter distinguishing between exploration (global search) and exploitation (local search). The special feature of this learning function is the evolution of the function values



(a) $\widetilde{q}_\alpha = 0$ (b) $\widetilde{q}_\alpha = 1$

Figure E.3: Expected improvement learning function $\widetilde{EI}_\alpha$ according to Eq. (E.11).

depending on the number of samples in the growing experimental design. In Srinivas et al. (2012), and after Dani, Hayes, & Kakade (2008), two formulations are suggested for the parameter β_N:

$$\beta_N^{(1)} = 2\,\log(|\mathcal{S}|\,N^2\,\pi^2/6\delta) \tag{E.14}$$

$$\beta_N^{(2)} = 2\,\log(M^2\,2\pi^2/(3\delta)) + 2M\,\log\left(N^2 M\,b\,r\sqrt{\log(4M\,c/\delta)}\right), \tag{E.15}$$

where $\mathcal{S}$ is a set of a Monte Carlo population of the input space, $|\mathcal{S}|$ is its cardinality, N is the current number of samples in the experimental design, $0 < \delta < 1$ is a tuning parameter, M is the number of dimensions in the input space and $\{b,\,c\}$ are additional tuning parameters. Note that β_N changes with every added sample, as N increases in every iteration of an adaptive meta-modelling technique. With increasing number of samples, the variance of the prediction points becomes more important compared to the prediction mean value and β_N becomes larger. Furthermore, the tuning parameters $\{b,\,c,\,\delta\}$ have a considerable effect on the result.

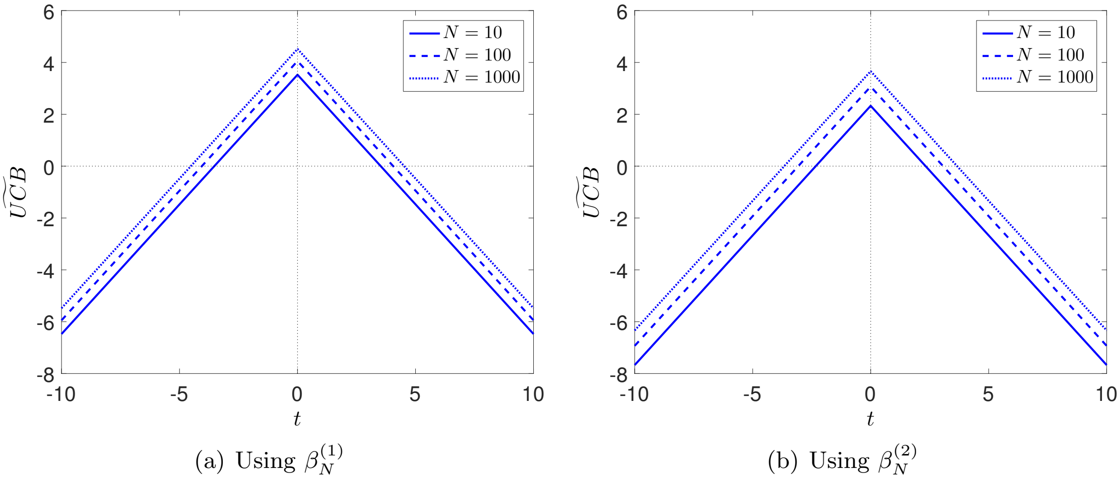
These two formulations of β_N originate from the context of decision planning for a bandit problem. The goal is to optimally decide what move to make in the next time step to minimize the regret, which is defined as the difference between maximum possible reward and the currently achieved reward ({regret} = {max. possible reward} − {achieved reward}). This resembles a global search for the maximum possible reward, *i.e.* smallest regret (Dani, Hayes, & Kakade 2008).

Let us consider UCB as a function of $\{t, s\}$:

$$\widetilde{UCB}(t,s) = -|t| + \sqrt{\beta_N}\cdot s. \tag{E.16}$$

Figure E.4 shows some realizations of the $\widetilde{UCB}$ learning function on a one-dimensional example. The following settings are chosen for Figure E.4(a): the learning function is applying $\beta_N^{(1)}$ defined in Eq. (E.14), $s=1$, the candidate sample set has a size of $|\mathcal{S}| = 10^4$ and $\delta = 1$. The different lines represent different number of samples N in the current experimental design. As the number of samples increases, the UCB-function becomes higher, thus allowing for more exploration.

Figure E.4(b) shows the analogous figure for $\beta_N^{(2)}$ (see Eq. (E.15)). The tuning parameters are set to $b=1$, $c=1$ and $\delta=1$. It is clear that the role of $\beta_N^{(2)}$ is similar to that of $\beta_N^{(1)}$.



(a) Using $\beta_N^{(1)}$ (b) Using $\beta_N^{(2)}$

Figure E.4: $\widetilde{UCB}$ learning function – comparison of the two β_N formulations ($s = 1$, $\delta = 1$, $|\mathcal{S}| = 10^4$, $b = 1$, $c = 1$).

E.6 Discussion

In the previous section different learning functions have been presented. In practice each of them will lead to a different best next point $\boldsymbol{x}^*$. In this section, the learning functions are compared on a specific example to highlight their behaviour.

For the following discussion, let us consider the so-called *hat function*, which is defined as (see also Eq. (5.23)):

$$y = f_{\text{hat}}(\boldsymbol{x}) = 20 - (x_1 - x_2)^2 - 8 \cdot (x_1 + x_2 - 4)^3, \tag{E.17}$$

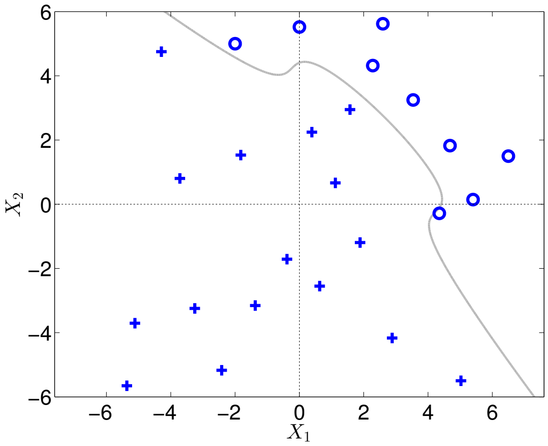
where the two-dimensional input parameters x_i are modelled by standard Gaussian variables, *i.e.* $X_i \sim \mathcal{N}(0, 1)$, $i = 1, 2$.

The function is meta-modelled with an ordinary Kriging model whose experimental design is shown in Figure E.5(a). This figure also shows the exact contour of the limit-state surface for $y = 0$. Figure E.5(b) shows the contour plot of the prediction mean values $\mu_{\widehat{Y}}(\boldsymbol{x})$ of the calibrated Kriging model and the current estimate of the limit-state function, whereas Figure E.5(c) shows the corresponding prediction standard deviation $\sigma_{\widehat{Y}}(\boldsymbol{x})$ of the same model.

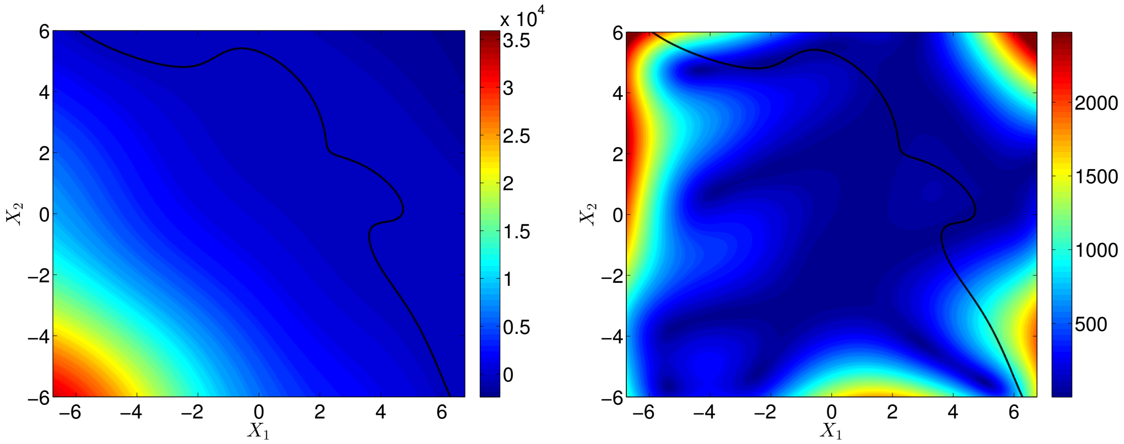
Figure E.6 shows the contour plot of the four previously discussed learning functions applied to this example. The blue markers represent the experimental design, in particular the "+" markers denote the samples of which $\mathcal{Y}^{(i)} > 0$ and the circles denote the samples of which $\mathcal{Y}^{(i)} < 0$. The solid dark line represents the current estimate of the limit-state surface, whereas the grey dashed line represents the exact limit-state surface. The learning functions are represented by the grey-scaled colour regions, where the darker regions are the regions of interest for finding the next point for enrichment. The green and red contour lines represent the zero values for $\{\mu_{\widehat{Y}}(\boldsymbol{x}) - 2 \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) = 0\}$ and $\{\mu_{\widehat{Y}}(\boldsymbol{x}) + 2 \cdot \sigma_{\widehat{Y}}(\boldsymbol{x}) = 0\}$ respectively and provide the confidence interval of the estimate of the limit-state surface.

Figures E.6(a) to E.6(d) display the EFF, U-function, expected improvement (EI) and the UCB learning functions, respectively. Note that for EFF, EI and UCB, the plotted values are normalized by the standard deviation of the output variable Y in order to get dimensionless values, *e.g.* $EFF/\text{Std}[Y]$. This results in comparable values for all four learning functions.

When comparing the four learning functions it becomes visible that the best next point $\boldsymbol{x}^*$ will be determined at different coordinates depending on the choice made. EFF, EI and UCB



(a) Experimental design and limit-state surface



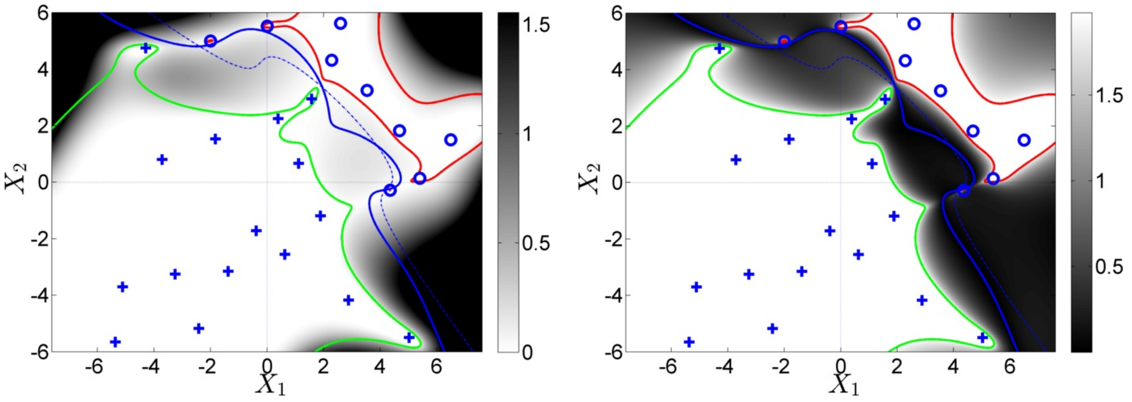
(b) Predition mean value $\mu_{\widehat{Y}}(\boldsymbol{x})$

(c) Prediction standard deviation $\sigma_{\widehat{Y}}(\boldsymbol{x})$

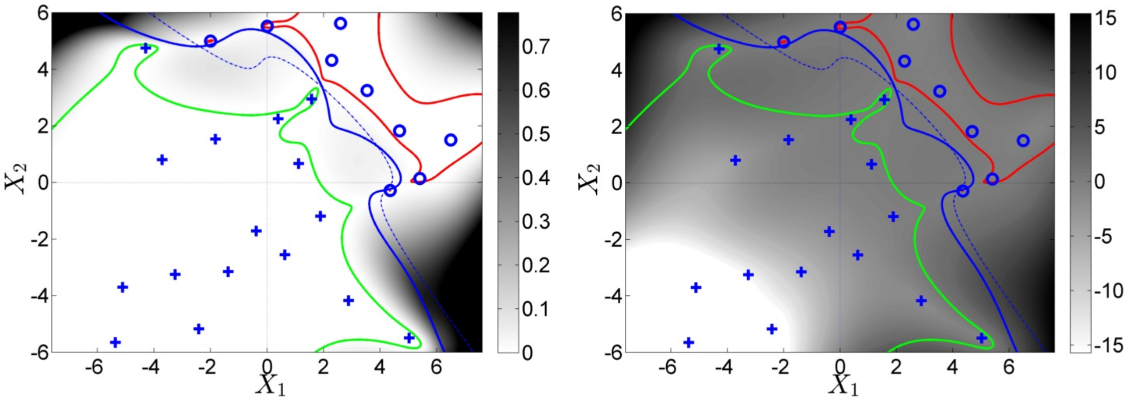
Figure E.5: Hat function – Kriging model.

tend to identify interesting samples along the boundaries of the presented domain in Figure E.6, whereas the U-function will select a point inside the domain, more specifically in the limit-state margin $\mathbb{M}_f$, which lies between the group of "+" and the group of "○".

