Matti Leinonen

Stochastic Galerkin Finite Element Method with Log-normal Random Field

School of Science

Master's thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology in the Degree Programme in Engineering Physics and Mathematics.

Espoo, November 14, 2011

Thesis supervisor:

Prof. Olavi Nevanlinna

Thesis instructors:

D.Sc. (Tech.) Harri Hakula

D.Sc. (Tech.) Nuutti Hyvönen

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE
MASTER'S THESIS

Author: Matti Leinonen

Title of thesis:
Stochastic Galerkin Finite Element Method with Log-normal Random Field

Title of thesis in Finnish:
Stokastinen Galerkinin elementtimenetelmä log-normaalin satunnaiskentän tapauksessa

Date: November 14, 2011 **Language**: English **Number of pages**: 5+58

Degree programme:
Degree Programme in Engineering Physics and Mathematics

Major subject: Mathematics
Minor subject: Engineering Physics

Chair: Mathematics **Code**: Mat-1

Supervisor: Prof. Olavi Nevanlinna

Instructors: D.Sc. (Tech.) Harri Hakula, D.Sc. (Tech.) Nuutti Hyvönen

Stochastic Galerkin finite element method (sGFEM) is introduced together with the associated mathematical tools, and implemented for a one-dimensional stochastic linear elliptic boundary value problem that can be considered as the stochastic conductivity equation reduced to one dimension. The conductivity coefficient of the model problem is assumed to be a log-normal random field with a known mean field and covariance function. The contact impedances of the model are assumed to be known log-normal random variables. According to numerical tests, the sGFEM is found to be a feasible choice when the variances of the stochastic parameters are not huge.

Keywords: stochastic Galerkin finite element method, polynomial chaos, Karhunen–Loève expansion, log-normal distribution, random field, sGFEM, sFEM

AALTO-YLIOPISTO
PERUSTIETEIDEN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Matti Leinonen

Työn nimi:
Stokastinen Galerkinin elementtimenetelmä log-normaalin satunnaiskentän tapauksessa

Työn nimi englanniksi:
Stochastic Galerkin Finite Element Method with Log-normal Random Field

Päivämäärä: 14. marraskuuta 2011 **Kieli**: Englanti **Sivumäärä**: 5+58

Tutkinto-ohjelma:
Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

Pääaine: Matematiikka
Sivuaine: Teknillinen fysiikka

Opetusyksikkö: Matematiikka **Koodi**: Mat-1

Valvoja: Prof. Olavi Nevanlinna

Ohjaajat: TkT Harri Hakula, TkT Nuutti Hyvönen

Työssä esitellään stokastinen Galerkinin elementtimenetelmä (sGFEM) yhdessä tarvittavien matemaattisten työkalujen kanssa. Menetelmä toteutetaan yksiulotteiselle stokastiselle lineaariselle elliptiselle reuna-arvo-ongelmalle, jonka voidaan ajatella kuvaavan stokastista johtavuusyhtälöä yhdessä ulottuvuudessa. Mallion-gelman johtavuuskertoimen oletetaan olevan log-normaali satunnaiskenttä tunnetulla odotusarvolla ja kovarianssifunktiolla. Mallin kontakti-impedanssien oletetaan olevan tunnettuja log-normaaleja satunnaismuuttujia. Numeeristen kokeiden perusteella menetelmän todetaan toimivan, jos parametrien varianssit eivät ole liian suuria.

Avainsanat: stokastinen Galerkinin elementtimenetelmä, polynomikaaos, Karhunen–Loève-hajotelma, log-normaalijakauma, satunnaiskenttä, sGFEM, sFEM

Acknowledgments

First, I thank my instructors D.Sc. (Tech.) Harri Hakula and D.Sc. (Tech.) Nuutti Hyvönen for introducing me to the subject of this thesis. Moreover, I thank them for their time and guidance. I thank my colleagues with whom I shared my office during the writing of this thesis and had many interesting discussions. Aalto University is thanked for providing a good working environment.

Finally, I want to thank Hanna for her support and encouragement.

This thesis was supported by the Finnish Funding Agency for Technology and Innovation TEKES (contract 40370/08).

Espoo, November 14, 2011

Matti Leinonen

Contents

Chapter 1

Introduction

The aim of numerical simulations is to model physical events and the behavior of engineered systems. However, the simulation of a physical phenomenon is often difficult even if there exists a good model for it. In many cases, the model is subject to significant data uncertainties causing meaningless results. For this reason, the uncertainties or the randomness must be realistically represented in the model. One way to take the uncertainties into account is to treat them as random processes or as random fields.

It is to be expected that physical problems that have been previously considered only in a non-probabilistic framework, that is, problems where no apparent probabilistic component is present, are going to be cast into a probabilistic framework. The justification for such a claim can be found in the rapid development of computer hardware and software as pointed out in [11]: the accuracy afforded on mathematical models exceeds the accuracy and reliability of the input data, and thus it has become important to quantify the effect of uncertainty in these problems.

Based on these facts, it is apparent that efficient methods to solve models with uncertainty in model parameters, initial values, and boundary conditions are required.

Monte Carlo sampling, or one of its variants, is probably the most common method in use to treat models with random parameters. In Monte Carlo sampling, the idea is to generate independent realizations of the random inputs in the model based on their prescribed probability distributions. For each realization, the parameters of the model are fixed, and thus the problem becomes deterministic. By solving the deterministic realizations of the problem using, for example, the finite element method, an ensemble of solutions is obtained. This ensemble can then be used to extract statistical information about the model, that is, the mean and the variance, for instance. Other methods in practical use include perturbation methods, moment equations, operator-based methods, and generalized polynomial chaos as pointed out in [20] and references therein.

The generalized polynomial chaos (gPC) is a relatively recently developed method which has become one of the most widely used techniques together with the Monte Carlo sampling. The idea of the gPC is to express the stochastic solutions as or-

thogonal polynomials of the input random parameters; different types of orthogonal polynomials can be chosen to optimize the convergence. The gPC is essentially a spectral representation of the solution in a random space, and the method exhibits fast convergence when the solution depends smoothly on the random parameters as mentioned in [20].

In this thesis, we introduce the stochastic Galerkin finite element method (sGFEM) that employs the gPC. Moreover, we implement the sGFEM for a certain stochastic linear elliptic boundary value problem and present numerical examples demonstrating the functionality of the sGFEM.

For readers interested in the gPC or sGFEM, we recommend the already twice cited book *Numerical methods for stochastic computations: a spectral method approach* by Dongbin Xiu together with the comprehensive reference list therein, and also the book *Stochastic finite elements: a spectral approach* by Roger G. Ghanem and Pol D. Spanos.

1.1 The model problem

In this work, we implement the sGFEM for the following stochastic linear elliptic boundary value problem: find a random function

$$u(\cdot,\cdot) : \Omega \times \bar{D} \to \mathbb{R} \tag{1.1}$$

such that the following equations hold P-almost surely:

$$\begin{cases} \frac{d}{d\mathrm{x}}(a(\omega,\mathrm{x})u_{\mathrm{x}}(\omega,\mathrm{x})) = 0, & \mathrm{x} \in D, \\ [u(\omega,\mathrm{x}) - z_1(\omega)a(\omega,\mathrm{x})u_{\mathrm{x}}(\omega,\mathrm{x})]_{\mathrm{x}=-L} = u_1, \\ [u(\omega,\mathrm{x}) + z_2(\omega)a(\omega,\mathrm{x})u_{\mathrm{x}}(\omega,\mathrm{x})]_{\mathrm{x}=L} = u_2, \end{cases} \tag{1.2}$$

where the domain of the problem, $D = (-L, L)$, is an open interval in $\mathbb{R}$, the set Ω is the sample space of the product probability space of the random variables a, z_1, and z_2, the variable ω belongs to the sample space Ω, the boundary values u_1 and u_2 are known real valued constants, and u_{x} denotes the derivative of u with respect to the spatial variable x.

In the model problem, the random variables z_1 and z_2 are assumed to follow known log-normal random distributions, and the random field a is assumed to be a log-normal random field. The underlying Gaussian random field for a is assumed to have a known finite mean field together with the commonly used exponential covariance function of the form $V_a(\mathrm{x}_1,\mathrm{x}_2) = \sigma^2 \exp(-|\mathrm{x}_1 - \mathrm{x}_2|/b)$, where the parameter b is the so-called *correlation length* and σ^2 is the variance of the random field. Moreover, the random variables are assumed to be independent of each other.

In the setting of electrical impedance tomography, the model problem can be considered as the conductivity equation reduced to one dimension which, at least on a conceptual level, can be considered to represent a rod whose end points are held at

potentials u_1 and u_2. The random variables z_1 and z_2 model the contact impedances at the end points of the rod, and the random field a represents the conductivity inside the rod. Notice that it is impossible to know exactly the contact impedances or the conductivity in real applications, and thus we are required to consider those properties in a probabilistic framework. We are interested in determining how the potential between the end points of the rod behaves, that is, what is the most probable value of the potential together with the corresponding variance at a specific point in the rod.

Since the conductivity and the contact impedances are by definition always non-negative, we use log-normal random variables to characterize them instead of widely used truncated normal distributions. On the negative side, log-normal random variables are more difficult to handle mathematically than plain normal distributions.

From a practical point of view, the one-dimensional case considered in this work is not as interesting as its two- or three-dimensional counterparts would be. However, the purpose of this thesis is merely to introduce the basic techniques used in the sGFEM and to prove the feasibility of the method by applying it to the one-dimensional conductivity equation. Based on this work, it is conceptually easy to implement the method for the two- or three-dimensional case and also for other types of problems with only minor changes and/or adjustments.

This thesis is divided into two main parts. In Chapters 2–4, we introduce the main necessary technical tools for understanding and implementing solvers based on the sGFEM employing the gPC, that is, the Karhunen–Loève expansion for the discretization of a random field and the generalized polynomial chaos expansion required to express the stochastic solution. In Chapters 5–7, we formulate the model problem, introduce and implement the stochastic Galerkin finite element method, and present numerical examples demonstrating its functionality. At the end of this thesis, we give some concluding remarks.

Chapter 2

Preliminaries

In this chapter, we briefly review the essential stochastic constructions and notations used throughout this thesis. The material is mostly based on references [4] and [15].

First, we construct a probability space and define a random variable which are mathematical tools used to model randomness. We proceed by introducing some important characteristics associated with random variables, such as the expectation and the variance. Moreover, we introduce Bochner spaces as they turn out to be the natural function spaces for the solutions of the model problem. Finally, we discuss random fields with the focus on Gaussian and log-normal behavior. We also review properties of (multivariate) normal and log-normal random variables as they are essential in developing the stochastic finite element method for the model problem considered in this work.

2.1 Probability space

To model an experiment or a real-world process consisting of states that occur randomly, a mathematical structure called probability space needs to be constructed. Probability spaces consist of three parts:

1. A sample space, which is the set of all possible outcomes of the experiment;
2. A set of events, where each event is a set containing zero or more outcomes of the experiment;
3. A function that associates a number called the *probability of the event* to each event.

An event can be understood as a property which can be observed to hold or to not hold after the experiment has been conducted, and the probability of an event can be understood as the limit of the frequency with which the event is realized if the experiment is repeated an infinite number of times.

Before giving the exact definition of a probability space, we give definitions for a sigma-algebra and for a probability measure. A sigma-algebra is used to describe what kind of properties the set of events for a given sample space has, and a probability measure is used to associate a probability to each event.

Definition 2.1 (Sigma-algebra)**.** Let Ω be a set with no special structure, and let 2^Ω denote the set of all subsets of Ω, including the empty set $\emptyset$ and Ω itself. Then a subset Σ of 2^Ω is a *sigma-algebra* on Ω if it satisfies the following conditions:

1. The set Σ is not empty;
2. If a set A is in Σ then its complement A^c is also in Σ;
3. If $\{A_i\}_{i=1}^\infty$ is a countable family of sets in Σ then their union

$$A = \bigcup_{n=1}^{\infty} A_n$$

is also in Σ.

Definition 2.2 (Probability measure)**.** A *probability measure* defined on a sigma-algebra Σ on a set Ω is a function $P : \Sigma \to [0, 1]$ such that:

1. The measure of the entire set is equal to one: $P(\Omega) = 1$;
2. For every countable collection $\{A_i\}_{i=1}^\infty$ of pairwise disjoint sets, that is, $i \neq j$, $A_i \cap A_j = \emptyset$, in Σ it holds that

$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i).$$

A probability space is defined using a sample space, a sigma-algebra, and a probability measure as follows:

Definition 2.3 (Probability space)**.** A *probability space* is a triplet (Ω, Σ, P), where the sample space Ω is an arbitrary non-empty set, Σ is a sigma-algebra on Ω, and P is a probability measure on Σ.

Since our model problem contains three stochastic parameters, countable products of probability spaces are needed. We review the definition of a product probability space following closely [4]. For more information about higher dimensional random variables, we refer the reader to [15].

Let $(\Omega_k, \Sigma_k, P_k)$, where $k \in \mathbb{N}$, denote a sequence of probability spaces. For a subset $J \subset \mathbb{N}$, we define the product

$$\Omega_J = \prod_{k \in J} \Omega_k$$

as the Cartesian product of the sample spaces Ω_k for which $k \in J$. When $J = \mathbb{N}$, we write

$$\Omega = \prod_{k\in\mathbb{N}} \Omega_k. \tag{2.1}$$

We define the projection operator

$$p_J : \Omega \to \Omega_J$$

as the restriction of $\omega \in \Omega$ to Ω_J and the product of sigma-algebras

$$\Sigma = \bigotimes_{k\in\mathbb{N}} \Sigma_k \tag{2.2}$$

as the smallest sigma-algebra such that any of the projections p_j is measurable. Moreover, the product measure of $\{P_k\}_{k\in\mathbb{N}}$, denoted by

$$P = \bigotimes_{k\in\mathbb{N}} P_k, \tag{2.3}$$

is defined as the unique measure on Σ such that for every finite subset $J \subset \mathbb{N}$ and arbitrary events $E_j \in \Sigma_j$ $(j \in J)$ it holds that

$$P\Big(p_J^{-1}\Big(\prod_{j\in J} E_j\Big)\Big) = \prod_{j\in J} P_j(E_j).$$

The *product probability space* (Ω, Σ, P) of the probability spaces $\{(\Omega_k, \Sigma_k, P_k)\}_{k\in\mathbb{N}}$ is then defined as

$$(\Omega, \Sigma, P) = \bigotimes_{k\in\mathbb{N}} (\Omega_k, \Sigma_k, P_k) = \Big(\prod_{k\in\mathbb{N}} \Omega_k, \bigotimes_{k\in\mathbb{N}} \Sigma_k, \bigotimes_{k\in\mathbb{N}} P_k\Big). \tag{2.4}$$

We recall how the independency of events is defined and what is meant by the P-almost sure property.

Definition 2.4 (Independency)**.** Let (Ω, Σ, P) be a probability space. Two events $A \in \Sigma$ and $B \in \Sigma$ are said to be *independent* if $P(A \cap B) = P(A)P(B)$. A possibly infinite collection of events $\{E_i\}_{i\in I} \subset \Sigma$ is an *(mutually) independent* collection if for every finite subset $J \subset I$ it holds that

$$P\Big(\bigcap_{i\in J} E_i\Big) = \prod_{i\in J} P(E_i). \tag{2.5}$$

Definition 2.5 (Negligible set)**.** Let (Ω, Σ, P) be a probability space. A *negligible set* for P is a subset A of Ω such that there exists a set B in Σ satisfying $A \subset B$ and $P(B) = 0$.

Definition 2.6 (Almost surely)**.** A property holds *almost surely* (a.s.) if it holds outside a negligible set.

By the definition of a negligible set, the almost surely property holds with respect to the underlying probability measure. If we want to emphasize the dependency on the measure P, we say P-almost surely, or P-a.s.

For the construction of probability measures on $\mathbb{R}$, we define the Borel sigma-algebra on $\mathbb{R}$ and the cumulative distribution function induced by a probability measure.

Definition 2.7 (Borel sigma-algebra on $\mathbb{R}$)**.** The *Borel sigma-algebra on* $\mathbb{R}$, denoted by $\mathcal{B}$, is the smallest sigma-algebra on $\mathbb{R}$ that contains all open sets of $\mathbb{R}$.

The Borel sigma-algebra on $\mathbb{R}$ can also be characterized with intervals of the form $(-\infty, a]$, where a is a real number; see [15] for a proof. A set is called a Borel set if it belongs to the Borel sigma-algebra on $\mathbb{R}$.

Definition 2.8 (Cumulative distribution function for a probability measure)**.** The function $F(\mathrm{x}) = P((-\infty, \mathrm{x}])$ induced by a probability measure P on $(\mathbb{R}, \mathcal{B})$ is the *cumulative distribution function for the probability measure* P.

The knowledge of the cumulative distribution function for a probability measure uniquely characterizes the probability measure; see [15] for a proof. Thus, in principle, we know the complete probability measure itself if we know the corresponding distribution function. In particular, for any Borel set A, the distribution function allows us to determine the probability $P(A)$.

The cumulative distribution function may often be written using its probability density function.

Definition 2.9 (Probability density function)**.** A function f is a *probability density function* for a cumulative distribution function F if $f \geq 0$, $\int_{-\infty}^{\infty} f(\mathrm{x})d\mathrm{x} = 1$, and it holds that

$$F(\mathrm{x}) = \int_{-\infty}^{\mathrm{x}} f(u)du. \tag{2.6}$$

Two important distributions used throughout this thesis are the normal or Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ and the log-normal distribution $\ln\mathcal{N}(\mu, \sigma^2)$ with parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. The probability density function for the normal distribution is

$$f(\mathrm{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(\mathrm{x} - \mu)^2/(2\sigma^2)], \tag{2.7}$$

and that for the log-normal distribution is

$$f(\mathrm{x}) = \begin{cases} 0, & \text{if} \quad \mathrm{x} \leq 0, \\ \frac{1}{\mathrm{x}\sqrt{2\pi\sigma^2}} \exp[-(\ln \mathrm{x} - \mu)^2/(2\sigma^2)], & \text{if} \quad \mathrm{x} > 0. \end{cases} \tag{2.8}$$

The normal distribution $\mathcal{N}(0, 1)$ is also known as the standard normal distribution. We extend the class of normal distributions $\mathcal{N}(\mu, \sigma^2)$ to include parameters $\mu \in \mathbb{R}$ and $\sigma^2 = 0$, where $\mathcal{N}(\mu, 0)$ denotes the law of the constant random variable equal to μ (Dirac's mass).

2.2 Random variable

A real-valued random variable on a probability space is defined as follows:

Definition 2.10 (Real-valued random variable)**.** Let (Ω, Σ, P) be a probability space. A function $X : \Omega \to \mathbb{R}$ is a *real-valued random variable* or simply a *random variable* (r.v.) if X is a measurable function with respect to the Borel sigma-algebra on $\mathbb{R}$, that is, $X^{-1}(B) \in \Sigma$ for all Borel sets $B \in \mathcal{B}$.

Definition 2.11 (Distribution measure)**.** If X is a r.v. then the *distribution measure* or *law* of X is the collection of the probabilities

$$P_X(B) = P(X \in B) = P(\{\omega : X(\omega) \in B\}) \tag{2.9}$$

for all $B \in \mathcal{B}$.

The distribution measure P_X of a real-valued random variable X is a probability measure on $\mathbb{R}$, and thus P_X is entirely characterized by the distribution function F_X of X:

$$F_X(\mathrm{x}) = P_X((-\infty, \mathrm{x}]) = P(X \leq \mathrm{x}).$$

If the random variable X follows the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, the random variable $Y = \exp(X)$ is said to be a log-normal random variable following the distribution given by (2.8), that is, the logarithm of a log-normal random variable is normally distributed.

