NUMERICAL STUDIES OF KORTEWEG-DE VRIES EQUATION WITH RANDOM INPUT DATA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET ALP ÜRETEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
SCIENTIFIC COMPUTING

SEPTEMBER 2018

Approval of the thesis:

NUMERICAL STUDIES OF KORTEWEG-DE VRIES EQUATION WITH RANDOM INPUT DATA

submitted by **MEHMET ALP ÜRETEN** in partial fulfillment of the requirements for the degree of **Master of Science in Scientific Computing Department, Middle East Technical University** by,

Prof. Dr. Ömür Uğur
Director, Graduate School of **Applied Mathematics** _____

Assist. Prof. Dr. Hamdullah Yücel
Head of Department, **Scientific Computing** _____

Assist. Prof. Dr. Hamdullah Yücel
Supervisor, **Scientific Computing, METU** _____

Prof. Dr. Ömür Uğur
Co-supervisor, **Scientific Computing, METU** _____

Examining Committee Members:

Assoc. Prof. Dr. Mustafa Yücel
Institute of Marine Sciences, METU _____

Assist. Prof. Dr. Hamdullah Yücel
Institute of Applied Mathematics, METU _____

Prof. Dr. Ömür Uğur
Institute of Applied Mathematics, METU _____

Assoc. Prof. Dr. Murat Uzunca
Department of Mathematics, Sinop University _____

Assist. Prof. Dr. Cüneyt Baykal
Department of Civil Engineering, METU _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: MEHMET ALP ÜRETEN

Signature :

ABSTRACT

NUMERICAL STUDIES OF KORTEWEG-DE VRIES EQUATION WITH RANDOM INPUT DATA

Üreten, Mehmet Alp
M.S., Department of Scientific Computing
Supervisor : Assist. Prof. Dr. Hamdullah Yücel
Co-Supervisor : Prof. Dr. Ömür Uğur

September 2018, 97 pages

Differential equations are the primary tool to mathematically model physical phenomena in industry and natural science and to gain knowledge about its features. Deterministic differential equations does not sufficiently model physically observed phenomena since there exist naturally inevitable uncertainties in nature. Employing random variables or processes as inputs or coefficients of the differential equations yields a stochastic differential equation which can clarify unnoticed features of physical events. Korteweg-de Vries (KdV) equation with the random input data is a fundamental differential equation for modeling and describing solitary waves occurring in nature. It can be represented by employing time dependent additive randomness into its forcing or space dependent multiplicative randomness into derivative of the solution. Since analytical solution of the differential equation with the random data input does not exist, quantifying and propagating uncertainty employed on the differential equation are done by numerical approximation techniques. This thesis will focus on numerical investigation of the Korteweg-de Vries equation with random input data by employing stochastic Galerkin in probability space, local discontinuous Galerkin method in spatial dimension, and theta (weighted average) method in temporal dimension. In numerical implementations, both additive noise and multiplicative noise cases are considered by comparing with other numerical techniques such as Monte Carlo and stochastic collocation methods for the probability space and finite differ-

ence method for the spatial discretization.

Keywords: uncertainty quantification, stochastic Galerkin method, Korteweg-de Vries equation, local discontinuous Galerkin method

ÖZ

RASTGELE GİRDİLERİ OLAN KORTEWEG-DE VRİES DENKLEMİNİN SAYISAL ÇALIŞMASI

Üreten, Mehmet Alp
Yüksek Lisans, Bilimsel Hesaplama Bölümü
Tez Yöneticisi : Dr. Öğr. Üyesi Hamdullah Yücel
Ortak Tez Yöneticisi : Prof. Dr. Ömür Uğur

Eylül 2018, 97 sayfa

Diferansiyel denklemler, endüstri ve doğa bilimlerindeki fiziksel fenomenleri matematiksel olarak modellemenin ve özellikleri hakkında bilgi edinmenin temel aracıdır. Deterministik diferansiyel denklemler, doğadaki kaçınılamaz belirsizliklerden dolayı fiziksel olarak gözlenen fenomenleri yeterince modelleyememektedir. Diferansiyel denklemlerin girdileri ya da katsayılarının rassal değişkenler olarak kullanılması, fiziksel olayların fark edilmeyen özelliklerini netleştirebilen bir stokastik diferansiyel denklemi ortaya çıkarır. Rassal girdili Korteweg-de Vries (KdV) denklemi doğada oluşan tekil dalgaların modellenmesi ve tanımlanması için temel bir diferansiyel denklemdir. Bu denklemler zamana bağlı toplama rassalığı veya uzaya bağlı çarpımsal rassallığı ile ifade edilebilir. Rassal girdili diferansiyel denklemlerin genellikle analitik çözümü bulunmadığından, diferansiyel denklem üzerinde kullanılan belirsizliğin nicelleştirilmesi ve yayılması sayısal yaklaşım teknikleriyle yapılır. Bu tez, olasılık uzayında stokastik Galerkin, mekansal boyutta yerel süreksiz Galerkin yöntemi ve zamansal boyutta teta (ağırlıklı ortalama) yöntemi kullanılarak Korteweg-de Vries denkleminin rassal girdi verileriyle sayısal olarak incelenmesine odaklanacaktır. Sayısal uygulamalarda, hem olasılıksal gürültü hem de çarpımsal gürültü durumları Monte Carlo ve stokastik kolokasyon gibi yöntemler ve uzaysal yaklaşım için sonlu farklar yöntemi gibi diğer sayısal tekniklerle karşılaştırılarak değerlendirilmektedir.

Anahtar Kelimeler: belirsizlik hesaplaması, stokastik Galerkin yöntemi, Korteweg-de Vries denklemi, yerel süreksiz Galerkin yöntemi

To my family

ACKNOWLEDGMENTS

I would like to express my absolute appreciation to my thesis supervisor Assist. Prof. Dr. Hamdullah Yücel. His patient guidance, encouragement strongly affected me during the development and preparation of this thesis. I also would to express my gratitude to Prof. Dr. Ömür Uğur for his enthusiasm and willingness to give his time and to share his experiences. I also want to thank Prof. Dr. Bülent Karasözen for introducing and suggesting me Scientific Computing Program during the senior year of my undergraduate degree.

I would also like to thank members of my thesis defense committee for their insightful comments and discussions.

I am grateful to colleagues, Abdullah Ali Sivas and Pelin Çiloğlu, for their great contributions to this study.

I am also thankful to everyone in the Institute of Applied Mathematics for providing pleasant and friendly environment.

Last but not least, I owe my most profound gratitude to my family for their continuous support through my entire life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Mathematical modeling with differential equations is the main tool in various fields of science and engineering for understanding crucial behaviors of complex physical phenomena. Since solving the differential equations analytically is quite impractical and sometimes impossible, scientific computing has become an essential tool to make numerical simulations of physical events efficiently and effectively. The primary goal of the scientific computing is to develop numerical techniques whose numerical errors are well established and to implement them to the problems induced by natural sciences. As the need of natural sciences grows considerably, understanding complex structure of interested models has become a necessity. Therefore, comprehension of how errors/uncertainties in model or data impact behaviors of the system have become a significant numerical tool in the past years. Uncertainty quantification (UQ) in the differential equations may uncover obscure features of the physical events.

The uncertainties can be classified into two different categories with respect to their sources [33]. The first one is related to inevitable random variability appearing in the nature, known as aleatory or statistical uncertainty. The second one, called epistemic uncertainty, arises from a lack of knowledge about the physical event involved and the assumptions made while modeling. Epistemic uncertainty can be reduced by making additional observations and by using improved measuring devices. Despite these processes augment on the knowledge about physical event, they are quite impractical to measure all features. These uncertainties can be studied as random inputs or operator of the mathematical model. In other words, the randomness can be introduced in the system by random variable or processes which depends on how sophisticated

uncertainties exist in the physical event. The most common problems interested in natural sciences and engineering mostly involve partial differential equations (PDEs) with random coefficients. Moreover, the random coefficients is classically defined as some idealized stochastic processes, which can be analyzed using elegant mathematical theory, such as Ito calculus [20]. In this case, the governing stochastic PDE has a solution as stochastic process which is generally not suitable to analytical solutions.

The treatment to stochastic equations consists of two steps, which are defining the properties of the stochastic system with random variables/processes having a proper probability distribution and solving corresponding differential equation to obtain their second order statistical information. In this thesis, our concern is mainly on the solving corresponding differential equation.

1.1 Model Equation

In this thesis, we will investigate the numerical solution of the Korteweg-de Vries (KdV) equation with random input data. Korteweg-de Vries equation is a mathematical model for solitary water waves with finite amplitudes, which was named after its discoverers Diederik Korteweg and Gustav de Vries [26]. John Scott Russel [38] made first experimental observation of such a solitary wave, then Boussinesq [5] developed first mathematical theory to support Russel's observations. In 1965, Zabusky and Kruskal's numerical study on the KdV equation revealed that solitary waves preserve their form in the collision with each other [56]. The KdV equation has been a primary tool to mathematically modeling and clarifying some events in the nature such as long internal waves in a ocean [41], ion acoustic waves in a plasma [25, 49]. However, deterministic KdV equation is not enough to clarify all the features of the physical events since uncertainties are inherently inevitable. Stochastic Korteweg-de Vries (sKdV) equation can be defined by presenting additive randomness in the forcing term or multiplicative randomness in the single derivative term of the KdV equation. Additive randomness in the KdV equation can be used to model traveling waves in noisy plasmas [47, 48], whereas KdV equation with multiplicative noise can be used to describe diffusive behavior of the solitons [19]. Moreover, dissipation and dispersion effects of uncertainties on the solitons are studied numerically in [23, 29].

1.2 Numerical Techniques

The widely used numerical techniques for quantifying uncertainty can be classically separated into two groups which are non-intrusive and intrusive methods [28]. While non-intrusive methods are based on generating identically independent samples of the random input with respect to their probability distribution, intrusive methods depend directly on spectral expansions of the solution on the predefined stochastic subspace.

One of the most popular class of the non-intrusive numerical techniques is Monte Carlo (MC) sampling methods [11] since it is straightforward to implement and can naturally be parallelized. The MC methods are built on generating pseudo-random samples of the random input of the system and on forming a set of realizations by solving corresponding deterministic system at each pseudo-random sample. Then, relevant statistical information can be derived from this set of realizations. The MC methods require only a recurrent deterministic solver, however, to obtain a conveniently small error in solution statistics takes a large number of samples to examine in this classical technique. The other popular non intrusive method is called Stochastic Collocation (SC) method [2, 52] which basically seeks an interpolation of the stochastic system output on the random sample space, where collocation points are chosen appropriately. The SC method requires only a deterministic solver and achieves fast convergence rate provided that stochastic outputs are sufficiently smooth on the random domain. However, the SC method suffers from the curse of dimensionality [53].

In this thesis, we mainly focus on one of the most powerful and recently developed intrusive approach called stochastic Galerkin method [14]. The formulation of this technique is a natural extension of the elementary idea behind Galerkin method in deterministic setting based on polynomial (homogeneous) chaos (PC) expansion. An important feature of this discretization technique for stochastic PDEs is the separation of the spatial and stochastic variables. This allows a reuse of established discretization techniques. The stochastic discretization, which is independent of the spatial one, results in a discretization of the range space of a collection of independent random variables, which are used for the representation of all random fields involved in the model. Stochastic Galerkin method inherits fast convergence from Galerkin method when the solution is smooth on its domain [28].

After the PDE with random input data is transformed into a deterministic setting, we need to discretize spatial and temporal dimensions by any appropriate discretization schemes. In this thesis, we use local discontinuous Galerkin (LDG) scheme [10] as a space discretization method. Local discontinuous Galerkin (LDG) method is one of several discontinuous Galerkin methods which are being vigorously studied, especially as applied to hyperbolic equations because of their applicability to a wide range of problems and their properties of local conservativity and higher degree of locality. The fundamental idea behind the LDG methods is to rewrite the differential equation with higher order derivatives into a first-order system. In the LDG method, the local conservativity holds because the conservation laws are weakly enforced element by element. In order to do that, suitable discrete approximations of the traces of the fluxes on the boundary elements are provided by the so-called numerical fluxes. These numerical fluxes enhance the stability of the method, and hence, the quality of its approximation. This is why the LDG method is strongly related to stabilized mixed finite elements.

On the other hand, the well-known weighted average (theta) method will be used for discretization of the temporal dimension. Moreover, the rational deferred correction method [17] will be introduced as a post-processing technique in order to enhance numerical solutions of time-dependent problem. This method can be considered as an extrapolation scheme which improves the accuracy of a low-order integrator iteratively. The idea is to construct representation of residual based on polynomial interpolation of the solution over the time interval.

The numerical techniques introduced above are displayed in the Figure 1.1.

1.3 Outline of Thesis

This thesis will focus on the implementation of Korteweg-de Vries equation having additive or multiplicative uncertainties in its input data. In Chapter 2 some basic concepts about polynomial approximation, probability and random field representation, which consists of Karhunen-Loève (KL) and polynomial chaos (PC) expansions, are outlined. Chapter 3 gives a brief introduction to deterministic/stochastic KdV equa-

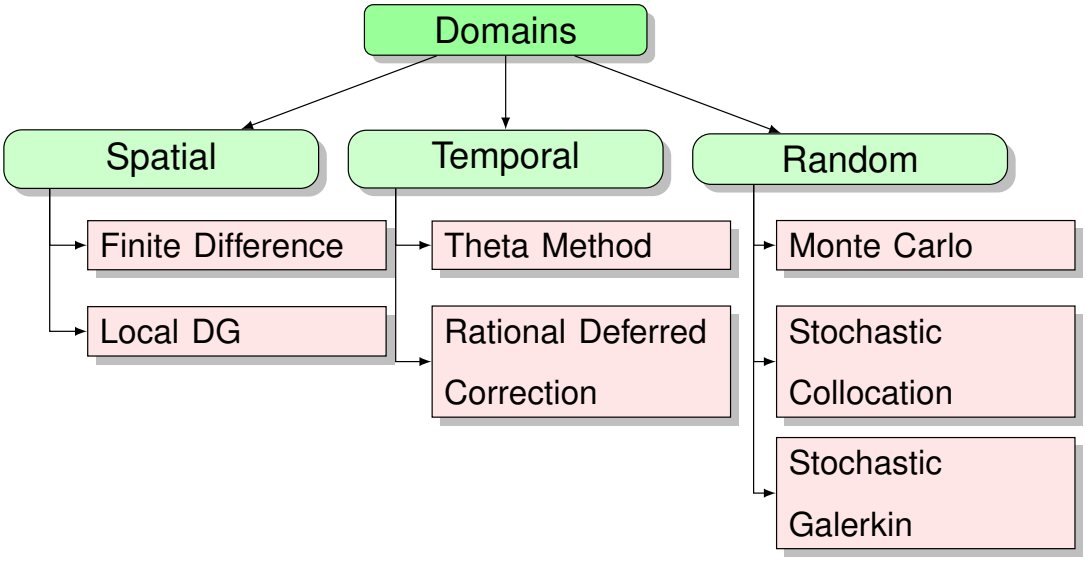

Figure 1.1: Numerical approaches.

tion and derivations of their exact solution if it exists. In Chapter 4, intrusive and non-intrusive methods are discussed, furthermore, Monte Carlo (MC) simulation, stochastic collocation (SC) and stochastic Galerkin (SG) methods are implemented to discretize the random domain of KdV equation with random input data. In Chapter 5, approximation of the KdV equation with random input data is completed by applying finite difference and local discontinuous Galerkin methods in spatial domain and the theta method in temporal domain. Moreover, some numerical examples are presented. Finally, the thesis ends with Chapter 6 which includes some discussions and prospective points for the future work.

CHAPTER 2

PRELIMINARIES

Uncertainty quantification (UQ) is a branch of scientific computing that quantifies uncertainties and characterizes the impacts of minor differences in both computational and real world systems. It tries to identify how likely outcomes occur if some features of the system are uncertain. The sources of uncertainty have distinct origins, which may be rooted in incomplete knowledge of numerical models and measurement errors. We refer to [14] and references therein for more information about the origin of uncertainty. Our starting point of numerical study will be a partial differential equation (PDE) formulation where input parameters, initial or boundary conditions are uncertain. Because of the nature of the uncertainty, random processes are used to characterize the uncertainties and a spectral series representation is then used to represent the solution to the problem of interest.

In this thesis, we focus on quantifying the uncertainty that propagates from the input of a mathematical model. Therefore, randomness is needed to be inserted into deterministic system. The fundamental elements of the probability theory and spectral methods help us in order to overcome this problem. In this chapter, we firstly introduce orthogonal polynomials that are used in spectral expansions and polynomial approximation theory. Then, the notion of randomness will be given for understanding in characterization of stochastic parameters presented in a numerical model. Lastly, basics of Karhunen-Loève expansion, well-known decomposition technique to represent processes as an expansion by linear combination of orthogonal functions, will be presented.

2.1 Orthogonal Polynomials

In this section, introductory definitions concerning the theory of orthogonal polynomials will be given. An interested reader can find further information on the subject in [7, 45] and references therein. First of all, a polynomial of degree n can be written in the general form:

$$Q_n(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0, \quad a_n \neq 0, \tag{2.1}$$

where a_n is called the leading coefficient. The monic version of this polynomial means that its leading coefficient equals to one. Therefore, monic form can be composed by dividing polynomial (2.1) with its leading coefficient a_n such that

$$P_n(x) = \frac{Q_n(x)}{a_n} = x^n + \frac{a_{n-1}}{a_n}x^{n-1} + \cdots + \frac{a_1}{a_n}x_1 + \frac{a_0}{a_n}, \quad a_n \neq 0.$$

A system of polynomials $\{Q_n(x), n \in \mathbb{N}_0\}$ is said to be an orthogonal system of polynomials with respect to some real positive measure α if it satisfies the following orthogonality relation

$$\int_{\mathcal{D}} Q_n(x)Q_m(x)d\alpha(x) = \gamma_n \delta_{mn}, \quad m, n \in \mathbb{N}_0, \tag{2.2}$$

where δ_{mn} is the Kronecker delta function, $\mathcal{D}$ is the support of the measure α, and γ_n is the normalization constant which obviously can be stated as

$$\gamma_n = \int_{\mathcal{D}} Q_n^2(x)d\alpha(x), \quad n \in \mathbb{N}_0.$$

A system of orthogonal polynomials is called orthonormal if normalization constant γ_n of every polynomial belonging to system equals to one. To normalize a set of orthogonal polynomials, every individual polynomial is divided by its corresponding normalization factor such that

$$\tilde{Q}_n(x) = \frac{Q_n(x)}{\sqrt{\gamma_n}}.$$

Here, if the measure $\alpha(x)$ on the set $\mathcal{D}$ is absolutely continuous, then the orthogonality relation (2.2) reduces to

$$\int_{\mathcal{D}} Q_n(x)Q_m(x)w(x)dx = \gamma_n \delta_{mn}, \quad m, n \in \mathbb{N}_0, \tag{2.3}$$

where $w(x)$ is a non-negative Lebesgue measurable function called as a weight function on $\mathcal{D}$. The weight function $w(x)$ is equal to zero outside of the support $\mathcal{D}$ and the total mass of the set $\mathcal{D}$ can be computed by $\int_{\mathcal{D}} w(x)dx$.

Knowing that $w(x)$ is the weight function on the support $\mathcal{D}$, a space of vectors defined by the set of real valued functions $f(x)$ belonging to the (weighted) Hilbert space L^2_w on $\mathcal{D}$ can be formulated as

$$L^2_w = \{f : \mathcal{D} \to \mathbb{R} \mid \int_D f^2(x)w(x)dx < \infty\}$$

and the corresponding norm is defined by the (weighted) inner product such that

$$\|f\|_w = \langle f, f\rangle_w^{1/2} = \left(\int_{\mathcal{D}} f^2(x)w(x)dx\right)^{1/2}.$$

As a result, the orthogonality relation (2.3) can be restated with the inner product as

$$\langle Q_n, Q_m\rangle_w = \int_{\mathcal{D}} Q_n(x)Q_m(x)w(x)dx = \gamma_n\delta_{mn}, \quad m, n \in \mathbb{N}_0,$$

where its normalization constant is

$$\gamma_n = \langle Q_n, Q_n\rangle_w = \|Q_n\|_w^2, \quad n \in \mathbb{N}_0.$$

2.1.1 Three-Term Recurrence Relation

A well known three-term recurrence relation can be formulated for system of orthogonal polynomials $\{Q_n(x)\}$ with $x \in \mathbb{R}$ by stating that

$$-xQ_n(x) = b_nQ_{n+1}(x) + a_nQ_n(x) + c_nQ_{n-1}(x), \quad n \geq 1,$$

where $b_n,\ c_n \neq 0$, and $\frac{c_n}{b_{n-1}} > 0$ as defined in [52]. The three-term recurrence relation is often rewritten in different equivalent form for system of monic orthogonal polynomials which is referred as Favard's theorem [7].

Theorem 2.1. *[7, Theorem 4.4] Let $\{P_n(x)\}$ be a system of monic orthogonal polynomials. For $n \geq 1$ we have a recurrence relation*

$$P_{n+1}(x) = (x - a_n)P_n(x) - b_nP_{n-1}(x)$$

together with $P_0(x) = 1$ *and* $P_{-1}(x) = 0$*, where* $\{a_n\}$*,* $\{b_n\}$ *are sequences of real numbers such that*

$$a_n = \frac{\langle xP_n, P_n\rangle}{\langle P_n, P_n\rangle}, \quad b_n = \frac{\langle P_n, P_n\rangle}{\langle P_{n-1}, P_{n-1}\rangle}.$$

In the implementation of the stochastic Galerkin method, the most frequently used classes of orthogonal polynomials are members of the Askey family of hypergeometric orthogonal polynomials [54]. Now, we will introduce the well-known members of the Askey family, which are Hermite and Legendre polynomials.

2.1.2 Hermite Polynomials

An important example of hypergeometric orthogonal polynomials is the class of Hermite polynomials. In the literature, there exist two widely used definitions of the Hermite polynomials defined on real line, i.e., $x \in \mathbb{R}$. The one which is mostly used in polynomial chaos expansion in order to represent the random quantities is denoted by $\{H_n(x)\}$ and the second one is denoted by $\{He_n(x)\}$.

The Hermite polynomials $H_n(x)$ of degree n are given as defined in [28] with its explicit representation such that

$$H_n(x) = \frac{1}{(-1)^n\, e^{-\frac{x^2}{2}}}\, \frac{d^n}{dx^n}\left[e^{\frac{-x^2}{2}}\right] = n! \sum_{k=0}^{\lfloor n/2\rfloor} (-1)^k\, \frac{1}{k!2^k(n-2k)!} x^{n-2k},$$

where $\lfloor n\rfloor$ denotes the greatest integer function. The three-term recurrence relation for polynomials $H_n(x)$ is

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x), \quad n > 0,$$

together with $H_0(x) = 1$ and $H_1(x) = x$. Moreover, they satisfy the following orthogonality relation

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)w(x)dx = n!\delta_{mn}, \quad n \in \mathbb{N}_0,$$

where $w(x)$ is the weight function for $H_n(x)$, that is,

$$w(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

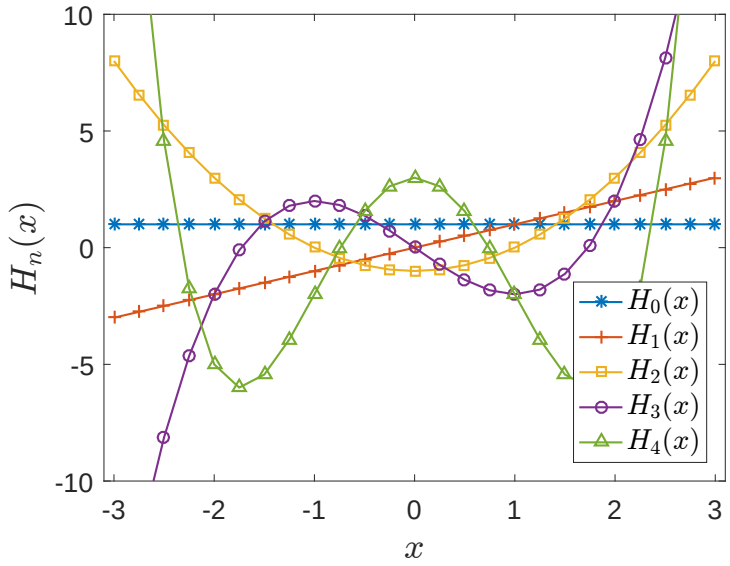

Figure 2.1: The first five Hermite polynomials, i.e, $H_n(x)$, $n = 0, \ldots, 4$.

Further, the first few Hermite polynomials $H_n(x)$, plotted in Figure 2.4, are

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= x, \\ H_2(x) &= x^2 - 1, \\ H_3(x) &= x^3 - 3x, \\ H_4(x) &= x^4 - 6x^2 + 3. \end{aligned} \tag{2.4}$$

The second definition of Hermite polynomials $He_n(x)$ of degree n has the following explicit representation:

$$He_n(x) = \frac{1}{(-1)^n e^{-x^2}} \frac{d^n}{dx^n} \left[e^{-x^2} \right] = n! \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \frac{1}{k!(n-2k)!} (2x)^{n-2k}.$$

The three-term recurrence relation for the polynomials $He_n(x)$ is formulated with

$$He_{n+1}(x) = 2xHe_n(x) - 2nHe_{n-1}(x), \quad n > 0$$

and they satisfy the following relation:

$$\int_{-\infty}^{\infty} He_n(x) He_m(x) \widetilde{w}(x) dx = 2^n n! \delta_{mn}, \quad n \in \mathbb{N},$$

where $\widetilde{w}(x)$ is the weight function for $He_n(x)$ defined as

$$\widetilde{w}(x) = \frac{1}{\sqrt{\pi}} e^{-x^2/2}.$$

Further, the first five Hermite polynomials $He_n(x)$ are

$$\begin{aligned} He_0(x) &= 1, \\ He_1(x) &= 2x, \\ He_2(x) &= 4x^2 - 2, \\ He_3(x) &= 8x^3 - 12x, \\ He_4(x) &= 16x^4 - 48x^2 + 12. \end{aligned} \tag{2.5}$$

2.1.3 Legendre Polynomials

Another frequently used member of the hypergeometric orthogonal polynomials is the class of Legendre polynomials. The class of Legendre polynomials $\{L_n(x)\}$ constructs an orthogonal basis for L^2_w on the interval $[-1, 1]$ with respect to the weight function $w(x) = 1/2$. The normalized Legendre polynomials $L_n(x)$ of degree n can be formulated as

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}[(x^2 - 1)^n] = \frac{1}{2^n} \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n}{k} \binom{2n - 2k}{n} x^{n-2k}.$$

Let first two of Legendre polynomials be $L_0(x) = 1$ and $L_1(x) = x$, then Legendre polynomials $L_n(x)$ satisfy the following three-term recurrence relation:

$$L_{n+1} = \frac{2n + 1}{n + 1} x L_n(x) - \frac{n}{n + 1} L_{n-1}(x), \quad n > 0.$$

The orthogonality relation with Legendre polynomials can be represented as

$$\int_{-1}^{1} L_n(x) L_m(x) w(x) dx = \frac{1}{2n + 1} \delta_{mn}, \quad n \in \mathbb{N}_0,$$

where the weight function in the orthogonality relation is a constant, i.e., $w(x) = 1/2$.