Note that for the UCB function in Figure E.6(d), samples around the existing experimental design are also of interest due to the formulation of the learning function and the relative importance of prediction mean $\mu_{\widehat{Y}}(\boldsymbol{x})$ and standard deviation $\sigma_{\widehat{Y}}(\boldsymbol{x})$. In contrast, the other three learning functions assess samples close to the experimental design as not interesting (this is seen from the white regions around the experimental design samples). This behaviour can also be observed when comparing learning functions and the red and green contour lines in the figures. EFF, U-function and expected improvement function focus on the region $\mathbb{M}_f$, *i.e.* on the region between the green and red contour lines. In the case of UCB, this distinction is not as clear as in the other three cases.



(a) $EFF/\mathrm{Std}[\mathrm{Y}]$ (b) U-function



(c) $EI_\alpha/\mathrm{Std}\,[Y]$ with $\alpha = 0$ (d) $UCB/\mathrm{Std}\,[Y]$ with $\beta_N^{(1)}$

Figure E.6: Hat function – Different learning functions of the same Kriging meta-model.

APPENDIX F

Random fields

This chapter introduces the theory of random fields and some basic algorithms to generate realizations of random fields. Furthermore, conditional random fields are defined in the case where spatial data is available. For this case, a modification of the conventional EOLE is discussed, which has been published in Schöbi & Sudret (2015a).

F.1 State of the art

Random fields (RF) theory is a mathematical concept which allows the modelling of uncertainties of physical properties in continuous media with respect to randomness and spatial variability (Sudret & Der Kiureghian 2000). Basic definitions and general properties can be found in Lin (1967) and E. Vanmarcke (1983).

In engineering disciplines, an important application of the RF theory are the so-called *stochastic finite element methods* (SFEM) which aim at solving a finite element model (FEM) subject to randomness in the formulation of the finite element model itself. SFEM is generally based on sampling-based approaches, where realizations of the random field are generated and the resulting FE simulation is carried out. Hence, an important topic in the literature on random fields is discretization procedures.

Discretization methods allow for representing random fields in a way that is suitable for a subsequent sampling. Indeed, random fields may be considered as an infinite set of random variables, which can thus not be sampled before some mathematical treatment. The discretization methods can be divided into three groups: point discretization, average discretization, and series expansion methods (C. Li & Der Kiureghian 1993; Ditlevsen 1996; Matthies et al. 1997). The point discretization methods include the *midpoint method* (MP) (Der Kiureghian & Ke 1988), the *shape function method* (SF) (Liu, Belytschko, & Mani 1986b,a), the *integration point method* (Matthies et al. 1997), and the *optimal linear estimation method* (OLE) (C. Li & Der Kiureghian 1993). The average discretization methods include the *spatial average method* (SA) (E. Vanmarcke 1983) and the *weighted integral method* (Deodatis 1990, 1991; Takada 1990a,b; Deodatis & Shinozuka 1991). And finally, the series expansion methods include the *Karhunen-Loève expansion method* (KL expansion) (Loève 1977), the *orthogonal series expansion methods* (OSE) (J. Zhang & Ellingwood 1994), and the *expansion optimal linear estimation* *method* (EOLE) (C. Li & Der Kiureghian 1993). A comprehensive comparison of these methods can be found in *e.g.* Sudret & Der Kiureghian (2000).

Recent studies discuss random fields apart from standard setups. Cho, Venturi, & Karniadakis (2013) discuss KL expansions for correlated random fields. Vorechovský (2008) discusses EOLE and KL expansions of correlated non-Gaussian random fields. Le (2014) uses the local average subdivision algorithm to sample correlated random fields.

The methods above assume that the value of the random field is unknown at every geographical location $\boldsymbol{z} \in \mathcal{D}_{\boldsymbol{Z}}$. However, when the value of the RF is known at some location, a *conditional random field* can be defined taking into account the additional information. Monteiro, Ramos, & Hatherly (2009) model random fields conditional of observations by direct modelling. E. H. Vanmarcke & Fenton (1991) and Hoshiya (1995) generate realizations of conditional random fields by Kriging. Kameda & Morikawa (1992, 1994) simulate conditional random fields by Fourier series expansion. Elishakoff, Ren, & Shinozuka (1994) discuss conditional simulation techniques for non-Gaussian random fields.

F.2 Theory

A random field $H(\boldsymbol{z}, \omega)$ is defined as a collection of random variables, which are indexed by the continuous parameter $\boldsymbol{z} \in \mathcal{D}_{\boldsymbol{Z}}$. The space $\mathcal{D}_{\boldsymbol{Z}} \subset \mathbb{R}^d$ describes the geometry of the system. In typical engineering applications, $\boldsymbol{z}$ describes geographical locations, *i.e.* coordinates. $\omega \in \Omega$ denotes the variability of the random field. Note that for a given location $\boldsymbol{z}^*$, the random field $H(\boldsymbol{z}^*, \omega)$ is a random variable, whereas for a given ω^*, the random field describes a curve in the domain $\mathcal{D}_{\boldsymbol{Z}}$.

A special case is the *Gaussian random field* where $H(\boldsymbol{z}, \omega)$ is a Gaussian random variable at any location $\boldsymbol{z} \in \mathcal{D}_{\boldsymbol{Z}}$. A Gaussian RF may be completely defined by its mean value $\mu(\boldsymbol{z})$, variance $\sigma^2(\boldsymbol{z})$, and autocorrelation function $R(\boldsymbol{z} - \boldsymbol{z}' | \boldsymbol{\rho})$. Due to some handy mathematical properties, the Gaussian RF is a popular choice in the literature and it is also the focus of this chapter.

F.3 Expansion optimal linear estimation method (EOLE)

The *expansion optimal linear estimation method* (EOLE) is an extension of the *optimal linear estimation method* (OLE), which was first published by C. Li & Der Kiureghian (1993). OLE can also be seen as a particular case of the Kriging method. It is a special case of the method of *regression of linear functionals* discussed in Ditlevsen (1996). EOLE is a discretization method that approximates the random field $H(\boldsymbol{z}, \omega)$ by some function $\widehat{H}(\boldsymbol{z}, \boldsymbol{\zeta})$, where $\boldsymbol{\zeta} = \{\zeta_i,\, i = 1, \ldots, n_\zeta\}$ is a finite set of random variables describing the randomness in the RF. The explicit function $\widehat{H}$ allows one to sample the random field H in a straightforward way by sampling $\boldsymbol{\zeta}$.

Assume that $H(\boldsymbol{z}, \omega)$ is a Gaussian random field and consider a vector of nodal points $\boldsymbol{\mathcal{Z}} = \left(\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(K)}\right)$ in the domain $\mathcal{D}_{\boldsymbol{Z}}$. The optimal linear estimation of the random field is then given by:

$$\widehat{H}(\boldsymbol{z}, \omega) = \boldsymbol{\mu}(\boldsymbol{z}) + \boldsymbol{\Sigma}_{H(\boldsymbol{z})Y}\, \boldsymbol{\Sigma}_{YY}^{-1}\, (\boldsymbol{Y} - \boldsymbol{\mu}_Y)\,, \tag{F.1}$$

where $\boldsymbol{Y} = \left(Y^{(1)}, \ldots, Y^{(K)}\right)$ is the set of correlated Gaussian variables associated to the points $\boldsymbol{\mathcal{Z}}$, and $\boldsymbol{\mu}_Y$ and $\boldsymbol{\Sigma}_{YY}$ are its mean value and covariance matrix, whose components are defined by:

$$\mu_{Y,k} = \mu\left(Y^{(k)}\right), \tag{F.2}$$

$$\Sigma_{YY,k,l} = \sigma\left(\boldsymbol{z}^{(k)}\right)\, \sigma\left(\boldsymbol{z}^{(l)}\right)\, R\left(\boldsymbol{z}^{(k)} - \boldsymbol{z}^{(l)}\right), \tag{F.3}$$

where $k = 1, \dots, K$ and $l = 1, \dots, K$. Further, $\boldsymbol{\Sigma}_{H(\boldsymbol{z})Y}$ is a correlation vector with components:

$$\Sigma_{H(\boldsymbol{z})Y,k} = \sigma\left(\boldsymbol{z}\right)\, \sigma\left(\boldsymbol{z}^{(k)}\right)\, R\left(\boldsymbol{z} - \boldsymbol{z}^{(k)}\right), \tag{F.4}$$

where $k = 1, \dots, K$.

Then, consider the eigenvalue decomposition of the covariance matrix:

$$\boldsymbol{\Sigma}_{YY} \boldsymbol{\phi}_i = \lambda_i\, \boldsymbol{\phi}_i, \tag{F.5}$$

where λ_i are the eigenvalues and $\boldsymbol{\phi}_i$ are the eigenfunctions. The eigenvalue decomposition results in K pairs of $\{\lambda_i,\, \boldsymbol{\phi}_i\}$. Note that the index i refers to the i-th largest eigenvalue of the set of K eigenvalues. Then, the EOLE approximation of the random field reads:

$$\widehat{H}\left(\boldsymbol{z}, \boldsymbol{\zeta}\left(\omega\right)\right) = \boldsymbol{\mu}\left(\boldsymbol{z}\right) + \sum_{i=1}^{n_\zeta} \frac{\zeta_i\left(\omega\right)}{\sqrt{\lambda_i}} \boldsymbol{\phi}_i^{\mathsf{T}} \boldsymbol{\Sigma}_{H(\boldsymbol{z})Y}, \tag{F.6}$$

where $\boldsymbol{\zeta}$ is a vector of n_ζ standard Gaussian variables. Commonly done in practice, only the $n_\zeta \leq K$ eigenvalues are used in the series expansion.

The quality of the EOLE approximation depends on the number of terms n_ζ in the series expansion. The variance of the approximation error in EOLE is (Sudret & Der Kiureghian 2000):

$$\operatorname{Var}\left[H\left(\boldsymbol{z}\right) - \widehat{H}\left(\boldsymbol{z}\right)\right] = \sigma^2\left(\boldsymbol{z}\right) - \sum_{i=1}^{n_\zeta} \frac{1}{\lambda_i} \left(\boldsymbol{\phi}_i^{\mathsf{T}} \boldsymbol{\Sigma}_{H(\boldsymbol{z})Y}\right)^2 = \operatorname{Var}\left[H\right] - \operatorname{Var}\left[\widehat{H}\right], \tag{F.7}$$

where σ^2 is the variance of the Gaussian RF and $n_\zeta \leq K$. Due to the fact that variances are always positive, the variance of the approximation field $\widehat{H}$ must be smaller or equal to the original RF H. Hence, EOLE always under-estimates the true variability of the Gaussian RF. Therefore, n_ζ should be chosen large enough in order to ensure a good approximation of H. In practice, $\operatorname{Var}\left[\widehat{H}\right] \geq 0.95 \cdot \operatorname{Var}\left[H\right]$ leads to good approximations with a reasonable number of variables n_ζ (see also Chapter 10).

F.4 Conditional random fields

F.4.1 Definition

Assume that in the continuous domain $\mathcal{D}_{\boldsymbol{Z}}$, there exists a set of locations $\mathfrak{Z} = \left\{\mathfrak{z}^{(1)}, \dots, \mathfrak{z}^{(N)}\right\}$ where the random field value is deterministically known, *i.e.* $\mathcal{Y}^{(i)} = H\left(\mathfrak{z}^{(i)}\right)$. These points are called *geographical evidence* or simply *observations*. Then, a *conditional* random field is a random field conditioned on these observations:

$$H_c\left(\boldsymbol{z}, \omega\right) = H\left(\boldsymbol{z}, \omega\right) \qquad \text{s.t.} \qquad H\left(\mathfrak{z}^{(i)}, \omega\right) = \mathcal{Y}^{(i)}, \qquad i = 1, \dots, N, \tag{F.8}$$

where $\mathcal{Y} = \left\{\mathcal{Y}^{(1)}, \dots, \mathcal{Y}^{(N)}\right\}$. Note that in the context of Kriging, $\mathfrak{Z}$ and $\mathcal{Y}$ are called the experimental design and the corresponding response values of the computational model.

The implicit formulation of the conditional RF makes it complex to generate realizations. Hence, in the following sections, three approaches are described to sample from the conditional RF: direct sampling, conditional EOLE, and Kriging.