The justification for the probability density function of a log-normal random variable (2.8) can be deduced from (2.7) by making the change of variables $\mathrm{y} = \exp(\mathrm{x})$. We have $\mathrm{x} = \ln \mathrm{y}$ and $d\mathrm{x} = d\mathrm{y}/\mathrm{y}$, and thus it holds for $a > 0$ that

$$\begin{aligned} P(Y \leq a) &= P(X \leq \ln a) \\ &= \int_{-\infty}^{\ln a} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(\mathrm{x} - \mu)^2/(2\sigma^2)] d\mathrm{x} \\ &= \int_0^a \frac{1}{\mathrm{y}\sqrt{2\pi\sigma^2}} \exp[-(\ln \mathrm{y} - \mu)^2/(2\sigma^2)] d\mathrm{y}. \end{aligned}$$

Usually we are interested to determine certain statistics about a random variable, such as the expected value, the variance, and higher moments. After we have acquired the relevant statistics, we make decisions based on this information. Definitions for the most often used statistics are given below. We assume that X is a (continuous) random variable with a probability density function f and that all the integrals converge.

Definition 2.12 (Expected value)**.** The *expected value* of X, denoted by μ_X or $\mathbb{E}[X]$, is defined as

$$\mu_X = \mathbb{E}[X] = \int_{\mathbb{R}} \mathrm{x} f(\mathrm{x}) d\mathrm{x}. \tag{2.10}$$

Definition 2.13 (Variance)**.** The *variance* of X, denoted by σ_X^2 or $\mathrm{Var}[X]$, is defined as

$$\sigma_X^2 = \mathrm{Var}[X] = \int_{\mathbb{R}} (\mathrm{x} - \mu_X)^2 f(\mathrm{x}) d\mathrm{x} = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \tag{2.11}$$

Definition 2.14 (Moment, centered moment)**.** For a non-negative integer n, the nth *moment* of X, denoted by $\mathbb{E}[X^n]$, is defined as

$$\mathbb{E}[X^n] = \int_{\mathbb{R}} \mathrm{x}^n f(\mathrm{x}) d\mathrm{x}, \tag{2.12}$$

and the nth *centered moment* of X, denoted by $\mathbb{E}[(X - \mu_{\mathrm{x}})^n]$, is defined as

$$\mathbb{E}[(X - \mu_{\mathrm{x}})^n] = \int_{\mathbb{R}} (\mathrm{x} - \mu_{\mathrm{x}})^n f(\mathrm{x}) d\mathrm{x}. \tag{2.13}$$

The expected value of an arbitrary real-valued function of X, $g(X)$, denoted by $\mathbb{E}[g(X)]$, is defined as the L^2-inner product of f and g:

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(\mathrm{x}) f(\mathrm{x}) d\mathrm{x}. \tag{2.14}$$

For a normally distributed random variable X with the probability density $\mathcal{N}(\mu, \sigma^2)$, the expected value is μ, the variance is σ^2, and the centered moments are obtained by evaluating the integral

$$\mathbb{E}[(X - \mu)^n] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\mathrm{x} - \mu)^n \exp[-(\mathrm{x} - \mu)^2/(2\sigma^2)] d\mathrm{x}.$$

When n is odd, the integrand is an odd function and the integral vanishes. When n is even, say $n = 2k$, we obtain by integrating by parts that

$$\mathbb{E}[(X - \mu)^{2k}] = \sigma^2(2k - 1)\mathbb{E}[(X - \mu)^{2(k-1)}],$$

and thus the centered moment for a normally distributed random variable X is

$$\mathbb{E}[(X - \mu)^n] = \begin{cases} 0, & \text{if } n \text{ is odd,} \\ 1 \cdot 3 \cdot 5 \cdots (n-1)\sigma^n, & \text{if } n \text{ is even.} \end{cases} \tag{2.15}$$

The moments of a log-normally distributed random variable X with the probability density $\ln \mathcal{N}(\mu, \sigma^2)$ can be calculated by using the standard formula

$$\mathbb{E}[X^n] = \int_0^{\infty} \mathrm{x}^{n-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(\ln \mathrm{x} - \mu)^2/(2\sigma^2)] d\mathrm{x}$$

which can be easily evaluated to give

$$\mathbb{E}[X^n] = \exp(n\mu + n^2\sigma^2/2) \tag{2.16}$$

by making the change of variables $y = \ln \mathrm{x}$, completing the square, and using the fact that the normal density function integrates to one. In particular, we obtain the expected value

$$\mu_X = \mathbb{E}[X] = \exp(\mu + \sigma^2/2) \tag{2.17}$$

and the variance

$$\sigma_X^2 = \mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = (\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2). \tag{2.18}$$

Because log-normal random variables can be described with the help of normal distributions, we are only required to consider the probability distribution function for the multivariate normal distribution in the case of finite number of independent standard Gaussian random variables in what follows. Let $\mathbf{X}$ denote a vector consisting of n independent standard Gaussian random variables $(X_1, \ldots, X_n)$. Then the probability density function for $\mathbf{X}$ is

$$f(\mathrm{x}) = \frac{1}{(2\pi)^{n/2}} \exp\Big(-\frac{1}{2}\mathrm{x}^{\mathrm{T}}\mathrm{x}\Big), \tag{2.19}$$

where x is a n-dimensional vector $(\mathrm{x}_1, \ldots, \mathrm{x}_n)^{\mathrm{T}} \in \mathbb{R}^n$.

To complete this section, we recall how independency, orthogonality, and orthonormality of random variables are defined.

Definition 2.15 (Independency of random variables)**.** Two random variables X and Y with probability densities $f_X(\mathrm{x})$ and $f_Y(\mathrm{y})$ are said to be *independent* if for every $a, b \in \mathbb{R}$, the events $\{X \leq a\}$ and $\{Y \leq b\}$ are independent events as defined in (2.5), that is, the joint density function of X and Y, $f_{X,Y}(\mathrm{x}, \mathrm{y})$, can be written in the form

$$f_{X,Y}(\mathrm{x}, \mathrm{y}) = f_X(\mathrm{x}) f_Y(\mathrm{y}). \tag{2.20}$$

Definition 2.16 (Orthogonality, orthonormality)**.** Random variables $\{X_m\}$, where m is a positive integer, are said to be *orthogonal* if for all positive integers m and n it holds that

$$\mathbb{E}[X_m X_n] = 0 \quad \text{when } m \neq n, \tag{2.21}$$

and *orthonormal* if it holds that

$$\mathbb{E}[X_m X_n] = \delta_{mn}, \tag{2.22}$$

where δ_{mn} is the Kronecker's delta:

$$\delta_{mn} = \begin{cases} 0, & \text{if } m \neq n, \\ 1, & \text{if } m = n. \end{cases}$$

Recall that Gaussian random variables are independent if and only if they are orthogonal; see [15] for a proof.

2.3 Bochner spaces

In addition to countable products of probability spaces introduced in Section 2.1, tensor products of Hilbert spaces and Bochner spaces are required to describe the stochastic parameters and the solution of the model problem. In this section, we review those concepts following closely [4].

Let $(H_1, \langle\cdot,\cdot\rangle)$ and $(H_2, \langle\cdot,\cdot\rangle)$ be two real Hilbert spaces with associated inner products. For each $\varphi_1 \in H_1$ and $\varphi_2 \in H_2$, let $\varphi_1 \otimes \varphi_2$ denote the bilinear form acting on $H_1 \times H_2$:

$$(\varphi_1 \times \varphi_2)\langle\psi_1, \psi_2\rangle = \langle\psi_1, \varphi_1\rangle_{H_1}\langle\psi_2, \varphi_2\rangle_{H_2}.$$

Let $\mathcal{E}$ be the set of all finite linear combinations of such forms, and define an inner product on $\mathcal{E}$ by extending

$$\langle\varphi_1 \otimes \varphi_2, \psi_1 \otimes \psi_2\rangle_{\mathcal{E}} = \langle\varphi_1, \psi_1\rangle_{H_1}\langle\varphi_2, \psi_2\rangle_{H_2}$$

through linearity to the whole of $\mathcal{E}$. It can be shown that $\langle\cdot,\cdot\rangle_{\mathcal{E}}$ is well defined and positive definite. The tensor product of H_1 and H_2, denoted by $H_1 \otimes H_2$, is defined as the completion of $\mathcal{E}$ with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathcal{E}}$. It is possible to show that if $\{\psi_k\}_{k\in\mathbb{N}}$ and $\{\varphi_l\}_{l\in\mathbb{N}}$ are orthonormal bases for H_1 and H_2, respectively, then $\{\psi_k \otimes \varphi_l\}_{k,l\in\mathbb{N}}$ is an orthonormal basis for $H_1 \otimes H_2$. This construction extends immediately to the tensor product of any finite number of Hilbert spaces.

Bochner spaces together with countable tensor products of probability spaces are the natural function spaces for the solutions of the model problem and also for the solutions of similar stochastic partial differential equations. Bochner spaces are a generalization of L^p-spaces to functions taking values in a Banach space which is not necessarily the space of real or complex numbers. Given a measure space (T, Σ, μ) and a Banach space $(X, \|\cdot\|_X)$, the Bochner space $L^p_\mu(T; X)$ is defined as

$$L^p_\mu(T; X) = \left\{ f : T \longrightarrow X : \int_T \|f(t)\|^p_X d\mu(t) < \infty \right\} \tag{2.23}$$

with the norm $\|f\|_{L^p_\mu(T;X)} = (\int_T \|f\|^p_X d\mu(t))^{1/p}$ when $1 \le p < \infty$ and as

$$L^\infty_\mu(T; X) = \left\{ f : T \longrightarrow X : \operatorname*{ess\,sup}_{t\in T} \|f(t)\|_X < \infty \right\} \tag{2.24}$$

with the norm $\|f\|_{L^\infty_\mu(T;X)} = \operatorname{ess\,sup}_{t\in T} \|f(t)\|_X$ when $p = \infty$.

If $p = 2$ and the underlying Banach space X is a separable Hilbert space, we have the decomposition

$$L^2_\mu(T; X) \backsimeq L^2_\mu(T) \otimes X,$$

where $\otimes$ denotes the tensor product between Hilbert spaces. In particular, the Bochner space $L^2_\mu(T; X)$ is itself again a Hilbert space.

2.4 Random fields

In many applications, uncertainties in physical quantities vary in time and/or in space. Thus, it is essential to be able to model varying uncertainties mathematically. Stochastic processes or random fields are one way to perform such modeling, where the phrase stochastic process is associated with uncertainty in time, and the phrase random field with uncertainties in space; mathematically these two terms are usually defined in the same way. In our model problem, we are mainly concerned with spatial behavior, and thus we talk about random fields in what follows.

In this section, we present definitions and properties of random fields which will be utilized in the subsequent chapters.

A random field in a given physical region of space can be understood as a set of random variables indexed by the possible positions in the region.

Definition 2.17 (Random field)**.** Let (Ω, Σ, P) be a probability space, and let $D \subset \mathbb{R}^n$ denote a bounded physical domain, where n is a positive integer. A *random field* is a jointly measurable function from $\Omega \times D$ to $\mathbb{R}$ with respect to the sigma-algebra Σ on the sample space Ω and the Borel sigma-algebra on the domains D and $\mathbb{R}$:

$$a(\cdot, \cdot) : \Omega \times D \to \mathbb{R}. \tag{2.25}$$

We assume that for a given random field, the mean field and the covariance function are known.

Definition 2.18 (Mean field, covariance function, variance)**.** For a random field a, the *mean field* is defined as

$$\mathbb{E}_a[\mathrm{x}] = \mathbb{E}[a(\cdot, \mathrm{x})] = \int_\Omega a(\omega, \mathrm{x}) dP(\omega) \tag{2.26}$$

and the *covariance function* as

$$\begin{aligned} V_a(\mathrm{x}_1, \mathrm{x}_2) &= \mathbb{E}[(a(\cdot, \mathrm{x}_1) - \mathbb{E}_a[\mathrm{x}_1])(a(\cdot, \mathrm{x}_2) - \mathbb{E}_a[\mathrm{x}_2])] \\ &= \int_\Omega (a(\omega, \mathrm{x}_1) - \mathbb{E}_a[\mathrm{x}_1])(a(\omega, \mathrm{x}_2) - \mathbb{E}_a[\mathrm{x}_2]) dP(\omega). \end{aligned} \tag{2.27}$$

The *variance* of the random field a is given by $\mathrm{Var}_a[\mathrm{x}] = V_a(\mathrm{x}, \mathrm{x})$.

For the mean field and the covariance function to exist in the L^2-sense, we must require the random field to have finite second moments, that is, $a \in L^2_P(\Omega; L^2(D))$.

Definition 2.19 (Square-integrable random field)**.** A random field a is a *square-integrable random field* if a belongs to the Bochner space $L^2_P(\Omega; L^2(D))$.

In engineering applications involving random fields, we usually have no exact knowledge about the underlying random field itself; instead, we have some samples from

the random field. In principle, we could use the samples to calculate an approximation for the mean field and for the covariance function. However, the number of available samples is usually quite small, and thus we are more likely to get a sufficient approximation only for the mean field, if any at all.

Luckily, besides the samples, we often have some a priori knowledge about the underlying covariance function. This knowledge can be used to device a suitable candidate for the covariance function. The prior information may have been obtained, for example, from some physical property relevant to the application.

Another possibility for knowing the covariance function is to have access to the two-point-correlation function of the random field.

Definition 2.20 (Two-point-correlation function)**.** For a random field a, the *two-point-correlation function* is defined as

$$C_a(\mathrm{x}_1, \mathrm{x}_2) = \int_\Omega a(\omega, \mathrm{x}_1) a(\omega, \mathrm{x}_2) dP(\omega). \tag{2.28}$$

An assumption equivalent to knowing the mean field $\mathbb{E}_a$ and the covariance function V_a is to know the mean field $\mathbb{E}_a$ and the two-point-correlation function C_a, since

$$V_a(\mathrm{x}_1, \mathrm{x}_2) = C_a(\mathrm{x}_1, \mathrm{x}_2) - \mathbb{E}_a[\mathrm{x}_1]\mathbb{E}_a[\mathrm{x}_2].$$

Both the covariance function and the two-point-correlation function can be understood to describe how two points in the random field are related.

Because the covariance function is usually assumed to have some specific form, it is important to know whether or not a given function is a valid covariance function. More precisely, we want to know whether there exists a random field with a given mean field and covariance function.

The concept of positive definiteness is fundamental when describing the class of covariance functions.

Definition 2.21 (Positive semidefinite function)**.** A function $V_a \in L^2(D \times D)$ is *positive semidefinite* on D if

$$0 \leq \sum_{k=1}^{n} \sum_{j=1}^{n} c_k V_a(\mathrm{x}_k, \mathrm{x}_j) \overline{c}_j \tag{2.29}$$

holds for any positive integer n, for any sequence of complex weights $\{c_i\}_{i=1}^n$, and for all $\mathrm{x}_k, \mathrm{x}_j \in D$. Note that, in particular, the right-hand side of (2.29) is required to be real.

The class of covariance functions on D can be shown to coincide with the the class of positive semidefinite functions on D; see [1] for a proof. We refer the reader to [1, 17] for an introduction to positive semidefinite functions and for examples of widely used covariance functions.

Gaussian random fields and log-normal random fields are important in engineering applications; they are reasonable models for many natural phenomena and they can be specified using only expectations and covariances.

Definition 2.22 (Gaussian random field)**.** A random field

$$a(\cdot,\cdot) : \Omega \times D \to \mathbb{R} \tag{2.30}$$

is said to be a *Gaussian random field* if the distribution of $a(\cdot, \mathrm{x})$ is Gaussian for every choice of $\mathrm{x} \in D$.

Since normal distributions are completely specified if the expectations and covariances are known, see [15] for a proof, it is enough to know the mean field and the covariance function of a Gaussian random field to characterize it.

Definition 2.23 (Log-normal random field)**.** A random field a is a *log-normal random field* if the logarithm of a is a Gaussian random field, that is, log-normal random fields are of the form $\exp(g(\omega, \mathrm{x}))$, where g is a Gaussian random field.

Notice that a log-normal random field can be characterized by defining the mean field and the covariance function of the underlying Gaussian random field g.

In our model problem, we are considering a one-dimensional rod $[-L, L]$, where the electrical conductivity is uncertain and can vary from point to point. The uncertain conductivity can be modeled using a random field. We make the assumption that the electrical conductivity at each point in the rod follows a log-normal distribution, that is, the random field is a log-normal random field. The underlying Gaussian random field is assumed to have a known finite mean together with the covariance function of the exponential form

$$V_a(\mathrm{x}_1, \mathrm{x}_2) = \sigma^2 \exp(-|\mathrm{x}_1 - \mathrm{x}_2|/b), \quad (\mathrm{x}_1, \mathrm{x}_2) \in [-L, L] \times [-L, L], \tag{2.31}$$

where σ^2 and b are positive parameters. The parameter b is called the correlation length, since it reflects the rate at which the correlation decays between two points.

Chapter 3

Karhunen–Loève expansion

In the previous chapter, we introduced random fields as a tool for taking uncertainties into account in a model. However, to handle a random field numerically, we need to expand it as a Fourier-type series, that is, discretize the random field in the stochastic dimension. Several such expansions exist based on a complete set of deterministic basis functions with corresponding random coefficients; the Karhunen–Loève expansion is one such expansion [16]. To be more precise, the Karhunen–Loève expansion is a representation of a random field as an infinite linear combination of the orthonormal eigenfunctions of the integral operator defined by the covariance function of the random field. Thus, the deterministic basis used in the Karhunen–Loève expansion depends on the random field in question.

The Karhunen–Loève expansion is our preferred choice for the expansion of the random field since it is an optimal linear approximation in the mean-square sense if truncated after the first, say M, terms. This result is well known; see [12].

In this chapter, we review the theory of the Karhunen–Loève expansion together with providing auxiliary results which will be used in the subsequent chapters. Before giving the definition for the Karhunen–Loève expansion, we state two theorems required for the expansion: Mercer's theorem and the Karhunen–Loève theorem which relies on the former.

Theorem 3.1 (Mercer's theorem)**.** *Let $V_a(\mathrm{x}_1, \mathrm{x}_2)$ be a continuous symmetric positive semidefinite kernel*

$$V_a : \bar{D} \times \bar{D} \to \mathbb{R}, \tag{3.1}$$

where $\bar{D} = [-L, L]$, and let T_K be the associated integral operator

$$(T_K\phi)(\mathrm{x}_1) = \int_D V_a(\mathrm{x}_1, \mathrm{x}_2)\phi(\mathrm{x}_2)d\mathrm{x}_2. \tag{3.2}$$

Then there exists an orthonormal basis $\{\phi_m\}_{m\geq 1}$ of $L^2(D)$ consisting of eigenfunctions of the operator T_K such that the corresponding sequence of eigenvalues $\{\lambda_m\}_{m\geq 1}$ is nonnegative. The eigenfunctions corresponding to the nonzero eigenvalues are

continuous on D, and V_a has the representation

$$V_a(\mathrm{x}_1, \mathrm{x}_2) = \sum_{m \geq 1} \lambda_m \phi_m(\mathrm{x}_1)\phi_m(\mathrm{x}_2), \tag{3.3}$$

where the convergence of the series is absolute and uniform in both variables.

For the proof of Mercer's theorem, see [22].

