

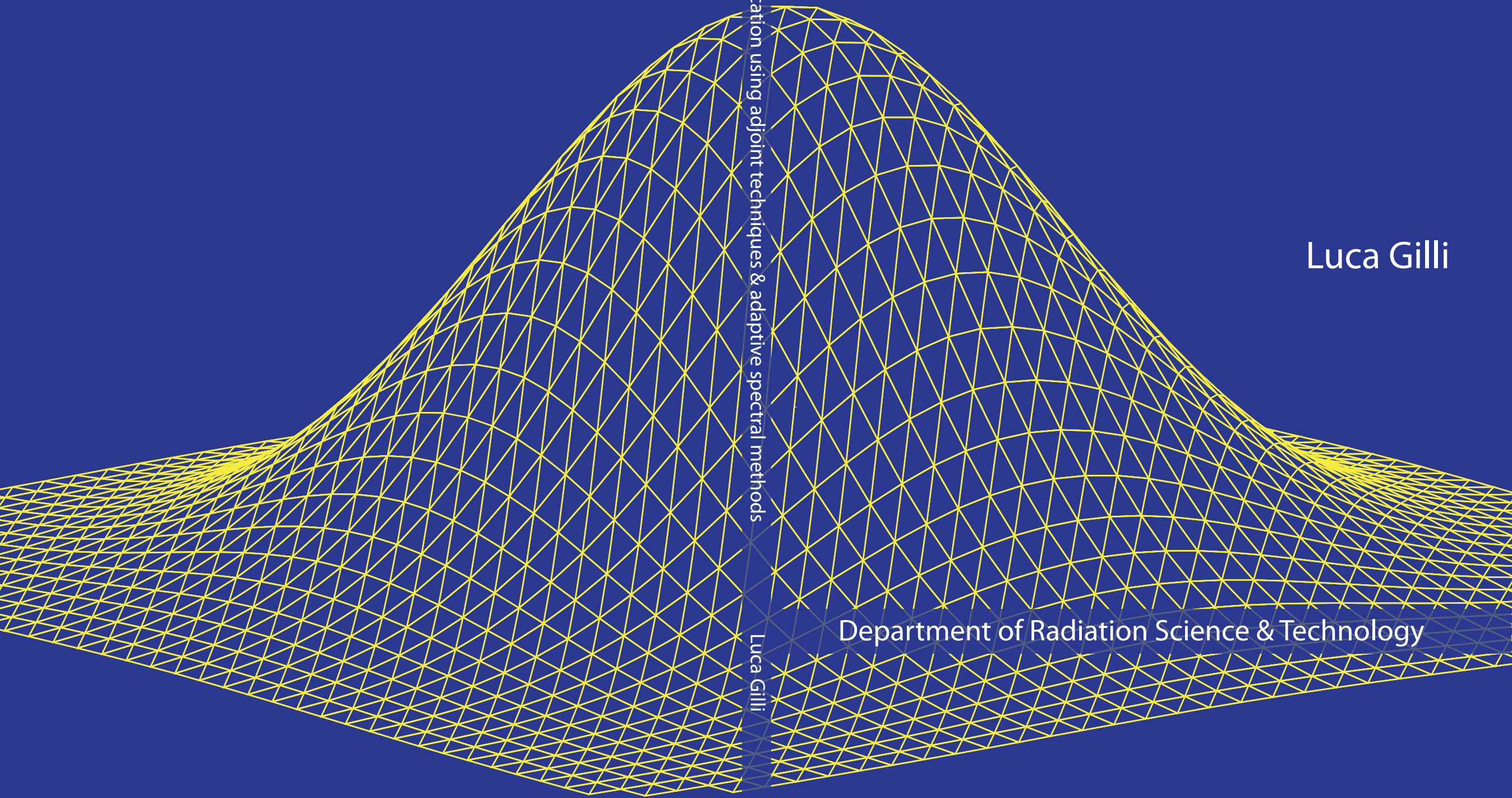
Uncertainty quantification
in reactor physics
using adjoint/perturbation techniques
and adaptive spectral methods

Luca Gilli

Uncertainty quantification using adjoint techniques & adaptive spectral methods

Luca Gilli

Department of Radiation Science & Technology



Propositions

Belonging to the PhD thesis of Luca Gilli

Uncertainty quantification in reactor physics using adjoint/perturbation techniques and adaptive spectral methods

1. The application of higher order adjoint perturbation techniques is not convenient for high-dimensional nonlinear stochastic problems.
2. Adaptive non-intrusive spectral techniques represent the best way to tackle the “curse of dimensionality” associated with spectral methods.
3. For large-scale stochastic problems characterized by a moderate number of input parameters, the adaptive non-intrusive spectral projection method is a faster and more accurate alternative to standard sampling.
4. Sampling based (Monte Carlo) uncertainty quantification techniques will always be the most convenient choice for commercial and large-scale applications due to their ease of implementation. For the same reason, any deterministic approach associated with extensive implementation efforts will be a fancy mathematical exercise rather than an effective tool.
5. Using a realistic risk aversion factor based on the current public perception of radiation effects would make any risk related to nuclear energy unacceptable no matter how low the probability of a radionuclide release.
6. The name Automatic Differentiation is misleading regarding the actual ease of application of the method to large-scale problems
7. Applying the adjective sustainable is an effective way to promote unsustainable practices.
8. The fundamental role that the availability of an affordable energy source has on society in terms of civil rights, family structure, and emancipation is usually ignored by energy saving campaigners.
9. The benefits of the World Wide Web on science are partially counter-balanced by the influence that the same medium has on fake-science.
10. The vehicle-to-vehicle interaction in Indian traffic could be modeled as a Coulomb-like scattering process where instead of the electric charge the honking intensity should be used.

These propositions are considered opposable and defensible and as such have been approved by the promotor, Prof. dr. ir. T.H.J.J. van der Hagen.

Stellingen

Behorende bij het proefschrift van Luca Gilli

Uncertainty quantification in reactor physics using adjoint/perturbation techniques and adaptive spectral methods

1. Het toepassen van hogere orde adjoint verstoringstechnieken is ongeschikt voor hoog-dimensionale niet-lineaire stochastische problemen.
2. Adaptieve niet-intrusieve spectrale technieken zijn de beste methode om de “curse of dimensionality” van spectrale methoden aan te pakken.
3. De adaptieve niet-intrusieve spectrale projectie methode is een sneller en nauwkeuriger alternatief voor standard sampling voor grootschalige stochastische problemen gekenmerkt door een beperkt aantal input parameters.
4. Sampling gebaseerde (Monte Carlo) onzekerheidskwantificeringstechnieken zullen - vanwege hun eenvoud van implementatie - altijd de gunstigste keuze zijn voor commerciële en grootschalige toepassingen. Om dezelfde reden zal elke deterministische aanpak die gepaard gaat met uitgebreide implementatie-inspanningen een fancy wiskundige oefening zijn, in plaats van een effectief instrument.
5. Het toepassen van een realistische risico-aversie factor gebaseerd op de huidige publieke perceptie van de effecten van straling zou elk risico verbonden aan kernenergie onaanvaardbaar maken, ongeacht hoe laag de kans op de uitstoot van een radionuclide ook is.
6. De naam *Automatic Differentiation* is misleidend met betrekking tot het werkelijke gemak van de toepassing van de methode voor grootschalige problemen.
7. Het toepassen van het adjectief duurzaam is een effectieve manier om niet-duurzame praktijken te bevorderen.
8. De fundamentele functie die de beschikbaarheid van een betaalbare energiebron voor de samenleving vervult op het gebied van burgerrechten, gezinsstructuur en emancipatie wordt meestal genegeerd door campagnevoerders voor energiebesparing.
9. De voordelen van het World Wide Web voor de wetenschap worden gedeeltelijk tenietgedaan door de invloed die dit medium heeft op nep-wetenschap.
10. De voertuig-tot-voertuig interactie in het Indiase verkeer kan worden gemodelleerd als een Coulomb-achtig verstrooiingsproces waarin de elektrische lading wordt vervangen door de intensiteit van het getoeter.

Deze stellingen worden oponeerbaar en verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotor, Prof. dr. ir. T.H.J.J. van der Hagen.

**Uncertainty quantification in reactor physics
using adjoint/perturbation techniques
and adaptive spectral methods**

Uncertainty quantification in reactor physics using adjoint/perturbation techniques and adaptive spectral methods

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op dinsdag 09 juli 2013 om 15:00 uur
door

Luca GILLI

Master of Science in Energy and Nuclear Engineering, Politecnico di Torino

geboren te Rivoli, Italië

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. T.H.J.J. van der Hagen

Copromotor:

Dr. ir. D. Lathouwers

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. T.H.J.J. van der Hagen,	Technische Universiteit Delft, promotor
Dr. ir. D. Lathouwers,	Technische Universiteit Delft, copromotor
Prof. dr. ir. drs. H. Bijl,	Technische Universiteit Delft
Prof. dr. ir. A. W. Heemink,	Technische Universiteit Delft
Prof. C. Pain,	Imperial College London, Verenigd Koninkrijk
Prof. dr. A. J. Koning,	Uppsala University, Zweden
Dr. V. Sanchez Espinoza,	Karlsruhe Institute of Technology, Duitsland

© 2013, Luca GILLI

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without prior permission from the copyright owner.

ISBN 978-90-8891-653-3

Keywords: Uncertainty Quantification, Reactor Physics, Adjoint and Spectral Methods

The research described in this thesis was performed in the section Nuclear Energy and Radiation Applications (NERA), of the department Radiation Science & Technology (RST), of the Delft University of Technology, Delft, The Netherlands.

Part of the work presented in this thesis was financed by the Nuclear Research and Consultancy Group (NRG), Petten, The Netherlands.

Printed by: Proefschriftmaken.nl || Uitgeverij BOXPress

Published by: Uitgeverij BOXPress, s-Hertogenbosch

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Uncertainty analysis within the simulation framework	1
1.3	Introduction to standard uncertainty quantification methodologies	5
1.4	Statistical methods	6
1.5	Deterministic methods	8
1.6	Aim of the thesis and overview	11
2	Adjoint sensitivity analysis methodologies	15
2.1	Introduction	15
2.2	Adjoint Sensitivity Analysis applied to multi-physics problems	16
2.3	Conclusion	29
3	Spectral Techniques for Uncertainty Quantification	31
3.1	Introduction	31
3.2	Basic aspects of spectral techniques for uncertainty quantification	32
3.3	Application to an illustrative time-dependent problem	41
3.4	NISP for a large scale problem using MC	51
3.5	Conclusions	54
4	A new adaptive algorithm for the application of NISP techniques	59
4.1	Introduction	59
4.2	Introduction to sparse grids	60
4.3	Definition of a new adaptive algorithm	65
4.4	Illustrative example	73

4.5	Conclusions	79
5	Cost reduction techniques and application to neutronics problems	81
5.1	Introduction	81
5.2	Cost reduction techniques for adaptive sparse grid quadrature algorithms	82
5.3	Application to a criticality problem	86
5.4	Application to a transport problem	93
5.5	Conclusions	98
6	Application to time-dependent multi-physics problems	99
6.1	Introduction	99
6.2	Definition of a reference multi-physics problem	99
6.3	Results: transient caused by a step reactivity insertion	106
6.4	Results: transient caused by a ramp reactivity insertion	107
6.5	Conclusions	114
7	Conclusions and Recommendations	117
	Bibliography	121
	A Appendix	129
	Summary	133
	Samenvatting	135
	Acknowledgements	139
	List of Publications	141
	Curriculum Vitae	143

CHAPTER 1

INTRODUCTION

1.1 Introduction

In this chapter we present a general introduction about the definition of uncertainty propagation and its importance within the reactor physics domain. First, we shortly discuss the meaning of uncertainty and its relevance for the simulation context. Then, a brief overview of two families of methodologies that can be employed to propagate uncertainties is presented. Finally, a summary of the work comprising the thesis is discussed.

1.2 Uncertainty analysis within the simulation framework

Dealing with a complex physical system, like a nuclear reactor, implies the need for the development of proper simulation tools in order to evaluate, for example, new designs and safety properties. These simulations need to be used to ensure a reliable prediction of the behaviour of such a system. For this purpose, a simplified mathematical model is usually derived to describe the system of interest, aiming to reproduce it in its most important features.

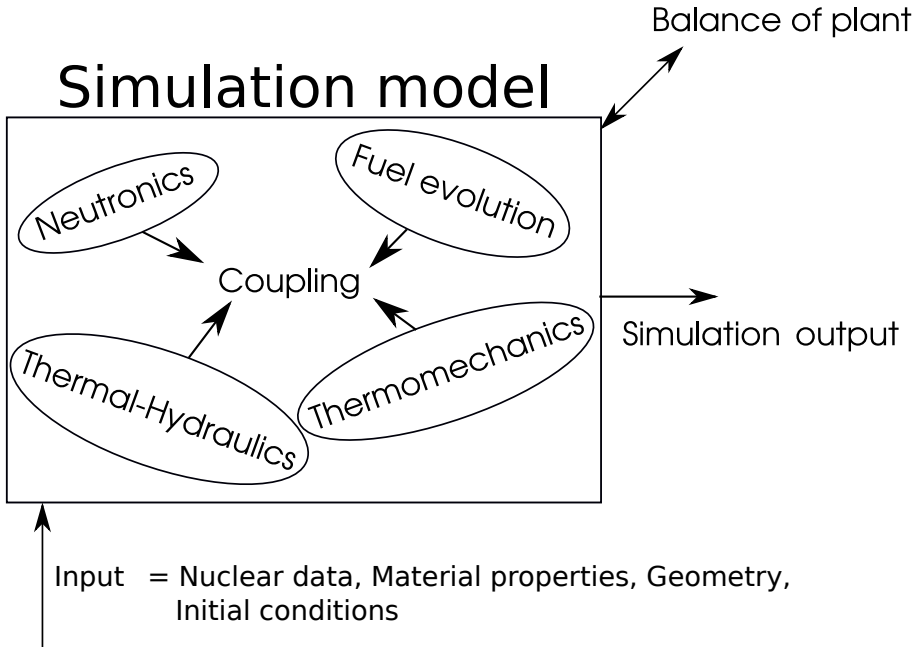


Figure 1.1: Illustrative simulation scheme for a nuclear reactor.

Several physical phenomena are involved while modeling the physics of nuclear reactors. A typical simulation scheme of a nuclear reactor calculation is shown in Figure 1.1. A nuclear reactor is a multi scale system in which different mechanisms determine its final behavior, each mechanism being coupled with each other. These mechanisms range, for example, from neutronics to thermal-hydraulics and mechanics, therefore several physical scales (both in space and in time) are present. As a result, this type of simulation is usually labelled as multi-physics. The development of simulation tools suitable to represent a multi-physics systems is associated with several challenges. The most important question is whether the outcome of these simulations is reliable and to which level of confidence we should trust it. This aspect plays a crucial role in the design and safety analysis of nuclear systems. Performing uncertainty analysis therefore corresponds to estimating the reliability of a simulation model.

Within the nuclear reactor physics field, its role is becoming increasingly

important and challenging as the physical realism of simulations is constantly increasing. Moreover, regulations used in reactor licensing have begun to allow the "best estimate plus uncertainty" simulation approach, increasing the need for reliable and precise uncertainty analysis methodologies (NEA, 2008) (D Auria et al., 2012) (Wilson, 2012).

The role of simulation and the consequent need of uncertainty analysis is crucial for any kind of reactor design, however its application is even more important for innovative ones. For example, the design of a numerical problems included within the Generation IV framework (Generation IV International Forum, 2002) require reliable numerical analysis which is going to be used during the design phase of the first prototypes. Many of these systems include characteristics which are intrinsically different from the current nuclear reactors. Fast reactors, for example, require the application of nuclear libraries of isotopes (especially actinides) whose data are associated with large uncertainties. Since many of these uncertainties are epistemic, performing uncertainty analysis will also indicate the libraries that need further experimental work in order to reduce their associated errors.

1.2.1 Sources of uncertainty

The uncertainty associated to a simulation can be influenced by many aspects which can be divided into the following categories

- Uncertainty due to the simplified mathematical model, caused by assumptions and approximations introduced to represent the physical system. The estimation of the validity of a mathematical model is also known as *validation*.
- Uncertainty caused by numerical approximations, like discretization and truncation errors. Determining whether the model is solved properly is called *verification*.
- Uncertainty introduced by the set of system inputs whose values are possible to determine in theory, but in practice associated to an error, usually because of the lack of proper experiments or data. This type of uncertainty is usually known as *epistemic*, and in reactor physics it is usually associated with cross sections (see Figure 1.2 for an example).

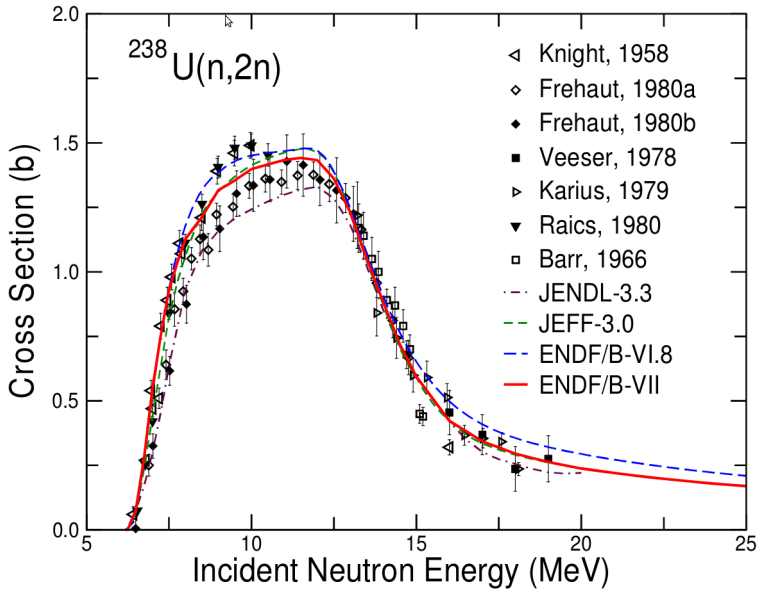


Figure 1.2: Example of uncertainty associated with a nuclear cross-section (from (Chadwick et al., 2006)). Figure contains values corresponding to several data libraries and measurements.

- Uncertainty caused by the set of system parameters whose exact values are practically impossible to determine because of their intrinsic stochasticity. For example, manufacturing errors or the impossibility to represent exact values (impossibility to represent reality). This type of uncertainty is known as *aleatoric*.

These kinds of uncertainties can be further grouped into two main types: the uncertainty generated while modeling the problem and the one associated to the input data used by the simulation. The study of the former requires the analysis of the numerical model itself while the estimation of the latter means that uncertainties entering in the system must be propagated to the simulation outputs of interest.

In this thesis we focus our attention on the propagation of the epistemic uncertainties associated with the input parameters of simulations. The main

objective of this work is the development and the application of methods suitable to tackle this particular problem. Performing uncertainty propagation requires the definition and the application of several statistical concepts, such as probability distributions, standard deviations, etc. In Appendix A an overview of the main statistical indicators used in this thesis is presented.

In the following section some of the standard methodologies used to determine the propagation of uncertainties existing within the input data of a simulation are presented. In general, it must be kept in mind that though this thesis focuses on the application to problems related to innovative nuclear reactors, the uncertainty methodologies that we are going to explore in the next chapters are completely general and can be applied to a wide range of problems.

1.3 Introduction to standard uncertainty quantification methodologies

As we discussed in the previous section, there are different types of uncertainties we have to deal with, however one of the most important challenges in Uncertainty Analysis (UA) is to handle the uncertainty associated with the input data of the problem (like the material properties or the geometric description of a system).

This corresponds to estimating how the lack of knowledge in the input data influences the simulation outputs used during the design and safety analysis. Many techniques have been implemented and used in the field so far, the main methods being statistical and deterministic. The main distinction between the two is that statistical methods are exact and require a large computational effort while deterministic methods usually rely on model approximations which make the technique faster compared to the first approach. In the following sections a more complete overview of these two different approaches is introduced. Before that, it is useful to define a generic mathematical problem which can be used as a reference system within the present chapter.

1.3.1 Definition of a reference mathematical problem

Before discussing uncertainty quantification approaches, it is useful to formulate a generic mathematical problem that can be used as a reference system within the following sections. We will assume in this chapter that we are interested in the application of UA techniques to a specific type of simulation, which is the solution of time-dependent multi-physics problems. We assume that all the spatial operators have been discretized by means of specific numerical techniques (i.e. Finite Element or Finite Volume), in this way we can model our differential system as a set of coupled Ordinary Differential Equations

$$\begin{aligned} \frac{d\mathbf{u}(\mathbf{t}, \boldsymbol{\theta})}{dt} &= L[\boldsymbol{\alpha}(\boldsymbol{\theta}), \mathbf{u}(\mathbf{t}, \boldsymbol{\theta})] \\ \mathbf{u}(0) &= \mathbf{U}_0(\boldsymbol{\theta}) \end{aligned} \tag{1.1}$$

where L is a generic nonlinear operator, \mathbf{u} the unknown solution of the problem, \mathbf{U}_0 the initial conditions, and $\boldsymbol{\alpha}$ the set of input parameters. We introduced the dependency of the input parameters and unknowns on the variable $\boldsymbol{\theta}$ which represents the random space characterizing the uncertainty problem. $\boldsymbol{\alpha}(\boldsymbol{\theta})$, according to this notation, is used to denote all the possible outcomes of the input set $\boldsymbol{\alpha}$. It must be stressed that we are using this system just for illustrative purposes, the application of the UQ techniques we are going to introduce in this chapter can be extended to other types of problems, such as steady-state or eigenvalue calculations.

In general, one may not only be interested in the solution of the previous system but also in a response $R(\boldsymbol{\alpha}, \mathbf{u}, \boldsymbol{\theta})$ which can be described as a functional of the solution and the input parameter set. In principle, this response R can be used to model any of the output quantities of interest during the uncertainty quantification phase. According to the notation just introduced performing UQ corresponds to estimating how the uncertainty present within the input parameter set $\boldsymbol{\alpha}(\boldsymbol{\theta})$ influences the outputs of interest, generically defined by the operator $R(\boldsymbol{\theta})$.

1.4 Statistical methods

Statistical methods, usually known as Monte Carlo (MC) or sampling approaches (Ronen, 1988), can be used to evaluate the statistical moments of

interest by performing a set of simulations of the problem defined by Eq. 1.1. The main concept behind statistical approaches is quite straightforward: the random input data set $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is sampled M times according to their random distributions until the statistical moments of the simulation output have converged sufficiently. Assuming that the number M is large enough to have a statistically meaningful collections of realizations for the response $R(\boldsymbol{\theta})$, the unbiased definition for its mean is

$$E(R) \equiv \frac{1}{M} \sum_{i=1}^M R(\boldsymbol{\theta}^i)$$

where $\boldsymbol{\theta}^i$ is a single realization of the random input set. The unbiased variance is expressed by the equation

$$\sigma^2(R) = \frac{1}{M-1} \sum_{i=1}^M (R(\boldsymbol{\theta}^i) - E(R))^2$$

the statistical error associated with these moments is proportional to $(1/M)^{1/2}$, which means that it will slowly reduce when increasing the number of realizations M . The advantage of MC methods is that they are extremely easy to implement, since their only requirement is to run the simulation for which we want to perform UQ for different values of the input data set. This means that this procedure does not require any substantial modification of the original code used to calculate the output quantities. In theory sampling techniques can be coupled to regression approaches in order to quantify the dependency of the stochastic outputs with respect to the stochastic inputs, however these regression approaches are usually characterized by poorly conditioned operators (Lockwood and Anitescu, 2012).

In general, using the original mathematical model for the evaluation of the statistical moments allows taking into account physical phenomena that would be neglected using approximated methods. The outcome of these approaches will therefore, provided that the number of realizations M is large enough, be close to exact, the only limit being the final statistical error. An example of this aspect can be found in the works of Rochman et al. (2009) and Koning and Rochman (2008) where the authors show that using a sampling technique for UQ of criticality problems is important to include phenomena, such as highly skewed distributions, that would be otherwise neglected.

We mentioned that the main issue associated to sampling methods is given by the poor convergence properties of the statistical estimators. This slow convergence is intrinsic to sampling techniques, however it is not strongly affected by the number of input parameters one is dealing with. For this reason, sampling techniques often represent the only feasible approach to perform UQ when dealing with highly dimensional random problems. The convergence of sampling approaches can be improved by introducing biasing techniques such as the stratified sampling or the Latin Hypercube Sampling (Helton and Davis, 2002) (Wyss and Jorgensen, 1998). These techniques are based on the generation of artificial realization sets which better cover the random domain, therefore improving the convergence rate of the statistical estimators. These realization sets can be built, for example, by assuming the realizations to be uniformly distributed along the stochastic domain or by collecting realizations which are rare from a probability point of view but important in terms of effects on the stochastic outputs.

1.5 Deterministic methods

Deterministic methods are usually based on the assumption that the stochastic quantities of interest can be represented by using a suitable approximation. This approximation is used to simplify the original stochastic problem and to reduce the computational requirements associated to uncertainty propagation. This approximation involves in many cases, unlike sampling approaches, the definition of a new mathematical problem and introduces therefore new computational challenges.

In this thesis we will focus our attention on two types of deterministic methods: perturbation methods and spectral methods. Perturbation methods are based on the representation of random outputs in terms of a lower order Taylor expansion while spectral methods use Fourier like expansions to approximate the same quantities. In the following two sections a more general overview of the two methods is presented.

1.5.1 Perturbation techniques

Perturbation techniques are based on the representation of a generic stochastic quantity by means of a Taylor expansion with respect to the input parameter set around a reference solution

$$\begin{aligned} \delta R = & \sum_{i=1}^K \left(\frac{\partial R}{\partial \alpha_i} \right)_{\alpha_0} \delta \alpha_i + \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K \left(\frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} \right)_{\alpha_0} \delta \alpha_i \delta \alpha_j + \\ & + \frac{1}{3!} \sum_{i=1}^K \sum_{j=1}^K \sum_{k=1}^K \left(\frac{\partial^3 R}{\partial \alpha_i \partial \alpha_j \partial \alpha_k} \right)_{\alpha_0} \delta \alpha_i \delta \alpha_j \delta \alpha_k + O(\delta \alpha^4) \end{aligned} \quad (1.2)$$

where α_0 is a reference value of the input data set (for example, its mean) and $\delta \alpha$ a variation around this value. Assuming the series has converged, knowing its coefficients corresponds to being able to determine any variation of the output δR with respect to any possible variation of the input parameters. We mentioned before that perturbation techniques rely on a model approximation, this approximation is usually obtained by truncating the Taylor series to a lower order. The most common perturbation techniques are first order which means taking into account only the first order derivatives on the right hand side of the equation. This reduction in information about our stochastic output will be of course associated to a reduction in computational cost needed to obtain it, on the other hand this operation could possibly introduce large errors in our predictions. A comprehensive comparison of the consequences of using perturbation techniques instead of sampling approaches was presented by (Rochman et al., 2011).

In general, obtaining derivative information of simulation outputs can be achieved in very different ways. The most straightforward method is the Finite Difference technique (Ronen, 1988) which consists of the evaluation of output derivatives by using the Finite Difference discretization technique. An alternative approach is represented by the concept of Automatic Differentiation (Griewank, 1992) which involves the definition of a new simulation which can be used to directly evaluate the derivatives of the problem. This new simulation is defined by “translating” the source code of the original problem using the concept of chain derivative, unfortunately it presents several challenges from the implementation point of view (Alexe et al., 2010).

Another possibility is the definition of a new mathematical model obtained by linearizing the reference problem defined in Eq. 1.1. This operation, also known as Forward Sensitivity Analysis (Cacuci, 2003), defines a new linear system which can be solved for each input perturbation to evaluate the first order perturbation of the unknowns. One can further reduce the computational requirements of this approach, if interested in a limited set of responses, by introducing an additional problem based on the definition of the adjoint operator (Williams, 1986). This additional adjoint problem, usually known as Adjoint Sensitivity Analysis Procedure, allows evaluating all the possible first order variations of a response with respect to any input perturbation by means of a single additional calculation. This family of methods, known as adjoint perturbation techniques, has been widely used in the field of reactor physics to perform sensitivity analysis and uncertainty propagation in simulations. Their main applications can be found in criticality calculations where such techniques can be used to calculate the perturbation of the fundamental eigenvalue (or any functional of the flux) following a perturbation in the input cross-sections (Williams, 1986),(Gandini, 1967). More recently this kind of techniques has been applied to perform uncertainty propagation for innovative reactor designs (Aliberti et al., 2006) or to perform sensitivity analysis of time-dependent thermal-hydraulics problems (Ionescu-bujor and Cacuci, 2004).

More recently, a new approach for the application of perturbation techniques has been presented by Abdel-Khalik et al. (2008). This method is based on the definition of a reduced subspace able to catch most of the perturbation information of the problem. How to perform the reduced subspace construction has been addressed in several works together with applications to standard reactor physics problems (Bang and Abdel-Khalik, 2012) (Bang et al., 2012). Though these techniques potentially represent an interesting alternative to perturbation approaches, their robustness has not been completely proven yet.

1.5.2 Spectral methods

Spectral techniques represent an alternative deterministic method for uncertainty propagation. They are based on a spectral representation of stochastic input and output quantities by means of a spectral, usually Fourier-like, expansion. The idea of representing a stochastic quantity by means of a spectral expansion was presented by Wiener (1938), who introduced for the first time the

concept of Polynomial Chaos Expansion (PCE). This idea has been extended successively and applied to uncertainty propagation problems by Ghanem and Spanos (1991). Since then, spectral techniques have been applied to different scientific fields, ranging from Computational Fluid Dynamics (Najm, 2009),(Mathelin et al., 2005),(Le Maitre et al., 2002) to structural mechanics (Ghanem and Spanos, 1997).

Two main PCE approaches, categorized as intrusive and non-intrusive, can be used to implement these spectral techniques. As the name suggests the main difference between the approaches is that with the former it is possible to use the original code as a "black box" while the latter involves the definition of a newly coupled problem which needs to be coded and solved. An extensive overview of practical applications of several intrusive and non-intrusive spectral techniques has been presented by Witteveen (2009),(Witteveen and Bijl, 2006).

Within the reactor physics field the application of an intrusive PCE approach was first presented for a neutron diffusion problem by Williams (2007) and later applied to the transport equation in two studies ((Williams, 2006),(Eaton and Williams, 2010)) for fixed source and eigenvalue problems. This concept has been also extended to spatially random problems and to non intrusive methods by Fichtl (2009) and Fichtl and Prinja (2011) while Roderick et al. (2010) presented the application of a PCE based regression technique to a coupled steady-state problem. Regarding time-dependent problems the only application in the nuclear field was proposed by Hagues et al. (2010) where an intrusive stochastic method is applied to a radionuclide dispersion model.

1.6 Aim of the thesis and overview

This thesis focuses on the development and the implementation of an uncertainty propagation algorithm based on the concept of spectral expansion. The motivation behind this development is that spectral techniques can potentially achieve the exactness of sampling approaches in a deterministic fashion. The main objective of this work is therefore to develop a method which would include the advantages in terms of accuracy and computational cost, of both the approaches. This would make such a method highly beneficial while performing uncertainty quantification of multi-physics simulations.

As mentioned in the previous section, there are several ways which can be

used to implement spectral techniques. Consequently, the definition of a new algorithm requires an extensive overview of all the possible approaches which can be employed to implement the method. Furthermore, the knowledge of the most important aspects characterizing uncertainty quantification problems is also needed. Therefore, the first part of the thesis is dedicated to the analysis of the main aspects characterizing uncertainty quantification of multi-physics problems, and to the discussion of the main differences between spectral approaches.

We first show that a convenient way to study uncertainty propagation problems is by implementing perturbation techniques. In **Chapter 2**, an overview of adjoint techniques for the application of sensitivity analysis is presented. This overview includes a discussion about employing adjoint methods for the evaluation of higher-order perturbation components together with an analysis of the method from the numerical point of view.

Chapter 3 focuses on spectral methods. A general overview of their mathematical background is presented together with a general description of the different techniques that can be used to apply them. Two examples that illustrate the application of spectral techniques to reactor physics problems are also included. In this chapter we show that non-intrusive spectral techniques represent a very convenient way to apply the method, especially while dealing with multi-physics problems.

After this overview, we present a new algorithm for the application of a non-intrusive technique to large scale multi-physics problems. **Chapter 4** presents the initial derivation of the new algorithm which can be used to apply non-intrusive spectral techniques in an adaptive fashion. This algorithm is applied to a reference problem and its robustness is tested by comparing it to the outcome of another method found in literature (Gerstner and Griebel, 2003). In **Chapter 5** two techniques used to further reduce the computational cost associated to the algorithm are also discussed. The uncertainty propagation algorithm is then applied to two reference reactor physics problems: a criticality calculation and a source-detector system modeled by using a transport formulation.

Chapter 6 is about the application of the adaptive non-intrusive algorithm to time-dependent multi-physics simulations. A multi-physics model, based on a fast reactor design, is introduced together with a reference uncertainty quantification problem which is solved by means of the adaptive spectral

algorithm.