F.4.2 Discretization

The conditional RF is discretized by a finite set of nodal points $\boldsymbol{\mathcal{Z}}$, as introduced before. Consider the combined vector of nodal points and observational points and denote it by $\mathbb{Y}$:

$$\mathbb{Y} = \begin{pmatrix} \boldsymbol{Y} \\ \mathcal{Y} \end{pmatrix} = \left(Y^{(1)}, \ldots, Y^{(K)}, \mathcal{Y}^{(1)}, \ldots, \mathcal{Y}^{(N)} \right)^{\mathsf{T}}. \tag{F.9}$$

Then, $\mathbb{Y}$ can be represented as a Gaussian vector $\mathbb{Y} \sim \mathcal{N}\left(\boldsymbol{\mu}_{\mathbb{Y}}, \boldsymbol{\Sigma}_{\mathbb{Y}\mathbb{Y}}\right)$ with mean and variance as:

$$\boldsymbol{\mu}_{\mathbb{Y}} = \begin{pmatrix} \boldsymbol{\mu}_{Y} \\ \boldsymbol{\mu}_{\mathcal{Y}} \end{pmatrix}, \tag{F.10}$$

$$\boldsymbol{\Sigma}_{\mathbb{Y}\mathbb{Y}} = \begin{pmatrix} \boldsymbol{\Sigma}_{YY} & \boldsymbol{\Sigma}_{Y\mathcal{Y}} \\ \boldsymbol{\Sigma}_{Y\mathcal{Y}}^{\mathsf{T}} & \boldsymbol{\Sigma}_{\mathcal{Y}\mathcal{Y}} \end{pmatrix}, \tag{F.11}$$

where $\boldsymbol{\Sigma}_{Y\mathcal{Y}}$ is the covariance matrix between the set of nodal points $\boldsymbol{\mathcal{Z}}$ and the set of observation points $\boldsymbol{\mathfrak{z}}$. Then, it can be shown that $\boldsymbol{Y}$ conditioned on the observations $\mathcal{Y}$ is obtained by:

$$\boldsymbol{Y}|\mathcal{Y} \sim \mathcal{N}\left(\boldsymbol{\mu}_{Y|\mathcal{Y}}, \boldsymbol{\Sigma}_{Y|\mathcal{Y}}\right), \tag{F.12}$$

where the mean and covariance are computed by:

$$\boldsymbol{\mu}_{Y|\mathcal{Y}} = \boldsymbol{\mu}_{Y} + \boldsymbol{\Sigma}_{Y\mathcal{Y}}\, \boldsymbol{\Sigma}_{\mathcal{Y}\mathcal{Y}}^{-1} \left(\mathcal{Y} - \boldsymbol{\mu}_{\mathcal{Y}}\right), \tag{F.13}$$

$$\boldsymbol{\Sigma}_{Y|\mathcal{Y}} = \boldsymbol{\Sigma}_{YY} - \boldsymbol{\Sigma}_{Y\mathcal{Y}}\, \boldsymbol{\Sigma}_{\mathcal{Y}\mathcal{Y}}^{-1}\, \boldsymbol{\Sigma}_{Y\mathcal{Y}}^{\mathsf{T}}. \tag{F.14}$$

F.4.3 Direct sampling

In order to generate realizations of the conditional random field, sampling the conditional distribution in Eq. (F.12) is the straightforward way. Hence, the standard Cholesky decomposition can be applied to the covariance matrix in Eq. (F.14), such that $\boldsymbol{\Sigma}_{Y|\mathcal{Y}} = LL^{\mathsf{T}}$, where L is a lower triangular matrix. This allows for a transformation of the correlated multivariate Gaussian distribution into an uncorrelated multivariate Gaussian distribution:

$$\boldsymbol{Y} = \boldsymbol{\mu}_{Y|\mathcal{Y}} + L\boldsymbol{\zeta}\left(\omega\right), \tag{F.15}$$

where $\boldsymbol{\zeta}$ is a K-dimensional vector of independent standard Gaussian variables.

The complexity of the direct sampling approach depends greatly on the Cholesky decomposition, *i.e.* on the size of the covariance matrix $\boldsymbol{\Sigma}_{Y|\mathcal{Y}}$. To bypass the problem of solving the Cholesky decomposition, Hoffmann & Ribak (1991) and Hoffmann (2009) proposed a two-step algorithm:

(i) Generate a realization of the unconditional RF $H\left(\boldsymbol{z}, \omega\right)$ ignoring the observations and denote the realization as:

$$\mathbb{Y}^{(0)} = \begin{pmatrix} \boldsymbol{Y}^{(0)} \\ \mathcal{Y}^{(0)} \end{pmatrix}. \tag{F.16}$$

(ii) Compute the corresponding realization of the conditional random field using as expression similar to Eq. (F.13):

$$\boldsymbol{Y} = \boldsymbol{Y}^{(0)} + \boldsymbol{\Sigma}_{Y\mathcal{Y}}\, \boldsymbol{\Sigma}_{\mathcal{Y}\mathcal{Y}}^{-1} \left(\mathcal{Y} - \mathcal{Y}^{(0)}\right). \tag{F.17}$$

Note that this is a deterministic transformation of the realization of the unconditional RF obtained in Step (i). This equation fulfils the condition $H\left(\mathfrak{z}^{(i)}, \omega\right) = \mathcal{Y}^{(i)}$, $i = 1, \ldots, N$, as proven by considering the case of predicting the variables $\boldsymbol{Y} = \mathcal{Y}$ at the corresponding geographical locations $\boldsymbol{\mathcal{Z}} = \mathfrak{Z}$:

$$\begin{aligned}
\mathcal{Y} &= \mathcal{Y}^{(0)} + \boldsymbol{\Sigma}_{\mathcal{Y}\mathcal{Y}} \boldsymbol{\Sigma}_{\mathcal{Y}\mathcal{Y}}^{-1} \left(\mathcal{Y} - \mathcal{Y}^{(0)}\right), && \text{(F.18)} \\
\mathcal{Y} - \mathcal{Y}^{(0)} &= \boldsymbol{I}_N \left(\mathcal{Y} - \mathcal{Y}^{(0)}\right), && \text{(F.19)} \\
\mathcal{Y} - \mathcal{Y}^{(0)} &= \mathcal{Y} - \mathcal{Y}^{(0)}, && \text{(F.20)}
\end{aligned}$$

where $\boldsymbol{I}_N$ is the identity matrix of size $N \times N$.

This two-step algorithm is particularly useful when realizations of the unconditional RF are available. In fact, the conditional realization is obtained by a simple post-processing scheme of the unconditional realization of the RF.

F.4.4 Conditional EOLE

Conditional EOLE is a combination of the EOLE approach in Section F.3 and the two-step algorithm in Section F.4.3 (Schöbi & Sudret 2015a). In particular, the realization of the unconditional RF in Step (i) of the two-step algorithm is computed by EOLE. Then, it is possible to compute realizations of the conditional random field efficiently. Moreover, realizations of the conditional RF can be generated conveniently as a function of n_ζ random variables and a set of observations $\{\mathfrak{Z}, \mathcal{Y}\}$.

When considering a random field as input of a computational model, the discretization procedure allows to reduce the input to n_ζ independent standard normal variables, which are simply collected with other input parameters of the problem into a single random vector. This, in turn, allows for meta-models to be applied in a straightforward manner, which allows for efficient estimations of the QoI. For practical examples in the context of geotechnical engineering, see also Chapter 10.

F.4.5 Kriging

Based on the mean value, variance, autocorrelation function of a Gaussian RF and a set of observations $\{\mathfrak{Z}, \mathcal{Y}\}$, Kriging is capable of predicting the Gaussian variable at any location $\boldsymbol{z} \in \mathcal{D}_{\boldsymbol{Z}}$. This allows the simulation of RF realizations for a given set of nodal points $\boldsymbol{\mathcal{Z}}$ by an iterative algorithm (Kameda & Morikawa 1992; Hoshiya 1995):

(i) Use the observations as experimental design for the calibration of a Kriging model as in Eqs. (3.27) and (3.28).

(ii) Choose one of the remaining unknown nodal points $\boldsymbol{\mathcal{Z}}$ and generate its value by random sampling the random field at that location, *i.e.* by random sampling of the Gaussian variable $y^* \sim \mathcal{N}\left(\mu_Y\left(\boldsymbol{z}^*\right), \sigma_Y\left(\boldsymbol{z}^*\right)\right)$, where $\mu_Y\left(\boldsymbol{z}^*\right)$ and $\sigma_Y\left(\boldsymbol{z}^*\right)$ are obtained through Eqs. (3.31) and (3.32), respectively.

(iii) Consider the samples value at $\boldsymbol{z}^*$ as a given observation and add it to the experimental design, *i.e.* $\mathfrak{Z} \leftarrow \{\mathfrak{Z}, \boldsymbol{z}^*\}$ and $\mathcal{Y} \leftarrow \{\mathcal{Y}, y^*\}$. Remove the sampled point from the set of nodal points, *i.e.* $\boldsymbol{\mathcal{Z}} \leftarrow \boldsymbol{\mathcal{Z}} \backslash \boldsymbol{z}^*$.

(iv) Iterate steps (i) to (iii) until the set of nodal points $\boldsymbol{\mathcal{Z}}$ is empty, *i.e.* till all samples have been sampled.

The main advantage of this algorithm is its simplicity. The algorithm mainly consists of a series of Kriging models with increasing size of the experimental design. Thus, the generation of conditional RF realizations is robust but inefficient when comparing to the conditional EOLE approach.

With respect to the application of meta-models, the number of input variables would be equal to the number of simulated points, *i.e.* the number of nodal points. When the set of nodal points is large, it becomes inefficient to represent the uncertainty at each nodal point by a separate random variable, comparable to the direct sampling approach in Eq. (F.15). That is why the adaptation of series expansion approaches, such as conditional EOLE, is more suitable for conditional random fields.

F.5 Toy example

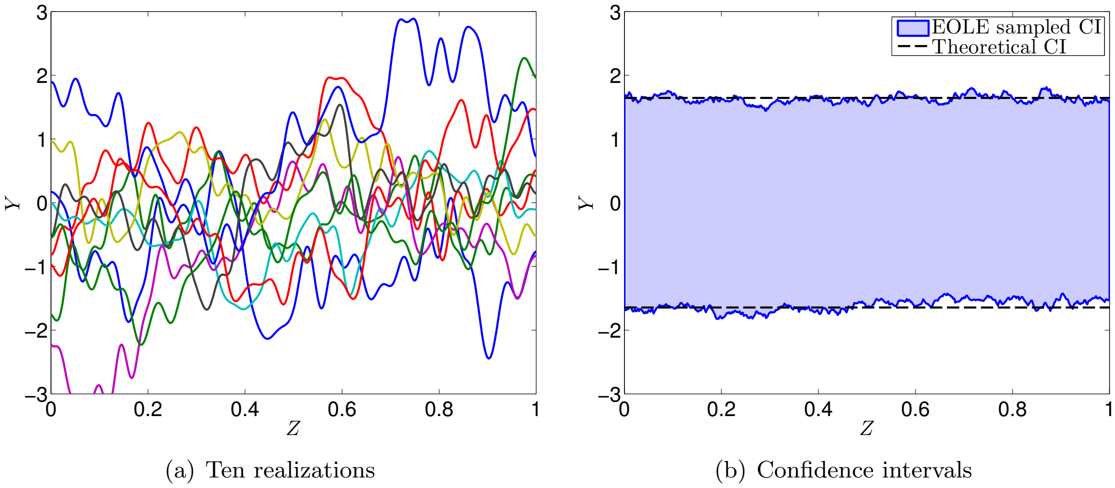
Consider a one-dimensional geographical domain $\mathcal{D}_Z = [0,\, 1]$. Assume that the spatial variability can be modelled by a one-dimensional Gaussian random field with mean value $\mu = 0$, variance $\sigma^2 = 1$, and an exponential autocorrelation function with a correlation length of $\rho = 0.2$ (see also Eq. (C.2)). Additionally, there are observations of the random field at three locations:

$$\left\{ \begin{array}{l} \boldsymbol{\mathfrak{Z}} = \left\{\mathfrak{z}^{(1)} = 0.22,\, \mathfrak{z}^{(2)} = 0.46,\, \mathfrak{z}^{(3)} = 0.92\right\} \\ \mathcal{Y} = \left\{\mathcal{Y}^{(1)} = 0.00,\, \mathcal{Y}^{(2)} = -0.15,\, \mathcal{Y}^{(3)} = 0.25\right\} \end{array} \right. . \tag{F.21}$$

Ignoring the observations, realizations of the unconditional RF are generated using the EOLE algorithm. In particular, the set of nodal points $\boldsymbol{\mathcal{Z}}$ consists of $K = 1000$ equally spaced points in $[0, 1]$. Ten realizations of the resulting discretized RF are shown in Figure F.1(a). Note that the number of random variables $\boldsymbol{\zeta}$ is set to $n_\zeta = 51$ in order to recover 98 % of the RF variance. The difference of response value Y between different realizations of the RF is apparent.

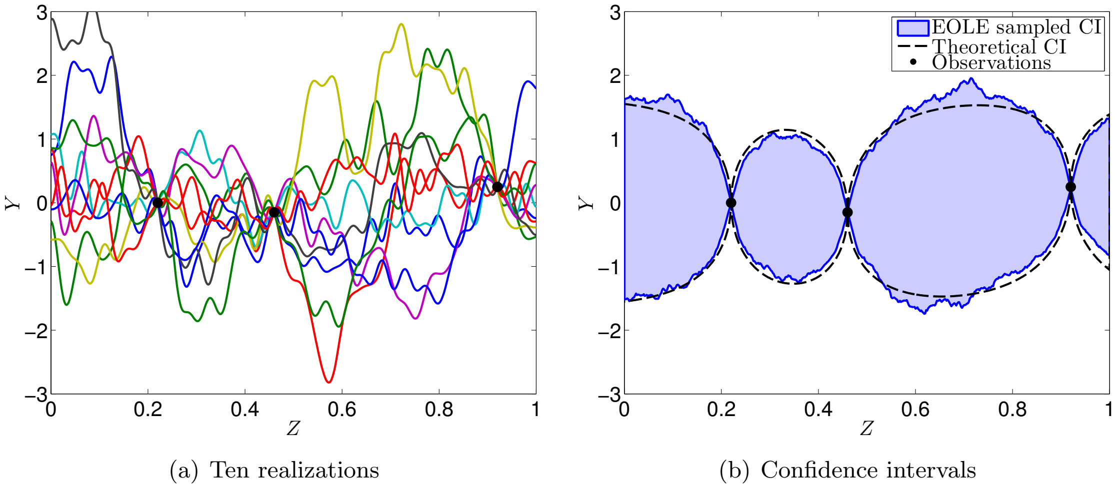
Figure F.1(b) illustrates the confidence intervals (CI) of the realizations of the EOLE algorithm to the theoretical values. The CI is set at 90 %. Note that the CI of the EOLE algorithm are obtained by computing the empirical quantiles of 1000 realizations of the unconditional RF. The overlapping of the CIs shows the ability of the EOLE algorithm to recover the variability of the original unconditional random field.