Let V_a be a continuous covariance function; a covariance function is symmetric and positive semidefinite by the definition. Hence, by Mercer's theorem, V_a has the decomposition

$$V_a(\mathrm{x}_1, \mathrm{x}_2) = \sum_{m \geq 1} \lambda_m \phi_m(\mathrm{x}_1)\phi_m(\mathrm{x}_2), \tag{3.4}$$

where $\{\lambda_m\}_{m \geq 1}$ are the positive eigenvalues and $\{\phi_m\}_{m \geq 1}$ are the corresponding eigenfunctions of the integral operator associated to the covariance function V_a. That is, $\{\lambda_m\}_{m \geq 1}$ and $\{\phi_m\}_{m \geq 1}$ are solutions to the integral equation

$$\int_D V_a(\mathrm{x}_1, \mathrm{x}_2)\phi(\mathrm{x}_2)d\mathrm{x}_2 = \lambda\phi(\mathrm{x}_1) \quad \text{for all } \mathrm{x}_1 \in D, \tag{3.5}$$

fulfilling the $L^2(D)$-orthonormality condition

$$\int_D \phi_m(\mathrm{x})\phi_n(\mathrm{x})d\mathrm{x} = \delta_{mn}, \tag{3.6}$$

where m and n are positive integers, and δ_{mn} is the Kronecker's delta.

The sequence of the positive eigenvalues $\{\lambda_m\}_{m \geq 1}$ used in the decomposition (3.4) is assumed to be enumerated with decreasing magnitude and is either finite or tends to zero as m approaches infinity.

Let a be a square-integrable random field with a finite mean field $\mathbb{E}_a$ and a continuous covariance function V_a. Then a can be written in the form

$$a(\omega, \mathrm{x}) = \mathbb{E}_a[\mathrm{x}] + \alpha(\omega, \mathrm{x}), \tag{3.7}$$

where α is a random field with a zero mean field and the covariance function V_a. The following Karhunen–Loève theorem gives an expansion for the zero mean random field α. The proof of the theorem can be found in [2].

Theorem 3.2 (The Karhunen–Loève theorem)**.** *Let α be a square-integrable random field with a zero mean field and a continuous covariance function V_a. Let $\{\phi_m\}_{m \geq 1}$ be an orthonormal basis for the space spanned by the eigenfunctions corresponding to the nonzero eigenvalues associated with the integral operator defined by the covariance function V_a with ϕ_m being an eigenfunction corresponding to the eigenvalue $\lambda_m > 0$. Then the random field α admits an expansion of the form*

$$\alpha(\omega, \mathrm{x}) = \sum_{m \geq 1} \sqrt{\lambda_m} Y_m(\omega)\phi_m(\mathrm{x}), \tag{3.8}$$

where the family of random variables $\{Y_m\}_{m\geq 1}$ *is determined by*

$$Y_m(\omega) = \frac{1}{\sqrt{\lambda_m}} \int_D \alpha(\omega, \mathrm{x}) \phi_m(\mathrm{x}) d\mathrm{x}. \tag{3.9}$$

The expansion converges to α *in* $L^2(\Omega)$ *uniformly over* D*, that is,*

$$\mathbb{E}\Big[(\alpha(\cdot, \mathrm{x}) - \sum_{m=1}^{n} \sqrt{\lambda_m} Y_m(\cdot)\phi_m(\mathrm{x}))^2\Big] \to 0 \text{ as } n \to \infty, \tag{3.10}$$

uniformly for x *in* D*. The family of random variables* $\{Y_m\}_{m\geq 1}$ *satisfies*

$$\mathbb{E}[Y_m] = 0 \text{ and } \mathbb{E}[Y_m Y_n] = \delta_{mn} \tag{3.11}$$

for all positive integers m *and* n*, that is* $\{Y_m\}_{m\geq 1}$ *have zero mean and they are orthonormal.*

The Karhunen–Loève theorem together with Mercer's theorem gives the Karhunen–Loève expansion for a square-integrable random field:

Definition 3.3 (The Karhunen–Loève expansion)**.** The *Karhunen–Loève expansion* for a square-integrable random field a with a finite mean field $\mathbb{E}_a$ and a continuous covariance function V_a is given by

$$a(\omega, \mathrm{x}) = \mathbb{E}_a[\mathrm{x}] + \sum_{m\geq 1} \sqrt{\lambda_m} Y_m(\omega)\phi_m(\mathrm{x}), \tag{3.12}$$

where the family of random variables $\{Y_m\}_{m\geq 1}$ is determined by

$$Y_m(\omega) = \frac{1}{\sqrt{\lambda_m}} \int_D (a(\omega, \mathrm{x}) - \mathbb{E}_a[\mathrm{x}])\phi_m(\mathrm{x}) d\mathrm{x}. \tag{3.13}$$

Because the Karhunen–Loève expansion is in general infinite, it is of little use in numerical calculations as such. Thus, to numerically handle the expansion, it has to be truncated after the first, say M, terms.

Definition 3.4 (The truncated Karhunen–Loève expansion)**.** Let a be a random field with the Karhunen–Loève expansion as given in Definition 3.3. Then the *truncated Karhunen–Loève expansion* for the random field a is given by

$$a_M(\omega, \mathrm{x}) = \mathbb{E}_a[\mathrm{x}] + \sum_{m=1}^{M} \sqrt{\lambda_m} Y_m(\omega)\phi_m(\mathrm{x}). \tag{3.14}$$

A natural question to ask is how many terms to include in the series. Before addressing this question, we state results which characterize the uniqueness and the error minimizing property of the expansion. The following two theorems are proved in [12].

Theorem 3.5 (Uniqueness of the expansion)**.** *The family of random variables appearing in an expansion of the type* (3.12) *are orthonormal if and only if the orthonormal functions* $\{\phi_m\}_{m\geq 1}$ *and the constants* $\{\lambda_m\}_{m\geq 1}$ *are respectively the eigenfunctions and the eigenvalues of the associated integral operator for the covariance kernel as given by* (3.5)*.*

Theorem 3.6 (Error minimizing property)**.** *The truncated Karhunen–Loève expansion is an optimal approximation for the random field* a *in the sense that the mean-square error* $\|a - a_M\|_{L^2_P(\Omega;L^2(D))}$ *resulting from a truncated linear representation of the random field* a *is minimized.*

The error minimizing property implies that for any other linear combination $\tilde{a}_M$ of M functions, the resulting error $\|a - \tilde{a}_M\|_{L^2_P(\Omega;L^2(D))}$ is not smaller than for the Karhunen–Loève expansion. However, nonlinear approximations may yield better approximations as noted in [16]. We can calculate the $L^2_P(\Omega; L^2(D))$-error caused by truncating the expansion as in (3.14) as follows:

$$\begin{aligned}\|a - a_M\|_{L^2_P(\Omega;L^2(D))} &= \mathbb{E}\left[\int_D (a(\omega,\mathrm{x}) - a_M(\omega,\mathrm{x}))^2 d\mathrm{x}\right] \\ &= \mathbb{E}\left[\int_D \left(\sum_{m>M} \sqrt{\lambda_m} Y_m(\omega)\phi_m(\mathrm{x})\right)^2 d\mathrm{x}\right] \\ &= \sum_{m>M} \lambda_M, \end{aligned} \tag{3.15}$$

where in the last step the orthonormality conditions (3.6) and (3.11) have been used. From (3.15) we notice that estimates for the eigenvalue decay in the Karhunen–Loève expansion are crucial to obtain good a priori control over the error imposed by truncating the expansion after M terms. It can be shown that for the pointwise error $\|a - a_M\|_{L^\infty(\Omega\times D)}$ information about the eigenfunctions is also required. Thus, the question about when to truncate the series is related to the properties of the eigenvalues and the eigenfunctions. For more information, see [4] and references therein.

From Definition 3.3, we see that a prerequisite for writing the Karhunen–Loève expansion is the knowledge of the mean field and the covariance function of the random field in question. Once we know the eigenvalues and the eigenfunctions associated with the integral operator defined by the underlying covariance function, probability density functions for the random variables $\{Y_m\}_{m\geq 1}$ in Definition 3.3 may be estimated from known sample input fields via (3.13). Thus, the calculation of the Karhunen–Loève expansion for a random field can be summarized as follows:

1. Sample input fields from the random field;
2. Estimate the mean field and the covariance function for the random field. For example, one may sample these functions from the available input fields and/or assume some specific form for them;

3. Solve the equation (3.5) for the first M largest eigenvalues and eigenfunctions;

4. Use the sample input fields, the calculated M largest eigenvalues, and the corresponding eigenfunctions to approximate the probability density functions for the random variables $\{Y_m\}_{m\geq 1}$ as defined by (3.13).

The last step is unnecessary if the random field is assumed to be Gaussian: from the properties of Gaussian random fields, we know that in this case the random variables $\{Y_m\}_{m\geq 1}$ are also Gaussian. Furthermore, since the random variables are orthonormal by the equation (3.11), the random variables must follow the standard normal distribution and be mutually independent. We emphasize that the random variables $\{Y_m\}_{m\geq 1}$ are not necessarily mutually independent if the random field is not Gaussian.

From the summary above, we see that to apply the Karhunen–Loève expansion as a general simulation tool, the eigenvalues and the eigenfunctions corresponding to the covariance function of the random field must be known. Thus, it is essential to be able to compute them efficiently and accurately. In practice, we are required to solve the equation (3.5) for a given covariance function. This can be done analytically for some covariance functions but if no analytic solution is possible, or if there is only numerical information about the underlying covariance function, one needs to resort to other methods to calculate the eigenvalues and the eigenfunctions. One possible technique is the finite element method; see [18] for more information.

Since the log-normal random field in our model problem is uniquely determined by a Gaussian random field, the sampling of the probability density functions for the random variables $\{Y_m\}_{m\geq 1}$ is not discussed here. See [16] and references therein for information about calculating statistics for random fields through sampling.

A log-normal random field $a(\omega, \mathrm{x}) = \exp(g(\omega, \mathrm{x}))$ is used in our model problem. Here g is a Gaussian random field with a mean field $\mathbb{E}_a^{\ln}[\mathrm{x}] = \mu(\mathrm{x})$ and a covariance function $V_a^{\ln}(\mathrm{x}_1, \mathrm{x}_2) = \sigma^2 \exp(-|\mathrm{x}_1 - \mathrm{x}_2|/b)$. In particular, g follows pointwise the normal distribution $\mathcal{N}(\mu(\mathrm{x}), \sigma^2)$. We could in principle write the Karhunen–Loève expansion directly for a, but due to the log-normality of a and the Gaussian nature of g, it is more convenient to work with the logarithm of a instead of a itself. Thus, we utilize the following expansion for the logarithm of the random field a:

$$g(\omega, \mathrm{x}) = \ln a(\omega, \mathrm{x}) = \mathbb{E}_a^{\ln}[\mathrm{x}] + \sum_{m\geq 1} \sqrt{\lambda_m} Y_m(\omega) \phi_m(\mathrm{x}), \tag{3.16}$$

where the random variables $\{Y_m\}_{m\geq 1}$ follow the standard normal distribution and are mutually independent as implied by the Gaussian assumption on the random field g. Moreover, $\{\lambda_m\}_{m\geq 1}$ and $\{\phi_m\}_{m\geq 1}$ are the eigenvalues and eigenfunctions corresponding to $V_a^{\ln}$.

Because we did not expand the random field a, but rather the logarithm of a, we are still required to define the truncated exponential Karhunen–Loève expansion which is defined similarly to Definition 3.4.

Definition 3.7 (The truncated exponential Karhunen–Loève expansion)**.** Let $\ln a$ be a random field with a Karhunen–Loève expansion as in (3.16). Then the *truncated exponential Karhunen–Loève expansion* for the random field a is defined as

$$a_M(\omega, \mathrm{x}) = \exp\left(\mathbb{E}_a^{\ln}[\mathrm{x}] + \sum_{m=1}^{M} \sqrt{\lambda_m} Y_m(\omega)\phi_m(\mathrm{x})\right), \tag{3.17}$$

where $\mathbb{E}_a^{\ln}$ is the mean field of $\ln a$.

The truncation errors associated with the truncated exponential Karhunen–Loève expansion are more involved than in the case of truncated Karhunen–Loève expansion; see [4] for more information.

The domain D of the model problem is assumed to be the one-dimensional interval $(-L, L)$, and hence the associated eigenvalue problem (3.5) adopts the following form:

$$\int_{-L}^{L} \sigma^2 \exp(-|\mathrm{x}_1 - \mathrm{x}_2|/b)\phi_m(\mathrm{x}_2)d\mathrm{x}_2 = \lambda_m \phi_m(\mathrm{x}_1) \quad \text{for all } \mathrm{x}_1 \in D. \tag{3.18}$$

This eigenvalue problem can be solved analytically [12]. The eigenvalues are

$$\lambda_m = 2\sigma^2 b \Big/ (\alpha_m^2 b^2 + 1), \tag{3.19}$$

and the corresponding eigenfunctions are

$$\phi_m(\mathrm{x}) = \begin{cases} \cos(\alpha_m \mathrm{x}) \Big/ \sqrt{L + \frac{\sin(2\alpha_m L)}{2\alpha_m}}, & \text{if } m \text{ is odd,} \\ \sin(\alpha_m \mathrm{x}) \Big/ \sqrt{L - \frac{\sin(2\alpha_m L)}{2\alpha_m}}, & \text{if } m \text{ is even,} \end{cases} \tag{3.20}$$

where α_{2k-1} corresponds to the kth solution of

$$1 - \alpha b \tan(\alpha L) = 0 \tag{3.21}$$

and α_{2k} corresponds to the kth solution of

$$\alpha b + \tan(\alpha L) = 0. \tag{3.22}$$

Thus, the calculation of the stochastic expansion for the random field used in the model problem is reduced to solving the equations (3.21) and (3.22).

Chapter 4

Polynomial chaos

We found out in the previous chapter that the knowledge of the mean field and the covariance function of a square-integrable random field is a prerequisite for the Karhunen–Loève expansion. In practice, such statistics can be assumed to be known for the stochastic input parameters of the model but clearly not for the response statistics. Thus, the Karhunen–Loève expansion cannot be used to describe the solution's random behavior. An alternative expansion is thus required.

One method to represent the solution is to use a series of suitably chosen orthogonal polynomials of the random variables appearing in the input parameters. For example, if the input parameters are Gaussian random variables, the orthogonal polynomials are the Hermite polynomials, and the corresponding expansion is known as the polynomial chaos expansion [21].

In this chapter, we introduce the polynomial chaos expansion for a random variable and show how to calculate approximations for the statistical properties of the random variable using such expansion. We start with general information about polynomial chaos before giving the exact representation of the corresponding expansion. We restrict our attention to results which will be required in the following chapters. For more information about the polynomial chaos expansion, we refer the reader to [12, 20]. The material presented here is based on the references [4, 12, 20, 21].

Wiener originally defined the polynomial chaos in terms of the Hermite polynomials of random variables. If some other orthogonal polynomials than the Hermite polynomials are used to construct the chaos, the term *generalized polynomial chaos* (gPC) is used instead of the polynomial chaos; see [21] for more information. We want to point out that based on the references we have used, there seems to be no standard convention for the terminology and notations used when discussing polynomial chaos expansions.

As the polynomial chaos expansion is in essence a projection onto the space spanned by the orthogonal polynomials of the input random variables, the rate of convergence of the expansion depends on the smoothness of the solution as a function of the input random variables. It is shown in [21] that when the used polynomials are orthogonal with respect to the probability density function of the random variables appearing in

Table 4.1: Correspondence between the distribution of the input random variable and the optimal generalized polynomial chaos basis polynomials. Parameters a, b are real numbers and N is a positive integer.

	Distribution	gPC basis polynomials	Support
Continuous	Gaussian	Hermite	$(-\infty, \infty)$
	Gamma	Laguerre	$[0, \infty)$
	Beta	Jacobi	$[a, b]$
	Uniform	Legendre	$[a, b]$
Discrete	Poisson	Charlier	$\{0, 1, 2, \ldots\}$
	Binomial	Krawtchouk	$\{0, 1, 2, \ldots N\}$
	Negative binomial	Meixner	$\{0, 1, 2, \ldots\}$
	Hypergeometric	Hahn	$\{0, 1, 2, \ldots N\}$

the input parameters, the convergence of the polynomial chaos expansion is optimal in the L^2-sense. Moreover, [21] demonstrates numerically that if for a certain process the preferred gPC basis polynomials are not used, the solution still converges but the rate of convergence can be substantially lower.

Table 4.1, cited from [20], lists orthogonal gPC basis polynomials for some standard probability distributions together with the supports of the corresponding densities. The table can be used to select the preferred basis polynomials if the distribution of the input parameters is known. If the distribution of the random input does not belong to the basic types of distributions listed in Table 4.1 or if it is not explicitly known, see [21] for the gPC representation of arbitrary random inputs.

4.1 Hermite polynomial chaos

In our model problem, the uncertainties are represented by log-normal random variables. Therefore, we may consider the (generalized) polynomial chaos expansion in the setting of orthonormal Gaussian random variables and use the original terminology introduced by Wiener, that is, we use the term polynomial chaos (PC) instead of generalized polynomial chaos (gPC).

Definition 4.1 (Polynomial chaos, homogeneous chaos)**.** Let $\{Y_m\}_{m=1}^{\infty}$ be a set of orthonormal Gaussian random variables with zero mean and unit variance. Let $\hat{H}_p$, where p is a non-negative integer, denote the space of all polynomials in $\{Y_m\}_{m=1}^{\infty}$ of degree not exceeding p. Then the *polynomial chaos of order p*, denoted by H_p, is the set of all polynomials in $\hat{H}_p$ which are orthogonal to $\hat{H}_{p-1}$, and the *homogeneous chaos of order p*, denoted by $\bar{H}_p$, is the space spanned by the polynomial chaos H_p.

In the following, we denote by $H_p(Y_{i_1}, \ldots, Y_{i_p})$ a member of the polynomial chaos of order p in the variables $(Y_{i_1}, \ldots, Y_{i_p})$. Due to the orthogonality of the Hermite polynomials with respect to the standard Gaussian density, the general expression for

the polynomial chaos of order p is given by the p-dimensional multivariate Hermite polynomial

$$H_p(Y_{i_1},\ldots,Y_{i_p}) = \exp\Big(\frac{1}{2}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\Big)(-1)^p\frac{\partial^p}{\partial Y_{i_1}\cdots Y_{i_p}}\exp\Big(-\frac{1}{2}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\Big), \tag{4.1}$$

where $\mathbf{Y}$ denotes the vector consisting of p Gaussian random variables $(Y_{i_1},\ldots,Y_{i_p})$ with zero mean and unit variance. The vector $\mathbf{Y}$ follows the multivariate Gaussian distribution as given by (2.19). The expected value of a p-dimensional multivariate Hermite polynomial is zero if the arguments are orthonormal Gaussian random variables, that is,

$$\mathbb{E}[H_p(Y_{i_1},\ldots,Y_{i_p})] = 0.$$

See [12] for the construction of the general expression (4.1).

It is proved in [9] that the system consisting of all polynomial chaoses is orthogonal and complete in $L^2_P(\Omega)$ and that every square integrable random variable $f \in L^2_P(\Omega)$ admits the following representation:

$$\begin{aligned} f(\omega) = c_0H_0 &+ \sum_{i_1=1}^{\infty} c_{i_1}H_1(Y_{i_1}(\omega)) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} c_{i_1i_2}H_2(Y_{i_1}(\omega),Y_{i_2}(\omega)) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} c_{i_1i_2i_3}H_3(Y_{i_1}(\omega),Y_{i_2}(\omega)Y_{i_3}(\omega)) \\ &+ \ldots, \end{aligned} \tag{4.2}$$

where $c_{i_1\cdots i_p}$ is a real-valued coefficient and the series converges in $L^2_P(\Omega)$. The representation given by (4.2) is known as the *polynomial chaos expansion of* f. For the succeeding development, we rewrite the polynomial chaos expansion of f in the form

$$f(\omega) = \sum_{j=0}^{\infty}\hat{c}_j\Psi_j(\mathbf{Y}(\omega)), \tag{4.3}$$

where a one-to-one correspondence exists between $\Psi_j(\mathbf{Y})$ and $H_p(Y_{i_1},\ldots,Y_{i_p})$ and also between the coefficients $\hat{c}_j$ and $c_{i_1\cdots i_p}$. The expansion (4.3) is assumed to be carried out in the order indicated by the expansion (4.2), that is, the contribution of the lower order polynomials is accounted for first.