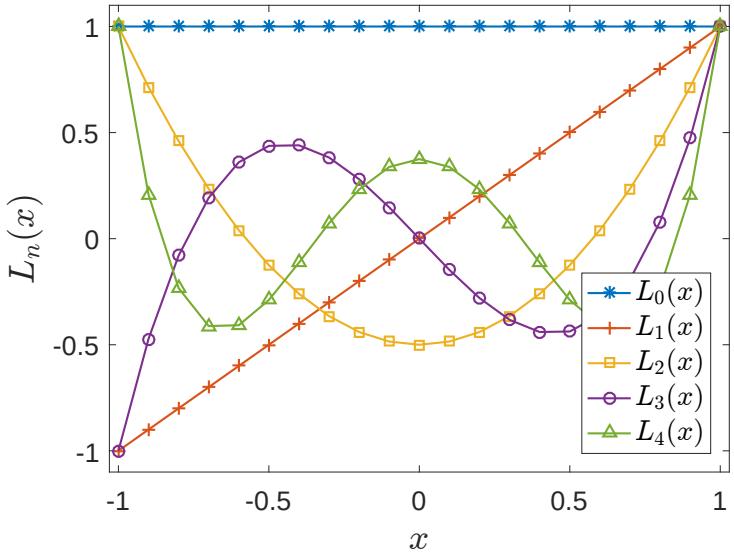

Figure 2.2: The first five Legendre polynomials, i.e, $L_n(x)$, $n = 0, \ldots, 5$.

The first few Legendre polynomials, plotted in Figure 2.2, are

$$\begin{aligned} L_0(x) &= 1, \\ L_1(x) &= x, \\ L_2(x) &= \frac{3}{2}x^2 - \frac{1}{2}, \\ L_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x, \\ L_4(x) &= \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}. \end{aligned}$$

It is noted that class of Legendre polynomials is a specific case of Jacobi polynomials $P_n^{\alpha,\beta}(x)$ with parameters $\alpha = \beta = 0$. Jacobi polynomials $\{P^{\alpha,\beta}(x)\}$ form a general class of orthogonal polynomials with respect to the weight function $w(x) = (1 - x)^\alpha (1 + x)^\beta$ on the interval $[-1, 1]$.

2.2 Polynomial Approximation

Approximation theory is concerned with approximating complex functions by using elementary functions and characterizing the approximation errors. Since polynomials are simple functions, it is natural to demand using them in approximations. The Weierstrass approximation theorem generalizes theoretically that polynomials can be

used for approximating continuous functions.

Theorem 2.2. *[46, Theorem 1.1] Let f be a continuous real-valued function defined on the bounded and closed interval $\mathcal{D}$. Then for any $\epsilon > 0$, there exists a polynomial $p \in \mathbb{P}_n$ and $n \in \mathbb{N}_0$ such that*

$$|f(x) - p(x)| < \epsilon, \quad \forall x \in \mathcal{D}.$$

One of the theoretical and practical outcome of the Weierstrass Theorem is that interpolation can be used as an approximation tool.

2.2.1 Interpolation

Interpolation is simply fitting some function f to given data set so that the function f satisfies the same given values at the given nodes. In one dimensional case, the general interpolation problem can be formed as seeking a function f satisfying

$$f(x_i) = y_i, \quad i = 1, \ldots, n$$

for given data set (x_i, y_i), $i = 1, \ldots, n$. The function f is called an interpolating function for the given data set.

The interpolating function is required to be a component of a function space that is spanned by basis functions $\{\phi_1, \phi_2, \ldots \phi_n\}$. Then, the interpolating function f is chosen to be a linear combination of these basis functions

$$f(x) = \sum_{j=1}^{n} c_j \phi_j(x),$$

where the parameters c_j are needed to be determined. Moreover, the function f has to satisfy the following system of equations

$$f(x_i) = \sum_{j=1}^{n} c_j \phi_j(x_i) = y_i, \quad i = 1, \ldots, n,$$

which can be expressed particularly as

$$\mathbf{A}\mathbf{c} = \mathbf{y},$$

where the entries of the matrix $\mathbf{A}$ are $A_{ij} = \phi_j(x_i)$, the vector of the known data values is $\mathbf{y} = (y_1, \ldots, y_n)^T$, and the vector of parameters to be determined is $\mathbf{c} = (c_1, \ldots, c_n)^T$.

One of the basic choice for basis function in interpolation is the monomial polynomials and they generate the space of polynomials, i.e., $\mathbb{P}_n = \text{span}\{1, x, x^2, \ldots, x^n\}$. For n data points (x_i, y_i), $i = 1, \ldots, n$, the interpolating polynomial with monomial basis

$$\phi_j(x) = x^{j-1}, \quad j = 1, \ldots, n$$

can be generally expressed as

$$P_{n-1}(x) = \sum_{j=1}^{n} c_j \phi_j(x) = c_1 + c_2 x + \ldots + c_n x^{n-1}.$$

Then, the suitable $n \times n$ linear system $\mathbf{Ac} = \mathbf{y}$ can be constructed such that

$$\begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

where the matrix $\mathbf{A}$ is called a Vandermonde matrix.

Another common choice for interpolation basis is the Lagrange polynomials. The Lagrange basis function are determined by

$$l_j(x) = \prod_{k=1, k \neq j}^{n} \frac{(x - x_k)}{(x_j - x_k)}.$$

More clearly, the basis functions are equivalent to

$$l_j(x_i) = \begin{cases} 1, & \text{if } \ i = j, \\ 0, & \text{if } \ i \neq j, \end{cases}$$

for the set of data points (x_i, y_i), $i = 1, \ldots, n$. The Lagrange interpolating polynomials can be given by

$$P_{n-1}(x) = \sum_{j=1}^{n} c_j l_j(x) = c_1 l_1(x) + c_2 l_2(x) + \ldots + c_n l_n(x),$$

thus, the matrix $\mathbf{A}$ of the corresponding linear system $\mathbf{Ac} = \mathbf{y}$ is the $n \times n$ identity matrix.

2.2.2 Orthogonal Projection

Any class of orthogonal polynomials can be used as basis functions for interpolation. Let $\{\phi_k(x),\ k \in \mathbb{N}_0\}$ be a set of orthogonal polynomials of degree at most n with respect to the positive weight $w(x)$ on the support $\mathcal{D}$. Interpolating polynomial $P_n f(x)$ for approximating a function $f(x)$ with given a set of data $\{x_i, f(x_i)\}, i = 0, \ldots, n$ can be similarly stated by

$$f(x) \approx P_n f(x) = \sum_{j=0}^{n} c_j \phi_j(x) = c_0\phi_0(x) + c_1\phi_1(x) + \ldots + c_n\phi_n(x).$$

Since sequence of orthogonal polynomials $\{\phi_k(x),\ k \in \mathbb{N}_0\}$ forms a basis for the weighted Hilbert space L^2_w on the support $\mathcal{D}$, the function $f \in L^2_w$ can be projected onto the space of polynomials of degree at most n, i.e., $\mathbb{P}_n$, with the help of the associated inner product such that

$$\begin{aligned}\langle f(x), \phi_k(x)\rangle &= \sum_{j=0}^{n} c_j \langle \phi_j(x), \phi_k(x)\rangle \\ &= c_k \langle \phi_k(x), \phi_k(x)\rangle, \quad k = 0, 1, \ldots, n.\end{aligned}$$

Thus, the coefficients c_k can be computed as

$$c_k = \frac{\langle f, \phi_k\rangle}{\langle \phi_k, \phi_k\rangle} = \frac{1}{\gamma_k}\int_{\mathcal{D}} f(x)\phi_k w(x)dx, \quad k = 0, 1, \ldots, n.$$

The coefficients $\{c_k\}$ are the generalized Fourier coefficients [12]. The interpolating polynomial $P_n f \colon L^2_w(\mathcal{D}) \to \mathbb{P}_n$ with the generalized Fourier coefficients is called orthogonal projection of function $f(x)$ onto $\mathbb{P}_n$ of L^2_w. Then, we have the following result.

Theorem 2.3. *[46, Theorem 2.1] For any $f \in L^2_w$ on a support $\mathcal{D}$ and any $n \in \mathbb{N}_0$, the orthogonal projection $P_n f$ of f is the best approximation in the associated L^2-norm such that*

$$\|f - P_n f\|_{L^2_w} = \inf_{\psi \in \mathbb{P}_n} \|f - \psi\|_{L^2_w}.$$

2.2.3 Gaussian Quadrature

Numerical quadrature is an approximation of definite integral of a function with a weighted sum. The integral we wish to compute can be written in the form of a n-point quadrature formula with $n \geq 1$

$$\int_{\mathcal{D}} f(x)d\alpha(x) = \int_{\mathcal{D}} f(x)w(x)dx \approx \sum_{i=1}^{n} f(x_i)w_i, \tag{2.6}$$

where x_i's are nodes, w_i's are integration weights, $i = 1, \ldots, n$. The general quadrature problem consists of finding weights w_i that satisfy the quadrature formula (2.6) when the nodes x_i are fixed.

Gaussian quadrature rules are based on the polynomial interpolation. Let $x_i \in \mathcal{D}$, $i = 1, \ldots, n$ be the nodes and l_i be the n-th degree Lagrange polynomials through the node x_i. Given an interpolating polynomial $\hat{f}$ of a function $f(x)$, then the definite integral can be approximated by a n-point Gaussian quadrature formula

$$\begin{aligned} \int_{\mathcal{D}} f(x)w(x)dx &\approx \int_{\mathcal{D}} \hat{f}w(x)dx \\ &= \int_{\mathcal{D}} \left[\sum_{i=1}^{n} f(x_i)l_i(x)\right] w(x)dx \\ &= \sum_{i=1}^{n} f(x_i) \int_{\mathcal{D}} l_i(x)w(x)dx. \end{aligned} \tag{2.7}$$

Thus, the weights w_i of the quadrature formula can be computed with

$$w_i = \int_{\mathcal{D}} l_i(x)w(x)dx, \quad i = 1, \ldots, n.$$

Remark 2.1. If the quadrature nodes x_i are chosen to be the zeros $\{z_k\}_{k=1}^{n}$ of orthogonal polynomial $\{Q_n, n \in \mathbb{N}\}$ with respect to $w(x)$ of degree n, then the Gaussian quadrature formula (2.7) is exact for approximating any polynomial $f(x) \in \mathbb{P}_n$ of degree less than $2n$ [12].

2.3 Basic Concept of Probability Theory

In this section, a brief introduction to some basic concepts of probability theory which are used in uncertainty quantification will be presented. Some essential definitions

about distributions and their statistics have to be introduced since stochastic parameters within equations will be investigated. Moreover, we would like to refer [30, 31] for detailed descriptions of the probability and measure theory and the references therein.

2.3.1 Random Variable and Probability

The uncertainty quantification discussed in this thesis generally involves random coefficients in the model problem. In other words, the stochastic system to be solved is formed by combining a deterministic system with stochastic input parameters. Mathematical treatment of this kind of systems requires a characterization of the stochastic parameters.

The stochastic parameters of a mathematical model can be classified as random variables. A random variable is a quantity on a set of random outcomes of an experiment. Let Ω be an abstract outcome space and $\omega \in \Omega$ be an outcome, a mapping from Ω to real space $\mathbb{R}$, i.e., $X(\omega) : \Omega \to \mathbb{R}$, is called a random variable. Moreover, it can be generalized as a random field $\alpha(i, \omega)$, which is a collection of random variables indexed by a parameter $i \in I \subset \mathbb{R}^n$. The random fields can often be analyzed by its properties such as expectation, variance, covariance, etc.

In order to study with random variables, it is insightful to collect a set of events, which are subsets of the outcome space. The relevant collection of events is called σ-algebra and denoted by Σ.

Definition 2.1. The collection Σ is called σ- algebra on Ω if the following conditions are satisfied

- Σ is not empty; $\emptyset \in \Sigma$ and $\Omega \in \Sigma$.
- If $A \in \Sigma$, then $A^c \in \Sigma$, where A^c is the complement of A.
- If $A_1, A_2, \ldots \in \Sigma$, then

$$\bigcup_{i=1}^{\infty} A_i \in \Sigma \quad \text{and} \quad \bigcap_{i=1}^{\infty} A_i \in \Sigma.$$

Two fundamental examples of σ-algebra on the space Ω are

$$\Sigma_1 = \{\emptyset, \omega\},$$
$$\Sigma_2 = 2^{\Omega} \equiv \{A : A \subset \Omega\},$$

where Σ_1 is the smallest and Σ_2 is the largest σ-algebra on the set Ω. Σ_2 contains all subsets of Ω, and therefore is called power set of Ω [52].

The other essential concept is the probability which is used to measure likelihood of occurrence of certain events. The probability measure can be defined as:

Definition 2.2. For σ-algebra Σ on a countable event space Ω, P is a probability measure if

- $0 \leq P(A) \leq 1, \ \forall A \in \Sigma$,
- $P(\Omega) = 1$,
- For $A_1, A_2, \ldots \in \Sigma$ and $A_i \cap A_j = \emptyset, \forall i, \neq j$ the following equality holds

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

Some elementary properties of the probability measures are outlined as; for events $A, B \in \Sigma$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$
$$P(A^c) = 1 - P(A), \quad P(\emptyset) = 0,$$

and

$$P(A) \leq P(B) \ \text{ for } \ A \subseteq B.$$

The outcome space Ω, σ-algebra Σ, and probability P constitute the triplet (Ω, Σ, P), which is called the probability space.

2.3.2 Distributions and Statistics

Basic definitions about distributions and statistics are need to be presented in order to understand and work on the stochastic problems accounted in this thesis. First of all,

a random variable X yields a collection of probabilities, which is called (cumulative) distribution function.

Definition 2.3. The collection of the probabilities defined as

$$F_X(x) = P(X \leq x) = P(\{\omega : X(\omega) \leq x\}), \quad x \in \mathbb{R}$$

is called as the distribution function F_X of X.

The cumulative distribution function holds that $0 \leq F_X \leq 1$ due to the definition of the probability. Moreover, it can be used to compute the probability of a random variable that belongs to any particular interval. That is, the probability of random variable X belongs to an interval $(a, b] \subset \mathbb{R}$ can be obtained by

$$P(\{\omega : a < X(\omega) \leq b\}) = F_X(b) - F_X(a).$$

2.3.2.1 Continuous Distribution

A continuous random variable maps infinitely many values. Probability of a continuous random variable in any particular value is 0, that is,

$$\forall x \in \mathbb{R}, \quad P(X = x) = 0.$$

Probability density function f_X of a continuous random variable interprets the probability value at any single point in the sample space. Furthermore, continuous distribution function F_X is defined with a probability density function f_X such that

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy, \quad x \in \mathbb{R},$$

where the probability density function f_X satisfies

$$\forall x \in \mathbb{R},\ f_X(x) \geq 0, \quad \int_{-\infty}^{\infty} f_X(y)dy = 1.$$

Some of important characteristics of a continuous random variable are its expectation,

variance, and mth moment which are

$$\mu_X = \mathbb{E}[X] = \int_{-\infty}^{\infty} x\ f_X(x)dx,$$
$$\sigma_X^2 = \mathrm{Var}[X] = \int_{-\infty}^{\infty} (x-\mu_X)^2\ f_X(x)dx,$$
$$\mathbb{E}[X^m] = \int_{-\infty}^{\infty} x^m\ f_X(x)dx, \quad m \in \mathbb{N}_0,$$

respectively, where f_X is the density function. For a real-valued function g, the expectation of $g(X)$ is defined in a similar way

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)\ f_X(x)dx.$$

The two of important continuous distributions are uniform distribution and normal distribution because of their powerful characteristics.

The uniform distribution is a probability distribution of events each of which has equal probability within a distribution support. The probability density function (PDF) of the uniform distribution $U(a,b)$, on the interval $[a,b]$, is defined as

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b], \\ 0, & \text{otherwise}, \end{cases}$$

and its cumulative distribution function is

$$F_X(x) = \begin{cases} 0, & x \in (-\infty, a), \\ \frac{x-a}{b-a}, & x \in [a,b), \\ 1, & x \in [b, \infty). \end{cases}$$

The PDF and CDF of the uniform distribution with support $[-1,1]$ are demonstrated in Figure 2.3.

On the other hand, the normal distribution, also called Gaussian distribution, is a probability distribution which is often used to represent random variables with unknown distributions by well-known result the central limit theorem [21]. The probability distribution function (PDF) of the normal distribution $\mathcal{N}(\mu,\sigma)$, where $\mu \in \mathbb{R}$ is the

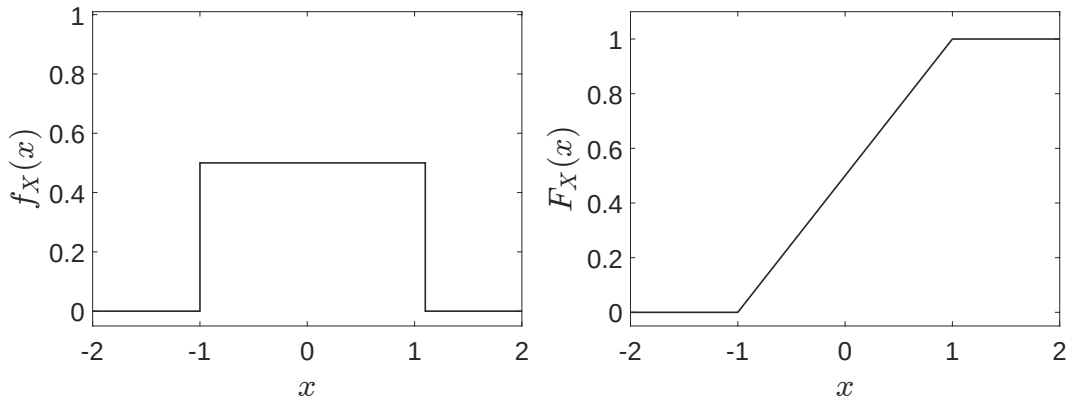

Figure 2.3: Probability distribution function (left) and cumulative distribution function (right) of the uniform distribution with $a = -1$, $b = 1$.

mean value, $\sigma^2 > 0$ is the variance, and σ is the standard deviation is defined as

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

The distribution is called as standard normal distribution if the parameters are $\mu = 0$ and $\sigma^2 = 1$, i.e., $\mathcal{N}(0, 1)$. The PDF and CDF of the standard normal distribution are displayed in Figure 2.4.

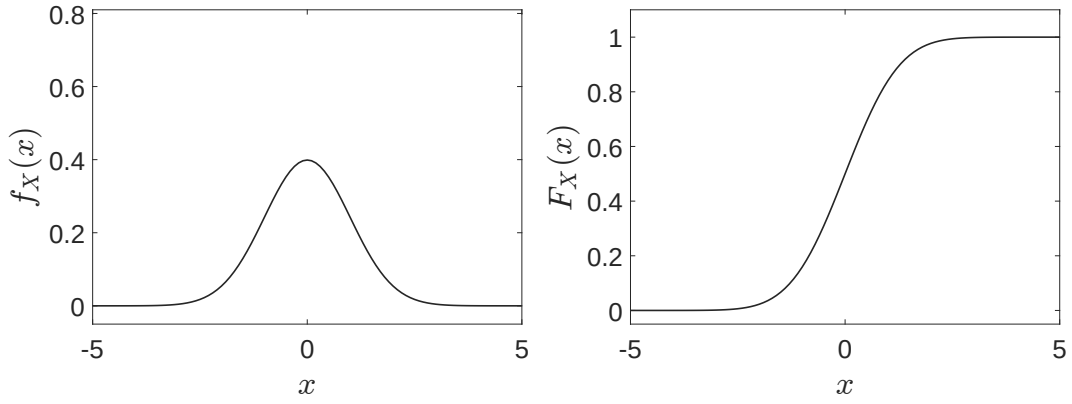

Figure 2.4: Probability distribution function (left) and cumulative distribution function (right) of the standard normal distribution with $\mu = 0$, $\sigma = 1$.

Remark 2.2 (Discrete Distribution). The set of outcomes is either finite or countably infinite. Unlike a continuous distribution, an experiment having a finite set of possible outcomes such as coin toss or roll of a dice is characterized by a discrete distribution. A discrete random variable maps only finite or countably infinite number of values

x_k with probabilities, $p_k = P(X = x_k)$, where $k = 1, \ldots, N$. Then, corresponding discrete distribution function can be defined as

$$F_X(x) = \sum_{k:x_k \leq x} p_k, \quad x \in \mathbb{R},$$

where

$$\forall k,\ 0 \leq p_k \leq 1, \quad \sum_{k=1}^{N} p_k = 1.$$

A discrete random variable also has important characteristics including its expectation, variance, and moments. For a discrete random variable X with probabilities p_k, $k = 0, \ldots, N$, $m \in \mathbb{N}$, the expectation, variance, and mth moment are defined as

$$\mu_X = \mathbb{E}[X] = \sum_{k=1}^{N} x_k\ p_k,$$

$$\sigma_X^2 = \mathrm{Var}[X] = \sum_{k=1}^{N} (x_k - \mu_X)^2\ p_k,$$

$$\mathbb{E}[X^m] = \sum_{k=1}^{N} x_k^m\ p_k,$$

respectively.

2.3.2.2 Multiple Dimensions

In multiple dimensions, random vectors can be defined as a collection of a finite number of random variables. Consequently, an n-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_n)$ is a vector whose components $X_1, \ldots, X_n$ are one-dimensional random variables. In the case of n-dimensional random vector, definition of a distribution function can be extended similarly as done in [52].

Definition 2.4. The collection of the probabilities defined as

$$F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$$

is called distribution function $F_{\mathbf{X}}$ of a random vector $\mathbf{X}$.

If the distribution of a random vector $\mathbf{X}$ has a joint probability density function $f_{\mathbf{X}}$, then distribution function $F_{\mathbf{X}}$ can be represented as

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_{\mathbf{X}}(y_1, \dots, y_n) dy_1 \cdots dy_n,$$

where the density of $\mathbf{X}$ satisfies

$$f_{\mathbf{X}}(\mathbf{x}) \geq 0, \quad \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$$

and

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\mathbf{X}}(y_1, \dots, y_n) dy_1 \dots dy_n = 1.$$

If a random vector $\mathbf{X}$ has density $f_{\mathbf{X}}$, then any component X_i has also a density function defined as

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(y_1, \dots, y_n) dy_1 \cdots dy_{i-1} dy_{i+1} \cdots dy_n,$$

which is called marginal density function of X_i. This definition can also be extended to vector of pairs (X_i, X_j), triples (X_i, X_j, X_k), and so on.

Random variables $X_1, X_2, \dots, X_n$ are said to be independent if their joint distribution function $F_{\mathbf{X}}$ can be written as a product of their marginal distribution functions

$$F_{\mathbf{X}}(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n).$$

Furthermore, expectation of a random vector $\mathbf{X}$ naturally is given by

$$\mu_{\mathbf{X}} = \mathbb{E}[\mathbf{X}] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_n])$$

and covariance between components X_i and X_j is defined as

$$\begin{aligned} \operatorname{cov}(X_i, X_j) &= \mathbb{E}[(X_i - \mu_{X_i})(X_j - \mu_{X_j})] \\ &= \mathbb{E}[X_i X_j] - \mu_{X_i} \mu_{X_j}. \end{aligned}$$

Since the variance of a random variable measures its variability, the covariance of two random variable measures their joint variability. The collection of covariances constructs the covariance matrix of random vector $\mathbf{X}$ such that

$$[\mathbf{C}_{\mathbf{X}}]_{ij} = \operatorname{cov}(X_i, X_j), \quad i, j = 1, \dots, n.$$

The covariance matrix generalizes the concept of covariance to multiple dimensions. Moreover, a covariance matrix has a basic but important property that it is symmetric and positive semidefinite [21]. It is also practical to define a statistical relationship between two or more random variables, called correlation coefficient, such that

$$\mathrm{corr}(X_1, X_2) = \frac{\mathrm{cov}(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}}.$$

The two random variables X_1, X_2 are called uncorrelated if $\mathrm{corr}(X_1, X_2) = 0$, and strongly correlated if $|\mathrm{corr}(X_1, X_2)| \approx 1$. It is also significant to know that if X_1 and X_2 have normal (Gaussian) distribution, then uncorrelated random variables are also independent [21].

2.3.3 Convergence

The collection of real valued random variables X defined on the probability space (Ω, Σ, P) such that

$$L^2(\Omega, \Sigma, P) = \{X : \mathbb{E}[X^2] < \infty\}$$

forms a vector space, called L^2 space [28]. In other words, for a random variable $X \in L^2(\Omega, \Sigma, P)$ it holds that

$$\mathbb{E}[X^2] = \int_{\mathcal{D}} X^2(s) dF_X(s) = \int_{\mathcal{D}} X^2(s) f_X(s) ds < \infty,$$

where $\mathcal{D}$ is the support of the probability measure. Furthermore, the associated norm $||X||_{L^2}$ is

$$||X||_{L^2} = \mathbb{E}[X^2]^{\frac{1}{2}} = \left(\int_{\mathcal{D}} X^2(s) f_X(s) ds \right)^{\frac{1}{2}} \tag{2.8}$$

and the expectation $\mathbb{E}[XY]$ defines an inner product on $L^2(\Omega, \Sigma, P)$ such that

$$\langle X, Y \rangle = \mathbb{E}[XY] = \int_{\mathcal{D}} X(s) Y(s) f_X(s) ds, \tag{2.9}$$

where $X, Y \in L^2(\Omega, \Sigma, P)$.

While working with the collection of stochastic variables, it is useful to introduce definitions of convergence for a sequence of random variables. These convergence definitions are based on the ones in [21].

Definition 2.5. Let $p > 0$, the sequence $\{X_n\}$ is said to converge in L^p to a random variable X, written as $X_n \xrightarrow{L^p} X$, if it holds that

$$\mathbb{E}[|X_n - X|^p] \to 0 \quad \text{as} \quad n \to \infty,$$

where $X, X_n \in L^p$ for all $n \in \mathbb{N}$.

Note that for the case $p = 2$ it is said that X_n converges to X in a mean square. It is a convergence in the Hilbert space which is $L^2 = L^2(\Omega, \Sigma, P)$ space equipped with the norm (2.8) and the associated inner product (2.9).

Definition 2.6. The sequence $\{X_n\}$ is said to converge in distribution to a random variable X, written as $X_n \xrightarrow{d} X$, if for all bounded and continuous functions f, it holds

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \quad \text{as} \quad n \to \infty.$$

The convergence in distribution holds if and only if there is a convergence in the distribution function. In other words, for all continuous points $x \in \mathcal{D}$ the relation

$$F_{X_n}(x) \to F_X(x) \quad \text{as} \quad n \to \infty$$

is satisfied. Convergence in distribution are often referred to as a weak convergence, whereas L^p convergence, namely mean square convergence, are referred to as a strong convergence.

2.3.4 Central Limit Theorem

Central limit theorem (CLT) plays a crucial role in many applications of probability theory and it is the main idea behind some numerical tools such as the Monte Carlo sampling, which will be introduced in Section 4.1.1.

Theorem 2.4. *[21, Theorem 21.1] Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (i.i.d.) random variables with their expectation $\mathbb{E}[X_i] = \mu$ and variance $\mathrm{Var}[X_i] = \sigma^2$. Then, the cumulative distribution function of*

$$Y = \frac{1}{n} \sum_{i=1}^{n} X_i$$

will converge to a Gaussian distribution $\mathcal{N}(\mu, \frac{\sigma}{\sqrt{n}})$ *as* $n \to \infty$. *Furthermore, distribution function of*

$$\bar{Y} = \sqrt{n}\left(\frac{Y-\mu}{\sigma}\right)$$

converges towards the standard Gaussian distribution, $\mathcal{N}(0,1)$.

This theorem can be read as the numerical average of a set of independent and identically distributed (i.i.d.) random variables X_i converges to a Gaussian distribution, as $n \to \infty$, where μ and σ^2 are the expectation and variance of the i.i.d random variables, respectively. The proof and detailed information about the central limit theorem can be found in [28, 30, 52].

