Wiener Chaos Expansion and Numerical Solutions of Stochastic Partial Differential Equations

Thesis by

Wuan Luo

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

(Defended May 2, 2006)

Contents

Acknowledgements **xiii**

Abstract **xv**

1 Introduction **1**

1.1 Overview 1

1.2 Stochastic Spectral Method Based on Polynomial Chaos Expansion 3

1.3 Wiener Chaos Method for SPDEs with Brownian Motion Forcing 5

1.4 Summary of the Thesis 7

1.4.1 Stochastic Elliptic Equations and Uncertainty Quantification 9

2 Hermite Polynomials and Wiener Chaos Expansions **13**

2.1 Hermite Polynomials 13

2.2 Fourier-Hermite Expansions of Functions of Gaussian Random Variables . . 19

2.3 Wiener Chaos Expansions of Functionals of Brownian Motions 22

2.4 Wiener Chaos Expansions: A Martingale Approach 28

3 WCE Methods for Stochastic Burgers Equations **33**

3.1 Stochastic Burgers Equations with Additive Random Forcing 34

3.1.1 Equations for the WCE Coefficients 34

3.1.2 Truncating the WCE Propagator 37

3.1.3 Numerical Solutions of the WCE Propagator 40

3.1.4 Convergence Verification of the WCE Method 48

3.2 Stochastic Burgers Equations with Multiplicative Random Forcing 54

3.2.1 Derivation of the WCE Propagator 55

3.2.2 Monte Carlo Simulation 57

3.2.3 Numerical Experiments 59
3.2.4 Convergence Verification of the WCE Method 61
3.3 Numerical Verification of Stationary Measure of Stochastic Burgers Equations 66
3.4 Random Forcing with Large Variability 73

4 WCE Methods for Stochastic Navier-Stokes Equations 75
4.1 The WCE method 77
4.2 Convergence Verification of the WCE Method 80

5 Error Analysis of the WCE Method 85
5.1 Convergence Analysis for SPDEs with Additive Forcing 85
5.2 Convergence Analysis for SPDEs with Multiplicative Forcing 94

6 Long Time Integrations of the WCE Method 97
6.1 Long Time Integrations of Stochastic Transport Equations 98
6.1.1 Full Sparse WCE Solutions 103
6.1.2 Aggressive Sparse WCE Solutions 105
6.2 Refining the WCE Solution by MC Simulations 108
6.2.1 Reducing the Variance of MC Simulations Using WCE Solutions . . 109
6.2.2 The Strongly Convergent Numerical Schemes for MC Simulations . . 114
6.2.3 Numerical Results of the WCE-MC Hybrid Method 115
6.3 Stochastic Burgers Equations with Large variability 118
6.4 Front Propagation Speeds of Stochastic Reaction-Diffusion Equations . . . 119
6.4.1 Problem Setup 121
6.4.2 Wiener Chaos Propagator for the R-D Equation 123
6.4.3 Numerical Simulations 124

7 WCE Methods and Upscaling for Stochastic Elliptic Equations 129
7.1 Karhunen-Loeve Expansion 131
7.2 WCE Methods for Stochastic Elliptic Equations 134
7.2.1 Property of the Truncated Elliptic Systems 141
7.3 Upscaling of the WCE Elliptic System 142

8 Uncertainty Quantification in Reservoir Modeling **147**

8.1 Introduction 147

8.2 The Model Equations and Problem Setting 149

8.3 Preconditioning MCMC Simulations Using Coarse-scale Models 152

8.3.1 The Preconditioned MCMC method 152

8.3.2 Analysis of the Preconditioned MCMC Method 158

8.3.3 Numerical Results 162

8.4 WCE Methods and Coarse Gradient Langevin Algorithms 173

8.4.1 Proposal Distributions Based on Langevin Diffusions 173

8.4.2 Constructing $\Pi^*(\theta)$ Explicitly by the Upscaled WCE Method 178

9 Concluding Discussions **181**

9.1 Future Work 182

A Stochastic Integrals **185**

B Monte Carlo Simulations of Stochastic Differential Equations **189**

B.1 Strong and Weak Convergence 189

B.2 Variance Reduction of MC Simulations 192

B.2.1 Antithetic Variables 194

B.2.2 Control Variates 195

C Coarse-scale Models Using Multiscale Finite Element Methods **197**

Bibliography **201**

List of Figures

3.1 Illustration of the sparse truncation 39
3.2 Statistic moments of the Burgers equation by the WCE method 43
3.3 L^2 norm of the WCE coefficients 44
3.4 Convergence of MC simulation in log-log scale 46
3.5 Convergence comparison for WCE and MC 47
3.6 Errors of the WCE method with different polynomial order 52
3.7 Errors of MC simulation 53
3.8 Statistics by WCE for multiplicative forcing 61
3.9 Convergence test of the WCE method 66
3.10 Errors of WCE with different polynomial orders 67
3.11 Errors of MC with different number of realizations 67
3.12 Time history of the statistics 71
3.13 Incremental change in time of the statistics 71
3.14 Time history of the statistics with a different initial state 72
3.15 Comparison of means and variances at $T = 6.0$ 73
3.16 Effects of the magnitude of the random forcing 74

4.1 Initial condition of vorticity and temperature 77
4.2 Statistic moments of the SNS equation 81

5.1 Error of WCE methods at sharp gradient regions 90

6.1 Stream function of cellular flow 100
6.2 L^2 norm of the WCE coefficients for the full sparse truncation 104
6.3 Mean and variance of the stochastic transport equation at $T = 10.0$ 106
6.4 Errors of the full sparse WCE solution 107

6.5 L^2 of the WCE coefficients for the aggressive sparse truncation 108
6.6 Error of the variance by the aggressive sparse WCE method 109
6.7 Errors of the WCE-MC hybrid solution 116
6.8 The convergence constants for estimating the mean by MC ensemble averages 116
6.9 The convergence constants for estimating the variance by MC ensemble averages 117
6.10 WCE-MC method for random forcing with large magnitude 118
6.11 Illustrations of the front propagation for R-D equations 120
6.12 Ensemble average front speed of the R-D equation 126

7.1 Eigenvalues of the KLE with a Gaussian covariance function 133
7.2 Eigenvalues of the KLE with the exponential covariance function 134
7.3 Eigenvalues of the exponential covariance function with larger correlation length 134

8.1 Typical fine and coarse scale fractional flows 154
8.2 Cross-plot between $E_k = \|F - F_k\|^2$ and $E_k^* = \|F - F_k^*\|^2$ 157
8.3 Eigenvalues of the KLE for a Gaussian covariance 164
8.4 Acceptance rate for single-phase flow 166
8.5 Fractional flows of sampled permeability fields 167
8.6 Fractional flow errors for the accepted permeability samples 168
8.7 Isotropic log permeability realizations of single phase flows 169
8.8 Acceptance rate for the anisotropic single-phase flow 170
8.9 Realizations of the anisotropic log permeability fields 171
8.10 Acceptance rate for two-phase flows 172
8.11 Fractional flow errors for the accepted permeability samples in two-phase flows 173

B.1 Asymptotical convergence rate of MC simulations 193

List of Tables

3.1 Errors of WCE and MC for an additive model equation 51
3.2 Errors of WCE and MC for a multiplicative model equation 66

4.1 Errors of WCE and MC for a model stochastic Navier-Stokes equation 83

Acknowledgements

I am deeply indebted to my advisor, Prof. Tom Hou, for his guidance and encouragement during my study at Caltech. It is Tom who attracted me to pursue my Ph.D. study at Caltech in the first place. With his passion and vision for scientific research, Tom sets a role model for his students, which we can hardly emulate. During our numerous research discussions and frequent personal contact, Tom keeps inspiring me both as an experienced mathematician and a good friend. Tom has always wanted the best for me, and supported every endeavor I have made. I consider myself very lucky to have had the opportunity working with him for the past five years.

I am truly grateful to Prof. Boris Rozovskii at USC for introducing me to the fascinating topic of Wiener chaos expansion and the field of stochastic differential equations. Boris's original works on Wiener chaos provide the starting point for this thesis. Prof. Haomin Zhou at Georgia Tech generously shared with me his insights and numerical codes for Wiener chaos expansion. His invaluable help is sincerely acknowledged. I would like to thank Prof. Jack Xin at UCI for stimulating discussions on reaction-diffusion problems and the references he provided me.

Special thanks go to Prof. Yalchin Efendiev at Texas A&M, for his insights and guidance in our enjoyable collaborations, and his hospitality during my visit at College Station. I am grateful to Prof. Oscar Bruno and Prof. Houman Ohwadi for many helpful discussions.

I would like to thank Prof. Kaushik Bhattacharya, Prof. Emmanuel Candes, and Prof. Houman Ohwadi for taking their precious time to read my thesis and serve on the committee.

During the first four years of my study at Caltech, I was supported by the Betty & Gordon Moore Fellowship. Their generosity is greatly appreciated.

I would like to thank my officemates through the years, Razvan Fetecau, Marco Latini, and Laurent Demanet, for the many happy conversations and laughter we had during those drowsy afternoons. I also want to thank all the other friends I have made in ACM, especially

Dr. Gaby Stredie, Dr. Andy Westhead, Dr. Xinwei Yu, Dr. Theofilos Strinopoulos, Andy Monro, Ke Wang, David Hoch, Nawaf Bou-Rabee, Hannes Helgason, Stephane Lintner, Jonathan Othmer, Lei Zhang, and Zhiyi Li, etc. Many thanks are due to Sheila Shull, Chad Schumtzer, and Sydney Garstang for being there whenever we need help.

Finally, I would like to thank my parents, and my brother and sister for constant love and support. My dear wife, and also my many-year classmate and friend, Lixiu Tian, deserves the most special thanks. Her company and support are the main reasons I have survived this strenuous Ph.D. journey.

Abstract

Stochastic partial differential equations (SPDEs) are important tools in modeling complex phenomena, and they arise in many physics and engineering applications. Developing efficient numerical methods for simulating SPDEs is a very important while challenging research topic. In this thesis, we study a numerical method based on the Wiener chaos expansion (WCE) for solving SPDEs driven by Brownian motion forcing. WCE represents a stochastic solution as a spectral expansion with respect to a set of random basis. By deriving a governing equation for the expansion coefficients, we can reduce a stochastic PDE into a system of deterministic PDEs and separate the randomness from the computation. All the statistical information of the solution can be recovered from the deterministic coefficients using very simple formulae.

We apply the WCE-based method to solve stochastic Burgers equations, Navier-Stokes equations and nonlinear reaction-diffusion equations with either additive or multiplicative random forcing. Our numerical results demonstrate convincingly that the new method is much more efficient and accurate than MC simulations for solutions in short to moderate time. For a class of model equations, we prove the convergence rate of the WCE method. The analysis also reveals precisely how the convergence constants depend on the size of the time intervals and the variability of the random forcing. Based on the error analysis, we design a sparse truncation strategy for the Wiener chaos expansion. The sparse truncation can reduce the dimension of the resulting PDE system substantially while retaining the same asymptotic convergence rates.

For long time solutions, we propose a new computational strategy where MC simulations are used to correct the unresolved small scales in the sparse Wiener chaos solutions. Numerical experiments demonstrate that the WCE-MC hybrid method can handle SPDEs in much longer time intervals than the direct WCE method can. The new method is shown to be much more efficient than the WCE method or the MC simulation alone in relatively

long time intervals. However, the limitation of this method is also pointed out.

Using the sparse WCE truncation, we can resolve the probability distributions of a stochastic Burgers equation numerically and provide direct evidence for the existence of a unique stationary measure. Using the WCE-MC hybrid method, we can simulate the long time front propagation for a reaction-diffusion equation in random shear flows. Our numerical results confirm the conjecture by Jack Xin [76] that the front propagation speed obeys a quadratic enhancing law.

Using the machinery we have developed for the Wiener chaos method, we resolve a few technical difficulties in solving stochastic elliptic equations by Karhunen-Loeve-based polynomial chaos method. We further derive an upscaling formulation for the elliptic system of the Wiener chaos coefficients. Eventually, we apply the upscaled Wiener chaos method for uncertainty quantification in subsurface modeling, combined with a two-stage Markov chain Monte Carlo sampling method we have developed recently.

Chapter 1

Introduction

1.1 Overview

Stochastic partial differential equations (SPDEs) are known to be an effective tool in modeling complex physical and engineering phenomena. Examples include wave propagation [82], diffusion through heterogeneous random media [83], randomly forced Burgers and Navier-Stokes equations (see e.g. [7, 23, 24, 25, 71, 85, 86, 92, 93] and the references therein). Additional examples can be found in materials science, chemistry, biology, and other areas. In those problems, the large structures and dominant dynamics are governed by deterministic physical laws, while the unresolved small scales, microscopic effects, and other uncertainties can be naturally modeled by stochastic processes. The resulting equations are usually PDEs with either random coefficients, random initial conditions, or random forcing. Unlike deterministic PDEs, solutions of stochastic PDEs are random fields. Hence, it is important to be able to study their statistical characteristics, e.g., mean, variance, and higher order moments.

Due to the complex nature of SPDEs, numerical simulations play an important role in studying this class of PDEs. For this reason, quite a few numerical methods have been developed for simulating SPDEs, such as the moment equations, probability density function (PDF) method, etc. The method of moment equations relies on deriving effective equations for the statistical moments. For nonlinear problems, however, the moment equations are not closed, with lower order moments depending on higher order moments. The hierarchy system for the different moments is quite complicated, and there is no easy way to tell whether the method will converge or not. Truncating the moment equations may cause serious numerical errors, or even yield unphysical results, such as negative second order

moments. So the moment equation method works only for limited applications.

The PDF method aims to derive a Fokker-Planck equation for the probability density function of the random solution. The PDF is a function of the spatial-temporal variables as well as the state variables for the random solution itself. It provides a detailed description of the statistical information of the random solution. For nonlinear equations, the Fokker-Planck equation for the PDF usually involves derivatives of conditional expectations [24, 23, 31], which are very difficult to handle in numerical computations. Moreover, the PDF is usually a high dimensional function. For example, the PDF of a stochastic Navier-Stokes equation in 2-D is a function with five dimensions, and in 3-D, it would be a function with seven dimensions. Even when there exists an approximate Fokker-Planck equation for the PDF, solving it numerically is still far from an easy job. So the PDF method is more useful for theoretical analysis than as a tool for numerical simulation.

Currently, Monte Carlo (MC) methods are still the most popular numerical method in simulating solutions of SPDEs [53, 84]. The idea of MC simulations is to sample the randomness in the SPDEs and solve the stochastic equations realization by realization. For each given realization of the randomness, the SPDEs become deterministic and can be solved by regular numerical methods. To estimate the statistical information of the random solution, one can solve the SPDEs many times with different realizations and compute the ensemble averages. To sample the source randomness in the SPDEs, MC simulations rely on pseudo random number generators. MC simulations are very general and robust, and can handle many different sources of randomness. MC simulations are also very easy to implement. Once we know how to sample the source randomness, we can use the existing deterministic code to simulate the stochastic equations.

Based on the law of large number, Monte Carlo ensemble averages converge at the rate of $1/\sqrt{M}$, where M is the total number of realizations. This is a rather slow convergence rate. To reduce the error by a factor of 2, we need to increase the number of realizations by a factor of 4. To compensate the slow convergence of MC simulations, various accelerating techniques have been developed, such as quasi-Monte Carlo methods, antithetic variables, control variates, importance sampling, etc. Please see [9, 53, 84] and the reference therein. Though these methods can be quite effective for numerical integration problems, their applications to simulating SPDEs, especially nonlinear SPDEs, are quite limited. In MC integration problems, the probability distributions are known explicitly. So we can tune

and engineer the acceleration technique based on the concrete distributions. For SPDEs, however, the distribution of the random solution is not known explicitly, due to the nonlinear dynamics of the SPDEs. The ensemble averages (expectations) are taken against an implicit distribution determined by the dynamical equations. Since the dynamics of the SPDEs will evolve and transform the randomness presented in the equation, the random solution may have very different distributions from the randomness source. It is not surprising that acceleration techniques based on the source randomness may not work very well for the random solution itself.

There exist a few other acceleration techniques designed specifically for simulating SPDEs, such as measure transformations, martingale representations, or conditional expectations, etc. These methods seek to reduce the variance of the MC simulations by exploiting the inherent stochastic structures of the SPDEs. They often rely on the analytical information of the random solutions, and therefore only work for simple stochastic ordinary differential equations. For complicated nonlinear stochastic partial differential equations, it is very difficult to apply those methods. Due to the above reasons, MC simulations combined with various acceleration techniques are still very expensive in simulating SPDEs. Alternative numerical methods are much desired for solving SPDEs with complicated randomness and dynamical structures.

1.2 Stochastic Spectral Method Based on Polynomial Chaos Expansion

In recent years, polynomial chaos expansion has received much attention as a promising numerical method in solving SPDEs. The polynomial chaos was originally developed by the great mathematician Norbert Wiener [95, 96] in 1938. Using Hermite polynomials, Wiener constructed an orthonormal random basis for expanding homogeneous chaos depending on white noise, and used it to study problems in statistical mechanics. Following Wiener's proposition, Meecham et al. [68] applied the Wiener-Hermite expansion to study turbulence solutions of Burgers equations. They assumed that the initial velocity field is a homogeneous Gaussian process, and expanded the turbulent solution as a second order Wiener-Hermite expansion. This approach for Burgers turbulence was critically examined by many authors [13, 16, 81] and it was pointed out that the second order Wiener-Hermite expansion cannot

faithfully represent the dynamics of the Burgers solution. However, including higher order terms was technically very difficult based on Wiener's formulation, and no such attempt had been made. The Wiener's formalism for studying turbulence phenomena had been largely abandoned.

Based on Wiener's original idea, Cameron and Martin [10] developed a more explicit and intuitive formulation for the Wiener-Hermite expansion. Their development is based on an explicit discretization of the white noise process through its Fourier expansion, which was missed in Wiener's formalism. The approach developed by Cameron and Martin is much easier to understand and more convenient to use, and hence replaced Wiener's original formulation. This Fourier-Hermite expansion was commonly called Wiener chaos expansion (WCE) thereafter. Since Martin and Cameron's elegant work, Wiener chaos expansion has become a useful tool in stochastic analysis involving white noise (Brownian motion) [43, 45]. Rozovskii et al. [70, 72, 71] derived Wiener chaos propagator equations for several important SPDEs driven by Brownian motion forcing, such as stochastic parabolic equations and Navier-Stokes equations. Their original work laid down a new framework for simulating those stochastic PDEs numerically. For the nonlinear filtering problem, Lototsky et al. [60, 61] proposed a new numerical method for solving the Zakai equation based on its Wiener chaos expansion. The convergence rate of the method was also proved.

Another incarnation of Wiener chaos (or polynomial chaos) expansions as a numerical method is largely due to the original work by Ghanem and Spanos [39], where they designed a new numerical method for solving elliptic equations with random coefficients. Using Karhunen-Loeve expansion (KLE), they first expanded the random coefficients (a Gaussian field in space) as a series of Gaussian random variables. Then the random solution was represented as a Hermite expansion of those Gaussian random variables. By projecting the stochastic elliptic equations onto the probability space, they solved the equations numerically within the framework of finite element method. This method is called stochastic finite element method. Ghanem et al. further applied this method in modeling transport in porous media [35, 34], solid mechanics and structures [37, 38]. Using Fourier-Hermite expansion for modeling non-Gaussian processes is also investigated [90, 36, 33]. Ghanem's approach for solving stochastic elliptic equations was further developed and generalized by many other researchers, notably Babuska et al. [5, 4, 3], Schwab et al. [91, 32] and Keese el al. [67, 51]. Le Maitre et al. [64, 65] extended the applications of polynomial chaos

to thermo-fluid modelings. Xiu and Karniadakis [102] generalized the Hermite polynomial expansion to include other orthogonal polynomials, and used it to study flow-structure interactions [101, 103] and linear convection problems [48]. Zhang et al. [104, 63] combined moment perturbation method with polynomial chaos expansion, and used it to study the saturation flows in heterogeneous porous media.

For most SPDEs considered in the KLE-based approach, the source randomness is usually represented by a fixed number of random variables. This assumption is introduced either directly, or via a truncated Karhunen-Loeve expansion. The dimension of the source randomness will not change in time. In most cases, the randomness originates from modeling uncertainty in certain physical quantities, such as the permeability field [39, 34, 5, 4, 104], diffusivity or viscosity constants [64, 65], stiffness and damping coefficients [101, 103], boundary conditions or initial conditions [100, 68, 81]. Hence the source of the randomness usually contains only a small number of random variables. For those cases, the polynomial chaos expansion based on Karhunen-Loeve expansion provides powerful tools for the uncertainty quantification purpose. However, if the source randomness keeps changing and has very short correlation length in time, such as white noise, then the KLE-based method may fail to capture the randomness correctly.

1.3 Wiener Chaos Method for SPDEs with Brownian Motion Forcing

In this thesis, we will study how to solve SPDEs driven by Brownian motion forcing based on Wiener chaos expansion (WCE). A SPDE with Brownian motion forcing has the general form

$$u_t = L(u) + \sigma(x, t, u) \cdot dW_t, \tag{1.1}$$

where $L(u)$ is a differential operator (linear or nonlinear) in space, and W_t is a Brownian motion (scaler or vector). We will focus on stochastic equations arising in fluid mechanics, such as stochastic Burgers equations, Navier-Stokes equations, and nonlinear reaction-diffusion equations, with either additive or multiplicative random forcing. SPDEs with Brownian motion forcing are generally much more difficult to solve than the SPDEs considered in the KLE based approaches. In the SPDE (1.1), independent increments of the Brownian

motion forcing keep entering the PDE system, which makes the problem complicated even on very short time intervals. If we want to use a polynomial chaos type expansion for the random solution, we have to deal with a constant influx of new randomness. This is an essential difference from the SPDEs considered in the previous KLE-based polynomial chaos approach. Since the random forcing in (1.1) is white in time and has delta covariance function, its Karhunen-Loeve expansion will diverge (see Section 7.1 for details about the convergence of Karhunen-Loeve expansion). For those reasons, the KLE based polynomial chaos method simply can not be applied to solve the SPDEs (1.1) directly.