The correspondence can be easily characterized by using multi-indices and graded lexicographic ordering (single-indices). Table 4.2, cited with small modifications from [20], illustrates the correspondence between multi-indices and single-indices in the case of four Gaussian random variables, that is, $\{Y_m\}_{m=1}^4$. See [20] for the construction of multi-indices and graded lexicographic ordering.

Table 4.2: An example of the graded lexicographic ordering of the multi-indices $\mathbf{i}$ in the case of four dimensions.

$\lvert\mathbf{i}\rvert$	Multi-index $\mathbf{i}$	Single-index j	$\Psi_j(Y_1, Y_2, Y_3, Y_4)$
0	(0 0 0 0)	0	1
1	(1 0 0 0)	1	Y_1
	(0 1 0 0)	2	Y_2
	(0 0 1 0)	3	Y_3
	(0 0 0 1)	4	Y_4
2	(2 0 0 0)	5	$Y_1^2 - 1$
	(1 1 0 0)	6	$Y_1 Y_2$
	(1 0 1 0)	7	$Y_1 Y_3$
	(1 0 0 1)	8	$Y_1 Y_4$
	(0 2 0 0)	9	$Y_2^2 - 1$
	(0 1 1 0)	10	$Y_2 Y_3$
	(0 1 0 1)	11	$Y_2 Y_4$
	(0 0 2 0)	12	$Y_3^2 - 1$
	(0 0 1 1)	13	$Y_3 Y_4$
	(0 0 0 2)	14	$Y_4^2 - 1$
3	(3 0 0 0)	15	$Y_1^3 - 3Y_1$
	(2 1 0 0)	16	$Y_1^2 Y_2 - Y_2$
	(2 0 1 0)	17	$Y_1^2 Y_3 - Y_3$
	$\vdots$	$\vdots$	$\vdots$

Due to the orthogonality of the homogeneous chaos of order p and the properties of the multivariate Hermite polynomials, the random variables $\{\Psi_j\}_{j=0}^{\infty}$ satisfy the following conditions:

$$\mathbb{E}[\Psi_j] = 0, \quad \mathbb{E}[\Psi_i \Psi_j] = 0, \quad i \neq j, \quad i, j = 1, 2, \ldots, \quad \Psi_0 \equiv 1. \tag{4.4}$$

As in the case of the Karhunen–Loève expansion, the infinite dimensional polynomial chaos expansion (4.3) has to be replaced with a finite expansion in practical computations. By restricting the number of Gaussian random variables used in the expansion to n and the maximum degree of the polynomial chaos to p, we define the n-dimensional polynomial chaos expansion of order p as follows:

Definition 4.2 (The n-dimensional polynomial chaos expansion of order p)**.** The *n-dimensional polynomial chaos expansion of order p* of $f \in L_P^2(\Omega)$ is the restriction of (4.3) containing only n of the uncorrelated random variables $\{Y_m\}_{m=1}^{\infty}$, say $\{Y_m\}_{m=1}^{n}$, and where the maximum degree of the polynomials of $\{Y_m\}_{m=1}^{n}$ is p.

In the limit, when both n and p tend to infinity, we recover the polynomial chaos expansion as defined in (4.3). The convergence of the n-dimensional polynomial chaos expansion depends on n as well as on the choice of the used subset $\{Y_m\}_{m=1}^{n}$. For our model problem, the subset will be composed of the first $n-2$ random variables in the Karhunen–Loève expansion for the random field describing the conductivity together with two random variables describing the contact impedances.

Example 4.3. The two-dimensional polynomial chaos of infinite order is written in a fully expanded form as

$$\begin{aligned} f(\omega) = {} & c_0 H_0 + c_1 H_1(Y_1) + c_2 H_2(Y_2) \\ & + c_{11} H_2(Y_1, Y_1) + c_{21} H_2(Y_2, Y_1) + c_{22} H_2(Y_2, Y_2) \\ & + c_{111} H_3(Y_1, Y_1, Y_1) + c_{211} H_3(Y_2, Y_1, Y_1) + c_{221} H_3(Y_2, Y_2, Y_1) \\ & + c_{222} H_3(Y_2, Y_2, Y_2) + \ldots \ . \end{aligned}$$

The total number of expansion terms in an n-dimensional polynomial chaos expansion of order p, denoted by $P+1$, is determined by the number of dimensions used, n, and by the highest polynomial order, p, of the polynomials $\{Y_m\}_{m=1}^{n}$, and can be given with the help of basic combinatorics as the sum

$$P + 1 = 1 + \sum_{s=1}^{p} \frac{(n+s-1)!}{s!(n-1)!},$$

which can be further simplified into the form

$$P + 1 = \frac{(n+p)!}{n!p!}. \tag{4.5}$$

Note that the total number of terms in the expansion increases faster than exponentially when n and/or p is increased. Thus, the size of the expansion can grow too large for any computer to handle even for reasonably small values of n and p.

Remark 4.4. It is known that the span of $\{X^k\}_{k\in\mathbb{N}_0}$ is dense in $L^2_P(\Omega)$ if and only if the (Hamburger) moment problem is determinate, that is, the distribution function F_X of X is uniquely defined by the sequence of its moments

$$\mathbb{E}[X^n] = \int_{\mathbb{R}} \mathrm{x}^n F_X(\mathrm{x}), \quad n = 0, 1, 2, \dots \ .$$

Thus the (generalized) polynomial chaos expansion in one basic random variable X converges if and only if the distribution function of X is determinate. For example, all distributions listed in Table 4.1 are determinate and, more specifically, the normal distribution is determinate. However, the log-normal distribution can be easily shown to be indeterminate by using Krein's condition [19]. For this reason, we consider log-normal random variables through normally distributed random variables.

4.2 Response statistics

A polynomial chaos expansion for a given function is an analytic representation of the function in terms of random variables. Therefore, the calculation of the response statistics is straightforward.

In the model problem, we are interested in calculating the mean field and the variance of the solution random field u. Let

$$\hat{u}(\omega, \mathrm{x}) = \sum_{j=0}^{P} \hat{c}_j(\mathrm{x})\Psi_j(\mathbf{Y}(\omega))$$

be the n-dimensional polynomial chaos expansion of order p for the random field u. Then the mean field of u can be approximated as

$$\mathbb{E}[u(\cdot, \mathrm{x})] \approx \mathbb{E}[\hat{u}(\cdot, \mathrm{x})] = \mathbb{E}\Big[\sum_{j=0}^{P} \hat{c}_j(\mathrm{x})\Psi_j(\mathbf{Y}(\cdot))\Big] = \hat{c}_0(\mathrm{x}), \tag{4.6}$$

and the corresponding covariance function as

$$\begin{aligned} V_u(\mathrm{x}_1, \mathrm{x}_2) &\approx V_{\hat{u}}(\mathrm{x}_1, \mathrm{x}_2) \\ &= \mathbb{E}[(\hat{u}(\cdot, \mathrm{x}_1) - \hat{c}_0(\mathrm{x}_1))(\hat{u}(\cdot, \mathrm{x}_2) - \hat{c}_0(\mathrm{x}_2))] \\ &= \sum_{j=1}^{P} \mathbb{E}[\Psi_j^2(\mathbf{Y}(\cdot))]\hat{c}_j(\mathrm{x}_1)\hat{c}_j(\mathrm{x}_2), \end{aligned} \tag{4.7}$$

where the orthogonality of the polynomials $\{\Psi_j\}_{j=0}^P$ was used. The approximation for the variance is thus

$$\mathrm{Var}_u[\mathrm{x}] \approx \sum_{j=1}^{P} \mathbb{E}[\Psi_j^2(\mathbf{Y}(\cdot))]\hat{c}_j^2(\mathrm{x}). \tag{4.8}$$

Other statistical properties of the random field can be approximated in an analogous way by applying the corresponding definitions to the polynomial chaos expansion of the random field.

Chapter 5

Problem setting

In this chapter, we formulate the stochastic boundary value problem considered in the rest of this work. We continue by giving the variational (weak) formulation for the problem and proving that it is well-posed in the sense of unique solvability.

Let the domain of the problem, $D = (-L, L)$, be an open interval in $\mathbb{R}$. Suppose that $(\Omega_i, \Sigma_i, P_i)$, $i = 1, 2, 3$, are given probability spaces, and let $(\Omega, \Sigma, P) = \otimes_{i=1}^{3}(\Omega_i, \Sigma_i, P_i)$ denote their product probability space as defined in (2.4). Moreover, let $z_i : \Omega_i \to \mathbb{R}$, $i = 1, 2$, be square-integrable random variables with finite expected values and known probability distributions. Finally, assume $a(\cdot, \cdot) : \Omega_3 \times \bar{D} \to \mathbb{R}$ to be a square-integrable random field with a finite mean field whose probability distribution is also known.

We consider the following stochastic linear elliptic boundary value problem: find a random function

$$u(\cdot, \cdot) : \Omega \times \bar{D} \to \mathbb{R} \tag{5.1}$$

such that the following equations hold P-almost surely:

$$\begin{cases} \frac{d}{d\mathrm{x}}(a(\omega, \mathrm{x})u_{\mathrm{x}}(\omega, \mathrm{x})) = 0, & \mathrm{x} \in D, \\ [u(\omega, \mathrm{x}) - z_1(\omega)a(\omega, \mathrm{x})u_{\mathrm{x}}(\omega, \mathrm{x})]_{\mathrm{x}=-L} = u_1, \\ [u(\omega, \mathrm{x}) + z_2(\omega)a(\omega, \mathrm{x})u_{\mathrm{x}}(\omega, \mathrm{x})]_{\mathrm{x}=L} = u_2, \end{cases} \tag{5.2}$$

where $\omega \in \Omega$, u_1 and u_2 are known real valued constants, and u_{x} denotes the derivative of u with respect to the variable x.

The random coefficients a, z_1, and z_2 are assumed to be independent and uniformly bounded from above and below, that is, for some positive real numbers $a_{\min}$, $a_{\max}$, $z_{\min}$, and $z_{\max}$ it holds that

$$P(\omega \in \Omega : a_{\min} < a(\omega, \mathrm{x}) < a_{\max} \text{ for all } \mathrm{x} \in \bar{D}) = 1 \tag{5.3}$$

and

$$P(\omega \in \Omega : z_{\min} < z_i(\omega) < z_{\max}) = 1, \quad i = 1, 2. \tag{5.4}$$

Remark 5.1. The conditions (5.3) and (5.4) do not hold for log-normal random variables as they can take values arbitrarily close to zero or infinity with a positive probability. Hence, if a, z_1, and z_2 were log-normal random variables or fields – as they are in our numerical tests – their behavior should in principle be restricted, or the assumptions (5.3) and (5.4) should be relaxed substantially. However, in our case this is not required: we are performing the calculations numerically with a computer, and thus the underlying hardware based limits can be assumed as the limits for the random variables. Moreover, we want to note that our decision to work with (unrestricted) log-normal random variables did not cause any numerical instabilities in the numerical studies of Section 6. See [3, 4, 14] for information about how the assumptions (5.3) and (5.4) can be relaxed and how the log-normal random variables can be suitably restricted.

We start by giving the variational formulation for the problem. Recall that the Hilbert space

$$\mathcal{H}_p = L^2_P(\Omega; H^1(D)) = \Big\{ v : \Omega \to H^1(D) : \int_\Omega \|v(\omega)\|^2_{H^1(D)} dP(\omega) < \infty \Big\}$$

is equipped with the norm

$$\|v\|^2_{L^2_P(\Omega;H^1(D))} = \int_\Omega \|v(\omega)\|^2_{H^1(D)} dP(\omega) = \mathbb{E}[\|v\|^2_{H^1(D)}],$$

where

$$\|v\|^2_{H^1(D)} = \|v\|^2_{L^2(D)} + \|v_x\|^2_{L^2(D)} = \int_D v(\mathrm{x})^2 d\mathrm{x} + \int_D v_\mathrm{x}(\mathrm{x})^2 d\mathrm{x}$$

and the derivative of v with respect to the variable x, that is, v_x, is understood in the weak sense. The Sobolev space $H^1(D)$ contains all $L^2(D)$ functions whose first weak derivative is also in $L^2(D)$. For more information about Sobolev spaces, see [10], for example.

By multiplying the first equation of (5.2) with a test function $v \in \mathcal{H}_P$, integrating by parts, using the given boundary conditions, and taking the expected value, we obtain the variational formulation for the problem: find $u \in \mathcal{H}_P$ such that for all $v \in \mathcal{H}_P$ it holds that

$$B(u, v) = F(v), \tag{5.5}$$

where

$$\begin{aligned} B(u,v) = \mathbb{E}\Big[\int_D a(\omega,\mathrm{x}) u_\mathrm{x}(\omega,\mathrm{x}) v_\mathrm{x}(\omega,\mathrm{x}) d\mathrm{x} \\ + \frac{1}{z_2(\omega)} u(\omega, L) v(\omega, L) + \frac{1}{z_1(\omega)} u(\omega, -L) v(\omega, -L) \Big] \end{aligned} \tag{5.6}$$

and

$$F(v) = \mathbb{E}\Big[\frac{1}{z_2(\omega)} u_2 v(\omega, L) + \frac{1}{z_1(\omega)} u_1 v(\omega, -L) \Big]. \tag{5.7}$$

We continue by proving that the variational formulation of the problem is well-posed in the sense of unique solvability. To this end, we must show the continuity of the linear functional F and the bilinear form B, i.e., that for some constant $C > 0$ it holds that

$$|F(v)| \leq C\|v\|_{\mathcal{H}_P} \quad \text{for all } v \in \mathcal{H}_P$$

and

$$|B(u,v)| \leq C\|u\|_{\mathcal{H}_P}\|v\|_{\mathcal{H}_P} \quad \text{for all } u, v \in \mathcal{H}_P,$$

and the coercivity of B, i.e., that for some constant $C > 0$ it holds that

$$B(v,v) \geq C\|v\|^2_{\mathcal{H}_P} \quad \text{for all } v \in \mathcal{H}_P.$$

To reach this goal, we need the algebraic inequality

$$|x| + |y| \leq \sqrt{2(x^2+y^2)}, \quad x, y \in \mathbb{R}, \tag{5.8}$$

and the so-called trace theorem.

Theorem 5.2 (The trace theorem)**.** *Let D be a bounded open set in $\mathbb{R}^n$ and suppose that D has a piecewise smooth boundary ∂D. In addition, suppose that D satisfies the cone condition. Then there exists a bounded linear mapping*

$$\gamma : H^1(D) \to L^2(\partial D), \quad \|\gamma v\|_{L^2(\partial D)} \leq C\|v\|_{H^1(D)}, \tag{5.9}$$

where C is a constant depending on D but not on v, such that $\gamma v = v_{|\partial D}$ for all $v \in C^1(\bar{D})$.

The proof of the trace theorem may be found in [7].

Let us first prove the continuity of F. By taking absolute values, estimating the random variables z_1 and z_2 from below by $z_{\min}$, and including all constants in $C > 0$, we obtain

$$\begin{aligned} F(v)^2 &\leq \mathbb{E}\Big[\frac{1}{z_{\min}}(|u_1| + |u_2|)(|v(\omega,-L)| + |v(\omega,L)|)\Big]^2 \\ &= C\mathbb{E}\Big[|v(\omega,-L)| + |v(\omega,L)|\Big]^2. \end{aligned}$$

The generic constant $C > 0$ may change during the subsequent steps. We continue by using (5.8) and the trace theorem to estimate the expectation from above, and finally, by using Jensen's inequality, we arrive at the desired result:

$$\begin{aligned} F(v)^2 &\leq C\mathbb{E}\Big[\|v(\omega,\cdot)\|_{L^2(\partial D)}\Big]^2 \leq C\mathbb{E}\Big[\|v(\omega,\cdot)\|_{H^1(D)}\Big]^2 \\ &\leq C\mathbb{E}\Big[\|v(\omega,\cdot)\|^2_{H^1(D)}\Big] \leq C\|v\|^2_{\mathcal{H}_P}. \end{aligned}$$

The continuity of B is shown in a similar way. By taking absolute values and using the assumptions on the random variables, we obtain

$$\begin{aligned} B(u,v)^2 &\le \mathbb{E}\Big[a_{\max}\int_D |u_{\mathrm{x}}(\omega,\mathrm{x})v_{\mathrm{x}}(\omega,\mathrm{x})|d\mathrm{x} \\ &\qquad + \frac{1}{z_{\min}}(|u(\omega,L)|+|u(\omega,-L)|)(|v(\omega,L)|+|v(\omega,-L)|)\Big]^2 \\ &\le C\mathbb{E}\Big[\|u_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)}\|v_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)} + \|u(\omega,\cdot)\|_{L^2(\partial D)}\|v(\omega,\cdot)\|_{L^2(\partial D)}\Big]^2 \end{aligned}$$

where in the last step the inequality (5.8) is used. The desired result follows now from the trace theorem and the Cauchy–Schwarz inequality:

$$\begin{aligned} B(u,v)^2 &\le C\mathbb{E}\Big[\|u_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)}\|v_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)} + \|u(\omega,\cdot)\|_{H^1(D)}\|v(\omega,\cdot)\|_{H^1(D)}\Big]^2 \\ &\le C\mathbb{E}\Big[\|u(\omega,\cdot)\|_{H^1(D)}\|v(\omega,\cdot)\|_{H^1(D)}\Big]^2 \\ &\le C\mathbb{E}\Big[\|u(\omega,\cdot)\|^2_{H^1(D)}\Big]\mathbb{E}\Big[\|v(\omega,\cdot)\|^2_{H^1(D)}\Big] \\ &\le C\|u\|^2_{\mathcal{H}_P}\|v\|^2_{\mathcal{H}_P}. \end{aligned}$$

Poincaré–Friedrichs inequality, given below for the one-dimensional case, is required for the coercivity of B.

Theorem 5.3 (Poincaré–Friedrichs inequality)**.** *Let D be as above. Then the relation*

$$\|v\|_{L^2(D)} \le 2L\|v_x\|_{L^2(D)} \tag{5.10}$$

holds for all $v \in H^1_0(D) = \{v \in H^1(D) \mid v(-L) = v(L) = 0\}$.

For the proof of Poincaré–Friedrichs inequality, see [7].

By estimating the bilinear form B from below with the help of the assumptions (5.3) and (5.4), we obtain

$$\begin{aligned} B(v,v) &\ge \mathbb{E}\Big[a_{\min}\|v_{\mathrm{x}}(\omega,\cdot)\|^2_{L^2(D)} + \frac{1}{z_{\max}}(|v(\omega,-L)|^2+|v(\omega,L)|^2)\Big] \\ &\ge C\mathbb{E}\Big[\|v_{\mathrm{x}}(\omega,\cdot)\|^2_{L^2(D)} + \|v(\omega,\cdot)\|^2_{L^2(\partial D)}\Big]. \end{aligned} \tag{5.11}$$

Due to an extension of the trace theorem [10], we known that there exists a function $h^v \in H^1(D)$ such that

$$h^v(\omega,-L) = v(\omega,-L), \quad h^v(\omega,L) = v(\omega,L), \tag{5.12}$$

and

$$\|h^v(\omega,\cdot)\|_{H^1(D)} \le C\|v(\omega,\cdot)\|_{L^2(\partial D)}, \tag{5.13}$$

where the constant C does not depend on v. (Notice that (5.13) holds in this form only in one dimension.)