Figure F.2 shows realizations and CI for the case of the conditional RF analogue to Figure F.1 for the unconditional RF. Note that the black bullet points represent the observations in Eq. (F.21). The realizations as well as the CI interpolate the observation points $\mathfrak{z}^{(i)}$, where the returned value is deterministically known. Additionally, the conditional EOLE recovers the variance of the conditional RF, as seen by the overlapping CI with the theoretical values.



(a) Ten realizations (b) Confidence intervals

Figure F.1: Toy function – unconditional random field – simulated by EOLE.



(a) Ten realizations (b) Confidence intervals

Figure F.2: Toy function – random field on observations in Eq. (F.21) – simulated by conditional EOLE.

Bibliography

Abbiati, G., R. Schöbi, B. Sudret, & B. Stojadinovic (2017). “Structural reliability analysis using deterministic hybrid simulations and adaptive Kriging metamodeling”. In: *16th World Conference of Earthquake (16WCEE)*. Santiago, Chile.

Ahlfeld, R., B. Belkouchi, & F. Montomoli (2016). “SAMBA: sparse approximation of moment-based arbitrary polynomial chaos”. In: *J. Comput. Phys.* 320, pp. 1–16.

Alibrandi, U. & C. G. Koh (2015). “First-Order Reliability Method for structural reliability analysis in the presence of random and interval variables”. In: *ASCE-ASME J. Risk Uncertain. Eng. Sys., Part B Mech. Engrg.* 1 (4), p. 041006.

Allen, D. (1971). *The prediction sum of squares as a criterion for selecting prediction variables.* Tech. rep. 23. Dept. of Statistics, University of Kentucky.

Alvarez, D. A. & J. E. Hurtado (2014). “An efficient method for the estimation of structural reliability intervals with random sets, dependence modeling and uncertain inputs”. In: *Comput. Struct.* 142, pp. 54–63.

Alvarez, D. A., J. E. Hurtado, & F. Uribe (2014). “Estimation of the lower and upper probabilities of failure using random sets and subset simulation”. In: *Vulnerability, Uncertain. Risk.* Ed. by M. Beer, S.-K. Au, & J. W. Hall. Liverpool, UK: ASCE, pp. 905–914.

An, J. & A. Owen (2001). “Quasi-regression”. In: *J. Complex.* 17 (4), pp. 588–607.

Archer, G., A. Saltelli, & I. Sobol’ (1997). “Sensitivity measures, ANOVA-like techniques and the use of bootstrap”. In: *J. Stat. Comput. Simul.* 58, pp. 99–120.

Arnaud, A., J. Bect, M. Couplet, A. Pasanisi, & E. Vazquez (2010). “Evaluation d’un risque d’inondation fluviale par planification séquentielle d’expériences”. In: *42èmes Journées de Statistique.* Marseille, France, pp. 1–6.

Au, S.-K. (2016). “On MCMC algorithm for Subset Simulation”. In: *Prob. Eng. Mech.* 43, pp. 117–120.

Au, S.-K. & J. L. Beck (2001). “Estimation of small failure probabilities in high dimensions by subset simulation”. In: *Prob. Eng. Mech.* 16 (4), pp. 263–277.

Ayyub, B. M. (2001). *Elicitation of Expert Opinions for Uncertainty and Risks.* CRC Press.

Ayyub, B. M. & G. J. Klir (2006). *Uncertainty Modeling and Analysis in Engineering and the Sciences.* Chapman & Hall.

Bachoc, F. (2013). “Cross Validation and Maximum Likelihood estimations of hyper-parameters of Gaussian processes with model misspecifications”. In: *Comput. Stat. Data An.* 66, pp. 55–69.

Bachoc, F. (2014). “Asymptotic analysis of the role of spatial sampling for hyper-parameter estimation of Gaussian processes”. In: *J. Multivariate Anal.* 125, pp. 1–35.

Bachoc, F., G. Bois, J. Garnier, & J.-M. Martinez (2014). “Calibration and improved prediction of computer models by universal Kriging”. In: *Nucl. Sci. Eng.* 176, pp. 91–97.

Balesdent, M., J. Morio, & L. Brevault (2016). “Rare event probability estimation in the presence of epistemic uncertainty on input probability distribution parameters”. In: *Meth. Comput. Appl. Probab.* 18 (1), pp. 197–216.

Balu, A. & B. Rao (2013). “High dimensional model representation for structural reliability bounds estimation under mixed uncertainties”. In: *Int. J. Struct. Eng.* 4 (3), pp. 251–272.

Barthe, D., A. Deraemaeker, P. Ladevèze, & G. Puel (2003). “On a theory of the quantification of the lack of knowledge (LOK) in structural computation”. In: *IMAC XXI Conf. Expo. Struct. Dyn. - Innov. Meas. Technol.*

Bect, J., D. Ginsbourger, L. Li, V. Picheny, & E. Vazquez (2012). “Sequential design of computer experiments for the estimation of a probability of failure.” In: *Stat. Comput.* 22 (3), pp. 773–793.

Beer, M., S. Ferson, & V. Kreinovich (2016). “Do we have compatible concepts of epistemic uncertainty?” In: *Proc. 6th Asian-Pacific Symp. Struct. Reliab. (APSSRA’2016), Shanghai.*

Beer, M. (2009). “Engineering quantification of inconsistent information”. In: *Int. J. Reliab. Safety* 3 (1/2/3), pp. 174–200.

Beer, M., S. Ferson, & V. Kreinovich (2013). “Imprecise probabilities in engineering analyses”. In: *Mech. Syst. Signal Process.* 37, pp. 4–29.

Beer, M., Z. Mingqiang, Q. S. Tong, & H. A. Jensen (2011). “Reliability analysis of offshore structures with imprecise corrosion effects”. In: *Comp. Stoch. Mech.* 1, pp. 66–79.

Beer, M., Y. Zhang, S. Tong, & K. Kwang (2012). “Reliability analysis with scarce information: comparing alternative approaches in a geotechnical engineering context”. In: *Structural Safety* 41, pp. 1–10.

Ben-Haim, Y. (2006). *Info-gap decision theory: decisions under severe uncertainty.* 2nd. Academic Press, London.

Beres, D. & D. Hawkins (2001). “Plackett-Burman techniques for sensitivity analysis of many-parametered models”. In: *Ecol. Mod.* 141 (1-3), pp. 171–183.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* New York: Springer-Verlag.

Berveiller, M., B. Sudret, & M. Lemaire (2006a). “A stochastic finite element procedure for moment and reliability analysis”. In: *Eur. J. Comput. Mech.* 15 (7-8), pp. 825–866.

Berveiller, M., B. Sudret, & M. Lemaire (2006b). “Stochastic finite elements: a non intrusive approach by regression”. In: *Eur. J. Comput. Mech.* 15 (1-3), pp. 81–92.

Beyer, H.-G. & B. Sendhoff (2007). “Robust optimization – a comprehensive survey”. In: *Comput. Methods Appl. Mech. Eng.* 196 (2007), pp. 3190–3218.

Bichon, B., M. Eldred, L. Swiler, S. Mahadevan, & J. McFarland (2008). “Efficient global reliability analysis for nonlinear implicit performance functions”. In: *AIAA Journal* 46 (10), pp. 2459–2468.

Bichon, B., J. McFarland, & S. Mahadevan (2011). “Efficient surrogate models for reliability analysis of systems with multiple failure modes”. In: *Reliab. Eng. Sys. Safety* 96 (10), pp. 1386–1395.

Bishop, A. (1955). “The use of slip circle in the stability analysis of slopes”. In: *Géotechnique* 5 (1), pp. 7–17.

Al-Bittar, T. & A.-H. Soubra (2013). “Bearing capacity of strip footings on spatially random soils using sparse polynomial chaos expansion”. In: *Int. J. Num. Anal. Meth. Geomech.* 37, pp. 2039–2060.

Al-Bittar, T. & A.-H. Soubra (2014). "Probabilistic analysis of strip footings resting on spatially varying soils and subjected to vertical or inclined loads". In: *J. Geotech. Geoenviron. Eng.* 140 (4), p. 04013043.

Blatman, G. (2009). "Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis". PhD thesis. France: Université Blaise Pascal, Clermont-Ferrand.

Blatman, G. & B. Sudret (2008). "Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach". In: *Comptes Rendus Mécanique* 336 (6), pp. 518–523.

Blatman, G. & B. Sudret (2010a). "An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis". In: *Prob. Eng. Mech.* 25 (2), pp. 183–197.

Blatman, G. & B. Sudret (2010b). "Efficient computation of global sensitivity indices using sparse polynomial chaos expansions". In: *Reliab. Eng. Sys. Safety* 95, pp. 1216–1229.

Blatman, G. & B. Sudret (2011). "Adaptive sparse polynomial chaos expansion based on Least Angle Regression". In: *J. Comput. Phys* 230 (6), pp. 2345–2367.

Bourinet, J. M., F. Deheeger, & M. Lemaire (2011). "Assessing small failure probabilities by combined subset simulation and support vector machines". In: *Structural Safety* 33 (6), pp. 180–191.

Box, G. E. P. & N. R. Draper (1987). *Empirical model building and response surfaces.* New York, NY: John Wiley & Sons.

Breitung, K. (1989). "Asymptotic approximations for probability integrals". In: *Prob. Eng. Mech.* 4 (4), pp. 187–190.

Byrd, R. H., M. E. Hribar, & J. Nocedal (1999). "An interior point algorithm for large-scale nonlinear programming". In: *SIAM J. Optim.* 9 (4), pp. 877–900.

Cacuci, D. (2003). *Sensitivity and uncertainty analysis: theory.* Vol. 1. Chapman & Hall/CRC, Boca Raton.

Cadini, F., F. Santos, & E. Zio (2014). "An improved adaptive Kriging-based importance technique for sampling multiple failure regions of low probability". In: *Reliab. Eng. Sys. Safety* 131, pp. 109–117.

Cartwright, H. (2015). *Artificial neural networks.* 2nd. New York: Humana Press.

Chebyshev, P. (1874). "Sur les valeurs limites des intégrales". In: *J. Math. Pures Appl., Ser 2* 19, pp. 157–160.

Chen, S., K.-R. G. Reyes, M. K. Gupta, M. C. McAlpine, & W. B. Powell (2015). "Optimal learning in experimental design using the knowledge gradient policy with application to characterizing nanoemulsion stability". en. In: *SIAM/ASA J. Uncer. Quant.* 3 (1), pp. 320–345.

Chkifa, A., A. Cohen, G. Migliorati, F. Nobile, & R. Tempone (2013). *Discrete least squares polynomial approximation with random evaluations - applications to parametric and stochastic elliptic PDEs.* Tech. rep. Lausanne, Switzerland: MATHICSE, EPFL.

Cho, H., D. Venturi, & G. Karniadakis (2013). "Karhunen-Loève expansion for multi-correlated stochastic processes". In: *Prob. Eng. Mech.* 34, pp. 157–167.

Choi, S., R. Grandhi, R. Canfield, & C. Pettit (2004). "Polynomial chaos expansion with Latin Hypercube sampling for estimating response variability". In: *AIAA Journal* 45, pp. 1191–1198.

Chowdhary, K. & P. Dupuis (2013). "Distinguishing and integrating aleatoric and epistemic variation in uncertainty quantification". In: *ESAIM Math. Model. Numer. Anal.* 47 (3), pp. 635–662.

Clarke, S., J. Griebsch, & T. Simpson (2003). “Analysis of support vector regression for approximation of complex engineering analyses”. In: *Proc. DETC’03, Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Chicago.*

COMSOL AB (2013). *Geomechanics Module User’s Guide, v. 4.4.* Tech. rep. Stockholm, Sweden: COMSOL AB.

Cotter, S. (1979). “A screening design for factorial experiments with interactions”. In: *Biometrika* 66 (2), pp. 317–320.

Crespo, L. G. & S. P. Kenny (2015). “Special Edition on Uncertainty Quantification of the AIAA Journal of Aerospace Computing, Information, and Communication”. In: *Journal of Aerospace Information Systems* 12 (1), pp. 9–9.

Crespo, L. G., S. P. Kenny, & D. P. Giesy (2013). *The NASA Langley multidisciplinary uncertainty quantification challenge.* Tech. rep. National Institute of Aerospace, Langley Research center, pp. 1–9.

Cukier, R. (1973). “Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I theory”. In: *J. Chem. Phys.* 59 (8), pp. 3873–3878.