2.4 Karhunen - Loève Expansion

In this section, fundamental aspects of representations of a random variable will be discussed. The main focus will be on a specific class of random process in L^2 and their representations in Fourier-like expansions that are convergent with respect to the norm associated with the corresponding inner product in L^2 space. The Karhunen-Loève (KL) expansion is a Fourier-like series for representing a stochastic process as a linear combination of orthogonal functions. It is also known as proper orthogonal decomposition (POD), which decomposes the random dimensions and the spatial dimensions of the stochastic process.

Let $\alpha(\mathbf{x}, \omega)$ be a zero mean random process where the vector $\mathbf{x}$ defined over the bounded domain $\mathcal{D}$, ω belongs to the space on random events Ω, and $C_\alpha(\mathbf{x}_1, \mathbf{x}_2)$ denotes its covariance function. By definition of the covariance function, it is bounded, symmetric, and positive definite [14]. As a result, its spectral decomposition can be defined as

$$C_\alpha(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=0}^{\infty} \lambda_i \phi_i(\mathbf{x}_1)\phi_i(\mathbf{x}_2), \tag{2.10}$$

where λ_i and $\phi_i(\mathbf{x})$ are, respectively, the eigenvalues and the eigenfunctions of the covariance operator. In other words, they are determined through the Fredholm integral

equation of the second kind

$$\int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, x_2) f_i(\mathbf{x}_1) d\mathbf{x}_1 = \lambda_i f_i(\mathbf{x}_2), \tag{2.11}$$

where the eigenfunctions ϕ_i are orthogonal and form a complete set since the covariance operator is symmetric and positive definite.

The stochastic process represented by $\alpha(\mathbf{x}, \omega)$ can be expanded in terms of eigenfunctions $\phi_i(\mathbf{x})$ and eigenvalues λ_i of the covariance function as done in [14]

$$\alpha(\mathbf{x}, \omega) = \sum_{i=0}^{\infty} \xi_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x}), \tag{2.12}$$

where $\xi_i(\omega)$ is a random variable. An explicit expression for a single random variable $\xi_i(\omega)$ can be obtained by multiplying the equation (2.12) with an eigenfunction $\phi_i(x)$ and integrating over the domain $\mathcal{D}$,

$$\xi_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{D}} \alpha(\mathbf{x}, \omega) \phi_i(\mathbf{x}) d\mathbf{x}. \tag{2.13}$$

Then, second order properties of the random variable ξ_i, i.e., mean and variance, can be derived such that

$$\langle \xi_i(\omega) \rangle = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{D}} \langle \alpha(\mathbf{x}, \omega) \rangle \, \phi_i(\mathbf{x}) d\mathbf{x} = 0,$$

$$\begin{aligned}
C_\xi(\xi_i(\omega), \xi_j(\omega)) &= \langle \xi_i(\omega), \xi_j(\omega) \rangle \\
&= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{\mathcal{D}} \int_{D} \langle \alpha_i(\mathbf{x}_1, \omega), \alpha_j(\mathbf{x}_2, \omega) \rangle \, \phi_i(\mathbf{x}_1) \phi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{\mathcal{D}} \int_{D} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_1) \phi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{\mathcal{D}} \phi_i(\mathbf{x}_1) \left(\int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_j(\mathbf{x}_2) d\mathbf{x_2} \right) d\mathbf{x_1} \\
&= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_{\mathcal{D}} \sqrt{\lambda_j} \phi_i(\mathbf{x}_1) \phi_j(\mathbf{x}_1) d\mathbf{x}_1 \\
&= \frac{1}{\sqrt{\lambda_i}} \sqrt{\lambda_i} \delta_{ij} \\
&= \delta_{ij}.
\end{aligned}$$

Assuming that $\bar{\alpha}(\mathbf{x})$ is the mean of a process $\alpha(\mathbf{x}, \omega)$, Karhunen-Loève expansion of the stochastic process $\alpha(\mathbf{x}, \omega)$ can be written as

$$\alpha(\mathbf{x}, \omega) = \bar{\alpha}(\mathbf{x}) + \sum_{i=0}^{\infty} \xi_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x}), \tag{2.14}$$

where

$$\langle \xi_i(\omega) \rangle = 0, \qquad \langle \xi_i(\omega), \xi_j(\omega) \rangle = \delta_{ij},$$

and λ_i, $\phi_i(\mathbf{x})$ are eigenvalues and eigenfunctions of the covariance operator C_α of the process $\alpha(\mathbf{x}, \omega)$, respectively.

2.4.1 Properties of Karhunen-Loève Expansion

Karhunen-Loève expansion provides an approximation for a stochastic process $\alpha(\mathbf{x}, \omega)$ by truncating the series with $M \in \mathbb{N}$ such that

$$\alpha(\mathbf{x}, \omega) \approx \sum_{i=0}^{M} \xi_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x}).$$

Then, the mean-square error of the approximation can be written as

$$\epsilon_M^2 = \langle \epsilon_M, \epsilon_M \rangle ,$$

where ϵ_M is its truncation error

$$\epsilon_M = \sum_{i=M+1}^{\infty} \xi_i(\omega) \sqrt{\lambda_i} \phi_i(\mathbf{x}).$$

As the number of terms used in the KL expansion increases, the mean-square error decreases monotonically at a rate depending on the decay of covariance C_α spectrum [28]. In other words, smaller number of terms are needed to achieve the desired error threshold when the stochastic process is more correlated.

Furthermore, the mean-square error resulting from the KL expansion is optimal,

which means that it is the minimum of all. That is,

$$\begin{aligned}
\epsilon_M^2 &= \langle \epsilon_M, \epsilon_M \rangle \\
&= \sum_{i=M+1}^{\infty} \sum_{j=M+1}^{\infty} \langle \xi_i(\omega), \xi_j(\omega) \rangle \sqrt{\lambda_i \lambda_j} \phi_i(\mathbf{x}) \phi_j(\mathbf{x}) \\
&= \sum_{i=M+1}^{\infty} \sum_{j=M+1}^{\infty} \phi_i(\mathbf{x}) \phi_j(\mathbf{x}) \int_{\mathcal{D}} \int_{\mathcal{D}} \langle \alpha(\mathbf{x}_1, \omega), \alpha(\mathbf{x}_2, \omega) \rangle \phi_i(\mathbf{x}_1) \phi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \sum_{i=M+1}^{\infty} \sum_{j=M+1}^{\infty} \phi_i(\mathbf{x}) \phi_j(\mathbf{x}) \int_{\mathcal{D}} \int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_1) \phi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2.
\end{aligned}$$

Integrating the equation above over the domain $\mathcal{D}$ and using the orthonormality of eigenfunctions, we obtain

$$\int_D \epsilon_M^2 = \sum_{i=M+1}^{\infty} \int_{\mathcal{D}} \int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2.$$

Thus, the problem remains to minimize $\int_{\mathcal{D}} \epsilon_M^2$ subject to the constraint that eigenfunctions $\phi_i(\mathbf{x})$ is normalized. That is constructed as,

$$F[\phi_i(\mathbf{x})] = \sum_{i=M+1}^{\infty} \int_{\mathcal{D}} \int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 - \lambda_m \left[\int_{\mathcal{D}} \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_1) d\mathbf{x}_1 - 1 \right].$$

Differentiating the equation $F[\phi_i(\mathbf{x})]$ with respect to $\phi_i(\mathbf{x})$ and setting it equal to zero yield

$$\frac{\partial F[\phi_i(\mathbf{x})]}{\partial \phi_i(\mathbf{x})} = \sum_{i=M+1}^{\infty} \int_{\mathcal{D}} \left[\int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_2) d\mathbf{x}_2 - \lambda_i \phi_i(\mathbf{x}_1) \right] d\mathbf{x}_1 = 0,$$

which is satisfied when $\phi_i(\mathbf{x})$ are chosen to be the eigenfunctions of the covariance operator $C_\alpha(\mathbf{x}_1, \mathbf{x}_2)$, that leads to the integral equation

$$\int_{\mathcal{D}} C_\alpha(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_2) d\mathbf{x}_2 = \lambda_i \phi_i(\mathbf{x}_1), \quad \forall i \in \mathbb{N}.$$

Hence, the truncated Karhunen–Loève expansion is the best approximation to the original process in the sense that the mean-square error of approximation is minimized.

Consider that a Gaussian process U is characterized by its variance σ_U^2 and its covariance function $C_U(\mathbf{x}_1, \mathbf{x}_2)$ which is given by the equation

$$C_U(\mathbf{x}_1, \mathbf{x}_2) = \sigma_U^2 e^{-|\mathbf{x}_1 - \mathbf{x}_2|/b}, \tag{2.15}$$

where $b > 0$ is the correlation length. The exact covariance function (2.15) is displayed in Figure 2.5.

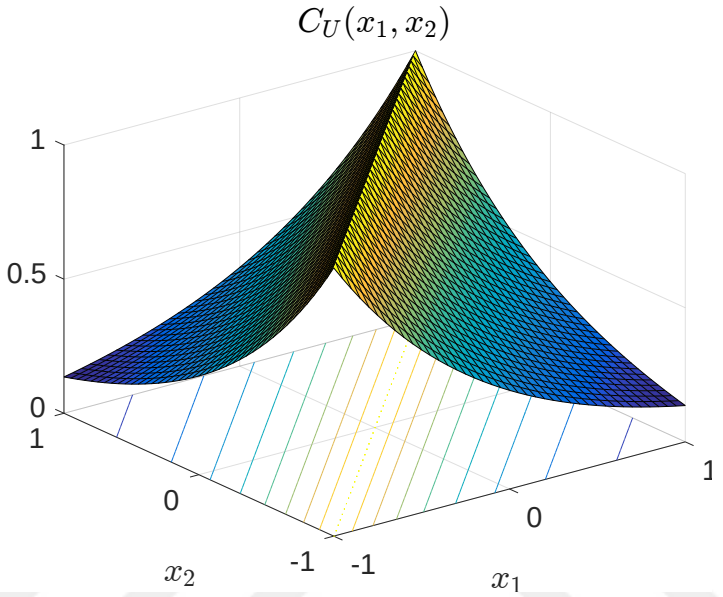

Figure 2.5: The covariance function of a Gaussian process on $[-1, 1]$ with $\sigma_U = 1$.

Assuming that a Gaussian process is defined over one dimensional interval $\mathcal{D} = [-a, a]$, the eigenfunctions and eigenvalues of the exponential covariance function are the solutions of the following integral equation

$$\int_{-a}^{a} \sigma_U^2 e^{-|\mathbf{x}_1 - \mathbf{x}_2|/b} \phi(\mathbf{x}_2) d\mathbf{x}_2 = \lambda \phi(\mathbf{x}_1). \tag{2.16}$$

Although this type of integral equations are generally solved with numerical methods, analytical expressions of eigenfunctions and eigenvalues can be found for this particular integral equation (2.16) in the case of Gaussian process. Moreover, the eigenfunctions and the corresponding eigenvalues are, respectively,

$$\phi_i(\mathbf{x}) = \begin{cases} \dfrac{\cos(\omega_i \mathbf{x})}{\sqrt{a + \frac{\sin(2\omega_i a)}{2\omega_i}}}, & \text{if } i \text{ is even,} \\ \dfrac{\sin(\omega_i \mathbf{x})}{\sqrt{a - \frac{2\sin(\omega_i a)}{2\omega_i}}}, & \text{if } i \text{ is odd,} \end{cases} \tag{2.17}$$

and

$$\lambda_i = \sigma_U^2 \frac{2b}{1 + (\omega_i b)^2}, \tag{2.18}$$

where ω_i are the ordered positive roots of the characteristic equation,

$$c(\omega_i) = \begin{cases} 1 - b\omega_i \tan(a\omega_i) = 0, & \text{if } i \text{ is even,} \\ b\omega_i + \tan(a\omega_i) = 0, & \text{if } i \text{ is odd.} \end{cases}$$

Moreover, Figure 2.6 shows the eigenfunctions (2.17) and eigenvalues (2.18) of the exponential covariance function defined in (2.15).

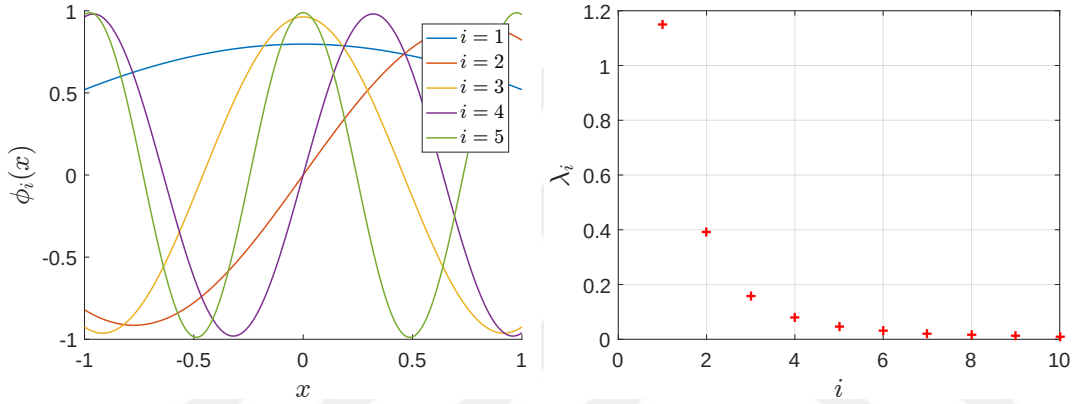

Figure 2.6: Eigenfunctions $\phi_i(x)$ (left) and eigenvalues λ_i (right) of the exponential covariance kernel (2.15) on $[-1, 1]$ with $b = 1$ and $\sigma_U = 1$.

Hence, the KL expansion to the exponential covariance function (2.15) can be constructed by truncating the series such that

$$C_U(\mathbf{x}_1, \mathbf{x}_2) \approx \widetilde{C}_U^M(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=0}^{M} \lambda_i \phi_i(\mathbf{x}_1)\phi_i(\mathbf{x}_2).$$

Moreover, the absolute error of the KL expansion can be defined as a function $R_U^N(x_1, x_2)$:

$$\begin{aligned} R_U^N(x_1, x_2) &= \left| \widetilde{C}_U^M(\mathbf{x}_1, \mathbf{x}_2) - C_U(\mathbf{x}_1, \mathbf{x}_2) \right| \\ &= \left| \sum_{i=0}^{M} \lambda_i \phi_i(\mathbf{x}_1)\phi_i(\mathbf{x}_2) - \sigma_U^2 e^{-|\mathbf{x}_1 - \mathbf{x}_2|/b} \right|. \end{aligned}$$

Unlike exponential covariance function, eigenvalues and eigenfunctions can not be defined explicitly for large number of covariance functions. Therefore, we need

some effective numerical techniques to approximate the values of the eigenvalues and eigenvectors of the covariance function, such as Collocation and Galerkin methods.

Using collocation methods is a decent choice for approximating the eigenpairs $\{\lambda_j, \phi_j\}$ resulting from the eigenvalue problem (2.11). Let x_i, $i = 1, \ldots, N$, be collocation points in domain $\mathcal{D}$ and define residual as

$$R_j = \int_{\mathcal{D}} C_\alpha(x, y)\phi_j(y)dy - \hat{\lambda}\hat{\phi}_j(x). \tag{2.19}$$

If the collocation points x_i are chosen to be the same as quadrature nodes whose weights are w_k, then the integral in (2.19) can be estimated as

$$\sum_{i=1}^{N} C_\alpha(x_k, y)\phi_j(x_i) \approx \int_{\mathcal{D}} C_\alpha(x_k, y)\phi_j(y)dy.$$

Then, setting the residual $R_j(x_k) = 0$ for all collocation points $k = 1, \ldots, N$ leads a system of equations

$$\sum_{i=1}^{N} C_\alpha(x_k, x_i)\phi_j(x_i) = \hat{\lambda}\hat{\phi}_j(x_k). \tag{2.20}$$

By solving the resulting system of equations (2.20) the numerical approximation of the eigenpairs $\{\hat{\lambda}_j, \hat{\phi}_j\}$ can be obtained.

Another convenient choice of numerical method for solving the corresponding eigenvalue problem is the Galerkin technique. Let V_N be a subspace of $L^2(\mathcal{D})$ with basis functions $\varphi_1(x), \ldots, \varphi_N$. Then, we can represent numerical approximation by a linear combination

$$\hat{\phi}_j = \sum_{k=1}^{N} a_k^j \varphi_j(x)dx, \quad a_k^j \in \mathcal{D}. \tag{2.21}$$

Multiplying the eigenvalue problem (2.11) with each basis function φ_i and integrating it over $\mathcal{D}$ result in a Galerkin system as

$$\int_{\mathcal{D}} \left(\int_{\mathcal{D}} C_\alpha(x, y)\hat{\phi}_j(y)dy \right) \varphi_i(x)dx = \hat{\lambda}_j \int_{\mathcal{D}} \hat{\phi}_j \varphi_i(x)dx. \tag{2.22}$$

Then, substituting (2.21) into (2.22) we obtain a system of equations

$$\sum_{k=1}^{N} a_k^j \int_{\mathcal{D}} \int_{\mathcal{D}} C_\alpha(x, y)\varphi_j(x)\varphi_i(x)dydx = \hat{\lambda}_j \sum_{k=1}^{N} a_k^j \int_{\mathcal{D}} \varphi_j(x)\varphi_i(x)dx. \tag{2.23}$$

In order to construct a compact matrix notation for the system of equation (2.23), we can form $N \times N$ stiffness matrix K and $N \times N$ mass matrix M such that

$$[K]_{ij} = \int_{\mathcal{D}} \int_{\mathcal{D}} C_{\alpha}(x, y)\varphi_j(x)\varphi_i(x)dydx, \quad [M]_{ij} = \int_{\mathcal{D}} \varphi_j(x)\varphi_i(x)dx.$$

Finally, we have the following matrix eigenvalue problem

$$Ka_j = \hat{\lambda}_j M a_j,$$

where the vector a_j contains investigated coefficients a_j.

Until now it can be seen that the Karhunen-Loève expansion is implemented for processes only whose their covariance function is known. However, profound information about most of the stochastic processes is not known priori. In such cases, an alternative spectral expansion, polynomial chaos expansion, is needed for approximating specifically the solution process [54].

2.5 Polynomial Chaos Expansion

The polynomial chaos expansion is a tool for representing second order stochastic processes whose covariance function is not known. This concept was introduced firstly by Wiener [50] as a series of nonlinear functionals of the Brownian motion. Then, Cameron and Martin [6] proved that a Fourier-like series with an orthogonal basis converges to these nonlinear functionals with Gaussian measure in L^2 sense.

Let $\{\xi_i(\omega)\}_{i=1}^{\infty}$ be a sequence of centered, orthonormal, Gaussian random variables defined on the space $L^2(\Omega, \Sigma, P)$ of second-order random variables. Consider the space $\widehat{\Gamma}_p$ of all polynomials of $\{\xi_i(\omega)\}_{i=1}^{\infty}$ having degree at most p. Let Γ_p denotes the set of polynomials in $\widehat{\Gamma}_p$ that are orthogonal to polynomials in $\widehat{\Gamma}_{p-1}$ and the space Γ_p is called the polynomial chaos of order p. Finally, let $\bar{\Gamma}_p$ denotes the space spanned by Γ_p which is called the p^{th} Homogeneous Chaos.

Based on the definitions, the polynomial chaos of order p consists of all orthogonal polynomials of order p which involves any possible combination of the random variables $\{\xi_i(\omega)\}_{i=1}^{\infty}$. A second-order random variable $U(\xi(\omega)) \in L^2(\Omega, \Sigma, P)$ can be

represented with a polynomial chaos (PC) expansion of the form [6]

$$\begin{aligned} U(\omega) = u_0\Gamma_0 &+ \sum_{i_1=1}^{\infty} u_{i_1}\Gamma_1(\xi_{i_1}(\omega)) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} u_{i_1 i_2}\Gamma_2(\xi_{i_1}(\omega),\xi_{i_2}(\omega)) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} u_{i_1 i_2 i_3}\Gamma_3(\xi_{i_1}(\omega),\xi_{i_2}(\omega),\xi_{i_3}(\omega)) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2}\sum_{i_4=1}^{i_3} u_{i_1 i_2 i_3 i_4}\Gamma_3(\xi_{i_1}(\omega),\xi_{i_2}(\omega),\xi_{i_3}(\omega),\xi_{i_4}(\omega)) \\ &+ \cdots, \end{aligned} \tag{2.24}$$

where Γ_p are successive polynomial chaoses of order p.

The original polynomial chaoses introduced by Wiener result on the Hermite polynomials, also called Hermite-Chaos, in terms of Gaussian random variables and the details related to construction of Hermite-Chaoses can be found in [14]. By construction, polynomial chaoses whose order greater than one have zero mean

$$\mathbb{E}[\Gamma_p] = 0,\ p > 0.$$

If there is a one-to-one correspondence between the functions $\Gamma_n(\xi_{i_1},\ldots,\xi_{i_n})$ and $\Psi_i(\xi_1,\xi_2,\ldots)$, then the equation (2.24) can be rewritten in the form

$$U(\omega) = \sum_{i=0}^{\infty} u_i\Psi_i(\boldsymbol{\xi}), \quad \boldsymbol{\xi} = \{\xi_1,\xi_2,\ldots\}, \tag{2.25}$$

where the deterministic coefficients u_i are simply called polynomial chaos (PC) coefficients. According to the definition proposed in [6] by Cameron-Martin, strong approximation of a functional $f(\boldsymbol{\xi})$ in L^2 can be defined as

$$P_N f = \sum_{i=0}^{N} f_i\Psi_i(\boldsymbol{\xi}),$$

where f_i are the PC coefficients

$$f_i = \frac{\langle f(\xi),\Psi_i(\boldsymbol{\xi})\rangle}{\langle\Psi_i(\boldsymbol{\xi}),\Psi_i(\boldsymbol{\xi})\rangle}.$$

Moreover, the PC representation of functional $f(\boldsymbol{\xi})$ converges to $f(\boldsymbol{\xi})$ in L^2 sense [6]

$$\lim_{N\to\infty} \mathbb{E}\left[(P_N f - f(\boldsymbol{\xi}))^2\right] = 0.$$

When the PC representation $P_N f$ converges in mean square sense, it implies that $P_N f$ converges in probability, $P_N f \xrightarrow{P} f$, then it consequently leads the convergence in distribution, $P_N f \xrightarrow{d} f$. Furthermore, the optimality property can be obtained by applying the best approximation theorem 2.3 for orthogonal approximations

$$\|f - P_N f\| = \inf_{g\in\mathbb{P}_N} \|f - g\|.$$

It means that the PC approximation $P_N f$ is optimal among the linear space $\mathbb{P}_N$ of all polynomials of degree at most N.

Remark 2.3. Besides the PC expansion is an effective tool for approximating stochastic processes with Gaussian inputs, it is also generalized to be convergent for non-Gaussian random inputs [54]. While the polynomial basis of PC expansions is the set of Hermite polynomials, the generalized polynomial chaos (gPC) expansions have their polynomial basis chosen from the Askey-scheme [54].

Askey scheme is a system of organizing the hypergeometric orthogonal polynomials [24]. The probability density functions of some types of random distributions are same as the weight functions of some particular types of orthogonal polynomials from the Askey-scheme. In practice, the type of polynomials in the Askey-scheme is chosen according to the type of distribution of independent random input as given in Table 2.1. Similar to expansion (2.25), we represent the general random process $U(\xi(\omega))$ in $L^2(\Omega, \Sigma, P)$ as

$$U(\xi) = \sum_{i=0}^{\infty} u_i \Phi_i(\xi), \tag{2.26}$$

where Φ_i are orthogonal polynomial basis from the Askey-scheme. Since each type of polynomials from the Askey-scheme form a complete orthogonal basis for the Hilbert space determined by their corresponding weight function, generalized result of theorem of Cameron-Martin can be used to obtain convergence. Therefore, each type of generalized Polynomial Chaos expansion converges to a functional $f(\xi)$ which is an element of the corresponding Hilbert space [6].

Table 2.1: Correspondence between gPC polynomials an Askey-scheme.

	Distribution of ξ	gPC basis $\{\Phi(\xi)\}$	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	$\{0, 1, \ldots, N\}$
	Negative Binomial	Meixner	$\{0, 1, 2, \ldots\}$
	Hypergeometric	Hahn	$\{0, 1, \ldots, N\}$

The original PC is a subset of the gPC since the basis of the original polynomial chaos corresponds to the Hermite polynomials in the Askey-scheme in order to represent Gaussian random variable.

2.5.1 Statistics

When the gPC expansion is used to represent stochastic process $f(Z)$, important statistical information of $f(Z)$ can be derived from the gPC representation. Consider a stochastic process $f(x, t, Z)$ with $x \in \mathcal{D}$, $t \in T$ and $Z \in \mathbb{R}^d$, then the N^{th} order gPC approximation can be expressed as

$$P_N f(x, t, Z) = \sum_{|\mathbf{i}| \leq N} f_i(x, t) \Phi_i(Z) \in \mathbb{P}_N^d$$

for any fixed $x \in D$ and $t \in T$, i.e., $f \approx P_N f$. Hence, the mean of gPC approximation is derived as

$$\begin{aligned}
\mu_f &= \mathbb{E}[f(x, t, Z)] \approx \mathbb{E}[P_N f(x, t, Z)] \\
&= \int \left(\sum_{|\mathbf{i}| \leq N} f_i(x, t) \Phi_i(Z) \right) dF_Z(Z) \\
&= \int \left(\sum_{|\mathbf{i}| \leq N} f_i(x, t) \Phi_i(Z) \right) \Phi_0(Z) dF_Z(Z) \\
&= f_0(x, t)
\end{aligned}$$

by using the orthogonality condition of the gPC basis. Similarly, the variance of gPC

approximation is

$$\begin{aligned}
\mathrm{Var}[f(x,t,Z)] &= \mathbb{E}[(f(x,t,Z)-\mu_f(x,t))^2] \\
&\approx \mathbb{E}[(P_N f(x,t,Z)-\mu_f(x,t))^2] \\
&= \int \left(\sum_{|\mathbf{i}|\le N} f_i(x,t)\Phi_i(Z) - f_0(x,t) \right)^2 dF_Z(Z) \\
&= \int \left(\sum_{0<|\mathbf{i}|\le N} f_i(x,t)\Phi_i(Z) \right)^2 dF_Z(Z) \\
&= \sum_{0<|\mathbf{i}|\le N} f_i^2(x,t)\gamma_{\mathbf{i}}.
\end{aligned}$$

Other statistical quantities of the stochastic process f can also be estimated by simply applying to the gPC approximation $P_N f$ to their analytic expressions.

CHAPTER 3

THE MODEL EQUATION

Many practices of applied sciences such as ion-acoustic waves in collisionless plasma, shallow-water gravity waves, long internal waves in the atmosphere, and ocean are described with solitons [25, 41, 49]. A soliton is a solitary wave packet that preserves its shape while it is traveling at a constant velocity in uniform direction and a soliton can interact strongly with other solitons and retain its identity. The first experimental observation of such solitary waves with similar characteristics was made in 1834 by John Scott Russell [38]. Solitons can be obtained from a widespread class of weakly nonlinear dispersive partial differential equations (PDEs) as their solutions.

Korteweg-de Vries (KdV) equation is a fundamental mathematical model for simulating long nonlinear wave propagation on the surface of a narrow and shallow channel. It was derived firstly by Boussinesq [5], then rediscovered by Diedrik Korteweg and Gustav de Vries [26] in the aim of modeling solitary waves that was discovered in nature by Russell. In 1965 Martin Zabusky and Norman Kruskal showed that the KdV equation can be solved numerically for a class of initial conditions which led to the discovery that solitary waves preserves their original form and speed in the interaction with each other [56]. Existence of infinitely many conservation laws for the KdV equation corresponding to specific properties of its solution was proved in [34]. For certain initial states, inverse scattering transform (IST) method was used to evaluate exact solutions of the KdV equation [13].