By using the triangle inequality, we can estimate the norm $\|v(\omega,\cdot)\|_{L^2(D)}$ from above to obtain

$$\|v(\omega,\cdot)\|_{L^2(D)} \le \|h^v(\omega,\cdot)\|_{L^2(D)} + \|(v-h^v)(\omega,\cdot)\|_{L^2(D)}.$$

Because the function $v - h^v$ belongs to $H_0^1(D)$, we can use Poincaré–Friedrichs inequality to get

$$\begin{aligned}\|v(\omega,\cdot)\|_{L^2(D)} &\le C\Big(\|h^v(\omega,\cdot)\|_{L^2(D)} + \Big\|\frac{d}{d\mathrm{x}}(v-h^v)(\omega,\cdot)\Big\|_{L^2(D)}\Big)\\ &\le C\Big(\|v_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)} + \|h^v(\omega,\cdot)\|_{L^2(D)} + \|h^v_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)}\Big)\\ &\le C\Big(\|v_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)} + \|v(\omega,\cdot)\|_{L^2(\partial D)}\Big),\end{aligned} \tag{5.14}$$

where the last two steps follow from the triangle inequality, the definition of the $H^1(D)$-norm, and (5.13). Using (5.14) together with (5.8), we thus obtain the estimate

$$\begin{aligned}\|v(\omega,\cdot)\|^2_{H^1(D)} &\le C(\|v_{\mathrm{x}}(\omega,\cdot)\|^2_{L^2(D)} + (\|v_{\mathrm{x}}(\omega,\cdot)\|_{L^2(D)} + \|v(\omega,\cdot)\|_{L^2(\partial D)})^2)\\ &\le C(\|v_{\mathrm{x}}(\omega,\cdot)\|^2_{L^2(D)} + \|v(\omega,\cdot)\|^2_{L^2(\partial D)}).\end{aligned} \tag{5.15}$$

By combining (5.15) with (5.11), we obtain the desired result, namely,

$$\mathbb{E}\Big[\|v(\omega,\cdot)\|^2_{H^1(D)}\Big] \le CB(v,v), \tag{5.16}$$

which proves the coercivity of B.

The unique solvability of (5.5) follows now from the Lax-Milgram lemma.

Theorem 5.4 (The Lax-Milgram lemma)**.** *Let B be a bounded, coercive bilinear form on a Hilbert space H. Then for every bounded linear functional $F \in H^*$, where the space H^* denotes the space of all bounded linear functionals on H, there exists a unique element $u \in H$ such that*

$$B(u,v) = F(v) \tag{5.17}$$

for all $v \in H$.

The proof of the Lax-Milgram lemma may be found in [13].

Remark 5.5. Notice that if the random field a is assumed to be a constant, that is, $a(\omega,\mathrm{x}) \equiv a$, it is straightforward to show that the model problem assumes the analytic solution

$$\begin{aligned}u(\omega,\mathrm{x}) = &\frac{u_2-u_1}{2L+(z_1(\omega)+z_2(\omega))a}\mathrm{x}\\ &+\frac{(u_1+u_2)L+(u_2z_1(\omega)+u_1z_2(\omega))a}{2L+(z_1(\omega)+z_2(\omega))a}.\end{aligned} \tag{5.18}$$

Chapter 6

Stochastic Galerkin finite element method

In this chapter, our aim is to introduce the main ideas behind the stochastic finite element method, sFEM, and apply it to the model problem presented in the previous chapter. We start by discussing the sFEM in general, and then continue by formulating the sFEM, or to be more precise, the stochastic Galerkin finite element method, sGFEM, for the model problem. We also point out some issues faced when implementing the sGFEM with a computer. Numerical examples illustrating the sGFEM are provided in the next chapter.

The (deterministic) finite element method, FEM, can be characterized to be a method for converting a continuum valued problem, such as a partial differential equation, into a discrete problem. The differential equation is first presented in a variational form, that is, the equation is required to hold only in a weak sense; an equation holds in a weak sense if it holds with respect to suitable "test vectors" or "test functions". Formulating the problem in the variational form is in essence the same as to require a solution to the original problem in the sense of distributions. Finally, some constraints are applied on the function space from which the weak solution is sought for in order to discretize the space with a finite set of basis functions.

For the subsequent development, we want to mention one specific FEM, namely the Galerkin finite element method, GFEM[1]. The GFEM is a widely used FEM whose key property is that the error of the weak solution is orthogonal to the corresponding FEM solution space in the sense of "energy inner product". For more information about the deterministic FEM, we refer to [7].

The idea of the sFEM is the same as that of the FEM: formulate the problem in a variational sense, use a finite set of basis functions to discretize the problem, and finally solve the discretized problem to obtain an approximate solution to the original problem. However, in sFEM the problem is discretized in both, random and spatial,

[1]In some settings, the term GFEM corresponds to generalized finite element method.

dimensions, or more precisely, in both parameter spaces: the discretization in the spatial dimension is done in the same way as in the deterministic FEM, and the discretization in the random dimension is usually performed using a (generalized) polynomial chaos expansion introduced in Chapter 4.

Two extensively used stochastic finite element methods exist: the stochastic Galerkin method, sGFEM, and the stochastic collocation method, sCFEM. The idea in the sGFEM is to build on top of the GFEM by adding the polynomial chaos expansion discretization of the random dimension over the standard discretization of the spatial dimension, whereas the idea in the sCFEM is to first sample the solution on a set of predetermined points in the random dimension and then interpolate the obtained probability densities to the whole random dimension. In this work, we consider only the sGFEM. For more information about the sFEM and the sCFEM, see [3, 6, 11, 12].

We proceed by formulating the sGFEM for the model problem given by (5.2).

From now on, we assume the random variables z_1 and z_2 to be log-normal. We know from Chapter 2 that z_1 and z_2 can be written in the form

$$z_1(\omega) = \exp(\mu_1 + \sigma_1 Y_1(\omega)) \quad \text{and} \quad z_2(\omega) = \exp(\mu_2 + \sigma_2 Y_2(\omega)),$$

where Y_1 and Y_2 follow the standard normal distribution. Furthermore, we assume that a is a log-normal random field for which the mean field $\mu_3(\mathrm{x})$ and the covariance function of the underlying Gaussian random field are known. Finally, we suppose that a can be approximated as a truncated exponential Karhunen–Loève expansion. Hence, according to Definition 3.7, we can write:

$$a(\omega, \mathrm{x}) \approx a_M(\omega, \mathrm{x}) = \exp\Big(\mu_3(\mathrm{x}) + \sum_{m=3}^{M+2} \sqrt{\lambda_m} Y_m(\omega)\phi_m(\mathrm{x})\Big), \tag{6.1}$$

where $\mu_3(\mathrm{x})$ corresponds to $\mathbb{E}_a^{\ln}[x]$, and the indexing starts from 3 to distinguish the random variables in the expansion from those used for z_1 and z_2. Recall that the random variables $\{Y_m\}_{m=3}^{M+2}$ are orthogonal and follow the standard normal distribution. Note also that the random variables a, z_1, and z_2 were assumed to be independent, and thus $\{Y_m\}_{m=1}^{M+2}$ are also independent.

From Chapter 4, we know that the random field $u \in \mathcal{H}_P$ can be approximated in the $L_P^2(\Omega)$ sense as

$$u(\omega, \mathrm{x}) \approx \hat{u}(\omega, \mathrm{x}) = \sum_{i=0}^{P} \hat{c}_i(\mathrm{x})\Psi_i(\mathbf{Y}(\omega)), \tag{6.2}$$

where $\mathbf{Y}$ is the random vector consisting of $\{Y_m\}_{m=1}^{M+2}$ and $\hat{c}_i(\mathrm{x})$ is a suitable weight function. We proceed by dividing the domain of the model problem $\bar{D} = [-L, L]$ into $N-1$ intervals with mesh size $h = 2L/(N-1)$ using $N \geq 2$ equally distributed points $\{\mathrm{x}_i\}_{i=1}^{N}$. We also introduce piecewise linear spatial basis functions $\{g_i\}_{i=1}^{N}$

whose value is one at x_i and zero at every x_j, $j \neq i$, that is,

$$g_i(x) = \begin{cases} \frac{\mathrm{x}-\mathrm{x}_{i-1}}{\mathrm{x}_i-\mathrm{x}_{i-1}}, & \text{if } \mathrm{x} \in [\mathrm{x}_{i-1}, \mathrm{x}_i], \\ \frac{\mathrm{x}-\mathrm{x}_{i+1}}{\mathrm{x}_i-\mathrm{x}_{i+1}}, & \text{if } \mathrm{x} \in [\mathrm{x}_i, \mathrm{x}_{i+1}], \\ 0, & \text{otherwise,} \end{cases}$$

and discretize the weight functions:

$$\hat{c}_i(\mathrm{x}) \approx \sum_{k=1}^{N} d_{ik} g_k(\mathrm{x}), \tag{6.3}$$

where $\{d_{ik}\}_{k=1}^{N} \subset \mathbb{R}$ are suitable constant coefficients. Thus, the field u is altogether approximated as

$$u(\omega, \mathrm{x}) \approx \tilde{u}(\omega, \mathrm{x}) = \sum_{i=0}^{P} \sum_{k=1}^{N} d_{ik} \Psi_i(\mathbf{Y}(\omega)) g_k(\mathrm{x}). \tag{6.4}$$

Besides discretizing the random field u, we have discretized the whole space $\mathcal{H}_P$; any random field $v \in \mathcal{H}_P$ can be written as a linear combination of the basis functions $\{g_k \Psi_i(\mathbf{Y})\}$ in the limit when M, N, and P tend to infinity [7, 20].

After discretizing the space $\mathcal{H}_P$, the problem (5.5) is transformed into the form: find constants $\{d_{ik}\} \subset \mathbb{R}$ such that the equation

$$B(\tilde{u}, \tilde{v}) = F(\tilde{v}), \tag{6.5}$$

where $\tilde{v} = g_l \Psi_j(\mathbf{Y})$, holds for all $l = 1, \ldots, N$ and $j = 0, \ldots, P$. Expanding the left-hand and right-hand sides of the equation (6.5) gives

$$\begin{aligned} B(\tilde{u}, \tilde{v}) = \mathbb{E}\Big[\int_{-L}^{L} a(\omega, \mathrm{x}) \sum_{i=0}^{P} \sum_{k=1}^{N} d_{ik} g_k'(\mathrm{x}) \Psi_i(\mathbf{Y}(\omega)) g_l'(\mathrm{x}) \Psi_j(\mathbf{Y}(\omega)) d\mathrm{x} \\ + \frac{1}{z_2(\omega)} \sum_{i=0}^{P} \sum_{k=1}^{N} d_{ik} g_k(L) \Psi_i(\mathbf{Y}(\omega)) g_l(L) \Psi_j(\mathbf{Y}(\omega)) \\ + \frac{1}{z_1(\omega)} \sum_{i=0}^{P} \sum_{k=1}^{N} d_{ik} g_k(-L) \Psi_i(\mathbf{Y}(\omega)) g_l(-L) \Psi_j(\mathbf{Y}(\omega)) \Big] \end{aligned}$$

and

$$F(\tilde{v}) = \mathbb{E}\Big[\frac{u_2}{z_2(\omega)} g_l(L) \Psi_j(\mathbf{Y}(\omega)) + \frac{u_1}{z_1(\omega)} g_l(-L) \Psi_j(\mathbf{Y}(\omega)) \Big].$$

By moving the expectations inside the integral and using the properties of the func-

tions $\{g_i\}_{i=1}^N$, we obtain

$$\begin{aligned} B(\tilde{u},\tilde{v}) = &\sum_{i=0}^{P}\sum_{k=1}^{N} d_{ik}\int_{-L}^{L} g_k'(\mathrm{x})g_l'(\mathrm{x})\mathbb{E}[a(\omega,\mathrm{x})\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))]d\mathrm{x} \\ &+\sum_{i=0}^{P}\mathbb{E}\Big[\frac{\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))}{z_2(\omega)}\Big]d_{iN}g_l(L) \\ &+\sum_{i=0}^{P}\mathbb{E}\Big[\frac{\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))}{z_1(\omega)}\Big]d_{i1}g_l(-L) \end{aligned}$$

and

$$F(\tilde{v}) = \mathbb{E}\Big[\frac{\Psi_j(\mathbf{Y}(\omega))}{z_2(\omega)}\Big]u_2 g_l(L) + \mathbb{E}\Big[\frac{\Psi_j(\mathbf{Y}(\omega))}{z_1(\omega)}\Big]u_1 g_l(-L).$$

Forming the equation (6.5) for all $l = 1,\dots,N$ and $j = 0,\dots,P$ produces a set of $N\times(P+1)$ algebraic equations which can be written in the form

$$\mathbf{A}\mathbf{d} = \mathbf{J}, \tag{6.6}$$

where $\mathbf{A}$ is a symmetric block matrix of dimension $N\times(P+1)$, and $\mathbf{d}$ and $\mathbf{J}$ are block vectors. The ijth block matrix of $\mathbf{A}$, $\mathbf{A}_{ij}$, is characterized by

$$\begin{aligned} \{\mathbf{A}_{ij}\}_{kl} = &\int_{-L}^{L} g_k'(\mathrm{x})g_l'(\mathrm{x})\mathbb{E}[a(\omega,\mathrm{x})\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))]d\mathrm{x} \\ &+\delta_{kN}\mathbb{E}\Big[\frac{\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))}{z_2(\omega)}\Big]\delta_{lN} \\ &+\delta_{k1}\mathbb{E}\Big[\frac{\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))}{z_1(\omega)}\Big]\delta_{l1}, \end{aligned} \tag{6.7}$$

the ith block of the vector $\mathbf{d}$ is given by $\mathbf{d}_i = (d_{i1},\dots,d_{iN})^{\mathrm{T}}$, and the jth block of the vector $\mathbf{J}$ by

$$\mathbf{J}_j = \Big(\mathbb{E}\Big[\frac{\Psi_j(\mathbf{Y}(\omega))}{z_1(\omega)}\Big]u_1, 0,\dots,0,\mathbb{E}\Big[\frac{\Psi_j(\mathbf{Y}(\omega))}{z_2(\omega)}\Big]u_2,\Big)^{\mathrm{T}}, \tag{6.8}$$

where $0\le i,j\le P$ and $1\le k,l\le N$.

Next, we turn our attention to evaluation of the expectations and integrals in (6.6). To this end, let $h(\omega,\mathrm{x})$ be an arbitrary random field that can be approximated using the truncated exponential Karhunen–Loève expansion,

$$h(\omega,\mathrm{x}) \approx \tilde{h}_M(\omega,\mathrm{x}) = \exp\Big(h_0(\mathrm{x}) + \sum_{m=1}^{M} Y_m(\omega)h_m(\mathrm{x})\Big), \tag{6.9}$$

where the random variables $\{Y_m\}$ are independent and follow the standard normal distribution. We want to use the expansion (6.9) to calculate an expectation of the

type $\mathbb{E}[h(\omega,\mathrm{x})\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))]$. We obtain

$$\mathbb{E}\Big[h(\mathrm{x})\Psi_i(\mathbf{Y})\Psi_j(\mathbf{Y})\Big] \approx \exp(h_0(\mathrm{x}))\mathbb{E}\Big[\exp\Big(\sum_{m=1}^{M} Y_m h_m(\mathrm{x})\Big)\Psi_i(\mathbf{Y})\Psi_j(\mathbf{Y})\Big]$$
$$= \exp(h_0(\mathrm{x}))(2\pi)^{-\frac{M}{2}}\int_{-\infty}^{\infty}\exp\Big(\sum_{m=1}^{M}(\mathrm{y}_m h_m(\mathrm{x}) - \frac{1}{2}\mathrm{y}_m^2)\Big)\Psi_i(\mathrm{y})\Psi_j(\mathrm{y})d\mathrm{y}.$$

By completing the square and making the change of variables

$$\tilde{\mathrm{y}}_m(\mathrm{x}) = \mathrm{y}_m - h_m(\mathrm{x}), \quad m = 1,\dots,M,$$

we obtain

$$\mathbb{E}\Big[h(\mathrm{x})\Psi_i(\mathbf{Y})\Psi_j(\mathbf{Y})\Big] \approx \exp\Big(h_0(\mathrm{x}) + \frac{1}{2}\sum_{m=1}^{M} h_m(\mathrm{x})^2\Big)$$
$$\times (2\pi)^{-\frac{M}{2}}\int_{-\infty}^{\infty}\exp\Big(-\frac{1}{2}\sum_{m=1}^{M}\tilde{\mathrm{y}}_m(\mathrm{x})^2\Big)\Psi_i(\tilde{\mathrm{y}}(\mathrm{x})+\mathbf{h}(\mathrm{x}))\Psi_j(\tilde{\mathrm{y}}(\mathrm{x})+\mathbf{h}(\mathrm{x}))d\tilde{\mathrm{y}}$$
$$= \exp\Big(h_0(\mathrm{x}) + \frac{1}{2}\sum_{m=1}^{M} h_m(\mathrm{x})^2\Big)\mathbb{E}\Big[\Psi_i(\tilde{\mathbf{Y}}(\mathrm{x})+\mathbf{h}(\mathrm{x}))\Psi_j(\tilde{\mathbf{Y}}(\mathrm{x})+\mathbf{h}(\mathrm{x}))\Big], \qquad (6.10)$$

where for each $\mathrm{x}\in[-L,L]$ each component of $\tilde{\mathbf{Y}}(\mathrm{x})$ follows the standard normal distribution and $\mathbf{h}(\mathrm{x}) = (h_1(\mathrm{x}),\dots,h_M(\mathrm{x}))$. Particularly, if the random field h can be written in the form

$$h(\omega,\mathrm{x}) = \exp(h_0(\mathrm{x}) + Y_1(\omega)h_1(\mathrm{x})),$$

it holds that

$$\mathbb{E}[h(\mathrm{x})\Psi_j(\mathbf{Y})] = \exp\Big(h_0(\mathrm{x}) + \frac{1}{2}h_1(\mathrm{x})^2\Big)\mathbb{E}[\Psi_j(\tilde{Y}_1(\mathrm{x}) + h_1(\mathrm{x}))]. \qquad (6.11)$$

Based on the formulas (6.10) and (6.11), the expectations in (6.7) and (6.8) are easy to evaluate. By using the truncated exponential Karhunen–Loève expansion for a, given by (6.1), we obtain that

$$\mathbb{E}[a(\omega,\mathrm{x})\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))] \approx$$
$$\exp\Big(\mu_3(\mathrm{x}) + \frac{1}{2}\sum_{m=3}^{M+2} a_m(\mathrm{x})^2\Big)\mathbb{E}\Big[\Psi_i(\tilde{\mathbf{Y}}(\mathrm{x})+\mathbf{a}(\mathrm{x}))\Psi_j(\tilde{\mathbf{Y}}(\mathrm{x})+\mathbf{a}(\mathrm{x}))\Big], \qquad (6.12)$$

where $\mathbf{a}(\mathrm{x}) = (0,0,a_3(\mathrm{x}),\dots,a_{M+2}(\mathrm{x}))$ and $a_m(\mathrm{x}) = \sqrt{\lambda_m}\phi_m(\mathrm{x})$. The expectation

$$\mathbb{E}\Big[\Psi_i(\tilde{\mathbf{Y}}(\mathrm{x})+\mathbf{a}(\mathrm{x}))\Psi_j(\tilde{\mathbf{Y}}(\mathrm{x})+\mathbf{a}(\mathrm{x}))\Big] \qquad (6.13)$$

can be calculated without any numerical evaluation of integrals: by expanding the product of the two polynomials, namely,

$$\Psi_i(\tilde{\mathbf{Y}}(\mathrm{x}) + \mathbf{a}(\mathrm{x}))\Psi_j(\tilde{\mathbf{Y}}(\mathrm{x}) + \mathbf{a}(\mathrm{x})), \tag{6.14}$$

and replacing the different powers of $\tilde{Y}_m(\mathrm{x})$ using (2.15), which is allowed since the random variables $\{\tilde{Y}_m(\mathrm{x})\}_{m=1}^{M+2}$ are independent, we obtain the value of (6.13). Since the value of (2.15) depends only on the power of $\tilde{Y}_m(\mathrm{x})$ (every $\tilde{Y}_m(\mathrm{x})$ has zero mean and unit variance), the calculation of (6.13) is reduced to a simple substitution rule: expand all terms in (6.14) and replace all occurrences of $\tilde{Y}_m(\mathrm{x})$ with the value given by (2.15) corresponding to the power of the $\tilde{Y}_m(\mathrm{x})$. Notice that the expectation (6.13) is a multivariate polynomial in terms of $\mathbf{a}$.