Cukier, R., H. Levine, & K. Shuler (1978). “Nonlinear sensitivity analysis of multiparameter model systems”. In: *J. Comput. Phys.* 26, pp. 1–42.

Dai, H., H. Zhang, K. Rasmussen, & W. Wang (2014). “Wavelet density-based adaptive importance sampling method”. In: *Structural Safety* 52, pp. 161–169.

Dani, V., T. P. Hayes, & S. M. Kakade (2008). “Stochastic linear optimization under bandit feedback”. In: *The 21st Annual Conference on Learning Theory (COLT 2008).*

De Angelis, M., E. Patelli, & M. Beer (2014). “Line Sampling for assessing structural reliability with imprecise failure probabilities”. In: *Vulnerability, Uncertain. Risk.* Ed. by M. Beer, S.-K. Au, & J. W. Hall. Liverpool, UK: ASCE, pp. 915–924.

De Angelis, M., E. Patelli, & M. Beer (2015). “Advanced Line Sampling for efficient robust reliability analysis”. In: *Structural Safety* 52, pp. 170–182.

De Cooman, G. (1997). “Possibility theory part I: measure- and integral-theoretic ground-work; part II: conditional possibility; part III: possibilistic independence”. In: *Int. J. Gen. Syst.* 25 (4), pp. 291–371.

De Cooman, G., D. Ruan, & E. Kerre (1995). “Foundations and applications of possibility theory”. In: *Proc. Fapt. Ghent, Belgium.* River Edge, NJ: World Scientific.

de Rocquigny, E. (2006a). “La maîtrise des incertitudes dans un contexte industriel : 1^e partie – une approche méthodologique globale basée sur des exemples”. In: *J. Soc. Française Stat.* 147 (3), pp. 33–71.

de Rocquigny, E. (2006b). “La maîtrise des incertitudes dans un contexte industriel : 2^e partie – revue des méthodes de modélisation statistique, physique et numérique”. In: *J. Soc. Française Stat.* 147 (3), pp. 73–106.

Deman, G., K. Konakli, B. Sudret, J. Kerrou, P. Perrochet, & H. Benabderrahmane (2016). “Using sparse polynomial chaos expansions for the global sensitivity analysis of groundwater lifetime expectancy in a multi-layered hydrogeological model”. In: *Reliab. Eng. Sys. Safety* 147, pp. 156–169.

Dempster, A. P. (1967). “Upper and lower probabilities induced by multivalued mapping”. In: *Ann. Math. Stat.* 38 (2), pp. 325–339.

Deng, W., X. Yang, L. Zou, M. Wang, Y. Liu, & Y. Li (2013). “An improved self-adaptive differential evolution algorithm and its application”. In: *Chemo. Intel. Lab.* 128, pp. 66–76.

Deodatis, G. (1990). “Bounds on response variability of stochastic finite element systems: effect of statistical dependence”. In: *Prob. Eng. Mech.* 5 (2), pp. 80–89.

Deodatis, G. (1991). "The weighted integral method I : stochastic stiffness matrix". In: *J. Eng. Mech.* 117 (8), pp. 1851–1864.

Deodatis, G. & M. Shinozuka (1991). "The weighted integral method II : response variability and reliability". In: *J. Eng. Mech.* 117 (8), pp. 1865–1877.

Der Kiureghian, A. & M. de Stefano (1990). *An efficient algorithm for second-order reliability analysis.* Tech. rep. UCB/SEMM-90/20. Dept of Civil and Environmental Engineering, University of California, Berkeley.

Der Kiureghian, A. & J. Ke (1988). "The stochastic finite element method in structural reliability". In: *Prob. Eng. Mech.* 3 (2), pp. 83–91.

Destercke, S., D. Dubois, & E. Chojnacki (2008a). "Unifying practical uncertainty representations. I: generalized p-boxes". In: *Int. J. Approx. Reason.* 49 (3), pp. 649–663.

Destercke, S., D. Dubois, & E. Chojnacki (2008b). "Unifying practical uncertainty representations. II: clouds". In: *Int. J. Approx. Reason.* 49 (3), pp. 664–677.

Dey, S., T. Mukhopadhyay, H. H. Khodaparast, & S. Adhikari (2016). "Fuzzy uncertainty propagation in composites using Gram–Schmidt polynomial chaos expansion". In: *Appl. Math. Model.* 40 (7), pp. 4412–4428.

Ditlevsen, O. (1996). "Dimension reduction and discretization in stochastic problems by regression method". In: *Mathematical models for structural reliability analysis, CRC Mathematical Modelling Series.* Ed. by F. Casciati & B. Roberts. Chap. 2, pp. 51–138.

Ditlevsen, O. & H. Madsen (1996). *Structural reliability methods.* J. Wiley and Sons, Chichester.

Dong, W. & H. Shah (1987). "Vertex method for computing functions of fuzzy variables". In: *Fuzzy Sets Syst.* 24 (1), pp. 65–78.

Doostan, A. & H. Owhadi (2011). "A non-adapted sparse approximation of PDEs with stochastic inputs". In: *J. Comput. Phys.* 230 (8), pp. 3015–3034.

Doostan, A., A. Validi, & G. Iaccarino (2013). "Non-intrusive low-rank separated approximation of high-dimensional stochastic models". In: *Comput. Methods Appl. Mech. Eng.* 263, pp. 42–55.

Dubois, D., H. Fargier, & J. Fortin (2004). "A generalized vertex method for computing with fuzzy intervals". In: *2004 IEEE Int. Conf. Fuzzy Syst. (IEEE Cat. No.04CH37542).* Vol. 1. IEEE, pp. 541–546.

Dubois, D. & H. Prade (1992). "On the combination of evidence in various mathematical frameworks". In: *Reliability data collection and analysis.* Ed. by J. Flamm & T. Luisi. Boston, MA: Kluwer, pp. 213–241.

Dubois, D. & H. Prade (1988). *Possibility theory: an approach to computerized processing of uncertainty.* New York: Plenum.

Dubois, D. & H. Prade (1998). *Handbook of defeasible reasoning and uncertainty management systems.* Ed. by D. Grabbay & P. Smets. Boston: Kluwer, pp. 169–226.

Dubois, D. & H. Prade (2005). "Interval-valued fuzzy sets, possibility theory and imprecise probability". In: *EUSFLAT - LFA 2005*, pp. 314–319.

Dubourg, V. (2011). "Adaptive surrogate models for reliability analysis and reliability-based design optimization". PhD thesis. Université Blaise Pascal, Clermont-Ferrand, France.

Dubourg, V. & B. Sudret (2014). "Meta-model-based importance sampling for reliability sensitivity analysis". In: *Structural Safety* 49, pp. 27–36.

Dubourg, V., B. Sudret, & J. M. Bourinet (2011). "Reliability-based design optimization using Kriging and subset simulation". In: *Struct. Multidisc. Optim.* 44 (5), pp. 673–690.

Dubourg, V., B. Sudret, & F. Deheeger (2013). "Metamodel-based importance sampling for structural reliability analysis". In: *Prob. Eng. Mech.* 33, pp. 47–57.

Dubrule, O. (1983). “Cross-validation of Kriging in a unique neighbourhood”. In: *Math. Geol.* 15 (6), pp. 687–698.

Duncan, J. (1996). “State of the art: limit equilibrium and finite-element analysis of slopes”. In: *J. Geotech. Engng.* 122 (7), pp. 577–596.

Duncan, J. & S. G. Wright (2005). *Soil Strength and Slope Stability.* Wiley.

Durrande, N., D. Ginsbourger, & O. Roustant (2012). “Additive covariance kernels for high-dimensional Gaussian process modeling”. In: *Ann. Fac. Sci. Toulouse Tome* 21 (3), pp. 481–499.

Durrande, N., D. Ginsbourger, & O. Roustant (2013). “ANOVA kernels and RKHS of zero mean function for model-based sensitivity analysis”. In: *J. Multivar. Anal.* 115, pp. 57–67.

Duvenaud, D., H. Nickisch, & C. Rasmussen (2012). “Additive Gaussian processes”. In: *Adv. Neur. Infor. Proc. Sys.* 25, pp. 1–8.

Echard, B., N. Gayton, & M. Lemaire (2011). “AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation”. In: *Structural Safety* 33 (2), pp. 145–154.

Echard, B., N. Gayton, M. Lemaire, & N. Relun (2013). “A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models”. In: *Reliab. Eng. Sys. Safety* 111, pp. 232–240.

Efron, B., T. Hastie, I. Johnstone, & R. Tibshirani (2004). “Least angle regression”. In: *Ann. Stat.* 32 (2), pp. 407–499.

Efron, B. & C. Stein (1981). “The Jackknife estimate of variance”. In: *Annals Statist.* 9 (3), pp. 586–596.

Eldred, M. (2009). “Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design”. In: *50th AIAA/ASME/ASCE/AHS/ASC Struct. Dyn. Mater. Conf.*

Eldred, M. & L. P. Swiler (2009). *Efficient algorithms for mixed aleatory-epistemic uncertainty quantification with application to radiation-hardened electronics part I: algorithms and benchmark results.* Tech. rep. SAND2009-5805. Sandia National Laboratories.

Eldred, M., C. Webster, & P. Constantine (2008). “Design under uncertainty employing stochastic expansion methods”. In: *12th AIAA/ISSMO Multidiscip. Anal. Optim. Conf.* Victoria, British Columbia, Canada.

Elishakoff, I., Y. Ren, & M. Shinozuka (1994). “Conditional simulation of non-Gaussian random fields”. In: *Eng. Struct.* 16 (7), pp. 558–563.

Engelund, S. & R. Rackwitz (1993). “A benchmark study on importance sampling techniques in structural reliability”. In: *Structural Safety* 12 (4), pp. 225–276.

Faure, H. (1982). “Discrépance de suites associées à un système de numération (en dimension s)”. In: *Acta Arith.* 41, pp. 337–351.

Fellenius, W. (1936). “Calculation of the stability of earth dams”. In: *Proc. 2nd Congr. Large Dams.* Vol. 4. Washington DC.

Feller, W. (1948). “On the Kolmogorov-Smirnov limit theorem for empirical distributions”. In: *Ann. Math. Stat.* 19, pp. 177–189.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications.* second. Vol. 2. New York: John Wiley and Sons.

Ferson, S. & L. R. Ginzburg (1996). “Different methods are needed to propagate ignorance and variability”. In: *Reliab. Eng. Sys. Safety* 54 (2-3), pp. 133–144.

Ferson, S. & W. Tucker (2008). “Probability boxes as info-gap models”. In: *NAFIPS 2008 - 2008 Annu. Meet. North Am. Fuzzy Inf. Process. Soc.* IEEE, pp. 1–6.

Ferson, S. & J. G. Hajagos (2004). “Arithmetic with uncertain numbers: rigorous and (often) best possible answers”. In: *Reliab. Eng. Sys. Safety* 85 (1-3), pp. 135–152.

Ferson, S., V. Kreinovich, L. Ginzburg, D. S. Myers, & K. Sentz (2003). *Constructing Probability Boxes and Dempster-Shafer Structures.* Tech. rep. Sandia National Laboratories.

Ferson, S. & W. Troy Tucker (2006). “Sensitivity analysis using probability bounding”. In: *Reliab. Eng. Sys. Safety* 91 (10-11), pp. 1435–1442.

Forrester, A., A. Sóbester, & A. Keane (2008). *Engineering design via surrogate modelling.* Wiley.

Fu, G., D. Butler, S.-T. Khu, & S. Sun (2011). “Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory”. In: *Water Resour. Res.* 47, W02534.

Fuchs, M. & A. Neumaier (2009). “Potential based clouds in robust design optimization”. In: *J. Stat. Theory Pract.* 3 (1), pp. 225–238.

Gautschi, W. (2004). *Orthogonal Polynomials: Computation and Approximation.* Numerical Mathematics and Scientific Computation. Oxford University Press.

Geisser, S. (1975). “The predictive sample reuse method with applications”. In: *J. Amer. Stat. Assoc.* 70, pp. 320–328.

Gelman, A., J. Carlin, H. Stern, & D. Rubin (2009). *Bayesian Data Analysis.* second. Chapman & Hall/CRC.

Ghanem, R. & P. Spanos (2003). *Stochastic Finite Elements: A Spectral Approach.* 2nd. Courier Dover Publications, Mineola.

Ghiocel, D. & R. Ghanem (2002). “Stochastic finite element analysis of seismic soil-structure interaction”. In: *J. Eng. Mech.* 128, pp. 66–77.

Ginsbourger, D., N. Durrande, & O. Roustant (2013). *Kernels and designs for modelling invariant functions: From group invariance to additivity.* Tech. rep. hal-00731657. INRIA.

Ginsbourger, D., B. Rosspopoff, G. Pirot, N. Durrande, & R. Renard (2013). “Distance-based Kriging relying on proxy simulation for inverse conditioning”. In: *Adv. Water Resources* 52, pp. 275–291.

Ginsbourger, D., O. Roustant, & N. Durrande (2013). “Invariances of random fields paths, with applications in Gaussian process regression”. In: *arXiv:1308.1359.*

Goldberg, G. H. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley Professional.

Griffiths, D. (2015). *Slope stability analysis by finite elements - a guide to the use of program slope64.* Geomechanics Research Center, Colorado School of Mines. Boulder, Colorado.