Finally, the concluding chapter contains an overview and a discussion about the application of the adaptive spectral method derived within this thesis.

CHAPTER 2

ADJOINT SENSITIVITY ANALYSIS METHODOLOGIES

2.1 Introduction

This chapter focuses on the application of adjoint techniques to time-dependent multi-physics problems. First we will introduce the main aspects regarding the application of the first order Adjoint Sensitivity Analysis Procedure to the generic problem defined in Eq. 1.1, then we will discuss its extension to higher order perturbation components. These techniques are applied to a reference nonlinear problem represented by a coupled point-kinetic/lumped parameters system. A general overview of the method and a discussion about its main numerical aspects is finally presented.

2.2 Adjoint Sensitivity Analysis applied to multi-physics problems

Let us start by redefining the problem introduced in Eq. 1.1 by eliminating the dependency with respect to the random variable θ

$$\begin{aligned}\frac{d\mathbf{u}}{dt} &= L(\boldsymbol{\alpha}, \mathbf{u}) \\ \mathbf{u}(0) &= \mathbf{u}_{in}(\boldsymbol{\alpha})\end{aligned}\tag{2.1}$$

where the same notation as in Eq. 1.1 is used. As we mentioned before, we are often interested not in the solution itself but in a response which depends on it. This is the case for example when dealing with integral parameters or minimum or maximum points such as averaged or peak temperatures. A general way to express these responses is by using the following functional

$$R = \langle \mathbf{f}(\boldsymbol{\alpha}, t), \mathbf{u}(t) \rangle\tag{2.2}$$

where we use the following Bra-ket notation to write the time integral

$$\langle \mathbf{a}, \mathbf{b} \rangle = \int_0^{t_f} \mathbf{a} \cdot \mathbf{b} dt$$

We assume for the sake of simplicity the response functional to be linear with respect to the solution and we introduce a weight function \mathbf{f} used to define the response of interest. We also assume that \mathbf{f} is linear with respect to the input parameters.

The first order variation of the response corresponding to a variation of any of the input parameters can be calculated using the Adjoint Sensitivity Analysis Procedure (ASAP). First step to implement the procedure is the introduction of the following perturbation notation

$$\begin{aligned}\mathbf{u} &= \mathbf{u}_0 + \delta\mathbf{u} \\ \boldsymbol{\alpha} &= \boldsymbol{\alpha}_0 + \delta\boldsymbol{\alpha}\end{aligned}$$

where the subscript 0 refers to the reference value of the quantity (which is usually considered to be the mean value of the input parameters and

corresponding solution). According to the theory the first order variation of the functional R can be defined, using the Bra-ket notation, as

$$\delta R \approx \left\langle \frac{\partial \mathbf{f}}{\partial \boldsymbol{\alpha}} \delta \boldsymbol{\alpha}, \mathbf{u}_0 \right\rangle + \left\langle \mathbf{u}^*, \frac{\partial L}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \delta \boldsymbol{\alpha} \right\rangle + \mathbf{u}^*(0) \cdot \delta \mathbf{u}_{in} \quad (2.3)$$

The first term on the right-hand side, known as the direct component, defines the variation caused by a perturbation of the weight function \mathbf{f} while the second term defines the contribution caused by the perturbation of the nonlinear operator. This depends on the operator $\frac{\partial L}{\partial \boldsymbol{\alpha}}$ which is the Gateaux Derivative (GD) of the nonlinear operator L taken in the direction of the parameters perturbation, and on \mathbf{u}^* which is an adjoint function solving the following problem

$$-\frac{d\mathbf{u}^*}{dt} = \left(\frac{\partial L}{\partial \mathbf{u}}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right)^* \mathbf{u}^* + \nabla_{\mathbf{u}} R \quad (2.4)$$

$$\mathbf{u}^*(t_f) = 0$$

Once this adjoint problem is solved for the specific response R (which defines the inhomogeneous term) one can calculate its variation by evaluating a simple time integral. The third term of the Equation 2.3 (known as the bilinear concomitant) includes the effects of a change in the initial boundary value of the time problem on the response: for linear systems and parameter independent initial conditions, this term is exact and defines a duality between the local value (at $t = 0$) of the adjoint solution and the response. One of the main issues associated to adjoint methods is the fact that a new adjoint solution is required for any additional response R . When the method is applied to time-dependent problems any additional response in time would introduce a new calculation, their choice should therefore be weighted carefully.

In general, the first order adjoint theory is derived by perturbing the reference problem and neglecting any term higher than first order, which means that there is a range of validity for the prediction of the variation of the response determined by the magnitude of the input perturbation. In the next part, a method to estimate the second order terms of the expansion using an equivalent adjoint approach used for the first order contribution is introduced.

2.2.1 Second order Adjoint theory

The aim of the second order adjoint theory is to define a way to obtain the second order perturbation components for the response of interest using the adjoint properties already introduced during the definition of the first order propagation technique. The first step is to express each of the second order mixed derivatives of the expansion as

$$\frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} = \frac{\partial}{\partial \alpha_i} \left[\frac{\partial R}{\partial \alpha_j} \right]$$

The first order adjoint theory allows to represent first order derivatives as functionals. It is therefore possible to write the second order information as a first order variation of these functionals (Greenspan et al., 1979). If we consider the adjoint formulation of the first order variation of the response with respect to the parameter α_i and we perturb this quantity in the direction $\delta \alpha_j$ by using the definition of GD, we are able to determine the derivative information we need. This corresponds to estimating the following variation

$$\begin{aligned} & \frac{\partial}{\partial \alpha_i} \left[\frac{\partial R}{\partial \alpha_j} \delta \alpha_j \right] \delta \alpha_i = \\ = & \frac{\partial}{\partial \alpha_i} \left[\left\langle \frac{\partial \mathbf{f}}{\partial \alpha_j} \delta \alpha_j, \mathbf{u}_0 \right\rangle + \left\langle \mathbf{u}^*, \frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right\rangle + \right. \\ & \left. + \mathbf{u}^* \cdot \frac{\partial \mathbf{u}_{in}}{\partial \alpha_j} \right] \delta \alpha_j \delta \alpha_i \end{aligned}$$

where we used Equation 2.3 to express the first order variation of R with respect to α_j . Taking into account that we are assuming the weight function f to be linear with respect to the input parameters, the GD of the previous

expression in the direction α_i can be written as

$$\begin{aligned} \frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} \delta \alpha_j \delta \alpha_i = & \left\langle \mathbf{u}^*, \frac{\partial}{\partial \alpha_i} \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right) \right\rangle \delta \alpha_j \delta \alpha_i + \\ & + \left\langle \delta \mathbf{u}^*(\delta \alpha_i), \frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right\rangle \delta \alpha_j + \\ & + \left\langle \mathbf{u}^*, \frac{\partial}{\partial \mathbf{u}} \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right) \delta \mathbf{u}(\delta \alpha_i) \right\rangle \delta \alpha_j + \\ & + \delta \mathbf{u}^*(\delta \alpha_i, t = 0) \cdot \frac{\partial \mathbf{u}_{in}}{\partial \alpha_j} \delta \alpha_j \end{aligned} \quad (2.5)$$

The unknown quantities in this expression, denoted by $\delta \mathbf{u}(\delta \alpha_i)$ and by $\delta \mathbf{u}^*(\delta \alpha_i)$, are, respectively, the first order variation of the forward and the adjoint solution caused by the variation $\delta \alpha_i$. Once these quantities are known (together with the Hessian of the linear operator with respect to the input parameters) the second order information can be calculated by performing the usual inner products. These first order variations can be evaluated by linearizing the forward and the adjoint problem and solving it with respect to the perturbation considered. If we consider, for example, the nonlinear forward problem, this linearization procedure leads to the following system

$$\begin{aligned} \frac{d\delta \mathbf{u}(\delta \alpha_i)}{dt} = & \left(\frac{\partial L}{\partial \mathbf{u}}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right) \delta \mathbf{u}(\delta \alpha_i) + \left(\frac{\partial L}{\partial \alpha_i}(\boldsymbol{\alpha}_0, \mathbf{u}_0) \right) \delta \alpha_i \\ & \delta \mathbf{u}(0) = \frac{\partial \mathbf{u}_{in}}{\partial \alpha_i} \delta \alpha_i \end{aligned} \quad (2.6)$$

which is defined by the partial GD of the original operator and by a source term and a set of initial conditions which depend on the perturbation considered. Due to its linearity and the time constant nature of the perturbations, this system needs to be solved once for each of the perturbed parameters independently of the perturbation magnitude. The solution of this problem is then used to evaluate the third term of Eq. 2.5. The same linearization has to be applied to the adjoint system; applying the GD to the system described in Eq. 2.4

with respect to α_i leads to

$$-\frac{d\delta\mathbf{u}^*(\delta\alpha_i)}{dt} = \left(\frac{\partial L}{\partial \mathbf{u}}(\boldsymbol{\alpha}_0, \mathbf{u}_0)\right)^* \delta\mathbf{u}^*(\delta\alpha_i) + \quad (2.7)$$

$$+ \frac{\partial}{\partial \alpha_i} \left[\left(\frac{\partial L}{\partial \mathbf{u}}(\boldsymbol{\alpha}_0, \mathbf{u}_0)\right)^* \mathbf{u}^* \right] \delta\alpha_i + \quad (2.8)$$

$$+ \frac{\partial}{\partial \mathbf{u}} \left[\left(\frac{\partial L}{\partial \mathbf{u}}(\boldsymbol{\alpha}_0, \mathbf{u}_0)\right)^* \mathbf{u}^* \right] \delta\mathbf{u}(\delta\alpha_i)$$

$$\delta\mathbf{u}^*(t_f) = 0$$

where we have introduced the GD of the first order adjoint problem. This operation defines a new system of ODEs used to estimate the variation of the adjoint solution with respect to any of the input perturbations. The main operator of this system is the same as the original adjoint problem 2.4, the only difference is the presence of two inhomogeneous terms arising from the application of the GD. Both terms depend on the adjoint solution while the second one introduces a dependency on the first order variation of the forward problem. The term is caused by the perturbation of the part of the adjoint operator which depends on the forward solution, this means that we can solve the system for each of the input perturbations only after having obtained the corresponding solution of problem 2.6.

After that problems 2.6 and 2.7 are solved for all the perturbed parameters, corresponding to a total of $2N$ additional calculations, we can substitute the solutions obtained in Eq. 2.5 to reconstruct the Hessian matrix containing the second order derivatives of the response. It is to be noted that, since the Hessian is a symmetric matrix, changing the indexes in Eq. 2.5 would not affect the value of the second order derivative. This expression is in principle equivalent to the one derived in other works (Gandini, 1978; Ozyurt and Barton, 2005) where the Hessian matrix for a time-dependent functional is obtained by performing $2N$ calculations. In the next section it will be shown that in presence of linear systems it is possible to obtain the Hessian matrix with N calculations, as has been previously proven for different linear applications (Greenspan and Gilai, 1978).

2.2.2 Application of the second order adjoint theory to linear time-dependent problems

In this section the second order adjoint theory presented in the previous Section is applied to linear systems by assuming that the operator L in Eq. 2.1 is linear. We start by rewriting Eq. 2.5, which in the presence of linear operators becomes

$$\begin{aligned} \frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} \delta \alpha_j \delta \alpha_i = & \left\langle \mathbf{u}^*, \frac{\partial}{\partial \alpha_i} \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0) \right) \mathbf{u}_0 \right\rangle \delta \alpha_j \delta \alpha_i + \\ & + \left\langle \delta \mathbf{u}^*(\delta \alpha_i), \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0) \right) \mathbf{u}_0 \right\rangle \delta \alpha_j + \\ & + \left\langle \mathbf{u}^*, \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0) \right) \delta \mathbf{u}(\delta \alpha_i) \right\rangle \delta \alpha_j + \\ & + \frac{\partial \mathbf{u}^*}{\partial \alpha_i} \delta \alpha_i \cdot \frac{\partial \mathbf{u}_{in}}{\partial \alpha_j} \delta \alpha_j \end{aligned} \quad (2.9)$$

where the only difference with respect to the original expression is that the operators are now independent of the reference solution. In a similar manner we can rewrite the system describing the perturbation of the adjoint solution as

$$\begin{aligned} -\frac{d\delta \mathbf{u}^*(\delta \alpha_i)}{dt} = L^* \delta \mathbf{u}^*(\delta \alpha_i) + \left(\frac{\partial L}{\partial \alpha_i}(\boldsymbol{\alpha}_0) \right)^* \mathbf{u}^* \delta \alpha_i \\ \delta \mathbf{u}^*(t_f) = 0 \end{aligned}$$

In this case, due to the linearity of the forward problem, the adjoint operator does not depend on the reference solution \mathbf{u}_0 and the third term on the RHS of Eq. 2.7 disappears. If we define the first order adjoint perturbation as

$$\delta \mathbf{u}^*(\delta \alpha_i) = \left(\frac{\partial \mathbf{u}^*}{\partial \alpha_i} \right) \delta \alpha_i = \hat{\mathbf{u}}^i \delta \alpha_i$$

where we denoted with $\hat{\mathbf{u}}^i$ the partial derivative of the adjoint solution with respect to α_i , we can rewrite the previous system using the linearity property

as

$$-\frac{d\hat{\mathbf{u}}^i}{dt} = L^* \hat{\mathbf{u}}^i + \left(\frac{\partial L}{\partial \alpha_i}(\boldsymbol{\alpha}_0) \right)^* \mathbf{u}^* \quad (2.10)$$

$$\hat{\mathbf{u}}^i(t_f) = 0$$

Moreover, the third term of Eq. 2.9 can be also reformulated when dealing with linear problems. Firstly, we rewrite the term containing the first order perturbation of the forward solution as

$$\left\langle \mathbf{u}^*, \left(\frac{\partial L}{\partial \alpha_j} \right) \delta \mathbf{u}(\delta \alpha_i) \right\rangle \delta \alpha_j = \left\langle \left(\frac{\partial L}{\partial \alpha_j} \right)^* \mathbf{u}^*, \delta \mathbf{u}(\delta \alpha_i) \right\rangle \delta \alpha_j$$

where we used the definition of the adjoint operator to restructure the integral. In this form we can look at the inner product as if it were a first order response variation whose weight function depends on the adjoint solution. This means that we can reconstruct the expression using the adjoint formulation of Eq. 2.3 as

$$\left\langle \left(\frac{\partial L}{\partial \alpha_j} \right)^* \mathbf{u}^*, \delta \mathbf{u}(\delta \alpha_i) \right\rangle \delta \alpha_j =$$

$$\left\langle \hat{\mathbf{u}}^j, \left(\frac{\partial L}{\partial \alpha_i} \right) \mathbf{u}_0 \right\rangle \delta \alpha_j \delta \alpha_i + \hat{\mathbf{u}}^j \delta \alpha_j \cdot \frac{\partial \mathbf{u}_{in}}{\partial \alpha_i} \delta \alpha_i$$

where $\hat{\mathbf{u}}^j$ solves Equation 2.10 considering a variation with respect to α_j . If we substitute this inner product in Equation 2.9 we obtain the following expression for the second order perturbation components

$$\frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} \delta \alpha_i \delta \alpha_j = \quad (2.11)$$

$$\left[\left\langle \mathbf{u}^*, \frac{\partial}{\partial \alpha_i} \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0) \right) \mathbf{u}_0 \right\rangle + \right.$$

$$+ \left\langle \hat{\mathbf{u}}^i, \left(\frac{\partial L}{\partial \alpha_j}(\boldsymbol{\alpha}_0) \right) \mathbf{u}_0 \right\rangle + \left\langle \hat{\mathbf{u}}^j, \left(\frac{\partial L}{\partial \alpha_i}(\boldsymbol{\alpha}_0) \right) \mathbf{u}_0 \right\rangle +$$

$$\left. + \hat{\mathbf{u}}^j \cdot \frac{\partial \mathbf{u}_{in}}{\partial \alpha_i} + \hat{\mathbf{u}}^i \cdot \frac{\partial \mathbf{u}_{in}}{\partial \alpha_j} \right] \delta \alpha_i \delta \alpha_j$$

The combination of inner products, derived using the linearity of the problem, builds a symmetric Hessian matrix, which is fully determined once the solution of the problem 2.10 is known for every perturbed parameter. The evaluation of this expression requires N adjoint calculations instead of the $2N$ needed in Equation 2.5.

It must be pointed out that it is possible to extend this procedure to obtain higher order sensitivity information in a similar fashion. For example, in order to obtain third order sensitivity coefficients we would need to apply the GD to equation 2.11 and use the properties of the adjoint problem in a similar way. This can be easily done in the presence of a single perturbed parameter, in which case it would be possible to obtain higher order components by applying Eq. 2.11 iteratively (for example using the second order adjoint solution as a source for the third order adjoint problem).

2.2.3 Application to a simplified coupled time-dependent problem

We now present the application of the techniques introduced so far to a coupled time-dependent problem. This example is used to illustrate the application of the method from a mathematical and numerical point of view. The problem considered consists of a system of coupled Ordinary Differential Equations (ODE) modeling the time-dependent behavior of a simplified reactor system. The model is derived by using a point-kinetic approximation for the neutron population (Duderstadt and Hamilton, 1976) together with a lumped parameter description of the reactor temperatures. These assumptions allow the elimination of the spatial dependencies and therefore to focus on the time-dependent part. The point-kinetic equations are

$$\begin{aligned} \frac{dP}{dt} &= \frac{\rho(T_f, T_c, t) - \beta}{\Lambda} P + \sum_{k=1}^6 \lambda_k C_k \\ \frac{dC_k}{dt} &= -\lambda_k C_k + \frac{\beta_k}{\Lambda} P \end{aligned} \quad (2.12)$$

where P is the reactor power, Λ the mean generation time, C_k the concentration of the k^{th} precursor group (in power terms), β_k and λ_k the delayed neutrons fraction and the decay constant for the k^{th} precursor group and β the total delayed neutron fraction. The thermo-kinetics/thermal-hydraulics equations,

needed to describe the removal of the heat by the coolant, are approximated using a lumped parameter model, i.e. averaging the unknown values over the whole domain. Assuming the reactor to be divided into a fuel (f) and a coolant (c) region, their (time-dependent) average temperatures are described by the equations

$$\begin{aligned} M_f c_{pf} \frac{dT_f}{dt} &= P + Ah(T_c - T_f) \\ M_c c_{pc} \left[\frac{dT_c}{dt} + v \frac{T_c - T_{in}}{L} \right] &= Ah(T_f - T_c) \end{aligned}$$

where M_f and M_c are the fuel and coolant mass respectively, h the heat transfer coefficient, A the heat transfer area, v the coolant flow velocity, L the channel length and T_{in} the inlet temperature of the coolant. The coupling between these two equations is given by the presence of the power production term P and by the time-dependent reactivity $\rho(t, T_f, T_c)$ in the point kinetic equation, defined as the contribution of three different terms

$$\rho(t, T_f, T_c) = \rho_{ext} + \alpha_D [T_f - T_f(0)] + \alpha_c [T_c - T_c(0)]$$

where ρ_{ext} represents an external reactivity insertion, α_D and α_c are the Doppler and the coolant reactivity coefficients respectively, and $T_f(0)$ and $T_c(0)$ are the initial system temperatures. We assume that the system starts from the following (initial) steady-state conditions

$$\begin{aligned} P(0) &= P_0 \\ C_k(0) &= \frac{\beta_k}{\lambda_k \Lambda} P_0 \\ T_f(0) &= T_c(0) + \frac{P_0}{Ah} \\ T_c(0) &= T_{in} + \frac{P_0 L}{M_c c_{pc} v} \end{aligned}$$

The input data needed by the simplified model (collected in Table 2.1) were obtained using the ERANOS suite (Rimpault et al., 2002) for the neutronics parameters and using (Waltar and Reynolds, 1981) as a reference for the heat removal parameters.

2.2. Adjoint Sensitivity Analysis applied to multi-physics problems

P (MW)	1800	Λ (s)	4×10^{-7}
M_f (kg)	9675	M_c (kg)	1168
c_{pf} (J/kgK)	500	c_{pc} (J/kgK)	1200
Ah (kW/K)	2.5×10^6	v (m/s)	7.5
α_d (pcm/K)	-0.687	α_c (pcm/K)	0.123
α_c (pcm/K)	0.123		

Table 2.1: Parameter values used in the coupled model.

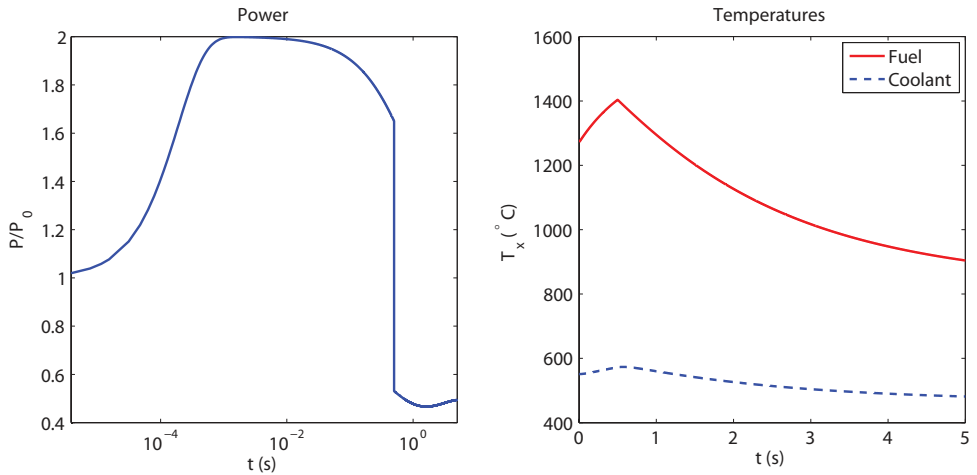


Figure 2.1: Reference transient describing the response of the system to a step reactivity insertion. The value of the step reactivity is $\rho_{ext} = 0.5\%$ for $0 < t \leq 0.5s$ and $\rho_{ext} = -1\%$ for $t > 0.5s$.

Starting from these initial conditions, we consider a transient triggered by a positive reactivity insertion. This external reactivity is described by a step function which is positive (0.5%) between the initial time and $t = 0.5$ and a negative (-1%) for ($t > 0.5s$). This reference problem has been solved by using a built-in ODE solver in Matlab. As it can be seen in the solution illustrated in Fig. 2.1, the power starts to increase after the reactivity insertion until the Doppler reactivity feedback is sufficiently strong to compensate the external amount. The introduction of the negative step (after 0.5s) then brings

the system to a power level lower than the initial value. Regarding the fuel temperature, this negative insertion represents the point when it reaches its maximum (with the coolant temperature following with a small delay).

This maximum temperature has been chosen as the response considered for the following perturbation analysis. We limited this analysis to five parameters: the two reactivity coefficients, the total heat transfer coefficient Ah , the inlet temperature T_{in} and the magnitude of the positive reactivity insertion ρ_{ext} .

The first order perturbation with respect to this set of parameters can be calculated by solving the following adjoint system

$$\begin{aligned}
 -\frac{dP^*}{dt} &= \frac{\rho_{ext} - \beta}{\Lambda} P^* + \frac{\alpha_d}{\Lambda} [T_f - T_f(0)] P^* + \\
 &\quad + \frac{\alpha_c}{\Lambda} [T_c - T_c(0)] P^* + \sum_{k=1}^6 \frac{\beta_k}{\Lambda} C_k^* \\
 -\frac{dC_k^*}{dt} &= -\lambda_k C_k^* + \lambda_k P^* \\
 -\frac{dT_f^*}{dt} &= \frac{\alpha_d}{\Lambda} P P^* + \frac{P^*}{M_f c_{pf}} + H \left(\frac{T_c^*}{M_c c_{pc}} - \frac{T_f^*}{M_f c_{pf}} \right) \\
 -\frac{dT_c^*}{dt} + v \frac{T_c^*}{L} &= \frac{\alpha_c}{\Lambda} P P^* + H \left(\frac{T_f^*}{M_f c_{pf}} - \frac{T_c^*}{M_c c_{pc}} \right)
 \end{aligned}$$

This adjoint system is subjected to the final conditions introduced in Eq. 2.4. As it has been explained earlier, the linear operator of the adjoint problem depends on the reference nonlinear solution. This means that, since the time integration of the adjoint problem is performed backwards in time (Williams, 1986), one would need to store this reference solution for every time step used during the adjoint integration procedure. When dealing with large scale problems the storage needed to build the adjoint operator would eventually become too large, therefore several techniques are usually employed in order to overcome this issue (van Rooijen and Lathouwers, 2008; Gilli et al., 2011). The left part of Figure 2.2 contains the first order sensitivity coefficients for this set of parameters calculated by means of the first order adjoint theory. The interpretation of these coefficients is quite straightforward and does not need an extensive discussion: increasing the heat transfer coefficient reduces the maximum fuel temperature during the transient while increasing the inlet temperature or the magnitude of the reactivity step would have the opposite effect. At the same time a stronger Doppler coefficient reduces the maximum

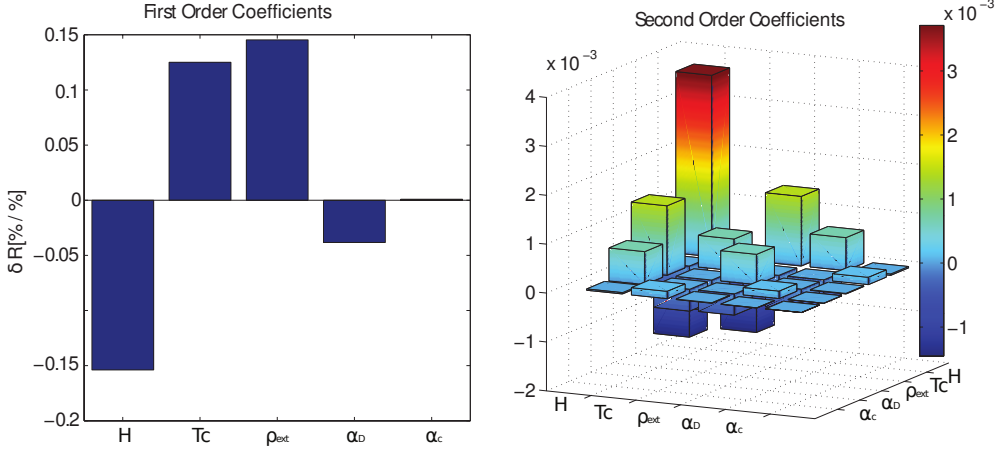


Figure 2.2: First (left) and second (right) order sensitivity coefficients for the fuel temperature at ($t = 0.5s$). The coefficients are expressed in terms of relative variation of the output corresponding to a relative variation of the inputs.

power reached during the transient (and therefore the temperature), while the contribution of the coolant coefficient is negligible.

For the calculation of the Hessian matrix of the response with respect to these 5 parameters, 10 further calculations are required (5 perturbed forward problems and 5 perturbed adjoint calculations). The second order adjoint calculations are performed, according to Eq. 2.7, by solving a set of new adjoint systems obtained by adding different source terms to the first order adjoint operator. For example, the second order adjoint solution with respect to the Doppler coefficient is determined by solving the second order adjoint system characterized by the following source term

$$S_{\hat{P}} = \frac{\alpha_d}{\Lambda} \delta T_f P^* + \frac{\alpha_c}{\Lambda} \delta T_c P^* + \frac{\delta \alpha_d}{\Lambda} [T_f - T_f(0)] P^* \quad (2.13)$$

$$S_{\hat{C}_k} = 0$$

$$S_{\hat{T}_f} = \frac{\delta \alpha_d}{\Lambda} P P^* + \frac{\alpha_d}{\Lambda} \delta P P^*$$

$$S_{\hat{T}_c} = \frac{\alpha_c}{\Lambda} \delta P P^*$$

which depends both on the first order adjoint solution (P^*) and on the linearized one (δT_f and δT_c). As a consequence, the computational requirements of the method further increase since these solutions need to be stored with an adequate time resolution. In presence of large scale systems the required storage space would therefore become too large and, as for the first order adjoint solution, the implementation of special numerical techniques would be necessary in order to solve this issue.

The right part of Figure 2.2 shows the Hessian matrix calculated by substituting these solutions into Eq. 2.5. If a numerical approximation (like Finite Difference) was used instead of the adjoint method to calculate the coefficients of the Taylor expansion one would need $N^2 + 2N + 1$ runs of the full model (Ronen, 1988) in order to estimate the first and second order sensitivity coefficients. This is in sharp contrast with the $2N + 2$ calculations needed by the adjoint theory while dealing with a response localized in time. However, for large perturbation sets the computational requirement would become eventually too expensive considering that each response analyzed would need a set of new calculations.

The matrix in Figure 2.2 defines the second order correction to the response variation given any input perturbation. The diagonal elements of the matrix correspond to the second order perturbation components caused by single input variations while the off-diagonal ones represent the cross-correlation between the different parameters. For example, the effect of perturbing the Doppler coefficient will be higher if the heat transfer coefficient is perturbed at the same time. This example shows that higher order perturbation components can be dominated by a limited set of input parameters, or in other words, that the evaluation of the full Hessian matrix would involve calculating many terms which are close to zero. This is in principle the concept behind the definition of reduced subspace modeling techniques, as the one presented by Bang and Abdel-Khalik (2012).

The problem when applying higher order perturbation theory would be determining "a priori" whose elements (and therefore calculations) can be discarded. Unfortunately, the comparison between the Hessian and the first order coefficients shows that in some cases the magnitude of second order perturbation components is not determined by the magnitude of the first order derivatives. For example in Figure 2.2 it can be seen that the second order sensitivity with respect to the inlet temperature is much larger than the one

with respect to the initial external reactivity, despite the fact that the first order sensitivities of the response with respect to these two input parameters are roughly the same. This can be explained considering the mathematical form of the second order adjoint problem: in presence of nonlinear systems the second order source term contains both the first order adjoint solution and a forward perturbation, as in Eq. 2.13. This means that the second order solution is not uniquely defined by the first order adjoint solution but it also depends on the outcome of a forward perturbation theory. A consistent ranking criterion would be therefore more difficult to implement. This, together with the computational issues associated to the adjoint solution for nonlinear problems, makes the application of second order adjoint techniques to large scale problems very challenging.

2.3 Conclusion

In this chapter we presented a general overview about adjoint methods. We also introduced the derivation of a useful implementation of a second order adjoint based perturbation technique for multi-physics time-dependent problems. The mathematical definition of this perturbation technique, based on the Adjoint Sensitivity Analysis Procedure, is very similar to the first order counterpart, the main difference being the increased computational cost required to reconstruct the Hessian matrix of a response. Applying second order adjoint perturbation techniques to nonlinear systems requires the solution of $2N$ additional problems (where N is the number of input parameters). We demonstrated from a mathematical point of view that, this requirement can be further reduced in the presence of linear time-dependent problems for which it is possible to obtain the second order information with N additional adjoint calculations. This result is consistent with other applications of second order adjoint theories performed in the past for different types of linear systems.

With the application of the technique to a reference problem we have shown the basic mathematical aspects associated with adjoint methods. The computational challenges linked to adjoint formulations are quite considerable, however the derivation can be useful for a better understanding, from a mathematical point of view, of uncertainty propagation problems.

In the next chapter, we are going to discuss about another family of deter-

2. *Adjoint sensitivity analysis methodologies*

ministic methods for uncertainty quantifications. These methods, known as spectral techniques, are based on the spectral representation of the stochastic quantities involved within our problem. The overview presented in the next chapter, will be followed by the derivation of a new spectral technique.

SPECTRAL TECHNIQUES FOR UNCERTAINTY QUANTIFICATION

3.1 Introduction

In this chapter we introduce and discuss the main concepts associated to spectral techniques for Uncertainty Quantification. These include the mathematical background of the methods together with some approaches that can be used for its application. These approaches can be divided into two main categories depending on their formulation. Intrusive techniques involve the definition of a new differential problem that has to be solved to apply the spectral theory, while non-intrusive methods can be used to collect spectral information by sampling the original stochastic problem.

Both methods and the underlying spectral theory will be introduced within the next section. This is followed by two illustrative applications. First both methods are applied to a non-linear time-dependent problem and their outcome compared to standard Uncertainty Quantification techniques. Then, a non-intrusive spectral approach is applied to a criticality benchmark characterized by the presence of a large number of random inputs.

3.2 Basic aspects of spectral techniques for uncertainty quantification

Spectral techniques are based on the idea of representing a generic stochastic process by means of a spectral expansion in random space. This spectral representation can be used to reproduce any of the processes involved in a mathematical model defining a physical problem of interest. There are many ways to implement this expansion depending on the type of basis used, one of the most known is the Polynomial Chaos Expansion (PCE) defined using multidimensional orthogonal polynomials. The basis for this spectral representation was first introduced by Wiener (1938) (as the Homogeneous Chaos) in order to reproduce Gaussian processes in a spectral fashion. Xiu and Karniadakis (2002) later presented a family of orthogonal polynomials, known as the Wiener-Askey scheme, that can be used to extend the theory to different families of stochastic processes.