In particular, the KdV is a prototypical model, formed as an exactly solvable nonlinear PDE, that provides a delicate balance between nonlinearity and dispersion effects. In this chapter, we firstly introduce the original KdV equation, then we present

the KdV equation with random input data by imposing some random parameters which is the main model of our study.

3.1 Korteweg de Vries Equation

In the original study [26], the KdV equation is defined within parameters, the depth of the water l and the elevation of the surface above equilibrium η, such that

$$\frac{\partial \eta}{\partial t} = \frac{3}{2}\sqrt{\frac{g}{l}}\frac{\partial}{\partial x}\left(\frac{1}{2}\eta^2 + \frac{2}{3}\alpha\eta + \frac{1}{3}\sigma\frac{\partial^2\eta}{\partial x^2}\right), \tag{3.1}$$

where $\sigma = \frac{1}{3}l^2 - \frac{Tl}{\rho g}$ depends on water density ρ, surface tension T, and an arbitrary constant α. The equation (3.1) can be rewritten in non-dimensional form by defining transformations

$$\eta = -\frac{2}{3}\alpha(6u+1), \quad \tau = \sqrt{\frac{2\alpha^3 g}{\sigma l}}t, \quad \xi = -\sqrt{\frac{2\alpha}{\sigma}}x.$$

When these transformations are applied on the equation (3.1), it becomes

$$u_\tau - 6uu_\xi + u_{\xi\xi\xi} = 0, \tag{3.2}$$

where the coefficient of nonlinear term 6 is a scaling factor for a complete integrability. All of the coefficients can be rescaled to arbitrary nonzero values through appropriate transformations. Then, an initial value problem (IVP) can be formed with the non-dimensional KdV equation (3.2)

$$\begin{gathered} u_t + \beta u u_x + u_{xxx} = 0, \quad \beta \neq 0, \\ u(x,0) = f(x). \end{gathered} \tag{3.3}$$

For appropriate class of initial conditions Sjöberg proved the existence and uniqueness of traveling wave solution of the KdV equation analytically [42].

Theorem 3.1. *[42, Theorem 1] If the nonlinear coefficient is nonzero, $\beta \neq 0$, and if the initial condition $f(x)$ is a periodic function where $f'(x) \in L^2$, then there exists a unique solution to the initial value problem* (3.3)*.*

Later, Lax showed that a localized traveling-wave solution of the KdV equation corresponds to the solitary waves described by Russel [27]. Since we are interested in traveling wave solution of the form $u(x,t) = f(\xi)$ such that $u, u_\xi, u_{\xi\xi} \to 0$ as $\xi \to \pm\infty$, where $\xi = x - ct$, then c denotes the speed of the wave. Substituting the trial solution into the KdV equation (3.2) leads to an ordinary differential equation (ODE)

$$-cf_\xi + -6ff_\xi + f_{\xi\xi\xi} = 0.$$

Integration of this equation gives us

$$-cf - 3f^2 + f_{\xi\xi} = A,$$

where A is a constant of integration. Multiplying the equation by f_ξ and integrating one more yield

$$\frac{1}{2}(f_\xi)^2 = f^3 + \frac{1}{2}cf^2 + Af + B,$$

where B is another constant of integration. Since $f, f_\xi, f_{\xi\xi} \to 0$ as $\xi \to \pm\infty$, it follows that $A = B = 0$. Thus, the differential equation takes its final form:

$$(f_\xi)^2 = f^2(2f + c).$$

It is obvious that $2f + c \geq 0$ must hold for the existence of a real solution. By separation of variables, the last equation can be written as

$$\int \frac{df}{f\sqrt{2f + c}} = \pm \int d\xi.$$

Here, integration can be done by the following transformation

$$f = -\frac{c}{2}\text{sech}^2\theta,$$

which yields the solution

$$f(\xi) = -\frac{c}{2}\text{sech}^2\left(\frac{1}{2}\sqrt{c}(\xi - \xi_0)\right).$$

Re-substituting $u(x,t) = f(\xi) = f(x - ct)$, the traveling wave solution corresponds to

$$u(x,t) = -\frac{c}{2}\text{sech}^2\left(\frac{1}{2}\sqrt{c}(x - ct - x_0)\right), \tag{3.4}$$

where the constant x_0 is the starting position of the wave. A traveling wave solution of the KdV equation (3.2) with a constant speed is displayed in the Figure 3.1.

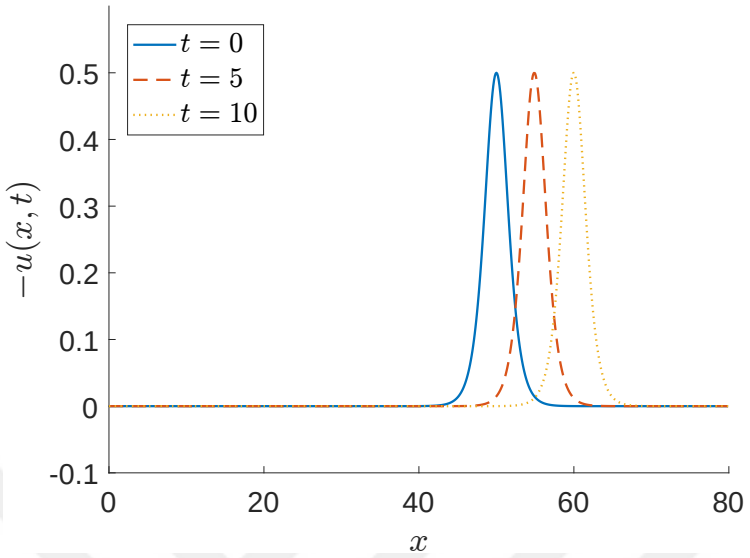

Figure 3.1: Single solitary wave with $c = 1$ at times: $t = 0$, $t = 5$, and $t = 10$.

An important characteristic of the KdV equation is that it has infinitely many integrals of motion which are constant [34]. A constant of motion is a quantity that is conserved throughout the motion which represents a physical constraint. First three integrals of motion (I_1, I_2, I_3) are

$$I_1 = \int u(x,t)dx, \quad I_2 = \int u^2(x,t)dx, \quad I_3 = \int \left[2u^3(x,t) - (u_x(x,t))^2\right] dx,$$

where they represent mass, momentum, and energy of the wave motion, respectively. Moreover, they are called as the first invariants of the KdV solution.

Although the KdV equation is a good model for describing physical wave phenomenons in the real-life applications, it cannot govern all the attributes of real world practices sufficiently because of incomplete knowledge of physical model and imprecise calculations of the data. Therefore, modeling and simulating the real-world studies by the KdV equations with random input data can establish a better understanding and anticipating the wave phenomena.

3.2 Korteweg de Vries Equation with Random Input Data

The general KdV equation with random input data forms an initial value problem with periodic boundary conditions such that

$$\begin{aligned} u_t + (-6u + \delta\alpha(x,\omega))u_x + u_{xxx} &= \epsilon\beta(t,\omega), \\ u(x,0) &= -\frac{c}{2}\mathrm{sech}^2\left(\frac{1}{2}\sqrt{c}x\right), \end{aligned} \tag{3.5}$$

where $\alpha(x,\omega)$ and $\beta(t,\omega)$ are the space and time dependent random noises with their amplitudes δ and ϵ. The noise $\beta(t,w)$ is called the additive noise, whereas $\alpha(t,w)$ is called as multiplicative noise. In this thesis, we consider three kind of cases: only additive noise ($\delta = 0$); only multiplicative noise ($\epsilon = 0$); both additive and multiplicative noises.

In the first case that only additive noise exists in the equation (3.5), $\delta = 0$, the IVP corresponds to

$$\begin{aligned} u_t - 6uu_x + u_{xxx} &= \epsilon\beta(t,\omega), \\ u(x,0) &= -\frac{c}{2}\mathrm{sech}^2\left(\frac{1}{2}\sqrt{c}x\right), \end{aligned} \tag{3.6}$$

where c is the wave speed.

The stochastic Korteweg-de Vries equation defined with additive noise describes effects of motion of solitons under the external noise. Wadati [47] showed that the stochastic KdV equation with additive Gaussian noise has an analytic solution for a single soliton whose mean propagates with its width increasing proportional to $t^{3/2}$ and its height decreasing proportional to $t^{-3/2}$ for a large t. Assume that the time-dependent additive noise $\beta(t,\omega)$ is integrable in time, i.e.,

$$W(t,\omega) = \epsilon\int_0^t \beta(\epsilon,t')dt', \tag{3.7}$$

and define

$$U(x,t,\omega) = u(x,t) - W(t,\omega). \tag{3.8}$$

Then, putting transformation (3.8) into the equation (3.6) leads to

$$U_t - 6\,W(t)\,U_x - 6\,U\,U_x + U_{xxx} = 0. \tag{3.9}$$

Set the Galilean transformation

$$X = x + m(t, \omega), \tag{3.10}$$

where

$$m(t, \omega) = 6 \int_0^t W(t_1, \omega) dt_1.$$

With the help of the Galilean transformation (3.10), the equation (3.9) can then be rewritten as

$$U_t(X, t) - 6\, U(X, t)\, U_X(X, t) + U_{XXX}(X, t) = 0. \tag{3.11}$$

Then, the equation (3.11) is similar to a standard KdV equation and the analytical solution is

$$U(X, t) = -\frac{c}{2} \mathrm{sech}^2 \left(\frac{1}{2} \sqrt{c} (x - ct) \right),$$

where c is the wave speed. Thus, the analytical single-soliton solution to the stochastic KdV equation (3.6) is

$$\begin{aligned} u(x, t, \omega) &= W(t, \omega) + U(X, t) \\ &= W(t, \omega) - \frac{c}{2} \mathrm{sech}^2 \left(\frac{1}{2} \sqrt{c} (x - ct) \right). \end{aligned}$$

If time-dependent noise is fully correlated, β can be described as a Gaussian random variable with zero mean and unit variance ξ, i.e., $\beta(t, \omega) = \xi(\omega)$. Thus, $W(t, \omega) = \epsilon \xi t$. Thus, the exact solution becomes

$$u(x, t, \omega) = \epsilon \xi t - \frac{c}{2} \mathrm{sech}^2 \left(\frac{1}{2} \sqrt{c} (x - ct - x_0) \right). \tag{3.12}$$

Then, the corresponding exact mean and variance of the solution (3.12), plotted in the Figure 3.2, are

$$\mathbb{E}[u(x, t, \omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(x, t, \omega) e^{-\frac{\xi^2}{2}} d\xi, \tag{3.13}$$

$$\mathrm{Var}[u(x, t, \omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} u^2(x, t, \omega) e^{-\frac{\xi^2}{2}} d\xi - \mathbb{E}[u(x, t, \omega)]^2. \tag{3.14}$$

In the case that time dependent noise is partially correlated Gaussian, β can be expressed with a random process $\beta(t, \omega)$ with zero mean. Assuming that covariance

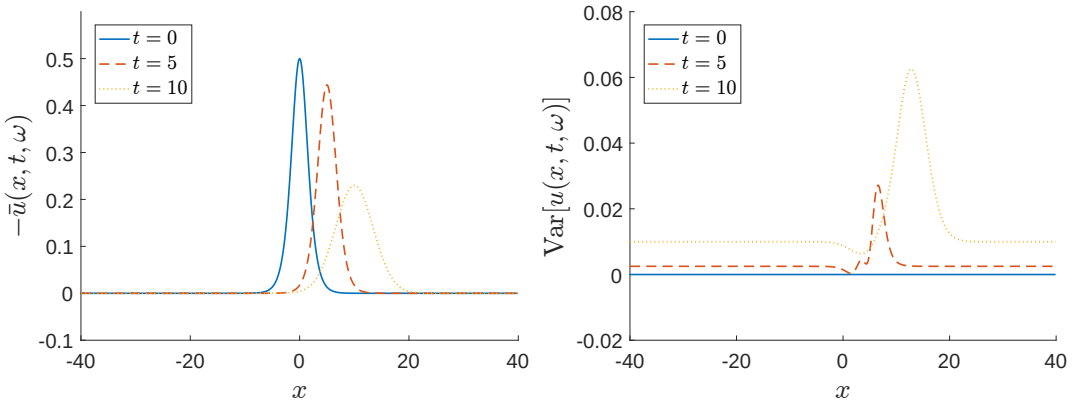

Figure 3.2: The mean values (left) and the variances (right) of $-u(x,t,\omega)$ with $c = 1$ and $\epsilon = 0.01$ at times: $t = 0$, $t = 5$, and $t = 10$.

function of the partially correlated Gaussian process is given by the exponential kernel, the stochastic input $\beta(t,\omega)$ can be approximated with the Karhunen-Loève expansion such that

$$\beta(t,\omega) = \sum_{i=1}^{N} \sqrt{\lambda_i}\phi_i(t)\xi_i, \tag{3.15}$$

where N is the number of random dimensions, ξ_i is the Gaussian random variable, λ_i and ϕ_i are the eigenvalues and eigenfunctions of the correlation function, respectively.

When the approximated Gaussian noise (3.15) is integrated in time, we obtain

$$W(t,\omega) = \sum_{i=1}^{N} \epsilon\sqrt{\lambda_i}\xi_i \int_0^t \phi_i(t')dt'.$$

Then, the analytic solution of a single soliton becomes

$$\begin{aligned} u(x,t,\omega) =& W(t,\omega) - \frac{c}{2}\mathrm{sech}^2\left(\frac{1}{2}\sqrt{c}(x-ct) + 3\sqrt{c}\int_0^t W(t',\omega)dt'\right) \\ =& \epsilon\sum_{i=1}^{N}\sqrt{\lambda_i}\xi_k \int_0^t \phi_i(t')dt' \\ & - \frac{c}{2}\mathrm{sech}^2\left(\frac{1}{2}\sqrt{c}(x-ct) + 3\sqrt{c}\epsilon\sum_{i=1}^{N}\sqrt{\lambda_i}\xi_k\int_0^t\int_0^{t_1}\phi_i(t')dt'dt_1\right). \end{aligned}$$

Hence, the exact mean is

$$\mathbb{E}[u(x,t,\omega)] = \frac{1}{(2\pi)^N}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x,t,\omega)e^{\frac{\xi_0^2+\cdots+\xi_N^2}{2}}d\xi_0\cdots d\xi_N,$$

whereas the exact variance is

$$\begin{aligned}\mathrm{Var}[u(x,t,\omega)] = \frac{1}{(2\pi)^N}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u^2(x,t,\omega)e^{\frac{\xi_0^2+\cdots+\xi_N^2}{2}}d\xi_0\cdots d\xi_N \\ - \mathbb{E}[u(x,t,\omega)]^2.\end{aligned}$$

In the second case, we consider the stochastic KdV equation (3.5) with only multiplicative noise, $\epsilon = 0$. Then, we obtain the following IVP problem

$$\begin{aligned} u_t + (-6u + \delta\alpha(x,\omega))u_x + u_{xxx} = 0, \\ u(x,0) = -\frac{c}{2}\mathrm{sech}^2\left(\frac{1}{2}\sqrt{c}x\right). \end{aligned} \tag{3.16}$$

The effect of multiplicative noise with long range correlation was studied by Iizuka theoretically in [19], and then Scalerandi and Romano studied numerically [39].

In the last case, we are interested in the stochastic KdV equation (3.5) with additive and multiplicative noises combined as mentioned earlier. In both multiplicative noise and combined cases, there is no analytical solution for the equations (3.16), (3.5), respectively. Therefore, numerical methods have to be used in order obtain solutions for such kind of the stochastic KdV equations. In the following chapter, we mention various numerical approaches such as non-intrusive methods (Monte Carlo and stochastic collocation) and intrusive methods (stochastic Galerkin) to solve this kind of problems numerically.

CHAPTER 4

SPECTRAL METHODS

The main objective of uncertainty quantification (UQ) is to study the effects of having uncertainty on the input variables to a mathematical model. Although mathematical models are not restricted with differential equations, partial differential equations (PDEs) with random inputs are one of the most powerful tools for modeling real world phenomenons. There have been studies recently on several numerical methods to approximate the solutions to such PDEs with the uncertain coefficients. Common numerical techniques can be divided into two categories which are non-intrusive and intrusive methods.

Non-intrusive methods are sampling based, in other words, they rely on a set of deterministic model resolutions corresponding to some specific samples of the random input. Here, the deterministic PDE corresponding to each realization could be discretized using any numerical methods for solving differential equations. One of the most traditional non-intrusive approach is Monte Carlo method which firstly generates a set of random realizations for the predefined random inputs and then utilizes repetitive deterministic solvers for each realization. Monte Carlo method has widespread applications because of its simplicity and natural parallelization. The other popular non-intrusive method is Stochastic Collocation, which principally seeks an interpolated polynomial representation of the random field generated by solutions of the stochastic PDE model. The implementation of this method is straightforward because it approximates the solution on a finite set of interpolation points on the random coefficient field. Although the non-intrusive methods are easy to implement, Monte Carlo method have insufficiently low convergence rate [11] and stochastic

collocation method suffers from curse of dimensionality [2].

On the other hand, intrusive methods seek for a spectral stochastic representation for the solution of the PDE with random inputs. As the name suggest, these methods appreciate the random field by predefined stochastic input while solving the PDE model. The stochastic Galerkin method is the principally major example of the intrusive methods. As similarly in the standard Galerkin methods, they are built on weighted residual formalism for determining coefficients of the spectral stochastic representation of the solution. Assuming that solutions of stochastic PDEs are sufficiently smooth in the random space, these methods result in fast convergence rates [28]. However, designing the resulting system of equation can be very complicated if the differential equation model has nontrivial and nonlinear forms.

In this chapter, the non-intrusive methods such as Monte Carlo sampling and stochastic collocation methods will be first introduced. Then, intrusive method, stochastic Galerkin method, will be presented by utilizing stochastic Korteweg de-Vries equation introduced in Section 3.2.

4.1 Non-Intrusive Methods

The non-intrusive methods generally rely on the set of realizations of the stochastic systems which are deterministic solutions of the model for some specific values of random inputs. This is the most significant feature of the non-intrusive methods for approximating stochastic models since it requires only deterministic solver without any particular adaptation. In addition, solving deterministic model for each specific values can be decoupled naturally so that computation can be done in parallel. However, the numerical cost of non-intrusive methods grows exponentially with the number of deterministic model resolution, which means that the non-intrusive methods are computationally intensive. Now, we introduce two non-intrusive methods; Monte Carlo simulation and stochastic collocation.

4.1.1 Monte Carlo Simulation

The Monte Carlo simulation (MCS) is one of the simplest and the most commonly used method for solving stochastic differential equations. The fundamental idea behind MCS is to generate a set of independent realizations of the random inputs based on their probability distribution. Then, the stochastic PDE becomes deterministic for each realization generated. Solving the deterministic problem for all set of independent realizations yields an ensemble of realizations of the random solution. Hence, statistical information, e.g., mean and variance, can be extracted from these set of realizations.

The general procedure of the MCS for solving the model problem (3.5), i.e., the KdV equation with random input data, is as follows:

1. Generate identically and independently distributed random numbers $\xi^{(i)} = \{\alpha^{(i)}, \beta^{(i)}\}$, $i = 1, \ldots, M$ with respect to their distributions;
2. Solve the corresponding deterministic PDE with its coefficients for each $i = 1, \ldots, M$ to form a sample set from solutions $u^{(i)}(x,t) = u(x,t,\xi^{(i)})$;
3. Approximate the required solution statistics by using proper schemes from the sample set $\{u^{(i)}(x,t)\}$.

For instance, the expectation of the stochastic solution, as a result of the Central Limit Theorem (CLT), see Theorem 2.4, can be immediately approximated as

$$\mathbb{E}[u] \approx \bar{u}(x,t) = \frac{1}{M}\sum_{i=1}^{M} u(x,t,\xi^{(i)}).$$

Inherently, error estimation of the mean derived by MCS also follows from the CLT given in Section 2.3.4. Since the solutions $\{u^{(i)}(x,t)\}$ are independent and identically distributed, the distribution of $\bar{u}(x,t)$ converges to a Gaussian distribution $\mathcal{N}(\mathbb{E}[u], \frac{\sigma_u}{\sqrt{M}})$ as $M \to \infty$. Hence, it is deduced that the convergence rate of the MCS statistics is $\mathcal{O}(M^{-1/2})$.

Although the MCS is one of the most powerful and the most flexible mathematical tool for solving the stochastic PDEs, the convergence rate, $\mathcal{O}(M^{-1/2})$, is the main

limitation of the MCS. In the aim of obtaining a decent accuracy of the statistical estimates, computational cost is dominated by increasing number of deterministic resolution of the governing equation. Herewith, there have been developments on sampling methods that improve their accuracy and efficiency such as quasi-Monte Carlo [35], multi-level Monte Carlo [15], and Latin hypercube sampling method [32].

4.1.2 Stochastic Collocation

The fundamental idea behind the deterministic collocation methods is to ensure the residue of the governing differential equation to be zero at the discrete nodes on the spatial domain, which are called collocation points. This definition can also be implemented to the random domain of the stochastic differential equations, which stands for stochastic collocation (SC) methods. Implementation of the SC methods is straightforward since it also naturally leads to uncoupled deterministic differential equations. Furthermore, these methods achieve exponential convergence under the assumption that the random coefficients are infinitely differentiable with respect to random variables [2].

Let us choose a set of collocation points $Z = \{Z^j\}_{j=1}^N$ in the random space of the stochastic KdV equation with the additive noise (3.6). Then the governing PDE with stochastic input which is needed to be solved for $j = 1, \ldots, N$, each collocation point, becomes

$$u_t(x,t,Z^j) - 6u(x,t,Z^j)u_x(x,t,Z^j) + u_{xxx}(x,t,Z^j) = \epsilon\beta(Z^j,t). \tag{4.1}$$

Since value of random parameter Z is fixed at each collocation point Z^j, this stochastic collocation method consists of solving N deterministic equations. Then solving the PDE system (4.1) for each collocation node results in a set of deterministic solutions $\{u(Z^j)\}_{j=1}^N$, where $u(Z^j) = u(x,t,Z^j)$ represents the solution of the model (4.1) at any collocation node Z^j. Then, the solution can be extracted from the set of realizations by using the interpolation approach.

The SC solution $u(Z)$ of the governing model (4.1) can be represented by using Lagrange interpolating approach, described in Section 2.2.1. Hence the solution $u(Z)$

can be approximated as

$$u(Z) \approx u_N(Z) = \sum_{j=1}^{N} u(Z^j) L_j(Z), \tag{4.2}$$

where $L_j(Z)$ are the Lagrange polynomials

$$L_j(Z^i) = \delta_{ij}, \quad 1 \leq i,\, j \leq N.$$

It is an important feature of the interpolation that solution $u_N(Z)$ is equal to the exact solution in each of the N collocation points. As in the Monte Carlo simulation, the statistical information of the stochastic solution $u_N(Z)$ can be derived from the set of N deterministic solutions computed for each the collocation points. For instance, the mean of the SC solution (4.2) can be computed as

$$\mathbb{E}[u_N(Z)] = \sum_{j=1}^{N} u(Z^j) \int_{I_Z} L_j(z) f_Z(z) dz, \tag{4.3}$$

where I_Z is the support of the random variable Z and $f_Z(z)$ is the probability density function of the distribution of the random variable Z. The evaluation of the expectation requires knowledge of the Lagrange polynomial and it can be computed by use of an inverted Vandermonde matrix as done in [53]. Besides computing the integral in (4.3) analytically, another common approach is approximating it by using quadrature rules. Approximation of the integrals can be done exactly by quadrature if its order is chosen to be equal to N Lagrange polynomials order such that

$$\mathbb{E}[u_N(Z)] = \sum_{j=1}^{N} u(Z^j) \sum_{k=1}^{N} L_j(z_k) f_Z(z_k) w_k, \tag{4.4}$$

where z_k are the quadrature points and w_k are the quadrature weights. Furthermore, if the quadrature nodes are chosen as same as the collocation points so that they represent the distribution of the random parameters as given in Table 2.1, and then the mean of the interpolation (4.4) can be analytically simplified. Choosing the collocation points as the quadrature points, i.e., $z_j = Z^j$, ensures that Lagrange interpolating functions $L_j(Z^j) = \delta_{ij}$ vanishes for $i, j \in [1, \ldots, N]$. Therefore, the equation (4.4)

reduces to

$$\begin{aligned}\mathbb{E}[u_N(Z)] &= \sum_{j=1}^{N} u(Z^j) \sum_{k=1}^{N} \delta_{jk} f_Z(Z^j) w_k \\ &= \sum_{j=1}^{N} u(Z^j) f_Z(Z^j) w_j.\end{aligned}$$

In a similar way, the variance of SC solution (4.2) can be derived such that

$$\begin{aligned}\mathrm{Var}[u_N(Z)] &= \mathbb{E}[(u_N(Z) - \mathbb{E}[u_N(Z)])^2] \\ &= \int_{I_Z} (u_N(Z) - \mathbb{E}[u_N(Z)])^2 \, f_Z(z)\, dz \\ &= \int_{I_Z} \left(\sum_{j=1}^{N} u(Z^j) L_j(z) - \mathbb{E}[u_N(Z)] \right)^2 f_Z(z)\, dz \\ &= \sum_{j=1}^{N} \left(\sum_{k=1}^{N} u(Z^j) L_j(z_k) - \mathbb{E}[u_N(Z)] \right)^2 f_Z(z_k)\, w_k, \qquad (4.5)\end{aligned}$$

where w_k and z_k are the quadrature weights and nodes, respectively. In order to simplify the variance equation (4.5), the quadrature points and collocation points are chosen to be same points, $z_j = Z^j$. Hence, the variance equation (4.5) reduces to

$$\begin{aligned}\mathrm{Var}[u_N(Z)] &= \sum_{j=1}^{N} \left(\sum_{k=1}^{N} u(Z^j) \delta_{jk} - \mathbb{E}[u_N(Z)] \right)^2 f_Z(Z^j) w_k \\ &= \sum_{j=1}^{N} \big(u(Z^j) - \mathbb{E}[u_N(Z)] \big)^2 f_Z(Z^j) w_j.\end{aligned}$$

Consider again the stochastic KdV equation under the effect of only additive noise (4.1), where β is the fully-correlated time dependent Gaussian noise, i.e., $\beta(Z^j, t) = Z^j$ with its amplitude ϵ, then Gauss-Hermite quadrature rules should be applied in order to establish corresponding probability space. The general procedure of SC method for solving the stochastic KdV equation (4.1) is outlined as:

1. Compute all quadrature points and weights $\{z_i, w_i\}$, $i = 1, \ldots, N$;
2. Evaluate random coefficient $\beta(t, z_i)$ at the quadrature nodes z_i;
3. Solve deterministic system to ensemble realizations $\{u(z_i)\}$ at the each node z_i, $i = 1, \ldots, N$;

4. Compute the expectation $\mathbb{E}[u_N(Z)] = \sum_{j=1}^{N} u(z_j) f_Z(z_j) w_j$;

5. Compute the variance $\mathrm{Var}[u_N(Z)] = \sum_{i=1}^{N} \left(u(z_i) - \mathbb{E}[u_N(Z)]\right)^2 f_Z(z_i) w_i$ from the set of realizations.

The other statistical information of the SC solution can be computed with this interpolation approach as well. Besides interpolation, there is another well established approach for the stochastic collocation methods which is called pseudo-spectral approach. The pseudo-spectral approach uses a discrete projection, about which any further information can be found in [52, 53].