The other expectations in (6.7) and (6.8) can be calculated in a similar way by using (6.11) and the fact that the reciprocal of a log-normal random variable $\ln\mathcal{N}(\mu, \sigma^2)$ can be given as $\exp(-\mu - \sigma Y)$, where Y follows the standard normal distribution.

Let us continue by discussing the integral

$$\int_{-L}^{L} g_k'(\mathrm{x})g_l'(\mathrm{x})\mathbb{E}[a(\omega, \mathrm{x})\Psi_i(\mathbf{Y}(\omega))\Psi_j(\mathbf{Y}(\omega))]d\mathrm{x}. \tag{6.15}$$

Notice first that the derivatives of the functions $\{g_i\}_{i=1}^{N}$ are piecewise constant and are supported only on specific intervals of the domain D. Thus, the term $g_k'(\mathrm{x})g_l'(\mathrm{x})$ is constant on each subinterval and reduces the integration limits in (6.15). Furthermore, we know from the previous paragraphs that the expectation in (6.15) is a multivariate polynomial in terms of $\mathbf{a}$. However, due to the exponential coefficient in front of the expectation (6.13) in (6.12), the approximate expectation (6.12) is not a multivariate polynomial but a sum of more complicated functions of $\mathbf{a}$.

The integral (6.15) can be integrated term by term. If all possible terms of the multivariate polynomial (6.13) are formed, multiplied with the corresponding exponential coefficient, integrated over all possible supports of the term $g_k'(\mathrm{x})g_l'(\mathrm{x})$, and tabulated beforehand, the calculation of (6.15) is reduced into a simple substitution task. Notice that when computing the required integrals, we are required to carry out the integrations numerically because of the exponential coefficient. Numerical integration is not discussed in this work; see [8] for information about the topic. We want to emphasize that most of the cells in $\mathbf{A}$ are zero with a non-trivial pattern. The sparsity and the symmetry of $\mathbf{A}$ should be taken into account in the implementation of the sGFEM.

After forming the matrix equation (6.6) and solving it for the constants $\{d_{ik}\}$, some post-processing is required to obtain the desired response statistics for the approximate solution $\tilde{u}$. By using equation (6.3), we form the functions $\{\hat{c}_i\}_{i=0}^{P}$. After that, the solution $\tilde{u}$ is written in the form of (6.2). Finally, the results presented in Section 4.2 are used to calculate the desired response statistics for the approximate solution $\tilde{u}$.

Numerical examples illustrating the sGFEM are presented in the next chapter.

Chapter 7

Numerical examples

We present five numerical examples about our model problem to demonstrate the functionality of the stochastic Galerkin finite element method.

In the first two examples, the main focus is to study the effect of the variance of the contact impedance random variables. We assume the random field a to be identically one, and hence, from Remark 5.5 on page 31, we know the exact solution to the model problem. Thus, we can validate the obtained sGFEM solutions in comparison to a ground truth.

In the last three examples, the random field a is assumed to be as introduced in the previous chapters. First, we justify the sGFEM by a simple example, where only the first term of the truncated exponential Karhunen–Loève expansion is taken into account. After that, we proceed by taking more terms into the expansion, that is, we increase the number of stochastic variables in the model problem. Finally, we consider large variances for some of the random parameters. Remark 5.5 cannot be used for the exact solution in these examples because the random field is not identical to a constant. Therefore, we are required to resort to Monte Carlo sampling, as explained in Chapter 1, to obtain "reference solutions" that we can use to evaluate the solutions given by the sGFEM.

The domain corresponding to $L = 1$ and the boundary conditions $u_1 = 1$ and $u_2 = 2$ were assumed for the model problem in all of our numerical examples. The choices were purely arbitrary; we could have chosen other parameters as well. Besides the domain and the boundary conditions, we are also required to specify the distributions of the random variables a, z_1, and z_2, the number of terms in the truncated exponential Karhunen–Loève expansion, M, and the discretization levels N and p for the spatial and stochastic dimensions, respectively, for each example. Recall that the parameter N is the number of piecewise linear functions used to discretize the spatial dimension and that the parameter p is the maximum order of the polynomial chaos expansion.

When the parameters M, N, and p have been specified, the dimension of the problem, or more precisely the size of the square matrix $\mathbf{A}$ in (6.6), is given by the

Table 7.1: The number of unknowns as given by (7.1) for some values of the parameters M, N, and p.

p	0	2	4	6	8	10	11
$N = 2, M = 0$	2	12	30	56	90	132	156
$N = 20, M = 1$	20	200	700	1680	3300	5720	7280
$N = 20, M = 4$	20	560	4200	18480	60060	160160	247520

formula

$$N \times (P + 1) = N \times \frac{(2 + M + p)!}{(2 + M)!p!} \tag{7.1}$$

as can be seen from (4.5). In Table 7.1, the dimension of the problem is shown for some values of the parameters M, N, and p. Notice that the size of the matrix $\mathbf{A}$, and thus the running time of the sGFEM, increases rapidly and becomes quite fast infeasible if the parameter M and/or p is increased. Hence, it would be desirable to be able to obtain good results with the sGFEM already when the parameters M and p are relatively small.

Recall that the matrix $\mathbf{A}$ is sparse and symmetric so the numbers in Table 7.1 do not tell the whole story. However, it is not trivial to discern whether a cell in the matrix $\mathbf{A}$ is zero or not before expanding the products of the polynomial chaos polynomials and imposing the substitution rules for the cells as explained in the previous chapter. Thus, at first, we are required to consider $\mathbf{A}$ as a full matrix.

In the first two examples, the random field is assumed to be identical to a constant, and thus, from Remark 5.5 we know that the exact solution is a first order polynomial in the spatial dimension and is given by (5.18). Hence, it is enough to use two linear functions to discretize the spatial dimension, that is, we use $N = 2$ in the first two examples.

Example 1

In the first example, we fix the log-normal random variables z_1 and z_2 to follow the probability density $\ln\mathcal{N}(0, 1/4)$. By specifying the parameter p, we can calculate the sGFEM solution for the problem as described in the previous chapter. Figure 7.1 shows the L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ as functions of p when the sGFEM solution is compared against the exact solution.

Figure 7.2 shows how the sGFEM mean field and the sGFEM variance function approach the corresponding exact solution. When $p \geq 1$, the exact mean field (and when $p \geq 2$, the variance function) is practically impossible to distinguish from the corresponding exact solution.

From Figures 7.1 and 7.2, we can deduce that in this first example the sGFEM works and that the corresponding numerical solution converges quickly to the exact solution. This is not the case in the next example where the variance of the random variable z_2 is increased.

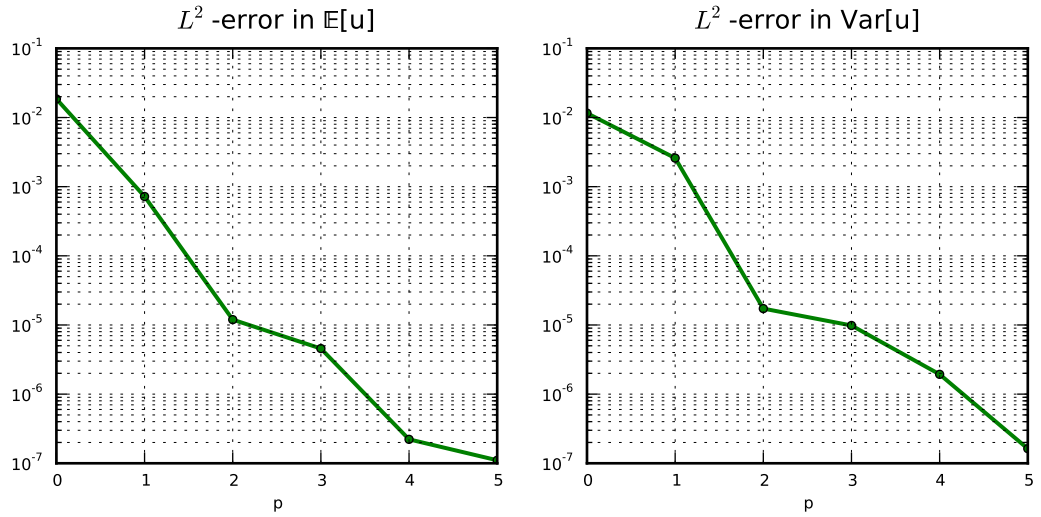

Figure 7.1: The L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ for the first example as functions of p when the sGFEM solution is compared against the exact solution.

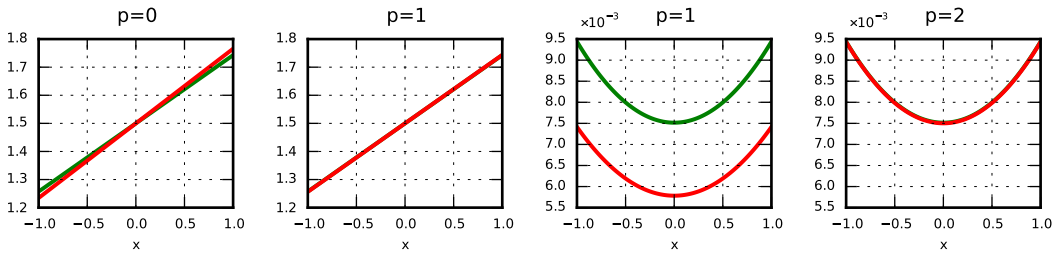

Figure 7.2: Left: the exact mean field (green) and the mean field given by the sGFEM (red) for the first example when $p = 0$ and $p = 1$. Right: the exact variance function (green) and the variance function given by the sGFEM (red) for the first example when $p = 1$ and $p = 2$.

Example 2

In the second example, we let z_1 still follow the log-normal probability density $\ln\mathcal{N}(0, 1/4)$, but we change the distribution of z_2 to be $\ln\mathcal{N}(0, 9)$, that is, we increase the variance and expected value of z_2 considerably. Figure 7.3 shows the L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ as functions of p when the sGFEM solution is compared against the exact solution. From Figure 7.3, we see that the convergence of the mean field and the variance function is much slower than in the first example. Moreover, the

convergence of the solution is periodical with alternating intervals of slow and fast convergence. The L^2-error in the variance oscillates even more and does not decrease much before $p > 2$.

Figure 7.4 shows how the sGFEM mean field (red) gradually approaches the exact mean field (green) as a function of p. The analogous convergence is shown for the variance function in Figure 7.5.

In order to understand why the convergence of the sGFEM is not as rapid as in the first example, we take a closer look at the solution function u. In Figure 7.6, a slice of the exact solution function $u(\omega, \mathrm{x})$ (green) is drawn as a function of Y_2 in the case when the random variable Y_1 has the value 0 and the spatial variable x is fixed to zero. Here Y_1 and Y_2 are the normally distributed random variables defining z_1 and z_2, respectively. The obtained sGFEM solution (red) is also drawn in the same figure for different values of p.

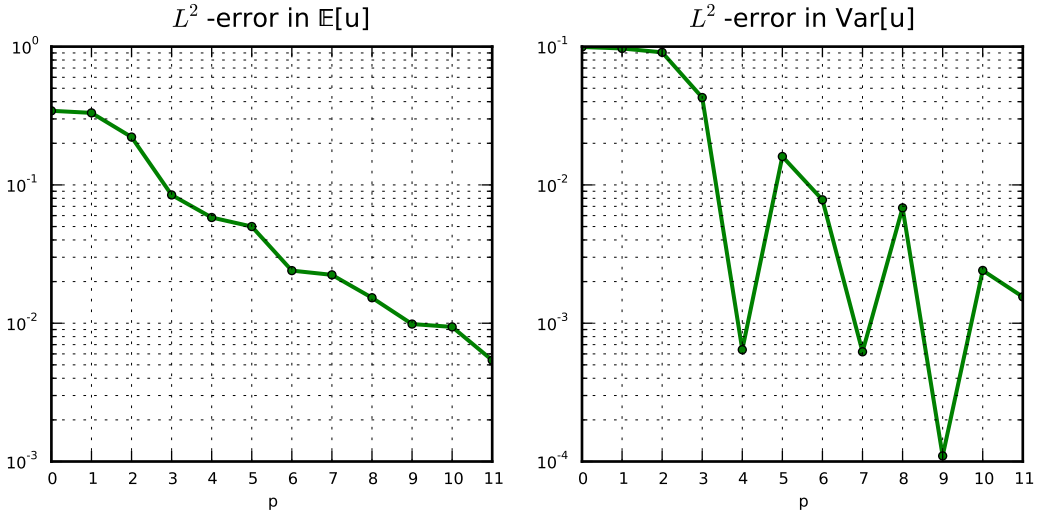

Figure 7.3: The L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ for the second example as functions of p when the sGFEM solution is compared against the exact solution.

From Figure 7.6 it is apparent that the sGFEM solution tries to approximate the exact solution as well as possible in the vicinity of zero, or more precisely on the interval $[-5, 0]$. Outside this interval, the obtained sGFEM solution is poor when compared to the exact solution. However, the situation is not as severe as Figure 7.6 would suggest: if we multiply the obtained sGFEM solution with the probability density function for the random variable Y_2, we obtain Figure 7.7 which is what we should examine, or at least what we should examine in the case of the mean field, because it gives the error in the sense of probability. Figure 7.7 demonstrates that the probability density function of Y_2 reduces the significance of the parts of the sGFEM solution for which the correspondence with the exact solution is poor. We may thus conclude that it is more important to approximate the exact solution well in the vicinity of zero, as is the case in Figure 7.6, since the "probability mass" of a given point decreases rapidly when one moves away from the origin.

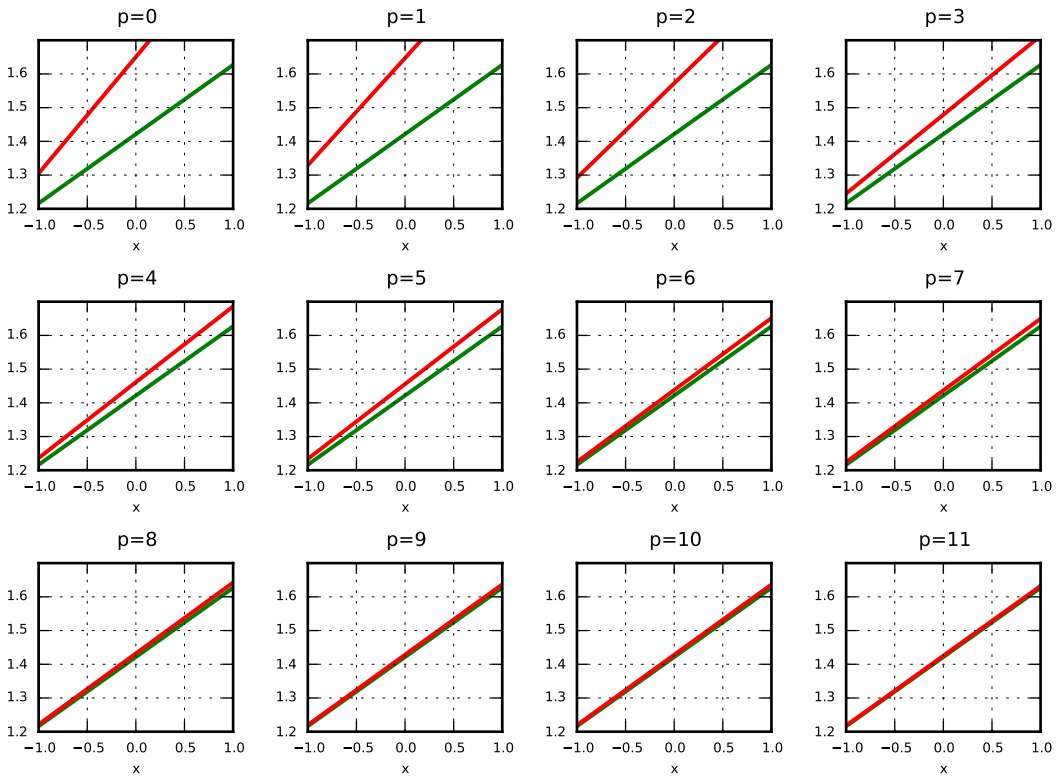

Figure 7.4: The sGFEM mean field (red) and the exact mean field (green) for the second example as functions of p.

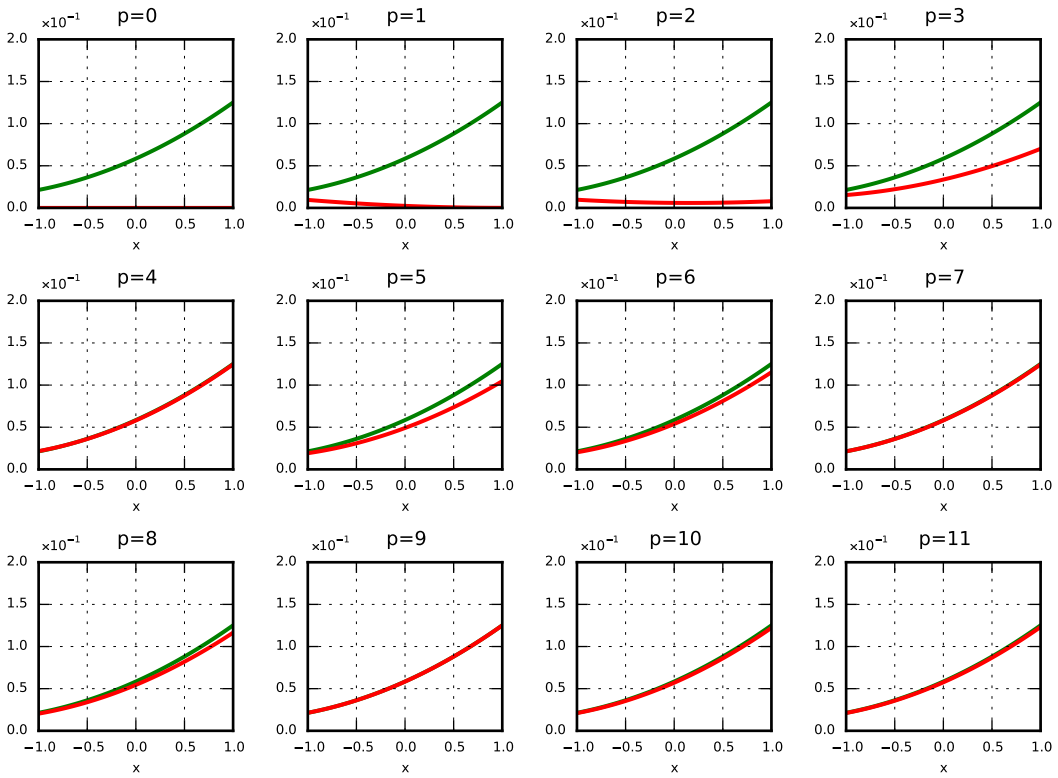

Figure 7.5: The sGFEM variance function (red) and the exact variance function (green) for the second example as functions of p.

In order to present the analogous visualization for the variance, one should first square the solution function and then multiply it with the probability density. However, the resulting image would look qualitatively the same as Figure 7.7 with only some difference in the size of the discrepancy between the exact and the sGFEM solutions, and thus we omit such a consideration.