The first step to implement the spectral expansion from a mathematical point of view is the introduction of a group of independent and identically distributed random variables $\boldsymbol{\xi} = [\xi_1(\theta), \dots, \xi_N(\theta)]$ (where θ is the domain of the stochastic problem) that can be used as a support for the construction of the spectral basis. The size of this variable set, N , is defined as the minimum number of independent random variables needed to describe the random input data set of the problem, which corresponds in most cases to the number of random input parameters. The distribution of these random variables is arbitrary and in principle one should choose it such that an optimal representation of the input data set is achieved (for example by using the same distribution family of the random inputs).

Assuming that these variables have been defined, it is possible, according to the Polynomial Chaos theory, to approximate any stochastic quantity (if characterized by a finite standard deviation) using a truncated expansion. For example, a generic stochastic quantity $R(\theta)$ can be expanded as

$$R(\boldsymbol{\theta}) = \sum_{i=0}^P r_i \Psi_i(\boldsymbol{\xi}(\boldsymbol{\theta}))$$

where a set of multidimensional polynomials Ψ_i (which are functionals since they depend on $\boldsymbol{\xi}$) belonging to the Wiener-Askey scheme is used, and P is the

number of terms in the expansion. This number depends on the highest polynomial degree used in the spectral representation. If we define the maximum polynomial order as p , the number of terms present in the spectral expansion is (Le Maitre and Knio, 2010)

$$P + 1 = \frac{(N + p)!}{N!p!}$$

It must be stressed that the maximum polynomial order p is the only approximation introduced by the PCE, in fact, choosing a hypothetical infinite order would generate an infinite expansion capable of representing any stochastic process of interest (provided that the process itself is characterized by a finite mean and standard deviation).

The most important characteristic for a polynomial basis Ψ_i to be usable in spectral techniques is its orthogonality property with respect to the following definition of the inner product

$$\langle \Psi_n, \Psi_m \rangle \equiv \int_{\Theta} \Psi_m(\boldsymbol{\xi}) \Psi_n(\boldsymbol{\xi}) w(\boldsymbol{\xi}) d\boldsymbol{\xi} = h_n^2 \delta_{mn} \quad (3.1)$$

where w is a weight function (measure) which depends on the polynomial expansion used, h_n a normalization constant, and Θ is the support of the stochastic domain. In (Xiu and Karniadakis, 2002), it is shown that some of the polynomials belonging to the Wiener-Askey scheme are more suitable than others when representing particular random distributions: Hermite polynomials, for example, are the best way to represent Gaussian random variables while Legendre polynomials are more suitable for uniform distributions. This implies that, if the proper polynomial basis is chosen, the representation of the random input data of the problem can be obtained by using the least number of terms in the spectral expansion.

When dealing with multi-dimensional random problems, the orthogonal basis is constructed by performing a tensorization of one dimensional polynomials. After introducing the multi-index $\gamma = (\gamma_1, \dots, \gamma_N)$ it is possible to define the following set

$$\lambda(k) = \left\{ \gamma : \sum_{i=1}^N \gamma_i = k \right\} \quad (3.2)$$

3. Spectral Techniques for Uncertainty Quantification

which can be used to formulate the k^{th} order polynomial expansion as

$$\Psi_k = \left\{ \bigcup_{\gamma \in \lambda(k)} \prod_{\gamma_1}^{\gamma_N} \psi_{\gamma_i}(\xi_i) \right\}$$

It must be pointed out that, since the multi-dimensional polynomial is obtained as a tensorization of one-dimensional functions, it is possible to adopt different polynomial families along different directions of the random problem without compromising the orthogonality condition (3.1).

Following the definition of PCE, we can represent any of the stochastic quantities involved in the Uncertainty Quantification problem, for example unknowns and responses, by using their spectral representations. Applying spectral techniques means determining the coefficients of the spectral expansion for the output quantities of interest once the spectral representations of the inputs are known. For any of these parameters it is possible to evaluate some of the statical moments of interest by using the following expression (Le Maitre and Knio, 2010)

$$\mathbf{E}[\mathbf{R}] = r_0$$

for the mean and

$$\sigma[\mathbf{R}] = \sqrt{\sum_{i=1}^P \langle \Psi_i, \Psi_i \rangle r_i^2} \quad (3.3)$$

for the standard deviation. The expression for the skewness is relatively more complicated. If we perform a similar manipulation to the one used to derive the standard deviation we can obtain the following expression

$$\gamma[\mathbf{R}] = \frac{1}{\sigma^3} \sum_{i,j,k=1}^P \langle \Psi_i, \Psi_j, \Psi_k \rangle r_i r_j r_k \quad (3.4)$$

which involves a triple summation over the PCE terms. This means that in presence of multi-dimensional problems an expression for the skewness would be non-trivial to implement. The probability density function of $\mathbf{R}(\theta)$ can also be obtained quite easily by sampling the spectral expansion itself, or, in other words, by sampling the independent random variables ξ which characterize the spectral expansion.

Different methodologies have been developed and applied so far for the evaluation of the output spectral coefficients, the main distinction being between intrusive and non-intrusive approaches (Le Maitre and Knio, 2010). The application of intrusive approaches requires the definition of a new mathematical problem whose solution is the set of spectral coefficients of the system unknowns. This new mathematical system is significantly larger than the original one and it is usually solved by using a separate solver. On the other hand, non-intrusive techniques are applied by collecting a set of realizations of the original system, using the original model as a “black box”. In the next section more details about the intrusive and the non-intrusive approaches are introduced.

3.2.1 Stochastic Galerkin Approach

Known as the Stochastic Galerkin approach, the following method provides a way to evaluate the coefficients of the spectral expansion based on a direct projection of the mathematical model defining the physical problem of interest. We consider, as an illustrative example, the time-dependent model introduced in Eq. 1.1. Firstly, we replace all the random quantities of the system (inputs and unknowns) with their spectral representations

$$\frac{d\left(\sum_{i=0}^P \mathbf{u}_i \Psi_i(\xi)\right)}{dt} = \mathbf{L}\left(\sum_{i=0}^P \alpha_i \Psi_i(\xi), \sum_{i=0}^P \mathbf{u}_i \Psi_i(\xi)\right) \quad (3.5)$$

Taking the inner product of this system with respect to each of the multi-dimensional polynomials defined by Eq. 3.2, leads to the definition of a new system of ODEs for each of the unknown coefficients

$$\langle \Psi_j, \Psi_j \rangle \frac{du_j}{dt} = \left\langle \mathbf{L}\left(\sum_{i=0}^P \alpha_i \Psi_i, \sum_{i=0}^P u_i \Psi_i\right), \Psi_j \right\rangle \quad (3.6)$$

The system of coupled ODEs generated by this operation corresponds to a system which is $P + 1$ times larger than the original one, obtained as the coupling between $P + 1$ subsystems. Each of the subsystems is associated to the j 'th coefficient of the spectral expansion of the unknowns.

As mentioned before, this dimension depends both on the number of random variables introduced and on the order of the polynomial approximation, hence

3. Spectral Techniques for Uncertainty Quantification

the solution may be problematic in presence of a multi-dimensional input data set. Furthermore, a large effort is also required from the implementation point of view since a new differential problem is introduced. In general, the system defined in Eq. 3.6 is rather complex because it involves operations between stochastic expansions. If, for example, we were dealing with a linear problem, the model defined by Eq. 3.6 would involve products between the expansions used to approximate the random parameters and the ones representing the stochastic unknowns. This would correspond to the following double summation

$$\alpha(\xi)u(\xi) = \sum_{i=0}^P \sum_{j=0}^P \alpha_i u_j \Psi_i(\xi) \Psi_j(\xi) = f(\xi)$$

The projection of this expression onto the k 'th polynomial can be reformulated as

$$f_k = \sum_{i=0}^P \sum_{j=0}^P \alpha_i u_j \mathbf{C}_{ijk} \quad (3.7)$$

where \mathbf{C}_{ijk} is a three dimensional tensor defined as

$$\mathbf{C}_{ijk} = \frac{\langle \Psi_i \Psi_j \Psi_k \rangle}{\langle \Psi_k \Psi_k \rangle}$$

This tensor represents the coupling which is present within the model 3.6 between different coefficients of the spectral expansion of the solution. Moreover, when dealing with nonlinear problems more complicated forms could arise. For example, in presence of first order nonlinearities the introduction of the spectral expansions would generate for each nonlinear term a triple product between a stochastic input quantity and two stochastic unknowns. This triple product can be rewritten, in a fashion similar to the one adopted for linear terms, as

$$[\alpha(\xi)u_1(\xi)u_2(\xi)]_m = \sum_{j,k,l=0}^P T_{jklm} \alpha_j u_{1k} u_{2l} \quad (3.8)$$

where the four dimensional tensor

$$\mathbf{T}_{jklm} = \frac{\langle \Psi_j \Psi_k \Psi_l \Psi_m \rangle}{\langle \Psi_m \Psi_m \rangle}$$

is used. This kind of operation has to be repeated for each term of the original system defined in Eq. 1.1. Although using these tensor forms facilitates the

projection process, one can easily understand how in presence of complex systems the definition of the Stochastic Galerkin system would become extremely complicated. For this reason, alternative ways which can be used to evaluate the coefficients of the spectral expansion without defining a new mathematical problem have been studied. These approaches are known as non-intrusive since their application does not involve any modification of the original mathematical problem. In the next section an overview of these approaches is presented.

3.2.2 Non-Intrusive Spectral Projection

Non-intrusive approaches rely on the orthogonality property of the PCE. Using this property it is possible to express the i^{th} coefficient of the PC expansion (as the one presented in Eq. 3.1) by projecting the stochastic quantity of interest onto the respective polynomial. For example, the i^{th} coefficient of the generic stochastic quantity R can be determined by performing the following projection

$$r_i = \frac{\langle R, \Psi_i \rangle}{\langle \Psi_i, \Psi_i \rangle} \quad (3.9)$$

If R was, for example, an unknown or a response of the mathematical model, solving the projection integral would require the knowledge of its integrand along the integration domain. Knowing the behavior of the integrand corresponds to determining the solution of the output R at different points of the Θ space. This operation can be in principle performed from the numerical point of view by solving the problem for a finite set of realizations of the input parameters. In this way the computer code which solves the mathematical problem could be used as a "black box" to determine the coefficients of the PCE. Problems arise when increasing the dimension of the integral. In this case the number of realizations required in order to have a meaningful estimate for the integral eventually becomes too large.

For high dimensional integrals, one may approximate the numerator of Equation 3.9 by using a Monte Carlo integration technique, randomly sampling the model inputs until statistical convergence for the integral is reached. Using an unbiased Monte Carlo integration technique the estimate for the projection integral would be

$$r_i \approx \frac{1}{N_{tot}} \frac{\sum_{j=1}^{N_{tot}} R(\xi_j) \Psi_i(\xi_j)}{\langle \Psi_i, \Psi_i \rangle} \quad (3.10)$$

3. Spectral Techniques for Uncertainty Quantification

where N_{tot} is the number of realizations collected and ξ_j their stochastic coordinates. The advantage of a Monte Carlo integration is that it is simple to implement even in presence of high-dimensionality problems but the disadvantage is given by the slow convergence rate which, in case of unbiased sampling, (reference to previous chapter) is proportional to $N^{-0.5}$. Furthermore, the presence of a polynomial of order i within the projection integral will also affect the convergence rate since it depends on the smoothness of the integrand. This implies that the truncation order p of the polynomial expansion will heavily affect the number of samples required to achieve statistical convergence for the PCE.

The deterministic alternative for the evaluation of the previous integral is the introduction of numerical integration techniques as quadrature formulae. In the case of a one dimensional random problem, the numerator of Equation 3.9 can be approximated using the following sum

$$\int \mathbf{R}(\xi_1) w_1(\xi_1) d\xi_1 = \sum_{i=1}^{nq} w_1^i \mathbf{R}(\xi_1^i) = Q_{lev}^1 \mathbf{R}$$

where ξ_1^i and w_1^i are respectively the quadrature points and weights associated with the one dimensional quadrature formula employed, and nq is the number of points. This number determines the accuracy of the quadrature formula which corresponds to a level index lev . The definition of these levels is arbitrary since its main purpose is to distinguish formulae with different orders of accuracy, see for example Figure 3.1 where different accuracy levels for a Gauss-Hermite rule are shown. In this case, increasing the accuracy level corresponds to using two more points within the quadrature rule. In the present work we limit our analysis to Gaussian quadrature rules (Abramowitz and Stegun, 1964) which can integrate with great accuracy many of the polynomial families defined by the Wiener-Askey scheme, however, a complete overview of univariate quadrature formulae is provided in reference (Le Maitre and Knio, 2010). The main characteristic of Gaussian quadrature formulae is that they can integrate exactly the corresponding polynomials up to order $2nq - 1$. In case of a multi-dimensional random input set, quadrature rules can be easily extended

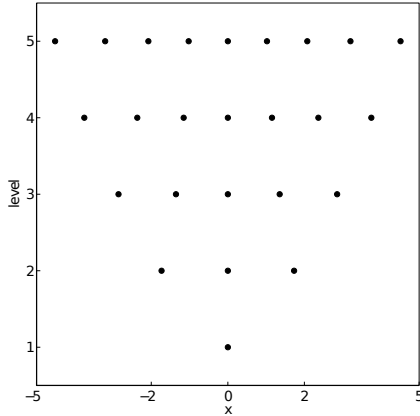


Figure 3.1: Quadrature points for different levels using a Gauss-Hermite rule.

by using a tensor product

$$Q^{(N)} \mathbf{R} = (Q_{lev}^1 \otimes \dots \otimes Q_{lev}^N) \mathbf{R} = \quad (3.11)$$

$$\sum_{i_1=1}^{n_{i_1}} \dots \sum_{i_N=1}^{n_{i_N}} \mathbf{R}(\xi_1^{i_1}, \dots, \xi_N^{i_N}) w_{i_1}^{i_1} \dots w_{i_N}^{i_N} \quad (3.12)$$

where it is assumed, for sake of simplicity, that the same quadrature formula Q_{lev} is used in each direction. The multivariate quadrature rule obtained by this direct tensorization has, along each direction, the same accuracy of the univariate formula used to construct it. This means that the multivariate formula will be exact for products of 1D polynomials for which univariate quadratures are exact. For example, a 2 dimensional quadrature rule obtained by tensorizing a univariate formula exact up to order 2 will be exact for multidimensional polynomials up to order 4, as long as these multidimensional polynomials are obtained as the tensorization of two order 2 polynomials. Also, the same quadrature formulae will not have the same accuracy for 2D polynomials of order 4 if they are obtained by using 1D polynomials of order greater than 2.

For tensor product formulae the number of nodes required to build the quadrature grid is equal to $(nq)^N$ where nq is the number of quadrature points

in each direction and N the dimension of the random space. It is clear that for high dimensional problems the method would have a computational cost comparable or even larger than a standard Monte Carlo technique, making the approach less appealing. These requirements can be explained as a consequence of the fact that quadrature tensorization formulae are, as mentioned before, also accurate for a limited set of higher order polynomials. Since this information is not used within the determination of the spectral coefficients it is in principle possible to reduce the number of points of tensorization formulae by discarding this higher order accuracy from the quadrature set. This concept led to the definition of sparse tensorization techniques for the construction of multivariate quadrature rules, known as sparse grid techniques, first introduced by the Russian mathematician Smolyak (1963) who defined an algorithm for sparse tensor product constructions which will be introduced in the following chapter.

3.2.3 Comparison between the different approaches

The spectral techniques introduced so far can be regarded as a way to achieve the exactness of Monte Carlo sampling in a deterministic fashion. Unfortunately, the computational cost required for the application of such methods increases with the dimension of the random problem, this drawback is usually known as the "curse of dimensionality" (Le Maitre and Knio, 2010). For example, it has already been mentioned that when using tensor product quadrature formulae the number of points for highly dimensional problem would increase exponentially, however various techniques can be adopted in order to tackle this problem. Furthermore the approach is relatively easy to apply from a numerical point of view, since it involves the collection of realization points obtained by solving the original problem as if it were a "black box". On the contrary the Stochastic Galerkin problem involves the definition of a new system which is $P + 1$ times larger than the original one, where P directly depends on the dimension and on the polynomial order used. This means, for example, that in the presence of 10 independent random variables a fourth order Stochastic Galerkin would involve the definition and the solution of a problem 1000 times larger than the original one. In Figure 3.2 it is shown the behavior of the number of spectral coefficients P with respect to the stochastic dimension and the truncation order. The figure is useful to visualize the fact that when dealing with large scale systems the application of the Stochastic Galerkin approach would become extremely unpractical.

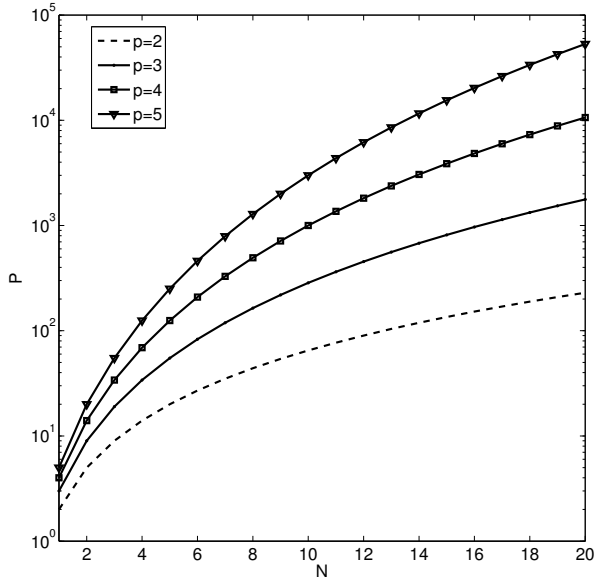


Figure 3.2: Size of the spectral expansion P with respect to the polynomial truncation order p and the dimension of the stochastic problem N .

Within the next two sections the spectral techniques introduced so far will be applied to two illustrative examples. Firstly, the Non-intrusive Spectral Projection and the Stochastic Galerkin approaches are going to be employed to perform Uncertainty Quantification of a nonlinear time-dependent problem. Secondly, a highly dimensional stochastic problem is introduced and the feasibility of a Non-intrusive approach is analyzed.

3.3 Application to an illustrative time-dependent problem

We now present the application of the techniques introduced so far to an illustrative problem. This example is used to show the differences between the two main spectral approaches and to compare them to standard UQ techniques. The problem considered is the same system of ODEs introduced in Section

2.2.3, used to model the time-dependent behavior of a simplified reactor system.

We considered, for the present example, the application of spectral and traditional uncertainty quantification techniques to a reference transient. This transient is the response of the system to a step reactivity insertion of 1\$. The uncertainty problem around this reference solution has been introduced considering the inserted reactivity ρ_{ext} and the Doppler coefficient α_D to be stochastic quantities (normally distributed with 5% and 10% relative standard deviation respectively). The data from Table 2.1 was used for their mean values.

In both cases the random inputs are assumed to be independent. This kind of distribution has been chosen in order to compare the results with the outcome of the first order uncertainty propagation despite the fact that the infinite realization support of the distribution could in theory lead to the generation of "undesired" transients (corresponding to a change of sign of the external reactivity or feedback coefficients). However, due to the small standard deviation of the input quantities, all the realizations needed to build the quadrature formulae for the NISP and the coupling tensors for the SG were within the desired limits.

The two main spectral approaches defined in the previous section have been implemented in order to solve this stochastic problem. In addition, the uncertainty problem has been solved by using a standard unbiased Monte Carlo sampling method and by implementing an adjoint based sensitivity technique based on Equation 2.3. The outcomes of these two additional methods will be compared with the results obtained by using spectral techniques.

The solution of the reference transient and the application of the UQ techniques have been performed using *MATLAB* and its built-in solvers. In order to implement the NISP and the Stochastic Galerkin methods we used a PCE expansion based on 2-dimensional Hermite polynomials. This expansion is the best choice in case of Gaussian processes since it can fully describe them within the first two terms. A fifth order polynomial expansion was used for both spectral techniques, this truncation order has been chosen in order to verify the convergence of the spectral expansion. The non-intrusive projection was obtained using the tensor product quadrature formula introduced in Eq. 3.12; the Galerkin problem derived for the intrusive approach is described in the following section.

3.3.1 Stochastic Galerkin formulation of the nonlinear point-kinetic model

In the previous section we explained that a new differential problem needs to be defined when applying the Stochastic Galerkin approach. We start this derivation by approximating the stochastic input quantities using the following spectral expansions

$$\begin{aligned}\rho_{ext} &= \sum_{i=0}^1 \rho_{ext}^i \Psi_i(\boldsymbol{\xi}) \\ \alpha_D &= \sum_{i=0}^1 \alpha_D^i \Psi_i(\boldsymbol{\xi}) \\ \alpha_c &= \sum_{i=0}^1 \alpha_c^i \Psi_i(\boldsymbol{\xi})\end{aligned}$$

where the sum is truncated after the first order term assuming that the quantities are Gaussian and that we are using Hermite polynomials. The same is done for the stochastic unknowns

$$\begin{aligned}Q &= \sum_{i=0}^P Q_i \Psi_i(\boldsymbol{\xi}) \\ C_k &= \sum_{i=0}^P C_{k,i} \Psi_i(\boldsymbol{\xi}) \\ T_f &= \sum_{i=0}^P T_{f,i} \Psi_i(\boldsymbol{\xi}) \\ T_c &= \sum_{i=0}^P T_{c,i} \Psi_i(\boldsymbol{\xi})\end{aligned}$$

The implementation of the Stochastic Galerkin formulation is done by substituting these expansions into equation 2.12 and projecting the resulting system on the p 'th polynomial of the basis. The first complication arises from the product between the stochastic external reactivity and the power

$$\rho_{ext} Q = \sum_{i=0}^1 \sum_{j=0}^P \rho_{ext}^i Q_j \Psi_i(\boldsymbol{\xi}) \Psi_j(\boldsymbol{\xi})$$

3. Spectral Techniques for Uncertainty Quantification

the projection on the p 'th polynomial of this expression is, using equation 3.7,

$$(\rho_{ext}Q)_p = \sum_{i=1}^P \sum_{j=0}^P \rho_{ext}^i Q_j \mathbf{C}_{ijp}$$

where we used the three dimensional tensor \mathbf{C}_{jlp} introduced before. Similarly we can project the nonlinear reactivity feedback term using the four dimensional tensor \mathbf{T}_{jlkp} as

$$(\alpha_D T_f Q)_p = \sum_{j=0}^1 \sum_{k,l=0}^P T_{jklp} \alpha_D^j T_{fk} Q_l$$

the tensors \mathbf{C}_{jlp} and \mathbf{T}_{jlkp} define the coupling between each of the coefficients of the spectral expansion. After some small manipulations, the final stochastic system can be written as

$$\begin{aligned} \frac{dQ_p}{dt} &= \sum_{j=0}^1 \sum_{l=0}^P \mathbf{C}_{jlp} \left[\frac{\rho_{ext}^j - \alpha_D^j T_f(0) - \alpha_c^j T_c(0)}{\Lambda} \right] Q_p + \\ &\quad + \sum_{k=1}^6 \left[\frac{\beta}{\Lambda} Q_p + \lambda_k C_{k,p} \right] + \\ &\quad + \sum_{j=0}^1 \sum_{l=0}^P \sum_{k=0}^P \mathbf{T}_{jlkp} \left[\frac{\alpha_D^j}{\Lambda} T_{f,l} Q_k + \frac{\alpha_c^j}{\Lambda} T_{c,l} Q_k \right] \\ \frac{dC_{k,p}}{dt} &= \left[\frac{\beta_k}{\Lambda} Q_p - \lambda_k C_{k,p} \right] \\ M_f c_{pf} \frac{dT_{f,p}}{dt} &= Q_p + Ah(T_{c,p} - T_{f,p}) \\ M_c c_{pc} \left[\frac{dT_{c,p}}{dt} + v \frac{T_c - T_{in}}{L} \right] &= Ah(T_{f,p} - T_{c,p}) \end{aligned}$$

If we assume the initial conditions to be deterministic we will have the same initial value as 2.12 for $p = 0$ and zero initial amplitude for the rest of the coefficients. In case of stochastic initial conditions the initial value for each coefficient p needs to be evaluated by performing the usual projection. The system just defined has been solved by using the same MATLAB solver used for the reference solution.

3.3.2 Results and Discussion

The results of the application of the different techniques introduced so far are presented in this section. The analysis was done for the transient and the associated stochastic problem described before. First, the mean value and the standard deviation of the reference solution, generated by considering the Doppler coefficient and the external reactivity to be stochastic quantities, have been calculated by using the Stochastic Galerkin formulation and the Non-intrusive Spectral Projection for a time ranging from 0 to 10 seconds. Figure 3.3 shows the time-dependent mean values and the associated uncertainty bands (given by adding and subtracting the standard deviation) of the reactor power and the system temperatures obtained by using the non-intrusive spectral projection approach. The figure shows that the large magnitude of the standard deviation causes a large uncertainty band around the power peak. The same solution calculated by using the Stochastic Galerkin approach did not present any substantial difference, only a small discrepancy due to the considerable difference between the two approaches was found. The first important thing to point out is that, since the probability density function of the output quantities can be non-symmetric, the mean value does not correspond to the solution obtained using the expectation values of the input data, as we would assume while performing a first order uncertainty propagation. A second remark is that obtaining these solutions with an adjoint sensitivity technique would require the evaluation of the adjoint problem at many points in time during the transient, thus increasing the computational cost. On the contrary, this is much easier with Monte Carlo and PCE techniques since the evolution of the time-dependent behavior of the uncertainty is a direct outcome of the method. While adjoint methods should be applied to estimate the uncertainties for few time-localized responses, spectral methods can be used to determine the most sensitive points during a transient.

The choice of using a fifth order truncated spectral expansion has been made in order to perform a convergence analysis of the solution with respect to the polynomial approximation used. Figure 3.4 shows the time dependent standard deviation of the power for the stochastic problem obtained using different expansion orders. A fourth order expansion was required to accurately calculate the statistical moments in every point of the time domain. However such a higher order expansion is necessary only to calculate the standard deviation for the time-region located around the maximum power peak, afterwards the

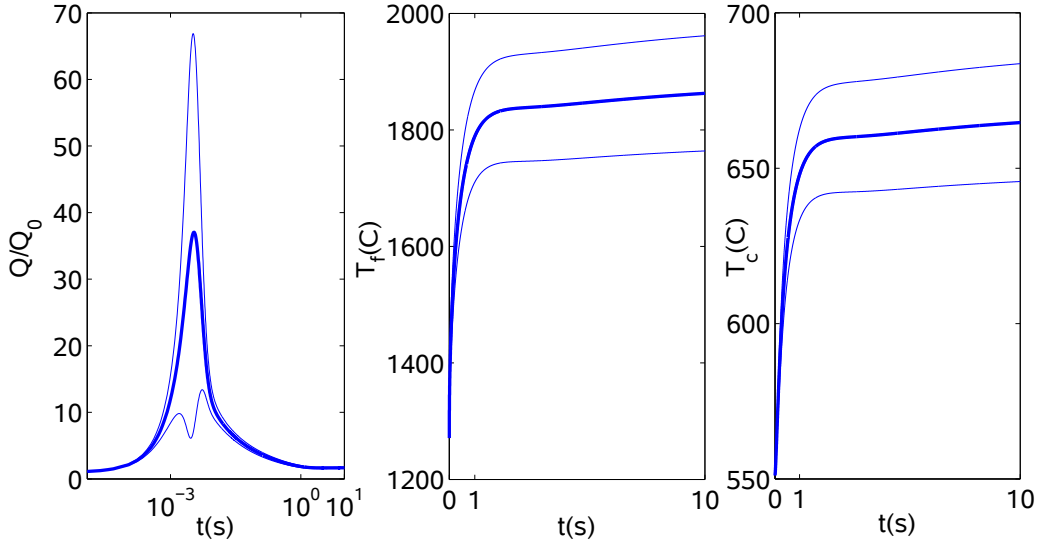


Figure 3.3: Mean value and uncertainty bands caused by the uncertainty on the reactivity insertion ρ_{ext} and on the Doppler coefficient α_d . Power plotted in a semi-logarithmic scale.

solution relaxes due to the feedback mechanisms and the discrepancy with respect to the first order approximation reduces considerably.

The computational requirements of both spectral methods are directly related to the PCE order employed, however it would be in principle possible, while applying a Non-intrusive Spectral Approach, to use different PCE truncation orders at different instants of the transient, whereas the size of the Stochastic Galerkin problem would be determined by the maximum order needed during the transient. In general, it is known that spectral representations of time-dependent stochastic processes may be affected by convergence issues (Wan and Karniadakis, 2006) requiring an increasing number of polynomial terms for their accurate representation during the time-evolution. Different techniques to overcome this problem have been proposed and they generally require a more complicated implementation (Geritsma et al., 2010) (Wan and Karniadakis, 2005). In this particular example, the presence of feedback mechanisms causes

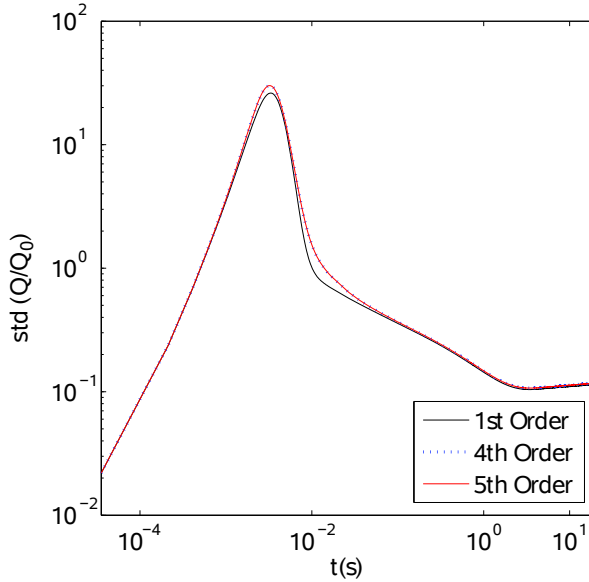


Figure 3.4: Time-dependent standard deviation of the power for the stochastic problem calculated with different PCE truncation orders

the solution to grow relatively slowly during the time interval following the power peak. For time-dependent problems reaching a steady state, PCE techniques have been proved to converge even for long transient behaviors (Hagues et al., 2010). In this case the feedback nature of the problem and the relatively short time interval make the PCE to converge to an accurate solution. Tables 3.1 and 3.2 contain the mean value and the uncertainty associated with the power and the fuel temperature at two different instants: in the proximity of the power peak ($t = 0.0034s$) and at $t = 1s$, when the standard deviation of the fuel and the coolant temperatures starts to increase. The table also includes the results obtained with the two PCE techniques (NISP and SG) together with the values calculated using a standard Monte Carlo sampling and the ASAP method. The NISP technique has been applied using two different quadrature grids (7×7 and 15×15). The right column contains the number of calculations needed to apply the method, the letter l means that the calculation is linear (2 linear calculations are needed to apply

3. Spectral Techniques for Uncertainty Quantification

	resp. = $Q/Q_0(-)$					
	t = 0.0034s		t=1s		N. Calc.	time [s]
	E	σ	E	σ		
NISP(7)	35.6535	29.5391	1.8603	0.1477	49	29
NISP(15)	35.6529	29.5293	1.8603	0.1477	225	133
ASAP	26.4778	18.0783	1.8405	0.1357	1+2 <i>l</i>	10
SG	35.6531	29.5395	1.8603	0.1477	1*	180
MC	35.6571	29.5212	1.8603	0.1477	10000	5400
	± 0.004	± 0.01	$\pm 5e-5$	$\pm 4e-5$		

Table 3.1: Mean and standard deviations (obtained with different techniques) of the power at 2 different instants ($t = 0.0034s$ and $t = 1s$). The columns on the right contain the number of calculations and the computational time needed to apply the method, the letter l means that the calculation is linear while the star (*) means that a system 20 times larger than the original one has been solved in order to apply the SG approach.

	resp. = $T_f(C)$					
	t = 0.0034s		t=1s		N. Calc.	time [s]
	E	σ	E	σ		
NISP(7)	1290.6895	12.5270	1789.0554	78.6719	49	29
NISP(15)	1290.6895	12.5276	1789.0554	78.6731	225	133
ASAP	1287.2495	8.4011	1779.5634	72.7214	1+2 <i>l</i>	10
SG	1290.6895	12.5271	1789.0553	78.6735	1*	180
MC	1290.6895	12.5276	1789.0501	78.6811	10000	5400
	$\pm 9e-2$	± 0.0025	± 0.018	± 0.015		

Table 3.2: Mean and standard deviations (obtained with different techniques) of the fuel temperature at 2 different instants ($t = 0.0034s$ and $t = 1s$).

the ASAP technique, 1 for each of the responses considered) while the star (*) indicates that a system 20 times larger than the original one has been solved in order to apply the SG approach. The calculation of the standard deviation at $t = 1s$ needed approximately 10 seconds using the ASAP technique, 29s using the 7×7 NISP approach, 180s using the SG and 1.5 hours using the unbiased Monte Carlo sampling.