Instead of using Karhunen-Loeve expansion, we adopt the Martin-Cameron version of the Wiener chaos expansion in designing our numerical methods. The solution of the SPDE (1.1) is a random process depending on the realizations of the Brownian motion forcing

$$u = u(x,\, t;\, W_0^t), \tag{1.2}$$

where W_0^t denote the Brownian motion paths up to time t. A Brownian motion can be decomposed as a linear combination of independent Gaussian random variables. Consequently, the random solution (1.2) can be expanded as a Fourier-Hermite series of those Gaussian random variables. More specifically, we can represent the random solution as an infinite series

$$u(x,\, t;\, W_0^t) = \sum_{\alpha} u_\alpha(x,t)\, T_\alpha, \tag{1.3}$$

where $u_\alpha(x,t)$ are deterministic functions, and T_α are multi-variable Hermite polynomials of Gaussian random variables. The Gaussian random variables are constructed by discretizing the Brownian motion W_0^t as a Fourier expansion with respect to a set of deterministic basis functions in $L^2([0,t])$.

The Wiener chaos expansion (1.3) can separate the deterministic effects from the randomness in the random solution. By plugging it in the original SPDE, we can derive a deterministic system for the Wiener chaos coefficients u_α, which we call the *WCE propagator* of the corresponding SPDEs. The WCE propagator is equivalent to the original SPDE and captures the deterministic dynamics responsible for evolving the inherent randomness. Once the WCE propagator is derived, standard deterministic numerical methods can be applied to solve it efficiently. The main statistics, such as the mean, covariance, and higher order statistical moments, can be calculated by simple formulae involving only the deter-

ministic WCE coefficients. In the WCE approach, there is no randomness directly involved in the computations. One does not have to rely on pseudo random number generators, and there is no need to solve the stochastic PDEs repeatedly for many realizations. Instead, the propagator system is solved only once. Furthermore, the WCE provides a detailed description about the stochastic structures of the random solution. Particularly, we can construct realizations of the random solution very easily based on its Wiener chaos expansion. These properties of Wiener chaos expansions can be very useful in certain applications.

1.4 Summary of the Thesis

Based on the framework developed by Rozovskii et al. [70, 72, 71], we have derived explicit WCE propagators for stochastic Burgers equations, stochastic Navier-Stokes equations, and nonlinear stochastic reaction-diffusion equations. Since the WCE propagator is an infinite system, we design a sparse truncation strategy based on the asymptotic decaying rate of the WCE coefficients. The sparse truncation can reduce the dimensions of the WCE propagator substantially while retain the asymptotic convergence rate of the method. Without the compression provided by the sparse truncations, the resulting WCE propagator may have too many equations to be numerically solvable. By building a generic and robust numerical solver, we can easily handle a nonlinear WCE propagator comprised of hundreds of PDEs. By applying the WCE method to a wide range of problems in fluid mechanics, we convincingly demonstrate that the numerical method based on WCE is much more efficient than the MC simulation for short to moderate time integrations.

When the Brownian motion forcing is spatially independent, we have discovered semi-analytical solutions for both stochastic Burgers equations and stochastic Navier-Stokes equations. By applying the WCE method to these model equations, we demonstrate the convergence rate of the truncated WCE method both numerically and theoretically. The WCE is an expansion with respect to an infinite number of Gaussian random variables and arbitrary order of Hermite polynomials. The WCE truncation is characterized by the number K of Gaussian random variables and order N of the Hermite polynomials retained. For a stochastic Burgers equation with spatially independent additive forcing, we established that the error of a (K, N)-truncated WCE solution in a time interval Δt is given by the

formula

$$O\left[\left(\frac{\sigma^2\left(\Delta t\right)^3}{N+1}\right)^{\frac{N+1}{2}}+\left(\frac{\sigma^2\left(\Delta t\right)^3}{K^3}\right)^{1/2}\right], \tag{1.4}$$

where σ^2 is the variance of the Brownian motion forcing. The error estimate (1.4) demonstrates that the WCE method converges exponentially fast in the polynomial order N, while only algebraically in the number K of retained Gaussian random variables. The analysis also quantifies precisely how the convergence rate depends on the length of the time interval and the variability of the random forcing, which provides very useful guidelines for truncating WCE solutions in general cases. The error estimate (1.4) also implies that the WCE method converges globally at the rate $O(\sqrt{\Delta t})$. To our best knowledge, no other analytical result exists on the WCE error analysis for nonlinear equations.

From both the numerical experiments and the error analysis (1.4), we find that WCE methods converge faster in short time intervals. However, as the SPDEs are solved in longer time intervals, the required number of WCE coefficients will increase very fast, and the resulting WCE propagator quickly becomes unmanageable numerically. To overcome these difficulties, we propose to combine WCE methods with MC simulations for long time integrations. We first solve the SPDE on the long time intervals by a sparse WCE truncation with only a few dozens of coefficients. The sparse WCE solution is equivalent to a coarse-scale discretization of the random solution in the probability space and hence is not very accurate locally. To capture the unresolved small scales in the WCE solution, we further use MC simulations to refine the computations. To do that, we subtract the WCE solution from the true solution, and use MC simulations to correct the errors in the WCE solution. This procedure is possible because the WCE solution realizations are very easy to construct for given Brownian motion paths. Since the error of the sparse WCE solution has relatively small variance, the MC correction step can be made very efficient. This WCE-MC hybrid method is a generalization of Chorin's estimator for Monte Carlo integrations [12]. The WCE-MC hybrid method can handle SPDEs in much longer time intervals than the direct WCE method. Numerical results demonstrate that the new WCE-MC hybrid method is scores of times more efficient than the direct WCE method or the MC simulations in solving SPDEs in relatively long time intervals.

However, the WCE-MC method is still sensitive to the length of the time intervals. For arbitrarily long time integrations, the error in the sparse WCE solution may be very large

and hence the performance of the WCE-MC will not be much better than that of MC simulations, especially when the overhead in the hybrid method is taken into account. Designing more efficient WCE-type numerical methods that are not sensitive to the integration time is still a challenging and open problem. Nevertheless, the sparse WCE truncation strategy and the WCE-MC hybrid method already enable us to study some interesting problems numerically, such as:

Verification of invariant measure

For stochastic Burgers equations with additive forcing, it has been proven by Sinai [92, 93] that there exists a unique invariant measure (stationary distribution) for the random solution, due to the balance between the viscosity term and the random forcing. Using the WCE method, we are able to resolve the invariant measure of the random solution numerically, which is represented by a stationary Wiener chaos solution. The WCE solutions provide direct numerical evidence for the existence of the invariant measure and can be used to study its property numerically.

Simulating reaction-diffusion front propagation

Using the WCE-MC hybrid method, we are able to simulate the long time front propagation of a class of nonlinear reaction-diffusion equations in random shear flows. This simulation is otherwise extremely expensive if only the WCE method or MC simulation is used. Our numerical results demonstrate that the front speed obeys the quadratic enhancing law in white shear flows with small magnitudes. This numerical result confirms the conjecture by Xin et al. [76] about the precise asymptotical relations of the front speed in white-in-time shear flows.

1.4.1 Stochastic Elliptic Equations and Uncertainty Quantification

In the second part of the thesis, we reexamine the KLE-based polynomial chaos approach for solving stochastic elliptic equations (Ghanem et al. [39]). The new ingredient of our analysis is an explicit and analytical formula for the stiffness coefficients. Based on the explicit derivation, we argue that there is no need to truncate the chaos expansions of the random permeability, which is contrary to the practice in previous work [4, 5, 32, 104]. Without truncating the random permeability in the stochastic elliptic equation, we can preserve the lower uniform elliptic condition and avoid a few serious technical difficulties.

To solve the resulting elliptic system for the polynomial chaos expansion, we generalize the upscaling formulation [19, 47] for scaler elliptic equations to elliptic systems, and obtain an upscaled elliptic system for the polynomial chaos expansion. Combining upscaling with polynomial chaos expansion is a novel practice. The upscaled formulation provides an efficient way for solving the stochastic elliptic equations by polynomial chaos.

The central work of the second part of the thesis is to develop an efficient multiscale strategy for uncertainty quantification in reservoir modeling. Uncertainties in the detailed description of reservoir permeability are major contributors to the errors in reservoir performance forecasting. To quantify these uncertainties, one can estimate the probability distribution of the model predictions based on a large number of reservoir realizations. It is essential that the permeability realizations adequately reflect the uncertainty in the reservoir properties. To reduce the uncertainty and improve the model prediction, one should sample the reservoir permeability field conditioned on dynamical production data.

The prediction of permeability fields based on dynamic data is a challenging problem because permeability fields are typically defined on a large number of grid blocks. The Markov chain Monte Carlo (MCMC) method and its modifications have been used previously to sample the posterior distribution of the permeability field. However, due to the large dimension of the permeability fields, the MCMC method has a very small acceptance rate. To accept just a few permeability samples, a large number of proposals need to be tested. Since each acceptance-rejection test involves a forward simulation of a nonlinear PDE system on the fine grid, the naive MCMC method is extremely expensive.

To sample the permeability fields conditioned on dynamical production data, we designed an efficient two-stage Markov chain Monte Carlo (MCMC) method. Using a coarse-scale model based on multiscale finite element, we first test each permeability proposal on coarse-grid, which only involves solving the nonlinear model equations on coarse-scale. If the permeability proposal is accepted by the coarse-grid test, then a fine-scale simulation will be performed at the second stage to determine its accurate acceptance rate. Otherwise, the proposal will be rejected by the coarse-scale test and new permeability sample will be generated. The coarse-scale model can effectively filter unacceptable permeability samples and avoid the expensive fine-scale simulations for them. The key step of the new method is to construct an effective coarse-scale model, which can be done using the idea of multiscale finite element method. Using the multiscale bases precomputed at the first step, we can

conduct the fine-scale simulation at the second stage very efficiently. The two-stage MCMC method seeks to exploit the multiscale structure of the problem and use coarse-scale models to precondition fine-scale simulations. Extensive numerical experiments are performed, and the new method is shown to be ten times more efficient than the regular MCMC sampling techniques for both single- and two-phase flow simulations.

Based on the upscaled polynomial chaos method for stochastic elliptic equations, we designed a new coarse-scale model for the reservoir model equations. By solving the upscaled polynomial chaos system numerically, we can construct an explicit coarse-scale distribution for the permeability fields. Using this coarse-scale distribution, we can generate problem-adapted permeability samples using the Langevin diffusion algorithm. As in the previous two-stage MCMC method, the Langevin proposals will be first tested by the coarse-scale model. The only difference is that no coarse-scale forward simulation is needed in this method, since the coarse-scale distribution is explicitly available from the solution of the upscaled polynomial chaos system. If the Langevin proposal is accepted by the coarse-scale test, then a fine-scale simulation will be conducted to estimate the accurate acceptance rate. The coarse-scale distribution serves as a preconditioner to the fine scale simulation, and plays the same role as the coarse-scale model based on multiscale finite element method. Compared with the previous two-stage MCMC method using a deterministic coarse-scale model, the Langevin algorithm-based two-stage MCMC method can avoid solving the nonlinear two-phase flow problems repeatedly on coarse scales. Instead, the upscaled stochastic system is solved only once. Preliminary numerical results show that this new method is quite promising, and further research is still going along this direction.

Chapter 2

Hermite Polynomials and Wiener Chaos Expansions

In this chapter, we will review the main facts of the Hermite polynomials and present the theory of the Wiener chaos expansions. A few useful lemmas are proved in this chapter, which are the main tools we will use in the later studies.

2.1 Hermite Polynomials

Consider functions on the real axis $\mathbf{R} = (-\infty, +\infty)$ equipped with the Gaussian measure

$$\mu(dx) = \rho(x)dx, \qquad \rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \tag{2.1}$$

where dx is the Lebesgue measure. We denote the space of square integrable functions with the Gaussian measure μ as

$$L^2(\mathbf{R}, \mu) = \left\{ f(x);\ \int_{-\infty}^{+\infty} f(x)^2 \mu(dx) < \infty \right\}.$$

The inner product on this space is defined as

$$(f,\ g)_\mu = \int_{-\infty}^{+\infty} f(x)g(x)\mu(dx) = \int_{-\infty}^{+\infty} f(x)g(x)\rho(x)dx.$$

Suppose ξ is a standard Gaussian random variable with distribution $N(0, 1)$, then

$$(f,\ g)_\mu = E[f(\xi)g(\xi)],$$

where E denotes the expectation operator. So the Hilbert space $L^2(\mathbf{R}, \mu)$ can also be interpreted as the space of functions of a unit Gaussian random variable with finite variance.

The un-normalized Hermite polynomials are defined as

$$P_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right), \qquad n = 0,\ 1,\ 2, \ldots . \tag{2.2}$$

$P_n(x)$ are orthogonal polynomials with respect to the Gaussian measure:

$$(P_n,\ P_m)_\mu = E[P_n(\xi) P_m(\xi)] = n!\, \delta_{n,m}.$$

Hence the normalized Hermite polynomials are defined as

$$H_n(x) = \frac{P_n(x)}{\sqrt{n!}} = (n!)^{-\frac{1}{2}} (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right). \tag{2.3}$$

It is well known [15, 94] that $\{H_n(x);\ n = 0, 1, \ldots\}$ are a complete orthonormal basis in the Hilbert space $L^2(\mathbf{R},\ \mu)$. Since $H_0(x) = 1$, particularly we have

$$E[H_n(\xi)] = \int_{-\infty}^{+\infty} H_n(x) d\mu(x) = (H_n,\ 1)_\mu = 0, \qquad \text{if} \qquad n \neq 0. \tag{2.4}$$

So a Hermite polynomial of a standard Gaussian random variable has zero mean if its order is bigger than zero.

Like most orthogonal polynomials, Hermite polynomials have a generating function [15]:

$$\psi(x,\ z) = e^{-\frac{z^2}{2} + xz}. \tag{2.5}$$

Expanding $\psi(x, z)$ into a Taylor series of variable z (treating x as a parameter), we have

$$\psi(x, z) = \sum_{n=0}^{\infty} \frac{\partial^n \psi(x, z)}{\partial z^n}\Big|_{z=0} \left(\frac{z^n}{n!} \right).$$

On the other hand,

$$\psi(x,\ z) = e^{\frac{x^2}{2}} e^{-\frac{(z-x)^2}{2}} = e^{\frac{x^2}{2}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right) z^n = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} z^n. \tag{2.6}$$

So the coefficients of the Taylor expansion of $\psi(x,\ z)$ are exactly the un-normalized Hermite

polynomials.

$$P_n(x) = \frac{\partial^n \psi(x,z)}{\partial z^n}\Big|_{z=0}. \tag{2.7}$$

The generating function (2.5) is a powerful tool in studying the properties of Hermite polynomials. We will use it repeatedly in our later derivations.

Differentiating (2.2) directly we get

$$P_n'(x) = x\, P_n(x) - P_{n+1}(x). \tag{2.8}$$

On the other hand, by differentiating both sides of (2.6) with respect to x we have

$$\frac{\partial}{\partial x}\psi(x,\ z) = z\psi(x,\ z) = \sum_{n=0}^{\infty} \frac{P_n(x)}{n!} z^{n+1} = \sum_{n=0}^{\infty} \frac{P_n'(x)}{n!} z^n.$$

Shifting the summation index and comparing the coefficients of z^n suggests that

$$P_n'(x) = nP_{n-1}(x). \tag{2.9}$$

Combining (2.8) and (2.9) together, we get the recursive relation of the un-normalized Hermite polynomials

$$P_{n+1}(x) - xP_n(x) + nP_{n-1}(x) = 0, \qquad n = 1,\ 2, \ldots, \tag{2.10}$$

with $P_{-1}(x) = 0,\ P_0(x) = 1$. From the recursive relation, we can easily get the first six un-normalized Hermite polynomials:

$$\begin{aligned}
P_0(x) &= 1, \\
P_1(x) &= x, \\
P_2(x) &= x^2 - 1, \\
P_3(x) &= x^3 - 3x, \\
P_4(x) &= x^4 - 6x^2 + 3, \\
P_5(x) &= x^5 - 10x^3 + 15x.
\end{aligned}$$

For the normalized Hermite polynomials, the recursive relation becomes

$$\sqrt{n+1}H_{n+1}(x) - xH_n(x) + \sqrt{n}H_{n-1}(x) = 0 \tag{2.11}$$

$$H_{-1}(x) = 0, \qquad H_0(x) = 1, \tag{2.12}$$

and the differential rule is

$$H_n'(x) = \sqrt{n}H_{n-1}(x). \tag{2.13}$$

As the reverse of formula (2.13), we have the following useful fact:

Lemma 2.1

$$\rho(x)H_{n+1}(x) = \frac{1}{\sqrt{n+1}}\frac{d}{dx}\left[\rho(x)H_n(x)\right], \tag{2.14}$$

where $\rho(x)$ is the Gaussian density (2.1).

Proof By direct differentiation and the recursive relation (2.11) we have

$$\begin{aligned}\frac{d}{dx}\left[\rho(x)\,H_n(x)\right] &= \rho'(x)\,H_n(x) + \rho(x)\,H_n'(x)\\ &= -x\,\rho(x)\,H_n(x) + \sqrt{n}\,\rho(x)\,H_{n-1}(x)\\ &= \sqrt{n+1}\,\rho(x)\,H_{n+1}.\end{aligned}$$

Dividing both sides by $\sqrt{n+1}$, we immediately get (2.14). □

By differentiation, the order of the Hermite polynomial decreases in formula (2.13), while it increases in formula (2.14). The formula (2.14) plays a crucial role in obtaining the decaying rate of the coefficients of the Fourier-Hermite expansion.

Since the product of two Hermite polynomials is still a polynomial, it can be expanded as a linear combination of Hermite polynomials. This simple fact turns out to be very useful in the Hermite spectral method for nonlinear equations, particularly in the Wiener chaos expansion method. Next we will derive an analytical formula for expanding the product of two arbitrary Hermite polynomials. The expansion of products of more than two Hermite polynomials can be obtained using this formula recursively.

Lemma 2.2 *For any nonnegative integer α and β, denote $\alpha \wedge \beta = \min\{\alpha,\ \beta\}$. We have*

$$H_\alpha(x)H_\beta(x) = \sum_{p\le\alpha\wedge\beta} B(\alpha,\ \beta,\ p)\ H_{\alpha+\beta-2p}(x), \tag{2.15}$$

where

$$B(\alpha,\ \beta,\ p) = \left[\binom{\alpha}{p}\binom{\beta}{p}\binom{\alpha+\beta-2p}{\alpha-p}\right]^{\frac{1}{2}}. \tag{2.16}$$

Proof From equation (2.6) we have

$$\psi(x,\ z)\psi(x,\ w) = \sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{\infty}\frac{P_\alpha(x)P_\beta(x)}{\alpha!\beta!}z^\alpha w^\beta. \tag{2.17}$$

On the other hand,

$$\begin{aligned}\psi(x,\ z)\psi(x,\ w) &= e^{-\frac{z^2+w^2}{2}+(z+w)x} = e^{zw}e^{\frac{x^2}{2}}e^{-\frac{(z+w-x)^2}{2}}\\ &= \sum_{p=0}^{\infty}\frac{(z\,w)^p}{p!}\sum_{k=0}^{\infty}\frac{P_k(x)}{k!}(z+w)^k\\ &= \sum_{p=0}^{\infty}\ \sum_{k=0}^{\infty}\frac{P_k(x)}{p!}\sum_{0\le m\le k}\frac{1}{k!}\binom{k}{m}\,z^{m+p}w^{k+p-m}.\end{aligned}$$

Let $k = m+\nu$, then $m \le k$ is equivalent to $\nu \ge 0$. The above formula can be rewritten as

$$= \sum_{p=0}^{\infty}\ \sum_{m=0}^{\infty}\ \sum_{\nu=0}^{\infty}\frac{P_{m+\nu}(x)}{p!\ m!\ \nu!}z^{m+p}w^{\nu+p}.$$

Denote $m+p=\alpha$, $\nu+p=\beta$. Since $m=\alpha-p\ge 0$, $\nu=\beta-p\ge 0$, we have $p\le\alpha\wedge\beta$. The above summation changes to

$$\begin{aligned}&= \sum_{\alpha=0}^{\infty}\ \sum_{\beta=0}^{\infty}\left(\sum_{\substack{m+p=\alpha\\ \nu+p=\beta}}\frac{P_{m+\nu}(x)}{p!\ m!\ \nu!}\right)z^\alpha w^\beta\\ &= \sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{\infty}\sum_{p\le\alpha\wedge\beta}\frac{P_{\alpha+\beta-2p}(x)}{p!\ (\alpha-p)!\ (\beta-p)!}z^\alpha w^\beta.\end{aligned}$$

So we have

$$\psi(x,\ z)\psi(x,\ w) = \sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{\infty}\sum_{p\le\alpha\wedge\beta}\frac{P_{\alpha+\beta-2p}(x)}{p!\ (\alpha-p)!\ (\beta-p)!}z^\alpha w^\beta. \tag{2.18}$$

Comparing the above equation with (2.17) we get

$$P_\alpha(x)P_\beta(x) = \sum_{p\le\alpha\wedge\beta} \frac{\alpha!\beta!}{p!\,(\alpha-p)!\,(\beta-p)!} P_{\alpha+\beta-2p}(x). \tag{2.19}$$

Plugging $P_n(x) = (n!)^{\frac{1}{2}} H_n(x)$ in the above formula immediately yields formula (2.15). □

As we have pointed out in (2.4), $E[H_n(\xi)] = 0$ if $n > 0$ and ξ is a standard Gaussian random variable. Next we prove a related result, which will be used in later applications.