Stochastic collocation method can also be extended to the case that the model equation has multivariate stochastic inputs by using tensor product approach. Let $Z = \{Z_1, \ldots, Z_d\}$ be a d-dimensional random vector and $\mathbf{Z} = Z_1 \otimes \ldots \otimes Z_d$ be the collocation points, where $Z_i = \{Z_i^j\}_{j=1}^{N_i}$ is the set of collocation points in the i^{th} random dimension, then the solution $u(\mathbf{Z})$ is approximated as

$$u(Z) \approx u_N(\mathbf{Z}) = u_{N_1}(Z_1) \otimes \cdots \otimes u_{N_d}(Z_d), \tag{4.6}$$

where $u_{N_i}(Z_i)$ is a one-dimensional SC approximation (4.2) for only i^{th} random dimension and $N = N_1 \times \cdots \times N_d$. Hence, the d-dimensional approximation (4.6) can be clarified as

$$u_N(\mathbf{Z}) = \sum_{j_1=1}^{N_1} \cdots \sum_{j_d=1}^{N_d} u(Z_1^{j_1}, \ldots, Z_d^{j_d}) L_{j_1}^1(Z_1) \ldots L_{j_d}^d(Z_d), \tag{4.7}$$

where L^i is the Lagrange polynomial for the i^{th} stochastic dimension. The statistical information, i.e., expectation and variance, can be accomplished in the same way as done in the one-dimensional case by using a multi-dimensional quadrature rule [28].

Let us choose a set of collocation points $\mathbf{Z} = Z_1 \otimes Z_2$, where $Z_1 = \{Z_1^1, \ldots, Z_1^{N_1}\}$ and $Z_2 = \{Z_2^1, \ldots, Z_2^{N_2}\}$ in the multidimensional random space of the stochastic KdV equation with two random input (3.5). Then, for each collocation point $j = 1, \ldots, N$, the following model equation is required to be solved

$$u_t(x,t,\mathbf{Z}^j) + (-6u(x,t,\mathbf{Z}^j) + \delta\alpha(Z_1^j, x))u_x(x,t,\mathbf{Z}^j) + u_{xxx}(x,t,\mathbf{Z}^j) = \epsilon\beta(Z_2^j, t),$$

where α and β are the fully-correlated space and time dependent Gaussian noises with their amplitude δ and ϵ, respectively. The general procedure of multi-dimensional SC method for solving a model equation is summarized as:

1. Compute the sets of quadrature points $z_1 = \{{z_1}^i\}_{i=1}^{N_1}$, $z_2 = \{{z_2}^j\}_{j=1}^{N_2}$;
2. Compute the sets of combined weights $w_1 = \{f_{Z_1}({z_1}^j){w_1}^i\}_{i=1}^{N_1}$, $w_2 = \{f_{Z_2}({z_2}^j)$ ${w_2}^j\}_{j=1}^{N_2}$, where ${w_i}^j$ is quadrature weight and $f_{Z_i}(z)$ is distribution weight of the random variable Z_i ;
3. Evaluate random coefficients $\alpha(x, {z_1}^i)$, $\beta(t, {z_2}^j)$ at each quadrature nodes ${z_1}^i$, ${z_2}^j$;
4. Compute the tensor product of the sets of the quadrature nodes $\mathbf{z} = z_1 \otimes z_2$ and sets of combined weights $\mathbf{w} = w_1 \otimes w_2$;
5. Solve deterministic system for $i = 1, \ldots, N$ in order to obtain a set of realization where $N = N_1 \times N_2$;
6. Compute the expectation $\mathbb{E}[u_N(\mathbf{Z})] = \sum_{j=1}^{N} u(\mathbf{z}_j)\mathbf{w}_j$;
7. Compute the variance $\mathrm{Var}[u_N(\mathbf{Z})] = \sum_{i=1}^{N} \left(u(\mathbf{z}_i) - \mathbb{E}[u_N(\mathbf{Z})]\right)^2 \mathbf{w}_i$ from the set of realizations.

The multi-dimensional stochastic collocation methods yield a numerical challenge that the total number of collocation points $N = N_1 \times \cdots \times N_d$ outgrows as the random dimension d increases, which is widely known as curse of dimensionality. In order to overcome this problem different strategies are proposed such as sparse grids [43, 53] and adaptive sparse grids [58].

4.2 Intrusive Methods

Intrusive methods rely on modifying the stochastic model with its corresponding random field in order to seek a spectral representation for the solution of a PDE with random inputs by modifying the initial model problem. In other words, these methods benefit from random and deterministic dimensions, simultaneously. The modification

of the initial stochastic model results in a system of equations that are analytically and numerically different from the usual deterministic system to be solved. Despite the extra analytical and numerical operations, the intrusive methods provide spectral, exponentially fast, convergence [28].

4.2.1 Stochastic Galerkin Method

The stochastic Galerkin (SG) method is a major intrusive spectral method that appreciates (generalized) polynomial chaos (PC) expansion to form a weighted residual formalism. As in the classic (deterministic) Galerkin method, the idea behind the SG method is to seek a solution for the model equation such that the residue is orthogonal to the space of polynomials, and then this notion leads that the SG solution has optimal error in the L^2 sense [28]. In addition, the discretization procedure performed in the SG method results in a system of coupled deterministic differential equations in which any spatial/time discretization scheme can be used.

Let model equation be the stochastic KdV equation (3.6) with a time dependent noise

$$u_t(x,t,\omega) - 6u(x,t,\omega)u_x(x,t,\omega) + u_{xxx} = \epsilon\beta(\omega,t), \tag{4.8}$$

where the initial condition is deterministic, periodic function and $\beta(t,\omega)$ is a Gaussian random variable with an amplitude ϵ. Since $\beta(t,\omega) \sim \mathcal{N}(\mu,\sigma)$ the PC basis $\{\Psi\}_{i=0}^{N}$ is chosen to be Hermite polynomials, see Table 2.1. The random input $\beta(t,\omega)$ can be expressed with N^{th} degree PC approximation

$$\beta_N(t,\omega) = \sum_{i=0}^{N} \beta_i \Psi_i(\xi(\omega)), \tag{4.9}$$

where $\beta_0 = \mu$, $\beta_1 = \sigma$, and $\beta_i = 0$ for $i > 1$. Moreover, the solution $u(x,t,\omega)$ is represented by a PC approximation

$$u(x,t,\omega) \approx u_N(x,t,\omega) = \sum_{i=0}^{N} u_i(x,t)\Psi_i(\xi(\omega)), \tag{4.10}$$

where the total number of PC basis is determined by the dimension M of the random

vector ξ and the highest order K of the basis polynomials Ψ_i such that

$$N+1=\frac{(M+K)!}{M!K!}.$$

Then, the PC approximation of the solution (4.10) and random input approximation (4.9) are substituted into the model equation (4.8), which results in

$$\sum_{i=0}^{N}(u_i)_t\Psi_i-6\sum_{i=0}^{N}\sum_{j=0}^{N}u_i(u_j)_x\Psi_i\Psi_j+\sum_{i=0}^{N}(u_i)_{xxx}\Psi_i=\epsilon\sum_{i=0}^{N}\beta_i\Psi_i, \tag{4.11}$$

where $u_i\equiv u_i(x,t)$ and $\Psi_i\equiv\Psi_i(\xi)$ for $i=0,\ldots,N$. Multiplying the equation (4.11) by a test function Ψ_k for $k=0,\ldots,N$ and taking (weighted) inner product, Galerkin procedure yields $(M+1)$ coupled deterministic equations

$$\sum_{i=0}^{N}(u_i)_t\left\langle\Psi_i\Psi_k\right\rangle-6\sum_{i=0}^{N}\sum_{j=0}^{N}u_i(u_j)_x\left\langle\Psi_i\Psi_j\Psi_k\right\rangle+\sum_{i=0}^{N}(u_i)_{xxx}\left\langle\Psi_i\Psi_k\right\rangle=\epsilon\sum_{i=0}^{N}\beta_i\left\langle\Psi_i\Psi_k\right\rangle.$$

With the help of orthogonality of PC basis polynomials, we obtain

$$\sum_{i=0}^{N}(u_i)_t\left\langle\Psi_k^2\right\rangle\delta_{ik}-6\sum_{i=0}^{N}\sum_{j=0}^{N}u_i(u_j)_x\left\langle\Psi_i\Psi_j\Psi_k\right\rangle+\sum_{i=0}^{N}(u_i)_{xxx}\left\langle\Psi_k^2\right\rangle\delta_{ik}=\epsilon\sum_{i=0}^{N}\beta_i\left\langle\Psi_k^2\right\rangle\delta_{ik}.$$

By the definition of Kronecker delta function δ_{ik}, we acquire

$$(u_k)_t\left\langle\Psi_k^2\right\rangle-6\sum_{i=0}^{N}\sum_{j=0}^{N}u_i(u_j)_x\left\langle\Psi_i\Psi_j\Psi_k\right\rangle+(u_k)_{xxx}\left\langle\Psi_k^2\right\rangle=\epsilon\beta_k\left\langle\Psi_k^2\right\rangle.$$

Then, the system of equations is simplified to

$$(u_k)_t-\frac{6}{\gamma_k}\sum_{i=0}^{N}\sum_{j=0}^{N}u_i(u_j)_xe_{ijk}+(u_k)_{xxx}=\epsilon\beta_k, \tag{4.12}$$

where $e_{ijk}=\left\langle\Psi_i\Psi_j\Psi_k\right\rangle$ and $\gamma_k=\left\langle\Psi_k^2\right\rangle$ for $i,j,k=0,\ldots,N$. Further, the system can be rewritten in a compact form by defining $(N+1)\times(N+1)$ matrices $\mathbf{E}^{(i)}$ whose elements are

$$\mathbf{E}_{kj}^{(i)}=\frac{1}{\gamma_k}e_{ijk}$$

and a $(N+1)$ vector $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_N)^T$ for each $j, k = 0, \ldots, N$. Hence the system of equations (4.12) can be defined conveniently in a matrix form as

$$\mathbf{u}_t - 6\left[\sum_{i=0}^{N} \bar{\mathbf{u}}^{(i)}\mathbf{E}^{(i)}\right]\mathbf{u}_x + \mathbf{u}_{xxx} = \epsilon\,\boldsymbol{\beta}, \tag{4.13}$$

where $\mathbf{u} = (u_0, \ldots, u_N)^T$ and

$$\bar{\mathbf{u}}^{(i)} = \begin{pmatrix} u_i & 0 & 0 & 0 \\ 0 & \ddots & 0 & \vdots \\ \vdots & 0 & u_i & 0 \\ 0 & \cdots & 0 & u_i \end{pmatrix}.$$

If the initial condition is assumed to be a deterministic function $u(x, 0, \omega) = u(x)$, the initial condition of the system of equation becomes vector of $N+1$ length such that $\mathbf{u}(x, 0) = (u(x), 0, \ldots, 0)$ by applying Galerkin procedure $\langle u(x)\Phi_k\rangle = u(x)$ for $k = 0, \ldots, N$. The general procedure of the SG method for solving the KdV equation with random input data (4.8) is summarized as follows:

1. Compute the random coefficient vector $\boldsymbol{\beta}$;
2. Compute the matrices of PC basis $\mathbf{E}^{(i)}$ for $i = 0, \ldots, N$;
3. Compute the initial condition $\mathbf{u}(x, 0)$;
4. Construct the system by using PC and spatial discretization matrices;
5. Solve the system by using a numerical solver to obtain the solution vector $\mathbf{u}$;
6. Compute the expectation $\mathbb{E}[u_N(x, t, \omega)] = u_0$;
7. Compute the variance $\mathrm{Var}[u_N(x, t, \omega)] = \sum_{i=1}^{N} \gamma_i u_i^2$.

Now we consider the stochastic KdV equation with multiplicative noise (3.16) as our second model

$$u_t + (-6u + \delta\alpha(x, \omega))u_x + u_{xxx} = 0, \tag{4.14}$$

where initial condition is deterministic, periodic function and $\alpha(x, \omega)$ is a Gaussian random variable with an amplitude δ. As additive noise model, the random input α

and the solution u are defined by the N^{th} PC approximations as

$$\begin{aligned} \alpha_N(x,\omega) &= \sum_{i=0}^{N} \alpha_i \Psi_i(\xi), \\ u_N(x,t,\omega) &= \sum_{i=0}^{N} u_i(x,t)\Psi_i(\xi), \end{aligned} \tag{4.15}$$

where $\{\Psi_i\}_{i=0}^N$ is set of Hermite polynomials since $\alpha(x,\omega) \sim \mathcal{N}(\mu,\sigma)$. Then, substituting the expansions (4.15) into equation (4.14), we obtain

$$\sum_{i=0}^{N}(u_i)_t\Psi_i - 6\sum_{i=0}^{N}\sum_{j=0}^{N} u_i(u_j)_x\Psi_i\Psi_j + \delta\sum_{i=0}^{N}\sum_{j=0}^{N}\alpha_j(u_i)_x\Psi_i\Psi_j + \sum_{i=0}^{N}(u_i)_{xxx}\Psi_i = 0.$$

Last, an application of the Galerkin procedure results in a system of $N+1$ equations

$$(u_k)_t - \frac{6}{\gamma_k}\sum_{i=0}^{N}\sum_{j=0}^{N} u_i(u_j)_x\, e_{ijk} + \frac{\delta}{\gamma_k}\sum_{i=0}^{N}\sum_{j=0}^{N}\alpha_i(u_j)_x\, e_{ijk} + (u_k)_{xxx} = 0, \tag{4.16}$$

where $e_{ijk} = \langle \Psi_i\Psi_j\Psi_k\rangle$ and $\gamma_k = \langle \Psi_k^2\rangle$ for $i,j,k = 0,1,\ldots,N$. The system of equation (4.16) can be stated with a matrix notation such that

$$\mathbf{u}_t - 6\left[\sum_{i=0}^{N}\bar{\mathbf{u}}^{(i)}\mathbf{E}^{(i)}\right]\mathbf{u}_x + \delta\mathbf{A}\,\mathbf{u}_x + \mathbf{u}_{xxx} = 0, \tag{4.17}$$

where $\mathbf{A}$ and $\mathbf{E}$ are $(N+1)\times(N+1)$ matrices which are defined as

$$\begin{aligned} \mathbf{A}_{kj} &= \frac{1}{\gamma_k}\sum_{i=0}^{N}\alpha_i e_{ijk}, \\ \mathbf{E}^{(i)}_{kj} &= \frac{1}{\gamma_k}e_{ijk}, \end{aligned}$$

and a vector $\mathbf{u} = (u_0,\ldots,u_N)^T$. For the stochastic KdV equation (3.5) under the effect of both additive and multiplicative noises, stochastic Galerkin system can simply be derived as a combination of both cases, (4.13) and (4.17). Moreover, statistical information, i.e., expectation and variance, of the SG solution can be computed by using formulas derived in Section 2.5.

Remark 4.1. The terms $\mathbf{u}_x$, $\mathbf{u}_{xxx}$ in the equations (4.13) (4.17) are defined as a vector consisting of first and third order derivatives of the PC coefficients in the solution vector $\mathbf{u}$. The first and third order derivatives of PC coefficients are required to be discretized with a spatial discretization techniques, as done in Section 5.1.

The stochastic Galerkin method is based on minimizing the residue of the stochastic governing equation. The error occurring in implementation of the SG method is optimal in mean-square sense which means that the SG method is preferred over intrusive methods for getting more accurate results [54]. However, implementation of the SG method requires analytical derivation of a large system of equation where the coefficients of the PC expansion are coupled. Moreover, solving the resulting coupled system of equations can be very cumbersome due to the large matrix system. . Decoupling strategies [51] can be utilized into the SG system since solving each equation for the PC coefficients is easier than solving coupled system or preconditioning [36] can be also used in order to overcome these numerical challenges.

CHAPTER 5

NUMERICAL RESULTS

In this chapter, the influences of two different types of uncertainty, which are additive (time dependent) and multiplicative (space dependent) noises, imposed on Korteweg-de Vries equation will be investigated. The analytical solution of stochastic KdV equation resulting from time dependent noise exists, whereas there does not exist any analytical solution of stochastic KdV equations containing space dependent noise. Therefore, special numerical methods are needed to perform the discretization of random dimension of the stochastic KdV equations. Our attention will be mainly on stochastic Galerkin (SG) method, and then SG method will be compared to Monte Carlo (MC) simulation and stochastic collocation (SC) methods in terms of their convergence rates and efficiency.

After the PDE with random input data is transformed into a deterministic form, we need to discretize spatial and temporal dimensions by any appropriate discretization schemes. In this thesis, the finite (central) difference (FD) and local discontinuous Galerkin (LDG) schemes are chosen to discretize spatial dimensions, whereas the weighted average (theta) method will be used for discretization of the temporal dimension. Moreover, the rational deferred correction method will be introduced to improve accuracy of the time stepping scheme. Now, the spatial and temporal discretization techniques will firstly be introduced upon the deterministic KdV equation. Then, fully discretized stochastic Galerkin system matrices will be constructed, and finally the computational results will be presented.

5.1 Spatial Domain Discretization

5.1.1 Finite Difference Method

The finite difference (FD) is one of the simplest and the oldest numerical method to approximate differential equation in the spatial dimension. The finite difference method approximates the solution of a differential equation by replacing the derivatives in the equation with their primitive depictions at each grid point. First the spatial domain $\Omega = [0, L]$ is needed to be partitioned in order to constitute a finite grid such that

$$x_i = i\Delta x, \qquad i = 0, 1, \ldots, M_x,$$

where $x_0 = 0$, $x_{M_x} = L$, and Δx is the spatial step size, i.e., $\Delta x = \frac{L}{M_x}$. Then, the central difference formulation of the first order derivative is given as

$$\frac{\partial u(x_i, t)}{\partial x} \approx \frac{u(x_{i+1}, t) - u(x_{i-1}, t)}{2\Delta x} \tag{5.1}$$

and the central difference formulation for approximating the third order derivative is presented as

$$\frac{\partial^3 u(x_i, t)}{\partial x^3} \approx \frac{u(x_{i+2}, t) - 2u(x_{i+1}, t) + 2u(x_{i-1}, t) - u(x_{i-2}, t)}{2\Delta x^3} \tag{5.2}$$

for each spatial grid points x_i, $i = 0, 1, \ldots, M_x$. Last, the finite difference formulations (5.1) and (5.2) are required to be substituted into the deterministic KdV equation (3.3) with periodic boundaries $u(0, t) = u(L, t)$, which results in, $i = 0, 1, \ldots, M_x$,

$$\frac{\partial u}{\partial t} + \frac{u(x_{i+2}, t) - 2u(x_{i+1}, t) + 2u(x_{i-1}, t) - u(x_{i-2}, t)}{2\Delta x^3} + \frac{1}{2}\frac{u^2(x_{i+1}, t) - u^2(x_{i-1}, t)}{2\Delta x} = 0.$$

Now, we will construct a semi–discrete matrix form by using the system of equations resulting from implementation of finite difference approximations. Let $\mathbf{x} = (x_0, \ldots, x_{M_x - 1})$ be the vector of spatial grid and $u(\mathbf{x}, t) \equiv (u(x_0, t), \ldots, u(x_{M_x - 1}, t))^T$ be the solution vector, then the semi-discrete form is

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} + D_{xxx} u(\mathbf{x}, t) + \frac{1}{2} D_x u^2(\mathbf{x}, t) = 0, \tag{5.3}$$

where D_x and D_{xxx} are a $(M_x \times M_x)$ matrices which are stencils of the first and third order derivatives in the forms of

$$D_x = \frac{1}{2\Delta x} \begin{bmatrix} 0 & 1 & & & -1 \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \\ 1 & 0 & & -1 & 0 \end{bmatrix} \tag{5.4}$$

and

$$D_{xxx} = \frac{1}{2(\Delta x)^3} \begin{bmatrix} 0 & -2 & 1 & & & -1 & 2 \\ 2 & 0 & -2 & 1 & & & -1 \\ -1 & \ddots & \ddots & \ddots & & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & 1 \\ 1 & & & \ddots & \ddots & \ddots & -2 \\ -2 & 1 & & & -1 & 2 & 0 \end{bmatrix}, \tag{5.5}$$

respectively.

5.1.2 Local Discontinuous Galerkin Method

Although the implementation of finite difference method is simple, it is ill-suited to deal with complex geometries, both in terms of general computational domains and internal discontinuities. Therefore, standard continuous finite element method [3] become popular due to its efficiency and high-order convergence rate. On the other hand, it does not satisfy local mass conservation, which is a crucial property in hyperbolic problems [9, 44].

In recent years, discontinuous Galerkin (DG) methods have been developed rapidly and applied into various areas such as computational fluid dynamics, environmental modeling, reservoir simulation and groundwater aquifer simulation, etc. Though these methods are known since the 1970s, much attention has been paid only in the past few years due to the availability of cheap computing resources. We would like to refer to [1, 18, 37] for more details of about discontinuous Galerkin methods. Local discontinuous Galerkin (LDG) method is one of several discontinuous Galerkin

methods which are being vigorously studied, especially as applied to hyperbolic equations because of their applicability to a wide range of problems and their properties of local conservativity and higher degree of locality. In addition, they may be advantageous because of the ease with which the method handles hanging nodes, elements of general shapes, and local spaces of different types.

The fundamental idea behind the LDG methods is to rewrite the differential equation with higher order derivatives into a first-order system. In the LDG method, the local conservativity holds because the conservation laws are weakly enforced element by element. In order to do that, suitable discrete approximations of the traces of the fluxes on the boundary elements are provided by the so-called numerical fluxes. These numerical fluxes enhance the stability of the method, and hence, the quality of its approximation. This is why the LDG method is strongly related to stabilized mixed finite elements. The stabilization is associated with the jump of the approximate solution across the element boundaries, see [1, 8].

Let $f(u(x,t)) = 3u^2$ and $p(x,t)$, $q(x,t)$ be auxiliary variables, then a first order system of Korteweg-de Vries equation (3.3) can be stated as

$$q - \frac{\partial u}{\partial x} = 0, \qquad p - \frac{\partial q}{\partial x} = 0, \qquad \frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} + \frac{\partial f(u)}{\partial x} = 0. \tag{5.6}$$

Let $\mathcal{T}_h$ denote a partition of the spatial domain $\Omega = [a,b]$, which is divided into N computational interval whose step sizes are $h_n = x_{n+1} - x_n$. Assuming that $h = \max\limits_{0\le n\le N} h_n$, a finite dimensional space $\mathcal{D}_k$ is defined as a space of piecewise polynomials of degree k in each cell such that

$$\mathcal{D}_k(\mathcal{T}_h) = \{v : v|_{I_n} \in \mathcal{P}_k(I_n), \forall n = 0, ..., N-1\}.$$

One sided-values of a scalar quantity $v = v(x)$ can be stated by

$$v(x_n^+) = \lim_{\epsilon\to 0} v(x_n + \epsilon) \qquad \text{and} \qquad v(x_n^-) = \lim_{\epsilon\to 0} v(x_n - \epsilon).$$

Then, the jump and average of v at the end points of each subintervals I_n are

$$[v(x_n)] = v(x_n^-) - v(x_n^+), \qquad [v(x_0)] = -v(x_0^+), \qquad [v(x_N)] = v(x_N^-),$$
$$\{v(x_n)\} = \frac{1}{2}(v(x_n^-) + v(x_n^+)), \qquad \{v(x_0)\} = v(x_0^+), \qquad \{v(x_N)\} = v(x_N^-),$$

respectively.

Now, we can apply formal Galerkin projection to the first order system (5.6) on each subinterval I_n in order to construct a variational form. For each test function v, w, $z \in \mathcal{D}_k(\mathcal{T}_h)$, we have

$$\int_{x_n}^{x_{n+1}} q\, v dx + \int_{x_n}^{x_{n+1}} u\, v' dx - \left[\widehat{u}(x_{n+1})v(x_{n+1}^-) - \widehat{u}(x_n)v(x_n^+)\right] = 0, \tag{5.7}$$

$$\int_{x_n}^{x_{n+1}} p\, w dx + \int_{x_n}^{x_{n+1}} q\, w' dx - \left[\widehat{q}(x_{n+1})w(x_{n+1}^-) - \widehat{q}(x_n)w(x_n^+)\right] = 0, \tag{5.8}$$

$$\begin{aligned}\int_{x_n}^{x_{n+1}} u_t\, z dx - \int_{x_n}^{x_{n+1}} p\, z' dx + \left[\widehat{p}(x_{n+1})z(x_{n+1}^-) - \widehat{p}(x_n)z(x_n^+)\right] \\ - \int_{x_n}^{x_{n+1}} f(u)z' dx + \left[\widehat{f}(u(x_{n+1}))z(x_{n+1}^-) - \widehat{f}(u(x_n))z(x_n^+)\right] = 0,\end{aligned} \tag{5.9}$$

where $\widehat{u}(x_n)$, $\widehat{p}(x_n)$, $\widehat{q}(x_n)$, $\widehat{f}(u(x_n))$ denote numerical fluxes. They have to be suitably defined in order to ensure the stability of the method and to enhance its accuracy. The numerical fluxes related to the dispersive term u_{xxx} are chosen to be based on upwind-scheme such that

$$\widehat{u}(x_n) = \begin{cases} u(x_n^-), & n = 1, ..., N-1, \\ u(x_0^-), & n = 0, \\ u(x_N^+), & n = N, \end{cases} \tag{5.10}$$

and

$$\widehat{q}(x_n) = \begin{cases} q(x_n^+), & n = 1, ..., N-1, \\ q(x_0^+), & n = 0, \\ q(x_N^-), & n = N, \end{cases} \tag{5.11}$$

and

$$\widehat{p}(x_n) = \begin{cases} p(x_n^+), & n = 1, ..., N-1, \\ p(x_n^+), & n = 0, \\ p(x_n^-), & n = N. \end{cases} \tag{5.12}$$

The upwind-scheme guarantees stability and convergence of the LDG method for KdV type equations [55]. On the other hand, the convective flux $\widehat{f}(u(x_n))$ is chosen simply as Lax-Friedrichs flux,

$$\widehat{f}(u(x_n)) = \frac{1}{2}(f(u(x_n^-)) + f(u(x_n^+)) - \alpha(u(x_n^+) - u(x_n^-)), \tag{5.13}$$

where $\alpha = \max_u |f'(u)|$. Since all the fluxes are defined, we can construct a variational formulation explicitly for the first order system of the KdV equation (5.6). On each subinterval $I_n = [x_n, x_{n+1}]$, $n = 0, 1, \ldots, N-1$, the discrete solutions are defined by

$$q_h = \sum_{j=0}^{k} q_j^n(t)\phi_j^n(x), \tag{5.14}$$

$$p_h = \sum_{j=0}^{k} p_j^n(t)\phi_j^n(x), \tag{5.15}$$

$$u_h = \sum_{j=0}^{k} u_j^n(t)\phi_j^n(x), \tag{5.16}$$

where $\phi_i^n \in \mathcal{D}_k(\mathcal{T}_h)$ are basis polynomials and k is maximum degree of basis polynomials. Find q_h, p_h, and $u_h \in \mathcal{D}_k(\mathcal{T}_h)$, for all test functions v, w, $z \in \mathcal{D}_k(\mathcal{T}_h)$, such that

$$\sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} q_h\, v dx + \sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} u_h\, v' dx - \sum_{n=0}^{N-1} \left[\widehat{u}_h(x_{n+1})v(x_{n+1}^-) - \widehat{u}_h(x_n)v(x_n^+)\right] = 0, \tag{5.17}$$

$$\sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} p_h\, w dx + \sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} q_h\, w' dx - \sum_{n=0}^{N-1} \left[\widehat{q}_h(x_{n+1})w(x_{n+1}^-) - \widehat{q}_h(x_n)w(x_n^+)\right] = 0, \tag{5.18}$$

$$\sum_{n=0}^{N-1}\int_{x_n}^{x_{n+1}}(u_h)_t\, z dx - \sum_{n=0}^{N-1}\int_{x_n}^{x_{n+1}} p_h\, z' dx + \sum_{n=0}^{N-1}\left[\widehat{p}_h(x_{n+1})z(x_{n+1}^-) - \widehat{p}_h(x_n)z(x_n^+)\right] - \sum_{n=0}^{N-1}\int_{x_n}^{x_{n+1}} f(u_h)z' dx + \sum_{n=0}^{N-1}\left[\widehat{f}(u_h(x_{n+1}))z(x_{n+1}^-) - \widehat{f}(u_h(x_n))z(x_n^+)\right] = 0. \quad (5.19)$$