As the results show, implementing the NISP using a 7 points quadrature instead of a 15 points leads to a very small error in the prediction of the statistical

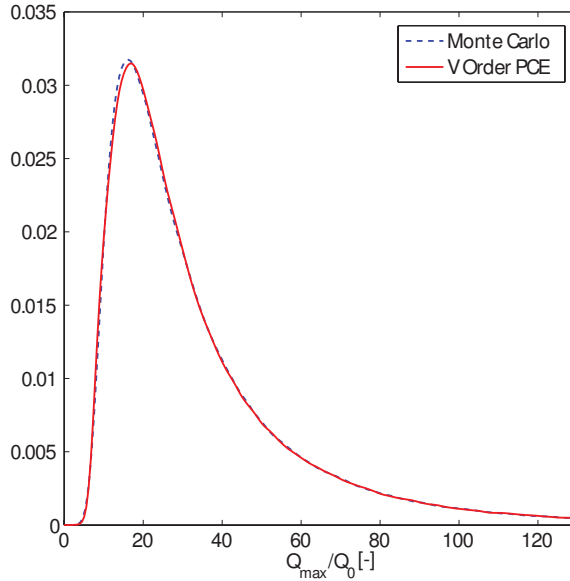


Figure 3.5: Probability density function of the peak power. Comparison between between PCE (continuous) and Monte Carlo (dashed).

moments. This proves the good convergence of the Gaussian quadrature rule, which is able to successfully solve the projection integral with a relatively small number of points, even in the presence of a higher order stochastic problem.

The values obtained with both spectral techniques compare very well to the Monte Carlo estimators, the first methods being much faster due to the small number of random parameters. This agreement can be also verified for the probability density functions of the problem, for example Figure 3.5 shows the pdf associated to the power at $t = 0.0034s$, obtained with the PCE approach and with a standard Monte Carlo sampling. It is apparent to see how the shape of this pdf is far from the Gaussian distribution (in other words it requires more terms for its spectral representation) which is the shape assumed by first order propagation techniques. The Tables show how, in fact, the first order approximation gives a considerable error in the prediction of many of the responses considered. The main reason for this is the large input uncertainty which places the problem in the region where the linear assumption ceases to be valid. The adjoint method largely fails to predict the uncertainty of the peak

3. Spectral Techniques for Uncertainty Quantification

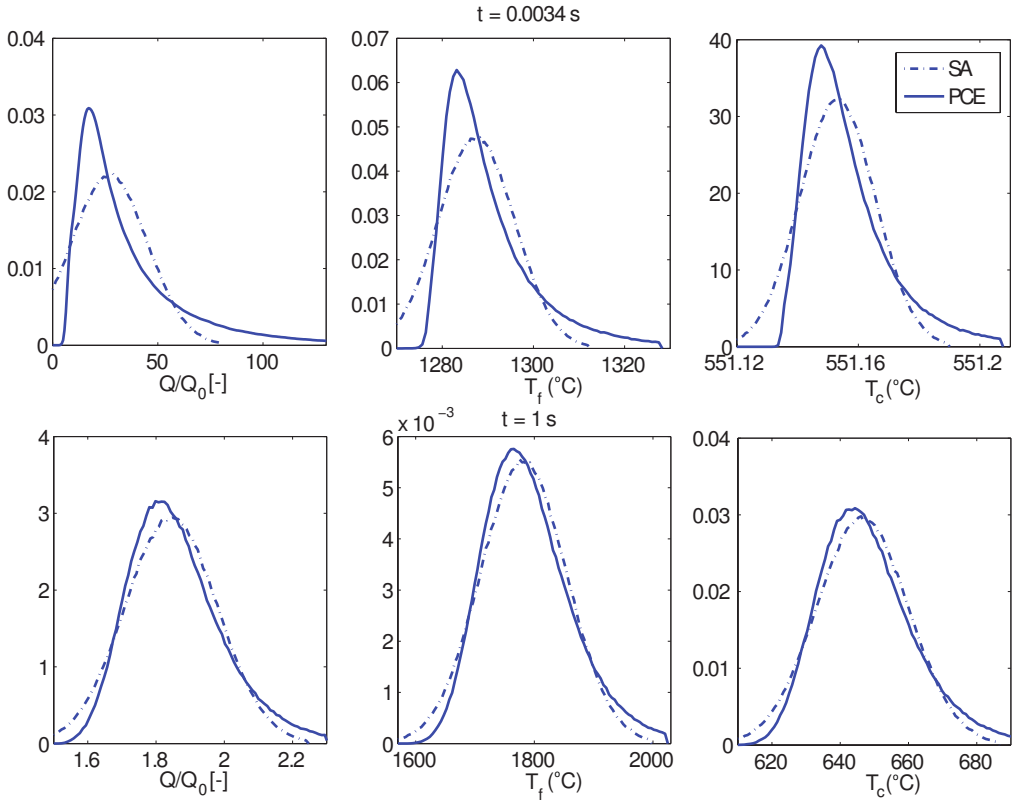


Figure 3.6: Comparison of the pdfs associated to the system unknowns obtained with a PCE expansion (continuous line) and with a first order adjoint technique (dashed line).

power. This discrepancy is shown in Figure 3.6 where the probability density functions obtained using the NISP approach are compared with the Gaussian prediction resulting from the adjoint method. This example clearly shows how, in some cases, the first order approximation introduced by perturbation techniques is not suitable to perform UA. On the contrary PCE techniques are able to capture higher order stochastic information in a deterministic way. This type of probability density profile could have important consequences from a probabilistic risk assessment point of view. In our example the probability density function for the fuel temperature has a tail toward higher temperature

values, which means that the probability of it being higher than a safety limit we may be interested in (for example, the melting temperature) is larger than the one calculated using a Gaussian distribution.

In general, we have shown with this example how spectral techniques can be employed to perform Uncertainty Quantification for problems associated with large nonlinearities. The application of the intrusive approach required the definition of a relatively complicated system despite the fact that the original model was fairly simple. Moreover, for time-dependent problems, its computational requirement directly depends on the maximum PCE truncation order needed during the transient but, as we have seen for this example, this higher order truncation is sometimes required only during a small part of the time-domain. On the contrary non-intrusive approaches are relatively easy when applying spectral methods to large scale problems since they can be regarded as an extension of standard sampling techniques. Furthermore their definition allows using different PCE orders at different instants of the transient, possibly reducing the computational cost of the method.

3.4 NISP for a large scale problem using MC

The aim of this second example is to show how spectral techniques can be used when dealing with highly dimensional stochastic problems. We have seen in the previous sections that both the quadrature based Non-intrusive Spectral Projection and the Stochastic Galerkin formulation have a maximum number of stochastic input parameters corresponding to which their application would become computationally too expensive. This number also depends on the spectral convergence of the random quantities of interest, nevertheless there will always be a computational threshold represented by the size of the random problem. When this threshold is reached, the only feasible way to perform uncertainty analysis would be to use a standard Monte Carlo approach, as discussed in Chapter 1.

By definition, using an unbiased Monte Carlo sampling technique for uncertainty quantification corresponds to the evaluation of a set of integrals associated to the statistical moments of interest. As we discussed before, calculating the estimator for the i^{th} coefficient of the spectral expansion introduced in Eq. 3.10 does not differ from a mathematical point of view. The only

difference is that the integrand is now multiplied by one of the polynomials of the PCE. For example, the numerical solution of the projection problem associated to the first term of the expansion (obtained by using the constant polynomial $\Psi_0 = 1$ in Eq. 3.10) is equivalent to the classical unbiased statistical estimator of the mean. If we are already using a Monte Carlo approach to perform uncertainty quantification we could therefore use the same realizations to evaluate the integrals characterizing the spectral expansion.

In theory, the evaluation of these PCE coefficients would not require any modification of standard sampling approaches, a part from the reformulation of the integrand. In general, because we are using a Monte Carlo integration approach for the evaluation of the spectral coefficients, the technique will be characterized by a slow convergence rate. Moreover, this rate will decrease when dealing with higher order coefficients because of the higher differentiability of the polynomial included within the integrand.

Assuming that the standard unbiased sampling approach has converged to a meaningful estimate, it might be of interest to quantify the amount of spectral information that is possible to evaluate by using the set of realizations already collected. For this purpose, we defined an illustrative example based on a criticality problem. A benchmark, taken from the handbook of criticality calculations (NEA, 2010), was selected to perform uncertainty propagation. This criticality benchmark, known as Jezebel (PU-MET-FAST-001), can be described as a highly enriched bare Plutonium sphere. The problem is characterized by a fast neutron spectra and its multiplication factor was calculated using the MCNP5.1 code (3000 neutron histories were used to evaluate the eigenvalue, corresponding to a statistical error of around 20pcm).

As an uncertainty propagation problem, we considered the average number of neutrons generated per fission ($\bar{\nu}(E)$) to be a random quantity affected by a 5% error. The objective of the analysis was to determine the influence of the introduction of this stochastic component on the fundamental eigenvalue of the system. The average number of neutrons per fission is an energy dependent quantity, each energy entry of ACE file used by MCNP has been assumed to be normally distributed and independent. This generated a stochastic problem defined by 2366 random inputs (corresponding to the number of discretized energy levels in the ACE file) which is, as discussed before, certainly too large for the application of the quadrature based non-intrusive spectral projection or the Stochastic Galerkin formulation. A standard unbiased sampling was first

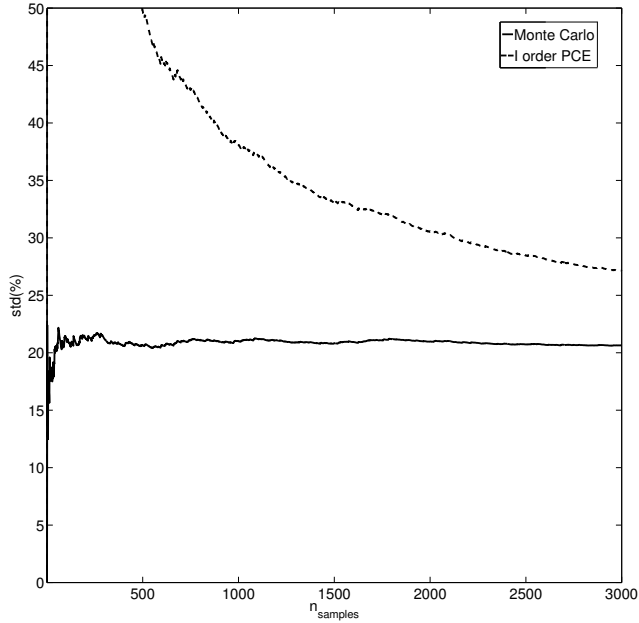


Figure 3.7: Convergence of the standard deviation evaluated using a standard sampling technique (continuous line) and a first order PCE (dashed line).

performed by collecting 3000 realizations of the multiplication factor. Then, a second order PCE (built by using Hermite polynomials) was introduced in order to spectrally represent this stochastic quantity. The coefficients of the PCE have been calculated by using Eq. 3.10 together with the 3000 realizations of the output already collected for the standard sampling. Figure 3.7 shows the convergence of the standard deviation obtained by using the standard sampling approach compared to the outcome obtained by using the spectral representation. It is clear from the figure that the estimate of the spectral method has not converged, unlike the one evaluated with the unbiased sampling. This can be easily explained since the error due to the Monte Carlo integration present in each of the spectral coefficients is added up in the final value of the standard deviation. This means that when the size of the stochastic problem is large, using a Monte Carlo approach for the evaluation of the spectral coefficients will eventually fail to predict the value of the statistical moments of interest within reasonable error.

However the figure does not tell much about the error which is present in the prediction of each of the PCE coefficients. Figure 3.8 shows the complete set of first order coefficients (corresponding to the whole energy range) for different sizes of the sample set used to evaluate the integrals. The peak present around higher energy ranges is expected since the system is characterized by a fast neutron spectrum, changing the number of neutrons produced by fission at the thermal energy range has therefore negligible influence on the fundamental eigenvalue. In the bottom right frame the outcome of the non-intrusive spectral projection is compared to a standard multilinear regression approach. In the latter case the error is much smaller, however a minimum of 2381 realizations (corresponding to one sample more than the number of random inputs) were required to evaluate the regression coefficients. Although the error present in the PCE coefficients is considerable, we can still obtain a meaningful prediction of their final value even with relatively small sample sets (as we can see, this peak is already noticeable with 500 samples).

Figure 3.9 shows a single PCE coefficient taken from the first order set (corresponding to $\bar{\nu}(E = 200\text{MeV})$, in the proximity of the peak) together with its associated statistical error. In this case the value of the coefficient has already converged to a reasonable estimate after 500 realizations of the output. This means that for the example considered it is possible to obtain first order derivative information for the most sensitive parameters of the problem with a reasonable amount of samples (a number comparable with the one required for the unbiased sampling method to converge). Since its application is cost free (assuming we are already applying a standard sampling approach), the non-intrusive spectral approach would represent a good way to obtain additional information of stochastic quantities of interest for large scale random problems.

3.5 Conclusions

In this chapter we presented an introduction to the most common approaches that can be used to apply spectral techniques for Uncertainty Quantification. Through illustrative examples, we have seen that the intrusive approach, known as the Stochastic Galerkin formulation, presents some challenges in terms of formulation and computational requirements that would be hard to tackle when dealing with large scale problems. On the contrary, non-intrusive approaches

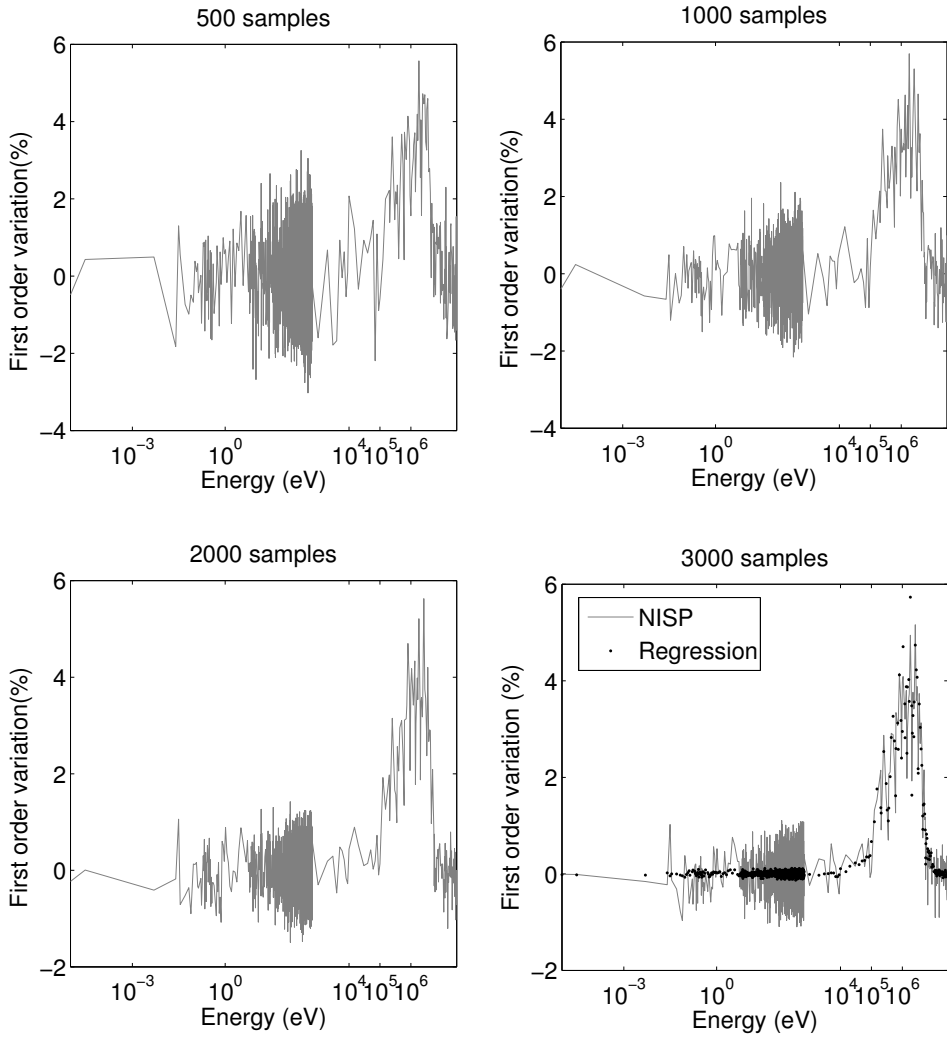


Figure 3.8: Convergence of the first order coefficients of the PCE with respect to the number of realizations collected. The bottom right frame contains the prediction for the same coefficients obtained by using a multilinear regression approach (based on 3000 samples).

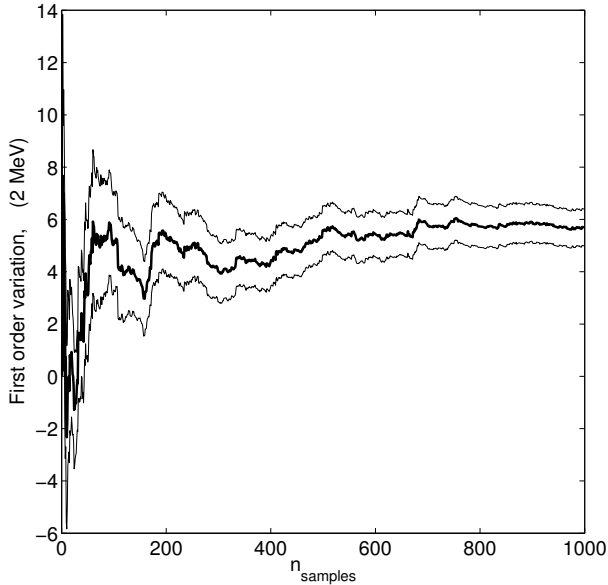


Figure 3.9: Convergence of the first order PCE coefficient corresponding to $E = 200\text{MeV}$. Thin lines represent the statistical error associated to the estimator.

are intrinsically easier to apply since they can be regarded as a way to perform an "ordinate" sampling of the stochastic quantities of interest.

We have shown with the second example how the Monte Carlo formulation could also represent a useful tool when dealing with large scale problems, providing an alternative way to estimate sensitivity coefficients. Moreover the method can be easily integrated with standard random sampling approaches.

Regarding the quadrature based Non-intrusive Spectral Projection, though the definition and the implementation are simpler than the intrusive counterpart, we discussed how they are heavily affected by the "curse of dimensionality" which makes non-intrusive methods defined by the tensorization of univariate formulae not applicable for most of the stochastic problems. Fortunately, it is possible to use more convenient quadrature formulae based on the definition of sparse grids. In the following chapter, we will focus on this type of quadrature

rule and on its optimization performed by developing an adaptive algorithm.

CHAPTER 4

A NEW ADAPTIVE ALGORITHM FOR THE APPLICATION OF NISP TECHNIQUES

4.1 Introduction

In the previous chapter we have seen how Non-Intrusive Spectral Projection approaches can potentially represent an optimal way to apply spectral techniques due to their simplicity and ease of formulation. However, standard tensorized quadrature techniques are affected by the "curse of dimensionality" which makes the technique often inapplicable when increasing the size of the stochastic problem. The requirements of tensor product formulae are caused by the fact that quadrature tensorization formulae are also accurate for a limited set of higher order polynomials which are not used within the projection process. It is therefore in principle possible to reduce the number of points of tensorization formulae by discarding this higher order accuracy from the quadrature set. This concept led to the definition of algorithms for the construction of multivariate quadrature rules, known as sparse grid techniques, first introduced by the Russian mathematician Smolyak (1963).

In this chapter we use this concept to develop an adaptive algorithm which can be used to build a sparse grid quadrature rule. In this case the adaptivity can be employed to achieve a further reduction of the computational requirements associated with sparse grid formulae. The main concepts behind the development of this adaptive algorithm are presented after a general introduction about sparse grids. This algorithm is derived taking into account the fact that a complex system is usually dominated by main effects and low-order interactions. From an uncertainty propagation point of view, this means, as we discussed in Chapter 2, that a response will be largely influenced by a limited set of input parameters and by lower order perturbation components.

The derivation of the algorithm is then followed by an illustrative application to a reference mathematical problem. In this example, the outcomes of the adaptive algorithm are compared to the results obtained with standard sparse grid rules.

4.2 Introduction to sparse grids

In this section we present a brief introduction to the concept of sparse grids, needed to define the adaptive algorithm further on. The first step for the definition of a sparse tensorization algorithm is the introduction of the following difference formula

$$\begin{aligned}\Delta_{lev}^{(1)}f &\equiv \left(Q_{lev}^{(1)} - Q_{lev-1}^{(1)}\right) f \\ Q_0^{(1)}f &\equiv 0\end{aligned}$$

where $\Delta_l^{(1)}f$ is a univariate quadrature formula defined as the difference between two consecutive quadrature levels lev . The set of abscissae of the difference formula presented in the previous equation is defined by the union of the abscissae of the two different quadrature formulae while the weights are obtained as their difference. Because of this definition, the difference formula corresponding to the $lev + 1$ level contains all the points defined within the previous level lev .

After the introduction of the difference formula, the second step to implement the sparse tensorization algorithm is the introduction of the multi-index set $\mathbf{l} = (l_1, \dots, l_N), \in \mathbb{N}^N$ which is a vector as large as the number of stochastic

directions, whose indexes l_i are used to associate an accuracy level to each direction. Furthermore, we define the norm of this multi-index as

$$|\mathbf{l}| \equiv \sum_{i=1}^N l_i$$

Using these definitions it is possible to write the sparse grid construction algorithm using the following summation notation (Le Maitre and Knio, 2010)

$$Q_{lev}^{(N)} f \equiv \sum_{|\mathbf{l}| \leq lev + N - 1} \left(\Delta_{l_1}^{(1)} \otimes \dots \otimes \Delta_{l_N}^{(1)} \right) f \quad (4.1)$$

according to which the final quadrature rule is built by adding a set of sub-grids, whose dimension d depends on the accuracy level used along each direction. Each sub-grid $\left(\Delta_{l_1}^{(1)} \otimes \dots \otimes \Delta_{l_N}^{(1)} \right)$ corresponds to a tensorization of a set of difference formulae whose levels are defined by the value of the multi-index \mathbf{l} . The set of multi-indexes having norm $|\mathbf{l}| \leq lev + N - 1$ (where lev is the accuracy level one wants to achieve for the multi-dimensional integral) is therefore used to generate the sub-grids used in the final rule. An important consequence of this norm constraint is that the maximum dimension for an admissible sub-grid is $lev - 1$ which is relatively low even when considering high dimensional problems. This is the main aspect determining the numerical convenience of a sparse grid with respect to a tensorized one.

The use of this sparse algorithm can reduce drastically the amount of quadrature points required, especially when dealing with high-dimensionality problems. Unfortunately, despite the drastic reduction of quadrature points associated with sparse quadrature rules, for high-dimensional problems the set of points (or, in other words, realizations) required to perform an accurate integration could still become larger, for the same accuracy, than the one required by standard Monte Carlo approaches. The number of quadrature points generated by the sparse tensorization depends directly on the univariate formula used to build it, however as it is explained in (Le Maitre and Knio, 2010) in presence of irregular integrands and highly dimensional problems the size of the grid would eventually become too large.

Several approaches, usually based on adaptive algorithms, have been presented in literature in order to overcome this issue. For example, adaptive techniques have been implemented by using multi-wavelet based approaches

(Le Maitre et al., 2004) or stochastic Finite Elements (Bieri and Schwab, 2009) (Witteveen et al., 2009). In these cases, the possibility of using adaptive and goal-oriented methods was successfully demonstrated, however their complexity makes their application difficult for problems characterized by large numbers of random inputs. Usually, dimension reduction techniques such as the ANalysis Of VAriance (Yang et al., 2012) are employed in order to transform large dimensional problems in a set of lower dimensional models which can be solved by using the adaptive techniques just mentioned (Foo et al., 2008).

An adaptive approach based on sparse grid quadrature was presented by Gerstner and Griebel (1998) who demonstrated that it is possible to adaptively build cubature rules in order to solve high dimensional integrals like the ones associated with spectral methods. These sparse grids can be also used to determine a set of collocation points which need to be evaluated in order to determine the stochastic outputs of interest (Ganapathysubramanian and Zabararas, 2007).

4.2.1 Adaptive Sparse Grid construction algorithms

The main idea behind the definition of an adaptive sparse grid algorithm is that the constraint acting on the set of multi-indexes used to build the sub-grids in Eq. 4.1 can be relaxed, therefore reducing the number of points used to generate the final quadrature formula. One of the first examples of an adaptive algorithm can be found in reference (Gerstner and Griebel, 1998) where the authors define a sequential way to construct the multi-index set by using the reduction of the integration error as a target. Some of the sub-grids included by a standard sparse grid algorithm might introduce a negligible contribution to the final quadrature formula, depending on the stochastic direction and the quantity which is being integrated.

The reduction of the multi-index set can be performed keeping in mind that Eq. 4.1 represents a sum of difference formulae which implies that a sub-grid can be introduced into the final rule only if all the previous levels of the difference formula have already been included. This translates to the following mathematical condition

$$\mathbf{1} - \mathbf{e}_j \in I, \text{ for } 1 \leq j \leq N, l_j > 1, \quad (4.2)$$

where I is the set of sub-grids already included in the sparse grid, the multi-

index \mathbf{l} corresponds to the sub-grid with respect to which the admissibility is checked and \mathbf{e}_j is the unity vector. According to this constraint, it is possible to build the quadrature rule starting from the sub-grid defined by the first multi-index ($\mathbf{l}_1 = (1, \dots, 1)$) and enriching it by progressively increasing the level of the multi-index along the directions which give the largest contribution to the integral.

Figure 4.1 depicts, as an illustrative example, a passage during the construction of a two dimensional sparse grid. The grid shown in the top part of the figure is enriched by including the subgrids corresponding to the indexes $\mathbf{l}_i = (1, 3), (2, 2)$ (the subgrids G1 and G2 represented in the central part). The resulting grid is shown in the bottom of the figure. The idea is to find, when dealing with high-dimensionality problem, “optimal” quadrature formulae which include only the set of multi-indexes \mathbf{l}_i giving a sensitive contribution to the final integral. This can be done while using a progressive multi-index approach by limiting the multi-index “enrichment” in the directions for which the sub-grid contribution to the final quantities of interest is larger than a specified tolerance. This firstly requires the definition of a marching algorithm to be used while sweeping throughout the set on multi-indexes, and secondly an indicator used to determine when the contribution of a set of subgrids is below the specified tolerance.

In references (Gerstner and Griebel, 1998) and (Gerstner and Griebel, 2003) an adaptive quadrature algorithm is proposed such that the new multi-indexes to be included within the final rule are taken from the neighborhoods of the indexes which have been already included and which cause the largest contribution to the integral. The same algorithm has been applied successfully by Ganapathysubramanian and Zabarar (2007) to propagate uncertainties in fluid mechanics problems.

In the following section we present an alternative algorithm that can be used to implement an adaptive sparse grid when applying non-intrusive spectral techniques for uncertainty quantification. This algorithm is based on the progressive addition of the sub-grids according to their size d . This allows projecting the information already obtained about lower dimensional sub-grids onto higher dimensional ones that still need to be analyzed. Moreover, the algorithm can be used to build a reduced spectral expansion based on the one-dimensional sub-grids, thereby reducing the computational cost associated to the quadrature process.

4. A new adaptive algorithm for the application of NISP techniques

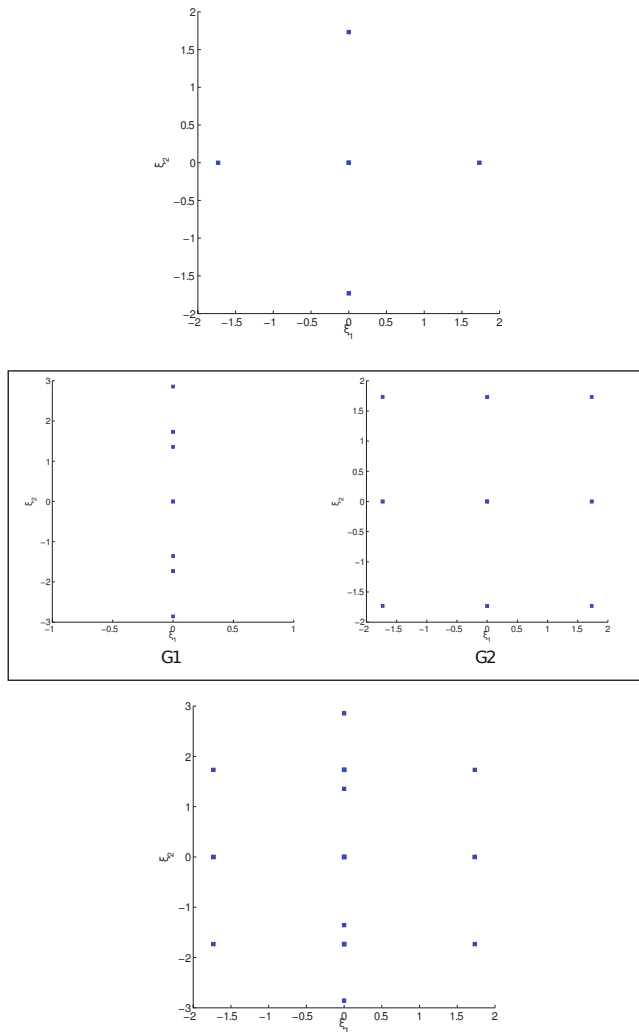


Figure 4.1: Example of a two dimensional adaptive sparse grid construction built by using a Gauss-Hermite univariate rule. The top grid is enriched by including the subgrids G1 and G2 corresponding to the indexes $\mathbf{l}_i = (1, 3), (2, 2)$ (represented in the central part). The resulting grid is showed in the bottom of the figure.

4.3 Definition of a new adaptive algorithm

In this section the adaptive algorithm developed for the application of the non-intrusive spectral projection is presented. The aim of the adaptive algorithm is to successively add admissible sub-grids until a desired integration tolerance is achieved. For this purpose, an indicator needs to be defined in order to determine if the contribution of an admissible sub-grid is small enough for the grid to be discarded. Once this condition is met and the sub-grid is dropped, the multi-index set can be updated by removing all the multi-indeces that do not satisfy the admissibility criterion. This procedure is repeated until all the sub-grids of the initial set have been either included or discarded from the final rule.

In order to implement the adaptive algorithm we first define the following estimator ϵ_i for the response R , associated to the i 'th subgrid

$$\epsilon_i \equiv \max \left[\left| \left(\frac{\delta E_i(R)}{E(R)} \right) \right|, \left| \left(\frac{\delta \sigma_i(R)}{\sigma(R)} \right) \right| \right] \quad (4.3)$$

where δE_i and $\delta \sigma_i$ are, respectively, the variation in the mean value and in the standard deviation of the spectral expansion of R due to the i 'th subgrid. If we keep in mind that each of the subgrids defined by the expansion 4.1 has an associated contribution ϵ_i , the aim of an adaptive algorithm would be to include in the final quadrature rule only the set of subgrids whose associated contribution is larger than a specified tolerance.

The algorithm we introduce is derived by taking into account the fact that a complex system is usually dominated by low-order interactions. This is known as the ‘‘sparsity of effects principle’’ (Wu and Hamada, 2000) and implies that in presence of large sets of input parameters the main influence on the stochastic response is caused by lower order propagation components. Since a sparse grid is defined as the sum of lower dimensional sub-grids (where the maximum dimension is determined by the accuracy level *lev*) the idea of a propagation dominated by lower order interactions is equivalent to stating that most of the contribution to the final integral will be given by lower dimensional subgrids (ex. 1, 2, or 3 dimensional grids).

An adaptive algorithm for the construction of a sparse grid formula has been developed taking these aspects into account. The main idea behind the algorithm is to add subgrids to the sparse grid by starting from one dimensional

grids and by progressively increasing their resolution and their size d . When the contribution ϵ_i associated with the N -dimensional grid i is smaller than a specified tolerance the refinement process is stopped and all the subgrids obtainable as a refinement of i are discarded from the admissible index set. In this context, the integration tolerance tol is the only parameter that needs to be specified.

Given the problem dimension N and a sufficiently large accuracy level (which can be increased if needed), the set of multi-indexes I characterizing the complete sparse grid (defined in Eq. 4.1) is generated. The aim of the algorithm is to select and evaluate from this starting set the sub-grids which cause the largest contribution to the final integral. As we explained before, the sub-grids present in this initial set I can be divided according to their dimension. For each sub-grid this dimension, denoted as d , depends on the univariate quadrature level used in each direction and ranges from 0 to $lev - 1$.

An important aspect to address before continuing with the description of the algorithm regards the truncation of the spectral expansion used to implement the spectral projection, or in other words, the number of terms employed in the PCE. The convergence of the stochastic response R should be verified while applying the adaptive algorithm in order to have a meaningful solution. One could in principle apply the adaptive algorithm by starting from a low order expansion and increasing the order till convergence is reached. In presence of highly anisotropic stochastic problems this procedure would be non-optimal since it would require the evaluation of a PCE whose truncation depends only on the most irregular stochastic direction. A possible way to overcome this problem is the definition of an anisotropic PCE whose higher order terms are included only for a limited set of stochastic directions.

In order to cover this aspect, the adaptive algorithm is initialized as follows. First the algorithm divides all the subgrids present in I according to their dimension d . The only zero-dimensional sub-grids, which corresponds to origin of the stochastic domain defined by the multi-index $\mathbf{1} = (1, \dots, 1)$, is then evaluated. After the subgrids are divided according to their dimension d the algorithm is applied in two separate steps.