Lemma 2.3 *Suppose $H_n(x)$ is the nth order normalized Hermite polynomial and ξ is a standard Gaussian random variable, then*

$$E[H_n(\xi + a)] = \frac{a^n}{\sqrt{n!}},$$

where a is an arbitrary constant.

Proof From the generating function (2.6) we have

$$\psi(x + a,\ z) = \sum_{n=0}^{\infty} \frac{P_n(x + a)}{n!} z^n.$$

On the other hand,

$$\begin{aligned} \psi(x + a, z) &= e^{-\frac{z^2}{2}+xz} e^{az} = \left(\sum_{i=0}^{\infty} \frac{a^i}{i!} z^i\right)\left(\sum_{j=0}^{\infty} \frac{P_j(x)}{j!} z^j\right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} \frac{a^i}{i!} \frac{P_{n-i}(x)}{(n-i)!}\right) z^n. \end{aligned}$$

Comparing the above two formulae we conclude that

$$P_n(x + a) = \sum_{i=0}^{n} \frac{a^i}{i!} \frac{n!}{(n-i)!} P_{n-i}(x).$$

Note that $E[P_k(\xi)] = \delta_{k,0}$, we have

$$E[H_n(\xi + a)] = \frac{E[P_n(\xi + a)]}{\sqrt{n!}} = \frac{a^n}{\sqrt{n!}}.$$

□

2.2 Fourier-Hermite Expansions of Functions of Gaussian Random Variables

Since Hermite polynomials are orthonormal bases in $L^2(\mathbf{R}, \mu)$, for any function $f(x) \in L^2(\mathbf{R}, \mu)$ there exists a Fourier-Hermite expansion

$$f(x) = \sum_{n=0}^{\infty} f_n H_n(x), \qquad f_n = \int_{-\infty}^{\infty} f(x)\, H_n(x) \mu(dx).$$

On the other hand, we can think of $f \in L^2(\mathbf{R}, \mu)$ as a function of a unit Gaussian random variable ξ with $E[f^2(\xi)] < +\infty$. So the random function $f(\xi)$ has the Fourier-Hermite expansion

$$f(\xi) = \sum_{n=0}^{\infty} f_n H_n(\xi), \qquad f_n = E[f(\xi) H_n(\xi)]. \tag{2.20}$$

By the definition of f_0, we have

$$E[f(\xi)] = f_0.$$

This is consistent with the Fourier-Hermite expansion (2.20) since $E[H_n(\xi)] = 0$ for $n > 0$. From Parseval's theorem we have

$$E[f^2(\xi)] = \sum_{n=0}^{\infty} |f_n|^2. \tag{2.21}$$

Denote the truncated Fourier-Hermite expansion of $f(\xi)$ as

$$f_N(\xi) = \sum_{n=0}^{N} f_n H_n(\xi). \tag{2.22}$$

The truncated Hermite expansion (2.22) converges very fast if the function $f(x)$ is very smooth. Traditionally, the Fourier-Hermite expansion is formulated for deterministic functions in the regular space $L^2(\mathbf{R})$ with Lebesgue measure. Define the Hermite functions as

$$\psi_n(x) = \rho^{1/2}(x) H_n(x),$$

where $\rho(x)$ is the Gaussian distribution (2.1). The Hermite functions $\{\psi_n(x)\}$ form an

orthonormal basis in $L^2(\mathbf{R})$. For any function $g(x) \in L^2(\mathbf{R})$, there exists the expansion

$$g(x) = \sum_n g_n \psi_n(x), \qquad f_n = \int_R g(x)\psi_n(x)dx. \tag{2.23}$$

The convergence rate of the Fourier-Hermite expansion (2.23) depends not only on the smoothness (location of singularities) of $g(x)$, but also upon the rate at which $g(x)$ decays to zero as $|x| \to \infty$. For example, if $g(x)$ is an entire function and decays exactly like a Gaussian kernel

$$g(x) = O\left(e^{-p|x|^2}\right),$$

then the expansion (2.23) has the fastest (geometric) convergence rate [8]

$$g_n = O\left(e^{-qn}\right).$$

To relate the expansion (2.20) to (2.23), for any function $f(x) \in L^2(\mathbf{R}, \mu)$ denote

$$F(x) = f(x)\rho^{1/2}(x),$$

then the expansion (2.20) is equivalent to

$$F(x) = \sum_n F_n \psi_n(x), \qquad F_n = \int_R F(x)\psi_n(x)dx.$$

Since the new function $F(x)$ is exactly a Gaussian-type function, the above expansion (and (2.20)) has the geometric convergence rate if $f(x)$ is also an entire function. For functions that are not entire, the convergence rate of the Fourier-Hermite coefficients depend on the smoothness of the function $f(x)$. As an illustration, we have the following simple results:

Lemma 2.4 *Suppose $f(x) \in C^k(\mathbf{R})$, then the Fourier-Hermite coefficients $f_n = (f, H_n)_\mu$ decay as*

$$f_n = \begin{cases} \frac{(-1)^n}{\sqrt{n!}} E_\mu\left[f^{(n)}(\xi)\right], & n \le k, \\ \frac{(-1)^k}{\sqrt{n\ldots(n-k+1)}} E_\mu\left[f^{(k)}(\xi)\right], & n > k, \end{cases} \tag{2.24}$$

where E_μ is the expectation with respect to the unit Gaussian measure μ. If the function

$f(x)$ is infinitely smooth, then f_n decays exponentially

$$f_n = O\left(e^{-qn}\right). \tag{2.25}$$

Proof Based on formula (2.14) we have

$$\begin{aligned} f_n &= \int_{-\infty}^{+\infty} f(x) H_n(x) \rho(x) dx \\ &= \frac{1}{\sqrt{n}} \int_{-\infty}^{+\infty} f(x) \frac{d}{dx} [H_{n-1}(x)\rho(x)] \\ &= \frac{-1}{\sqrt{n}} \int_{-\infty}^{+\infty} f'(x) H_{n-1}(x) \rho(x) dx. \end{aligned}$$

By induction we can easily get formula (2.24). If $f(x)$ is infinitely smooth, using the Stirling formula

$$n! = O\left(n^{n+1/2}\, e^{-n}\right),$$

we can get the estimate (2.25). □

The convergence rate of the Fourier-Hermite expansion is quite complicated in general. John Boyd gives a nice survey of the convergence results of Fourier-Hermite expansions in his classic book [8]. Using the convergence rate of the Fourier-Hermite coefficients, it is not difficult to obtain the mean convergence rate for the truncated expansion (2.22).

The Fourier-Hermite expansion (2.20) can be easily extended to multi-dimensions. For a finite index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ with nonnegative integer components, define the multi-variable Hermite polynomial by the tensor product

$$H_\alpha(x) = \prod_{i=1}^{d} H_{\alpha_i}(x_i).$$

Then $\{H_\alpha(x)\}$ is an orthonormal basis in the Hilbert space $L^2(R^d, \mu^d)$, where μ^d is the $d-$multiple Gaussian measure on R^d. Denote $\xi = (\xi_1, \ldots, \xi_d)$ as the standard Gaussian random vector with independent components. Suppose $u(\xi)$ is a function of the random variables ξ with $E|u^2(\xi)| < +\infty$, then $u(\cdot) \in L^2(R^d, \mu^d)$ and there exists a Fourier-Hermite expansion

$$u(\xi) = \sum_\alpha u_\alpha \, H_\alpha(\xi) \quad \text{where} \quad u_\alpha = E[u(\xi) H_\alpha(\xi)]. \tag{2.26}$$

Furthermore, we have

$$E[u(\xi)] = u_{\mathbf{0}}, \qquad E|u^2(\xi)| = \sum_{\alpha} |u_\alpha|^2.$$

This expansion is valid for any finite dimensional space $L^2(R^d, \mu^d)$ with $d < \infty$.

2.3 Wiener Chaos Expansions of Functionals of Brownian Motions

In this thesis, we are interested in a class of stochastic partial differential equations (SPDEs)

$$u_t(x,t) = L(u) + \sigma\, \dot{W}(t), \tag{2.27}$$

where $L(u)$ is a general linear or nonlinear differential operator in spatial variables, and $W(t)$ is a Brownian motion. The derivative of the Brownian motion $\dot{W}(t)$ is usually understood as a model of the white noise. The solution of (2.27) not only depends on the spatial-time variable (x,t), but also the Brownian motion path. In other words, $u = u(x,t)$ is a functional of $\{W(s), 0 \le s \le t\}$, which comprises all the possible Brownian motion realizations up to time t. We want to generalize the finite dimensional Fourier-Hermite expansion to the stochastic solution $u(x,t)$ with respect to its random dependence. Although Brownian motion $\{W(s), 0 \le s \le t\}$ is by no means a finite number of Gaussian random variables, the goal can be achieved in a relatively straightforward way.

For any fixed time $t > 0$, assume $m_i(s)$, $i = 1, 2, \ldots$, are a set of complete orthonormal bases in the Hilbert space $L^2([0,\ t])$. Define the Ito integral (see Appendix A for the definition)

$$\xi_i = \int_0^t m_i(s) dW(s), \qquad i = 1, 2, \ldots. \tag{2.28}$$

Next we use W_s to denote $W(s)$ when no confusion is incurred. When we want to refer to the derivative of $W(s)$, we always write explicitly as $\dot{W}(s)$. Naively, we can interpret (2.28) as the projection of the white noise $\dot{W}(s)$ onto the L^2 basis function $m_i(s)$, because $\xi_i = (m_i,\ \dot{W}_s)$. Obviously, ξ_i, $i = 1, 2, \ldots$ are Gaussian random variables and $E(\xi_i) = 0$.

Furthermore, from the isometry property (A.5) of the Ito integrals, we have

$$E[\xi_i\,\xi_j] = \int_0^t m_i(s)m_j(s)ds = \delta_{i,j}.$$

So ξ_i, $i = 1, 2, \ldots$, are independent standard Gaussian random variables.

As a useful fact, the Brownian motion path $\{W(s), 0 \le s \le t\}$ can be decomposed as a linear combination of the Gaussian random variables ξ_i. Suppose the orthonormal bases $m_i(s)$ in $L^2([0,\ t])$ are chosen as the trigonometric functions

$$m_1(s) = \frac{1}{\sqrt{t}}, \qquad m_k(s) = \sqrt{\frac{2}{t}}\cos\left(\frac{(k-1)\pi s}{t}\right), \quad k = 2, 3, \ldots, \quad 0 \le s \le t. \tag{2.29}$$

We have the following result:

Theorem 2.1 *The Brownian motion* $\{W_s;\ 0 \le s \le t\}$ *has the Fourier expansion*

$$W(s) = \sum_{i=1}^{\infty} \xi_i \int_0^s m_i(\tau)d\tau, \qquad 0 \le s \le t, \tag{2.30}$$

and the expansion (2.30) *converges in the mean square sense for all* $s \le t$*:*

$$E\left[W(s) - \sum_{i=1}^{N} \xi_i \int_0^s m_i(\tau)d\tau\right]^2 \le \frac{t}{\pi N}. \tag{2.31}$$

Proof Denote the characteristic function of the interval $[0,\ s]$ by $\chi_{[0,\ s]}(\tau)$. Obviously $\chi_{[0,\ s]}(\tau) \in L^2([0,\ t])$ and has the expansion

$$\chi_{[0,\ s]}(\tau) = \sum_{i=1}^{\infty} (\chi_{[0,s]},\ m_i)\, m_i(\tau) = \sum_{i=1}^{\infty} m_i(\tau) \int_0^s m_i(\tau)d\tau.$$

It follows that

$$\begin{aligned} W(s) &= \int_0^s dW_\tau = \int_0^t \chi_{[0,\ s]}(\tau)dW_\tau \\ &= \int_0^t \left(\sum_{i=1}^{\infty} m_i(\tau) \int_0^s m_i(\tau)d\tau\right) dW_\tau \\ &= \sum_{i=1}^{\infty} \xi_i \int_0^s m_i(\tau)d\tau. \end{aligned}$$

Denote the mean square error of the finite-term truncation of (2.30) as

$$I(N) = E\left[W(s) - \sum_{k=1}^{N} \xi_k \int_0^s m_k(\tau) d\tau\right]^2.$$

Since ξ_i are independent unit Gaussian variables, we have

$$\begin{aligned} I(N) &= \sum_{k=N+1}^{\infty} \left(\int_0^s m_k(\tau) d\tau\right)^2 = \sum_{k=N}^{\infty} \frac{2t}{k^2\pi^2} \sin^2\left(\frac{k\pi s}{t}\right) \\ &\leq \sum_{k=N}^{\infty} \frac{2t}{k^2\pi^2} \leq \frac{2t}{\pi^2}\left(\frac{1}{N^2} + \int_N^{\infty} \frac{dx}{x^2}\right) \\ &\leq \frac{t}{\pi N}. \end{aligned}$$

So the expansion (2.30) converges in the mean square sense uniformly for all $s \leqslant t$. □

If $m_i(s), i = 1, 2, \ldots$, are chosen as Haar wavelets, then the expansion (2.30) is exactly the Levy-Ciesielski construction [27] of the Brownian motion, and (2.30) converges uniformly in s for almost all Brownian motion realizations.

Since the random solution $u(x,t)$ of (2.27) is a functional of the Brownian motion path $\{W(s);\ 0 \leq s \leq t\}$, which in turn is a linear combination of unit Gaussian random variables $\xi_i, i = 1, 2, \ldots$, we can formally interpret $u(x,t)$ as a function of the standard Gaussian random variables ξ_i:

$$u(x,t) = U(x,t;\ \xi_1, \xi_2, \ldots, \xi_n, \ldots).$$

Analogous to the finite dimensional Fourier-Hermite expansion (2.26), the solution $u(x,t)$ admits a similar expansion. To make this analogy more precise and rigorous, we need to introduce some notations first.

Denote a set of multi-indices as

$$\mathcal{J} = \Big\{\alpha = (\alpha_i,\ i \geq 1) \ \Big|\ \alpha_i \in \{0, 1, 2 \ldots\}, \quad |\alpha| = \sum_{i=1}^{\infty} \alpha_i < \infty\Big\}.$$

So $\mathcal{J}$ is the set of integer indices with only finite number of nonzero components. For $\alpha \in \mathcal{J}$, define the Hermite polynomial of $\xi_i, i = 1, 2, \ldots$ by the tensor product

$$T_\alpha(\xi) = \prod_{i=1}^{\infty} H_{\alpha_i}(\xi_i). \tag{2.32}$$

Since $|\alpha| < \infty$, the above product comprises only finite number of factors. The random functions $T_\alpha(\xi)$ are often called *Wick polynomials*. The order of the polynomial $T_\alpha(\xi)$ is defined as $|\alpha|$. For convenience, we set $T_\alpha(\xi) = 0$ if the index α has negative components $\alpha_i < 0$.

We shall summarize some of the important properties of the Wick polynomials. For any $\alpha, \beta \in \mathcal{J}$, denote $\alpha \wedge \beta = (\min\{\alpha_i,\ \beta_i\}, i \geq 1)$, and $\alpha \vee \beta = (\max\{\alpha_i,\ \beta_i\}, i \geq 1)$. We say $\beta \leq \alpha$ if $\beta_i \leq \alpha_i$ for all $i \geq 1$. The operation $\alpha \pm \beta$ is also defined component-wise. As a convention, we denote $\alpha! = \prod_i \alpha_i!$. We have the following facts:

(i) $\{T_\alpha(\xi), \quad \alpha \in \mathcal{J}\}$ are orthonormal bases:

$$E(T_\alpha T_\beta) = \begin{cases} 0 & \text{if} \quad \alpha \neq \beta, \\ 1 & \text{if} \quad \alpha = \beta. \end{cases}$$

(ii) Since $T_{\mathbf{0}}(\xi) = 1$, where $\mathbf{0} = (0, 0, \ldots)$ is the zero index, we have $E(T_{\mathbf{0}}) = 1$, and

$$E(T_\alpha) = E(T_\alpha T_{\mathbf{0}}) = 0 \quad \text{if} \quad \alpha \neq \mathbf{0}.$$

(iii) As a generalization of Lemma 2.2, it can be shown that

$$T_\alpha(\xi) T_\beta(\xi) = \sum_{p \leq \alpha \wedge \beta} B(\alpha,\ \beta,\ p)\, T_{\alpha+\beta-2p}(\xi). \tag{2.33}$$

As a nontrivial generalization of the finite Fourier-Hermite expansion, Cameron and Martin [10] proved the following theorem, which forms the theoretical foundation of the numerical algorithm we will discuss later.

Theorem 2.2 (Cameron-Martin) *[10] Assume that for any x and $s \leq t$, the solution $u(x, s)$ of (2.27) is a functional of the Brownian motion $\{W_s;\ 0 \leq s \leq t\}$ with $E\,|u(x, s)|^2 < \infty$, then $u(x, s)$ has the following Fourier-Hermite expansion:*

$$u(x, s) = \sum_{\alpha \in \mathcal{J}} u_\alpha(x, s) T_\alpha(\xi), \qquad u_\alpha(x, s) = E[u(x, s) T_\alpha(\xi)], \tag{2.34}$$

where $T_\alpha(\xi)$ are the random Wick polynomials defined by (2.28) and (2.32). Furthermore,

the first two statistical moments of $u(x,s)$ are given by: the mean

$$E[u(x,s)] = u_0(x,s), \tag{2.35}$$

and the variance

$$E[u^2(x,s)] = \sum_{\alpha\in\mathcal{J},\, \alpha\neq\mathbf{0}} |u_\alpha(x,s)|^2. \tag{2.36}$$

When constructing the Wick polynomials $T_\alpha(\xi)$, we first fix a time $t > 0$, then project the Brownian motion $\{W(s), 0 \le s \le t\}$ onto $L^2([0,t])$ and obtain the Gaussian random variables $\xi_i, i = 1, 2, \ldots$. So the Wick polynomials $T_\alpha(\xi)$ always carry a parameter t implicitly. The expansion (2.34) is valid only for $s \le t$.

The Fourier-Hermite series (2.34) is usually called the Wiener chaos expansion (WCE) of $u(x,s)$. The WCE is a spectral expansion of the stochastic solution in the probability space. It represents the randomness of the solution analytically by a set of random bases with deterministic coefficients. If we have a way to compute the deterministic WCE coefficients $u_\alpha(x,s)$ dynamically, we can recover all the probability information of u. For example, we can compute its statistical moments, construct pathwise solutions for given Brownian motion realizations, or even derive its probability density function.

For index α with $|\alpha| = 1$, the Wick polynomials $T_\alpha(\xi) = \xi_i$ are Gaussian. We can separate the WCE (2.34) into two parts:

$$u(x,s) = \sum_{|\alpha|\le 1} u_\alpha(x,s)T_\alpha(\xi) + \sum_{|\alpha|\ge 2} u_\alpha(x,s)T_\alpha(\xi). \tag{2.37}$$

The first part is the Gaussian approximation of $u(x,s)$, while the second part is the higher order non-Gaussian terms. Gaussian approximations are widely used in literatures for studying SPDEs. It would be interesting to see how the Gaussian approximations fit in the framework of the Wiener chaos expansion. We will demonstrate by numerical examples that the Gaussian approximation is usually a bad approximation for nonlinear problems, especially when high order statistics are concerned.