Griffiths, D. & G. A. Fenton (2004). “Probabilistic slope stability analysis by finite elements”. en. In: *J. Geotech. Geoenviron. Eng.* 130 (5), pp. 507–518.

Griffiths, D., J. Huang, & G. A. Fenton (2009). “Influence of spatial variability on slope reliability using 2-D random fields”. en. In: *J. Geotech. Geoenviron. Eng.* 135 (10), pp. 1367–1378.

Griffiths, D., J. Huang, & G. A. Fenton (2011). “Probabilistic infinite slope analysis”. In: *Comput. Geotech.* 38 (4), pp. 577–584.

Griffiths, D. & P. a. Lane (1999). “Slope stability analysis by finite elements”. In: *Geotechnique* 49 (3), pp. 387–403.

Grosof, B. N. (1986). “An inequality paradigm for probabilistic knowledge: the logic of conditional probability intervals”. In: *Uncertainty in Artificial Intelligence.* Ed. by L. N. Kanal & J. F. Lemmer. Amsterdam: Elsevier Science Publisher.

Guan, J. W. & D. A. Bell (1991). *Evidence theory and its applications, vols 1 and 2.* New York, North-Holland: Elsevier Science Inc.

Gunn, S. (1998). *Support vector machines for classification and regression.* Tech. rep. ISIS-1-98. UK: Dpt. of Electronics and Computer Science, University of Southampton.

Guo, J. & X. Du (2012). “Sensitivity analysis with mixture of epistemic and aleatory uncertainties”. en. In: *AIAA J.* 45 (9), pp. 2337–2349.

Hall, J. (2006). “Uncertainty-based sensitivity indices for imprecise probability distributions”. In: *Reliab. Eng. Sys. Safety* 91 (10-11), pp. 1443–1451.

Halton, J. H. (1960). “On the efficiency of certain quasi-random sequences of point in evaluating multi-dimensional integrals”. In: *Numer. Math.* 2, pp. 84–90.

Hammersley, J. & D. Handscomb (1964). *Monte Carlo methods.* London, New York: Methuen & Co Ltd.

Harper, W. V. & S. K. Gupta (1983). *Sensitivity/uncertainty analysis of a borehole scenario comparing Latin hypercube sampling and deterministic sensitivity approaches.* Tech. rep. No. BMI/ONWI-516. Columbus, OH, USA: Battelle Memorial Inst. - Office of Nuclear Waste Isolation.

Hasofer, A.-M. & N.-C. Lind (1974). “Exact and invariant second moment code format”. In: *J. Eng. Mech.* 100 (1), pp. 111–121.

Hastie, T., R. Tibshirani, & J. Friedman (2001). *The elements of statistical learning: data mining, inference and prediction.* Springer, New York.

He, Y., M. Mirzargar, & R. M. Kirby (2015). “Mixed aleatory and epistemic uncertainty quantification using fuzzy set theory”. In: *Int. J. Approx. Reason.* 66, pp. 1–15.

Helton, J., J. Johnson, & W. Oberkampf (2004). “An exploration of alternative approaches to the representation of uncertainty in model predictions”. In: *Reliab. Eng. Sys. Safety* 85 (1-3), pp. 39–71.

Helton, J., J. Johnson, W. Oberkampf, & C. Sallaberry (2006). “Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty”. In: *Reliab. Eng. Sys. Safety* 91 (10-11), pp. 1414–1434.

Helton, J., J. Johnson, C. Sallaberry, & C. Storlie (2006). “Survey of sampling-based methods for uncertainty and sensitivity analysis”. In: *Reliab. Eng. Sys. Safety* 91 (10-11), pp. 1175–1209.

Helton, J. & W. Oberkampf (2004). “Alternative representations of epistemic uncertainty”. In: *Reliab. Eng. Sys. Safety* 85 (1-3), pp. 1–10.

Hoffmann, Y. (2009). “Gaussian fields and constrained simulation of the large-scale structure”. In: *Data Anal. Cosmol. Lect. Notes Phys.* Berlin: Springer Verlag, pp. 565–583.

Hoffmann, Y. & E. Ribak (1991). “Constrained realizations of Gaussian fields: a simple algorithm”. In: *Astrophys. J.* 380, pp. L5–L8.

Hohenbichler, M., S. Gollwitzer, W. Kruse, & R. Rackwitz (1987). “New light on first- and second order reliability methods”. In: *Structural Safety* 4, pp. 267–284.

Homma, T. & A. Saltelli (1996). “Importance measures in global sensitivity analysis of non linear models”. In: *Reliab. Eng. Sys. Safety* 52, pp. 1–17.

Hoshiya, M. (1995). “Kriging and conditional simulation of Gaussian field”. en. In: *J. Eng. Mech.* 121 (2), pp. 181–186.

Hu, Z. & X. Du (2015). “A random field approach to reliability analysis with random and interval variables”. In: *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B Mech. Eng.* 1 (4), p. 041005.

Hurtado, J. E. & D. A. Alvarez (2001). “Neural-network-based reliability analysis: a comparative study”. In: *Comput. Method. Appl. M.* 191 (1-2), pp. 113–132.

Hurtado, J. E. (2013). “Assessment of reliability intervals under input distributions with uncertain parameters”. In: *Prob. Eng. Mech.* 32, pp. 80–92.

Isukapalli, S. S. (1999). “Uncertainty Analysis of Transport-Transformation Models”. PhD thesis. The State University of New Jersey.

Jakeman, J., M. Eldred, & K. Sargsyan (2015). “Enhancing ℓ_1-minimization estimates of polynomial chaos expansions using basis selection”. In: *J. Comput. Phys.* 289, pp. 18–34.

Jakeman, J., M. Eldred, & D. Xiu (2010). “Numerical approach for quantification of epistemic uncertainty”. In: *J. Comput. Phys.* 229 (12), pp. 4648–4663.

Jala, M. (2013). “Plans d’expériences adaptatifs pour le calcul de quantiles et application à la dosimétrie numérique”. PhD thesis. Paris, France: Institut Mines-Télécom.

Janbu, N. (1968). *Slope stability computations.* Soil Mech. Found. Engng. Report. Trondheim: Technical University.

Jones, D. R., M. Schonlau, & W. J. Welch (1998). “Efficient global optimization of expensive black-box functions”. In: *J. Glob. Optim.* 13 (4), pp. 455–492.

Kameda, H. & H. Morikawa (1992). “An interpolating stochastic process for simulation of conditional random fields”. In: *Prob. Eng. Mech.* 7 (4), pp. 243–254.

Kameda, H. & H. Morikawa (1994). “Conditioned stochastic processes for conditional random fields”. en. In: *J. Eng. Mech.* 120 (4), pp. 855–875.

Kanno, Y. & I. Takewaki (2006a). “Robustness analysis of trusses with separable load and structural uncertainties”. In: *Int. J. Solids Struct.* 43 (9), pp. 2646–2669.

Kanno, Y. & I. Takewaki (2006b). “Sequential semidefinite program for maximum robustness design of structures under load uncertainty”. In: *J. Optim. Theory Appl.* 130 (2), pp. 265–287.

Kaymaz, I. (2005). “Application of Kriging method to structural reliability problems”. In: *Structural Safety* 27 (2), pp. 133–151.

Keese, A. & H. Matthies (2005). “Hierarchical parallelisation for the solution of stochastic finite element equations”. In: *Comp. Struct.* 83, pp. 1033–1047.

Kersaudy, P., B. Sudret, N. Varsier, O. Picon, & J. Wiart (2015). “A new surrogate modeling technique combining Kriging and polynomial chaos expansions – application to uncertainty analysis in computational dosimetry”. In: *J. Comput. Phys* 286, pp. 103–117.

Klir, G. J. (2006). *Uncertainty and information - foundations of generalized information theory.* Wiley & Sons.

Kolmogoroff, A. (1941). “Confidence limits for unknown distribution function”. In: *Ann. Math. Stat.* 12 (4), pp. 461–463.

Koutsourelakis, P. (2004). “Reliability of structures in high dimensions, part II: theoretical validation”. In: *Prob. Eng. Mech.* 19, pp. 419–423.

Koutsourelakis, P., H. Pradlwarter, & G. Schuëller (2004). “Reliability of structures in high dimensions, part I: algorithms and applications”. In: *Prob. Eng. Mech.* 19, pp. 409–417.

Kozine, I. O. & L. V. Utkin (2005). “Constructing imprecise probability distributions”. In: *Int. J. Gen. Syst.* 34 (4), pp. 401–408.

Krige, D. (1951). “A statistical approach to some basic mine valuation problems on the Witwatersrand”. In: *J. of the Chem., Metal. and Mining Soc. of South Africa* 52 (6), pp. 119–139.

Krzykacz-Hausmann, B. (2006). “An approximate sensitivity analysis of results from complex computer models in the presence of epistemic and aleatory uncertainties”. In: *Reliab. Eng. Sys. Safety* 91 (10-11), pp. 1210–1218.

Kucherenko, S., M. Rodriguez-Fernandez, C. Pantelides, & N. Shah (2009). “Monte Carlo evaluation of derivative-based global sensitivity measures”. In: *Reliab. Eng. Sys. Safety* 94, pp. 1135–1148.

Ladevèze, P., G. Puel, & T. Romeuf (2006). “Lack of knowledge in structural model validation”. In: *Comput. Methods Appl. Mech. Eng.* 195 (37-40), pp. 4697–4710.

Lamboni, M., B. Iooss, A.-L. Popelin, & Gamboa (2013). “Derivative-based global sensitivity measures: general links with Sobol’ indices and numerical analysis”. In: *Math. Comput. Simul.* 87, pp. 45–54.

Lang, H.-J., J. Huder, P. Amann, & A. M. Puzrin (2007). *Bodenmechanik und Grundbau.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Lataniotis, C., S. Marelli, & B. Sudret (2015). *UQLab user manual – Kriging.* Tech. rep. Report # UQLab-V0.9-105. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich.

Le Maître, O., M. Reagan, H. Najm, R. Ghanem, & O. Knio (2002). “A stochastic projection method for fluid flow – II. random process”. In: *J. Comput. Phys.* 181, pp. 9–44.

Le, T. M. H. (2014). “Reliability of heterogeneous slopes with cross-correlated shear strength parameters”. en. In: *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* 8 (4), pp. 250–257.

Lebrun, R. & A. Dutfoy (2009a). “A generalization of the Nataf transformation to distributions with elliptical copula”. In: *Prob. Eng. Mech.* 24 (2), pp. 172–178.

Lebrun, R. & A. Dutfoy (2009b). “An innovating analysis of the Nataf transformation from the copula viewpoint”. In: *Prob. Eng. Mech.* 24 (3), pp. 312–320.

Lee, S. & B. Kwak (2006). “Response surface augmented moment method for efficient reliability analysis”. In: *Structural Safety* 28, pp. 261–272.

Lemaire, M. (2009). *Structural reliability.* Wiley.

Li, C. & A. Der Kiureghian (1993). “Optimal discretization of random fields”. In: *J. Eng. Mech.* 119 (6), pp. 1136–1154.

Li, C. & S. Mahadevan (2016a). “Relative contributions of aleatory and epistemic uncertainty sources in time series prediction”. In: *Int. J. Fatigue* 82, pp. 474–486.

Li, C. & S. Mahadevan (2016b). “Role of calibration, validation, and relevance in multi-level uncertainty integration”. In: *Reliab. Eng. Sys. Safety* 148, pp. 32–43.

Li, D.-Q., K.-B. Shao, Z.-J. Cao, X.-S. Tang, & K.-K. Phoon (2016). “A generalized surrogate response aided-subset simulation approach for efficient geotechnical reliability-based design”. In: *Comput. Geotech.* 74, pp. 88–101.

Lin, Y.-K. (1967). *Probabilistic theory of structural dynamics.* McGraw-Hill.

Liu, W., T. Belytschko, & A. Mani (1986a). “Probabilistic finite elements for non linear structural dynamics”. In: *Comput. Methods Appl. Mech. Eng.* 56, pp. 61–86.

Liu, W., T. Belytschko, & A. Mani (1986b). “Random field finite elements”. In: *Int. J. Numer. Meth. Engng.* 23 (10), pp. 1831–1845.

Loève, M. (1977). *Probability theory.* Springer Verlag, New-York.

Lophaven, S. N., H. B. Nielsen, & J. Sondergaard (2002). *DACE - A Matlab Kriging Toolbox.* Tech. rep. Informatics and mathematical modelling, Technical University of Denmark.

Lowe, J. & L. Karafiath (1960). “Stability of earth dams upon drawdown”. In: *Proc. 1st Pan-Am. Conf. Soil Mech. Found. Engng.* Pp. 537–552.

Mallat, S. & Z. Zhang (1993). “Matching pursuits with time-frequency dictionaries”. In: *IEEE T. Signal. Proces.* 41 (12), pp. 3397–3415.

Mallows, C. L. (1956). “Generalizations of Tchebycheff’s inequalities”. In: *J. R. Stat. Soc.* 18, pp. 139–176.

Mara, T. A. (2009). “Extension of the RBD-FAST method to the computation of global sensitivity indices”. In: *Reliab. Eng. Sys. Safety* 94 (8), pp. 1274–1281.

Marelli, S., C. Lamas, & B. Sudret (2015). *UQLab user manual – Sensitivity analysis.* Tech. rep. Report # UQLab-V0.9-106. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich.