Firstly, the adaptive sparse grid formula is enriched by adding exclusively one-dimensional grids and by using a one-dimensional PCE for each of the stochastic dimensions. This allows having an initial estimate of the convergence of the PCE along each direction which can be used to define the order of the

final spectral expansion. Secondly the algorithm is continued by adding higher dimensional subgrids, using the PCE truncation determined within the previous steps. In the next section these two steps are described in more detail.

4.3.1 First step: Integration over 1D sub-grids and definition of a reduced Polynomial Chaos Expansion

During the first part of the adaptive sparse grid generation, the construction of the grid is performed by including quadrature points exclusively along the main axes of the stochastic domain. This is done by introducing a one-dimensional random problem for each of the stochastic random directions and by separately determining the convergence of the corresponding one-dimensional numerical quadrature.

The scope of this convergence analysis is to determine in each direction both the number of quadrature points needed to evaluate the stochastic integral, and the truncation order required for the 1 dimensional PCE to converge.

This corresponds to the following process repeated for each of the stochastic directions. Starting from the initial point associated to the first multi-index $\mathbf{l} = (1, \dots, 1)$ the grid is refined by adding the smallest one-dimensional subgrid taken from the admissible multi-index set (corresponding for example to $\mathbf{l} = (2, \dots, 1)$ when dealing with the first stochastic direction). The indicator introduced in Eq. 4.3 is then evaluated for this grid using a first and a second order PCE. If, after the evaluation of this initial grid, the difference in the standard deviation obtained by using the first and the second order PCE is smaller than the specified tolerance the refinement is stopped and the PCE is assumed to be first order along the direction. If, on the contrary, the contribution of the second order term is not negligible the grid is refined after increasing the PCE truncation order by adding the subsequent subgrids (corresponding for example to $\mathbf{l} = (3, \dots, 1)$, $\mathbf{l} = (4, \dots, 1)$, etc) until the indicator is smaller than the specified tolerance. A convergence test for the one-dimensional PCE is then performed, and, if the spectral expansion has not converged, the previous step is repeated after the truncation order has been increased. The reference flow-chart for this algorithm is presented in Figure 4.2. The outcome of this first phase is an initial collection of subgrids located along the main axes of the stochastic domain (see for example Fig. 4.3) which can be used as a starting basis for the construction of the final quadrature rule. Furthermore, the truncation

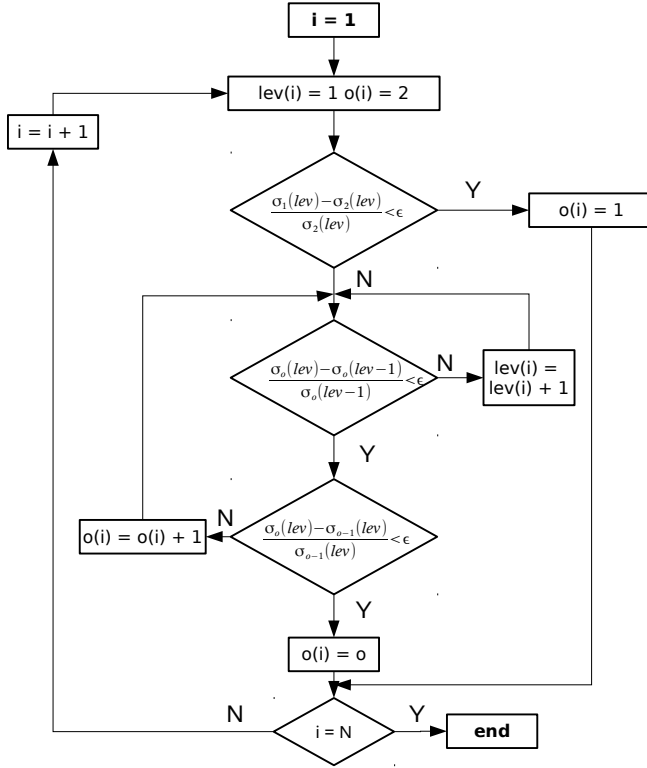


Figure 4.2: Description of the first part of the adaptive algorithm. The index i refers to the stochastic direction while $o(i)$ and $lev(i)$ are, respectively, the corresponding PCE order and grid resolution.

orders required to achieve convergence of the one-dimensional PCEs along each direction are also known after completing this first part. We define a vector $\mathbf{p} = (p_1, \dots, p_N)$ where each entry p_i represents the maximum order needed to represent the stochastic output when considering the corresponding variable to be the only stochastic quantity of the problem. Using this information about the convergence of the spectral expansion along each of the stochastic directions it is possible to define a reduced multi-dimensional PCE by modifying the

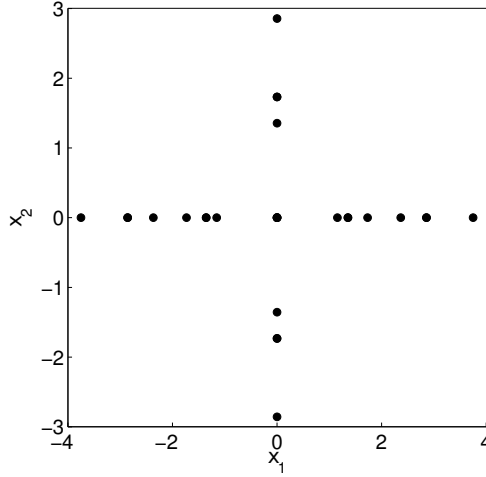


Figure 4.3: Final set of points of the quadrature formula along the main axes for a 2 dimensional case.

constraint on the polynomial basis introduced in Eq. 3.2. We introduce a new definition for the multi-index set λ (originally defined in Eq. 3.2)

$$\lambda(k) = \left\{ \gamma : \left(\sum_{i=1}^N \gamma_i = k \right) \wedge (\gamma_i \leq p_i, \text{ for } 1 \leq i \leq N) \right\} \quad (4.4)$$

The definition of the new multi-index set is similar to the one presented in Eq. 3.2, the only difference being the constraint on the maximum polynomial order along the i' th direction. This order is now determined by the value p_i , obtained while adding the one-dimensional subgrids to the quadrature formula. We assume that the multi-dimensional PCE obtained by using this new constraint is converged and that can be used in the following steps for the construction of the adaptive quadrature algorithm.

4.3.2 Second step: integration over the higher dimensional sub-grids

As we explained when we first discussed the mathematical definition of sparse grid, each of the lower dimensional sub-grids used to build the final rule can be divided according to its size d and norm $|\mathbf{I}|$. When the algorithm described within the previous section is completed, all the one-dimensional sub-grids (characterized by different norms) have been either included or discarded from the final quadrature rule. We can therefore continue the algorithm by considering higher dimensional sub-grids. According to our algorithm, the order these sub-grids are evaluated is firstly determined by their size d and, then, by their norm $|\mathbf{I}|$.

Figure 4.4 shows a scheme representing the order according to which the sub-grids are evaluated. Each block in this figure represents a set of sub-grids characterized by the same size and norm. The blocks in the left column represent the union of all the possible one-dimensional sub-grids (which have been already evaluated within the first part of the algorithm), moving horizontally corresponds to increasing the size of the sub-grids while moving vertically increases their norm (which correspond to a higher “resolution”). After evaluating the set of one-dimensional sub-grids (first column of the scheme) we proceed by analyzing the sub-grids defined by the bottom block of the second column. Once the contribution of each of these two-dimensional sub-grids has been calculated we then continue by moving vertically on the scheme, therefore analyzing the two-dimensional grids associated with a higher norm. After reaching the top block of the column we move to the third column, starting from its bottom block. This procedure is repeated until the top right block has been evaluated.

The adaptivity within this process is implemented by enforcing, each time that a sub-grid is evaluated, the condition defined in Eq. 4.2. This means that if the indicator 4.3 associated with the current sub-grid is below the specified tolerance all the sub-grids that require it for their existence are discarded from the final quadrature rule. This operation introduces another condition for the algorithm to end, which is when every higher dimensional sub-grid have been discarded from the quadrature rule.

It must be kept in mind that each time that a higher-dimensional sub-grid is discarded from the final quadrature rule the polynomials associated with

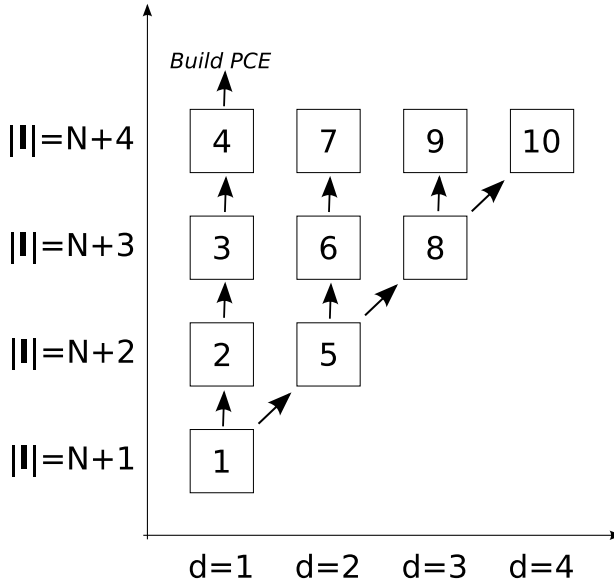


Figure 4.4: Order according to which the different sets of sub-grids are evaluated. Each set contains the sub-grids characterized by size d and norm $||I||$. N is the number of random inputs characterizing the stochastic problem.

the same multi-index must be also taken out from the PCE. This is because the grids required to evaluate the corresponding coefficients are left out from the final quadrature rule.

4.3.3 Overview of the method

In the previous sections we introduced an algorithm which can be used to evaluate the projection coefficients of the spectral expansion in an adaptive fashion. The algorithm is based on the concept of sparse grids and is structured in two main parts: firstly the sparse quadrature points are collected along the main axes of the stochastic domain and a reduced PCE is built, secondly the integration is extended to higher dimensional grids until a convergence for the stochastic quantities of interest is reached.

The convergence rate of this adaptive algorithm depends mainly on how many

grids are discarded from the admissible multi-index set at the initial stages, or, in other words, on the smoothness of the response surface of the problem. If, for example, we deal with a completely linear problem the evaluation of two dimensional grids would already exclude most of the multi-index set from the final quadrature rule. On the contrary an extremely irregular output could in principle require the addition of every single grid defined by the initial set I .

The minimum requirements of the algorithm in terms of quadrature points largely depend on the number of two-dimensional subgrids added during the second phase. Assuming every single input parameter has an influence on the stochastic output of interest, this number can be expressed in the following combinatorial fashion

$$n_{2D} = \binom{N}{2} \quad (4.5)$$

Although the adaptive script can be used to drastically reduce the requirements associated to sparse grid integration, increasing the number of input parameters N will eventually make the method inapplicable because of the starting set of 2 dimensional grids that needs to be evaluated. This number could be in principle used as a rough criterion to estimate whether the application of the adaptive algorithm is feasible (depending on the number of random inputs N).

Depending on the differentiability of the response surface, 15-20 parameters would therefore represent the number above which the method would be computationally more expensive than standard random sampling. This upper constraint could in theory be relaxed by finding a way to discard two-dimensional sub-grids based on the evaluation of the one-dimensional ones.

In theory, the computational cost associated with spectral techniques could be further reduced by exploiting the gradient of the stochastic response of interest with respect to the input parameters. This gradient information could be obtained by using perturbative approaches or, when dealing with functional responses, by implementing adjoint techniques such as the Adjoint Sensitivity Analysis Procedure (Cacuci, 2003). For example, Alekseev et al. (2011) presented a way to use the information associated with the gradient of a functional when dealing with non-intrusive PCE methods.

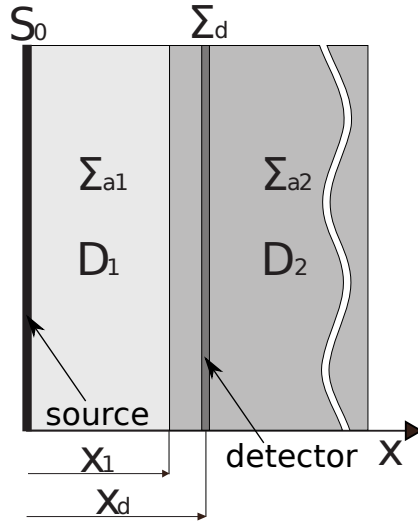


Figure 4.5: Geometric used for the neutron diffusion problem.

4.4 Illustrative example

We now present the application of the adaptive sparse grid algorithm defined in the previous section. We consider a simple one dimensional neutron diffusion problem in a non-multiplying medium, modeled by the following equation

$$(-\nabla \cdot D\nabla + \Sigma_a) \phi(x) = S$$

where D is the diffusion coefficient of the medium, Σ_a the absorption cross-section and S the source term. The geometry analyzed, represented in Figure 4.5, is a semi-infinite one dimensional slab divided into two different regions whose properties are listed in Table 4.1. The boundary between the two regions is located at $x_1 = 20\text{cm}$ while at $x = 0\text{cm}$ a plane source S_0 introduces 10^6 neutrons per second in the system. We also assume that a 1 cm thick neutron detector present in the second region at $x_d = 25\text{cm}$ whose cross-section is $\Sigma_d = 0.01\text{cm}^{-1}$. Finally, we consider the reaction rate of this detector to be the response of our mathematical problem.

$$R = \int_{V_d} \Sigma_d \phi(\mathbf{r}) d\mathbf{r} \quad (4.6)$$

4. A new adaptive algorithm for the application of NISP techniques

where V_d is the volume of the detector.

	Region 1	Region 2				
Σ_a [cm]	$2 \cdot 10^{-2}$	$3 \cdot 10^{-5}$	S_0 [n/s]	10^6	Σ_d [cm^{-1}]	0.01
D [cm^{-1}]	$2 \cdot 10^{-2}$	$3 \cdot 10^{-5}$	x_1 [cm]	20	x_d [cm]	25

Table 4.1: Parameter values used in the diffusion model

We decided to use such a simple model because it is possible to derive an analytical solution which can be used to better evaluate the performance of the spectral method. After solving the diffusion equation, the flux in the second region can be written as (Duderstadt and Hamilton, 1976)

$$\phi_2(x) = A \exp\left(-\frac{x}{L_2}\right)$$

where we introduce the diffusion length $L_i \equiv \sqrt{D_i/\Sigma_{ai}}$. The constant A is equal to

$$A = \frac{S_0 L_1}{2D_1} \frac{D_1 L_2 \cosh\left(\frac{x_1}{2L_1}\right) + D_2 L_1 \sinh\left(\frac{x_1}{2L_1}\right)}{D_2 L_1 \cosh\left(\frac{x_1}{2L_1}\right) + D_1 L_2 \sinh\left(\frac{x_1}{2L_1}\right)}$$

This solution can be used in the integral 4.6 to evaluate our response of interest.

We now introduce a stochastic component to the problem by considering the absorption cross-sections, the diffusion coefficients, the detector cross-section, and the source term to be random quantities. These quantities are assumed to be normally distributed with 10% relative standard deviations. As we explained in Chapter 3 it is possible to represent these quantities by using the following expansions

$$\begin{aligned} S_0(\boldsymbol{\xi}) &= \mu(S_0) + \xi_1 \sigma(S_0) \\ D_1(\boldsymbol{\xi}) &= \mu(D_1) + \xi_2 \sigma(D_1) \\ D_2(\boldsymbol{\xi}) &= \mu(D_2) + \xi_3 \sigma(D_2) \\ \Sigma_{a1}(\boldsymbol{\xi}) &= \mu(\Sigma_{a1}) + \xi_4 \sigma(\Sigma_{a1}) \\ \Sigma_{a1}(\boldsymbol{\xi}) &= \mu(\Sigma_{a1}) + \xi_5 \sigma(\Sigma_{a1}) \\ \Sigma_d(\boldsymbol{\xi}) &= \mu(\Sigma_d) + \xi_6 \sigma(\Sigma_d) \end{aligned}$$

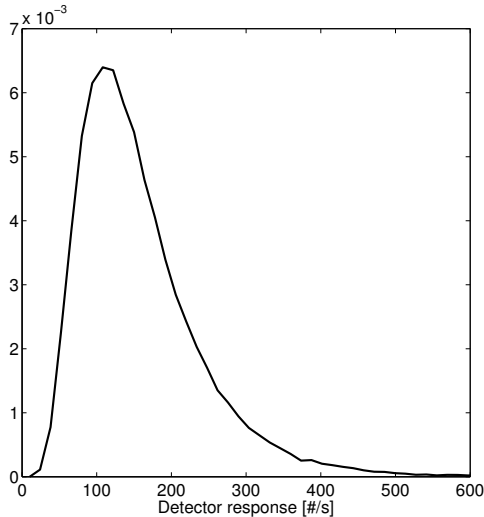


Figure 4.6: Probability density function of the response obtained by using a standard sampling approach (based on 10^5 samples).

where $\mu(\times)$ and $\sigma(\times)$ define, respectively, the mean and the standard deviation of the stochastic quantity and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_6)$ are 6 independent normally distributed (having zero mean and unity standard deviation) variables. Once the random input quantities are defined it is possible to represent any stochastic quantity by using a spectral expansion, in this case obtained by adopting 6-dimensional Hermite polynomials (chosen because of the distribution of the random inputs). In order to have a reference solution for the stochastic problem, the probability density function of the response (shown in Figure 4.6) and its statistical moments were first obtained by using a standard unbiased sampling technique. The adaptive algorithm defined in the first part of this chapter was then applied to this stochastic problem. Applying the adaptive algorithm requires the definition of a relative tolerance ϵ that needs to be used in two separate stages of the process. Firstly this tolerance is used when evaluating the convergence of the PCE, secondly while adding subgrids to the final quadrature formula.

The algorithm (ASG) was first compared with the outcomes of using full sparse grid rules (SG) and standard sampling techniques (MC). Furthermore,

4. A new adaptive algorithm for the application of NISP techniques

the results obtained by applying the adaptive algorithm presented by Gerstner and Griebel (1998) were also included. A fixed order PCE was employed in all cases, in order to have a fair comparison between the different algorithms. The outcome of this analysis is collected in Table 4.2.

	E	σ	N	order
MC	150.8071 ± 0.270	84.4135 ± 2.54	100000	-
SG(lev=4)	150.3001	83.5329	533	4
SG(lev=5)	150.3001	83.5322	2381	4
Gerstner et al. ($\epsilon = 10^{-2}$)	150.3001	83.5324	319	4
Gerstner et al. ($\epsilon = 10^{-3}$)	150.3001	83.5313	399	4
Gerstner et al. ($\epsilon = 10^{-4}$)	150.3001	83.5316	601	4
ASG($\epsilon = 10^{-2}$)	150.3002	83.5087	257	4
ASG($\epsilon = 10^{-3}$)	150.3001	83.5313	399	4
ASG($\epsilon = 10^{-4}$)	150.3001	83.5320	605	4

Table 4.2: Statistical moments of the response obtained using a standard Monte Carlo sampling approach, different sparse grid quadratures, and the adaptive algorithm with different tolerances. Right column contains the PCE order used to construct the spectral expansion.

The table shows that both adaptive techniques converge to the values obtained by using full sparse grid rules and by randomly sampling the model. In all cases the number of points employed by the two different adaptive algorithms is comparable, although the algorithm introduced in this chapter presents a considerable reduction. For the $\epsilon = 10^{-2}$ case this reduction is associated with an increase in the integration error, however this final error is well within the specified tolerance. At this point it must be stressed that all these comparisons were performed by using a fixed PCE order, therefore excluding from the algorithm the benefits of using the reduced PCE introduced in Eq. 4.4.

Table 4.3 contains the results of the application of the algorithm together with the reduced PCE. The last column of the table contains the PCE order \mathbf{p} (along each direction) obtained after the first phase of the algorithm according to the definition introduced in Eq. 4.4.

	E	σ	N	p
SG(lev=4)	150.3001	83.5329	533	4
SG(lev=5)	150.3001	83.5322	2381	4
ASG($\epsilon = 10^{-2}$)	150.3002	83.0095	97	121211
ASG($\epsilon = 10^{-3}$)	150.3001	83.4510	127	121211
ASG($\epsilon = 10^{-4}$)	150.3001	83.5310	333	121421
ASG($\epsilon = 10^{-5}$)	150.3001	83.5318	429	121421
ASG($\epsilon = 10^{-6}$)	150.3001	83.5322	847	133431

Table 4.3: Statistical moments of the response obtained using different adaptive algorithm tolerances. Right column contains the PCE expansion order along each direction.

The results collected in the table show how the number of points needed for the convergence of the algorithm are considerably reduced if compared to Table 4.2. This is of course expected since a lower order spectral expansion is employed within the projection problem. If one compares the outcome of the algorithm corresponding to $\epsilon = 10^{-3}$ and to $\epsilon = 10^{-6}$ it is possible to notice how, though the prediction in the first case is within reasonable error, the number of points is considerably lower. This means that the adaptive component of the algorithm could potentially represent a considerable improvement, in terms of number of quadrature points needed for convergence, if we sacrifice some precision (accuracy which would be however still higher than the one associated to sampling techniques). One should take into account the fact that when performing uncertainty propagation we are usually not interested in determining the statistical moments of interest within a many digits precision, in many cases a 1% error would be more than acceptable.

Overall, these results clearly show how the adaptive algorithm is able to reconstruct the statistical moments of interest, within the specified tolerance, with a considerable lower number of points compared to standard sparse grids. Furthermore the reduced PCE allows neglecting a subset of polynomials whose evaluation would increase the number of quadrature points required. For the present example this reduced expansion captures the fact that the response is linear with respect to the source term (first random variable) and the detector cross-section (last random variable) and therefore higher order PCE terms along these two directions are discarded.

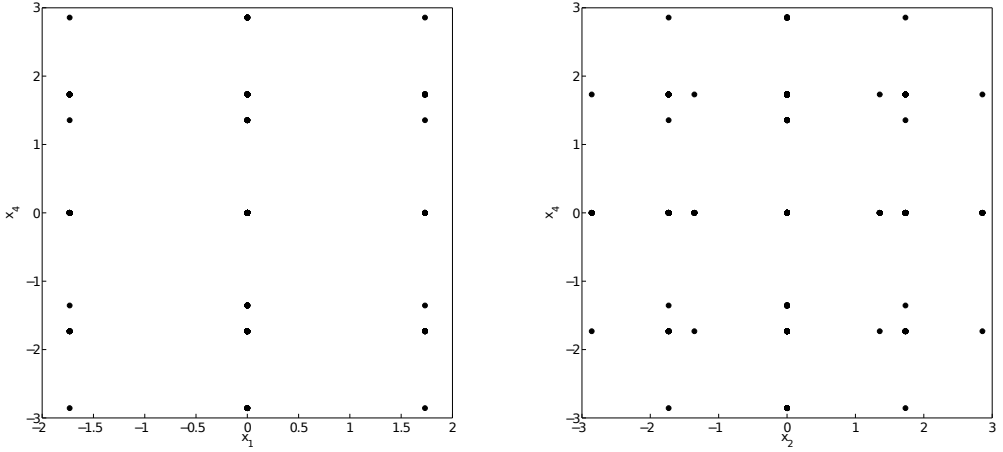


Figure 4.7: Final quadrature points over 2 planes of the stochastic domain obtained by using a $\epsilon = 10^{-3}$ tolerance. Left plane contains the ξ_1 and ξ_4 directions while the right one ξ_2 and ξ_4 .

Figure 4.7 shows the quadrature points obtained by using the adaptive technique and a relative tolerance $\epsilon = 10^{-3}$. The quadrature points shown in this figure are much coarser compared to the other directions, along the first one (corresponding to the source terms) because of the linearity with respect to that particular parameter. Finally, Figure 4.8 compares the probability density functions obtained with the adaptive algorithm, using different tolerances, with the one obtained using the standard sampling technique. The probability density function corresponding to the lowest accuracy of the adaptive algorithm is in very good agreement with the one calculated with Monte Carlo. This means that the error introduced by the adaptive script does not affect the evaluated PCE, on the other hand the number of evaluations needed to build this probability density function is considerably smaller (corresponding about one third of the points).

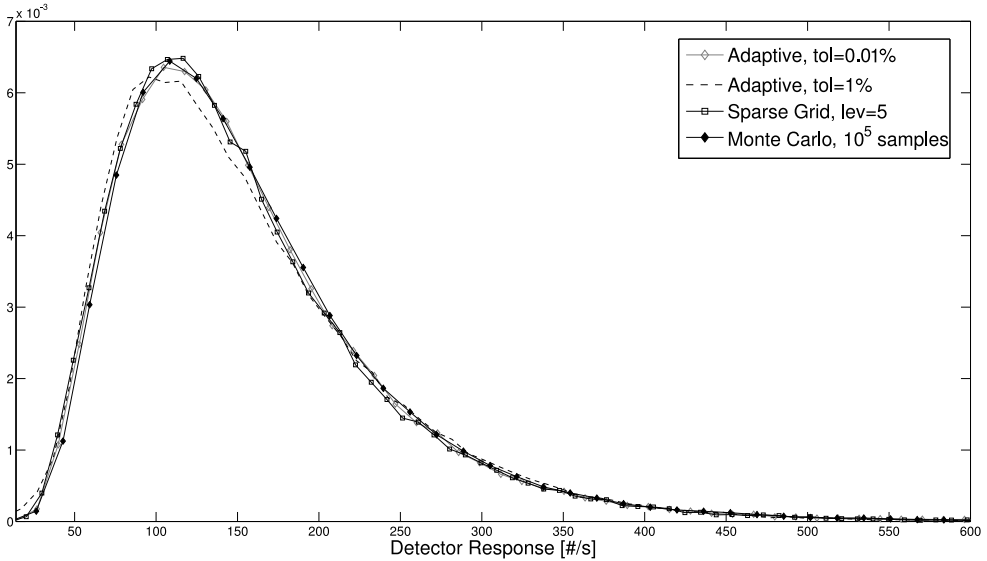


Figure 4.8: Comparison of the probability density function obtained with a standard sampling technique with the one obtained by using the adaptive script with different tolerances.

4.5 Conclusions

In this chapter we introduced a new algorithm for the implementation of an adaptive quadrature rule, based on the notion of a sparse grid. The technique is based on the sequential construction of a sparse grid by the addition of lower dimensional subgrids which are progressively increased in resolution and size. The construction process is divided into two main steps. First, the algorithm adds quadrature points exclusively along the main axes of the stochastic domain. During this phase the convergence of the PCE is assessed and a reduced multi-dimensional PCE is defined. This reduced PCE is then used within the second part of the algorithm which focuses on the addition of higher dimensional sub-grids to the final quadrature rule. The main approximation introduced by this algorithm is that higher dimensional grids are not included in the final rule if obtained as the tensorization of lower dimensional grids associated to a negligible contribution.

The adaptive sparse grid algorithm has been tested for a reference stochastic case defined using a simple source-detector problem. A comparison was also performed between the new algorithm and another one already present in literature. The results obtained by applying the new algorithm show that by acting on the integration tolerance it is possible to strongly reduce the number of quadrature points preserving the stochastic information of the system. When using very small tolerances the final quadrature set built by the adaptive algorithm converges to a standard sparse grid rule, however by relaxing this constraint it is possible to achieve a considerable reduction of the quadrature points preserving the statistical moments and the probability density function of the outputs of interest.

The problem analyzed within the present chapter was relatively simple, however the application of the method to larger scale problems is relatively straightforward because it does not involve any modification of its mathematical models. In the next two chapters this application is performed for more complex systems: in Chapter 5 the algorithm is going to be applied to criticality and source-detector problems, while in Chapter 6 the same is done for coupled time-dependent models.

CHAPTER 5

COST REDUCTION TECHNIQUES AND APPLICATION TO NEUTRONICS PROBLEMS

5.1 Introduction

In the previous chapter, a new algorithm to implement non-intrusive spectral techniques has been introduced and applied to a problem with a known analytical solution. In this chapter, we present the application of the algorithm to more complex models characterized by more random inputs. Before presenting these applications we discuss two cost reduction techniques which can be used to further reduce the computational cost associated with the method. These techniques are applied to two typical reactor physics problems: a criticality calculation, solved by using the diffusion approximation, and a source-detector problem modeled using transport theory.

5.2 Cost reduction techniques for adaptive sparse grid quadrature algorithms

The adaptive sparse grid algorithm introduced in the previous chapter can be employed to reduce the amount of realizations needed to evaluate the PCE expansion representing a stochastic quantity of interest. Unfortunately, the presence of a large number of input parameters or highly irregular stochastic outputs results in a drastic increase of points needed for the method to converge. Ultimately the application of the method will become more expensive than performing a standard random sampling approach. In order to tackle this issue, we propose in this section two ways to further reduce the computational requirements of the adaptive algorithm. These techniques are based on the particular implementation of the adaptive algorithm we presented in the previous chapter.

5.2.1 First cost reduction technique: neglecting part of the cross-correlations between input parameters

In the previous chapter we explained that one of the main constraints associated to the method is given by the number of two-dimensional sub-grids included within the final quadrature rule. This number, determined by the binomial coefficient introduced in Eq. 4.5, directly depends on the number of input quantities that have an effect on the stochastic outputs of the model. These two-dimensional sub-grids are evaluated to determine the cross-correlation between all possible pairs of input parameters. The cross-correlation between two inputs can be used to estimate how the first order derivative of a stochastic output with respect to the first input is influenced by a change in the second input.

When analyzing the spectral expansion of a stochastic output, this cross-correlation is given by the coefficients for which $\frac{\partial^2 \Psi_k}{\partial \xi_{\alpha_1} \partial \xi_{\alpha_2}} \neq 0$ (where α_1 and α_2 are two arbitrary inputs). According to the algorithm defined in the previous chapter, if the partial derivative of a random output with respect to any input parameter is non-zero, we automatically assume the presence of a cross-correlation between this parameter and the rest of the input set. This naturally corresponds to more quadrature points needed to evaluate the

5.2. Cost reduction techniques for adaptive sparse grid quadrature algorithms

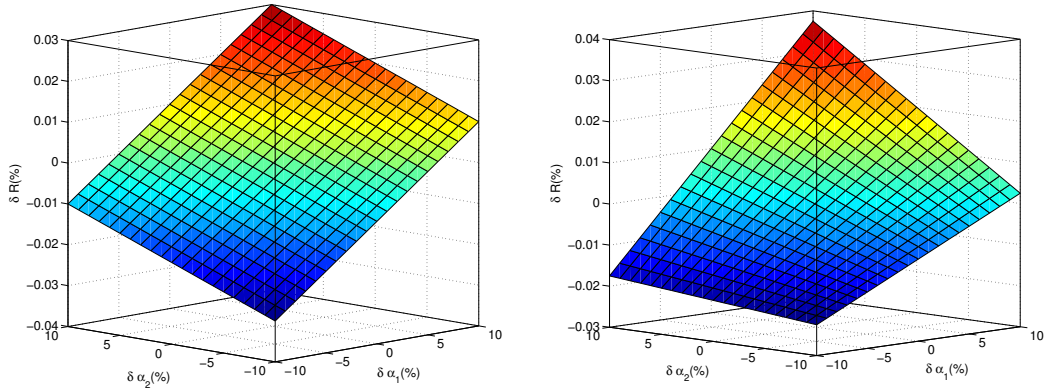


Figure 5.1: Example of uncorrelated and cross-correlated parameters (respectively left and right part of the figure)

spectral expansion.

Figure 5.1 can be used to better understand the meaning of cross-correlation between inputs. If two input quantities are not correlated, the response surface with respect to these quantities is a flat plane, which can be built by using their first order derivatives. In general, as we have also seen in Chapter 2, cross-correlation between parameters is not always present. During the derivation of an adjoint-based expression for the Hessian matrix of a stochastic response, we have seen that in many cases its off-diagonal terms are zero, especially if the diagonal elements of the matrix are already small (i.e., when the dependency of the output with respect to those directions is linear). We could therefore think of using this aspect to reduce the number of quadrature points needed to build the spectral expansion.

As a further step within the algorithm defined in the previous chapter, we can take advantage of the fact that after the evaluation of the one-dimensional sub-grids is completed, we know how the stochastic outputs behave along each of the stochastic directions. This knowledge is represented by the vector \mathbf{p} which is used to generate the reduced PCE presented in Eq. 4.4.

Before evaluating each of the two dimensional sub-grids (represented by the second column in Figure 4.4) we can exclude from its starting set (and from the final multi-index set I) all the sub-grids for which $p_i = 1$ along both the

defining directions. This operation is equivalent to neglecting the effect on the stochastic outputs given by the cross-correlation between parameters which have a linear influence.

Discarding a set of two-dimensional sub-grids from the multi-index set I , according to the information gathered while evaluating the one-dimensional sub-grids, could potentially correspond to a considerable reduction in terms of quadrature points, especially when dealing with problems characterized by sparse Hessian matrices (as the one discussed in Chapter 2). This approximation is used when dealing with the examples presented in the final part of this chapter.