The WCE is a spectral expansion in the probability space. Usually, a spectral method is most convenient for linear problems. However, the following theorem enables us to handle polynomial nonlinearities rather easily:

Theorem 2.3 *Suppose u, v have Wiener chaos expansions*

$$u = \sum_{\alpha\in\mathcal{J}} u_\alpha\, T_\alpha(\xi), \qquad v = \sum_{\beta\in\mathcal{J}} v_\beta\, T_\beta(\xi).$$

If $E(|uv|^2) < \infty$, then the product $u\,v$ has the Wiener chaos expansion

$$uv = \sum_{\theta\in\mathcal{J}} \Big(\sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\theta} C(\theta,\ \beta,\ p)\, u_{\theta-\beta+p}\, v_{\beta+p} \Big)\, T_\theta(\xi), \tag{2.38}$$

where

$$C(\theta,\ \beta,\ p) = \left[\binom{\theta}{\beta} \binom{\beta+p}{p} \binom{\theta-\beta+p}{p} \right]^{\frac{1}{2}}. \tag{2.39}$$

Proof From property (2.33), we have

$$\begin{aligned} uv &= \sum_{\alpha\in\mathcal{J}} \sum_{\beta\in\mathcal{J}} u_\alpha v_\beta T_\alpha T_\beta \\ &= \sum_{\alpha\in\mathcal{J}} \sum_{\beta\in\mathcal{J}} u_\alpha\, v_\beta \sum_{p\le\alpha\wedge\beta} \binom{\alpha}{p} \binom{\beta}{p} p! \frac{\sqrt{(\alpha+\beta-2p)!}}{\sqrt{\alpha!\,\beta!}} T_{\alpha+\beta-2p}. \end{aligned}$$

Let $\tilde\alpha = \alpha - p$, $\tilde\beta = \beta - p$, then $p \le \alpha\wedge\beta$ is equivalent to $\tilde\alpha,\ \tilde\beta \ge 0$. Alternatively, $\alpha = \tilde\alpha + p$, $\beta = \tilde\beta + p$ and the above summation can be rewritten as

$$= \sum_{\tilde\alpha\in\mathcal{J}} \sum_{\tilde\beta\in\mathcal{J}} \sum_{p\in\mathcal{J}} u_{\tilde\alpha+p} u_{\tilde\beta+p} \binom{\tilde\alpha+p}{p} \binom{\tilde\beta+p}{p} p! \frac{\sqrt{(\tilde\alpha+\tilde\beta)!}}{\sqrt{(\tilde\alpha+p)!\,(\tilde\beta+p)!}} T_{\tilde\alpha+\tilde\beta}.$$

For simplicity, we still denote $\alpha = \tilde\alpha$ and $\beta = \tilde\beta$. Let $\theta = \alpha + \beta$, then $\alpha = \theta - \beta \geqslant 0$ and $0 \leqslant \beta \le \theta$. The above summation is equivalent to

$$\begin{aligned} &= \sum_{\theta\in\mathcal{J}} \sum_{\alpha+\beta=\theta} \sum_{p\in\mathcal{J}} u_{\alpha+p}\, u_{\beta+p} \binom{\alpha+p}{p} \binom{\beta+p}{p} p! \frac{\sqrt{\theta!}}{\sqrt{(\alpha+p)!\,(\beta+p)!}} T_\theta \\ &= \sum_{\theta\in\mathcal{J}} \sum_{0\le\beta\le\theta} \sum_{p\in\mathcal{J}} u_{\theta-\beta+p}\, u_{\beta+p} \binom{\theta-\beta+p}{p} \binom{\beta+p}{p} p! \frac{\sqrt{\theta!}}{\sqrt{(\theta-\beta+p)!\,(\beta+p)!}} T_\theta \\ &= \sum_{\theta\in\mathcal{J}} \Big(\sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\theta} C(\theta,\ \beta,\ p)\, u_{\theta-\beta+p}\, u_{\beta+p} \Big)\, T_\theta, \end{aligned}$$

which completes the proof. □

Theorem 2.3 is very useful because it provides an *analytical* formula for the WCE of polynomial nonlinear terms. It is a powerful tool in deriving the governing equations for the WCE coefficients in nonlinear equations. The Cameron-Martin theorem gives explicit formula for the mean and variance of the random solution in terms of its WCE coefficients. As a simple application of Theorem 2.3, we can also derive formulae for the higher order statistics of the random solution in terms of its Wiener Chaos coefficients.

Lemma 2.5 *Suppose* $u(x,\ s) = \sum_\alpha\ u_\alpha(x,\ s)T_\alpha$. *If* $E(u^3)$ *and* $E(u^4)$ *exist, then they have the following expressions:*

$$Eu^3(x,\ s) \ = \ \sum_{\alpha\in\mathcal{J}}\Big(\sum_{p\in\mathcal{J}}\sum_{0\le\beta\le\alpha} C(\alpha,\ \beta,\ p)\ u_{\alpha-\beta+p}u_{\beta+p}\Big)\ u_\alpha, \tag{2.40}$$

$$Eu^4(x,\ s) \ = \ \sum_{\alpha\in\mathcal{J}}\Big(\sum_{p\in\mathcal{J}}\sum_{0\le\beta\le\alpha} C(\alpha,\ \beta,\ p)\ u_{\alpha-\beta+p}\ u_{\beta+p}\Big)^2. \tag{2.41}$$

Proof From Theorem 2.3 we have

$$u^2 = \sum_{\theta\in\mathcal{J}}\Big(\sum_{p\in\mathcal{J}}\sum_{0\le\beta\le\theta} C(\theta,\ \beta,\ p)\ u_{\theta-\beta+p}\ v_{\beta+p}\Big)\ T_\theta(\xi).$$

Note that $E(u^3) = E(u\cdot u^2)$ and $E(u^4) = E(u^2\cdot u^2)$. From the orthogonality of $T_\alpha(\xi)$, we immediately get formula (2.40) and (2.41). □

2.4 Wiener Chaos Expansions: A Martingale Approach

While presenting the general theory of the Wiener chaos expansion in Section 2.3, we intentionally avoided the peculiar stochastic calculus language. In this section, we will study the WCE in a more analytical way and derive a few interesting results for the Wick polynomials. Those results are powerful tools in deriving the governing equations for the WCE coefficients.

Denote $\mathcal{F}_s^W$ as the σ-algebra generated by the Brownian motion $\{W(\tau);\ 0\le\tau\le s\}$. Obviously $\mathcal{F}_{s_1}^W \subset \mathcal{F}_{s_2}^W$ if $s_1 < s_2$. So the σ-algebra family $\mathcal{F}_s^W$ is increasing in time s. For fixed $t > 0$, denote the probability space of $\{W_s;\ \ 0\le s\le t\}$ as $(\Omega,\ \mathcal{F}_t^W,\ P)$, where Ω is the sample space and P is the probability measure. Define the set of all square integrable

functionals of the Brownian motion $\{W_s;\ 0 \le s \le t\}$ as

$$L^2(\Omega,\ \mathcal{F}_t^W,\ P) = \left\{ f(W_s;\quad 0 \le s \le t) \ \middle|\ E(f^2) < \infty \right\}. \tag{2.42}$$

The functional space $L^2(\Omega,\ \mathcal{F}_t^W,\ P)$ is a Hilbert space with respect to the Brownian motion measure P. Based on Cameron-Martin Theorem 2.2, the Wick polynomials $\{T_\alpha(\xi);\ \alpha \in \mathcal{J}\}$ is a complete orthonormal basis in $L^2(\Omega,\ \mathcal{F}_t^W,\ P)$.

Consider the general SPDE again

$$u_t(x,t) = L(u) + \sigma\, \dot{W}(t). \tag{2.43}$$

Denote $\{T_\alpha^t(\xi),\ \alpha \in \mathcal{J}\}$ as the orthonormal basis of $L^2(\Omega,\ \mathcal{F}_t^W,\ P)$ for a fixed time $t > 0$. Suppose $u(x,\ s)$ is a solution of (2.43) for $0 \le s \le t$. We assume that $u(x,\ s)$ is a well-defined random process and has second order moment for any fixed x and $s \leqslant t$, which is usually true from the existence results. So $u(x,\ s) \in L^2(\Omega,\ \mathcal{F}_s^W,\ P)$. Since $\mathcal{F}_s^W \subset \mathcal{F}_t^W$ for $s \le t$, it follows that $L^2(\Omega,\ \mathcal{F}_s^W,\ P) \subset L^2(\Omega,\ \mathcal{F}_t^W,\ P)$. So $u(x,s)$ also belongs to the space $L^2(\Omega,\ \mathcal{F}_t^W,\ P)$ and hence has the Wiener Chaos expansion:

$$u(x,\ s) = \sum_\alpha E\left[u(x,\ s) T_\alpha^t(\xi)\right]\ T_\alpha^t(\xi). \tag{2.44}$$

Note that the expansion (2.44) is valid for all $s \le t$ and $u(x,s)$ has a universal Wiener chaos expansion on the interval $s \in [0,t]$. When plugging the expansion (2.44) into the SPDE (2.43), we can derive the governing equations for the WCE coefficients.

Take expectation of (2.44) with respect to the σ-algebra $\mathcal{F}_s^W$. Since $u(x,\ s)$ only depends on $\{W_\tau;\quad 0 \le \tau \le s\}$, we have

$$u(x,\ s) = \sum_\alpha E\left[u(x,\ s) T_\alpha^t(\xi)\right]\ E\left[T_\alpha^t(\xi)\,|\,\mathcal{F}_s^W\right].$$

Define

$$T_\alpha(s) := E(T_\alpha^t \,\big|\, \mathcal{F}_s^W), \qquad 0 \le s \le t. \tag{2.45}$$

$T_\alpha(s)$ is the filtered version of $T_\alpha^t(\xi)$ against the σ-algebra family $\{\mathcal{F}_s^W\}_{0 \le s \le t}$, and $T_\alpha(s) \in$

$L^2(\Omega,\ \mathcal{F}_s^W,\ P)$. Moreover, the coefficients of the expansion

$$\begin{aligned} E\left[u(x,\ s)T_\alpha^t(\xi)\right] &= E\,E\left[u(x,\ s)T_\alpha^t(\xi)\ \middle|\mathcal{F}_s^W\right] \\ &= E\left\{u(x,\ s)\,E\left[T_\alpha^t(\xi)\ \middle|\mathcal{F}_s^W\right]\right\} \\ &= E\left[u(x,\ s)\,T_\alpha(s).\right] \end{aligned}$$

So we also have

$$u(x,\ s) = \sum_\alpha E\left[u(x,\ s)T_\alpha(s)\right]\ T_\alpha(s).$$

As a result, we have the following conclusion:

Theorem 2.4 *$\{T_\alpha(s);\ \alpha\in\mathcal{J}\}$ defined by (2.45) is a complete set of basis in the space $L^2(\Omega,\ \mathcal{F}_s^W, P)$ for any $s\leqslant t$.*

However, $T_\alpha(s)$ are no longer orthogonal to each other. It can be shown that $T_\alpha(s)$ defined by (2.45) is a martingale with respect to the σ-algebra family $\mathcal{F}_s^W$ and can be rewrite as an Ito integral. Please see Appendix A for the definition of martingale and its relations with Ito integrals. The following result belongs to Boris et al. [60, 70, 71].

Theorem 2.5 *$T_\alpha(s)$ defined by (2.45) satisfies the stochastic differential equation*

$$d\,T_\alpha(s) = \sum_{i=1}^{\infty} m_i(s)\sqrt{\alpha_i}\,T_{\alpha_{(i)}^-}(s)\,dW_s, \qquad 0\leq s\leq t, \tag{2.46}$$

or equivalently

$$T_\alpha(s) = I_{\{\alpha=0\}} + \int_0^s \sum_{i=1}^{\infty} m_i(\tau)\sqrt{\alpha_i}\,T_{\alpha_{(i)}^-}(\tau)\,dW_\tau, \tag{2.47}$$

where $I_{\{\alpha=0\}}=1$ if $\alpha=0$ and is zero otherwise; and $\alpha_{(i)}^-$ is a multi-index defined as

$$\alpha_{(i)}^-(j) = \begin{cases} \alpha_j, & j\neq i, \\ \alpha_i-1, & j=i. \end{cases}$$

Theorem 2.5 is very useful in deriving the governing equations of the WCE coefficients when applying the WCE to solve SPDEs. To prove Theorem 2.5, we need to prove two lemmas first.

Suppose $m_i(s), i = 1, 2, \ldots$ is a set of orthonormal basis in $L^2([0,t])$, where t is a fixed time instance. For any real sequence $z = (z_1, z_2, \ldots,)$ with finite number of nonzero components, define

$$X(z,s) = \int_0^s \sum_{k=1}^{\infty} z_k m_k(\tau) dW_\tau - \frac{1}{2}\int_0^s \left|\sum_{k=1}^{\infty} z_k m_k(\tau)\right|^2 d\tau. \tag{2.48}$$

Denote

$$P_s(z) = e^{X(z,s)}, \qquad 0 \leqslant s \leqslant t. \tag{2.49}$$

The following lemma says that $P_{s=t}(z)$ is the generating function of the Wick polynomials.

Lemma 2.6

$$T_\alpha(\xi) := \prod_{i=1}^{\infty} H_{\alpha_i}(\xi_i) = \frac{1}{\sqrt{\alpha!}} \frac{\partial^\alpha}{\partial z^\alpha} P_t(z)\Big|_{z=0}. \tag{2.50}$$

Proof From the orthogonality of $m_k(s)$ and the definition (2.28), it is easy to show that

$$X(t) = \sum_{k=1}^{\infty} z_k \int_0^t m_k(\tau)\, d\, W_\tau - \sum_{k=1}^{\infty} \frac{1}{2} z_k^2 = \sum_{k=1}^{\infty} \left(z_k \xi_k - \frac{1}{2} z_k^2 \right).$$

Hence

$$P_t(z) = \exp\left[\sum_{k=1}^{\infty} \left(z_k \xi_k - \frac{1}{2} z_k^2 \right)\right] = \prod_{k=1}^{\infty} \exp\left(-\frac{z_k^2}{2} + z_k \xi_k \right).$$

So $P_t(z)$ is the generating function of the multi-variable Hermite polynomials of $\xi_1, \xi_2, \ldots$, and Lemma 2.6 follows immediately from property (2.7). □

The $P_s(z)$ defined by (2.49) is a martingale and can be expressed as an Ito integral.

Lemma 2.7

$$dP_s(z) = \sum_{k=1}^{\infty} z_k m_k(s)\, P_s(z)\, dW(s), \qquad 0 \leq s \leq t.$$

So $P_s(z)$ is a martingale with respect to $\mathcal{F}_s^W$ and

$$E\left(P_t(z) \mid \mathcal{F}_s^W\right) = P_s(z) \quad \text{for} \quad s \leq t.$$

Proof Note that

$$dX(s) = \sum_{k=1}^{\infty} z_k m_k(s) dW_s - \frac{1}{2}\left|\sum_{k=1}^{\infty} z_k m_k(s)\right|^2 ds.$$

Then from Ito's formula (A.8) we have

$$\begin{aligned} dP_s(z) = d\left(e^{X(s)}\right) &= e^{X(s)}dX(s) + \frac{1}{2}e^{X(s)}\left[dX(s)\right]^2 \\ &= e^{X(s)}\left[\sum_{k=1}^{\infty} z_k m_k(s)\, dW_s - \frac{1}{2}\left|\sum_{k=1}^{\infty} z_k m_k(s)\right|^2 ds\right] + \frac{1}{2}e^{X(s)}\left|\sum_{k=1}^{\infty} z_k m_k(s)\right|^2 ds \\ &= \sum_{k=1}^{\infty} z_k m_k(s) P_s(z) dW_s. \end{aligned}$$

$P_s(z)$ is a martingale because it can be expressed as an Ito integral. □

Now we are ready to prove Theorem 2.5.

Proof From the definition of $T_\alpha(s)$ and Lemma 2.6, we have

$$\begin{aligned} T_\alpha(s) &= E\left(\frac{1}{\sqrt{\alpha!}}\frac{\partial^\alpha}{\partial z^\alpha}P_t(z)\Big|_{z=0} \;\Big|\; \mathcal{F}_s^W\right) \\ &= \frac{1}{\sqrt{\alpha!}}\frac{\partial^\alpha}{\partial z^\alpha}E\left(P_t(z) \;\Big|\; \mathcal{F}_s^W\right)\Big|_{z=0} \\ &= \frac{1}{\sqrt{\alpha!}}\frac{\partial^\alpha}{\partial z^\alpha}P_s(z)\Big|_{z=0}. \end{aligned}$$

The above formula gives the precise expression of $T_\alpha(s)$. From Lemma 2.7 we have

$$\begin{aligned} d\,T_\alpha(s) &= \frac{1}{\sqrt{\alpha!}}\frac{\partial^\alpha}{\partial z^\alpha}\frac{\partial P_s}{\partial s}(z)\Big|_{z=0} \\ &= \frac{1}{\sqrt{\alpha!}}\frac{\partial^\alpha}{\partial z^\alpha}\left(\sum_{i=1}^{\infty} z_i m_i(s) P_s(z) dW_s\right)\Bigg|_{z=0} \\ &= \sum_{i=1}^{\infty} m_i(s)\left(\frac{\alpha_i}{\sqrt{\alpha!}}\frac{\partial^{\alpha_i^-}}{\partial z^{\alpha_i^-}}P_s(z)dW_s\right)\Bigg|_{z=0} \\ &= \sum_{i=1}^{\infty} m_i(s)\sqrt{\alpha_i}\, T_{\alpha_i^-}(s)\, dW_s, \end{aligned}$$

which completes the proof. □

Chapter 3

WCE Methods for Stochastic Burgers Equations

Stochastically forced Burgers equations have been an active research subject in recent years. Unlike the unforced Burgers equations, which have purely dissipative solutions, the randomly forced Burgers equations have much richer structures. Based on the Cole-Hopf transformation [52], the unforced Burgers equation is mapped to a heat equation, while the forced Burgers equations is mapped to a kind of imaginary-time Schrödinger equation with the primitive function of the random forcing as the potential [6]. The randomly forced Burgers equations appear frequently in nonequilibrium statistical physics where strong nonlinear effects are present, such as vortex lines in superconductors, directed polymers, and kinetic roughening of growing surfaces. As a testing model, randomly forced Burgers equations are also studied intensively in the context of turbulence, either for their own sake as Burgers turbulence, or in the spirit of forced Navier-Stokes equations. For theoretical discussions about the stochastic Burgers equations and their applications, please see [6, 23, 24, 25, 56, 92, 93] and the references therein.

In this chapter, we will use the stochastic Burgers equations as a test model for the Wiener chaos method. We will consider the stochastic Burgers equations with additive forcing or multiplicative forcing. By comparing the numerical results of the WCE method with those by Monte Carlo simulations, we will demonstrate the efficiency and accuracy of the WCE method. Our numerical results also verify the existence of an invariant measure for the stochastic Burgers equation with additive forcing.

3.1 Stochastic Burgers Equations with Additive Random Forcing

In this section, we consider the following 1-D stochastic Burgers equation

$$\begin{cases} u_t + \frac{1}{2}(u^2)_x = \mu u_{xx} + \sigma(x)\dot{W}(t), \\ u(x,\,0) = u_0(x), \qquad u(0,t) = u(1,t) \end{cases} \qquad (t,x) \in (0,T] \times [0,1], \tag{3.1}$$

where $W(t)$ is a Brownian motion. For simplicity, we assume that the initial condition $u_0(x)$ is deterministic and limit our discussion to the periodic case. It is known (see e.g., [86]) that if $\|u_0\|_{L_2} < \infty$ and $\|\sigma\|_{L_2} < \infty$, then equation (3.1) has a unique solution with finite second order moments. Therefore, the solution of (3.1) admits a Wiener chaos expansion. Below, we will apply the WCE to the stochastic Burgers equation (3.1) and derive a deterministic PDE system for its WCE coefficients. The resulting nonlinear PDE systems can be solved numerically by the well-developed deterministic algorithm.

3.1.1 Equations for the WCE Coefficients

First we need to derive the governing equations for the WCE coefficients, so that we can compute them numerically. We choose an arbitrary orthonormal basis in $L^2([0,T])$, such as the trigonometric functions (2.29). Define the Gaussian random variables $\xi_k = \int_0^T m_k(t)dW(t)$ and the Wick polynomials $\{T_\alpha(\xi),\, \alpha \in \mathcal{J}\}$. In practice, we don't really need to compute the variables ξ_k and form $T_\alpha(\xi)$. As we will see later, only the L^2 base functions $m_k(t)$ are involved in the coefficient equations explicitly.

Since the stochastic solution $u(x,\,t)$ of (3.1) is well defined and has second order moment for any fixed x and t, from Theorem 2.2, $u(x,\,t)$ admits the Wiener chaos expansion

$$u(x,\,t) = \sum_{\alpha\in\mathcal{J}} u_\alpha(x,\,t)\, T_\alpha, \qquad u_\alpha(x,\,t) = E\big[u(x,\,t)T_\alpha\big]. \tag{3.2}$$

This expansion is valid for all $t \le T$. Rewriting the equation (3.1) in integral form, we have

$$u(x,\,t) = u_0(x) + \int_0^t \left[\mu u_{xx}(x,\,\tau) - \frac{1}{2}u_x^2(x,\,\tau)\right] d\tau + \sigma(x)W(t).$$

Multiplying both sides by T_α and taking expectations, we get

$$u_\alpha(x,\, t) = u_0(x) I_{\{\alpha=\mathbf{0}\}} + \int_0^t \Big[\mu \partial_{xx} u_\alpha - \frac{1}{2}\partial_x E\big(u^2 T_\alpha\big)\Big] d\tau + \sigma(x)\, E[W(t) T_\alpha], \tag{3.3}$$

where the indicator function $I_{\{\alpha=\mathbf{0}\}}$ is defined as

$$I_{\{\alpha=\mathbf{0}\}} = \begin{cases} 1 & \alpha = \mathbf{0}, \\ 0 & \text{otherwise.} \end{cases}$$

Set $v = u$ in Theorem 2.3, and use the orthogonality of T_α, we obtain that

$$E[u^2 T_\alpha] = \sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\alpha} C(\alpha,\, \beta,\, p)\; u_{\alpha-\beta+p}\, u_{\beta+p}. \tag{3.4}$$

Recalling the expansion (2.30) for $W(t)$, we have

$$E\big[W(t) T_\alpha\big] = \sum_{i=1}^{\infty} \int_0^t m_i(\tau) d\tau E\big(\xi_i\; T_\alpha\big).$$

Note that $H_1(x) = x$, so $\xi_i = T_{\{\alpha_j=\delta_{ij}\}}$. From the orthogonality of T_α we get

$$E\big[W(t) T_\alpha\big] = \sum_{i=1}^{\infty} I_{\{\alpha_j=\delta_{ij}\}} \int_0^t m_i(\tau) d\tau. \tag{3.5}$$

Plugging (3.4) and (3.5) into equation (3.3), we arrive at

$$\begin{aligned} u_\alpha(x,\, t) = u_0(x) I_{\{\alpha=\mathbf{0}\}} &+ \int_0^t \mu \partial_{xx} u_\alpha(x,\, \tau) d\tau + \sigma(x) \sum_{i=1}^{\infty} I_{\{\alpha_j=\delta_{ij}\}} \int_0^t m_i(\tau) d\tau \\ &- \frac{1}{2} \sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\alpha} C(\alpha,\, \beta,\, p) \int_0^t \partial_x (u_{\alpha-\beta+p} u_{\beta+p})(x,\, \tau)\, d\tau. \end{aligned}$$

Hence the WCE coefficients satisfy the PDE system

$$\begin{aligned} \frac{\partial}{\partial t} u_\alpha(x,\, t) + \frac{1}{2} \sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\alpha} C(\alpha,\, \beta,\, p) \frac{\partial}{\partial x} (u_{\alpha-\beta+p}\, u_{\beta+p})(x,\, t) \\ = \mu \frac{\partial^2}{\partial x^2} u_\alpha(x,\, t) + \sigma(x) \sum_{i=1}^{\infty} I_{\{\alpha_j=\delta_{ij}\}} m_i(t). \end{aligned} \tag{3.6}$$

The coefficient $C(\alpha, \beta, p)$ is defined in (2.39). Since the initial condition $u_0(x)$ is determin-

istic, the initial conditions for the WCE coefficients are

$$u_\alpha(x,\,0) = \begin{cases} u_0(x) & \alpha = \mathbf{0}, \\ 0 & \alpha \neq \mathbf{0}. \end{cases} \tag{3.7}$$

Clearly, all the WCE coefficients satisfy the periodic boundary condition $u_\alpha(0,t) = u_\alpha(1,t)$. The PDE system (3.6) is valid for $0 \le t \le T$. If we solve it to an intermediate time $t < T$, we can still recover the statistics of $u(x,\,t)$. We call (3.6) the *WCE propagator* of the corresponding stochastic Burgers equation.