Marelli, S., C. Lamas-Fernandes, R. Schöbi, & B. Sudret (2015). *UQLab user manual – Reliability analysis.* Tech. rep. Report # UQLab-V0.9-107. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich.

Marelli, S. & B. Sudret (2014). “UQLab: a framework for uncertainty quantification in MATLAB”. In: *Advances in Engineering Software.* in preparation.

Marelli, S. & B. Sudret (2015). *UQLab user manual – Polynomial chaos expansions.* Tech. rep. Report # UQLab-V0.9-104. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich.

Marelli, S. & B. Sudret (2016). “Bootstrap-polynomial chaos expansions and adaptive design for reliability analysis”. In: *Proc. 6th Asian-Pacific Symp. Struct. Reliab. (APSSRA’2016).*

Marrel, A., B. Iooss, F. Van Dorpe, & E. Volkova (2008). “An efficient methodology for modeling complex computer codes with Gaussian processes”. In: *Comput. Stat. Data An.* 52, pp. 4731–4744.

Massey, F. J. (1951). “The Kolmogorov-Smirnov test for goodness of fit”. In: *J. Am. Stat. Assoc.*

Matsui, T. & K.-C. San (1992). “Finite element slope stability analysis by shear strength reduction technique”. In: *Soils Found.* 32 (1), pp. 59–70.

Matthies, H., C. Brenner, C. Bucher, & C. Guedes Soares (1997). “Uncertainties in probabilistic numerical analysis of structures and solids - stochastic finite elements”. In: *Structural Safety* 19 (3), pp. 283–336.

McClave, J. & T. Sincish (2013). *A first course in statistics.* 11th Edition. Pearson.

McKay, M. (1995). *Evaluating prediction uncertainty.* Tech. rep. NUREG/CR-6311. Los Alamos National Laboratory.

McKay, M., R. Beckman, & W. Conover (1979). “A comparison of three methods for selecting values of input variables in the analysis of output from a computer code”. In: *Technometrics* 2, pp. 239–245.

Melchers, R. (1989). “Importance sampling in structural systems”. In: *Structural Safety* 6, pp. 3–10.

Melchers, R. (1999). *Structural reliability analysis and prediction.* John Wiley & Sons.

Metropolis, N. & S. Ulam (1949). “The Monte Carlo method”. In: *J. Am. Stat. Assoc.* 44 (247), pp. 335–341.

Migliorati, G., F. Nobile, E. von Schwerin, & R. Tempone (2014). “Analysis of discrete L2 projection on polynomial spaces with random evaluations”. In: *Found. Comput. Math.* 14 (3), pp. 419–456.

Miller, L. H. (1956). “Table of percentage points of Kolmogorov statistics”. In: *J. Amer. Stat. Soc.* 51, pp. 111–121.

Mockus, J., V. Tiesis, & A. Zilinskas (1978). “The application of Bayesian methods for seeking the extremum”. In: *Towards Global Optimization.* Noth Holland, Amsterdam.

Möller, B. & M. Beer (2004). *Fuzzy Randomness.* Springer.

Möller, B. & M. Beer (2008). “Engineering computation under uncertainty – capabilities of non-traditional models”. In: *Comput. Struct.* 86 (10), pp. 1024–1041.

Monteiro, S. T., F. Ramos, & P. Hatherly (2009). “Conditional random fields for rock characterization using drill measurements”. In: *2009 Int. Conf. Mach. Learn. Appl.* IEEE, pp. 366–371.

Moore, R. (1966). *Interval analysis.* Englewood Cliffs, NJ: Prentice Hall.

Morgenstein, N. & V. Price (1965). “The analysis of stability of general slip surfaces”. In: *Géotechnique* 15 (1), pp. 79–93.

Morio, J. & M. Balesdent (2016). “Estimation of a launch vehicle stage fallout zone with parametric and non-parametric importance sampling algorithms in presence of uncertain input distributions”. In: *Aerosp. Sci. Technol.* 52, pp. 95–101.

Morio, J., M. Balesdent, D. Jacquemart, & C. Vergé (2014). “A survey of rare event simulation methods for static input-output models”. In: *Simul. Model. Pract. Theory* 49, pp. 287–304.

Morris, M. D. (1991). “Factorial sampling plans for preliminary computational experiments”. In: *Technometrics* 33 (2), pp. 161–174.

Morris, M. D. (1993). “Bayesian design and analysis of computer experiments - use of derivatives in surface prediction”. In: *Technometrics* 35 (3), pp. 243–255.

Moustapha, M. (2016). “Métamodèles adaptifs pour l’optimisation fiable multi-prestations de la masse de véhicules”. PhD thesis. Clermont-Ferrand, France: Université Blaise Pascal - Clermont II.

Moustapha, M., B. Sudret, J. M. Bourinet, & B. Guillaume (2016). “Quantile-based optimization under uncertainties using adaptive Kriging surrogate models”. In: *Struct. Multidisc. Optim.* In press.

Mulvey, J. M., R. J. Vanderbei, & S. A. Zenios (1995). “Robust optimization of large-scale systems”. In: *Oper. Res.* 43 (2), pp. 264–281.

Muscolino, G., R. Santoro, & A. Sofi (2015). “Explicit reliability sensitivities of linear structures with interval uncertainties under stationary stochastic excitation”. In: *Structural Safety* 52, pp. 219–232.

Muscolino, G. & A. Sofi (2012). “Stochastic analysis of structures with uncertain-but-bounded parameters via improved interval analysis”. In: *Prob. Eng. Mech.* 28, pp. 152–163.

Nagel, J. & B. Sudret (2016a). “A unified multilevel framework for Bayesian inversion under uncertainty”. In: *Prob. Eng. Mech.* 43, pp. 68–84.

Nagel, J. & B. Sudret (2016b). “Spectral likelihood expansions for Bayesian inference”. In: *J. Comput. Phys.* 309, pp. 267–294.

Nagel, J. & B. Sudret (2014). “A Bayesian multilevel framework for uncertainty characterization and the NASA Langley multidisciplinary UQ challenge”. In: *16th AIAA Non-Deterministic Approaches Conf.* Reston, Virginia: American Institute of Aeronautics and Astronautics.

Nelsen, R. (2006). *An introduction to copulas.* 2nd. Vol. 139. Lecture Notes in Statistics. Springer-Verlag, New York.

Neumaier, A. (2004). “Clouds, fuzzy sets, and probability intervals”. In: *Reliab. Comput.* 10 (4), pp. 249–272.

Niederreiter, H. (1988). “Low-discrepancy and low-dispersion sequences”. In: *J. Number Theory* 30, pp. 51–70.

Oakley, J. E. & A. O’Hagan (2004). “Probabilistic sensitivity analysis of complex models: a Bayesian aproach”. In: *J. R. Stat. Soc. B* 66, pp. 751–769.

Oberguggenberger, M. & W. Fellin (2005). “Assessing the sensitivity of failure probabilities: a random set approach”. In: *Proc. 9th Int. Conf. Struct. Safety and Reliability (ICOSSAR’2005), Rome, Italy.* Ed. by G. Augusti, G. Schuëller, & M. Ciampoli. Millpress, Rotterdam, pp. 1755–1760.

Oberguggenberger, M. & W. Fellin (2008). “Reliability bounds through random sets: nonparametric methods and geotechnical applications”. In: *Comput. Struct.* 86 (10), pp. 1093–1101.

Oberguggenberger, M. (2014). “Analysis and computation with hybrid random set stochastic models”. In: *Structural Safety* 52 (Part B), pp. 233–243.

Oberguggenberger, M., J. King, & B. Schmelzer (2009). “Classical and imprecise probability methods for sensitivity analysis in engineering: a case study”. In: *Int. J. Approx. Reason.* 50 (4), pp. 680–693.

Papaioannou, I. & A. Der Kiureghian (2010). “Reliability-based design of slope angle considering spatial variability of soil material”. In: *Comput. Stoch. Mech. - Proc. 6th Int. Conf.* Ed. by G. Deodatis & P. D. Spanos, pp. 1–10.

Patelli, E., D. A. Alvarez, M. Broggi, & M. de Angelis (2015). “Uncertainty management in multidisciplinary design of critical safety systems”. In: *Journal of Aerospace Information Systems* 12, pp. 140–169.

Patelli, E., M. Broggi, & M. de Angelis (2014). “An integrated and efficient numerical framework for uncertainty quantification: application to the NASA Langley multidisciplinary uncertainty quantification challenge”. In: *16th AIAA Non-Deterministic Approaches Conf.* National Harbor, Maryland.

Pati, Y., R. Rezaiifar, & P. Krishnaprasad (1993). “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition”. In: *Proc. 27th Asilomar Conference on Signals, Systems and Computers*, pp. 40–44.

Peng, J., J. Hampton, & A. Doostan (2014). “A weighted-minimization approach for sparse polynomial chaos expansions”. In: *J. Comput. Phys.* 267, pp. 92–111.

Phoon, K.-K. (2008). *Reliability-based design in geotechnical engineering: computations and applications.* CRC Press.

Pierce, S., K. Worden, & G. Manson (2006). “A novel information-gap technique to assess reliability of neural network-based damage detection”. In: *J. Sound Vib.* 293 (1-2), pp. 96–111.

Rackwitz, R. & B. Fiessler (1978). “Structural reliability under combined load sequences”. In: *Computers & Structures* 9, pp. 489–494.

Ranjan, P., D. Bingham, & G. Michailidis (2012). “Sequential experiment design for contour estimation from complex computer codes”. In: *Technometrics* 50 (4), pp. 527–541.

Rasmussen, C. & C. Williams (2006). *Gaussian processes for machine learning.* Internet. Adaptive computation and machine learning. Cambridge, Massachusetts: MIT Press, p. 266.

Rastrigin, L. (1974). *Systems of extremal control.* Zinatne, Riga.

Rosenbrock, H. (1960). “An automatic method for finding the greatest or least value of a function”. In: *Comput. J.* 3, pp. 175–184.

Roustant, O., D. Ginsbourger, & Y. Deville (2012). “DiceKriging, DiceOptim: two R package for the analysis of computer experiments by Kriging-based metamodeling and optimisation”. In: *J. Stat. Softw.* 51, pp. 1–55.

Roustant, O., D. Ginsbourger, & Y. Deville (2013). *Package DiceKriging - Kriging methods for computer experiments.*

Rowe, N. C. (1988). “Absolute bounds on the mean and standard deviation of transformed data for constant-sign-derivative transformations”. In: *SIAM J. Sci. Stat. Comput.* 9, pp. 1098–1113.

Sacks, J., W. Welch, T. Mitchell, & H. Wynn (1989). “Design and analysis of computer experiments”. In: *Stat. Sci.* 4, pp. 409–435.

Saltelli, A., K. Chan, & E. Scott, eds. (2000). *Sensitivity analysis.* J. Wiley & Sons.

Saltelli, A. & I. Sobol’ (1995). “About the use of rank transformation in sensitivity of model output”. In: *Reliab. Eng. Sys. Safety* 50, pp. 225–239.

Saltelli, A., S. Tarantola, & K. Chan (1999). “A quantitative, model independent method for global sensitivity analysis of model output”. In: *Technometrics* 41 (1), pp. 39–56.

Sankararaman, S. & S. Mahadevan (2013). “Separating the contributions of variability and parameter uncertainty in probability distributions”. In: *Reliab. Eng. Sys. Safety* 112, pp. 187–199.

Santner, T., B. Williams, & W. Notz (2003). *The Design and Analysis of Computer Experiments.* Springer, New York.

Saporta, G. (2006). *Probabilités, analyse des données et statistique.* Technip.

Sarangi, S., K. K. Bodla, S. V. Garimella, & J. Y. Murthy (2014). “Manifold microchannel heat sink design using optimization under uncertainty”. In: *Int. J. Heat Mass Transfer* 69, pp. 92–105.

Sargsyan, K., C. Safta, H. Najm, B. Debusschere, D. Ricciuto, & P. Thornton (2014). “Dimensionality reduction for complex models via Bayesian compressive sensing”. In: *Int. J. Uncer. Quant.* 4 (1), pp. 63–93.

Schöbi, R., S. Marelli, & B. Sudret (2016). *UQLab user manual – PC-Kriging.* Tech. rep. Report # UQLab-V0.9-109. Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich.

Schöbi, R. & B. Sudret (2014a). *Combining Polynomial Chaos Expansions and Kriging.* Tech. rep. RSUQ-2014-01. Switzerland: Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, pp. 1–48.

Schöbi, R. & B. Sudret (2014b). “Combining polynomial chaos expansions and Kriging for solving structural reliability problems”. In: *Proc. 7th Int. Conf. on Comp. Stoch. Mech (CSM7), Santorini, Greece.* Ed. by P. Spanos & G. Deodatis.

Schöbi, R. & B. Sudret (2014c). “PC-Kriging: a new meta-modelling method and its application to quantile estimation”. In: *Proc. 17th IFIP WG7.5 Conf. Reliab. Optim. Struct. Syst. Huangshan, China.* Ed. by Y. Li, J and Zhao. 1. Taylor & Francis.

Schöbi, R. & B. Sudret (2014d). “PC-Kriging: a new metamodelling method combining Polynomial Chaos Expansions and Kriging”. In: *Proc. 2nd Int. Symp. Uncert. Quant. Stoch. Mod.* Rouen, France.