5.2.2 Second cost reduction technique: Sub-grid ranking

Another help for the reduction of the computational cost associated to the algorithm comes from the particular order which is used to include the sub-grids within the final quadrature rule. Since these sub-grids are added according to their size and resolution, we can use the information we obtained from the sub-grids already included to have a better estimate of the indicator presented in Eq. 4.3 and associated to the grids that still need to be included. The idea is to take advantage, when evaluating a subset of grids having a certain size and norm, of the information we already know about the lower-dimensional and lower-norm grids.

As we have shown in the previous chapter (Figure 4.4), the sub-grids constituting the multi-index set I are added according to their norm $|\mathbf{l}|$ and size d . Each time we start evaluating a new set of sub-grids we can take advantage of the information obtained while evaluating the set of sub-grids characterized by a lower norm or a lower size by implementing the following algorithm:

- If $|\mathbf{l}| = N + d$ (which means that the sub-grids belonging to the set are built by using level 1 rules) then associate to each of the sub-grids a score given by the maximum indicator among the lower dimensional sub-grids contained in it. For example, the score given to the sub-grid represented in Figure 5.2 is determined as the largest indicator among the three two-dimensional grids highlighted in gray. Otherwise if $|\mathbf{l}| > N + d$ each sub-grid is assigned with a score determined by the indicator associated with the sub-grid defined on the same directions having norm $|\mathbf{l}| - 1$.

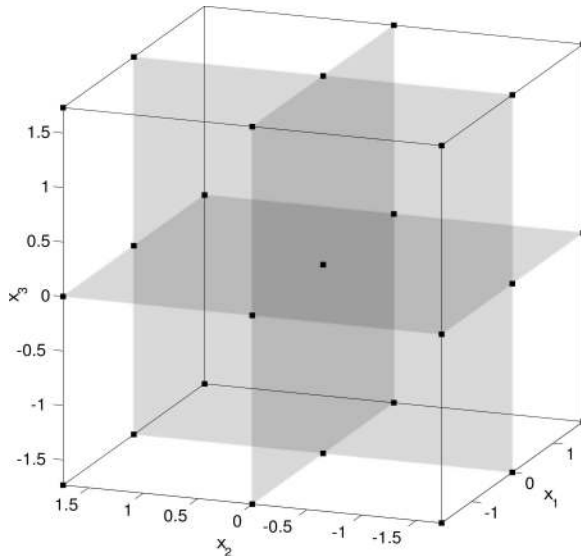


Figure 5.2: Example of three-dimensional grid. The grey planes lie on the two-dimensional grids already evaluated.

- Once this process is finished, group the sub-grids according to their score. Start evaluating the sub-grids contained in the group associated with the highest score.
- If the indicator associated with the group currently being evaluated is smaller than the algorithm tolerance multiplied by the number of sub-grids left, finish evaluating the sub-grids belonging to the current group and discard the remaining grids. Otherwise proceed to the next group until all the sub-grids having the same norm and size are evaluated.

This algorithm is based on the assumption that the ranking of the sub-grids, performed during the second step, is accurate. If this is true, the sub-grids that still need to be evaluated will be characterized by an indicator which is smaller than or equal to the last indicator calculated. Discarding sub-grids based on this assumption will considerably reduce the number of points needed to reach convergence. This is because every time a sub-grid is discarded, the admissibility criterion is enforced on the higher dimensional grids, therefore

excluding a considerable amount of points from the final quadrature rule. In the next two sections the adaptive algorithm is applied to two reactor physics problems. The effect of using the two optimization techniques just introduced is also discussed.

5.3 Application to a criticality problem

We first present the application of the adaptive algorithm and the optimization techniques just introduced to a criticality problem. This problem is modeled by using a 2 group diffusion model (Duderstadt and Hamilton, 1976) defined by the following set of equations

$$\begin{aligned} (-\nabla \cdot D_g \nabla + \Sigma_{rg}) \phi_g(\mathbf{r}) - \sum_{g' \neq g} \Sigma_s(g' \rightarrow g) \phi_{g'}(\mathbf{r}) &= \\ &= \frac{\chi_g}{k_{eff}} \sum_{g'} \nu \Sigma_{fg'}(\mathbf{r}) \phi_{g'}(\mathbf{r}) \end{aligned}$$

$g = 1, 2$

This model is implemented for a 240 cm thick heterogeneous slab reactor (similar to the one introduced in (Favorite and Stacey, 1995)) constituted by a 120 cm central part (region 1) surrounded by a 60 cm outer zone (region 2) on each side. The cross-sections and the transversal bucklings B_t^2 of the problem are provided in Table 5.1. This problem was solved, within the MATLAB environment, by using a Finite Volume approximation for the diffusion operator while the power method (Duderstadt and Hamilton, 1976) was employed for the calculation of the fundamental eigenvalue k_{eff} . Regarding the uncertainty propagation problem, we considered the macroscopic fission and capture cross-sections of the system to be normally distributed stochastic quantities. We assumed the cross-sections in the two regions to be independent, in this case a total of 8 uncorrelated random quantities can be used to model the stochastic problem. In terms of PCE, these quantities are represented in the following

Table 5.1: Mean values for the 2 group diffusion problem.

$D_1(cm)$	1.695497
$D_2(cm)$	0.4097145
$\Sigma_s(1 \rightarrow 2)(cm^{-1})$	0.01641351
$\Sigma_{c1}(cm^{-1})$	0.01375321
$\Sigma_{c2}(cm^{-1})$	0.2613659
$\nu\Sigma_{f1}(cm^{-1})$	0.01949058
$\nu\Sigma_{f2}(cm^{-1})$	0.497857
$B_t^2(\text{region 1})(cm^{-2})$	0.01090
$B_t^2(\text{region 2})(cm^{-2})$	0.01200

way

$$\begin{aligned}
 \Sigma_{f1}(\text{region1}) &= E(\Sigma_{f1}) + std(\Sigma_{f1})\xi_1 \\
 \Sigma_{c1}(\text{region1}) &= E(\Sigma_{c1}) + std(\Sigma_{c1})\xi_2 \\
 \Sigma_{f2}(\text{region1}) &= E(\Sigma_{f2}) + std(\Sigma_{f2})\xi_3 \\
 \Sigma_{c2}(\text{region1}) &= E(\Sigma_{c2}) + std(\Sigma_{c2})\xi_4 \\
 \Sigma_{f1}(\text{region2}) &= E(\Sigma_{f1}) + std(\Sigma_{f1})\xi_5 \\
 \Sigma_{c1}(\text{region2}) &= E(\Sigma_{c1}) + std(\Sigma_{c1})\xi_6 \\
 \Sigma_{f2}(\text{region2}) &= E(\Sigma_{f2}) + std(\Sigma_{f2})\xi_7 \\
 \Sigma_{c2}(\text{region2}) &= E(\Sigma_{c2}) + std(\Sigma_{c2})\xi_8
 \end{aligned}$$

where ξ_i is a vector constituted by a set of normally distributed (zero mean and unity variance) independent variables. Once these quantities are defined the uncertainty propagation is performed by evaluating the coefficients of the following expansion

$$k_{eff} = \sum_0^P k_k \Psi_k(\xi)$$

where Ψ_k is the set of Hermite polynomials (chosen since the input processes are Gaussian) which can be used to represent the behavior of the fundamental eigenvalue in the stochastic space.

5. Cost reduction techniques and application to neutronics problems

Table 5.2: Statistical moments (expressed in pcm) of the fundamental eigenvalue obtained by using a standard sparse grid algorithm (SG) and the adaptive one (ASG). 1% relative standard deviation case. Comparison with Monte Carlo (MC) sampling included.

Input std	1%			
	$E(k_{eff})$	$\sigma(k_{eff})$	N	Trunc. order \mathbf{p}
MC	-27.96 ± 2.2	704.17 ± 1.4	10^5	-
SG (l=3)	-26.33	706.90	161	22222222
SG (l=4)	-25.88	705.24	1105	22222222
SG (l=5)	-25.86	705.41	6097	22222222
ASG ($tol = 1\%$)	-26.33	706.90	161	21211121
ASG ($tol = 0.1\%$)	-26.33	706.90	161	22222222
ASG ($tol = 0.01\%$)	-25.99	705.84	448	22222222
ASG ($tol = 0.001\%$)	-25.91	705.55	1368	22222222

The adaptive sparse grid algorithm has been applied to evaluate the coefficients of the previous spectral expansion. Two reference cases were initially considered: a first case obtained by considering all the microscopic cross-sections to have 1% relative standard deviation and a second case for which all these standard deviations have been increased to 5%. These two configurations have been analyzed in order to better understand the convergence properties of the adaptive sparse grid algorithm. Tables 5.2 and 5.3 list the values of the statistical moments obtained by using this adaptive algorithm, compared to the values obtained by using a standard sparse grid quadrature and a Monte Carlo statistical sampling (Table 5.2 contains the results for the 1% input uncertainty case, while Table 5.3 refers to the 5% one). The adaptive algorithm (ASG) has been applied in both cases without using the reduction cost techniques introduced at the beginning of the chapter. A second order PCE was required to reach convergence for most of the adaptive cases, the same order was thus used to evaluate the moments using the complete sparse grids (SG). The convergence of the PCE can be seen in Figure 5.3 where a comparison between the probability density function of the fundamental eigenvalue obtained by sampling the second order PCE with the one obtained by collecting 10^5 random realizations of the model is shown. In Tables 5.2 and 5.3 the results obtained using different tolerances ($tol = 1\%$, $tol = 0.1\%$

Table 5.3: Statistical moments (expressed in pcm) of the fundamental eigenvalue obtained by using a standard sparse grid algorithm (SG) and the adaptive one (ASG). 5% relative standard deviation case. Comparison with Monte Carlo (MC) sampling included.

Input std	5%			
	$E(k_{eff})$	$\sigma(k_{eff})$	N	Trunc. order \mathbf{p}
MC	1237.30 ± 11.1	3742.97 ± 7.0	10^5	-
SG (l=3)	956.88	4086.95	161	22222222
SG (l=4)	1347.80	3666.39	1105	22222222
SG (l=5)	1240.87	3746.57	6097	22222222
ASG ($tol = 1\%$)	1320.80	3689.46	617	22222222
ASG ($tol = 0.1\%$)	1245.60	3745.46	3541	22222222
ASG ($tol = 0.01\%$)	1241.13	3746.54	5729	23222223
ASG ($tol = 0.001\%$)	1240.71	3746.66	6097	23222223

and $tol = 0.01\%$), and the corresponding number of evaluations required for convergence, are compared to the values obtained by using the standard sparse grid approach with different accuracy levels lev (3,4,5). It should be pointed out that, when using a standard sparse grid quadrature, it is not possible to determine the convergence of the quadrature formula without evaluating one accuracy level higher than the one actually required for convergence. On the contrary, the sequential construction of the adaptive algorithm makes the convergence analysis possible within the sparse grid construction procedure. Of course, an increase in the specified tolerance corresponds to a reduction of the number of quadrature points needed for the convergence of the algorithm, one can therefore define an optimal tolerance, depending on the problem considered, giving a good compromise between the final error and the computational cost.

The comparison between the results obtained for the 1% and 5% standard deviation cases shows an important aspect of the sparse grid quadrature approach. In each line the values of the mean and standard deviation are accompanied by the number of realizations N required to reach convergence and, when applicable, by the truncation order \mathbf{p} needed to spectrally represent the eigenvalue.

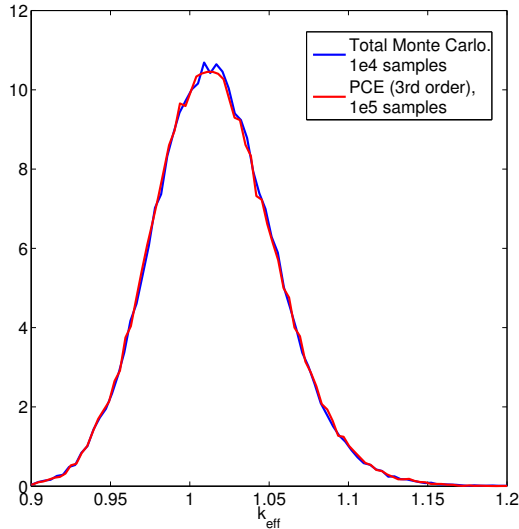


Figure 5.3: Comparison between probability density functions obtained by a direct sampling of the model and by using the PCE.

As seen in Table 5.3 the number of points required to converge increases considerably when increasing the standard deviation of the stochastic input quantities. This can be expected since the accuracy level required by the sparse grid formula directly depends on the differentiability of the function (Le Maitre and Knio, 2010). In Figure 5.4 it can be seen how increasing the standard deviation of the inputs corresponds to moving to regions where the differentiability of the response surfaces increases. This aspect represents a drawback of non-intrusive spectral methods since it poses a limit, in terms of number of realizations required, on the maximum number of parameters with respect to which the output can present an irregular behavior.

As we discussed at the beginning of the chapter, we can further reduce the computational requirements associated to the method by using two cost reduction techniques: first, the neglect of cross-correlation when in presence of linear responses, and second the sub-grids ranking performed when applying the adaptive algorithm. Table 5.4 collects the outcome of using these two cost reduction techniques for the very same problem introduced in Tables 5.2 and 5.3. It is immediate to see how in both cases the numbers of realizations needed

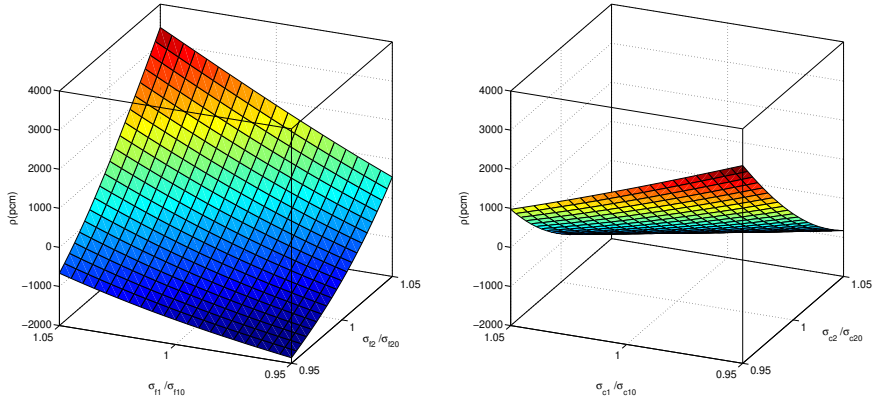


Figure 5.4: Variation of the fundamental eigenvalue with respect to the capture and the fission cross-sections of region 1.

Table 5.4: Statistical moments (expressed in pcm) for a second order spectral expansion of the fundamental eigenvalue obtained by using the adaptive algorithm together with the two cost reduction techniques defined in this chapter.

Input std	1%			5%		
	$E(k_{eff})$	$\sigma(k_{eff})$	N	$E(k_{eff})$	$\sigma(k_{eff})$	N
ASG ($tol = 1\%$)	-25.71	701.27	137	1243.63	3750.52	401
ASG ($tol = 0.1\%$)	-26.33	706.90	161	1213.83	3762.24	2053
ASG ($tol = 0.01\%$)	-26.10	706.18	289	1241.13	3746.54	5729

for convergence are considerably lower than the ones presented in Tables 5.2 and 5.3. For the 5% input uncertainty case, using these techniques proved to be very effective since up to two thirds of the original realizations were required for convergence.

The effectiveness of this application can be explained analyzing the validity of the sub-grid ranking approach. Figure 5.5 can be used to show how this

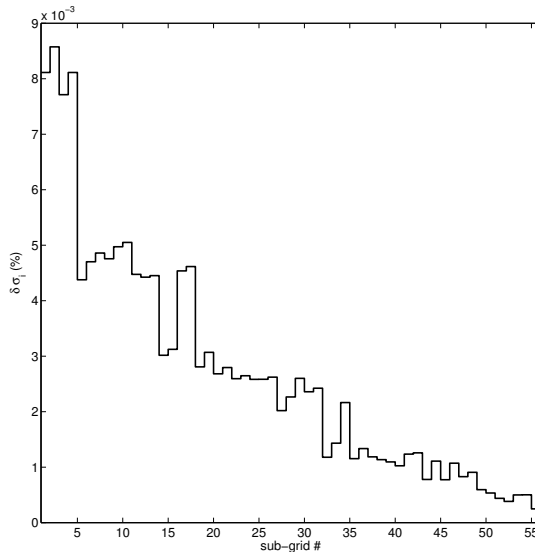


Figure 5.5: Ranking (left to right) of the three-dimensional level 1 sub-grids performed by using the method introduced in the first section and actual contribution to the final standard deviation corresponding to each of these grids.

ranking has been correctly performed. Here, the contribution on the standard deviation of each of the level 1 three-dimensional sub-grids is plotted. These sub-grids are ordered along the abscissa axis according to the order (from left to right) they have been evaluated while applying the adaptive algorithm. The Figure shows how the ranking performed before the evaluation is fairly accurate, with the grids causing the largest contribution being evaluated first. This ranking allows the script to discard part of these grids thus reducing the required number of realizations considerably.

Although a considerable reduction can be achieved by using optimization techniques, the final number of realizations needed to have an accurate result is still relatively large. Because of the reason mentioned before, there are cases for which spectral techniques will present slow convergence properties, even in their adaptive form. It must be pointed out that each of the input parameters of the example considered has a large influence on the stochastic eigenvalue

due to the simplified form of the problem. In general one could expect, for more complex systems, the dependency of the stochastic outputs with respect to the input parameters to be less strong.

5.4 Application to a transport problem

The second example is the application of the method to a neutron transport problem. In this example the uncertainty propagation is performed for a non-multiplying system modeled by the steady-state Boltzmann equation (Prinja and Larsen, 2010)

$$\begin{aligned} \boldsymbol{\Omega} \cdot \nabla \psi(\mathbf{x}, \boldsymbol{\Omega}, E) + \sigma_t(\mathbf{x}, E)\psi(\mathbf{x}, \boldsymbol{\Omega}, E) = \\ = \int_0^\infty \int_{4\pi} \sigma_s(\mathbf{x}, \boldsymbol{\Omega}' \cdot \boldsymbol{\Omega}, E' \rightarrow E)\psi(\mathbf{x}, \boldsymbol{\Omega}', E')d\boldsymbol{\Omega}'dE' + S(\mathbf{x}, \boldsymbol{\Omega}, E) \end{aligned}$$

where ψ is the angular flux, $\boldsymbol{\Omega} \cdot \nabla$ is the streaming operator, σ_t and σ_s are respectively the total and the scattering macroscopic cross-sections, and S the source term of our problem. This is an integro-differential equation whose unknown is defined in space (\mathbf{x}), angle ($\boldsymbol{\Omega}$), and energy (E). There are several techniques to solve such equation, for this work we used an inhouse code whose details have first been presented by Kophazi and Lathouwers (2012). This code uses a Discontinuous Galerkin formulation to deal with the spatial operator and the Discrete Ordinate method for the angular domain. A multi-group formulation was used to solve the energy problem.

The Boltzmann equation was used to model a two-dimensional source-detector problem whose geometry is presented in Figure 5.6. The energy domain was divided into 2 groups, the capture cross-sections and the scattering matrices for each different material are included in Table 5.5.

The source in the system is isotropic, introducing 1 neutron per second in the first energy group. The two following responses were introduced

$$\begin{aligned} R_1 &= \frac{1}{V_d} \int_{V_d} \int_{4\pi} \psi_1(\mathbf{x}, \boldsymbol{\Omega})d\boldsymbol{\Omega}d\mathbf{x} \\ R_2 &= \frac{1}{V_d} \int_{V_d} \int_{4\pi} \psi_2(\mathbf{x}, \boldsymbol{\Omega})d\boldsymbol{\Omega}d\mathbf{x} \end{aligned}$$

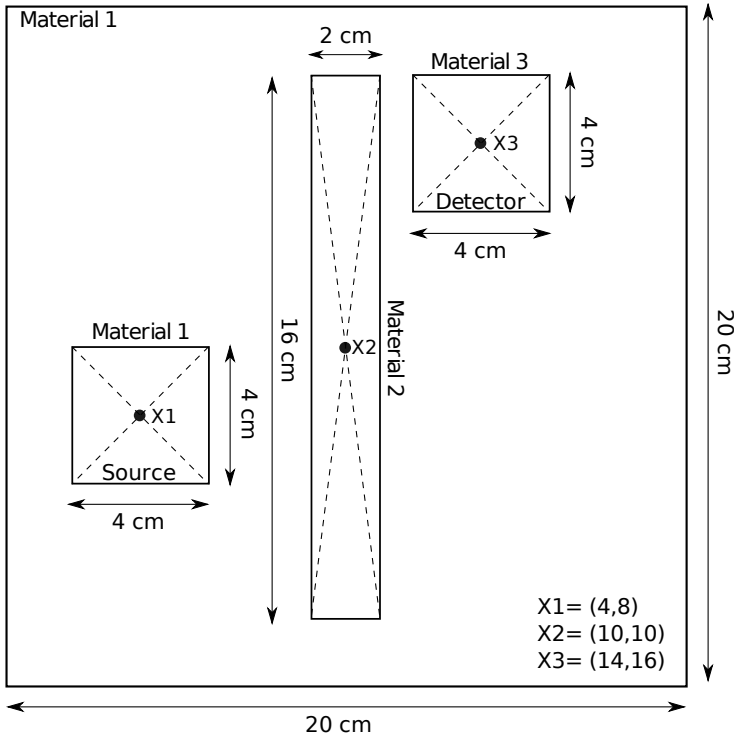


Figure 5.6: Geometry used for the source-detector problem.

where V_d is the area of the detector. These expressions represent the scalar flux, for the two energy groups, averaged over the detector region. The numerical solution of this problem was calculated using an S4 approximation (corresponding to 12 angular directions) and a mesh containing 6742 elements.

Regarding the uncertainty propagation problem, we considered all cross-sections included in Table 5.5 to be stochastic quantities. These cross-sections are assumed to be normally distributed with a 5% relative standard deviation, for a total of 15 random inputs. The influence of these random inputs on the two responses defined above is the final aim of the current uncertainty propagation application.

First, the uncertainty quantification problem was tackled by using a standard Monte Carlo sampling technique. 5000 realizations of the random inputs

Table 5.5: Cross-sections for each different material used in the transport model

	Material 1	Material 2	Material 3
σ_{c1} [cm]	0.05	0.2	0.04
σ_{c2} [cm]	0.01	0.9	0.1
$\sigma_{s1 \rightarrow 1}$ [cm ²]	0.34	0.1	0.3
$\sigma_{s1 \rightarrow 2}$ [cm ²]	0.01	0.1	0.01
$\sigma_{s2 \rightarrow 2}$ [cm ²]	0.99	0.1	0.9

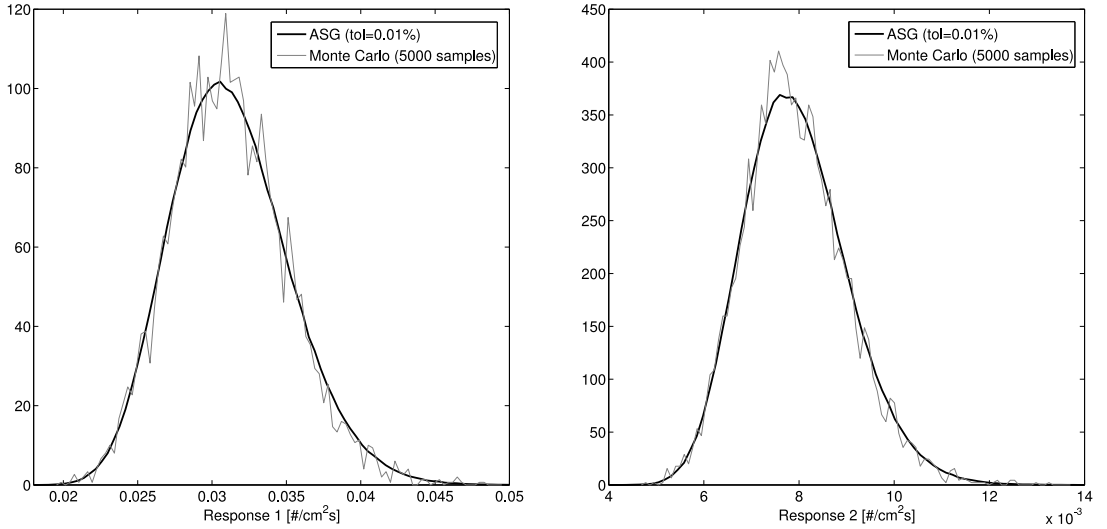


Figure 5.7: Comparison between probability density functions obtained with the adaptive algorithm using different tolerances.

were collected to reconstruct the probability density functions of the two responses. Then, the adaptive algorithm was applied by including the two approximation techniques introduced at the beginning of the chapter. Figure 5.7 shows the comparison of the probability density functions of the two responses obtained with the sampling approach (MC) and with the adaptive algorithm (ASG) applied using a 0.001% tolerance. The comparison shows a very good agreement between the spectral technique and the sampling approach. Furthermore, a considerable positive skewness is present in the probability

5. Cost reduction techniques and application to neutronics problems

Table 5.6: Statistical moments of the first detector response (in neutrons/cm²s).

	E	σ	N
MC	$0.007792 \pm 7.5 \times 10^{-6}$	$5.324 \times 10^{-4} \pm 4.5 \times 10^{-6}$	5000
ASG (<i>tol</i> = 0.01%)	0.031054	0.0039778	159
ASG (<i>tol</i> = 0.1%)	0.031054	0.0039648	67
ASG (<i>tol</i> = 1%)	0.031054	0.0039474	55

Table 5.7: Statistical moments of the second detector response (in neutrons/cm²s).

	E	σ	N
MC	$0.007878 \pm 1.5 \times 10^{-5}$	$0.0010905 \pm 3.5 \times 10^{-5}$	5000
ASG (<i>tol</i> = 0.01%)	0.007884	0.0010880	307
ASG (<i>tol</i> = 0.1%)	0.007884	0.0010887	215
ASG (<i>tol</i> = 1%)	0.007884	0.0010840	119

density functions. Tables 5.6 and 5.7 include the values of the first two statistical moments (of the two responses) obtained by running the algorithm with different tolerances and approximations. The number of realizations needed to build the final quadrature rule is relatively low for both responses, even when using small tolerances. In this case, evaluating the full sparse grid provided to be computationally unfeasable due to the large number of input parameters. Even with a 1% tolerance (corresponding to a considerably small number of realization) the statistical moments were accurately predicted.

In order to understand if this was the case for higher order statistical moments, the probability density functions of the two responses have also been evaluated. In Figure 5.8 a comparison between these functions obtained by using different tolerances (1% and 0.1%) is shown. The distributions built by using a 0.1% tolerance are basically overlapping the ones obtained by using the smaller tolerance. A small discrepancy is present for the distribution of the first response when using a 1% tolerance, however this difference is within the specified error. The important thing is that in both cases the skewness

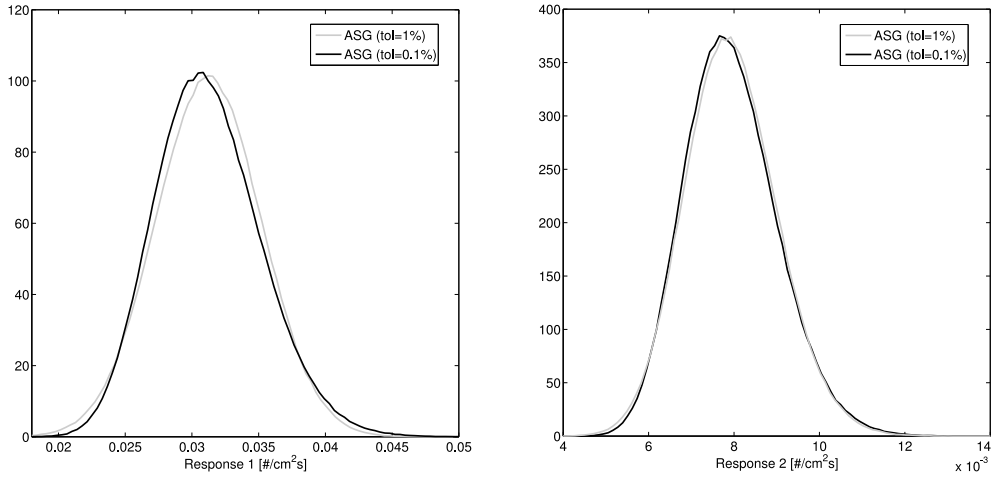


Figure 5.8: Comparison between probability density functions obtained by including and excluding the approximations from the algorithm.

of each of the density functions is correctly represented even when using the 1% tolerance. As we discussed above, this corresponds to a remarkably low number of realizations needed to reconstruct the stochastic responses.

5.5 Conclusions

In this chapter we discussed the application of optimization techniques which can be used to further reduce the computational cost associated with non-intrusive spectral methods. These techniques, whose application involves specific approximations, take advantage of the sequential sparse grid construction presented in the previous chapter.

Two illustrative uncertainty propagation examples have been considered. First a multigroup diffusion problem has been introduced using the microscopic cross-sections of the system as a stochastic input set. Then, a fixed-source transport problem has been defined for which a set of cross-sections have been assumed to be randomly distributed. For both problems the adaptive sparse grid algorithm proved to be an effective way to reduce the number of realizations required to evaluate the spectral expansion. It has been shown, with the first example, that the convergence rate of the adaptive quadrature algorithm directly depends on the differentiability of the response surface.

In both cases the use of the two techniques introduced in this chapter represents a considerable reduction in terms of the number of realizations required for convergence.

In general the adaptive spectral technique could represent a suitable uncertainty propagation technique, even in the presence of complex models.

CHAPTER 6

APPLICATION TO TIME-DEPENDENT MULTI-PHYSICS PROBLEMS

6.1 Introduction

In this chapter we present the application of the algorithm introduced in the previous two chapters to a multi-physics problem. The model characterizing this type of problem is first described from the mathematical and numerical point of view. Then, we define an uncertainty propagation case and we solve it by means of the adaptive spectral technique. The outcome of this application is finally discussed in the last section.

6.2 Definition of a reference multi-physics problem

This section describes a nonlinear time-dependent problem to be used as benchmark for the adaptive spectral technique. This type of problems is commonly encountered in the reactor physics field when performing safety

analysis. In general, the uncertainty associated with such simulations is quantified by using traditional sampling techniques (D Auria et al., 2012), it is therefore interesting to compare this approach with the adaptive spectral method defined in this thesis.

6.2.1 Description of the model

The problem considered for the application of the adaptive algorithm defined in Chapter 4 is a coupled system which can be used to model the time-dependent behavior of a sodium fast reactor. The reference configuration used within the present analysis is the BN800, a sodium cooled fast breeder reactor (IAEA, 1994), with main characteristics summarized in Table 6.1. The neutronics part of the model is described using a point-kinetic model (Duderstadt and Hamilton, 1976)

$$\begin{aligned}\frac{dQ}{dt} &= \frac{\rho(t, T_f, T_c) - \beta}{\Lambda} Q + \sum_{k=1}^K \lambda_k C_k \\ \frac{dC_k}{dt} &= -\lambda_k C_k + \frac{\beta_k}{\Lambda} Q\end{aligned}$$

where Q is the reactor power, Λ the mean generation time, C_k the concentration of the k^{th} precursor group (in power terms), β_k and λ_k the delayed neutrons fraction and the decay constant for the k^{th} precursor group respectively, and β the total delayed neutrons fraction. The reactivity ρ is considered as the sum of three different contributions

$$\rho(t, T_f, T_c) = \rho_{ext}(t) + \delta\rho_D(T_f) + \delta\rho_C(T_c)$$

the external reactivity is provided to the model while the Doppler (D) and coolant (C) reactivities represent the temperature dependent feedback mechanisms. These reactivities are determined by spatially averaging the temperature field

$$\begin{aligned}\delta\rho_D &= \alpha_D \frac{1}{V_f} \int \int dr dz 2\pi r [\varphi_D(r, z) (T_f(r, z) - T_{rf})] \\ \delta\rho_C &= \alpha_C \frac{1}{H} \int dz [\varphi_C(r, z) (T_c(r, z) - T_{rc})]\end{aligned}$$

φ_D and φ_C are the spatial weighing functions used to evaluate the integral and α_D and α_C the first order reactivity coefficients, modeled around a reference

temperature $T_{r,f,c}$. The kinetic parameters used in the point-kinetic model and the reactivity coefficients have been obtained using the ERANOS2.2 code (Rimpault et al., 2002) employing the heterogeneous 3D model described in (IAEA, 1994).