The WCE propagator (3.6) shares the same nonlinearity of the original stochastic problem. The random bases $\{T_\alpha, \alpha \in \mathcal{J}\}$ do not appear explicitly in the WCE propagator. But their properties are implicitly built into the system. The system (3.6) is completely deterministic. The effect of the random forcing is captured by the L^2 functions $m_i(t)$. More specifically, the first WCE coefficient $u_{\mathbf{0}}(x,t)$, which is the mean of the stochastic solution, satisfies

$$(u_{\mathbf{0}})_t + \frac{1}{2}(u_{\mathbf{0}}^2)_x + \frac{1}{2}\sum_{|\alpha|\neq\mathbf{0}} (u_\alpha^2)_x = \mu(u_{\mathbf{0}})_{xx}. \tag{3.8}$$

The mean $u_{\mathbf{0}}$ is not driven by the random forcing directly, since the forcing has zero mean. However, because of the nonlinearity, the mean is forced by higher order WCE coefficients, which represent the randomness of the solution. For indices $\{\alpha,\ \alpha_j = \delta_{ij}\}$, $i = 1, 2, \ldots$, the corresponding WCE bases $T_\alpha = \xi_i$ are Gaussian random variables. Their WCE coefficients satisfy

$$(u_\alpha)_t + \sum_{p\in\mathcal{J}} (p_i+1)^{\frac{1}{2}} (u_p u_{p+\alpha})_x = \mu(u_\alpha)_{xx} + \sigma m_i(t), \qquad i = 1, 2, \ldots.$$

The coefficients $\{u_\alpha, |\alpha| = 1\}$ of the Gaussian part are forced by the random input directly. That is how the random effect enters the WCE propagator. For higher order coefficients u_α with $|\alpha| > 1$, they are not driven by the stochastic forcing directly. Instead they are forced by the Gaussian part $\{u_\alpha, |\alpha| = 1\}$. Note that the initial condition $u_\alpha(x,0) = 0$ for all $\alpha \neq \mathbf{0}$. If $\sigma = 0$, then the Gaussian part $\{u_\alpha, |\alpha| = 1\}$ will never grow from zero. As a result, all the higher order coefficients will also stay zero. Then the WCE propagator reduces to the unforced deterministic equation.

The hierarchy structure of the WCE propagator (3.6) has a profound impact on numerical computations. Since (3.6) is an infinite system, in practice we need to truncate it at

finite terms. Naturally, the efficiency of the WCE method highly depends on how many coefficients we need to keep in the computation. In the WCE propagator, higher order coefficients are driven by lower order coefficients, and at the bottom, the Gaussian coefficients are driven by the random forcing directly. So the lower order coefficients are usually dominant in magnitude, and the coefficients u_α decay fast as the order of T_α increases. In practice, even we keep only a few dozens of leading order coefficients, we can still capture quite a significant part of the information of the random solution. The truncated WCE system can be regarded as a closure of the moment equation for the nonlinear problem.

3.1.2 Truncating the WCE Propagator

The WCE propagator (3.6) is an infinite PDE system. In practice we need to truncate it at finite terms. Recall that WCE is a double infinite expansion in the number of Gaussian random variables ξ_i and the order of the Wick polynomials:

$$u(x,t) = \sum_{\alpha\in\mathcal{J}} u_\alpha(x,t) \prod_{i=1}^{\infty} H_{\alpha_i}(\xi_i).$$

Therefore we need to truncate the expansion in both directions. Suppose we want to keep K Gaussian random variables and maximum Nth order Wick polynomials in the WCE approximation. Define the truncated index set

$$\mathcal{J}_{K,N} = \{\alpha = (\alpha_1, \ldots, \alpha_K);\ |\alpha| \le N\}.$$

Then the truncated WCE can be denoted as

$$u_{K,N}(x,t) = \sum_{|\alpha|\le N} u_\alpha(x,t) \prod_{i=1}^{K} H_{\alpha_i}(\xi_i) = \sum_{\alpha\in\mathcal{J}_{K,N}} u_\alpha(x,t) T_\alpha. \tag{3.9}$$

It is shown in Theorem 5.1 that for a model problem the error of the WCE truncation (3.9) is

$$O\left[\frac{\left(\sigma T^{3/2}\right)^{N+1}}{\sqrt{(N+1)!}} + \sigma\left(\frac{T}{K}\right)^{3/2}\right], \tag{3.10}$$

where σ is the magnitude of the forcing, and T is the size of the time interval. We can choose the truncation parameter N and K accordingly based on the estimate (3.10) and the error tolerance.

The simple truncation (3.9) has altogether

$$\sum_{n=0}^{N} \binom{K+n-1}{n} = \frac{(K+N)!}{K!N!}$$

terms. The number of terms will increase very fast with both N and K. For a moderate truncation such as $K = 6$ and $N = 4$, the finite WCE propagator will have 210 coefficients. Though there is no problem for our numerical solver to handle such a large PDE system, it is neither efficient nor necessary to retain so many WCE coefficients in the computation. The simple truncation (3.9) is far from optimal, since a lot of the coefficients in the truncations are very small and contribute little to the final solution. We will prove in Theorem 5.2 that the WCE coefficients u_α, $\alpha = (\alpha_1, \alpha_2, \ldots)$ decay asymptotically as

$$u_\alpha = O\left[\frac{1}{\sqrt{\alpha!}} \prod_{k=2}^{\infty} \left(\frac{1}{k-1}\right)^{2\alpha_k}\right],$$

where the factor $\frac{1}{\sqrt{\alpha!}}$ reflects the decaying rate of the expansion in terms of the Wick polynomial order, and the factor $\prod_{k=2}^{\infty} \left(\frac{1}{k-1}\right)^{2\alpha_k}$ stems from the decaying rate of the Gaussian expansion (2.30) of the Brownian motion. With the same polynomial order $|\alpha|$, the Wick base $T_\alpha(\xi)$ depending on ξ_k with higher subscript k is less important than the one depending on ξ_k with lower subscript k. Instead of using Hermite polynomials with the same order for all ξ_k, $k \le K$, it is more advisable to use lower order polynomials for ξ_k with higher subscripts. To introduce such an adaptivity in the WCE truncation, we define an adaptive index

$$r = (r_1, r_2, \ldots, r_K) \qquad \text{with} \quad N = r_1 \ge r_2 \ge \ldots \ge r_K,$$

and define the sparse index truncation as

$$\mathcal{J}^r_{K,N} = \{(\alpha_1, \ldots, \alpha_K);\ |\alpha| \le N,\ \alpha_i \le r_i\}.$$

Instead of using the simple truncation (3.9), we truncate the WCE adaptively by the sparse truncation index:

$$u^r_{K,N}(x,t) = \sum_{\alpha \in \mathcal{J}^r_{K,N}} u_\alpha(x,t) T_\alpha. \tag{3.11}$$

W call this type of truncation a *sparse truncation*. The idea of the sparse truncation is

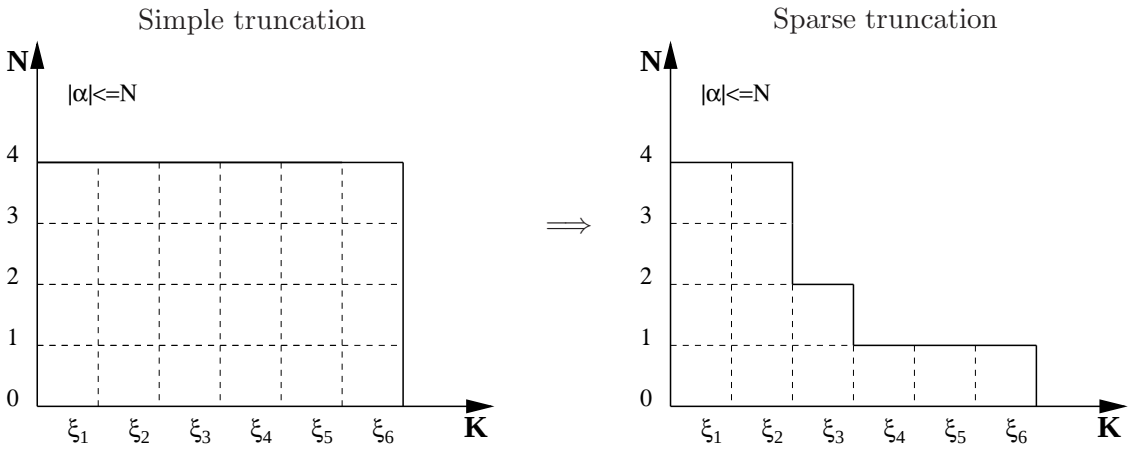

Figure 3.1: Illustration of the sparse truncation for $K = 6$ and $N = 4$. The left is the simple truncation. The right is the sparse truncation with the adaptive index as $r = (4, 4, 2, 1, 1, 1)$.

similar to the one used by Schwab [32] in the context of stochastic finite elements for random elliptic problems. We will show (Theorem 5.3) that the optimal adaptive index r is given by balance conditions

$$r_1 = N, \qquad \frac{1}{\sqrt{r_k!}} \left(\frac{1}{k-1} \right)^{2r_k} \simeq \frac{1}{\sqrt{N!}}.$$

For this optimal choice of r, the resulting sparse truncation (3.11) has similar asymptotic convergence rate as the simple truncation (3.9). For the truncation with $K = 6$ and $N = 4$, the optimal choice of the control index is $r = (4, 4, 2, 1, 1, 1)$. The resulting sparse truncation is illustrated in Figure 3.1. Using the sparse truncation technique, we can reduce the dimension of the truncated WCE system dramatically. For the hypothetical case $K = 6$ and $N = 4$, the number of coefficients will be reduced from 210 to 57. The larger the parameters N, K, the more effective the sparse truncation.

The idea of sparse truncation can be pursued even further in practice. Since the WCE decays both in the number of retained Gaussian random variables and the order of the Wick polynomials, we can include less Gaussian random variables in constructing higher order Wick polynomials. In the sparse truncation $K = 6$, $N = 4$, and $r = (4, 4, 2, 1, 1, 1)$, six Gaussian random variables are included in constructing the Wick polynomials. However, we don't need to use all the six Gaussian random variables in constructing higher order Wick polynomials. We can, for example, use only the first three ξ_i in the third order Wick polynomials, and the first two ξ_i in the fourth order Wick polynomials. This practice will reduce the number of the retained WCE bases further by eliminating unimportant crossing

products between different ξ_i. To implement this local adjustment, we only need to adjust the length of the adaptive index $r = (r_1, r_2, \ldots, r_K)$. For the current hypothetical case, the composite sparse truncation can be implemented as below.

Sparse Truncation of the WCE index α

1. Set $K = 6$ and $N = 4$,
2. Set the adaptive index as $r = (4, 4, 2, 1, 1, 1)$,
   (i) For $|\alpha| = 3$, set $r = (4, 4, 2)$,
   (ii) For $|\alpha| = 4$, set $r = (4, 4)$.

In summary, the idea of the sparse truncation is to put favor in the Gaussian variables ξ_i with smaller subscript i and trim the WCE bases by using lower order polynomials for ξ_i with bigger subscript i. These sparse truncation techniques can be implemented automatically and systematically when we generate the index α by induction. Using the sparse truncation techniques, we can bring down the total number of the WCE coefficients dramatically.

3.1.3 Numerical Solutions of the WCE Propagator

Next we discuss how to solve the WCE propagator (3.6) numerically. Since the random solution $u(x, t)$ is periodic in $[0, 1]$, so are all its WCE coefficients $u_\alpha(x, t)$. We approximate each WCE coefficient u_α by its truncated Fourier series

$$u_\alpha(x,\, t) = \sum_{k=-M/2+1}^{M/2} \hat{u}_\alpha(k,\, t) e^{2\,\pi\, i\, k\, x}.$$

Then the Fourier coefficients $\hat{u}_\alpha(k,\, t)$ are determined by the following ODE systems:

$$(\hat{u}_\alpha)_t + i\,\pi\,k \sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\alpha} C(\alpha,\, \beta,\, p)\widehat{u_{(\beta+p)}u_{(\alpha-\beta+p)}} = -\mu 4\pi^2 k^2 \hat{u}_\alpha + \hat{\sigma} \sum_{i=1}^{\infty} I_{\{\alpha_j=\delta_{ij}\}} m_i(t). \tag{3.12}$$

To compute the convolution $\widehat{u_\gamma u_\theta}$, we first transform the Fourier coefficient $\hat{u}_\gamma$, $\hat{u}_\theta$ to physical space and do the multiplication there, then transform the product $u_\gamma u_\theta$ back to the Fourier space. Such a method is called pseudo-spectral method. The diffusion term $-\mu 4\pi^2 k^2 \hat{u}_\alpha$

poses stiff stability conditions for solving the ODE system (3.12) numerically. To avoid extremely small time steps, we can integrate the diffusion term analytically by the linear propagator method [46]. Denote $\lambda = 4\pi^2 k^2 \mu,$, then the ODE system (3.12) can be rewritten as

$$\frac{d}{dt}\left[\hat{u}_\alpha(k,\, t)e^{\lambda t}\right] = e^{\lambda t}\left(\hat{\sigma}\sum_{i=1}^{\infty} I_{\{\alpha_j=\delta_{ij}\}} m_i(t) - i\,\pi\,k\sum_{p\in\mathcal{J}}\sum_{0\le\beta\le\alpha} C(\alpha,\, \beta,\, p)\widehat{u_{(\beta+p)}u_{(\alpha-\beta+p)}}\right) \tag{3.13}$$

To solve the ODE system (3.13) numerically, we use the fourth order Runge-Kutta method.

In the following numerical example, the initial condition of the stochastic Burgers equation (3.1) is chosen as

$$u_0(x) = \frac{1}{2}(e^{\cos 2\pi x} - 1.5)\sin 2\pi(x + 0.37). \tag{3.14}$$

We choose such a nontrivial initial condition because it has multi-modes in the Fourier space. We take the spatial part of the random forcing as

$$\sigma(x) = \frac{1}{2}\cos(4\pi x),$$

and set the viscosity $\mu = 0.005$. We first solve the problem (3.1) to $T = 0.8$ by the WCE method. Based on Theorem 5.1, the error induced by truncating the Gaussian expansion of the Brownian motion is $O\left[\sigma\left(\frac{T}{K}\right)^{3/2}\right]$, where σ is the magnitude of the random forcing, T is the length of the time interval, and K is the number of Gaussian random variables. For a given error tolerance ϵ, we can approximately set

$$K \simeq T\left(\frac{\sigma}{\epsilon}\right)^{2/3}.$$

In the current case, $T = 0.8$ and $|\sigma| \le 0.5$. Following the above guideline, we set $K = 8$ and project the Brownian motion $\{W(t),\, 0 \le t \le T\}$ onto eight cosine bases. The error for such a truncation is approximately $O(10^{-2})$. For the Wick polynomial order, we set $N = 1, 2, 3$, that is, we truncate the WCE propagator to the first order, second order, and third order, respectively. The total numbers of coefficients in the simple WCE truncation is 9, 45, and 165. However, using the sparse truncation strategy, the WCE coefficients are reduced to 9, 19, and 35, respectively. The centered statistical moments computed from the different order

of WCE approximations are compared in Figure 3.2. As we mentioned before, the first order WCE approximation is a Gaussian approximation. With only nine coefficients, the mean of the solution is captured quite well. That is because the Gaussian modes are dominant in magnitude and hence provide the leading order correction to the mean equation (3.8). Since the stochastic solution of a nonlinear equation is by no means a Gaussian process, the first order WCE approximation is inadequate for higher order moments. For example, the third order moment of the Gaussian approximation is zero, which is obviously incorrect. Including second order coefficients improves the numerical solutions significantly, especially for the variance and third order moment. The numerical results of the third order WCE approximation are very accurate and almost identical to the MC solutions with 100,000 realizations.

Figure 3.3 is the L_2 norm of the WCE coefficients with third order truncation. We order the multi-indices in the following way: if $|\alpha| \neq |\beta|$, then the index with smaller summation is listed first. Otherwise, the first component is compared. The index α will be listed ahead of β if $\alpha_1 > \beta_1$. If $\alpha_1 = \beta_1$, then the second component is compared, and so on and so forth. In Figure 3.3, the first coefficient is the mean, and coefficients 2 to 9 correspond to the Gaussian part of the solution. The Gaussian coefficients decay quickly and the coefficients of ξ_8 is already very small. This confirms the truncation choice $K = 8$ as predicted by Theorem 5.1. The coefficients of $\xi_k,\ k = 5, 6, 7, 8$ are relatively small comparing with those of $\xi_i,\ i = 1, 2, 3, 4$, which justifies the argument for the sparse truncation. The coefficients 10 to 19 correspond to the second order Wick polynomials. They are quite important in magnitude and provide significant correction to the first order WCE approximation. The coefficients 20 to 35 correspond to the third order Wick polynomials. The numerical results show that high order WCE coefficients are important in resolving strongly nonlinear problems, such as Burgers equations. Figure 3.3 shows that the WCE coefficients decay quickly as the order of Wick polynomials increases. However, the rate of decay will depend on the variability σ and the length T of the time interval, as indicated by our error analysis in Section 5. Larger variability and longer time interval will lead to slower decay in the WCE coefficients.