Putting the discrete solutions (5.14)–(5.16) and the numerical fluxes (5.10)–(5.12) into (5.17)–(5.19) and summing over all subintervals I_n, we obtain the following system:

$$\mathbf{M}\, q(t) + \mathbf{L}\, u(t) - \mathcal{O}^1\, u(t) = \mathcal{J}_D^1, \quad (5.20)$$

$$\mathbf{M}\, p(t) + \mathbf{L}\, q(t) - \mathcal{O}^2\, q(t) - \mathcal{H}^2\, q(t) = 0, \quad (5.21)$$

$$\mathbf{M}\, \partial u(t) + \mathcal{O}^3\, u(t) + \mathcal{H}^3\, u(t) - \mathbf{L}\, u(t) + \mathcal{O}^{\mathrm{NL}}(u) - \mathbf{NL}(u) - \mathcal{J}_D^{\mathrm{NL}} = 0, \quad (5.22)$$

where mass $\mathbf{M}_n$, advection $\mathbf{L}_n$, nonlinear $\mathbf{NL}_n$ matrices, for each subinterval I_n, are

$$(\mathbf{M}_n)_{ij} = \int_{I_n} \phi_j^n \phi_i^n dx,$$
$$(\mathbf{L}_n)_{ij} = \int_{I_n} \phi_i^n (\phi_j^n)' dx,$$
$$(\mathbf{NL_n})_{ij} = \int_{I_n} 3(\phi_j^n(x))^2(\phi_i(x)) dx,$$

interior/exterior boundary matrices $\mathcal{O}_n^1$, $\mathcal{O}_n^2$, $\mathcal{O}_n^3$, $\mathcal{O}_n^{\mathrm{NL}}$, $\mathcal{H}_n^2$, $\mathcal{H}_n^3$ are

$$(\mathcal{O}_n^1)_{ij} = \phi_j(x_n^-)\phi_i(x_n^-) - \phi_j(x_n^-)\phi_i(x_n^+),$$
$$(\mathcal{O}_n^2)_{ij} = \phi_j(x_n^+)\phi_i(x_n^-) - \phi_j(x_n^+)\phi_i(x_n^+),$$
$$(\mathcal{O}_n^3)_{ij} = \phi_j(x_n^+)\phi_i(x_n^-) - \phi_j(x_n^+)\phi_i(x_n^+),$$
$$(\mathcal{H}_n^2)_{ij} = \phi_j(x_N^-)\phi_i(x_N^-) - \phi_j(x_0^+)\phi_i(x_0^+),$$
$$(\mathcal{H}_n^3)_{ij} = \phi_j(x_N^-)\phi_i(x_N^-) - \phi_j(x_0^+)\phi_i(x_0^+),$$
$$(\mathcal{O}^{\mathrm{NL}})_{ij} = \frac{1}{2}[3(\phi_j(x_n^-))^2 + 3(\phi_j(x_n^+))^2 - \alpha(\phi_j(x_n^+) - \phi_j(x_n^-))]\phi_i(x_n^-) - \frac{1}{2}[3(\phi_j(x_n^-))^2 + 3(\phi_j(x_n^+))^2 - \alpha(\phi_j(x_n^+) - \phi_j(x_n^-))]\phi_i(x_n^+),$$

and boundary vectors $\mathcal{J}_D^1$, $\mathcal{J}_D^{\mathrm{NL}}$ are

$$(\mathcal{J}_D^1)_i = u(x_N^-)\phi_i(x_N^-) - u(x_0^-)\phi_i(x_0^+),$$
$$(\mathcal{J}_D^{\mathrm{NL}})_i = \widehat{f}(u(x_N))\phi_j(x_N^-) - \widehat{f}(u(x_0))\phi_j(x_0^+).$$

Hence, a semi-discrete form of the KdV equation (5.6), discretized by LDG method, can be written in a matrix form as

$$\begin{bmatrix} 0 \\ 0 \\ \mathbf{M}\partial_t \mathrm{u} \end{bmatrix} + \begin{bmatrix} \mathbf{M} & & \mathbf{L} - \mathcal{O}^1 \\ \mathbf{L} - \mathcal{O}^2 - \mathcal{H}^2 & \mathbf{M} & \\ & \mathcal{O}^3 + \mathcal{H}^3 - \mathbf{L} & \end{bmatrix} \begin{bmatrix} \mathrm{q} \\ \mathrm{p} \\ \mathrm{u} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \mathcal{O}^{\mathrm{NL}}(\mathrm{u}) - \mathbf{NL}(\mathrm{u}) \end{bmatrix} = \begin{bmatrix} \mathcal{J}_D^1 \\ 0 \\ -\mathcal{J}_D^{\mathrm{NL}} \end{bmatrix}, \quad (5.23)$$

where q, p, u are the discretized solutions. Let discrete solution vector $\mathbf{Y} \in \mathbb{R}^{3N(K+1)}$, matrix $\mathbf{A}, \mathbf{W} \in \mathbb{R}^{3N(K+1)\times 3N(K+1)}$ and right-hand-side vector $\mathbf{J} \in \mathbb{R}^{3N(K+1)}$ be

$$\mathbf{Y} = \begin{bmatrix} \mathrm{q} \\ \mathrm{p} \\ \mathrm{u} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathbf{M} & & \mathbf{L} - \mathcal{O}^1 \\ \mathbf{L} - \mathcal{O}^2 - \mathcal{H}^2 & \mathbf{M} & \\ & \mathcal{O}^3 + \mathcal{H}^3 - \mathbf{L} & \end{bmatrix},$$
$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbf{M} \end{bmatrix}, \qquad \mathbf{J} = \begin{bmatrix} \mathcal{J}_D^1 \\ 0 \\ -\mathcal{J}_D^{\mathrm{NL}} \end{bmatrix},$$

and the matrix $\mathbf{N} \in \mathbb{R}^{3N(K+1)\times 3N(K+1)}$ corresponding to nonlinear part in KdV equation be

$$\mathbf{N} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathcal{O}^{\mathrm{NL}}(\mathrm{u}) - \mathbf{NL}(\mathrm{u}) \end{bmatrix}.$$

Then, we can rewrite the system of equations (5.23) as

$$\mathbf{W}\partial_t \mathbf{Y} + \mathbf{A}\mathbf{Y} + \mathbf{N}(\mathbf{Y}) = \mathbf{J}. \quad (5.24)$$

5.2 Temporal Domain Discretization

5.2.1 The Weighted Average (Theta) Method

Now, we will discretize temporal space by the weighted average (theta) method. The temporal domain $[0, T]$ is partitioned as

$$t_j = i\Delta t, \qquad j = 0, 1, \ldots, N_t,$$

where $t_0 = 0$, $t_{N_t} = T$, and Δt is the temporal step size, i.e., $\Delta t = \frac{T}{N_t}$. The weighted average (theta) method is formed by assembling forward and backward Euler time discretization schemes with a weight $\theta \in [0, 1]$.

The semi-discretized KdV equation (5.3) obtained by finite difference scheme in space can be approximated in temporal dimension such that

$$\begin{aligned}&\frac{u(\mathbf{x}, t_{j+1}) - u(\mathbf{x}, t_j)}{\Delta t} + (1-\theta)\left[D_{xxx}u(\mathbf{x}, t_j) + \frac{1}{2}D_x u^2(\mathbf{x}, t_j)\right] \\ &+ \theta\left[D_{xxx}u(\mathbf{x}, t_{j+1}) + \frac{1}{2}D_x u^2(\mathbf{x}, t_{j+1})\right] = 0\end{aligned}$$

for a given initial conditions $u(x, t_0)$.

On the other hand, the discrete system (5.24) obtained via the LDG discretization in space can be stated simply as, $\theta \in [0, 1]$,

$$\mathbf{W}\,\frac{\mathbf{Y}^{t+1} - \mathbf{Y}^t}{\Delta t} + (1-\theta)\left[\mathbf{A}\mathbf{Y}^t + \mathbf{N}(\mathbf{Y}^t)\right] + \theta\left[\mathbf{A}\mathbf{Y}^{t+1} + \mathbf{N}(\mathbf{Y}^{t+1})\right] - \mathbf{J} = 0.$$

5.2.2 Rational Deferred Correction Method

Rational deferred correction framework is an efficient post-processing technique that enhance numerical solutions of time-dependent problems. The method can be considered as an extrapolation scheme which improves the accuracy of a low-order integrator iteratively. The idea is to construct representation of residual based on polynomial interpolation of the solution over the time interval $[0, T]$ [16].

Let us assume that a semi-discretized initial value problem (IVP) is defined as

$$\mathbf{u}'(t) = \mathbf{f}(t, \mathbf{u}), \qquad \mathbf{u}(0) = \mathbf{u}_0, \tag{5.25}$$

where $\mathbf{u}$ is a vector valued function, i.e., $\mathbf{u} : [0, T] \to \mathbb{R}^N$ and $\mathbf{u}_0$ is the initial condition. Moreover, approximations $\mathbf{u}_j$ at partitioned temporal domain $0 = t_0 < t_1 < \ldots < t_n = T$ are assumed to be available and the interpolant is stated as

$$\widetilde{\mathbf{u}} = \sum_{j=0}^{n} l_j(t)\mathbf{u}_j,$$

where $l_j(t)$ are differentiable Lagrange functions, i.e., $l_j(t_i) = \sigma_{ij}$.

With the help of the Picard formulation on the semi-discretized IVP (5.25), we obtain

$$\mathbf{u}(t) = \mathbf{u}_0 + \int_0^t \mathbf{f}(\tau, \mathbf{u}(\tau))d\tau. \tag{5.26}$$

Equivalently, the equation (5.26) can be rewritten as

$$\widetilde{\mathbf{u}}(t) + \mathbf{e}(t) = \mathbf{u}_0 + \int_0^t \mathbf{f}(\tau, \widetilde{\mathbf{u}} + \mathbf{e}(\tau))d\tau, \tag{5.27}$$

where $\mathbf{e}(t)$ corresponds to error. Then, the residual can be defined by

$$\mathbf{r}(t) = \mathbf{u}_0 + \int_0^t \mathbf{f}(\tau, \widetilde{\mathbf{u}}(\tau))d\tau - \widetilde{\mathbf{u}}(t). \tag{5.28}$$

Using (5.27) and (5.28), it is easy see that

$$\begin{aligned}\mathbf{e}(t) &= \mathbf{r}(t) + \int_0^t \mathbf{f}(\tau, \widetilde{\mathbf{u}}(\tau) + \mathbf{e}(\tau)) - \mathbf{f}(\tau, \widetilde{\mathbf{u}}(\tau))d\tau \\ &= \mathbf{r}(t) + \int_0^t \mathbf{h}(\tau, \mathbf{e}(\tau))d\tau, \end{aligned} \tag{5.29}$$

where $\mathbf{h}(\tau, \mathbf{e}(\tau)) = \mathbf{f}(\tau, \widetilde{\mathbf{u}} + \mathbf{e}(\tau)) - \mathbf{f}(\tau, \widetilde{\mathbf{u}}(\tau))$. Now we can represent the residual $\mathbf{r}(t)$ by integrating a smooth interpolant $\mathbf{f}_j = \mathbf{f}(t_j, \mathbf{u}_j)$ such that

$$\mathbf{r}(t) \approx \mathbf{u}_0 + \sum_{j=0}^{n} \mathbf{f}_j \left(\int_0^t l_j(\tau)d\tau \right) - \widetilde{\mathbf{u}}(t).$$

The estimation to the residual $\mathbf{r}(t_j)$ at times t_j can be written as

$$[\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_n] = \mathbf{u}_0[1, 1, \ldots, 1] + [\mathbf{f}_0, \mathbf{f}_1, \ldots, \mathbf{f}_n]\mathbf{C} - [\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_n],$$

where $\mathbf{C} \in \mathbb{R}^{(n+1)\times(n+1)}$ is a collocation matrix whose entries are

$$c_{i,j} = \int_0^{t_j} l_i(\tau)d\tau, \qquad i, j = 0, 1, \ldots, n.$$

Since the integrand (5.29) is assumed to be smooth, a time-stepping method for the error $\mathbf{e}(t)$ at the time points t_j can be written by applying a low-order quadrature rule to the integral. Starting with $\mathbf{e}_0 = 0$ an implicit Euler scheme can be stated as

$$\mathbf{e}_j = \mathbf{e}_{j-1} + (\mathbf{r}_j - \mathbf{r}_{r-1}) + (t_j - t_{t-1})\,\mathbf{h}(t_j, \mathbf{e}_j), \qquad j = 1, \ldots, n. \tag{5.30}$$

Finally, one deferred correction step is concluded with an update on the temporal estimation, i.e., $\widetilde{\mathbf{u}} = \widetilde{\mathbf{u}} + \mathbf{e}$. The general procedure of the rational deferred correction scheme is summarized as follows:

1. Initialize $\mathbf{u}_j = \mathbf{u}_0$ for all $j = 0, 1, \ldots, n$;
2. Compute the residuals $\mathbf{r}_j$ of $\widetilde{\mathbf{u}}(t)$ by applying a quadrature rule to
$$\mathbf{r}(t) = \mathbf{u}_0 + \int_0^t \mathbf{f}(\tau, \widetilde{\mathbf{u}}(\tau) d\tau - \widetilde{\mathbf{u}}(t));$$
3. Compute the errors $\mathbf{e}_j$ as defined in (5.30);
4. Update the approximations $\mathbf{u}_j = \mathbf{u}_j + \mathbf{e}_j$;
5. If error criterion is satisfied, stop. Otherwise, go to step 2.

Now, we investigate effects of the rational deferred correction method on linear KdV equation, see Tables 5.1 and 5.2. The solution accuracy is improved significantly by employing only one deferred correction sweep without a need of substantial amount of time steps.

Table 5.1: L_2 and L_∞ errors in solutions, discretized by the LDG scheme, of linear KdV equation at $t = 1$ with $\theta = 1$.

n	$\|\cdot\|_{L^2(\Omega)}$	$\|\cdot\|_{L_\infty(\Omega)}$
10	$8.261e-02$	$5.008e-02$
20	$3.741e-02$	$2.238e-02$
40	$1.650e-02$	$9.768e-03$
80	$7.049e-03$	$4.513e-03$
160	$4.191e-03$	$2.690e-03$

Table 5.2: L_2 and L_∞ errors in solutions, discretized by LDG scheme post-processed with the rational deferred correction method, of linear KdV equation at $t = 1$ with $n = 10$ and $\theta = 1$.

#sweeps	$\|\cdot\|_{L^2(\Omega)}$	$\|\cdot\|_{L_\infty(\Omega)}$
0	$8.261e-02$	$5.008e-02$
1	$4.364e-03$	$3.129e-03$

5.3 Additive Noise

In this section, the numerical results of stochastic KdV equation with time dependent additive noise, formulated as (3.6), will be presented. In the numerical simulations, the random domain is estimated by three different methodologies which are Monte Carlo (MC), stochastic collocation (SC), and stochastic Galerkin (SG) methods, while the spatial domain is discretized with local discontinuous Galerkin (LDG) and the second order finite difference (FD) methods. Moreover, the temporal domain is approximated with weighted average (theta) scheme.

In the following simulations, the number of computational subintervals used in LDG method is $M_x = 800$, the number of spatial grid points used in FD method is $M_x = 1600$ and the temporal domain $[0, 1]$ is discretized with weighted average method where $\theta = 0.5$ and step sizes $\Delta t = 10^{-2}$.

Figure 5.1 shows the exact mean of the KdV equation with the additive noise (3.13) whose amplitude is $\epsilon = 0.5$ and wave speed is $c = 2$.

The error between exact and numerical solutions occurs due to the spatial, the temporal discretizations, and the random space approximations. In this study, L^2 norm is used to compute the error in the numerical solutions for the convergence analysis of the stochastic approximations. Since there exists the analytical solution (3.12) for the stochastic KdV equation with additive noise, the L^2 error in the mean and the variance at time t are computed by

$$\text{Error}(\bar{u}) = \sqrt{\frac{(\bar{u}^t_{\text{num},i} - \bar{u}^t_{\text{exact},i})^2}{M_x}},$$

$$\text{Error}(\text{Var}[u]) = \sqrt{\frac{\left(\text{Var}[u]^t_{\text{num},i} - \text{Var}[u]^t_{\text{exact},i}\right)^2}{M_x}},$$

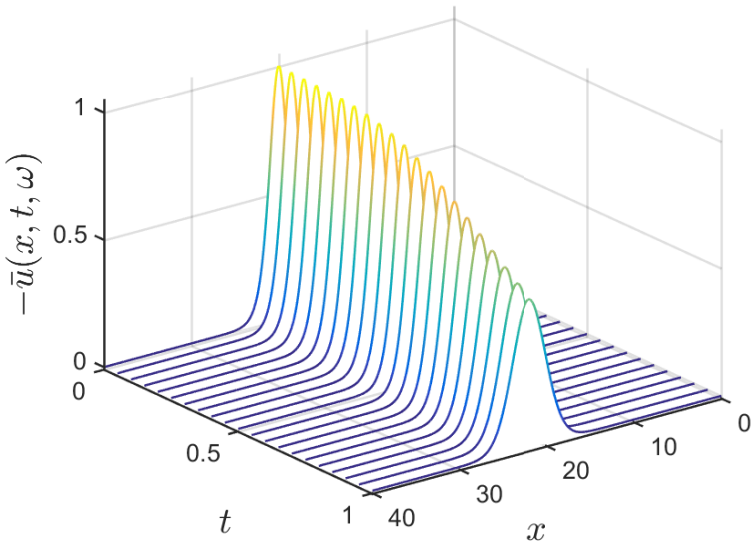

Figure 5.1: The exact mean values of $u(x,t,\omega)$ with the additive noise amplitude $\epsilon = 0.5$ and the wave speed $c = 2$.

where M_x is the total number of computational subinterval used in the LDG method or the total number of grid points used in the FD method. The simulations for the L^2 error are done for different number of grid points and subintervals in order to observe the influence of the error due to the spatial/temporal discretizations on the total error.

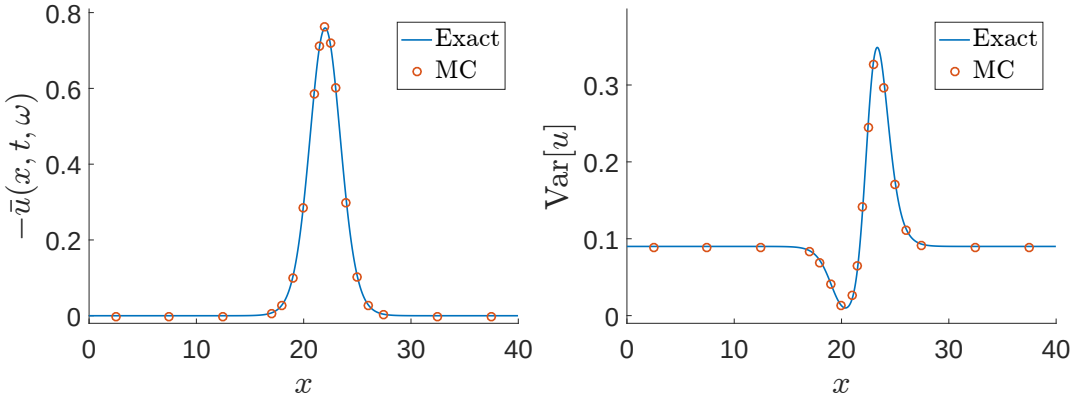

Figure 5.2: The mean value (left) and the variance (right) of $u(x,t,\omega)$ obtained by MC method at time $t = 1$ with the speed $c = 2$ and the amplitude $\epsilon = 0.3$.

In Figure 5.2, the exact and the approximated statistics of the solutions $u(x,t,\omega)$ at $t = 1$ are plotted together. The dots indicate the results approximated from the set

generated by Monte Carlo simulations with 2500 samples. A sufficient agreement between the MC and the exact solutions is observed. Moreover, the theoretical convergence rate of the Monte Carlo simulations $\mathcal{O}(M^{-1/2})$ is observed through Figure 5.3, where M is the number of the samples.

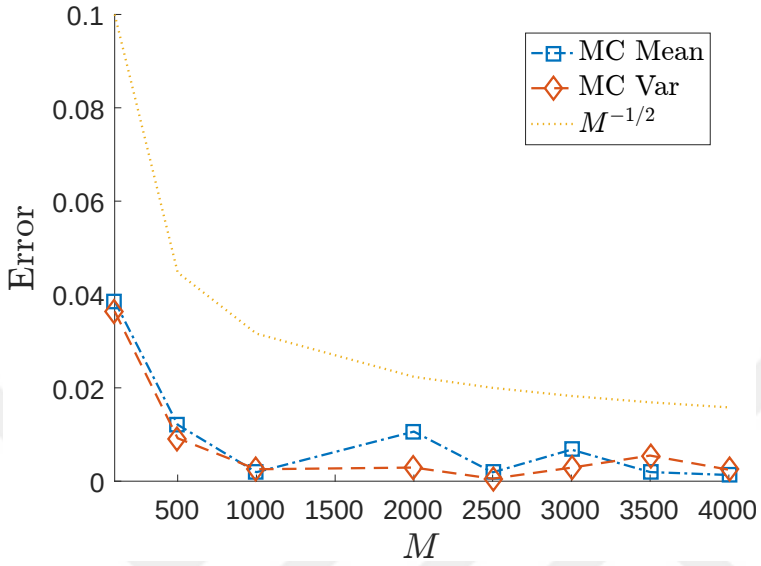

Figure 5.3: Convergence rate of L^2-error of the mean and variance with $\epsilon = 0.3$ at $t = 1$ with respect to sample size M.

Then, the stochastic collocation method with 5 collocation points is implemented to approximate the stochastic KdV equation with additive noise (3.6), the amplitude $\epsilon = 0.3$. The resulting deterministic KdV systems for all collocation points are solved with the second order FD method with $M_x = 1600$. In Figure 5.4, the approximated statistics with SC method are plotted with the exact statistics. When the collocation points N are increased, the SC approximation becomes more accurate.

In Figure 5.5, the exponential convergence rate is realized with respect to the collocation order N. It can be noticed that the convergence rate of the variance is not affected deeply when the number of the spatial grid points is increased, while the convergence rate of the mean is affected. Therefore, it can be inferred that the error in the mean obtained by the SC approximation of random domain can be decreased when the number of M_x is increased.

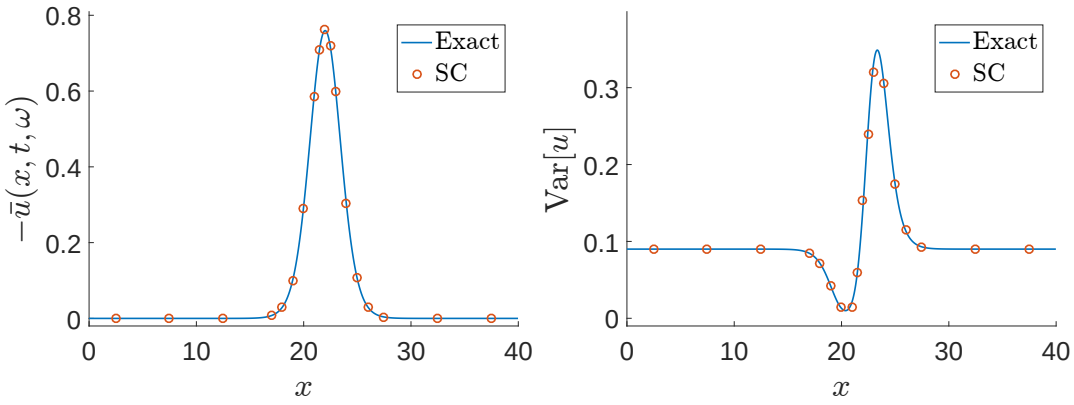

Figure 5.4: The mean value (left) and the variance (right) of $u(x,t,\omega)$ with $c = 2$ and $\epsilon = 0.3$ at time $t = 1$, obtained by the SC method.

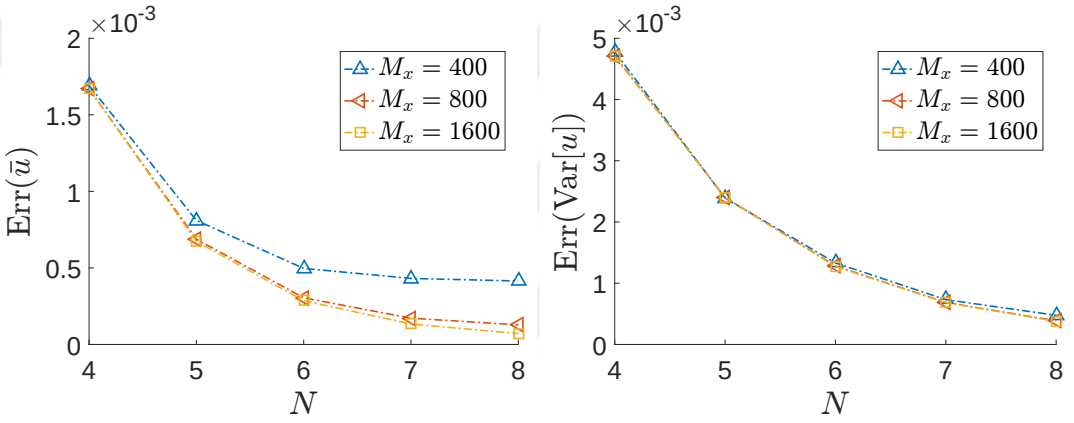

Figure 5.5: Convergence rate of L^2-error in the mean (left) and variance (right) computed with the SC method at $t = 1$, with $\epsilon = 0.3$, as a function of collocation order N for different number of grid points M_x.

The mean and the variance of the KdV equation with additive noise with the amplitude $\epsilon = 0.3$ approximated with the stochastic Galerkin method is displayed in Figure 5.6 at $t = 1$. The SG method is implemented with a fifth order PC expansion of Hermite polynomials, the LDG method with the number of subintervals $M_x = 800$, and the weighted average method with $\theta = 0.5$. When the number of spatial grid is chosen sufficiently high, it can be seen that exponential convergence of the mean and variances of the SG solutions is achieved as in the theory [54], see Figure 5.7. It can also be said that the spatial and the temporal discretization errors can easily dominate the total L^2 error in the mean obtained by SG method. On the contrary, the L^2 error in the variance is not strongly affected by the these discretization errors as obtained

in the SC case.