In order to calculate the temperature fields required for the feedbacks, the reactor domain is represented by an equivalent pin geometry. In this pin the volumetric heat produced by the fission processes is transferred across the cylindrical fuel and cladding to the sodium coolant, which is supposed to have a fixed inlet temperature. Axial conduction along the pin is neglected which means the heat transfer process is modeled by using a set of radial energy conservation equations. In Figure 6.1 the reference geometry for the thermo-kinetic problem is shown. We use the following equation for the heat transfer within the fuel

$$\rho_f c_{p,f} \frac{\partial T_f^z}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(r k_f (T_f^z) \frac{\partial T_f^z}{\partial r} \right) + f^z(Q)$$

where the superscript z refers to the position along the vertical axis and $f^z(Q)$ is an axial shape function used to describe the distribution of the volumetric power. This function is supposed to be distributed as the fundamental neutronics solution of an equivalent homogeneous cylindric core. Similarly, for the cladding we have

$$\rho_{cl} c_{p,cl} \frac{\partial T_{cl}^z}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(r k_{cl} (T_{cl}^z) \frac{\partial T_{cl}^z}{\partial r} \right)$$

The thermodynamic and heat transport properties of the fuel and the cladding have been characterized using the correlations present in (Waltar and Reynolds, 1981). The boundary conditions used for the thermo-kinetic problem are

$$\begin{aligned} \left. \frac{\partial T_f^z}{\partial r} \right|_{r=0} &= 0 \\ k_f(T_f^z) \left. \frac{\partial T_f^z}{\partial r} \right|_{R_f} &= h_g(T_{cl}^z|_{R_f+\delta_g} - T_f^z|_{R_f}) \\ R_f k_f(T_f^z) \left. \frac{\partial T_f^z}{\partial r} \right|_{R_f} &= (R_f + \delta_g) k_{cl}(T_{cl}^z) \left. \frac{\partial T_{cl}^z}{\partial r} \right|_{R_f+\delta_g} \\ k_{cl}(T_{cl}^z) \left. \frac{\partial T_{cl}^z}{\partial r} \right|_{R_s} &= h_c(T_c(z))(T_c(z) - T_{cl}^z|_{R_s}) \end{aligned}$$

6. Application to time-dependent multi-physics problems

where R_s is the pin radius, R_f the radius of the fuel pellet, δ_g the size of the fuel-cladding gap, h_g the heat transfer coefficient across the gap and h_c the heat transfer coefficient of the coolant (obtained using the correlation for liquid metals introduced in (Waltar and Reynolds, 1981)). The coolant temperature is determined by the following energy conservation equation, used for the equivalent channel

$$A_c \rho c_{pc} \left[\frac{\partial T_c}{\partial t} + v_{in} \frac{\partial T_c}{\partial z} \right] - 2\pi R_s n_p h_c(T_c)(T_{cl}(R_s, z) - T_c) = 0$$

where A_c is the equivalent flow area and n_p the total number of fuel pins in the reactor. A fixed coolant velocity is considered and the boundary condition is given by a fixed coolant inlet temperature. All the parameters defined in the model are summarized in Table 6.1.

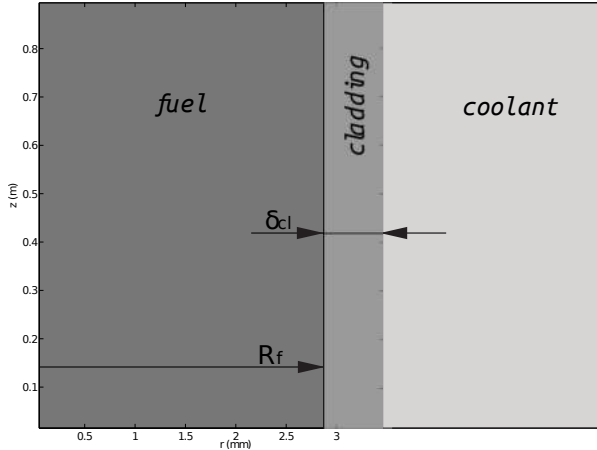


Figure 6.1: Reference geometry for the thermo-kinetic problem.

6.2.2 Numerical solution of the model

The numerical code used to solve the coupled problem employs the CVODE time-integration modules included within the SUNDIALS suite (Hindmarsh et al., 2005). These modules make use of a Backward Differential Formula, which is adaptive both in terms of time stepping and order used during the

Table 6.1: Main parameters used for the coupled model

Λ (s)	4×10^{-7}	Mean generation time of the reactor
R_f (cm)	0.293	Fuel pellet radius
α_d (pcm/K)	-0.68747	Doppler reactivity coefficient
α_c (pcm/K)	0.12251	Coolant reactivity coefficient
δ_{clad} (cm)	0.055	Clad thickness
v_{in} (m/s)	10	Coolant inlet velocity
H(m)	0.91	Coolant channel height
Q_0 (MW)	2100	Reactor thermal power
δ_g (cm)	10^{-6}	Fuel-cladding gap width
h_g ($W/m^2 K$)	10^4	Fuel-Cladding gap heat transfer coefficient
T_{in} (K)	628.15	Coolant inlet temperature
λ_1 (1/s)	0.0124	Decay constant of precursor 1
λ_2 (1/s)	0.0124	Decay constant of precursor 2
λ_3 (1/s)	0.111	Decay constant of precursor 3
λ_4 (1/s)	0.301	Decay constant of precursor 4
λ_5 (1/s)	1.14	Decay constant of precursor 5
λ_6 (1/s)	3.01	Decay constant of precursor 6
β_1 (-)	0.00009	Delayed neutron fraction for nuclide 1
β_2 (-)	0.000853	Delayed neutron fraction for nuclide 2
β_3 (-)	0.0007	Delayed neutron fraction for nuclide 3
β_4 (-)	0.0014	Delayed neutron fraction for nuclide 4
β_5 (-)	0.0006	Delayed neutron fraction for nuclide 5
β_6 (-)	0.00055	Delayed neutron fraction for nuclide 6

time integration. The spatial discretization of the model has been performed by implementing a Finite Volume scheme, a linear interpolation was used to determine the surface values between volumes.

Since the conduction along the axial direction was considered to be negligible, the heat transfer problem was formulated by dividing the fuel pin into 30 axial regions. In each of these regions a one-dimensional radial heat conduction problem was defined and numerically solved by uniformly dividing the radial domain into 50 volumes. The boundary condition of each of these problems,

given by the heat transferred between the clad and the coolant, determined the coupling between the heat transfer equations and the energy conservation equation along the coolant channel. This equation was finally solved by dividing the channel into 150 axial volumes.

In order to reduce the numerical error associated with the time integration process, the relative error used by the CVODE module has been set to 10^{-7} .

6.2.3 Definition of an uncertainty quantification problem

Based on this coupled model, we introduce a reference UQ problem whose solution can be obtained by using the adaptive spectral algorithm introduced in the previous chapters. This reference problem is the transient caused by the insertion of an external reactivity term ρ_{ext} . Two different types of insertion have been modeled. Firstly, a 1\$ step insertion and secondly a ramp characterized by a slope of .75\$/s. In both cases the external reactivity term lasts until $t_s = 1s$ and is zero afterwards. Furthermore, we consider the system to be at a steady state condition when the external reactivity is introduced at $t = 0$.

In order to introduce the reference uncertainty problem, some of the input parameters of the coupled model have been considered to be stochastic. A total of 10 parameters were assumed to be normally distributed, their list and the associated uncertainties are presented in Table 6.2. These were the total heat transfer coefficient h_g of the gap between the fuel pellet and the cladding, the radius of the pin R_s , the thickness of the cladding δ_{clad} , the inlet velocity and temperature v_{in} and T_{in} , the magnitude of the external reactivity insertion ρ_{ext} (the slope for the ramp insertion) and the time at which the external reactivity is removed t_s .

The fact that these parameters are normally distributed could in theory lead to non-physical realizations of the input quantities, however due to the relatively small uncertainties no such realization occurred while applying spectral and sampling approaches.

The UQ problem arising from the presence of these stochastic input parameters has first been solved by using a standard unbiased random sampling approach using 10000 realizations and then by applying the adaptive spectral approach. Because all the input parameters were normally distributed, Hermite

Table 6.2: Input parameter (relative) uncertainties used for the definition of the reference problem. A 5% relative standard deviation of ρ_{ext} was considered for the step insertion case while a 1% relative standard deviation was used for the ramp.

Param	R_s	δ_{clad}	h_g	ρ_{ext}	α_d	α_c	t_s	v_{in}	T_{in}	Λ
σ	1%	1%	10%	5%/1%	5%	5%	1%	5%	2%	1%

polynomials were used to build a 10 dimensional PCE suitable to represent the stochastic quantities of the problem. Using this set of polynomials it is possible to spectrally represent the stochastic solution by using the following expansions

$$\begin{aligned}
 Q(t, \boldsymbol{\xi}) &= \sum_{i=0}^P Q_i(t) \Psi(\boldsymbol{\xi}) \\
 T_f(\mathbf{x}, t, \boldsymbol{\xi}) &= \sum_{i=0}^P T_{f,i}(\mathbf{x}, t) \Psi(\boldsymbol{\xi}) \\
 T_{cl}(\mathbf{x}, t, \boldsymbol{\xi}) &= \sum_{i=0}^P T_{cl,i}(\mathbf{x}, t) \Psi(\boldsymbol{\xi}) \\
 T_c(\mathbf{x}, t, \boldsymbol{\xi}) &= \sum_{i=0}^P T_{c,i}(\mathbf{x}, t) \Psi(\boldsymbol{\xi})
 \end{aligned}$$

Where $\boldsymbol{\xi}$ is a vector containing the 10 independent normal variables that are used to define the stochastic parameters of Table 6.2.

The algorithm used to implement the non-intrusive spectral method was implemented as a MATLAB script which was used to adaptively generate the stochastic realizations corresponding to the quadrature points, run the solver, and collect the outputs of interest. In the next section the outcomes are presented and discussed.

6.3 Results: transient caused by a step reactivity insertion

In this section we analyze the response of the coupled system to a step reactivity insertion. This test problem was chosen because of the characteristic probability density functions associated to the presence of a large uncertainty in the magnitude of the inserted reactivity. As shown in Chapter 3, this type of transient can lead to highly skewed, non-gaussian probability density functions of the system outputs, especially considering the maximum power reached during the transient. For this reason, the magnitude of the external reactivity was assumed to have a 5% relative standard deviation. The values included in Table 6.2 were used for the rest of the parameters. We considered the maximum power reached during the transient as the response of interest for this stochastic problem.

Figure 6.2 shows the probability density function of this response obtained by using a standard unbiased sampling technique (MC), by applying the adaptive approach (ASG) using three different tolerances ($\epsilon = 1\%$, $\epsilon = 0.1\%$ and $\epsilon = 0.0001\%$) together with the cost reduction techniques presented in Chapter 5. Table 6.3 collects the mean and standard deviation of the response corresponding these different cases. The number of realizations needed for the algorithm to converge is also included, together with the final PCE order employed to represent the stochastic response, as defined in Eq. 4.4. The comparison between the probability density functions in Figure 6.2 shows that reducing the algorithm tolerance corresponds to a convergence to the distribution obtained by using a sampling approach. The skewness of these probability density functions is quite large, this means that standard perturbative approaches would have failed in the prediction of the statistical moments of interest.

The distribution built by using the largest tolerance ($\epsilon = 1\%$) is considerably different from the Monte Carlo one, which is a consequence of the fact that the algorithm convergence is verified only for the mean and standard deviation of the output: the values in Table 6.3 show how these values are in good agreement with the Monte Carlo prediction even when using the largest tolerance. In theory one could improve this aspect by introducing higher statistical moments within the indicator defined in Eq. 4.3. Unfortunately it is not possible to evaluate higher order statistical moments by using a simple expression like

Table 6.3: Mean and Standard deviation of the maximum power reached during the transient. Comparison between the application of the adaptive script applied with different tolerances and Monte Carlo sampling. The # column contains the number of realizations needed to reach convergence while the \mathbf{p} column refers to the truncation order used along each direction of the PCE. The order of the parameters represented by \mathbf{p} is the same as the one presented in Table 6.2.

	Mean[-]	Std[-]	#	\mathbf{p}
ASG (tol=1%)	21.4145	17.4934	63	1131111111
ASG (tol=.1%)	21.4163	17.5460	83	2141111112
ASG (tol=.0001%)	21.4185	17.5528	4577	5151231145
MC	21.30 ± 0.17	17.43 ± 0.13	10000	-

the one presented in Eq. 3.3 for the standard deviation. When applying the algorithm without any of the optimizations introduced in Chapter 5, it emerges (see Table 6.3) that there is clearly a value of the algorithm tolerance ($\epsilon = 0.1\%$) which represents the best compromise between the number of realizations needed and the statistical moments. When reducing this tolerance the number of points needed to converge increases considerably despite the fact that the additional stochastic information obtained is negligible.

The column including the values of \mathbf{p} shows the convenience of using a truncated PCE, as defined in Eq. 4.4. This originates from the fact that many directions can be reasonably represented by using a first order expansion. This implies that the algorithm will converge faster since the projection integrals become easier to solve along a subset of directions. This specific example shows that the adaptive technique introduced so far is very well suited for problems where most of the output perturbation is caused by a small number of input parameters.

6.4 Results: transient caused by a ramp reactivity insertion

The adaptive sparse grid algorithm was then applied to the transient triggered by the ramp insertion. We first analyzed a limited set of stochastic outputs.

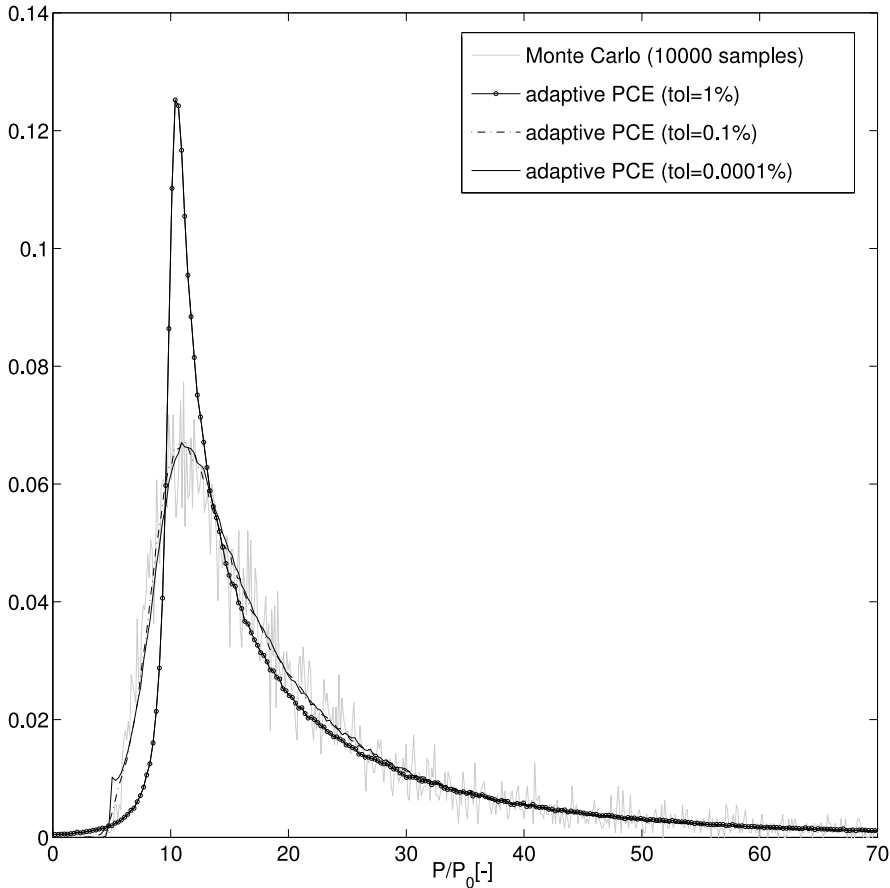


Figure 6.2: Probability density function of the maximum power reached during the transient obtained by using standard Monte Carlo sampling and the adaptive algorithm with three different tolerances. These density functions were built by directly sampling the Polynomial Chaos Expansion of the response.

These were the maximum power, the maximum temperature of the fuel, cladding, and coolant reached during the transient and the values of the same quantities at $t = 10s$. This set of responses was used to check the convergence of the adaptive sparse grid quadrature with several tolerances used for this purpose.

Tables 6.4 and 6.5 collect the outcome, in terms of standard deviation of

Table 6.4: Standard deviation of the maximum values reached during the transient for 4 stochastic quantities of interest. Comparison between the application of the adaptive script applied with different tolerances and Monte Carlo sampling. N is the number of realizations needed to evaluate the statistical moments.

	$\epsilon = 10^{-2}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	Sparse Grid	MC
	$N = 21$	$N = 203$	$N = 777$	$N = 12981$	$N = 10000$
$\sigma(\max[T_f])[K]$	12.8376	12.8544	12.8632	12.8667	12.82 ± 0.08
$\sigma(T_{cl})[K]$	9.7369	9.7593	9.7597	9.7597	9.83 ± 0.06
$\sigma(T_{outlet})[K]$	12.8178	12.8165	12.8170	12.8083	12.88 ± 0.08
$\sigma(P/P_0)$	0.07578	0.07598	0.07592	0.07597	0.077 ± 0.001

the responses, of the application of the adaptive spectral technique (ASG line in the Tables). Three different tolerances have been used ($\epsilon = 1\%$, $\epsilon = 0.1\%$ and $\epsilon = .01\%$), which correspond to different amounts of realizations needed to reach convergence. A reference calculation obtained by using a complete sparse grid (built according to the definition introduced in Eq. 4.1) was also performed. Table 6.4 includes the standard deviation corresponding to the maximum value of the responses during the transient, while Table 6.5 presents the value of the standard deviations at $t = 10s$. Both tables also contain the output of a standard Monte Carlo sampling approach for UQ. All the values obtained with the spectral technique are in agreement with their corresponding statistical estimators.

It is immediately noticeable how the use of different tolerances corresponds to very different numbers of realizations needed for the algorithm to converge. In every case, the difference between the final predictions obtained is relatively small, especially considering the additional amount of computations required when using smaller tolerances. The tables show that a ($\epsilon = 0.1\%$) tolerance leads to reasonable results despite the fact that the number of realizations is considerably lower compared to the other cases.

6. Application to time-dependent multi-physics problems

Table 6.5: Standard deviation of 4 stochastic quantities of interest at $t = 10s$. Comparison between the application of the adaptive script applied with different tolerances and Monte Carlo sampling. N is the number of realizations needed to evaluate the statistical moments.

	$\epsilon = 10^{-2}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	Sparse Grid	MC
	$N = 49$	$N = 189$	$N = 937$	$N = 12981$	$N = 10000$
$\sigma(\max[T_f])[K]$	15.6292	15.6551	15.6840	15.6850	15.60 ± 0.1
$\sigma(T_{cl})[K]$	8.1824	8.1955	8.1955	8.1943	8.27 ± 0.06
$\sigma(T_{outlet})[K]$	11.3042	11.3248	11.3257	11.3231	11.41 ± 0.08
$\sigma(P/P_0)$	0.029871	0.029956	0.029945	0.029938	0.0302 ± 0.0004

This important aspect can be explained by keeping in mind that the specified tolerance not only acts on the building process of the final sparse grid, but also on the definition of the reduced PCE expansion introduced in Eq. 4.4. Using a small tolerance for the convergence check of the PCE corresponds to introducing an additional set of higher order polynomials within the final expansion. As a consequence, the projections onto this new set of polynomials will require additional quadrature points for the integration procedure to convergence. Higher order polynomials that are contributing only marginally to the final spectral expansion will cause an increase in quadrature points which is not compensated in terms of final accuracy of our estimators. Table 6.6 presents the order \mathbf{p} (as defined in Eq. 4.4) of the final PCE expansions used to represent the outputs at $t = 10s$ for different tolerances used within the adaptive script. Though the exact convergence of the PCE requires higher order polynomials (third column), it is possible to accurately represent our stochastic outputs by using a PCE built with first and second order polynomials (first column). This naturally translates to a reduced number of quadrature points, as discussed in the previous paragraph.

As Table 6.6 shows, the outputs with respect to many of the input variables have a first order dependency. According to the approximation we defined in Chapter 5, this corresponds to eliminating a considerable amount of two-dimensional sub-grids from the quadrature formula.

The fact that it is possible to represent these stochastic responses with a low polynomial order also suggests that the error associated with using first order perturbation techniques would be relatively low. If a forward perturbation

Table 6.6: Final PCE order \mathbf{p} at $t = 10s$ corresponding to using different tolerances within the adaptive algorithm.

Variable	PCE order		
	$\epsilon = 10^{-2}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$
R_s	1	2	2
δ_{clad}	1	1	1
ρ_{ext}	1	1	2
t_s	1	1	1
Λ	1	1	1
α_d	2	2	2
α_c	2	2	2
T_{in}	1	1	2
h_g	1	3	3
v_{in}	1	2	2

approach was used (since we are dealing with many responses localized at different times and a relatively low number of parameters) we would be able to evaluate the first order perturbation of the solution by running one nonlinear calculation and ten (corresponding to the number of input parameters) linearized ones. However, the implementation of such an approach would present two main difficulties. Firstly, the solution of the linearized problems would need the implementation of a separate mathematical model and a dedicated numerical solver. Secondly, some of the perturbations introduced for this UQ problem correspond to geometric (R_s) and time-dependent (t_s) parameters: their modeling, though possible in theory, is not straightforward and would introduce additional errors as explained in (Favorite and Bledsoe, 2010).

A second test was performed on the same problem. This time every unknown of the coupled system was considered to be a stochastic response whose PCE needed to be evaluated. Figures 6.3 and 6.4 show the standard deviation of the temperature field of the system at $t = 1s$ and $t = 10s$ obtained in both cases by using a 0.1% tolerance within the adaptive algorithm. In both cases the number of realizations needed for the convergence of the complete solution was the same required for the convergence of the small subset of responses included in Table 6.4. This can be explained since this subset, being a set of

6. Application to time-dependent multi-physics problems

local maxima, represents the most demanding outputs from the convergence point of view. In both figures a very large standard deviation is visible close

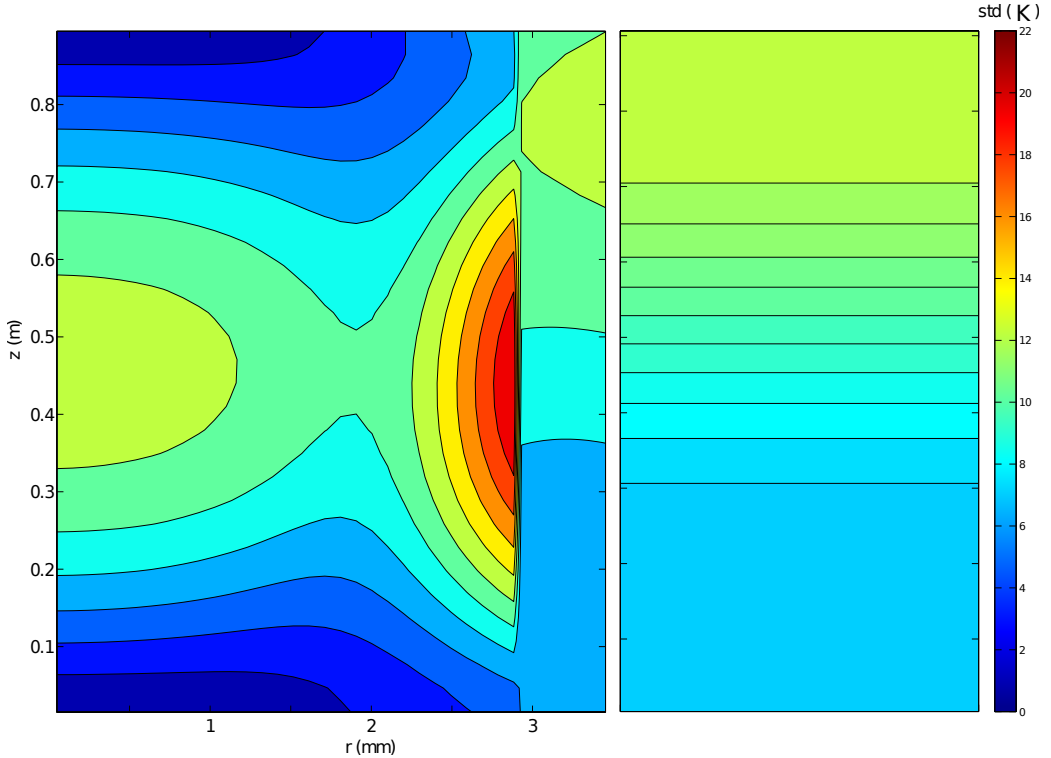


Figure 6.3: Standard deviation of the system temperature field at $t = 1s$, obtained by using $\epsilon = 0.1\%$. 329 realizations were needed to reach convergence.

to the fuel-cladding gap, this is expected since its heat transfer coefficient presents a large uncertainty (10%). While at the beginning of the transient the uncertainty profile is rather symmetrical with respect to the axial direction, after 10 seconds the uncertainty tends to be larger along the axial direction probably due to the uncertainties associated with the coolant inlet temperature and velocity.

Finally, a remarkable feature of spectral methods is that, once the spectral solution is known, it is possible to have a break down of the stochastic effects arising from each of the stochastic inputs. This is because our solution is now expressed as a function of the input parameters, which makes it easy to perform

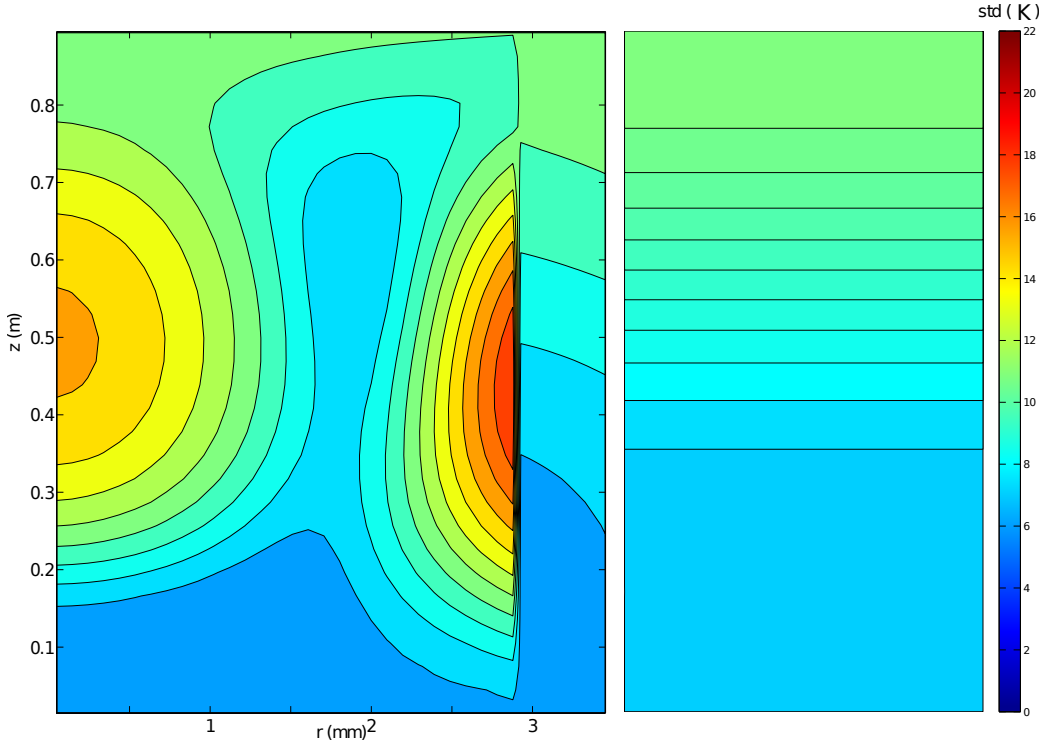


Figure 6.4: Standard deviation of the system temperature field at $t = 10s$, obtained by using $\epsilon = 0.1\%$. 269 realizations were needed to reach convergence.

a sensitivity analysis. Evaluating the coefficients of the PCE presented in Eq. 3.1 corresponds to determining the dependency of the stochastic outputs of interest with respect to the input random variables ξ used to model the random inputs. Once these coefficients are known we can estimate the effects of any single input parameter by considering the corresponding random variable ξ_i as the only non-zero term within the expansion 3.1.

For example, Figure 6.5 shows the influence on the solution at $t = 10s$ of the uncertainty introduced by the heat transfer coefficient of the gap. In the Figure the presence of a peak in the proximity of the gap is seen, which is responsible for the large standard deviation shown in Figure 6.4. At the same manner, it is possible to reconstruct first and higher order sensitivity coefficients from the spectral expansion, therefore allowing the possibility to perform a sensitivity

analysis when applying spectral techniques.

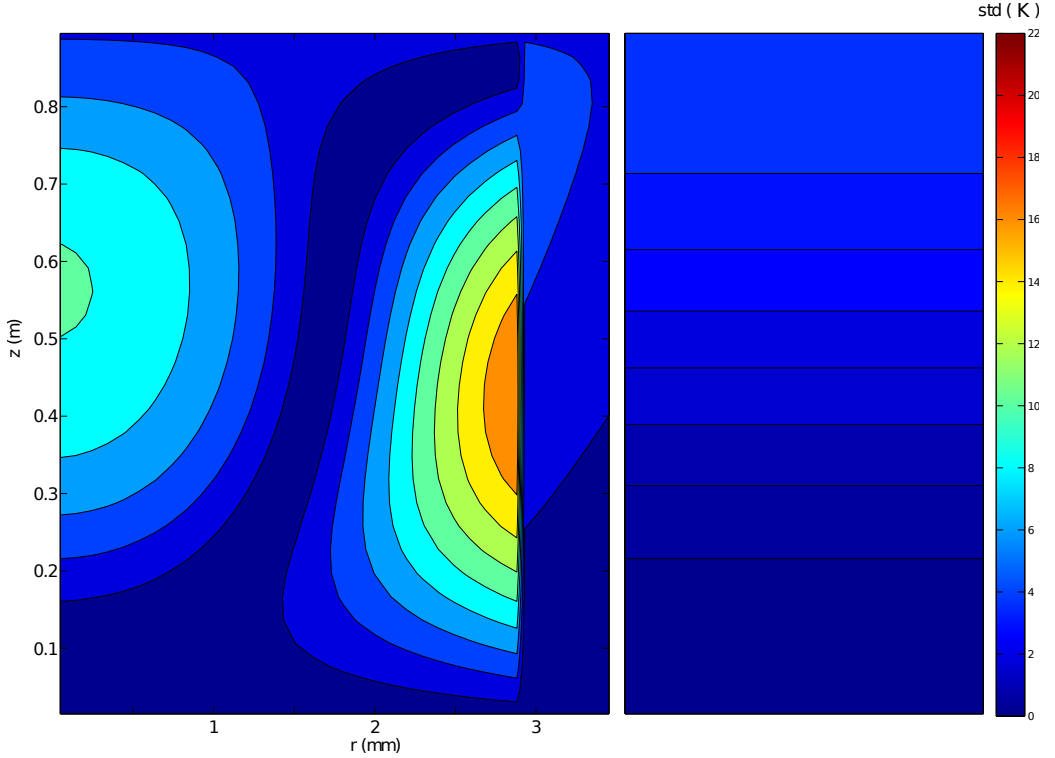


Figure 6.5: Standard deviation of the temperature field at $t = 10s$ obtained by considering the heat transfer coefficient of the gap as the only stochastic variable of the PCE.

6.5 Conclusions

In this chapter we presented the application of the adaptive sparse grid algorithm to a time-dependent multi-physics problem. This problem was formulated in order to reproduce a system that arises when performing safety analysis of nuclear reactors. In this context, the algorithm has been applied to propagate uncertainties for two reference transients simulating an accident scenario.

Although the stochastic input set was relatively large (10 random variables), in both cases it was possible to evaluate the spectral expansion with a number of realizations which is considerably lower than the one required by sampling approaches. This was because the most important contributions on the stochastic outputs were caused by a limited set of input parameters, thus enabling the adaptive algorithm to reproduce the stochastic information by collecting a limited amount of points in the right directions. The statistical moments of interest can also be reconstructed quite accurately with this reduced number of realizations. Furthermore the information associated with the final outputs allows to perform sensitivity analysis of the problem.

The optimization techniques introduced in Chapter 5 have been successfully used as well. This corresponds to an accurate reconstruction of the stochastic outputs, associated with a further reduction in the computational cost of the method. These results finally suggest that the adaptive sparse grid algorithm developed in this thesis could be applied to more complex cases, including larger sets of input quantities.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The main idea of this thesis is the development of a cost effective yet accurate uncertainty quantification technique suitable for reactor physics applications and multi-physics problems.

The role of uncertainty quantification is becoming increasingly important within the reactor physics field, especially when using “Best Estimate” models. Unfortunately, due to the complexity of reactor physics simulations, performing uncertainty propagation is either associated with large computational requirements or with the development of additional large scale adjoint problems. Finding a way to perform this task in an accurate and computationally cheap way is therefore a challenge.

The key contributions presented with this thesis can be summarized as follows:

- It is demonstrated that the extension of the adjoint formalism to higher order perturbation components corresponds to a considerable increase

in terms of computational requirements especially when dealing with nonlinear problems. This is because of the fact that, unlike for linear problems, it is not possible to perform a reliable parameter ranking of the Hessian matrix components. However, this preliminary analysis concludes that adjoint techniques can be useful tools to demonstrate how uncertainty propagation is dominated by low-order interactions and by dominant effects.