For comparison purpose, we also solve the stochastic Burgers equation (3.1) by MC simulation. We first discretize the spatial derivatives by the pseudo-spectral method. The

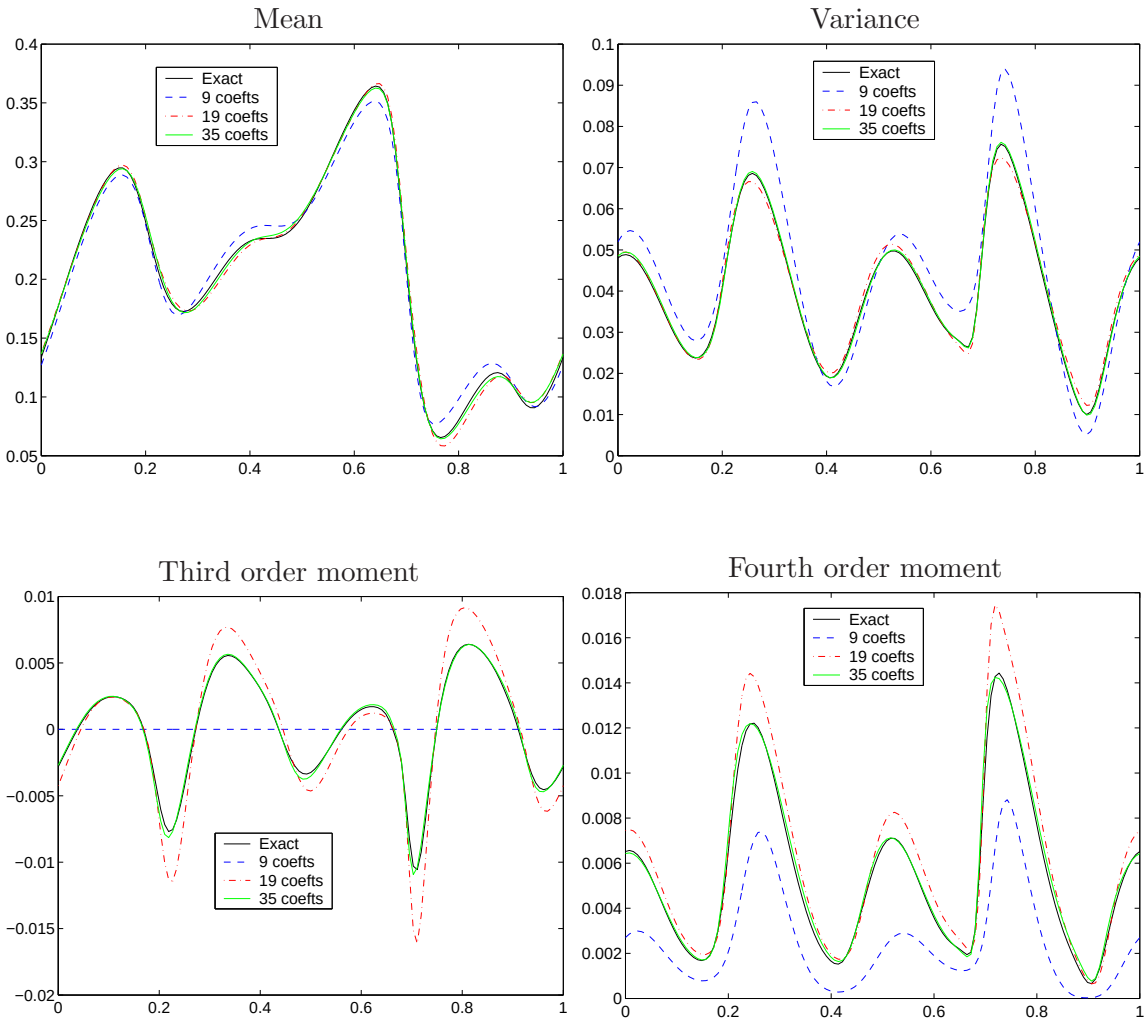

Figure 3.2: Statistical moments computed by WCE method with 9, 19, and 35 coefficients, respectively, corresponding to first order, second order, and third order WCE approximations. The exact solution is computed by MC simulation with 100,000 realizations. The first order WCE approximation is a piecewise Gaussian approximation. It is obviously not accurate, especially for high moments. Including second order correction improves the results significantly. The results by third order truncation are almost identical to the exact solutions.

resulting equations for the Fourier coefficients are

$$\hat{u}_t + i\pi k\widehat{uu} = -\mu 4\pi^2 k^2 \hat{u} + \hat{\sigma}\dot{W}(t).$$

After integrating the diffusion term analytically, we get a simple ODE system

$$d\big(\hat{u}e^{\lambda t}\big) = -i\pi k e^{\lambda t}\widehat{uu}\,dt + e^{\lambda t}\hat{\sigma}\,dW(t), \qquad \lambda = 4\pi^2 k^2 \mu. \tag{3.15}$$

In MC simulation we need to solve the stochastic Burgers equation (3.1) realization by real-

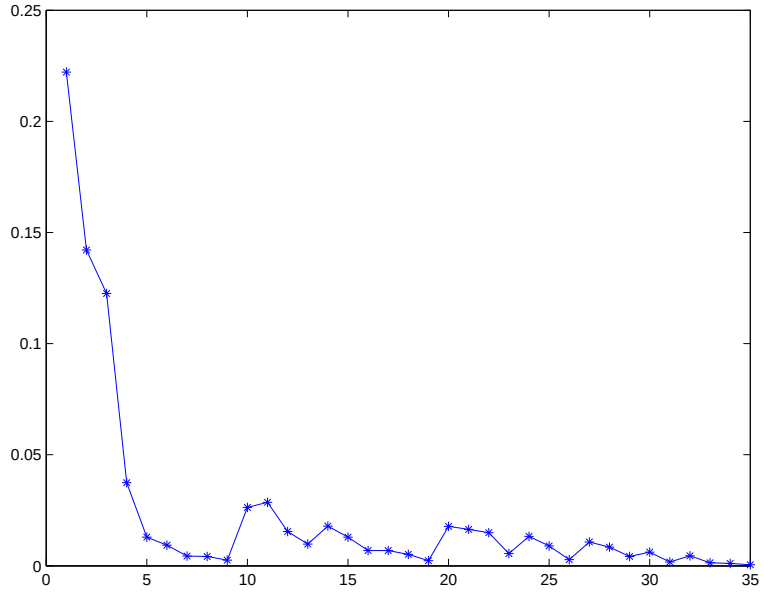

Figure 3.3: L^2 norm of the WCE coefficients at third order truncation. The coefficients decay very fast. The first 9 coefficients correspond to the Gaussian part of the solution; coefficients 10–19 correspond to the second order Wick polynomials; coefficients 20–35 correspond to the third order Wick polynomials. The coefficients of the Gaussian part are dominant in magnitude. However, higher order coefficients are also important.

ization, which requires discretizing the stochastic ordinary differential equations (SDE). Due to the peculiar differential rules of the Brownian motion calculus, the numerical schemes for discretizing SDE are usually much more complicated than their deterministic counterparts. In Appendix B we will provide a review of the numerical schemes for discretizing SDEs in MC simulation. In this thesis, we choose to use the weakly second order Runge-Kutta method (B.4) in all the MC simulations (unless specified otherwise). For the additive stochastic Burgers equation, the weakly second order Runge-Kutta scheme is reduced to the simple modified Euler scheme.

To compute realizations of the stochastic solution, we need to sample the Brownian motion path by generating Gaussian random variables $W(t_{n+1}) - W(t_n) = \sqrt{\Delta t} N(0,1)$ repeatedly. We adopt the standard Gaussian random number generator *gasdev* from *Numerical Recipes in C* [87], which is based on an analytical transformation of a uniform random variable in $[0,1]$. We use *ran2*, also from [87], as the source of uniform generator for *gasdev*. The pseudo random number generator *ran2* is believed to be able to generate perfect random numbers within the limit of floating-point precision. It also has a very long period ($> 2 \times 10^{18}$). For the problems under consideration, there is no need to worry about

the exhaustion of the pseudo random numbers in the period.

To compute the statistic moments of the random solution, we need to average many realizations in the MC simulation. For example, the mean $E(u)$ of the random solution is estimated by the MC ensemble average

$$u_M^{MC}(x,t) = \frac{1}{M}\sum_{k=1}^{M} u(x,t,\omega_k),$$

where $u(x,t,\omega_k)$ is a realization computed by the MC simulation, and M is the total number of realizations. To assess the convergence behavior of the MC simulation, we compute the relative errors of the statistic moments for each $M \in [10,\ 10,000]$. For example, the relative error of the mean is computed as

$$\varepsilon_u(M) = \frac{\|u_M^{MC} - E(u)\|_2}{\|E(u)\|_2},$$

where $E(u)$ is the benchmark mean computed by MC simulation with 100,000 realizations, and $\|\cdot\|_2$ is the L^2 norm in x variable. The relative error of the other moments are computed similarly. Since the MC ensemble average u_M^{MC} is a random variable itself, so is the relative error $\varepsilon_u(M)$. Based on the central limit theorem [30], for large M

$$\varepsilon_u(M) \sim M^{-1/2} N(0,1),$$

where $N(0,1)$ is the standard Gaussian distribution. We repeat the MC simulation 20 times to compute the root mean square (RMS) error $E[\varepsilon_u^2(M)]^{1/2}$ for each realization number M. The RMS errors of the MC estimations for other statistical moments can be computed similarly. Figure 3.4 is the RMS error (in log-log scale) of the MC estimations with different realization numbers. The RMS errors decay slowly as the number of the realizations increases. For all the statistic moments computed, the RMS errors are reduced by only one order when the realization number increases by two orders (from 100 to 10,000). So the convergence rate of the MC estimations is proportional to $O\left(\frac{1}{\sqrt{M}}\right)$ for all the statistical moments, which is consistent with the Central Limit Theorem. With 10,000 realizations, the RMS errors of the MC ensemble averages are 0.8%, 1.4%, 6.4%, and 3.2% for the mean, variance, and third and fourth order moments, respectively. The errors are slightly different for difference statistical moments, which is caused by the different proportional constants in

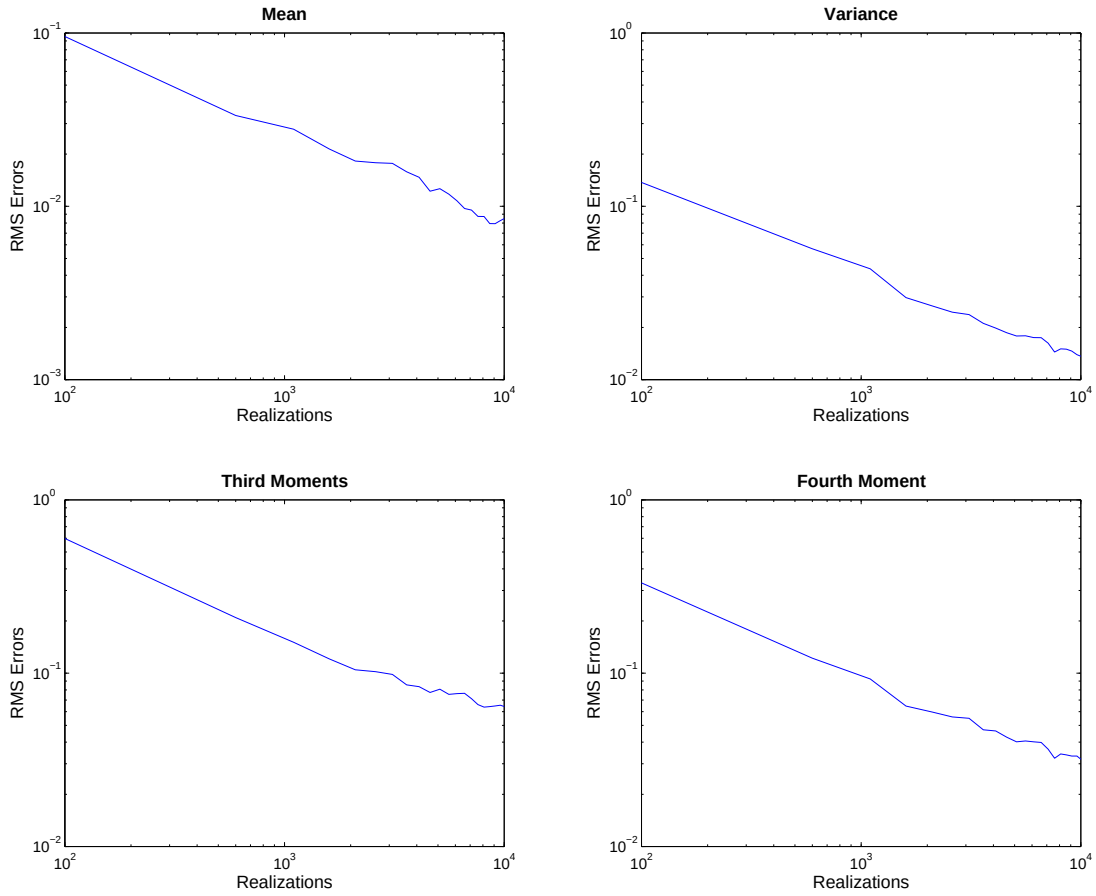

Figure 3.4: Log-log convergence of the MC simulation with different realizations. The relative RMS errors are computed by repeating the simulations 20 times for each realization number. The "exact" solutions are computed by MC with 100,000 realizations. With 10,000 realizations, the RMS relative errors of the MC simulation are 0.8%, 1.4%, 6.4%, and 3.2% for the mean, variance, and third and fourth order moments, respectively.

the convergence rate $O\left(\frac{1}{\sqrt{M}}\right)$. From the convergence study, we find that 10,000 realizations are required for the MC simulation to reach an accuracy of $O(10^{-2})$.

Figure 3.5 is the comparison between the WCE method with 35 coefficients and the MC simulation with 10,000 realizations. The numerical results by the WCE method have very similar accuracy as those by MC simulation. However, the CPU time is only 2 seconds for the WCE method, as opposed to 315 seconds for MC simulations. Therefore, the WCE method is more efficient than MC simulation in this case. In this thesis, all the numerical experiments are conducted on a PC with a 2.60 GHz CPU.

In the MC simulations, we also tried various kind of acceleration techniques (see Appendix B). However, most of the variance reduction techniques are not readily applicable

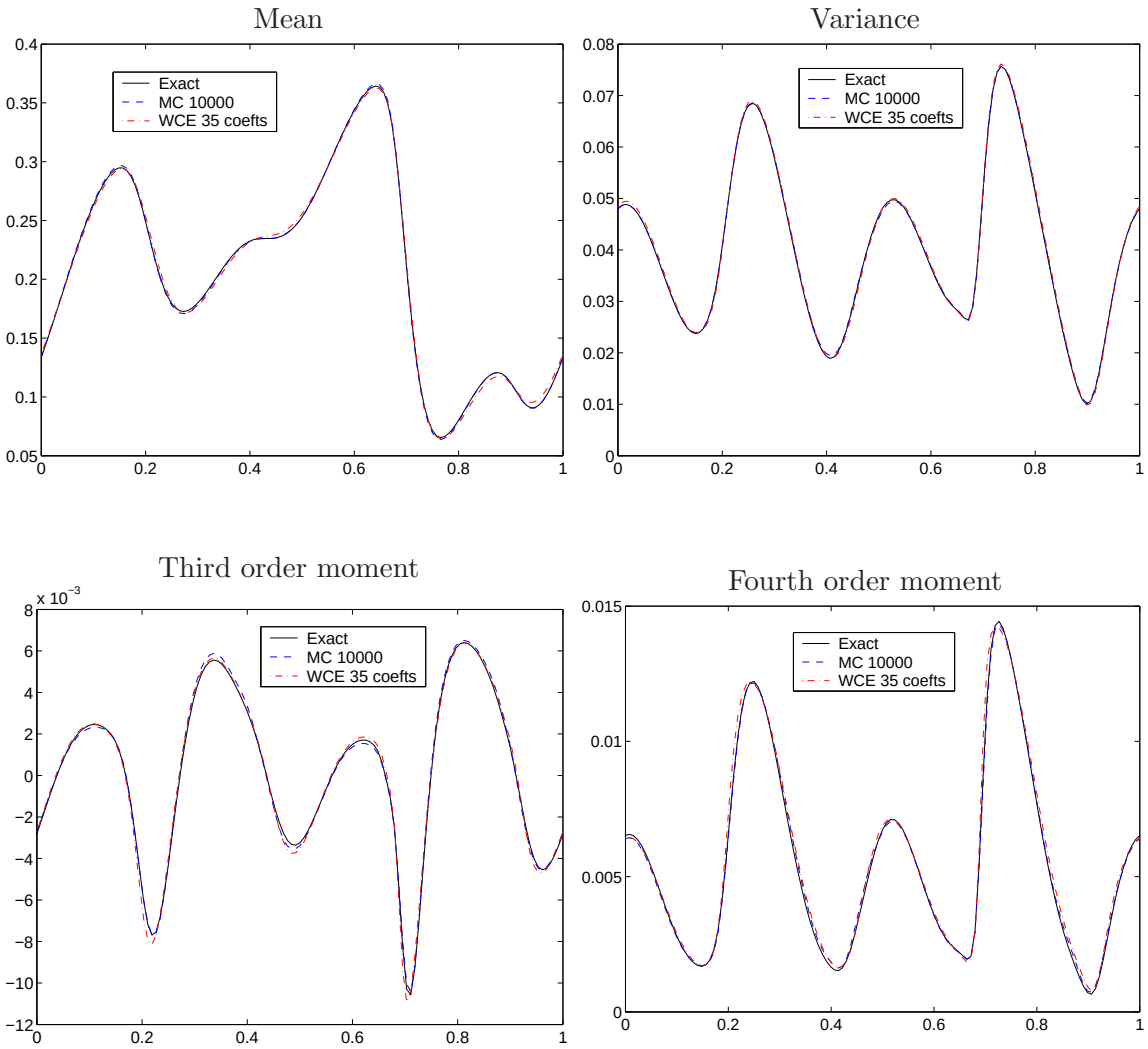

Figure 3.5: Convergence comparison between WCE method and MC simulation. The WCE method with 35 coefficients has the same accuracy as MC simulation with 10,000 realizations. However, the CPU time is only 2 seconds for the WCE method, as opposed to 315 seconds for the MC simulation.

here since they require prior knowledge of the distribution of the stochastic solutions. We did implement antithetic variables in our MC simulations. The motivation of using antithetic variables is to preserve the symmetry of the Brownian motion and reduce the noise in the samples of the forcing. We found that the antithetic variable can only improve the convergence of the MC simulation slightly, usually no more than a couple of times. This is because the distribution of the random solution is quite different from that of the Brownian motion forcing, and in particular, it is not symmetric. The antithetic variable based on the Brownian motion forcing may not be able to reduce the variance of the random solution effectively. With the improvement provided by the acceleration techniques, the MC simulation is still far less efficient than the WCE method.

3.1.4 Convergence Verification of the WCE Method

To demonstrate the convergence of the WCE method, we compare its numerical solutions with those of MC simulation. However, it is difficult to obtain accurate solutions by MC simulation, due to its slow and nonmonotone convergence. In this subsection, we will test both WCE method and MC simulation on a model equation, where a semi-analytical solution is available. The semi-analytical solution can be computed accurately by deterministic numerical algorithms, without the need of generating random numbers. Using it as the benchmark, we will compare the performance of both WCE method and MC simulation quantitatively.

In stochastic Burgers equation (3.1), suppose the spatial part $\sigma(x)$ of the random forcing is a constant. In this particular case the solution of the stochastic Burgers equation can be obtained by a nonlinear stochastic perturbation of its deterministic counterpart. More specifically, the following theorem holds:

Theorem 3.1 *If σ is a constant, then the solution of* (3.1) *is given by*

$$u(x,\, t) = v\left(x - \sigma \int_0^t W(s)ds,\; t\right) + \sigma W(t), \tag{3.16}$$

where $v(x,\, t)$ is a solution of the unforced deterministic Burgers equation

$$\begin{cases} v_t + \frac{1}{2}(v^2)_x = \mu v_{xx}, \\ \\ v(x,0) = u_0(x),\; v(0,t) = v(1,t). \end{cases} \tag{3.17}$$

Proof Let

$$X(x,t) = x - \sigma \int_0^t W(s)ds,$$

then $u(x,t) = v(X,t) + \sigma W(t)$ and

$$u_x(x,t) = v_x(X,t), \qquad u_{xx}(x,t) = v_{xx}(X,t).$$

Since v satisfies equation (3.17), we have

$$v_t(X,t) = \mu v_{xx}(X,t) - v(X,t)v_x(X,t).$$

It follows that

$$\begin{aligned} u_t(x,t) &= v_x(X,t)\, X_t + v_t(X,t) + \sigma\, \dot{W}(t) \\ &= -\sigma W(t)\, v_x(X,t) + \mu v_{xx}(X,t) - v(X,t)\, v_x(X,t) + \sigma\, \dot{W}(t) \\ &= \mu v_{xx}(X,t) - \big[v(X,t) + \sigma W(t)\big]\, v_x(X,t) + \sigma\, \dot{W}(t) \\ &= \mu u_{xx}(x,t) - u(x,t)\, u_x(x,t) + \sigma\, \dot{W}(t). \end{aligned}$$

So formula (3.16) satisfies the differential equation of (3.1). It is easy to check that it also satisfies the initial and boundary condition. So (3.16) is indeed the solution of the stochastic Burgers equation (3.1). □

Theorem 3.1 states that the random solution of the stochastic Burgers equations is the unforced deterministic solution compounded with a random perturbation along the flow characteristics. Formula (3.16) provides a convenient way to compute the statistical moments of u numerically. Instead of solving the stochastic equation (3.1) by MC simulations with many realizations, we can first solve the deterministic equation (3.17) and then compute the statistical moments by numerical integrations.

In formula (3.16) denote

$$Y(\omega,\, t) = \sigma W(t), \qquad Z(\omega,\, t) = \sigma \int_0^t W(s) ds.$$

Obviously, both Y and Z are Gaussian random variables with zero mean. By simple stochastic calculus, we can show that

$$E(Y^2) = \sigma^2 t, \qquad E(Z^2) = \sigma^2 t^3/3, \qquad E(Y \cdot Z) = \sigma^2 t^2/2.$$

Since the joint distribution of Gaussian random variables is determined uniquely by their means and covariance, we have

$$(Y,\ Z) \sim N(\mathbf{0},\ \Sigma), \qquad \text{with} \qquad \Sigma = \sigma^2 \begin{bmatrix} t & t^2/2 \\ t^2/2 & t^3/3 \end{bmatrix},$$

where $N(\mathbf{0},\ \Sigma)$ denotes the Gaussian distribution with mean zero and covariance Σ. Denote

$\mathbf{y} = (y, z)^T$, then the joint probability density function of random variables $(Y,\, Z)$ is

$$\begin{aligned}\rho(y,\, z) &= \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}\mathbf{y}'\Sigma^{-1}\mathbf{y}\right] \\ &= \frac{\sqrt{3}}{\pi\sigma^2 t^2} \exp\left(-\frac{2y^2}{\sigma^2 t} + \frac{6yz}{\sigma^2 t^2} - \frac{6z^2}{\sigma^2 t^3}\right).\end{aligned}$$

From (3.16) it is easy to see that

$$E\left[u(x,\, t)\right]^n = \int_{\mathbf{R}^2} \left[v(x - z,\, t) + y\right]^n \rho(y,\, z)\, dydz. \tag{3.18}$$

Integrating the variable y analytically, we can further get

$$Eu(x,\, t) = \int_{-\infty}^{+\infty} v(x - z,\, t)\rho(z)dz, \tag{3.19}$$

$$Eu^2(x,\, t) = \int_{-\infty}^{+\infty} \left[v^2(x - z,\, t) + v(x - z,\, t)\frac{3z}{t}\right] \rho(z)dz + \sigma^2 t, \tag{3.20}$$

$$Eu^3(x,\, t) = \int_{-\infty}^{+\infty} \left[v^3 + v^2\left(\frac{9z}{2t}\right) + v\left(\frac{3\sigma^2 t}{4} + \frac{27z^2}{4t^2}\right)\right] \rho(z)dz, \tag{3.21}$$

$$\begin{aligned}Eu^4(x,\, t) = \int_{-\infty}^{+\infty} \Bigg[&v^4 + v^3\left(\frac{6z}{t}\right) + v^2\left(\frac{3\sigma^2 t}{2} + \frac{27z^2}{2t^2}\right) \\ &+ v\left(\frac{9\sigma^2 z}{2} + \frac{27z^3}{2t^3}\right)\Bigg]\rho(z)dz + 3\sigma^4 t^2,\end{aligned} \tag{3.22}$$

where $\rho(z)$ is the marginal probability density function of random variable Z

$$\rho(z) = \frac{1}{\sqrt{2\pi}\gamma} \exp\left(-\frac{z^2}{2\gamma^2}\right), \qquad \gamma^2 = \sigma^2 t^3/3. \tag{3.23}$$

To compute the statistical moments of u, we first solve the deterministic Burgers equation (3.17) and obtain the solution v. Then we compute the integral (3.19)–(3.22) by numerical quadratures. Since all the computations are deterministic, we can quantify and hence control the numerical error at each step. By carefully choosing the deterministic algorithms, we can compute the statistical moments of u very accurately.