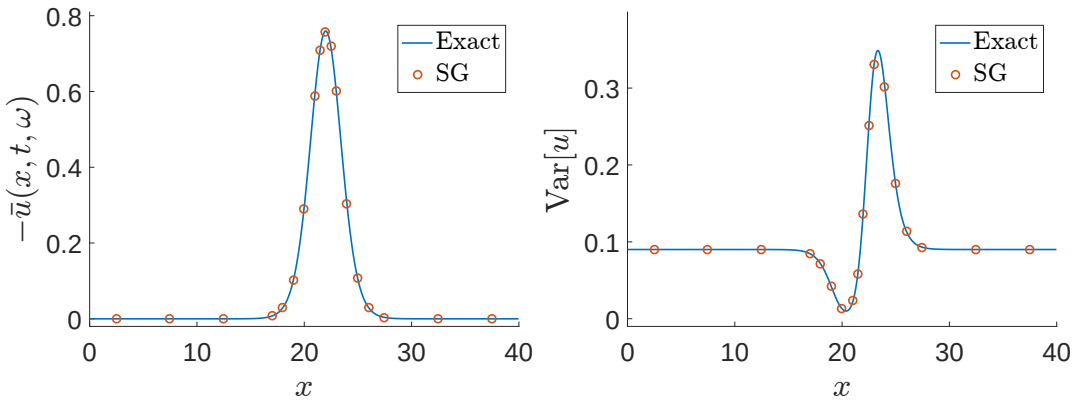

Figure 5.6: The mean value (left) and the variance (right) of $u(x,t,\omega)$ with $c = 1$ and $\epsilon = 0.3$ at time $t = 1$, obtained by the SG method.

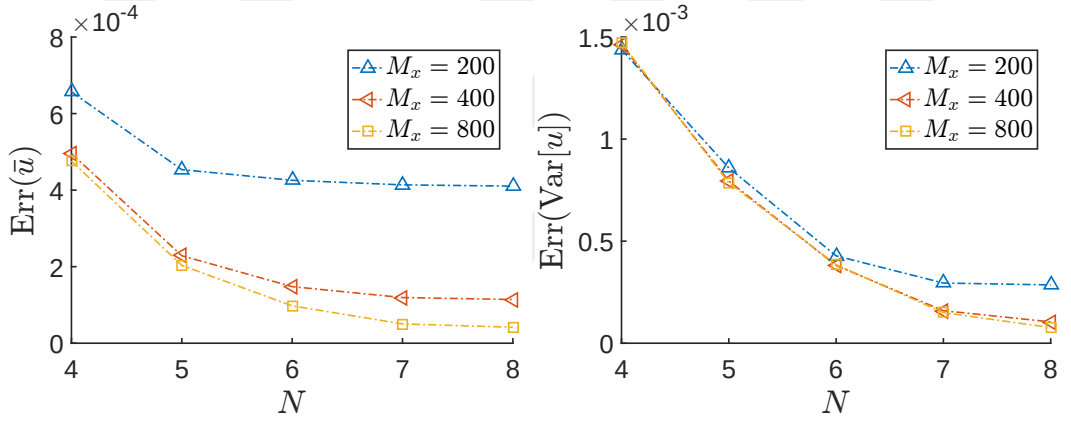

Figure 5.7: Convergence rate of L^2-error in the mean (left) and variance (right) computed with SG method at $t = 1$, with $\epsilon = 0.3$, as a function of PC expansion order N for different number of subintervals M_x.

The solitons obtained from the deterministic KdV equation travel along the spatial domain without changing its form. On the contrary, the mean of the single soliton solution of the stochastic KdV equation is variating its form through time. In other words, the height of the mean decreases, while the width of the mean increases over time. In order to analyze the variation in the form of the mean of the solitons, the mean soliton height (MSH) is defined to be $\max_x |\bar{u}(x,t,\omega)|$ and also the mean soliton width (MSW) is defined to be $|x_1 - x_2|$, where x_1 and x_2 satisfy $\bar{u}(x_1,t,\omega) = \bar{u}(x_2,t,\omega) = \frac{\max_x|\bar{u}(x,t,\omega)|}{2}$ and $x_1 \neq x_2$. In Figure 5.8, MSH and MSW variating over time t is displayed for increasing values of the amplitude of the noise

ϵ. The displayed MSH and MSW are computed by using SG method with 5th order PC expansion and $M_x = 800$ subintervals. Furthermore, Figure 5.9 indicates a strong correlation between MSH and MSW for different amplitudes.

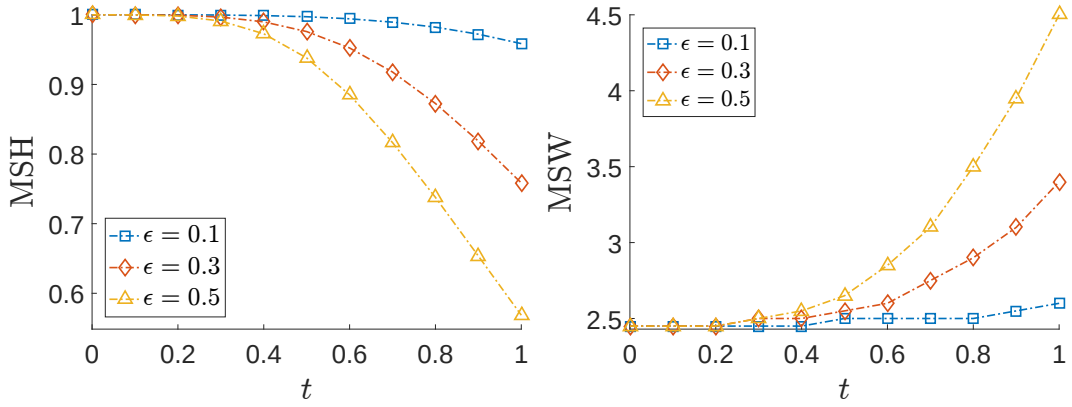

Figure 5.8: MSH (left) and (MSW) (right) as a function of time t with different noise amplitudes $\epsilon = 0.1, 0.3, 0.5$.

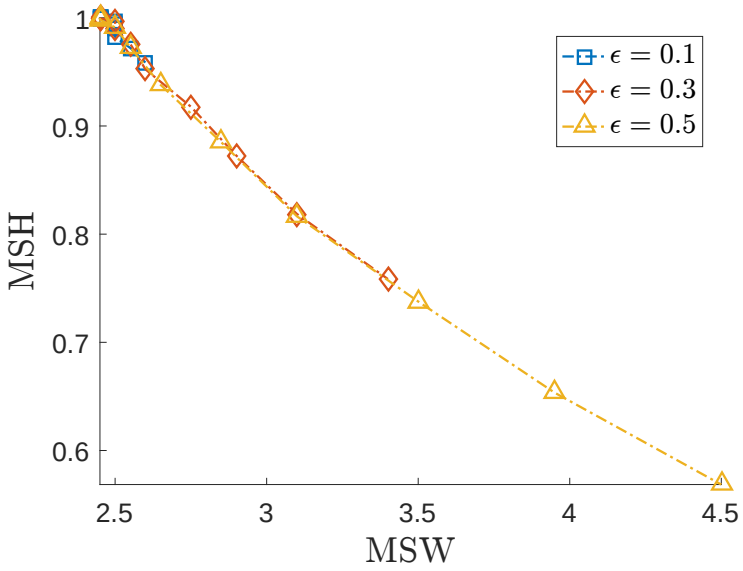

Figure 5.9: Correlation between MSH and MSW obtained by SG methods for time up to $t = 1$ with different noise amplitudes $\epsilon = 0.1, 0.3, 0.5$.

Now, we investigate the effect of the distribution on the mean and variance of the solutions of the stochastic KdV equation with time dependent noise. Figure 5.10 exhibits the results for the Gaussian and uniform distribution at $t = 1$. As the height of mean obtained under the effect of Gaussian noise decreases more deeply, its variance also propagates accordingly.

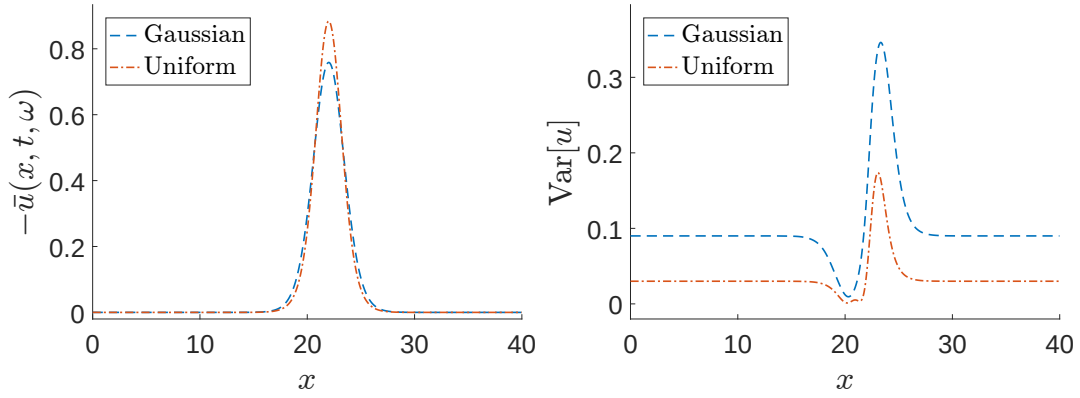

Figure 5.10: The mean (left) and variance (right) of $u(x,t,\omega)$ computed by the SG method at $t=1$ with the speed $c=2$ and the amplitude $\epsilon=0.3$ for the Gaussian and uniform distribution.

Now, we discuss error estimates of stochastic mean and deterministic solutions at $t=1$ obtained by the LDG and the second order FD methods by L_∞–norm:

$$\|u\|_{L_\infty} = \max |u_{\text{exact}} - u_{\text{num}}|.$$

The L_∞ errors in deterministic solutions and stochastic mean of the KdV equation with additive random input (3.6) are presented in Table 5.3 and Table 5.4. It can be concluded that higher accuracy can be achieved by using LDG method for discretizing the spatial dimension of stochastic Galerkin approach.

Table 5.3: L_∞ errors in solutions, discretized by FD scheme, of KdV equation (3.2) and KdV equation with additive random input (3.6) with $\epsilon=0.3$, at $t=1$, $\theta=1$.

Number of Vertices	100	200	300
Deterministic	$4.700e-03$	$1.300e-03$	$8.4384e-04$
Stochastic	$3.000e-03$	$1.100e-03$	$9.3720e-04$

Table 5.4: L_∞ errors in solutions, discretized by LDG scheme, of KdV equation (3.2) and KdV equation with additive random input (3.6) with $\epsilon=0.3$, at $t=1$, $\theta=1$.

Number of Vertices	100	200	300
Deterministic	$1.039e-03$	$5.127e-04$	$4.113e-04$
Stochastic	$2.100e-03$	$9.520e-04$	$7.482e-04$

5.4 Multiplicative Noise

In this section, the numerical results of the stochastic KdV equation under the effect of only space dependent multiplicative noise, formulated as (3.16), will be presented.

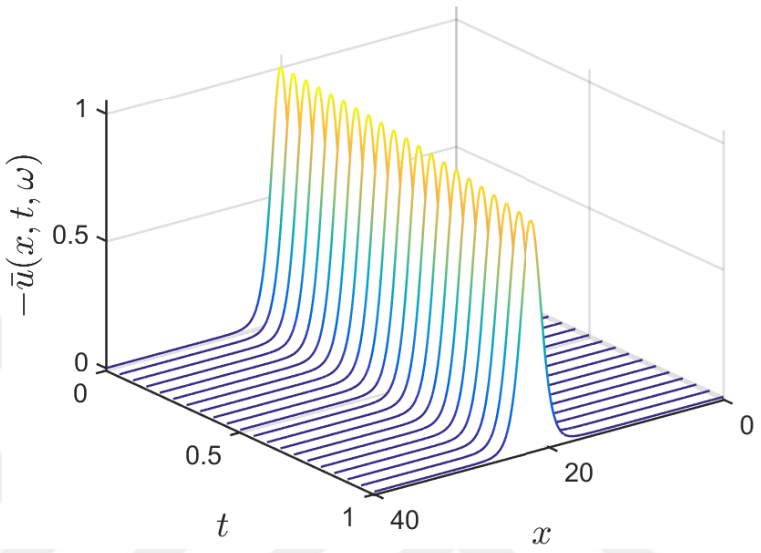

Figure 5.11: The mean values of $u(x,t,\omega)$ obtained by SG method having 5th order PC expansion with the multiplicative noise amplitude $\delta = 0.5$ and the wave speed $c = 2$.

Similarly to the additive case, the spatial domain is estimated with the local discontinuous Galerkin (LDG), the second order finite difference (FD) and the temporal domain is estimated with weighted average. Moreover, the random domain is approximated by Monte Carlo (MC), stochastic collocation (SC), and stochastic Galerkin (SG) methods. As in the additive noise case, the number of spatial grid points used in FD method is $M_x = 1600$, the number of computational subintervals used in the LDG method is $M_x = 800$ and the weighted average method with $\theta = 0.5$, and $\Delta t = 10^{-2}$ is used to discretize the temporal domain $[0, 1]$ in all of the simulations performed.

Figure 5.11 demonstrates the mean of the sKdV (3.16), denoted by $\bar{u}(x,t,\omega)$, with the multiplicative noise amplitude $\delta = 0.5$ and $c = 2$, computed by the SG method with 5th order PC expansion. The height of the wave descends and the width of the wave grows while the wave is traveling along the x-axis in the time.

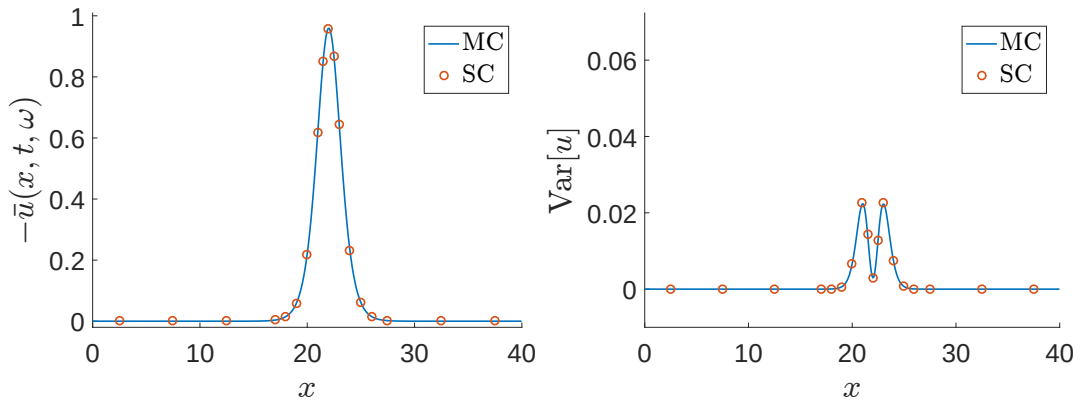

Figure 5.12: The mean value (left) and the variance (right) of $u(x,t,\omega)$ with $c = 2$ and $\delta = 0.3$ at time $t = 1$, obtained by the SC method.

Since the analytical solution for sKdV (3.16) does not exist, Monte Carlo simulation with 5000 samples implemented with the LDG and FD discretization composes bases for comparing the numerical solutions obtained by other stochastic domain discretizations. As a result, the L^2 error in the mean and the variance at time t are computed by

$$\mathrm{Error}(\bar{u}) = \sqrt{\frac{(\bar{u}^t_{\mathrm{num},i} - \bar{u}^t_{\mathrm{mc},i})^2}{M_x}},$$

$$\mathrm{Error}(\mathrm{Var}[u]) = \sqrt{\frac{\left(\mathrm{Var}[u]^t_{\mathrm{num},i} - \mathrm{Var}[u]^t_{\mathrm{mc},i}\right)^2}{M_x}},$$

where M_x is the total number of subintervals used in the LDG method or grid points used in the FD method. The simulations for the L^2 error are done for different number of grid points so that the effect of the spatial and temporal discretization errors on the total error will be studied.

Now, we will solve the stochastic KdV equation (3.16) having a multiplicative noise with the amplitude $\delta = 0.3$, using the stochastic collocation method with 5 collocation points. The deterministic system originated from each collocation points is solved by FD method with $M_x = 1600$. The approximated statistics of the solutions $u(x,t,\omega)$ at $t = 1$ by the SC and MC methods are compared in the Figure 5.12. Moreover, the convergence rate computed between the SC and MC approximations with respect to

the collocation order N is displayed in the Figure 5.13. While the total convergence rate of the variance is not affected deeply by the spatial discretization error, the total convergence rate of the mean is affected deeply. Thus, the total error in the mean obtained by the SC method can be reduced considerably by increasing the number of grid points after 4th collocation point.

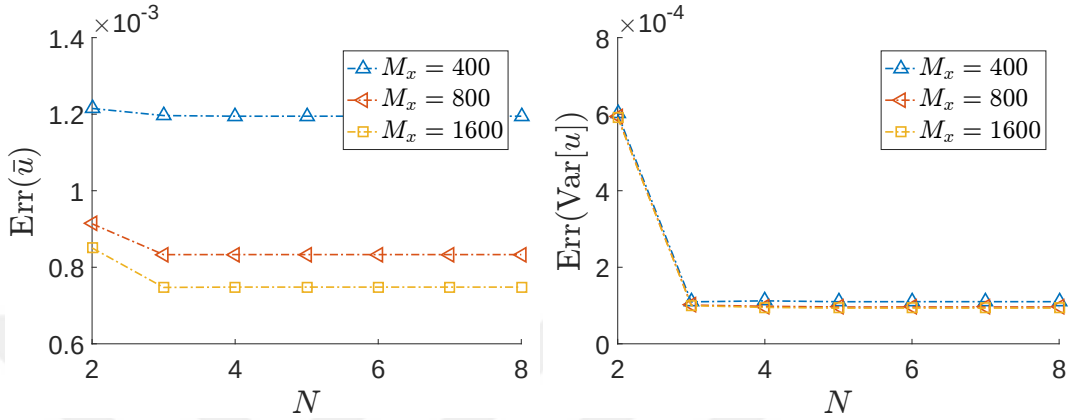

Figure 5.13: Convergence rate of L^2-error in the mean (left) and variance (right) computed with the SC method at $t = 1$, with $\delta = 0.3$, as a function of collocation order N for different number of grid points M_x.

In Figure 5.14, we compare the mean and variance of the problem obtained by using the SG method and the reference solution based on the MC simulation. The SG method is applied with 5th order PC expansion of Hermite polynomials, the LDG method with $M_x = 800$, and the weighted average method with $\theta = 0.5$. Moreover, Figure 5.15 displays the convergence rate in the variance and in the mean as a function of PC expansion order N. It can be seen that the spatial and the temporal discretization errors dominate the total L^2 error in the mean after the 3rd PC basis order while the total L^2-error in the variance is not strongly affected by the spatial/temporal discretization errors.

The mean of the soliton solutions of the stochastic KdV equation varies as an uncertainty is enforced into the KdV equation. In other words, the height of the soliton descends and the width of the solution grows as the soliton is traveling along x–axis. The mean soliton height (MSH) and the mean soliton width (MSW) are defined as in Section 5.3. The MSH and MSW as a function of time t are showed for different

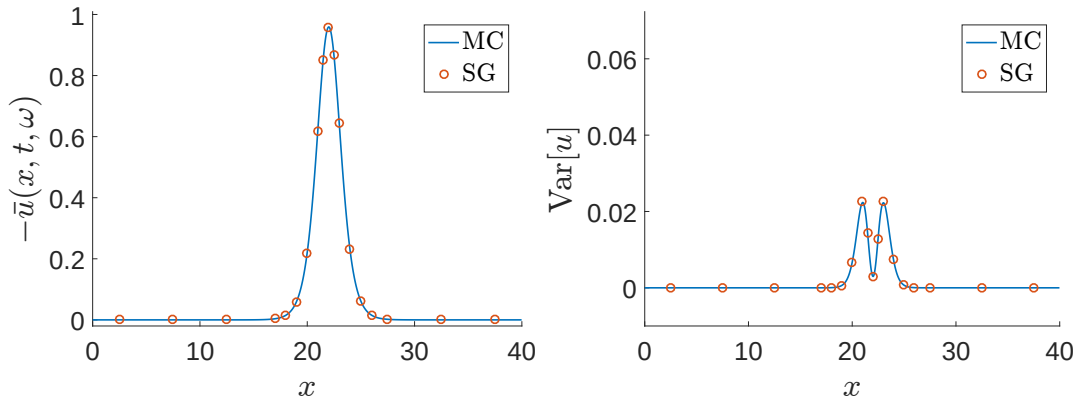

Figure 5.14: The mean value (left) and the variance (right) of $u(x,t,\omega)$ with $c = 1$ and $\delta = 0.3$ at $t = 1$, obtained by the SG method.

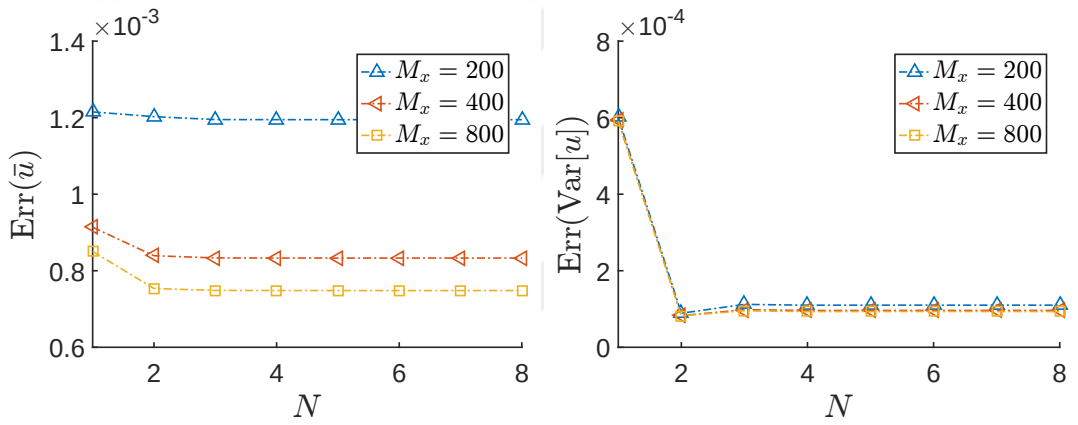

Figure 5.15: Convergence rate of L^2-error in the mean (left) and variance (right) computed with the SG method at $t = 1$, with $\delta = 0.3$, as a function of PC expansion order N for different number of grid points M_x.

values of the multiplicative noise amplitude in Figure 5.16. The demonstrated MSH and MSW are obtained by using the SG method with 5th order PC expansion. Moreover, a strong correlation between MSH and MSW for different multiplicative noise amplitudes is displayed in the Figure 5.17. The correlation obtained in the case of multiplicative noise is more linear than obtained in the case of the additive noise.

In the Figure 5.18, the effect of the distribution with noise amplitude $\delta = 0.3$ on the mean and variance of the solutions of the stochastic KdV equation with the multiplicative noise is investigated. The mean of the solution with uniform noise changes

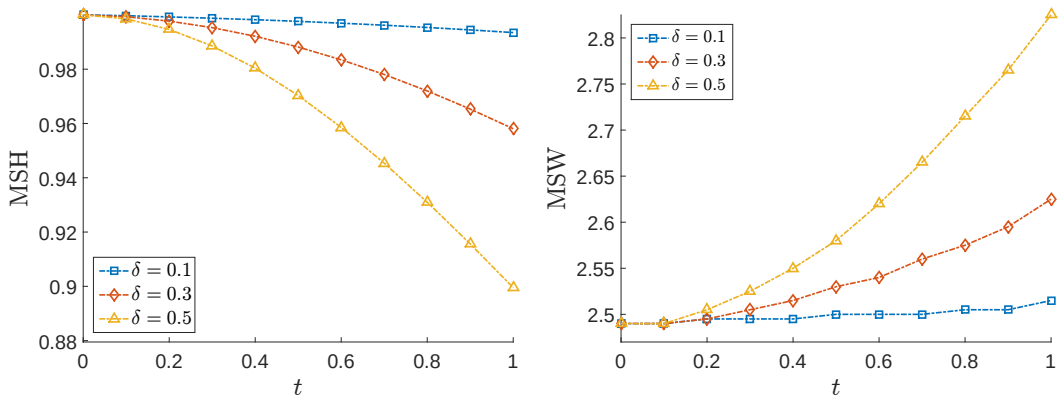

Figure 5.16: MSH (left) and (MSW) (right) as a function of time t with different noise amplitudes $\delta = 0.1, 0.3, 0.5$.

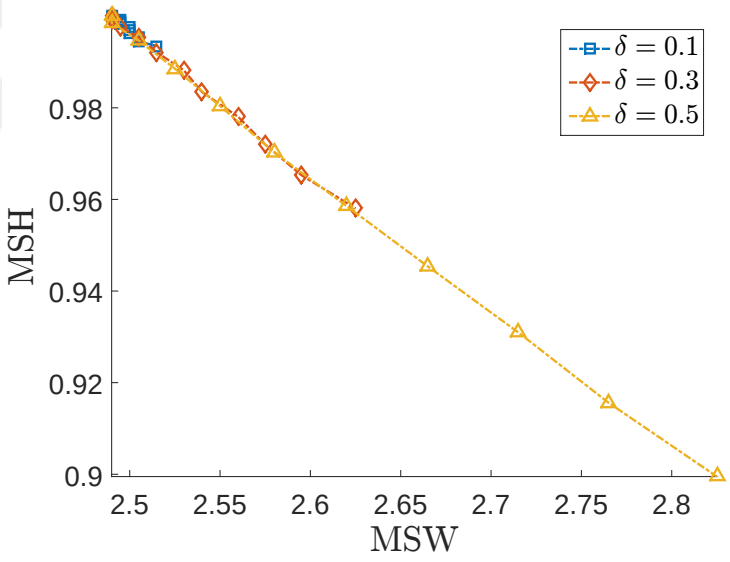

Figure 5.17: Correlation between MSH and MSW obtained by SG methods for time up to $t = 1$ with different noise amplitudes $\delta = 0.1, 0.3, 0.5$.

more slower than the mean of solution with Gaussian noise. Moreover, the variance computed for the uniform noise is much less than the variance computed for the Gaussian noise.

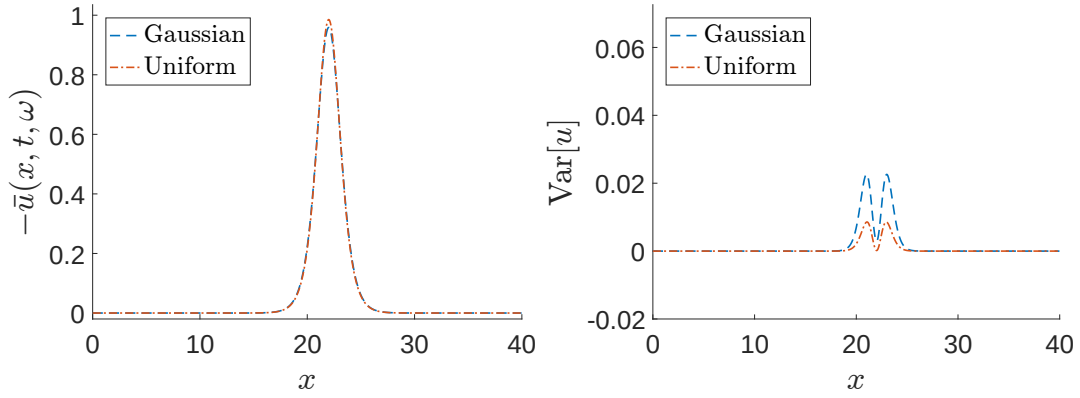

Figure 5.18: The mean (left) and variance (right) of $u(x,t,\omega)$ computed by the SG method at $t = 1$ with the speed $c = 2$ and the amplitude $\delta = 0.3$ for the Gaussian and uniform distribution.