- From the preliminary analysis of spectral techniques performed in the third chapter we conclude that non-intrusive methods represent, due to their mathematical formulation, the most flexible way to implement this type of techniques for multi-physics problems. This is because non-intrusive implementations allow to adapt the spectral expansion according to the type of response one is evaluating. We show that for coupled time-dependent problems higher order expansions are usually needed only for short intervals while the rest of the stochastic solution can be represented by using a lower order truncation. These convergence properties are demonstrated for a coupled point-kinetic/thermal-hydraulics problem which also implies that for this type of transient the convergence of the spectral expansion is assured, which is not the case for every time-dependent problem.
- We demonstrate that the adaptive algorithm developed within the thesis can represent, in many situations, a cheap and accurate alternative to standard sampling approaches. It is shown how the particular construction of the spectral basis, based on a convergence check performed considering each random direction to be independent, can further reduce the number of realizations needed to build the spectral outputs.
- Two optimization techniques have also been presented in this thesis. These techniques are based on the particular way the sparse grid is built and can be used to reduce the computational requirement of the algorithm when in presence of parameters characterized by a linear dependency with respect to the stochastic outputs. Neglecting the cross-correlation between these type of parameters can be used to exclude *a priori* a considerable set of realizations from the final quadrature rule. The most important aspect, when using these techniques, is that the adaptive spectral method can be computationally more convenient than traditional sampling even when dealing with moderately large sets of

input parameters. Unfortunately, this convenience is strongly dependent on the regularity of the response surface of the problem.

- The adaptive method developed in this thesis can also represent a computationally convenient tool to perform uncertainty propagation of time-dependent problems which may be encountered while performing safety analysis. Since this type of problem is usually associated with a moderate number (10-20) of random input parameters, we presented the application of the adaptive spectral method to an illustrative case characterized by a comparable number of stochastic inputs. The results show that for this type of problem spectral techniques can provide a deterministic stochastic solution with a fraction of the realizations needed to reach convergence with Monte Carlo sampling.

In theory, though the algorithm is generally proven to be effective for multi-physics applications, it still presents possible ways of improvement. An overview of this topic is presented in the next section.

Recommendations

The main disadvantage of the adaptive spectral method developed in this thesis is that its computational cost directly depends on the regularity of the output response surface. This means that, depending on properties of the surface associated with a stochastic response, the method could become computationally more expensive than random sampling even when dealing with a relatively low number of inputs. One could already determine whether a standard sampling approach is computationally more convenient during the first part of the algorithm, after the solution of the one-dimensional random problems has been obtained. An additional check could be introduced after this step to eventually stop the sparse grid construction and start a sampling procedure.

For all the tests presented in this thesis, the indicator used to evaluate the sub-grids constituting the final sparse grid was based on the evaluation of the mean and the standard deviation of the response. In some cases this led to results which were accurate in terms of the first two statistical moments, but inaccurate when reproducing the probability density function of the response.

In theory, one could introduce the evaluation of the skewness as an additional check. However, as explained in Chapter 3, this implementation is not trivial since the analytical expression of the skewness involves the evaluation of a very large set of triple products.

Furthermore, the algorithm convergence could be also improved by using the derivative information associated with the response of interest. Assuming this information is available, for example, by means of adjoint codes, one could find a way of exploiting it within the adaptive algorithm.

The computational cost associated with the method could also be reduced by implementing an adaptive polynomial basis construction. The algorithm developed within this thesis is based on the definition of a reduced polynomial expansion whose truncation is determined during the first part of the procedure. This means that the rest of the projection process is performed using a fixed basis which contains higher order multi-dimensional polynomials. However, as many of the higher polynomials characterizing this basis give a negligible contribution to the stochastic outputs, it would be advantageous to find a way to avoid evaluating their projections. This could be done, for example, by defining a method to adaptively build the polynomial expansion within the spectral algorithm.

In general, the work done in this thesis represents a solid basis for future developments of adaptive spectral techniques. Our conclusions and recommendations suggest that there is a limit in the number of random inputs below which the method proposed in this thesis will be more accurate and computationally more convenient than standard sampling approaches. The recommendations hitherto discussed are an example of how to increase this limit for future applications.

BIBLIOGRAPHY

- H. S. Abdel-Khalik, P. J. Turinsky and M. A. Jessee. Efficient Subspace Methods-Based Algorithms for Performing Sensitivity , Uncertainty , and Adaptive Simulation of Large-Scale Computational Models. *Nuclear Science and Engineering*, **159**, 256, 2008.
- M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1964.
- A. Alekseev, I. Navon and M. Zelentsov. The estimation of functional uncertainty using polynomial chaos and adjoint equations. *International Journal for Numerical Methods in Fluids*, **67**, 328, 2011.
- M. Alexe et al. Using Automatic Differentiation in Sensitivity Analysis of Nuclear Simulation Models. *ANS Winter meeting*, **1**, 235, 2010.
- G. Aliberti et al. Nuclear data sensitivity, uncertainty and target accuracy assessment for future nuclear systems. *Annals of Nuclear Energy*, **33**, 700, 2006.
- Y. Bang and H. S. Abdel-Khalik. Projection-based second order perturbation theory. *Annals of Nuclear Energy*, 2012.
- Y. Bang, C. Wang and H. S. Abdel-Khalik. State-Based Adjoint Method for Reduced Order Modeling. *Transport Theory and Statistical Physics*, **41**, 101, 2012.

- M. Bieri and C. Schwab. Sparse high order fem for elliptic spdes. *Comput. Methods Appl. Mech.*, **198**, 1149, 2009.
- D. Cacuci. *Sensitivity and Uncertainty Analysis, Volume 1: Theory*. Chapman and Hall CRC, 2003.
- M. Chadwick et al. ENDF/B-VII.0: Next Generation Evaluated Nuclear Data Library for Nuclear Science and Technology. *Nuclear Data Sheets*, **107**, 2931, 2006.
- F. D Auria, C. Camargo and O. Mazzantini. The Best Estimate Plus Uncertainty (BEPU) approach in licensing of current nuclear reactors. *Nuclear Engineering and Design*, **248**, 317, 2012.
- J. Duderstadt and L. Hamilton. *Nuclear Reactor Analysis*. John Wiley and sons, 1976.
- M. Eaton and M. M. R. Williams. A probabilistic study of the influence of parameter uncertainty on solutions of the neutron transport equation. *Progress in Nuclear Energy*, **52**, 580, 2010.
- J. Favorite and K. Bledsoe. Eigenvalue sensitivity to system dimensions. *Annals of Nuclear Energy*, **37**, 522, 2010.
- J. Favorite and M. Stacey. Variational Estimates of Point-Kinetics Parameters. *Nuclear Science and Engineering*, **121**, 353, 1995.
- E. Fichtl. *Stochastic methods for Uncertainty Quantification in radiation transport*. PhD thesis. University of New Mexico, 2009.
- E. D. Fichtl and A. K. Prinja. The stochastic collocation method for radiation transport in random media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **112**, 646, 2011.
- J. Foo, X. Wan and G. Karniadakis. The multi-element probabilistic collocation method: error analysis and simulation. *Journal of Computational Physics*, **227**, 9572, 2008.
- B. Ganapathysubramanian and N. Zabararas. Sparse grid collocation schemes for stochastic natural convection problems. *Journal of Computational Physics*, **225**, 652, 2007.

- A. Gandini. A generalized perturbation method for bi-linear functionals of the real and adjoint neutron fluxes. *J. of Nuclear Energy*, **21**, 755, 1967.
- A. Gandini. Higher order time-dependent generalized perturbation theory. *Nuclear Science and Engineering*, **67**, 91, 1978.
- Generation IV International Forum. A Technology Roadmap for Generation IV Nuclear Systems. Technical Report GIF-002-00, U.S. DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum, 2002.
- M. Geritsma et al. Time dependent generalized polynomial chaos. *Journal of Computational Physics*, **229**, 8333, 2010.
- T. Gerstner and M. Griebel. Numerical integration using sparse grids. *Numer. Algorithms*, **18**, 209, 1998.
- T. Gerstner and M. Griebel. Dimension-Adaptive Tensor-Product Quadrature. *Computing*, **71**, 65, 2003.
- R. Ghanem and P. Spanos. *Stochastic Finite Elements: a Spectral Approach*. Dover publications, 1991.
- R. Ghanem and P. Spanos. Spectral techniques for stochastic finite elements. *Archives of Computational Methods in Engineering*, **4**, 63, 1997.
- L. Gilli et al. Sensitivity analysis to a coupled neutronic/thermal-hydraulics transient in a fast reactor using adjoint techniques. In *International Conference on Mathematics, Computational Methods and Reactor Physics (M&C 2011)*. American Nuclear Society, Rio de Janeiro, Brazil, 2011.
- E. Greenspan and D. Gilai. Second-order generalized perturbation theory for source-driven systems. *Nuclear Science and Engineering*, **68**, 1, 1978.
- E. Greenspan, Y. Kami and D. Gilai. Higher order effects in cross-section sensitivity analysis. Technical Report ORNL/RSIC-42, Oak Ridge National Laboratory, 1979.
- A. Griewank. Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optimization Methods and Software*, **1**, 35, 1992.

- A. Hagues, M. M. R. Williams and M. Eaton. A probabilistic study of the effect of retardation factor uncertainty using a compartment model for radionuclide release into the biosphere. *Annals of Nuclear Energy*, **37**, 1197, 2010.
- J. Helton and F. Davis. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. Technical Report SAND2001-0417, Sandia National Laboratories, 2002.
- A. C. Hindmarsh et al. SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers. *ACM Transactions on Mathematical Software*, **31**, 363, 2005.
- IAEA. Evaluation of benchmark calculations on a fast power reactor core with near zero sodium void effect. Technical Report IAEA-TECDOC-731, IAEA, 1994.
- M. Ionescu-bujor and D. G. Cacuci. A Comparative Review of Sensitivity and Uncertainty Analysis of Large-Scale Systems I : Deterministic Methods. *Nuclear Science and Engineering*, **147**, 189, 2004.
- A. Koning and D. Rochman. Towards sustainable nuclear energy: Putting nuclear physics to work. *Annals of Nuclear Energy*, **35**, 2024, 2008.
- J. Kophazi and D. Lathouwers. Three-dimensional transport calculation of multiple alpha modes in subcritical systems. *Annals of Nuclear Energy*, **50**, 167, 2012.
- O. Le Maitre and O. Knio. *Spectral Methods for Uncertainty Quantification*. Springer, 2010.
- O. Le Maitre et al. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *Journal of Computational Physics*, **197**, 502, 2004.
- O. P. Le Maitre et al. A Stochastic Projection Method for Fluid Flow. *Journal of Computational Physics*, **181**, 9, 2002.
- B. A. Lockwood and M. Anitescu. Gradient-Enhanced Universal Kriging for Uncertainty Propagation. *Nuclear Science and Engineering*, **170**, 168, 2012.
- L. Mathelin, M. Hussaini and T. Zang. Stochastic approaches to uncertainty quantification in CFD simulations. *Numerical Algorithms*, **38**, 209, 2005.

- H. Najm. Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics. *Annu. Rev. Fluid Mech.*, **41**, 35, 2009.
- NEA. *Uncertainty and target accuracy assessment for innovative systems using recent covariance data evaluations*. OECD-NEA, 2008.
- NEA. *International handbook of evaluated criticality safety benchmark experiments*, volume NEA/NSC/DOC(95)03. OECD-NEA, 2010.
- D. Ozyurt and P. I. Barton. Cheap second order directional derivatives of stiff ode embedded functionals. *SIAM J. SCI. COMPUT.*, **26**, 1725, 2005.
- A. Prinja and E. Larsen. *General Principles of Neutron Transport*. Handbook of Nuclear Engineering. Springer, 2010.
- G. Rimpault et al. The ERANOS code and data system for fast reactor neutronic analyses. In *Proceedings of the 2002 PHYSOR CONFERENCE*, 2002.
- D. Rochman, A. Koning and S. van der Marck. Uncertainties for criticality-safety benchmarks and keff distributions. *Annals of Nuclear Energy*, **36**, 810, 2009.
- D. Rochman et al. Nuclear data uncertainty propagation: Perturbation vs. Monte Carlo. *Annals of Nuclear Energy*, **38**, 942, 2011.
- O. Roderick, M. Anitescu and P. Fischer. Polynomial Regression approaches using derivative information for Uncertainty Quantification. *Nuclear Science and Engineering*, **164**, 122, 2010.
- Y. Ronen. *Uncertainty Analysis*. Chapman and Hall CRC, Boca Raton, United States, 1988.
- W. van Rooijen and D. Lathouwers. Sensitivity analysis for delayed neutron data. *Annals of Nuclear Energy*, **35**, 2186, 2008.
- S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, **4**, 240, 1963.
- A. Waltar and A. Reynolds. *Fast Breeder Reactors*. Pergamon Press, 1981.

- X. Wan and G. E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, **209**, 617, 2005.
- X. Wan and G. E. Karniadakis. Long term behavior of polynomial chaos in stochastic flow simulations. *Comput. Methods Appl. Mech. Engrg.*, **195**, 5582, 2006.
- N. Wiener. The Homogeneous Chaos. *American Journal of Mathematics*, **60**, 897, 1938.
- M. L. Williams. *CRC Handbook for Nuclear Reactors Calculations vol III - Perturbation Theory for Nuclear Reactor Analysis*. CRC Press, 1986.
- M. M. R. Williams. Polynomial Chaos functions and stochastic differential equations. *Annals of Nuclear Energy*, **33**, 774, 2006.
- M. M. R. Williams. Polynomial Chaos functions and neutron diffusion. *Nuclear Science and Engineering*, **155**, 109, 2007.
- G. E. Wilson. Historical insights in the development of Best Estimate Plus Uncertainty safety analysis. *Annals of Nuclear Energy*, 2012.
- J. Witteveen, G. Loeven and H. Bijl. An adaptive stochastic finite elements approach based on Newton-Cotes quadrature in simplex elements. *Computers and Fluids*, **38**, 1270, 2009.
- J. A. S. Witteveen. *Efficient and Robust Uncertainty Quantification for Computational Fluid Dynamics and Fluid-Structure Interaction*. Ph.D. thesis, Technische Universiteit Delft, 2009.
- J. A. S. Witteveen and H. Bijl. An Efficient Monomial Chaos Approach or UQ in Nonlinear Computational Models. In *AIAA*, 2006.
- C. Wu and M. Hamada. *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley and sons, 2000.
- G. D. Wyss and K. H. Jorgensen. A User's Guide to LHS : Sandias Latin Hypercube Sampling Software. Technical Report February, Sandia National Laboratory, 1998.

- D. Xiu and G. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, **24**, 619, 2002.
- X. Yang et al. Adaptive ANOVA decomposition of stochastic incompressible and compressible flows. *Journal of Computational Physics*, **231**, 1587, 2012.

APPENDIX A

APPENDIX

Appendix: elements of probability theory

In this appendix we provide a brief overview about the main concepts of probability theory used within this thesis. We will first introduce the notion of probability density function and its mathematical definition. Then, we will discuss about the statistical moments associated with a stochastic variable.

Elements of probability theory: the probability density function of a random variable

Given a random variable R defined over the stochastic support θ , its probability density function is a distribution used to quantify the likelihood for this random variable to assume a certain value θ_i . This distribution is defined such that the probability for the stochastic variable to fall within a particular region is given by the integral of the probability density function over this region

$$P_R[a \leq \theta \leq b] = \int_a^b p_R(\theta) d\theta$$

Where $P_R[a < \theta < b]$ is the probability for R to be characterized by a value between a and b and p_R is the probability density function associated with the random variable R. This probability distribution is characterized by two important mathematical aspects: firstly, it must be nonnegative all over the realization space, and secondly the outcome of integrating this function over the whole stochastic support must be equal to one.

If a random variable admits a probability density function the knowledge of this distribution will give us a complete insight on how this random variable behaves over the stochastic domain on which it is defined. Usually, quantitative measures based on the definition of probability density function are introduced in order to better describe the features of a random variable. These measures are usually known as statistical moments, in the next section a brief description of these moments is presented.

Statistical moments of a random variable

If we are dealing with a random variable R which admits a probability density function, its statistical moments are quantitative measures based on the behavior of the probability density function, used to better characterize the random variable. For example, we can measure the tendency for the random variable to assume a certain value by introducing its mean (or expected) value μ in the following manner

$$\mu = E[R] = \int_{-\infty}^{+\infty} \theta p_R(\theta) d\theta$$

This measure is also known as the first statistical moment of the random process and defines the stochastic realization which is bounded to happen more often in case the random process was repeated for an infinite amount of times. When in presence of an integral which does not have an analytical solution we can evaluate the mean of a random variable by using the statistical estimator defined in Section 1.4 within the first chapter of this thesis.

The second statistical moment of a random variable is used to measure how its realizations are “spread” over the stochastic domain. Within the framework of this thesis the measure of this spread defines the uncertainty associated with a random variable characterized by the mean value μ and it is therefore of

fundamental importance. This statistical moment is defined as follows

$$Var = E[(R - \mu)^2]$$

Alternatively, the measure of the variance of a variable is often substituted by its standard deviation $\sigma^2 = Var$. The fact that the moment is also defined as the expected value of a functional which depends on the random variable means that it can also be evaluated by using the statistical estimator defined in Section 1.4 once the mean value of the random variable is known. Intuitively, the realization of a variable associated with a low standard deviation will be more likely to be close to its mean value.

Finally, we introduce another statistical moment, known as the skewness. This moment can be introduced by using a definition similar to the one used to define the variance.

$$\gamma = E\left[\left(\frac{R - \mu}{\sigma}\right)^3\right]$$

The skewness of a random variable defines how “skewed” its probability density function is. A positive skewness corresponds to saying that it is more likely for a realization of the variable to be larger than the mean. Within the uncertainty quantification framework this aspect can have important implications considering that we might be interested in the probability of the variable being above a certain value rather than in the mean and standard deviation.

Summary

This thesis presents the development and the implementation of an uncertainty propagation algorithm based on the concept of spectral expansion. The first part of the thesis is dedicated to the study of uncertainty propagation methodologies and to the analysis of spectral techniques. The concepts introduced within this preliminary analysis are successively used for the derivation of the spectral algorithm.

In Chapter 2 we discuss the application of higher order adjoint perturbation theory for coupled problems. This method is relatively easy to implement once the first order adjoint problem is defined, however it is computationally expensive. It is shown, for example, that the number of additional adjoint calculations needed to build the Hessian matrix of a response corresponds, for nonlinear problems, to two times the number of input parameters. It is also shown that for linear problems this number can be halved. It is also discussed that for linear problems it is possible to perform a ranking of the higher order perturbation components, while for nonlinear ones this is not the case. In general, higher order adjoint perturbation theory can be a useful tool to understand uncertainty propagation phenomena.

In Chapter 3 an overview of spectral techniques for uncertainty quantification is presented. The mathematical backgrounds of two approaches, defined as intrusive and non-intrusive, are discussed. These approaches are applied to perform uncertainty quantification of a simplified coupled time-dependent problem. The illustrative example shows how non-intrusive approaches are relatively easy to apply while intrusive approaches are quite challenging from the implementation point of view. The curse of dimensionality affecting spectral techniques is also discussed. The example also demonstrates that for time-dependent problems, the convergence of spectral expansions required to represent stochastic outputs varies considerably during the transient. From this point of view, non-intrusive approaches allow the usage of different expansion orders at different times, thereby reducing the computational requirements. Using these initial conclusions as a starting point, an algorithm based on the definition of Polynomial Chaos Expansion is developed. Chapter 4 introduces this new algorithm for the application of quadrature based spectral techniques. This algorithm is based on the notion of sparse grid and its application is divided into two main steps. Firstly, the algorithm adds quadrature points

exclusively along the main axes of the stochastic domain. During this phase the convergence of the PCE is assessed and a reduced multi-dimensional PCE is defined. Secondly, this reduced PCE is then used within the second part of the algorithm which focuses on the addition of higher dimensional sub-grids to the final quadrature rule. The adaptive sparse grid algorithm is tested for a reference stochastic case defined by using a simple source detector problem. The algorithm is first validated by comparing it to another sparse grid integration approach found in literature. It is successively shown how the particular construction of the spectral basis, based on a convergence check performed considering each random direction to be independent, can further reduce the number of realizations needed to build the spectral outputs.

In Chapter 5 two cost reduction techniques which take advantage of the peculiar definition of the algorithm are presented. These techniques are proven to be effective in the reduction of quadrature points needed to reach convergence. Two uncertainty propagation examples are also considered. The method has been proven to be particularly effective for reactor physics applications, mainly because of the fact that higher order propagation phenomena are usually dominated by a limited set of input parameters. It is also shown, with the first example, that the convergence rate of the adaptive quadrature algorithm directly depends on the differentiability of the response surface.

Chapter 6 shows another application of the adaptive sparse grid algorithm, this time to a time-dependent multi-physics problem. This problem is formulated in order to reproduce the type of system that arises when performing safety analysis. Two reference transients simulating an accident scenario of fast reactors are considered. Even in this case the adaptive algorithm proves to be very effective, being capable of reproducing all the stochastic outputs of interest with a relatively low number of realizations.

In conclusion, adaptive spectral methods represent a computationally efficient uncertainty quantification technique when in presence of a moderately large set of random input parameters. However, this number strongly depends on the regularity of the response surface. Several strategies could be adopted in order to increase this number and make the method more appealing for a larger set of problems. An overview of these possibilities is presented in the final recommendation section of the thesis.

Samenvatting

Dit proefschrift bespreekt de ontwikkeling en de implementatie van een onzekerheidspropagatie algoritme, gebaseerd op het concept van spectrale expansie. Het eerste deel van het proefschrift is gewijd aan de bestudering van diverse onzekerheidspropagatie methoden en aan de analyse van spectrale technieken. De begrippen geïntroduceerd binnen deze eerste analyse worden achtereenvolgens gebruikt voor de afleiding van het spectrale algoritme.

In Hoofdstuk 2 wordt de toepassing van hogere orde adjoint verstoringstechnieken besproken voor gekoppelde problemen. Deze methode is relatief eenvoudig te implementeren zodra de eerste orde adjoint probleem is gedefinieerd, maar vereist veel rekenkracht. Analyse laat zien dat het aantal extra adjoint berekeningen nodig voor het opzetten van een Hessian matrix, in geval van niet-lineaire problemen, overeenkomt met tweemaal het aantal input parameters. We tonen ook aan datvoor lineaire problemen dit aantal kan worden gehalveerd. Ook wordt besproken hoe het voor lineaire problemen in principe mogelijk is om een rangorde van de hogere orde verstoring onderdelen aan te brengen terwijl het voor niet-lineaire problemen niet het geval is. In het algemeen kunnen hogere orde adjoint perturbatietheorie een nuttig instrument zijn om onzekerheidspropagatie te begrijpen.

In Hoofdstuk 3 wordt een overzicht van spectrale technieken voor onzekerheidskwantificering gepresenteerd. De wiskundige achtergrond van twee benaderingen genaamd intrusief en niet-intrusief, worden besproken. Deze twee aanpakken worden toegepast om de onzekerheid kwantificering van een vereenvoudigd gekoppeld tijdsafhankelijk probleem uit te voeren. Een illustratief voorbeeld laat zien hoe de niet-intrusieve aanpak relatief eenvoudig uit te voeren zijn, terwijl intrusieve methoden zeer uitdagend zijn vanuit het implementatie oogpunt. De invloed van de dimensionaliteit op spectrale technieken wordt hier ook besproken. Het voorbeeld toont ook aan dat voor tijdsafhankelijke problemen, de convergentie van spectrale uitbreidingen vereist voor het representeren van stochastische uitgangssignalen aanzienlijk varieert tijdens de tijdsafhankelijke simulatie. Vanuit dit standpunt, maken niet-intrusieve benaderingen het mogelijk om gebruik te maken van verschillende expansie orden op verschillende tijden, waardoor de benodigde rekenkracht vermindert. Met deze eerste conclusie als uitgangspunt, is een algoritme op basis van Polynomial Chaos Expansion ontwikkeld.

In Hoofdstuk 4 wordt dit nieuwe algoritme voor de toepassing van de kwadratuur gebaseerde spectrale technieken gintroducteerd. Dit algoritme is gebaseerd op het principe van sparse grid en de toepassing daarvan is onderverdeeld in twee stappen. Gedurende de eerste staop voegt het algoritme uitsluitend kwadratuur punten toe langs de hoofdassen van de stochastische domein. Tijdens deze fase wordt de convergentie van de PCE beoordeeld en een gereduceerde multidimensionale PCE wordt gedefinieerd. Deze gereduceerde PCE wordt dan gebruikt in de tweede stap van het algoritme dat zich richt op de toevoeging van hoger- dimensionale sub-roosters aan de uiteindelijke kwadratuurformule. Het resulterende adaptieve algoritme is getest op een referentie stochastisch geval gedefinieerd met behulp van een eenvoudig bron detector probleem. Het algoritme is eerst gevalideerd door het te vergelijken met een andere sparse grid integratie benadering, aanwezig in de literatuur. Volgens wordt aangetoond hoe de specifieke constructie van de spectrale basis, gebaseerd op een convergentie controle die uitgevoerd is met als beschouwing dat elke willekeurige richting onafhankelijk is, tot een verdere verlaging van het aantal benodigde realisaties voor het bouwen van spectrale outputs kan leiden.

In Hoofdstuk 5 worden twee kostreductie technieken voorgesteld die profiteren van de specifieke definitie van het algoritme. Deze technieken hebben hun effectiviteit bewezen in het verminderen van het aantal benodigde kwadratuur punten voor convergentie. Twee onzekerheidspropagatie voorbeelden worden gepresenteerd. De methode blijkt bijzonder effectief voor toepassingen op reactor fysica, vooral vanwege het feit dat hogere orde propagatie verschijnselen meestal gedomineerd worden door een beperkt aantal invoerparameters. Het is ook aangetoond, d.m.v. het eerste voorbeeld, dat de convergentie snelheid van de adaptieve kwadratuur algoritme direct afhankelijk is van de differentieerbaarheid van de respons oppervlak.

Hoofdstuk 6 toont een andere toepassing van het adaptieve sparse grid algoritme, ditmaal een tijdsafhankelijke multi-fysica probleem. Dit probleem lijkt wat betreft structuur op op het type systeem dat ontstaat bij het uitvoeren van een veiligheidsanalyse. Twee referentie tijdsafhankelijke simulaties die een ongeval scenario van snelle reactoren simuleren worden beschouwd. Zelfs in dit geval blijkt het adaptieve algoritme zeer effectief te zijn. Het is in staat alle belangrijke uitkomsten met een relatief klein aantal realisaties te kunnen reproduceren.

Concluderend zijn adaptieve spectrale methoden een rekenkundig efficiënte

onzekerheidskwantificering techniek voor gebruik met een matig grote set van willekeurige invoerparameters. Dit aantal is sterk afhankelijk van de regelmatigheid van het respons-oppervlak. Verschillende strategieën kunnen worden aangenomen om dit aantal verder te vergroten en de werkwijze aantrekkelijker te maken voor problemen van bredere aard. Een overzicht van deze mogelijkheden wordt gepresenteerd in het uiteindelijke aanbevelingen van dit proefschrift.

(Dutch translation provided by Mr. Pierre Willem Lenoir)

Acknowledgements

First of all I would like to thank Tim van der Hagen and Jan Leen Kloosterman for giving me the opportunity to perform my PhD research in Delft. Their decision to hire me turned out to be not only a great career opportunity but also a choice that changed my life, reason for which I will be forever grateful to them. I would like to thank Danny Lathouwers for the supervision he provided me during these four years and for all the insightful comments and feedback I promptly got from his many reviews of my work. I am grateful to Arjan Koning and Dimitri Rochman for the productive meetings we had during my PhD and for sharing their knowledge on the uncertainty propagation field. I need to thank Zoltan for the many useful conversations we had about perturbation and spectral methods and Casper for coping with me as a supervisor and helping me pursue my research on spectral techniques. Also, thanks to Jozsef for his help and assistance with the S_n code and, together with Christophe, for showing me how brilliant researchers function. A big thanks to Ine for all the help and assistance she gave me every time I had to deal with some bureaucratic issue.

I should thank all the members of PNR for the share of fun I had during my time in Delft. Starting from Stuart, my office mate for the last four years, who has always patiently coped with my Italian English and taught me the rules of cricket. The same gratitude goes to all the colleagues and students I had the pleasure to meet at PNR and in the Koepeltje, in random order: Bart, Gert Jan, Frank, Jurriaan, Johan and Jitka, Stavros, Norbert, Dimitrios, Karoly, Wim, Johan, Leon, and Edith.

Living in Delft has been great also from the social point of view, on this regard I should really thank some members of the large Italian community in Delft like Samy, Vito, Donato, Paola, Francesca and Filippo. The latter two have been also particularly important to me personally since they are the reason I met my wife.

I cherish great memories of the house in Pijnacker where I lived with Marco, Marc and Markus, thank you guys for the good times. A hearty thanks to the Pootstraat gang as well: Pierre, Nath, Felix, Patsie, Nick and Muge. I had great fun during the time we spent together and I hope this will continue for the years to come.

Acknowledgements

Last but not least I should thank my family. Devo ringraziare la mia famiglia: Gisella, Claudio, Gianfranco, e Andrea per tutto l'amore e il supporto incondizionato che mi hanno sempre dato e per avere accettato serenamente il fatto che sia così lontano senza avermelo fatto mai pesare.

Lo stesso vale per i miei amici e famiglia estesa in Italia: grazie a Stefano, Matteo, Riccardo, Federico, Roberto, Alessandra, Silvia, Andrea e Giulia, Guido, Patrizia, Angelica, Mila, Fabio e Monica per avermi fatto sentire nuovamente a casa ogni volta che sono tornato a Torino.

My family now includes my in-laws in Chittoor as well: thank you Radhagaru and Dayananda-garu for accepting me as if I was your son and for making me feel home even when I am on the other side of the planet.

And of course, my Bangaru Konda. I cannot properly express in words how happy I am with you and how much your love and support are making me a better person. The way our random walks met in Delft makes me believe that I am one of the luckiest men on the planet.

List of Publications

L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Application of Sensitivity Analysis to a Coupled Neutronics/Thermal-Hydraulics Transient in a Fast Reactor using Adjoint Techniques. In *Proceedings of the International Conference on Mathematics and Computational Methods applied to Nuclear Science and Engineering*, Rio de Janeiro, Brazil, 2011.

L. Gilli, D. Lathouwers and J.L. Kloosterman. Uncertainty Analysis of Radiation Transport Problems Using Non-Intrusive Polynomial Chaos Techniques. In *Proceedings of the International Conference on Transport Theory*, Portland, US, 2011.

C. T'Joel, L. Gilli, M. Rodhe. Sensitivity analysis of numerically determined linear stability boundaries of a supercritical heated channel. *Nuclear Engineering and Design*, **241**, 3879, 2011.

L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Performing uncertainty analysis of a nonlinear Point-kinetics/Lumped Parameters problem using Polynomial Chaos techniques. *Annals of Nuclear Energy*, **40**, 35, 2012.

L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Second Order Perturbation Theory for Nonlinear Time-dependent Problems: Application to a Simplified Coupled Problem. In *Proceedings of PHYSOR 2012*, Knoxville, US, 2012.

L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Ap-

plying second-order adjoint perturbation theory to time-dependent problems. *Annals of Nuclear Energy*, **53**, 9, 2013.

L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Uncertainty Quantification for criticality problems using non-intrusive and adaptive Polynomial Chaos techniques. *Annals of Nuclear Energy*, **56**, 71, 2013.

L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Development of an adaptive non-intrusive spectral technique for uncertainty quantification and application to a multi-physics problem. To appear in *Nuclear Science and Engineering*

Z. Perko, L. Gilli, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Adaptive Polynomial Chaos Techniques for Uncertainty Quantification of a Gas Cooled Fast Reactor Transient. In *Proceedings of the International Conference on Mathematics and Computational Methods applied to Nuclear Science and Engineering*, Sun Valley, USA, 2013.

L. Gilli, C. Hoogwerf, D. Lathouwers, J.L. Kloosterman and T.H.J.J. van der Hagen. Uncertainty Quantification of a Radionuclide Release Model Using an Adaptive Spectral Technique. In *Proceedings of the International Conference on Mathematics and Computational Methods applied to Nuclear Science and Engineering*, Sun Valley, USA, 2013.

Curriculum Vitae

Luca Gilli was born in Rivoli (Italy) on the 15th of August 1982. He grew up in the alpine suburbs of Turin where he attended the Charles Darwin lyceum in Rivoli. After graduating from high school he enrolled in the Energy Engineering bachelor program at Politecnico di Torino in 2002 where his passion for energy and nuclear engineering was kindled. He received his bachelor degree in 2006 with a thesis on the use of Laplace transform techniques for reactor kinetics. In 2006 he continued with his masters' education in Energy and Nuclear Engineering at Politecnico di Torino, which culminated in 2008 with a thesis on plasma instabilities in a Reverse Field Pinch reactor. His masters' degree was awarded with the *cum laude* distinction.

In November 2008 he joined the Faculty of Applied Sciences as a PhD student at the Reactor Institute Delft. The current thesis is the outcome of the 4 year research performed at the Physics of Nuclear Reactors group (currently NERA) in Delft.

In February 2013 he started working at the Nuclear Research and consultancy Group (NRG) within the Fuel Management Services division. He currently spends his time traveling between the Netherlands, India and Italy.