To solve the unforced deterministic equation (3.17), we use the pseudo-spectral method in space and the fourth order Runge-Kutta method in time. To compute the integrals

	CPU seconds	Mean	Variance	Third order	Fourth order
WCE	3	0.08%	1.3%	4.3%	5.2%
MC	319	0.18%	2.1%	3.8%	4.1%

Table 3.1: Relative L^2 error of the WCE method with 40 coefficients and MC simulation with 10,000 realizations

(3.19)–(3.22) numerically, we first truncate the infinite integral domain by $[-M, M]$, where M is an integer about ten times large as the variance of $\rho(z)$. Since the Gaussian kernel $\rho(z)$ decays exponentially, such a truncation is very accurate. For the same reason, the integrands in (3.19)–(3.22) can be treated as compact supported and hence periodic in $[-M, M]$. Then we use the trapezoidal rule to compute the integrals in the truncated domain. It is well known [57] that the trapezoidal rule has spectral accuracy for periodic and infinitely smooth integrands. Since both the solution v of equation (3.17) and the Gaussian kernel $\rho(z)$ are infinitely smooth, the trapezoidal integrator has spectral accuracy for the integrals (3.19)–(3.22). As we can expect, such computed statistical moments will be very accurate.

In the following numerical test, we consider the Burgers equation (3.1) with the same setup as in Section 3.1.3. The only difference is that the spatial part of the random forcing is chosen as a constant $\sigma = 0.1$. We first solve the problem to $T = 0.8$ by the semi-analytical approach and obtain the benchmark statistics of u. Then we solve the same problem by the WCE method. We project the Brownian motion in $[0, T]$ onto the first eight cosine bases, and truncate the WCE expansion at fourth order Wick polynomials. The resulting WCE propagator has 40 coefficients. According to Theorem 5.1, the error for such a truncation is approximately $O(10^{-2})$. For comparison, we also solve the problem by MC simulation with 10,000 realizations. In all the three methods, we use the same spatial mesh $N = 128$ and time step $\Delta t = 0.001$. We compare the numerical solutions of the WCE method and MC method in Table 3.1. In the table, the relative error for the WCE Mean is defined as $\|E(u^{WCE}) - E(u)\|_2 / \|E(u)\|_2$, where u is the semi-analytical solution, and $\|\cdot\|_2$ is the L^2 norm in x variables. The errors for the other moments are defined similarly. To reach a comparable accuracy, the WCE method takes only 3 seconds, while MC simulation takes 319 seconds. So the WCE method is much faster than MC simulation for that 1-D stochastic Burgers equation. Furthermore, we indeed observed an accuracy of $O(10^{-2})$ for the WCE method, as predicted by Theorem 5.3 for the sparse WCE truncation.

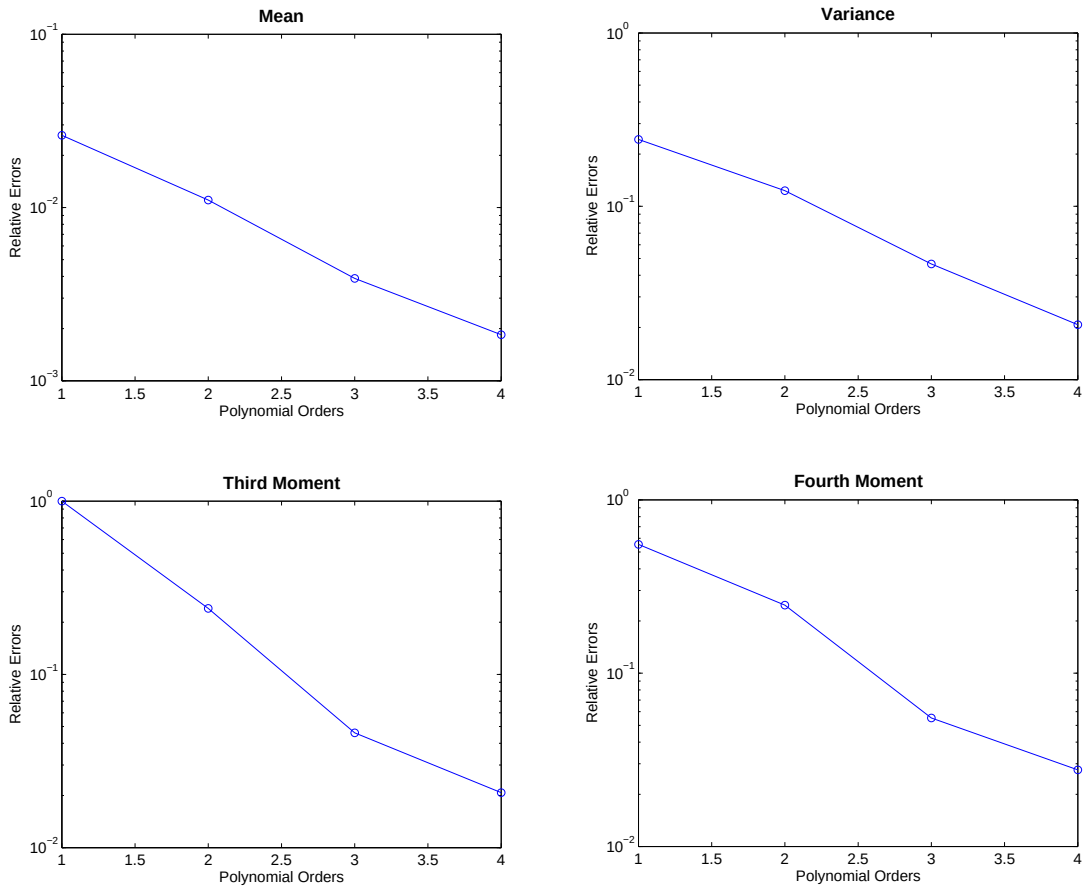

Figure 3.6: Log plot of the relative error of the WCE method with different polynomial orders. The errors decrease exponentially when the Wick polynomial order increases.

To study the convergence of the WCE truncations, we solve (3.1) again by the WCE method with first order, second order, third order, and fourth order Wick polynomials sequentially. In all the cases, eight Gaussian random variables are retained. Figure 3.6 is the log plot of the relative errors of the statistical moments computed by the WCE method with different polynomial orders. The errors decay linearly in log scale when the polynomial order increases. So the WCE method converges exponentially in terms of the Wick polynomial order, which is consistent with Theorem 5.1.

For comparison, we also compute the RMS error of the MC simulation with 500, 1,000, 2,000, 4,000, 8,000 realizations, respectively. We repeat the simulations 20 times to estimate the RMS errors for each realization number. Figure 3.7 is the log-log plot of the RMS errors. The numerical results show that the MC simulation converges quite slowly and the errors decay linearly in the log-log scale. It is not a surprise that MC simulations require 10,000 realizations to achieve the similar accuracy as the WCE method with only 40 coefficients.

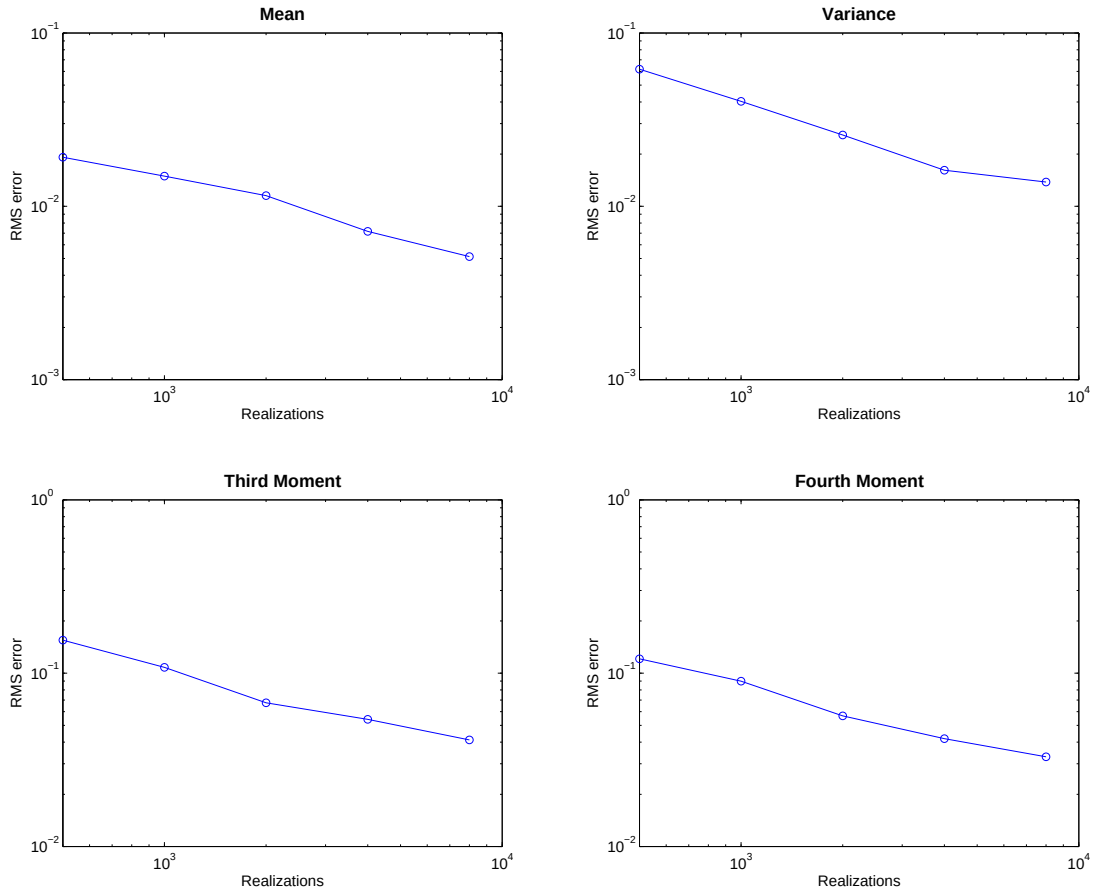

Figure 3.7: Log-log plot of the relative RMS errors of the MC simulation with 500, 1,000, 2,000, 4,000, 8,000 realizations, respectively. Twenty simulations are performed to estimate the RMS errors for each realization number.

It is worth pointing out that the performance of the WCE method is not sensitive to the choice of the L^2 basis function $m_k(t)$, as long as they are orthonormal. In addition to the cosine base (2.29), we also tried the Haar wavelets [17] and the sine basis

$$m_k(t) = \sqrt{\frac{2}{T}} \sin \frac{(k-1/2)\pi t}{T}, \quad 0 \le t \le T, \qquad k = 1, 2, \ldots \tag{3.24}$$

in our numerical experiment. The numerical results are almost the same for the different choice of $m_k(t)$.

3.2 Stochastic Burgers Equations with Multiplicative Random Forcing

Stochastic equations with multiplicative random forcing are usually much more difficult to handle than those with additive forcing. A random forcing is classified as multiplicative if it depends on the solution itself. Due to the peculiar differential rules of the stochastic calculus, SPDEs with multiplicative forcing require special care. In this section we consider the following stochastic Burgers equation:

$$\begin{cases} du + uu_x dt = \mu u_{xx} dt - \sigma u_x \circ dW(t), \\ \\ u(x,0) = u_0(x), \quad u(0,t) = u(1,t), \end{cases} \tag{3.25}$$

where the stochastic integral $\sigma u_x \circ dW(t)$ is defined in the Stratonovich's sense (A.4). The random forcing in (3.25) involves the unknown function u_x, so the SPDE (3.25) has a multiplicative type random forcing. We can formally rewrite (3.25) as

$$u_t + \left(u + \sigma \dot{W}\right) \circ u_x = \mu u_{xx}, \tag{3.26}$$

which reveals that the equation has a stochastic convection velocity. We propose the stochastic forcing in Stratonovich's sense because it always gives rise to a well-posed equation. In the next section, we will give an example to show that if we propose the random forcing in Ito's sense, we may have an ill-posed problem. For the definitions and differential rules of Stratonovich integrals, please see Appendix A, or the reference [77].

Since the Stratonovich integral is not a martingale, we prefer to rewrite it in Ito's form. Denote

$$b(x, \partial_x) = -\sigma\, \partial_x,$$

then the random forcing in (3.25) can be formally rewritten as

$$-\sigma u_x \circ dW(t) = b\,(x,\, \partial_x)\, u \circ dW(t).$$

From the relations between Ito integrals and Stratonovich integrals, we can *formally* get

$$\begin{aligned} b\,(x,\,\partial_x)\,u \circ dW(t) &= b\,(x,\,\partial_x)\;u\;\cdot dW(t) + \frac{1}{2}\,b\,(x,\,\partial_x)\;b\,(x,\,\partial_x)\;u dt \\ &= -\sigma\,u_x\;\cdot dW(t) + \frac{1}{2}\,\sigma(\sigma\,u_x)_x\,dt. \end{aligned}$$

Then the equation (3.26) can be rewritten in Ito's form:

$$u_t + \left[u + \sigma\dot{W}(t)\right]\cdot u_x = \mu u_{xx} + \frac{1}{2}\,\sigma(\sigma u_x)_x. \tag{3.27}$$

The equivalence between (3.26) and (3.27) can be established rigorously based on the definition of Stratonovich integrals. From the form (3.27), it becomes clear that the random perturbation is along the characteristic velocity field. This consideration is motivated by Kraichnan's turbulence model [54] for transport equations.

3.2.1 Derivation of the WCE Propagator

In this section we derive the WCE propagator for the stochastic Burgers equation (3.27). Suppose the solution of (3.27) has a Wiener chaos expansion for $t \in [0, T]$

$$u(x,t) = \sum_{\alpha\in\mathcal{J}} u_\alpha(x,t)T_\alpha(\xi)$$

with

$$u_\alpha(x,t) = E[u(x,t)T_\alpha(\xi)] = E[u(x,t)T_\alpha(t)],$$

where

$$T_\alpha(t) = E\big(T_\alpha(\xi)\;\big|\;\mathcal{F}_t^W\big)$$

is the Wick polynomials filtered by the σ-algebra $\mathcal{F}_t^W$. Recall that $T_\alpha(t)$ is a martingale and satisfies the differential equation (Theorem 2.5):

$$dT_a(t) = \left(\sum_k m_k(t)\sqrt{\alpha_k}\;T_{\alpha_{(k)}^-}(t)\right)\;d\,W(t).$$

Rewriting the equation (3.27) in differential form

$$du = \left[\mu\, u_{xx} + \frac{1}{2}\sigma\,(\sigma\, u_x)_x - uu_x\right]\, dt - \sigma\, u_x\, dW(t), \tag{3.28}$$

we have

$$\begin{aligned} d[uT_\alpha(t)] &= T_\alpha(t)\, d\, u + u\, d\; T_\alpha(t) + du \cdot d\, T_\alpha(t) \\ &= T_\alpha(t)\left[\mu u_{xx} + \frac{1}{2}\sigma\,(\sigma\, u_x)_x - u\, u_x\right] dt - T_\alpha(t)\,\sigma\, u_x\, dW(t) \\ &\quad + u\left(\sum_k m_k(t)\sqrt{\alpha_k}\; T_{\alpha^-_{(k)}}(t)\right)\; d\, W(t) \\ &\quad - \sigma\, u_x\left(\sum_k m_k(t)\sqrt{\alpha_k}\; T_{\alpha^-_{(k)}}(t)\right) dt. \end{aligned}$$

Taking expectation of the above formula, the terms involving Ito integrals will disappear since they are mean zero. Plugging in the expression $u_\alpha = E[u\, T_\alpha(t)]$ we have

$$\begin{aligned} du_\alpha &= \left\{\mu(u_\alpha)_{xx} + \frac{1}{2}\,\sigma[\sigma(u_\alpha)_x]_x - E[u\, u_x\, T_\alpha(t)]\right\} dt \\ &\quad - \sigma\left(\sum_k m_k(t)\sqrt{\alpha_k}\; \frac{\partial}{\partial x} u_{\alpha^-_{(k)}}\right) dt. \end{aligned}$$

Expanding the nonlinear term $E[u\, u_x\, T_\alpha(t)]$ by formula (2.38), we get the Wiener chaos propagator of equation (3.27):

$$\begin{aligned} (u_\alpha)_t =& \mu(u_\alpha)_{xx} + \frac{1}{2}\,\sigma[\sigma(u_\alpha)_x]_x - \sigma\sum_k m_k(t)\sqrt{\alpha_k}\; \frac{\partial}{\partial x}\left(u_{\alpha^-_{(k)}}\right) \\ & - \frac{1}{2}\sum_{p\in\mathcal{J}}\sum_{0\leqslant\beta\leqslant\alpha} C(\alpha,\beta,p)(u_{\alpha-\beta+p}u_{\beta+p})_x. \end{aligned} \tag{3.29}$$

The multi-index $\alpha^-_{(k)}$ is defined as

$$\alpha^-_{(k)}(j) = \begin{cases} \alpha_j & \text{if } j \neq k, \\ \alpha_k - 1 & \text{if } j = k. \end{cases} \tag{3.30}$$

The index $\alpha^-_{(k)}$ can be obtained from α by subtracting 1 from the component α_k. For example, if

$$\alpha = (4, 3, 0, 1),$$

then

$$\alpha^-_{(1)} = (3, 3, 0, 1), \quad \alpha^-_{(2)} = (4, 2, 0, 1), \quad \alpha^-_{(3)} = (4, 3, -1, 1), \quad \alpha^-_{(4)} = (4, 3, 0, 0).$$

The index $\alpha^-_{(3)}$ has a negative component so it is not a valid index. If $\alpha \notin \mathcal{J}$, we simply set the corresponding WCE coefficient as zero $u_\alpha \equiv 0$.

Since the initial condition in (3.25) is deterministic, the corresponding initial conditions for the WCE coefficients are

$$u_\alpha(x, 0) = \begin{cases} u_0(x) & \text{if} \quad \alpha = \mathbf{0}, \\ 0 & \text{if} \quad \alpha \neq \mathbf{0}. \end{cases} \tag{3.31}$$

Since the solution of (3.25) is periodic, all the WCE coefficients also satisfy the periodic boundary condition

$$u_\alpha(0, t) = u_\alpha(1, t). \tag{3.32}$$

The deterministic PDE system (3.29) augmented with the initial condition (3.31) and boundary condition (3.32) is equivalent to the random equation (3.25). By solving a truncated version of the WCE propagator (3.29), we can obtain a finite term Wiener chaos solution of (3.25), which represents the random solution as a functional of the white noise forcing. In particular, we can compute the statistic moments of the random solution based on Theorem 2.2 and Lemma 2.5.

3.2.2 Monte Carlo Simulation

For comparison purpose, we will also solve the multiplicative stochastic Burgers equation (3.27) by MC simulation. The MC simulation for SPDEs with multiplicative forcing is much more complicated than the one for additive random forcing. We will use the weakly second order Runge-Kutta scheme (B.4) in the MC simulation. Higher order numerical schemes exist but are all extremely complicated and have little practical use. Please see Appendix B for a detailed discussion about the numerical schemes for discretizing stochastic differential

equations in MC simulation.