Nevertheless, Figure 7.6 gives us an idea of what may be causing the stalling and oscillation in the L^2-error of the variance function: near the origin the exact solution decreases rapidly, and thus high order Hermite polynomials, that is, larger p, are required to properly approximate the solution. Figure 7.8, where the same slice is drawn for the first example, suggests that our hypothesis is accurate: when the variance of the random variable z_2 is reduced, the slope of the exact solution as a function of Y_2 is much smaller than in the second example, and thus the sGFEM solution agrees well with the exact solution in the vicinity of zero already when $p = 2$.

We expect to encounter same kind of phenomenon in the following examples where a is a proper random field, that is, the speed of convergence for sGFEM is expected to slow down when the variance of z_1 and/or z_2 is increased.

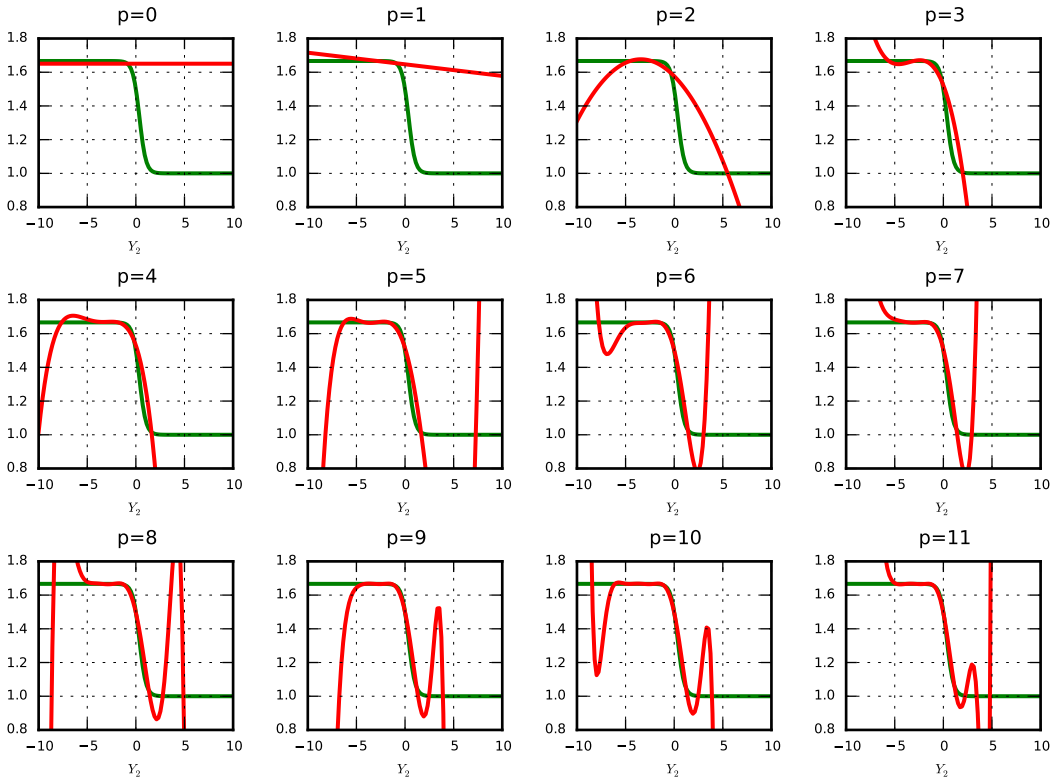

Figure 7.6: The exact solution u (green) and the corresponding sGFEM solution (red) for the second example as functions of Y_2 and p when $x = 0$ and $Y_1 = 0$.

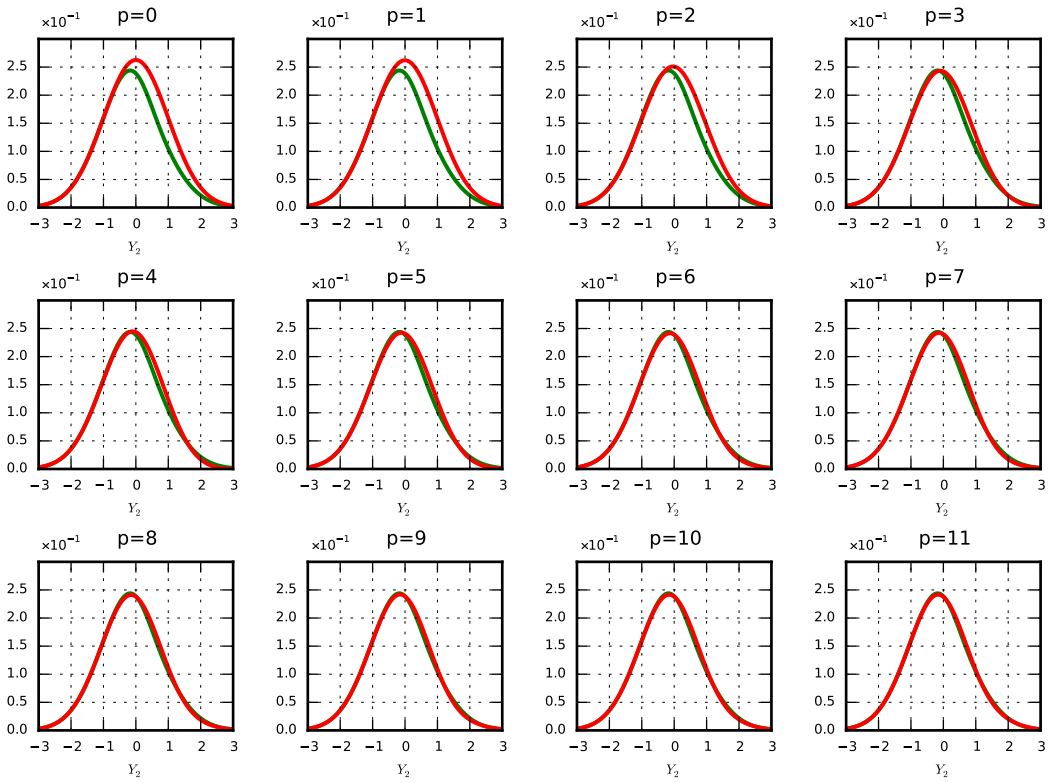

Figure 7.7: The exact solution u multiplied with the probability density function of the random variable Y_2 (green) and the corresponding sGFEM solution (red) for the second example as functions of Y_2 and p when $x = 0$ and $Y_1 = 0$.

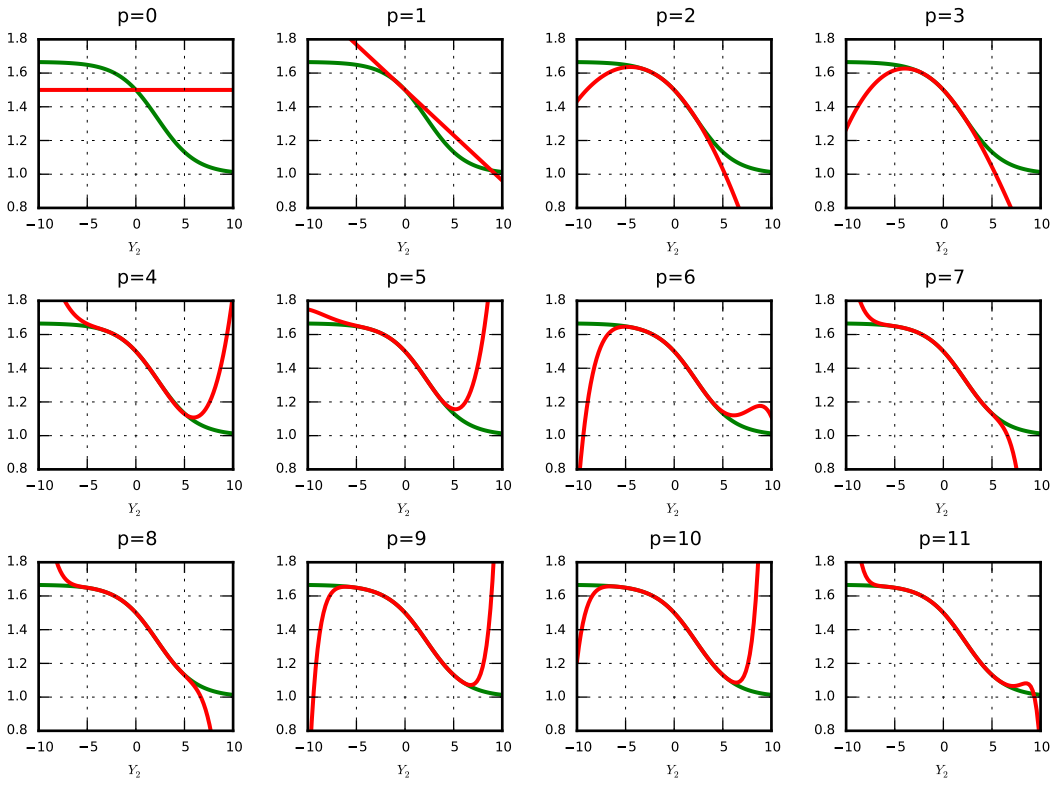

Figure 7.8: The exact solution u (green) and the corresponding sGFEM solution (red) for the first example as functions of Y_2 and p when $x = 0$ and $Y_1 = 0$.

Examples 3 and 4

In the following two examples and in the one presented in the next section, the random field a is not anymore assumed to be identical to a constant. Thus, we do not have the analytical solution for the problem at our disposal as we did in the previous two examples. Therefore, we evaluate the sGFEM via comparison with "exact" solutions obtained by using extensive Monte Carlo simulations with the same M and N.

Since there is no guarantee that the solution would anymore be a first order polynomial in the spatial dimension, the number of piecewise linear functions used to discretize the interval $[-1, 1]$ is fixed to $N = 20$, instead of the previous $N = 2$, for the remaining three examples. Moreover, the correlation length of the covariance function for the random field a is assumed to be one, that is, $b = 1$. Recall that the covariance function for the underlying Gaussian random field of a is

$$V_a^{\ln}(\mathrm{x}_1, \mathrm{x}_2) = \sigma^2 \exp(-|\mathrm{x}_1 - \mathrm{x}_2|/b).$$

Table 7.1 gives the number of unknowns for the examples to be considered in the following.

In the third example, we fix the random variables z_1 and z_2 to follow the probability densities $\ln\mathcal{N}(0, 1/4)$ and $\ln\mathcal{N}(0, 1)$, respectively. The Gaussian random field determining the log-normal field a is assumed to be distributed pointwise as $\mathcal{N}(0, 4)$. For simplicity, the mean field of a is set to a constant. Figure 7.9 shows the trend of the eigenvalues in the exponential Karhunen–Loève expansion for the random field a in this specific case. Due to (3.19), we know that if the variance of the underlying Gaussian random field of a is modified, the eigenvalues shown in Figure 7.9 are only multiplied by a constant, that is, the form of the figure stays the same. In this example, we take into account only the first eigenvalue of the exponential Karhunen–Loève expansion for the random field a, that is, the parameter M is assumed to be one. Thus, the total number of random parameters in this example is three.

The obtained results are shown in Figures 7.10–7.12. Figure 7.10 shows the L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ as functions of p when the sGFEM solution is compared against an accurate solution obtained via extensive Monte Carlo simulation. Figure 7.11 demonstrates how the sGFEM mean field approaches the Monte Carlo solution as p increases and the same behavior is shown for the variance function in Figure 7.12.

The L^2-error in $\mathbb{E}[u]$ decreases rapidly until, say $p = 4$, after which increasing p does not seem to affect the error considerably. The same phenomena is eventually also seen in the L^2-error for $\mathrm{Var}[u]$ as the rate of the convergence slowly decreases when p is increased. This may be a consequence of the interplay between the stochastic and spatial errors; recall that N is held fixed.

We proceed by taking more eigenvalues from the Karhunen–Loève expansion into account. Based on Figure 7.9, we use the parameter $M = 4$ for the fourth example, meaning that the total number of the random parameters in the next example is

six. The size of the problem is substantially larger than in the previous cases as can be seen from Table 7.1.

Figures 7.13–7.15, which are arranged analogously to Figures 7.10–7.12, show the results for the fourth example. These figures show that the sGFEM works well, and that the convergence of the mean field and the covariance function is close to being exponential. We do not show results beyond $p = 6$, as we did in the previous examples, because the running time for our implementation of the sGFEM starts to be frustrating long for larger values of p.

Notice that the shapes of the mean fields in Figures 7.11 and 7.14 and those of the variances in Figures 7.12 and 7.15 differ quite a bit from each other. Consequently, choosing $M = 1$, as in Example 3, is not a feasible option from a practical point of view as it obviously results in a too large discretization error in the stochastic dimension.

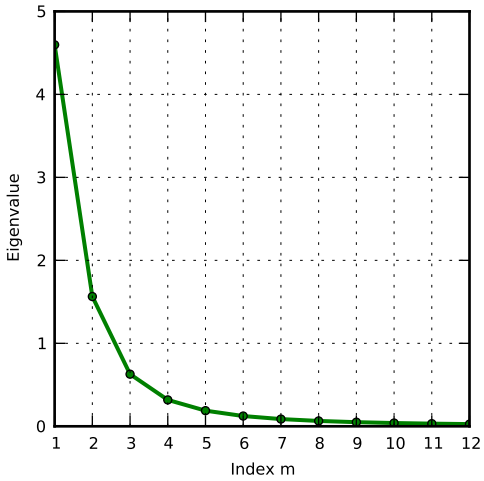

Figure 7.9: Trend of the eigenvalues in the exponential Karhunen–Loève expansion for the random field a as given by (3.19) in the case when $\sigma = 2$ and $b = 1$.

Example 5

In this last example, we see that the increase of the variance of the random field changes the behavior of the sGFEM solution dramatically. We return to consider only a single eigenvalue in the Karhunen–Loève expansion, that is, we fix $M = 1$, and thus the total number of random parameters in this example is three. The number of the piecewise linear functions used for the discretization of the spatial dimension and the probability densities of the random variables z_1 and z_2 are as in the third example. We increase the variance of the random field a by choosing the probability density $\mathcal{N}(0, 25)$ for the underlying Gaussian random field. To sum up, this last example is identical to the third one apart from the considerably higher

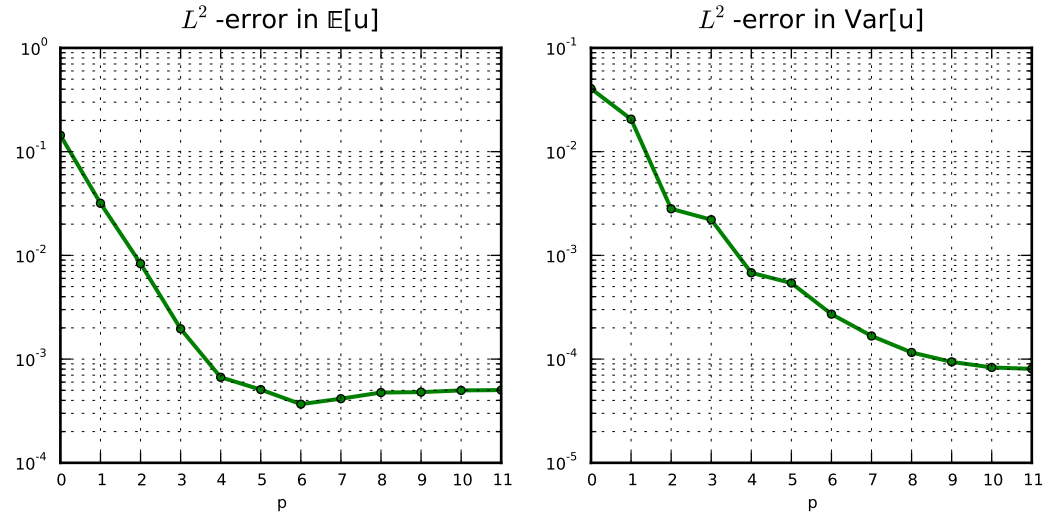

Figure 7.10: The L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ for the third example as functions of p when the sGFEM solution is compared against an extensive Monte Carlo simulation. ($M = 1$, $N = 20$)

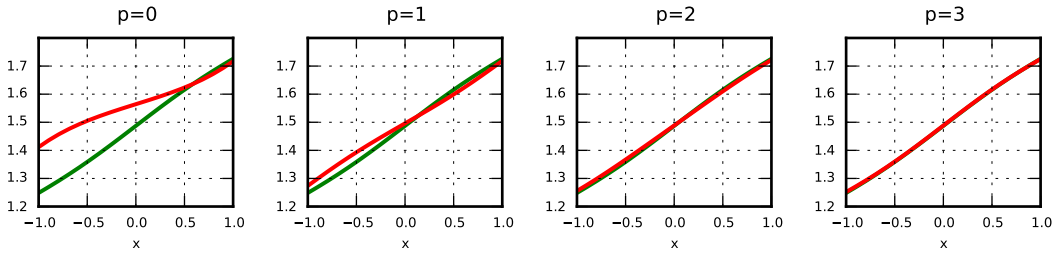

Figure 7.11: The sGFEM mean field (red) and the Monte Carlo solution (green) for the third example as functions of p.

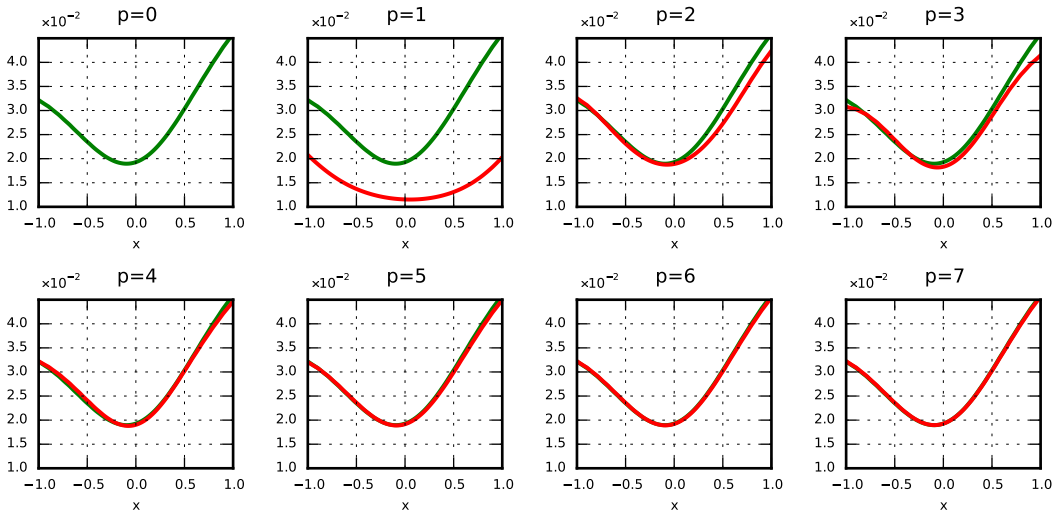

Figure 7.12: The sGFEM variance function (red) and the Monte Carlo solution (green) for the third example as functions of p.