Schöbi, R. & B. Sudret (2014e). *PC-Kriging in adaptive experimental design algorithms.* Tech. rep. Switzerland: Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich.

Schöbi, R. & B. Sudret (2015a). “Application of conditional random fields and sparse polynomial chaos expansions to geotechnical problems”. In: *Proc. 5th Int. Sympo. Geotech. Safety Risk (ISGSR).* Rotterdam, The Netherlands.

Schöbi, R. & B. Sudret (2015b). “Imprecise structural reliability analysis using PC-Kriging”. In: *Proc. 25th European Safety and Reliability Conference (ESREL2015), Zurich, Switzerland.*

Schöbi, R. & B. Sudret (2015c). “Propagation of uncertainties modelled by parametric p-boxes using sparse polynomial chaos expansions”. In: *Proc. 12th Int. Conf. App. Stat. Prob. Civil Eng. (ICASP12), Vancouver, Canada.*

Schöbi, R. & B. Sudret (2016a). “Comparing probabilistic and p-box modelling in structural reliability analysis”. In: *Proc. 18th IFIP WG7.5 Conf. Reliab. Optim. Struct. Syst. Pittsburgh PA, USA.*

Schöbi, R. & B. Sudret (2016b). “Multi-level meta-modelling in imprecise structural reliability analysis”. In: *Proc. 6th Asian-Pacific Symp. Struct. Reliab. (APSSRA’2016).* Shanghai, China.

Schöbi, R. & B. Sudret (2017a). “PCE-based imprecise Sobol’ indices”. In: *Proc. 12th Int. Conf. Struct. Safety and Reliability (ICOSSAR’2017), Vienna, Austria.*

Schöbi, R. & B. Sudret (2017b). "Uncertainty propagation of p-boxes using sparse polynomial chaos expansions". In: *J. Comput. Phys.* accepted.

Schöbi, R., B. Sudret, & S. Marelli (2016). "Rare event estimation using Polynomial-Chaos-Kriging". In: *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* P. D4016002.

Schöbi, R., B. Sudret, & J. Wiart (2015). "Polynomial-chaos-based Kriging". In: *Int. J. Uncer. Quant.* 5 (2), pp. 171–193.

Schueremans, L. & D. Van Gemert (2005a). "Benefit of splines and neural networks in simulation based structural reliability analysis". In: *Structural Safety* 27 (3), pp. 246–261.

Schueremans, L. & D. Van Gemert (2005b). "Use of Kriging as meta-model in simulation procedures for structural reliability". In: *Proc. 9th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2005), Rome, Italy*, pp. 2483–2490.

Sentz, K. & S. Ferson (2002). *Combination of Evidence in Dempster-Shafer Theory.* Technical Report SAND2002-0835. Albuquerque, New Mexico: Sandia National Laboratories.

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton, NJ: Princeton University Press.

Shah, H., S. Hosder, & T. Winter (2015). "Quantification of margins and mixed uncertainties using evidence theory and stochastic expansions". In: *Reliab. Eng. Sys. Safety* 138, pp. 59–72.

Smirnov, N. (1939). "On the estimation of the discrepancy between empirical curves of distribution for two independent samples". In: *Bulletin de l'Université de Moscou, Série internationale (Mathématiques)* 2.

Smith, J. E. (1995). "Generalized Chebychev inequalities: theory and applications in decision analysis". In: *Oper. Res.* 43, pp. 807–825.

Smola, A. & B. Schölkopf (2006). "A tutorial on support vector regression". In: *Stat. Comput.* 14, pp. 199–222.

Sobol', I. (1967). "The distribution of point in a cube and the approximate evaluation of integrals". In: *USSR. Comput. Math. Math. Phys.* 7 (4), pp. 86–112.

Sobol', I. (1993). "Sensitivity estimates for nonlinear mathematical models". In: *Math. Modeling & Comp. Exp.* 1, pp. 407–414.

Sobol', I. (2001). "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates". In: *Math. Comput. Simul.* 55 (1-3), pp. 271–280.

Sobol', I. & S. Kucherenko (2009). "Derivative based global sensitivity measures and their link with global sensitivity indices". In: *Math. Comput. Simul.* 79 (10), pp. 3009–3017.

Sobol', I. & S. Kucherenko (2010). "A new derivative based importance criterion for groups of variable and its link with global sensitivity indices". In: *Comp. Phys. Commun.* 181, pp. 1212–1217.

Soize, C. & R. Ghanem (2004). "Physical systems with random uncertainties: chaos representations with arbitrary probability measure". In: *SIAM J. Sci. Comput.* 26 (2), pp. 395–410.

Spencer, E. (1967). "A method of analysis of the stability of embankments assuming parallel interslice forces". In: *Géotechnique* 17 (1), pp. 11–26.

Srinivas, N., A. Krause, S. Kakade, & M. Seeger (2012). "Information-theoretic regret bounds for Gaussian process optimization in the bandit setting". In: *IEEE Trans. Inform. Theory* 58 (5), pp. 3250–3265.

Stein, M. (1999). *Interpolation of Spatial Data: some Theory for Kriging.* Springer-Verlag, New York.

Stolfi, J. & L. H. De Figueiredo (2003). "An introduction to affine arithmetic". In: *TEMA Tend. Mat. Appl. Comput.* 4 (3), pp. 297–312.

Stone, M. (1974). "Cross-validatory choice and assessment of statistical predictions". In: *J. R. Stat. Soc., Ser. B* 36, pp. 111–147.

Storlie, C., L. Swiler, J. Helton, & C. Sallaberry (2009). "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models". In: *Reliab. Eng. Sys. Safety* 94 (11), pp. 1735–1763.

Storn, R. & K. Price (1997). "Differential evolution, a simple and efficient heuristic for global optimization over continuous spaces". In: *J. Global Optim.* 11 (4), pp. 341–359.

Sudret, B. (2006). "Global sensitivity analysis using polynomial chaos expansions". In: *Proc. 5th Int. Conf. on Comp. Stoch. Mech (CSM5), Rhodos, Greece.* Ed. by P. Spanos & G. Deodatis.

Sudret, B. (2007). *Uncertainty propagation and sensitivity analysis in mechanical models – contributions to structural reliability and stochastic spectral methods.* Habilitation à diriger des recherches. Université Blaise Pascal, Clermont-Ferrand, France.

Sudret, B. (2008). "Global sensitivity analysis using polynomial chaos expansions". In: *Reliab. Eng. Sys. Safety* 93, pp. 964–979.

Sudret, B. (2012). "Meta-models for structural reliability and uncertainty quantification". In: *Proc. 5th Asian-Pacific Symp. Struct. Reliab. (APSSRA'2012), Singapore.* Ed. by K. Phoon, M. Beer, S. Quek, & S. Pang. Keynote lecture, pp. 53–76.

Sudret, B. (2015). "Polynomials chaos expansions and stochastic finite element methods". In: *Risk Reliab. Geotech. Eng.* Ed. by K.-K. Phoon & J. Ching. Taylor and Francis. Chap. 6.

Sudret, B., G. Blatman, & M. Berveiller (2007). "Quasi random numbers in stochastic finite element analysis – Application to global sensitivity analysis". In: *Proc. 10th Int. Conf. App. Stat. Prob. Civil Eng. (ICASP10), Tokyo, Japan.*

Sudret, B., G. Defaux, & M. Pendola (2004). "Finite element reliability analysis of cooling towers submitted to degradation". In: *Proc. 2nd Int. ASRANet Colloquium, Barcelona, Spain.*

Sudret, B. & A. Der Kiureghian (2000). *Stochastic finite elements and reliability: a state-of-the-art report.* Tech. rep. UCB/SEMM-2000/08. CA, USA: University of California, Berkeley.

Sudret, B. & C. Mai (2015). "Computing derivative-based global sensitivity measures using polynomial chaos expansions". In: *Reliab. Eng. Sys. Safety* 134, pp. 241–250.

Sudret, B. & R. Schöbi (2017). "Application of conditional random fields and sparse polynomial chaos expansions in structural reliability analysis". In: *Proc. 12th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2017), Vienna, Austria.*

Suen, C. & W. F. Kuhfeld (2005). "On the construction of mixed orthogonal arrays of strength two". In: *J. Stat. Plan. Infer.* 133, pp. 555–560.

Takada, T. (1990a). "Weighted integral method in multidimensional stochastic finite element analysis". In: *Prob. Eng. Mech.* 5 (4), pp. 158–166.

Takada, T. (1990b). "Weighted integral method in stochastic finite element analysis". In: *Prob. Eng. Mech.* 5 (3), pp. 146–156.

Tatang, M., W. Pan, R. Prinn, & G. McRae (1997). "An efficient method for parametric uncertainty analysis of numerical geophysical models". In: *J. Geophys. Research* 102 (D18), pp. 21925–21932.

The MathWorks Inc. (2016). *Matlab R2016a.* Natick, Massachusetts.

Tibshirani, R. (1996). "Regression shrinkage and selection via LASSO". In: *J. Roy. Stat. Soc. B* 58, pp. 267–288.

Tonon, F. (2004). "Using random set theory to propagate epistemic uncertainty through a mechanical system". In: *Reliab. Eng. Sys. Safety* 85 (1-3), pp. 169–181.

Vanmarcke, E. (1983). *Random fields: analysis and synthesis.* Cambridge, Massachusetts: MIT Press.

Vanmarcke, E. H. & G. A. Fenton (1991). "Conditioned simulation of local fields of earthquake ground motion". In: *Structural Safety* 10 (1-3), pp. 247–264.

Vapnik, V. (1995). *The nature of statistical learning theory.* Springer, p. 314.

Vazquez, E. & E. Walter (2003). "Multi-output support vector regression". In: *Proc. 13th IFAC Symposium on System Identification, Rotterdam*, pp. 1820–1825.

Vorechovský, M. (2008). "Simulation of simply cross correlated random fields by series expansion methods". In: *Structural Safety* 30 (4), pp. 337–363.

Waarts, P.-H. (2000). "Structural reliability using finite element methods: an appraisal of DARS: Directional Adaptive Response Surface Sampling". PhD thesis. Technical University of Delft, The Netherlands.

Walley, P. (2000). "Towards a unified theory of imprecise probability". In: *Int. J. Approx. Reason.* 24 (2-3), pp. 125–148.

Wan, X. & G. Karniadakis (2005). "An adaptive multi-element generalized polynomial chaos method for stochastic differential equations". In: *J. Comput. Phys.* 209, pp. 617–642.

Wan, X. & G. Karniadakis (2006). "Beyond Wiener-Askey expansions: handling arbitrary PDFs". In: *J. Sci. Comput.* 27, pp. 455–464.

Witteveen, J., S. Sarkar, & H. Bijl (2007). "Modeling physical uncertainties in dynamic stall induced fluid-structure interaction of turbine blades using arbitrary polynomial chaos". In: *Computers & Structures* 85, pp. 866–878.

Xiu, D. (2009). "Fast numerical methods for stochastic computations: a review". In: *Comm. Comput. Phys.* 5 (2-4), pp. 242–272.

Xiu, D. & J. Hesthaven (2005). "High-order collocation methods for differential equations with random inputs". In: *SIAM J. Sci. Comput.* 27 (3), pp. 1118–1139.

Xiu, D. & G. Karniadakis (2002). "The Wiener-Askey polynomial chaos for stochastic differential equations". In: *SIAM J. Sci. Comput.* 24 (2), pp. 619–644.

Xu, C. & G. Gertner (2008). "Uncertainty and sensitivity analysis for models with correlated parameters". In: *Reliab. Eng. Sys. Safety* 93, pp. 1563–1573.

Zaki, M. & W. Meira (2014). *Data Mining and Analysis: fundamental Concepts and Algorithms.* Cambridge University Press.

Zhang, H. (2012). "Interval importance sampling method for finite element-based structural reliability assessment under parameter uncertainties". In: *Structural Safety* 38, pp. 1–10.

Zhang, H., R. L. Mullen, & R. L. Muhanna (2010). "Interval Monte Carlo methods for structural reliability". In: *Structural Safety* 32, pp. 183–190.

Zhang, H., H. Dai, M. Beer, & W. Wang (2013). "Structural reliability analysis on the basis of small samples: an interval quasi-Monte Carlo method". In: *Mech. Syst. Signal Process.* 37 (1-2), pp. 137–151.

Zhang, J. & B. Ellingwood (1994). "Orthogonal series expansions of random fields in reliability analysis". In: *J. Eng. Mech.* 120 (12), pp. 2660–2677.

Zhang, Y. (2007). "Orthogonal arrays obtained by repeating-column difference matrices". In: *Discret. Math.* 307, pp. 246–261.

Zhang, Y. & A. Der Kiureghian (1995). "Two improved algorithms for reliability analysis". In: *Proc. 6th IFIP WG7.5 on Reliability and Optimization of Structural systems, Assisi, Italy.* Ed. by R. Rackwitz, G. Augusti, & A. Bori. Chapman & Hall, London.

Zhang, Z., C. Jiang, X. Han, D. Hu, & S. Yu (2014). "A response surface approach for structural reliability analysis using evidence theory". In: *Adv. Eng. Softw.* 69, pp. 37–45.

Zhang, Z., C. Jiang, G. Wang, & X. Han (2015). “First and second order approximate reliability analysis methods using evidence theory”. In: *Reliab. Eng. Sys. Safety* 137, pp. 40–49.