5.5 Additive and Multiplicative Noises

Last, the numerical results of the stochastic KdV equation under the effect of both additive time dependent and multiplicative space dependent noises, formulated as (3.5), will be discussed. In the numerical simulations, the spatial domain is discretized with the local discontinuous (LDG) method with $M_x = 800$ and the second order finite difference method (FD) with $M_x = 1600$. The temporal domain is discretized with the weighted average (theta) method with the step size $\Delta t = 10^{-2}$ and $\theta = 0.5$. Since the stochastic KdV equation (3.5) has two different random inputs, the random domain generated by the problem is two dimensional. Furthermore, the random space is estimated by implementing standard Monte Carlo (MC) simulation, stochastic collocation (SC), and stochastic Galerkin (SG) methods defined for multi-dimensional random spaces.

Figure 5.19 displays how the mean of the sKDV (3.5), denoted by $\bar{u}(x,t,\omega)$, evolves over time t which is obtained by implementing the SG method with 5th order PC expansion, where the additive and multiplicative noise amplitudes are $\delta = 0.5$ and $\epsilon = 0.5$, respectively.

The stochastic collocation method with 5 collocation points in each random dimension is implemented to the sKdV equation with both noises having amplitudes $\epsilon = 0.3$

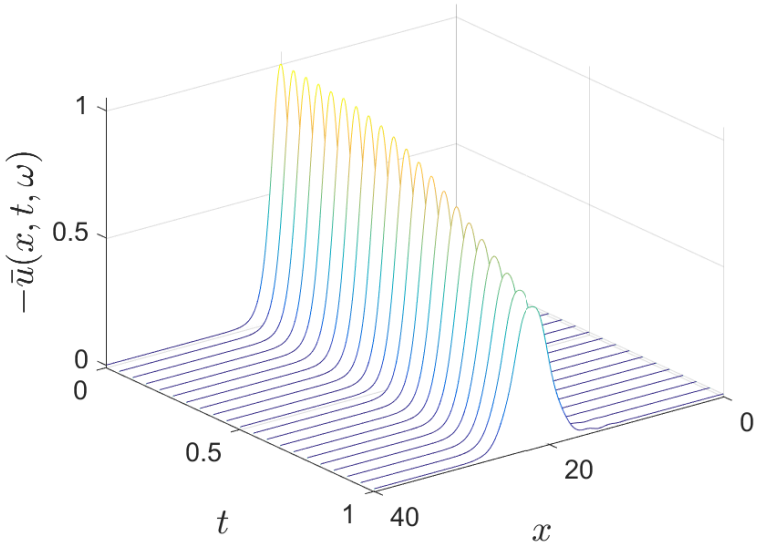

Figure 5.19: The mean values of $u(x,t,\omega)$ obtained by SG method having 5th order PC expansion with the multiplicative noise amplitude $\delta = 0.5$, the additive noise amplitude $\epsilon = 0.5$, and the wave speed $c = 2$.

and $\delta = 0.3$. Since the analytical solution of sKdV (3.5) with both noises does not exist, Monte Carlo simulation having 5000 samples is implemented with the FD and LDG scheme in order to constitute bases for verifying the other stochastic domain approximations. In Figure 5.20 the approximated statistics of the solutions $u(x,t,\omega)$ at $t = 1$ obtained by the 5th order SC method are compared with the statistics obtained by the MC simulation.

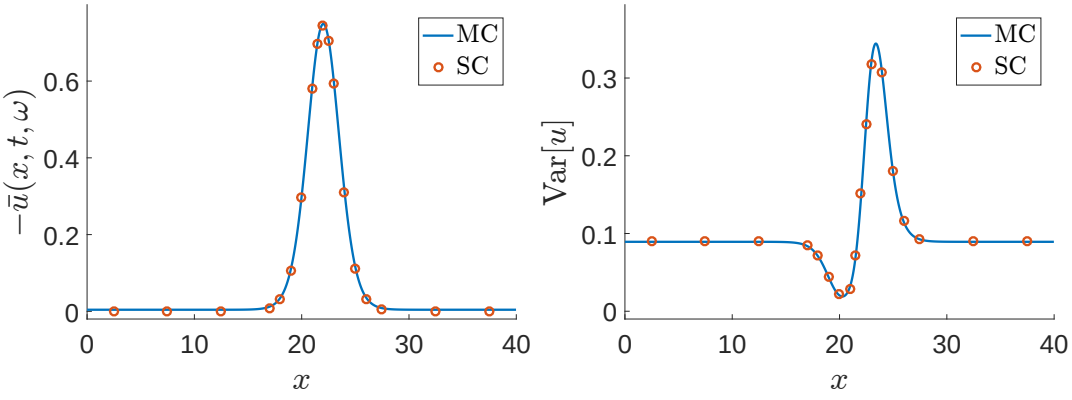

Figure 5.20: The mean value (left) and the variance (right) of $u(x,t,\omega)$ with $c = 2$, $\delta = 0.3$, and $\epsilon = 0.3$ at time $t = 1$, obtained by the SC method.

Now, the effect of the spatial and temporal discretization error in the numerical sim-

ulations is studied through L^2 error. Moreover, the L^2 error in the mean and the variance at time t are computed with respect to solutions obtained by the Monte Carlo simulation, as done in the Section 5.4. The convergence rate computed between the SC and MC approximations with respect to the collocation order N in a single random dimension is displayed in the Figure 5.21. In the SC approximation of the mean, it can be deduced that error arising from random space discretization dominantly affects the total error since refinement over number of spatial subintervals does not reduce the total error efficiently. As in the previous cases the variance approximation is not affected by improvement of the number of spatial subintervals.

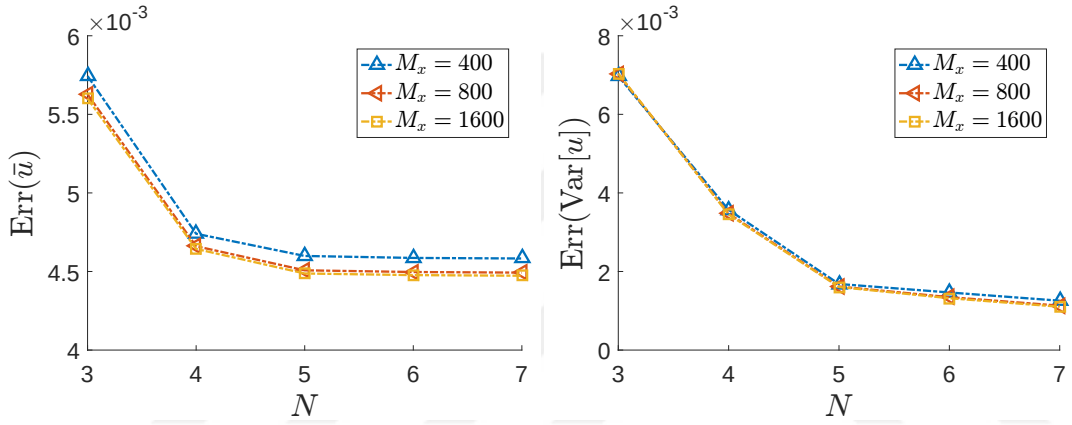

Figure 5.21: Convergence rate of L^2-error in the mean (left) and variance (right) computed with the SC method at $t = 1$, when the noise amplitudes are $\epsilon = 0.3$ and $\delta = 0.3$, as a function of collocation order N for different number of grid points M_x.

The stochastic Galerkin method with a 5th order PC expansion generated by 2–dimensional Hermite polynomials is applied to the sKdV equation with the noise amplitudes $\epsilon = 0.3$ and $\delta = 0.3$. The mean and variance approximated by the SG method and MC simulations are displayed together in Figure 5.22. Moreover, L^2 error in the mean and in the variance are computed with respect to the Monte Carlo approximation and displayed in Figure 5.23 as a function of PC expansion order N for different number of grids. After the 6th order of PC basis, the total L^2 error in SG approximation of the mean can be reduced by refining the spatial grid. However, the random space estimation has a large influence on total error in the variance so that it can be reduced influentially by increasing the SG approximation order.

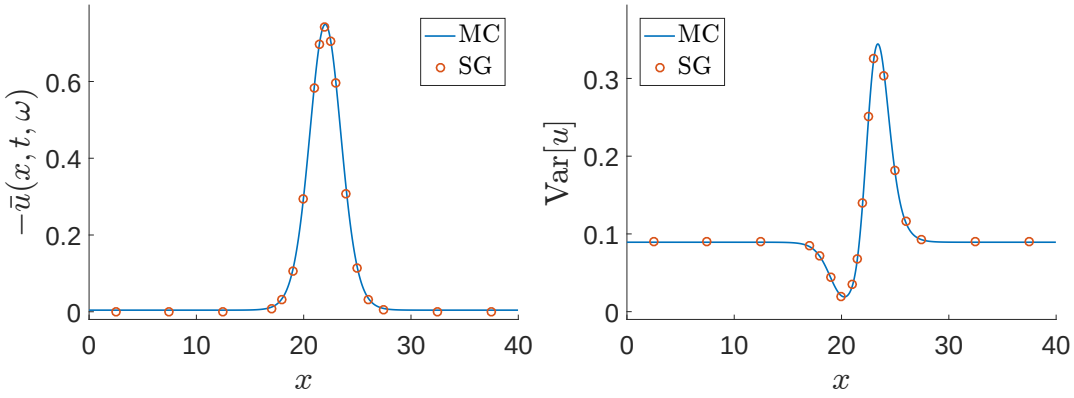

Figure 5.22: The mean value (left) and the variance (right) of $u(x,t,\omega)$ with $c = 1$, $\epsilon = 0.3$, and $\delta = 0.3$ at $t = 1$, obtained by the SG method.

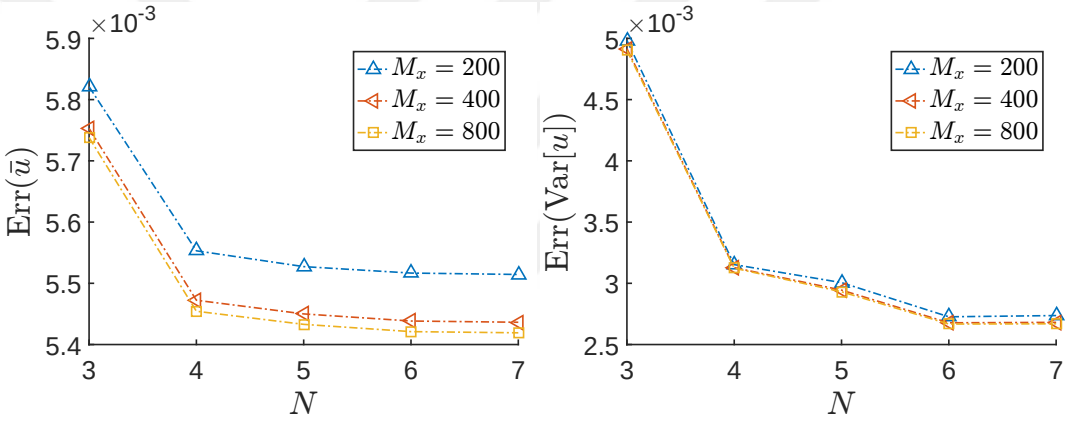

Figure 5.23: Convergence rate of L^2-error in the mean (left) and variance (right) computed with the SG method at $t = 1$, when the noise amplitudes are $\epsilon = 0.3$ and $\delta = 0.3$, as a function of PC expansion order N for different number of grid points M_x.

In Figure 5.24, the change in the single soliton through time which are solution of the stochastic KdV equation under the effect of both the additive and multiplicative noises is analyzed. The mean soliton height (MSH) and the mean soliton width (MSW) are obtained from the solitons approximated by the SG method with 5th order PC expansion. The MSH and MSW are computed as defined in the Section 5.3. It is observed that the MSH decays and MSH increases faster than the other two cases. The relationship obtained between MSH and MSW, in Figure 5.25, is conveniently correlated also when both additive and multiplicative noises are enforced into the

KdV equation.

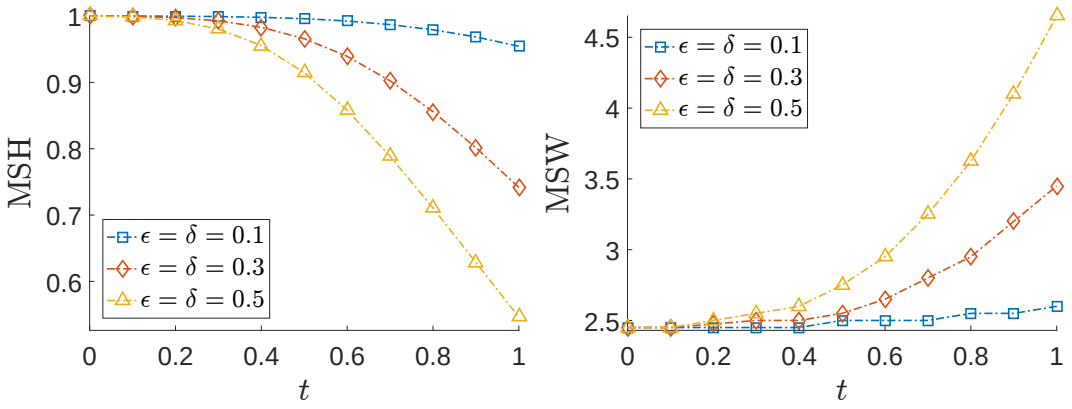

Figure 5.24: MSH (left) and (MSW) (right) as a function of time t with different noise amplitudes $\epsilon = \delta = 0.1, 0.3, 0.5$.

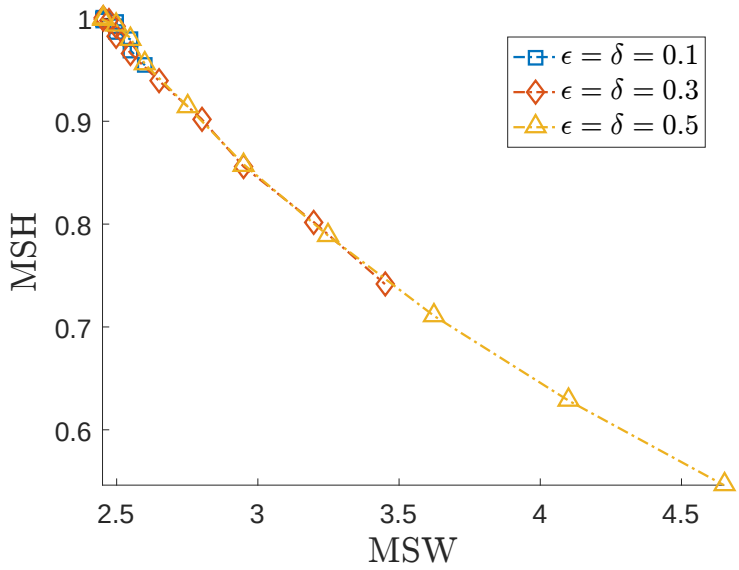

Figure 5.25: Correlation between MSH and MSW obtained by SG methods for time up to $t = 1$ with different noise amplitudes $\epsilon = \delta = 0.1, 0.3, 0.5$.

Lastly, the effect of the distribution type on the uncertainty is investigated. The mean and variance of the solutions of the sKdV with additive and multiplicative inputs having Gaussian and uniform distributions are displayed in Figure 5.26. While the height of the soliton reduces more slowly with uniform noises, the width stays almost same. Moreover, the variance obtained for the uniform inputs is considerably less than the variance computed for the Gaussian inputs.

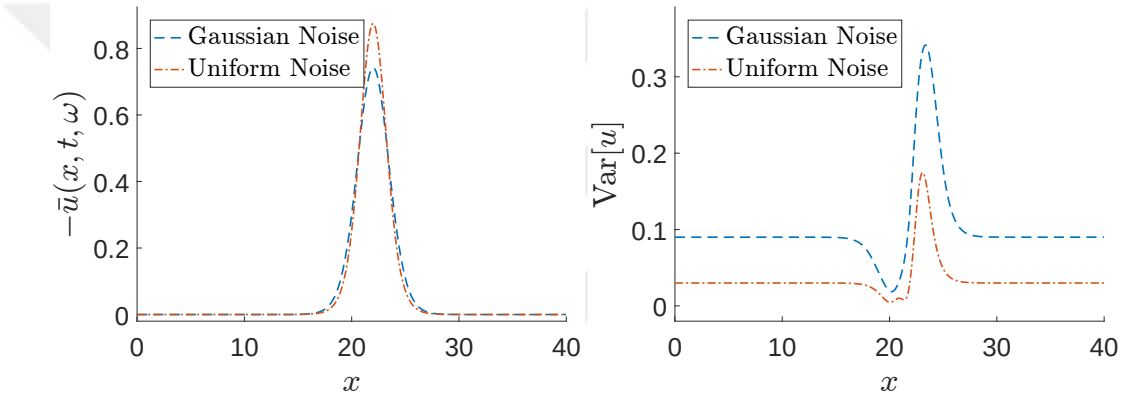

Figure 5.26: The mean (left) and variance (right) of $u(x,t,\omega)$ computed by the SG method at $t = 1$ with the speed $c = 2$, the noise amplitudes $\epsilon = 0.3$ and $\delta = 0.3$ for the Gaussian and uniform distributions.

CHAPTER 6

CONCLUSION AND FUTURE WORK

In this thesis, we have investigated the numerical solutions of the Korteweg-de Vries (KdV) equation with the random input data. The randomness is considered in the right-hand side of the equation as additive noise and in the coefficients of the single derivative term as the multiplicative noise. Then, the probability space emerged from the random input data has been handled by mainly using Stochastic Galerkin method however Monte Carlo and Stochastic Collocation methods are also used to make comparison in the numerical implementations. On the other hand, the local discontinuous Galerkin method is chosen to discretize the physical domain due to its local conservativity property, which is crucial for hyperbolic PDEs.

The numerical results show that the effect of spatial discretization error on the mean of solution becomes dominant over the total error as random domain discretization gets finer. In other words, when the order of Stochastic Galerkin (or Stochastic Collocation) method is taken as fixed, the global error obtained can be decreased deeply by refining the spatial mesh. However, we have deduced that the global error in the variance is not affected by the spatial discretization error as strong as the global error in the mean.

It is well known that the soliton solutions of the KdV equation under the effect of uncertainties do not preserve their original form as they travel along spatial axis. Therefore in the thesis, we have examined the dissipation and dispersion effects of random inputs on the solitons by computing their height and width. We have observed that the height of mean of solitons decreases while their width increases, and there exists a strong correlation between them for both additive and multiplicative

noises. Moreover, the rate of the decrease/increase in the height/width of solitons is the heaviest in the combined case, moderate in the additive case, and the modest in the multiplicative case. Additionally, we have noticed dissipation and dispersion effects of uncertainties differ with different kinds of probability distributions. The decrease in height of mean is more severe with Gaussian distribution than with uniform distribution, whereas their increase in width is almost similar.

Polynomial chaoses (PC) basis for Stochastic Galerkin methods yields a increasing dimensionality of the system as the number of random variables increases. Therefore, as future work, the implementation of SG method with sparse high order polynomial chaos basis can overcome this problem [4]. Moreover, the numerical results of linear KdV equation, obtained by using the rational deferred correction method as a post-processing method, are promising in order to improve the accuracy using minimum number of time steps. Further, the strategies applied to one-dimensional KdV equation with the random input data in this thesis can be extended to two dimensional nonlinear PDEs such as Kadomtsev–Petviashvili (KP) equation [22, 40] and Zakharov–Kuznetsov (ZK) equation [57] including random input data.

REFERENCES

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM Journal on Numerical Analysis, 39, pp. 1749–1779, 2006.

[2] I. Babuška, F. Nobile, and R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM Journal on Numerical Analysis, 45(3), pp. 1005–1034, 2007.

[3] I. Babuška, R. Tempone, and G. E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM Journal on Numerical Analysis, 42(2), pp. 800–825, 2007.

[4] M. Bieri and C. Schwab, Sparse high order FEM for elliptic sPDEs, Computer Methods in Applied Mechanics and Engineering, 198, pp. 1149–1170, 2009.

[5] J. Boussinesq, Théorie de l'intumescence liquide, appelée onde solitaire ou de translation, se propageant dans un canal rectangulaire, Comptes rendus de l'Académie des sciences, 72, pp. 755–759, 1871.

[6] R. H. Cameron and W. T. Martin, The orthogonal development of nonlinear functionals in a series of Fourier-Hermite functionals, Annals of Mathematics, 48(2), pp. 358–392, 1947.

[7] T. S. Chihara, *An Introduction to Orthogonal Polynomials*, Gordon and Breach, 1978.

[8] B. Cockburn, Devising discontinuous Galerkin methods for non-linear hyperbolic conservation laws, Journal of Computational and Applied Mathematics, 128, pp. 187–204, 2001.

[9] B. Cockburn, G. E. Karniadakis, and C.-W. Shu, *The development of discontinuous Galerkin methods, Discontinuous Galerkin Methods: Theory, Computations and Applications*, Springer, 2000.

[10] B. Cockburn and C. W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM Journal of Numerical Analysis, 35, pp. 2440–2463, 1998.

[11] G. S. Fishman, *Monte Carlo Concepts, Algorithms and Applications*, Springer-Verlag, 1996.

[12] D. Funaro, *Polynomial Approximation of Differential Equations*, Springer-Verlag, 1992.

[13] C. Gardner, J. Greene, M. Kruskal, and R. M. Miura, Method for solving the Korteweg-de Vries equation, Physical Review Letters, 19(19), pp. 1095–1907, 1967.

[14] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements:A Spectral Approach*, Springer-Verlag, 1991.

[15] I. G. Graham, R. Scheichl, and E. Ullmann, Mixed finite element analysis of lognormal diffusion and multilevel Monte Carlo methods, Stochastics and Partial Differential Equations Analysis and Computations, 4(1), pp. 41–75, 2016.

[16] S. Güttel and K. Georges, Efficient high-order rational integration and deferred correction with equispaced data, Electronic Transactions on Numerical Analysis, 41, pp. 443–464, 2014.

[17] S. Güttel and J. W. Pearson, A rational deferred correction approach to parabolic optimal control problems, IMA Journal of Numerical Analysis, Accepted in 2017.

[18] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin methods: Analysis, Algorithms, and Applications*, Springer, 2008.

[19] T. Iizuka, Anomalous diffusion of solitons in random systems, Physics Letters A, 181(1), pp. 39–42, 1993.

[20] K. Itó, Multiple Wiener integral, Journal of Mathematical Society of Japan, 3(1), pp. 157–169, 1951.

[21] J. Jacod and P. Protter, *Probability essentials*, Springer, 2nd edition, 2004.

[22] B. B. Kadomtsev and V. I. Petviashvili, The stability of solitary waves weakly dispersing media, Soviet Physics Doklady, 15, pp. 539–541, 1970.

[23] H. Kim, Y. Kim, and D. Yoon, Dependence of polynomial chaos on random types of forces of KdV equations, Applied Mathematical Modelling, 36, pp. 3080–3093, 2012.

[24] R. Koekoek, P. A. Lesky, and R. F. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q-Analogues*, Springer-Verlag, 2010.

[25] K. Konno and Y. H. Ichikawa, A modified Korteweg de Vries equation for ion acustic waves, Applied Mathematical Modelling, 37(6), pp. 1631–1636, 1974.

[26] D. J. Korteweg and G. D. Vries, On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves, Philosophical Magazine Series, 39(5), pp. 422–443, 1895.

[27] P. D. Lax, Integrals of nonlinear equations of evolution and solitary waves, Communications on Pure and Applied Mathematics, 21(5), pp. 467–490, 1968.

[28] O. Le Maître and O. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer, 2010.

[29] G. Lin, L. Grinberg, and G. E. Karniadakis, Numerical studies of the stochastic Korteweg-de Vries equation, Journal of Computational Physics, 213, pp. 976–703, 2006.

[30] M. Loève, *Probability Theory I*, Springer, 4th edition, 1977.

[31] M. Loève, *Probability Theory II*, Springer, 4th edition, 1978.

[32] W. Loh, On Latin hypercube sampling, The Annals of Statistics, 24(5), pp. 2058–2080, 1996.

[33] G. Matheron, *Estimating and Choosing: An Essay on Probability in Practice*, Springer-Verlag, 1989.

[34] R. M. Miura, C. S. Gardner, and M. D. Kruskal, Korteweg-de Vries equation and generalizations. II. existense of conservation laws and constants of motion, Journal of Mathematical Physics, 9(8), pp. 1204–1209, 1968.

[35] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, Society For Industrial and Applied Mathematics, 1992.

[36] C. E. Powell and U. Elisabeth, Preconditioning stochastic Galerkin saddle point systems, SIAM Journal on Matrix Analysis and Applications, 31(5), pp. 2813–2840, 2010.

[37] B. Rivière, *Discontinuous Galerkin Methods for solving elliptic and parabolic equations: Theory and Implementation*, Frontiers in Applied Mathematics, 2008.

[38] J. S. Russell, Report on waves, 14th Meeting of The British Association for the Advancement on Science, 1, p. 311, 1844.

[39] M. Scalerandi and A. Romano, Korteweg-de Vries solitons under additive stochastic perturbations, Pyhsical Review E, 58(4), pp. 4166–4173, 1998.

[40] H. Segur and A. Finkel, An analytical model of periodic waves in shallow water, Studies in Applied Mathematics, 73, pp. 183–220, 1985.

[41] H. Segur and J. L. Hammack, Soliton models of long internal waves, Journal of Fluid Mechanics, 118, pp. 285–304, 1982.

[42] A. Sjöberg, On the Korteweg-de Vries equation: Existence and uniqueness, Journal of Mathematical Analysis and Applications, 29, pp. 569–579, 1970.

[43] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Mathematics Doklady, 4, pp. 240–243, 1963.

[44] S. Sun and J. Liu, A locally conservative finite element method based on piecewise constraint enrichment of the continuous Galerkin method, SIAM Journal on Scientific Computing, 41, pp. 2528–2548, 2009.

[45] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society, 4th edition, 1975.

[46] A. F. Timan, *Theory of Approximation of Functions of a Real Variable*, Pergamon Press, 1963.

[47] M. Wadati, Stochastic Korteweg-de Vries equation, Journal of the Physical Society of Japan, 52(8), pp. 2642–2648, 1983.

[48] M. Wadati and Y. Akutsu, Stochastic Korteweg-de Vries equation with and without damping, Journal of the Physical Society of Japan, 53(10), pp. 3342–3350, 1984.

[49] H. Washimi and T. Taniuti, Propagation of ion-acoustic solitary waves of small amplitude, Physical Review Letters, 17, pp. 996–998, 1966.

[50] N. Wiener, The homogeneous chaos, American Journal of Mathematics, 60(4), pp. 897–936, 1938.

[51] D. Xiu, Efficient collocational approach for parametric uncertainty analysis, Communications in Computational Physics, 2(2), pp. 293–209, 2007.

[52] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.

[53] D. Xiu and J. S. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM Journal on Scientific Computing, 27(3), pp. 1118–1139, 2005.

[54] D. Xiu and G. E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM Journal on Scientific Computing, 24(2), pp. 619–644, 2002.

[55] J. Yan and C. W. Shu, A local discontinuous Galerkin method for KdV type equations, SIAM Journal of Numerical Analysis, 40, pp. 769–791, 2002.

[56] N. J. Zabusky and M. D. Kruskal, Interaction of "solitons" in a collisionless plasma and the recurrence of initial states, Physical Review Letters, 15(6), pp. 240–243, 1965.

[57] V. E. Zakharov and E. A. Kuznetsov, Three-dimensional soliton, Soviet Physics–JETP, 39, pp. 285–286, 1974.

[58] G. Zhang, D. Lu, M. Ye, M. Gunzburger, and C. Webster, An adaptive sparse-grid high-order stochastic collocation method for Bayesian inference in groundwater reactive transport modeling, Water Resources Research, 49(10), pp. 6871–6892, 2013.