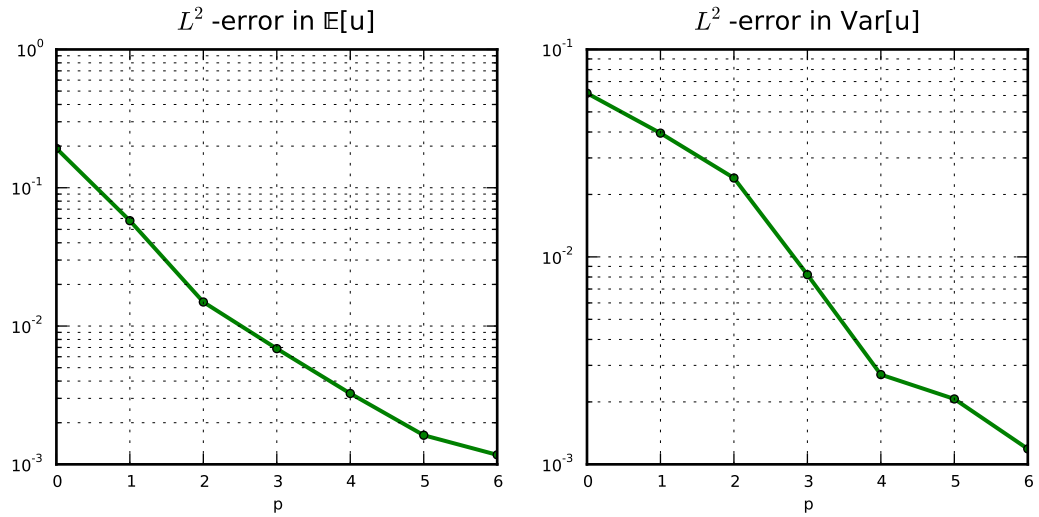

Figure 7.13: The L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ for the fourth example as functions of p when the sGFEM solution is compared against an extensive Monte Carlo simulation. ($M = 4$, $N = 20$)

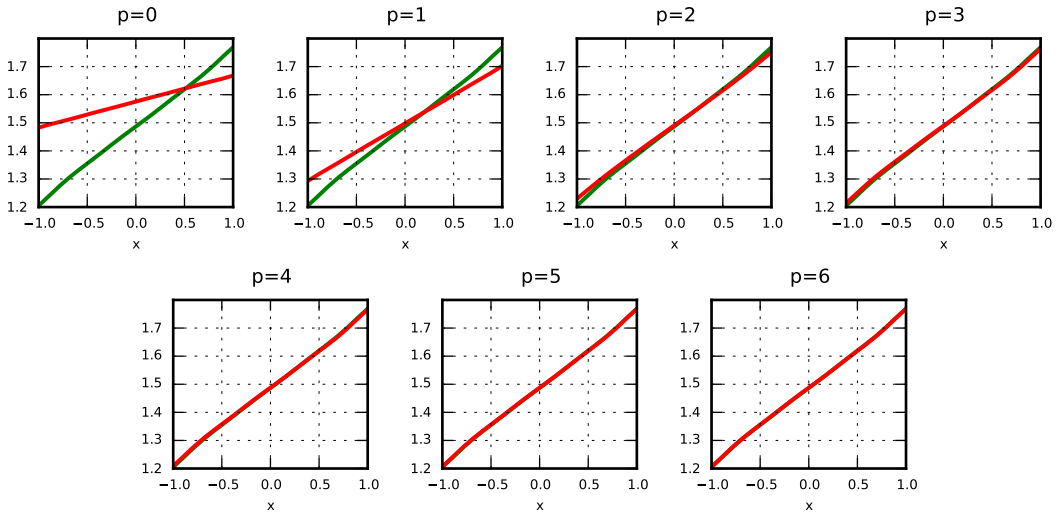

Figure 7.14: The sGFEM mean field (red) and the Monte Carlo solution (green) for the fourth example as functions of p.

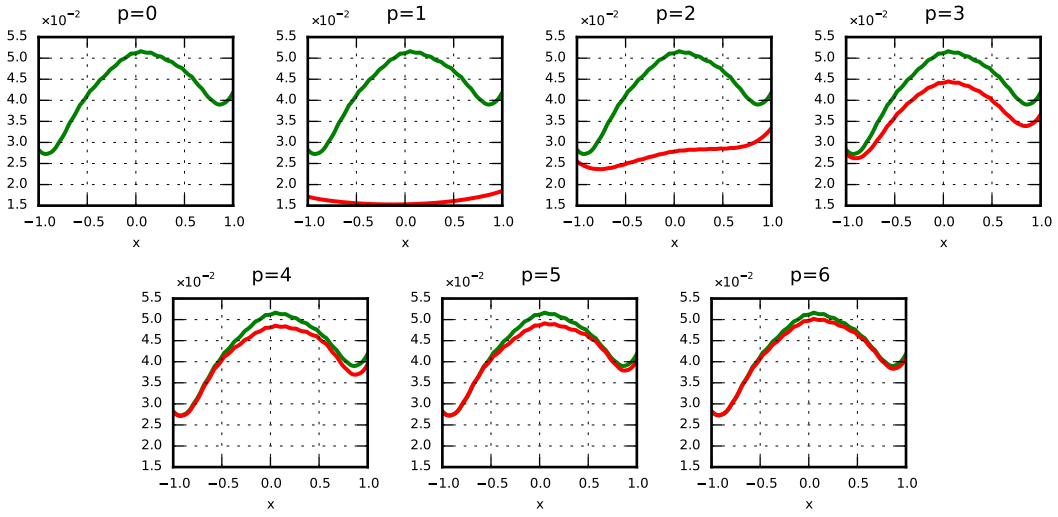

Figure 7.15: The sGFEM variance function (red) and the Monte Carlo solution (green) for the fourth example as functions of p.

variance for a.

The numerical results for the last example are visualized in Figures 7.16–7.18 in the same way as for the previous two examples. These figures show that in this case the sGFEM solution does not converge satisfactorily. To begin, the L^2-error in $\mathbb{E}[u]$ decreases as in the previous cases. However, after $p = 6$, the convergence of the error slows down, starts to oscillate and finally begins to increase. The situation with the error in the variance is even more severe: after $p = 4$, the error starts to increase fast.

Figure 7.18 demonstrates that the variance of the sGFEM solution starts to oscillate in an uncontrolled manner when p increases. This could indicate that the culprit for the failing of the sGFEM is some kind of polynomial interpolation error taking place in the numerical algorithm. However, we were not able to track down the exact reason for the bad behavior of the sGFEM in this case. The same phenomenon is eventually also seen if the variances of the other random variables, that is, z_1 or z_2 are increased. However, we were not able to reproduce the phenomenon in the case when the random field a is assumed to be identical to a constant. We also used different number of functions to discretize the spatial dimension with no meaningful effect. Deducing the reason causing the failing of the sGFEM in the high variance case is left for future studies.

The reader should note, however, that the variance of the random field a is enormous in this example: according to (2.17) and (2.18) the expected value and variance of a log-normal random variable defined by the distribution $\ln\mathcal{N}(0, 25)$ are $268 \cdot 10^3$ and $5.18 \cdot 10^{21}$, respectively. As a consequence, the failure of the sGFEM algorithm cannot be considered as a disaster, and moreover one cannot even completely trust the Monte Carlo simulations under these circumstances.

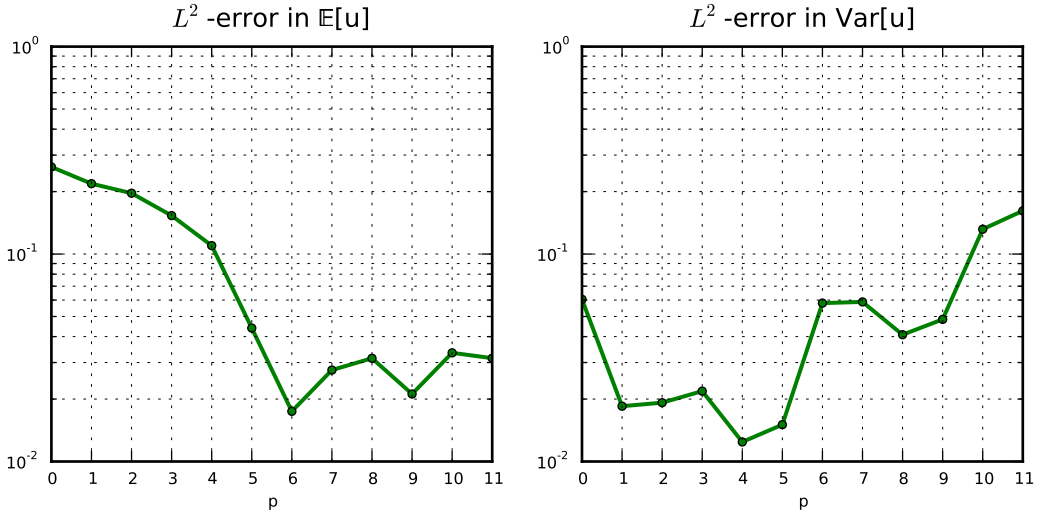

Figure 7.16: The L^2-errors in $\mathbb{E}[u]$ and in $\mathrm{Var}[u]$ for the last example as functions of p when the sGFEM solution is compared against an extensive Monte Carlo simulation. ($M = 1$, $N = 20$)

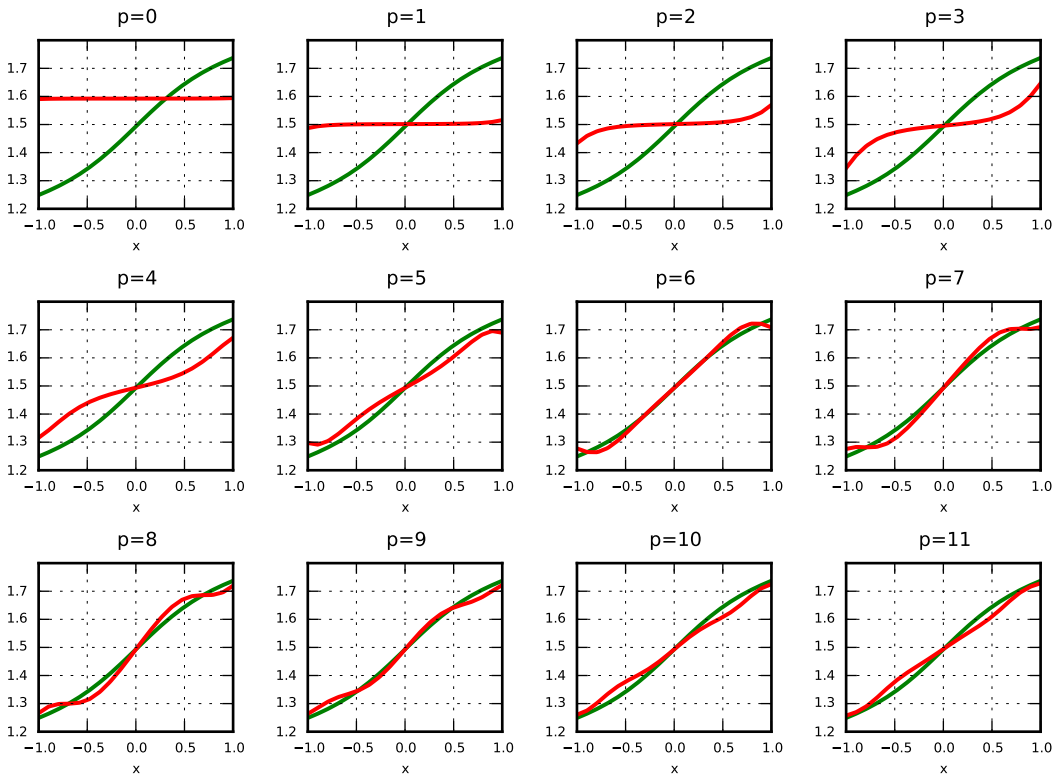

Figure 7.17: The sGFEM mean field (red) and the Monte Carlo solution (green) for the last example as functions of p.

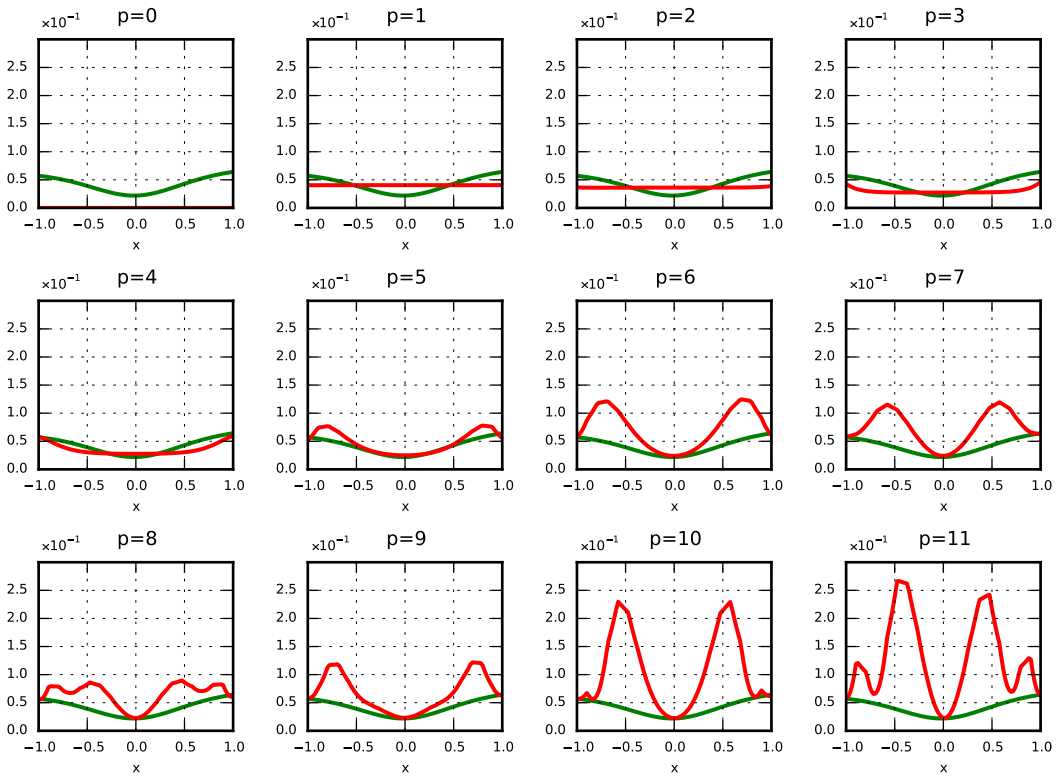

Figure 7.18: The sGFEM variance function (red) and the Monte Carlo solution (green) for the last example as functions of p.

Chapter 8

Conclusions

In this work, we introduced the stochastic Galerkin finite element method (sGFEM) and implemented it for the conductivity equation in one spatial dimension. More specifically, our model problem was a one-dimensional stochastic linear elliptic boundary value problem that can be considered (in the setting of electrical impedance tomography) as the stochastic conductivity equation reduced to one dimension. The conductivity coefficient of the model was assumed to be a log-normal random field with a known mean field and covariance function. The contact impedances of the model were also assumed to be stochastic and to follow known log-normal probability densities.

Even though we implemented the sGFEM only for the one-dimensional conductivity equation, we want to emphasize that the methods presented in this work can be readily applied to higher dimensional problems with only minor changes and/or adjustments required.

First, a brief review of the basic stochastic constructions and notations required in the work was given. After that, we continued by presenting a method to discretize random fields, namely, the Karhunen–Loève expansion, to numerically treat the random field describing the stochastic conductivity coefficient. While it was possible to use the Karhunen–Loève expansion to discretize the conductivity, it turned out that other techniques were needed to describe the random behavior of the solution potential. Thus, we were required to introduce the polynomial chaos expansion. Subsequently, we gave the precise formulation of the model problem and proved the corresponding variational formulation to be well-posed in the sense of unique solvability. Finally, the sGFEM was introduced and applied to the model problem. Numerical examples were given to illustrate the feasibility of the technique.

According to the numerical tests, the sGFEM approach was found to be a feasible choice when the variances of the initial stochastic parameters were not too large. Finding the exact reason for the failure of the method in the high variance case was left for future studies.

In order to use the sGFEM as a general numerical tool, the number of unknowns required by the method should be substantially reduced: in the current implemen-

tation the system matrix becomes rapidly unfeasibly large when the number of stochastic parameters or the maximum order of the polynomial chaos expansion is increased. Luckily, some algorithms have already been proposed to substantially reduce the number of unknowns; see [5], for example. The investigation of such methods is also left for future studies.

In the numerical tests, we did not address the running time of the stochastic Galerkin finite element method compared to other methods such as Monte Carlo based techniques, since for a fair comparison, we should have optimal implementations of the methods in hand. However, based on literature references, the stochastic finite element method should outperform the Monte Carlo methods especially in the case when there are many stochastic parameters in the model; see [5].

As was mentioned in the introduction, we are mainly interested in the two- or three-dimensional cases in real life applications involving the conductivity equation. Since the stochastic dimension of the sGFEM is added over a standard finite element solver, the transition to higher dimensional problems does not involve new mathematical theory in the stochastic dimension from what is already presented in this work. However, the introduction of a two- or three-dimensional finite element solver for the method in a higher dimensional case increases substantially the workload. Because of this, the implementation of the sGFEM in more practical settings is also left for future studies.

References

[1] Abrahamsen, P. A review of gaussian random fields and correlation functions, Apr 1997. Second edition.

[2] Ash, R. B., and Gardner, M. F. *Topics in Stochastic Processes*, vol. 27 of *Probability and mathematical statistics*. Academic press, 1975.

[3] Babuska, I., Nobile, F., and Tempone, R. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal. 45*, 3 (2010), 1005–1034.

[4] Bieri, M. *Sparse tensor discretizations of elliptic PDEs with random input data*. PhD thesis, ETH Zürich, 2009. Diss. ETH No. 18598.

[5] Bieri, M., Andreev, R., and Schwab, C. Sparse tensor discretization of elliptic sPDEs. *SIAM J. Sci. Comput. 31*, 6 (2009), 4281–4304.

[6] Bieri, M., and Schwab, C. Sparse high order fem for elliptic sPDEs. *Comput. Methods Appl. Mech. Engrg. 198* (2008), 1149–1170.

[7] Braess, D. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, second ed. Cambridge University Press, 2001.

[8] Bulirsch, R., and Stoer, J. *Introduction to Numerical Analysis*. New York: Springer-Verlag, 1991.

[9] Da Prato, G., and Zabczyk, J. *Second Order Partial Differential Equations in Hilbert Spaces*. Cambridge University Press, 2002.

[10] Dautray, R., and Lions, J.-L. *Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 2. Springer-Verlag Berlin Heidelberg New York, 1988.

[11] Ghanem, R. Ingredients for a general purpose stochastic finite elements implementation. *Comput. Methods Appl. Mech. Engrg. 168* (1999), 19–34.

[12] Ghanem, R. G., and Spanos, P. D. *Stochastic Finite Elements: A Spectral Approach*. Dover Publications, 2003.

[13] Gilbarg, D., and Trudinger, N. S. *Elliptic Partial Differential Equations of Second Order*, vol. 224 of *A Series of Comprehensive Studies in Mathematics*. Springer-Verlag Berlin Heidelberg New York, 1977.

[14] Gittelson, C. J. Stochastic galerkin discretization of the log-normal isotropic diffusion problem. *Mathematical Models and Methods in Applied Sciences 20*, 2 (2010), 237–263.

[15] Jacod, J., and Protter, P. *Probability Essentials*. Springer-Verlag Berlin Heidelberg New York, 2004.

[16] Keese, A. A review of recent developments in the numerical solution of stochastic partial differential equations (stochastic finite elements). Technical Report 2003-06, Institute of Scientific Computing, Technical University Braunschweig, Oct 2003.

[17] Schlather, M. Introduction to positive definite functions and to unconditional simulation of random fields. Technical Report ST-99-10, Department of Mathematics and Statistics, Lancaster University, UK, Jul 1999.

[18] Schwab, C., and Todor, R. A. Karhunen–Loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics 217* (2006), 100–122.

[19] Simon, B. The classical moment problem as a self-adjoint finite difference operator. *Advances in Math 137* (Jun 1998), 82–203.

[20] Xiu, D. *Numerical methods for stochastic computations: a spectral method approach*. Princeton University Press, 2010.

[21] Xiu, D., and Karniadakis, G. E. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput. 24*, 2 (2002), 619–644.

[22] Zaanen, A. C. *Linear Analysis*. North-Holland Publishing Co., 1960.