Rewrite the differential form (3.28) of the stochastic Burgers equation in an abbreviated form

$$d\,u = a(x,u)\,d\,t + b(x,u)\,d\,W(t),$$

where

$$a(x,\,u) = \mu\,u_{xx} + \frac{1}{2}\sigma\,(\sigma\,u_x)_x - uu_x$$

and

$$b(x,u) = -\sigma\,\partial_x u.$$

Here the variable x is treated as a parameter. From Appendix B, the weakly second order Runge-Kutta method (B.4) has the following form:

$$\begin{aligned} u_{n+1} = u_n &+ \frac{1}{2}[a(x,u_n) + a(x,\hat{u}_{n+1})]\Delta t \\ &+ \frac{1}{4}[2b(x,u_n) + b(x,u_n^+) + b(x,u_n^-)]\Delta W_n \\ &+ \frac{1}{4}[b(x,u_n^+) - b(x,u_n^-)]\,\{(\Delta W_n)^2 - \Delta t\}/\sqrt{\Delta t}, \end{aligned} \tag{3.33}$$

where

$$\hat{u}_{n+1} = u_n + a(x,u_n)\Delta t + b(x,u_n)\Delta W_n,$$

and

$$u_n^{\pm} = u_n + a(x,u_n)\Delta t \pm b(x,u_n)\,\sqrt{\Delta t}.$$

Based on the specific form of $b(x,u)$, the Runge-Kutta scheme (3.33) can be simplified as

$$\begin{aligned} u_{n+1} = u_n \quad &+ \quad \frac{1}{2}[a(x,u_n) + a(x,\hat{u}_{n+1})]\Delta t \\ &+ \quad b(x,u_n)\Delta W_n - \frac{1}{2}\,\sigma\partial_x[a(x,u_n)]\Delta t\Delta W_n \\ &+ \quad \frac{1}{2}\,\sigma\,\partial_x\,[\sigma\,\partial_x\,(u_n)]\{(\Delta W_n)^2 - \Delta t\}. \end{aligned}$$

The above numerical scheme has a weakly second order accuracy in time. Comparing with the weakly first order Euler scheme, the second order Runge-Kutta scheme involves the extra terms $\Delta t\Delta W_n$ and $(\Delta W_n)^2 - \Delta t$, which have the order of $(\Delta t)^{\frac{3}{2}}$. That is why the numerical scheme can achieve second order accuracy $O(\Delta t)^2$ in time discretization. For

stochastic differential equations with additive forcing, those extra terms will disappear and the numerical scheme degenerates to the regular modified second order Euler schemes. That is the main reason that SDEs with additive forcing is much easier to solve than SDEs with multiplicative forcing. Note that the solution of equation (3.25) is periodic in space; we discretize the spatial derivatives by the spectral method using fast Fourier transform.

An additional complexity in solving multiplicative SPDEs by MC simulation is the special CFL condition. For the stochastic Burgers equation (3.27), the CFL condition requires that

$$\frac{|\sigma \Delta W_n|}{\Delta x} \leqslant C, \tag{3.34}$$

where ΔW_n is the increment of the Brownian motion at the nth step. However, the Gaussian random variable $\Delta W_n \sim N(0, \Delta t)$ does not have a compact support. In practice it is usually enough to bound ΔW_n by four times its standard deviation, that is, assume $|\Delta W_n| \leqslant 4\sqrt{\Delta t}$ since the probability $P(|\Delta W_n| > 4\sqrt{\Delta t}) = O(10^{-5})$. So the CFL condition (3.34) can be approximated by

$$4|\sigma| \frac{\sqrt{\Delta t}}{\Delta x} \leqslant C. \tag{3.35}$$

The condition (3.35) requires that $\Delta t \sim O(\Delta x)^2$, which is an order smaller than the CFL condition $\Delta t \sim O(\Delta x)$ for the additive stochastic Burgers equation.

3.2.3 Numerical Experiments

Next we solve the multiplicative stochastic Burgers equation (3.25) by the WCE method, and compare its numerical results with those by MC simulation. We take the initial condition of (3.25) as

$$u_0(x) = \frac{1}{2}(e^{\cos\, 2\pi x} - 1.5) \sin[2\pi(x + 0.37), \tag{3.36}$$

and set $\mu = 0.01$. The spatial part of the random forcing is chosen as $\sigma(x) = 0.1 \sin(2\pi x)$. We solve the stochastic Burgers equation (3.25) to $T = 0.8$ by the WCE method.

The WCE propagator (3.29) is an infinite system and we need to truncate it for computational purpose. We will show in Theorem 5.4 that for the model equation if we keep K number of Gaussian random variables and maximum Nth order Wick polynomials in the

finite WCE, the error of the truncation would be

$$O\left[\frac{\left(\sigma T^{1/2}\right)^{N+1}}{\sqrt{(N+1)!}}+\sigma\left(\frac{T}{K}\right)^{1/2}\right],$$

where σ is the magnitude of the random forcing, and T is the size of the time interval for the solution. This error estimate is very similar to the one for the additive stochastic Burgers equation. Guided by this error estimate, we project the Brownian motion $\{W(t); 0 \leqslant t \leqslant T\}$ onto eight cosine functions and truncate the WCE at maximum fourth order Wick polynomials. That is, we set $K = 8$, $N = 4$ and truncate the WCE of the solution as

$$u_{K,N} = \sum_{|\alpha|\leqslant N} u_\alpha(x,t) \prod_{k=1}^{K} H_{\alpha_k}(\xi_k). \tag{3.37}$$

The error for such a simple truncation is expected to be $O(10^{-2})$.

Similar to the additive stochastic Burgers equation, we can apply the sparse truncation strategy for the multiplicative stochastic Burgers equation. Based on Theorem 5.6, we set the adaptive index as

$$r = (4, 4, 3, 2, 1, 1, 1, 1).$$

The optimal adaptive index r is decided by the error balance condition (5.29). Define the sparse index set

$$\mathcal{J}^r_{K,N} = \{(\alpha_1, \ldots, \alpha_K);\ |\alpha| \leqslant N,\ \alpha_i \leqslant r_i\}.$$

Then the sparse WCE truncation of the solution can be denoted as

$$u^r_{K,N} = \sum_{\alpha\in\mathcal{J}^r_{K,N}} u_\alpha(x,t) T_\alpha(\xi). \tag{3.38}$$

We will see later (Theorem 5.6) that the sparse truncation (3.38) has the same asymptotic convergence rate as the simple truncation (3.37).

Since the WCE decays both in the number of retained Gaussian random variables and the order of the Hermite polynomials, we should include less Gaussian random variables in constructing higher order Wick polynomials. For example, we can include only the first four ξ_k in the third order Wick polynomials, and the first three ξ_k in the fourth order Wick polynomials. This practice is based on the asymptotic decaying rate of the WCE coefficients

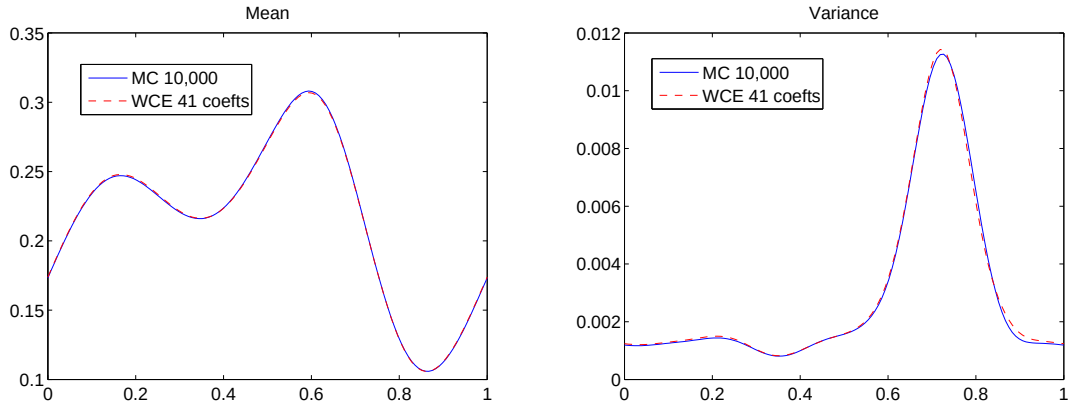

Figure 3.8: The mean and variance computed by the WCE method using 41 coefficients and MC simulation with 10,000 realizations. The results by the two different methods agree with each other quite well. However, the CPU time for the WCE method is only 25 seconds, as opposed to 1,466 seconds for the MC simulations.

(Theorem 5.5). By eliminating the unimportant crossing products between different ξ_k at higher polynomial order, we can reduce the dimension of the truncated WCE further. Using these sparse truncation techniques, the resulting WCE propagator has only 41 coefficients, as opposed to 495 coefficients in the simple truncation (3.37). We solve the finite WCE propagator by spectral method using fourth order Runge-Kutta method.

For comparison, we also solve the stochastic Burgers equation (3.25) by MC simulation using the weakly second order Runge-Kutta method (B.4). We compute the mean and variance of the solution by averaging 10,000 MC realizations. Figure 3.8 is the comparison of the numerical results obtained by the WCE method and the MC simulation. The variance of the solution is much larger near the sharp layers than in the smooth region of the solution. With only 41 coefficients, the results of the WCE method agree quite well with those by the MC simulation with 10,000 realizations. However, the WCE method takes only 25 seconds, as opposed to almost 25 minutes required by the MC simulation.

3.2.4 Convergence Verification of the WCE Method

Since the MC simulation converges very slowly, it is quite difficult to assess the accuracy of the results by the MC method. Next, we will test the convergence of the WCE method and the MC simulation in a special case, where a semi-analytical solution exists and can be computed very accurately.

In the stochastic Burgers equation (3.25), suppose the spatial part $\sigma(x)$ of the random

forcing is a constant. In this particular case the solution of the stochastic Burgers equation (3.25) can be obtained by a stochastic perturbation of its deterministic counterpart. More specifically, we have the following theorem:

Theorem 3.2 *If the spatial part σ of the random forcing in equation (3.25) is a constant, then its solution has the expression*

$$u(x,\ t) = v\big(x - \sigma W(t),\ t\big), \tag{3.39}$$

where $v(x,\ t)$ is the solution of the unforced deterministic Burgers equation

$$\begin{cases} v_t + \frac{1}{2}(v^2)_x = \mu v_{xx}, \\ v(x,0) = u_0(x), \quad v(0,t) = v(1,t). \end{cases} \tag{3.40}$$

Proof Let

$$X(x,t) = x - \sigma W(t),$$

then the formula (3.39) changes to $u(x,t) = v(X,t)$. We have

$$u_x(x,t) = v_x(X,t), \qquad u_{xx}(x,t) = v_{xx}(X,t).$$

Since v satisfies the differential equation (3.40), we have

$$v_t(X,t) = \mu v_{xx}(X,t) - v(X,t)v_x(X,t).$$

From the differential rules of Stratonovich integrals, we have

$$\begin{aligned} du(x,t) &= v_t(X,t)\ dt - \sigma\, v_x(X,t) \circ d\, W(t) \\ &= \big[\mu\, v_{xx}(X,t) - v(X,t)\ v_x(X,t)\big] dt - \sigma\, v_x\,(X,t) \circ d\, W(t) \\ &= \big[\mu\, u_{xx}(x,t) - u(x,t)\ u_x(x,t)\big] dt - \sigma\, u_x\,(x,t) \circ d\, W(t). \end{aligned}$$

So formula (3.39) indeed satisfies the differential equation of (3.25). It is easy to check that (3.39) also satisfies the initial condition and the periodic boundary condition. □

Note As we emphasized before, the multiplicative forcing in the stochastic Burgers

equation (3.25) should be proposed in Stratonovich's sense rather than in Ito's sense. Otherwise, the resulting SPDEs can be ill-posed. That is a subtle but important difference between stochastic partial differential equations and stochastic ordinary differential equations. It reflects some of the extra difficulties in studying SPDEs. In stochastic ordinary differential equations, the stochastic integral can be defined in either way, and it is very easy to change from one form to the other. For SPDEs, the transformation between Ito integrals and Stratonovich integrals is a little bit more tricky due to the presence of the spatial derivatives.

Based on Theorem 3.2, we can construct a simple example to demonstrate the necessity of proposing multiplicative SPDEs in Stratonovich's sense. Consider the following 1-D stochastic heat equation in the real axis

$$d\,u(x,t) = \mu u_{x\,x}(x,t)dt + \sigma u_x \cdot d\,W(t), \tag{3.41}$$

where σ is a constant and the stochastic integral $\sigma u_x \cdot d\,W(t)$ is in Ito's sense. We can rewrite (3.41) in Stratonovich's sense

$$d\,u(x,t) = \left(\mu - \frac{1}{2}\sigma^2\right) u_{x\,x}(x,t)dt + \sigma u_x \circ d\,W(t).$$

Similar to Theorem 3.2, the solution of (3.41) has the analytical expression

$$u(x,t) = v\big(x - \sigma\,W(t),\ t\big), \tag{3.42}$$

where v satisfies the deterministic heat equation

$$v_t(x,t) = \left(\mu - \frac{\sigma^2}{2}\right) v_{x\,x}(x,t). \tag{3.43}$$

If $\sigma^2/2 > \mu$, then the deterministic heat equation (3.43) has negative diffusion and is ill-posed. As a result, the solution of (3.41) is also ill-posed because it is a stochastic translation of the ill-posed solution v.

However, if the equation (3.41) is proposed in Stratonovich's sense

$$d\,u(x,t) = \mu u_{x\,x}(x,t)dt + \sigma u_x \circ d\,W(t),$$

then the formula (3.42) still holds, but v satisfies a different heat equation

$$v_t(x,t) = \mu\, v_{xx}(x,t),$$

which is well-posed for all positive μ. In this example, the Stratonovich integral automatically provides a diffusion mechanism to balance its strong "dispersion" along the characteristic velocity, which is critical for multiplicative SPDEs.

Theorem 3.2 states that for spatially independent random forcing, the stochastic solution can be obtained by a random translation of the corresponding deterministic solution. The explicit formula (3.39) provides a very convenient way to compute the statistic moments of the solution of (3.25). Define the random variable $Y = \sigma W(t)$, then $Y \sim N(0, \sigma^2 t)$ and its probability density function is

$$\rho(y) = \frac{1}{\sigma\sqrt{2\pi t}}\, e^{-\frac{y^2}{2t\sigma^2}}.$$

Then the statistical moments of u has the explicit form

$$E[u^n(x,t)] = E[v^n(x-Y,t)] = \int_{-\infty}^{+\infty} v^n(x-y,t)\rho(y)\, dy.$$

Since $v(x,t)$ is periodic, we have

$$\begin{aligned} E[u^n(x,t)] &= \sum_{n=-\infty}^{\infty} \int_n^{n+1} v^n(x-y,t)\rho(y)\, dy \\ &= \int_0^1 v^n(x-y,t) \sum_{n=-\infty}^{\infty} \rho(y+n)\, dy. \end{aligned}$$

Denote

$$\rho_d(y) = \sum_{n=-\infty}^{\infty} \rho(y+n)\,,$$

then

$$E[u^n(x,t)] = \int_0^1 v^n(x-y,t)\, \rho_d(y)\, d\, y. \tag{3.44}$$

Because of the symmetry of the Gaussian kernel $\rho(y)$, it can be shown that $\rho_d(y)$ is periodic in the interval $[0,1]$. Hence the integration (3.44) is a convolution of two periodic functions, which can be computed by FFT very easily. Note that $\rho_d(y)$ is defined by an infinite

summation, we need to truncate it at finite terms. Since $\rho(y)$ is a Gaussian kernel and decays exponentially, the truncation can be made very accurate.

Next we test the convergence of the WCE method and the MC simulation on the model equation using the semi-analytical formula (3.44). We choose the same initial condition (3.36) and set $\sigma = 0.1$. The viscosity $\mu = 0.01$ remains the same. We compare the numerical solutions of the WCE method and MC simulation at $T = 0.8$. We first solve v from the deterministic equation (3.40) by the spectral method in space and the fourth order Runge-Kutta in time. Then we compute the mean and variance of the random solution based on formula (3.44) by FFT. Such computed mean and variance do not involve random number generations and are very accurate.

Then we solve the same problem by the WCE method. Based on Theorem 5.4, we project the Brownian motion onto eight cosine bases and use maximum fifth order Hermite polynomials. The expected error for such a truncation would be $O(10^{-2})$. Based on Theorem 5.6, we choose the optimal adaptive index as $r = (5, 4, 3, 2, 1, 1, 1, 1)$. As before, we include only the first four ξ_k in the third order Wick polynomial, and the first three ξ_k in the fourth and fifth order Wick polynomial. Such a sparse truncation can reduce the number of WCE coefficients from 1287 to 51. We solve the resulting finite WCE propagator by the fourth order Runge-Kutta method. In the MC simulation, we average 10,000 realizations to compute the mean and variance. Figure 3.9 is the mean and variance computed by the WCE method and the MC simulation. Comparing with the semi-analytical solutions, the numerical results of the WCE method and MC simulation are both quite accurate. Table 3.2 is the relative error of the two methods. For both the mean and variance, the relative error of the WCE solution is about $O(10^{-2})$, which is consistent with the error prediction based on Theorem 5.6. With 51 coefficients, the WCE method has similar accuracy as the MC simulation with 10,000 realizations. However, the WCE method only takes 43 seconds, as opposed to the 1,466 seconds for the MC simulation. So the WCE method is much more efficient than the MC simulation at this case.

In Figure 3.10, we plot the errors of the WCE method truncated at first, second, third, fourth, and fifth order Wick polynomials, respectively. The relative error is plotted in log scale using the semi-analytical solution as the benchmarks. The errors decays linearly in the log scale. So the WCE truncation converges exponentially in terms of the Wick polynomial orders, which is consistent with the error estimate given in Theorem 5.4.

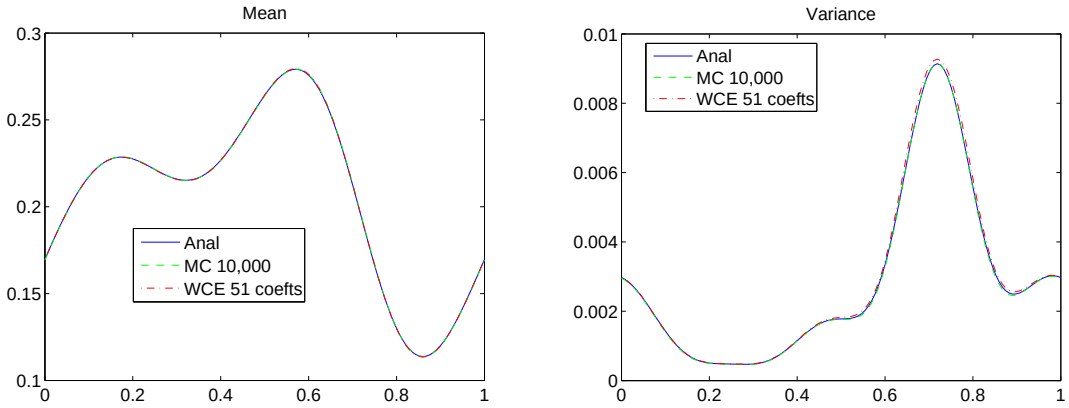

Figure 3.9: The mean (left) and variance (right) computed by the WCE method with 51 coefficients and MC simulation with 10,000 realizations. Both methods agree quite well with the semi-analytical solution. However, the CPU time for the WCE method is only 43 seconds, as opposed to 1,466 seconds for MC simulations.

For comparison purpose, we also compute the RMS errors for MC simulations with 500, 1,000, 2,000, 4,000, 8,000 realizations, respectively. We repeat the simulations 20 times to estimate the RMS errors for each realization number. Figure 3.11 is the log-log plot of the RMS errors with respect to the different realization numbers. Based on the decaying rate of the RMS error, the MC simulations indeed require 10,000 realizations to achieve the same accuracy as the WCE method with a fifth order truncation.

	CPU Time (s)	Mean Error	Variance Error
WCE	43	0.09 %	1.8 %
MC	1466	0.46 %	1.2 %

Table 3.2: Relative errors of the mean and variance by the WCE method with 51 coefficients and MC simulation with 10,000 realizations. The WCE method has similar accuracy as the MC simulation, but is more efficient in terms of CPU time. The WCE method is more accurate in computing the mean than in computing the variance.

3.3 Numerical Verification of Stationary Measure of Stochastic Burgers Equations

In the above numerical experiments, we only presented the numerical results by the WCE method up to time $T = 0.8$. Next we solve the stochastic Burgers equation (3.1) by the sparse WCE method in longer time intervals. A sparse WCE truncation is equivalent to a coarse-grid discretization of the random solution in the probability space. We will

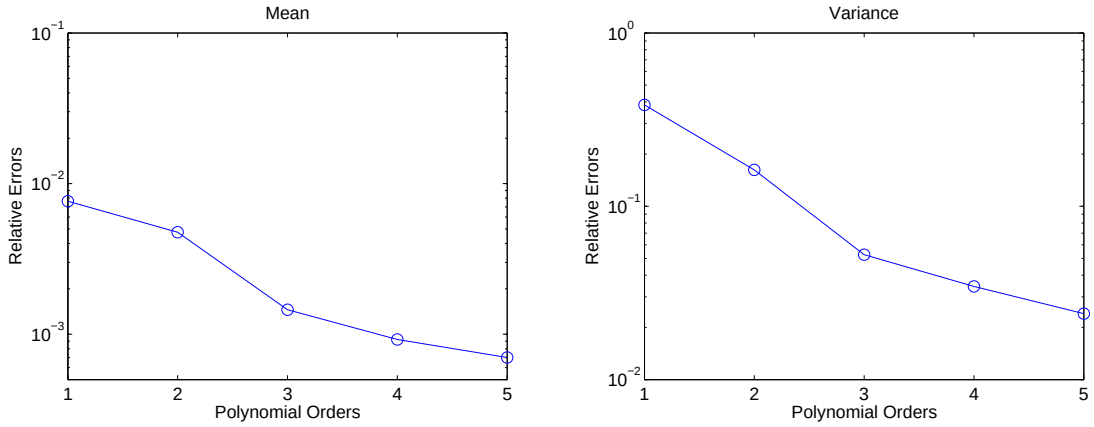

Figure 3.10: Log plot of the relative error of the WCE method with first, second, third, fourth, and fifth order Wick polynomials, respectively. The convergence is approximately exponential in terms of the Wick polynomial orders, which is consistent with the error estimate Theorem 5.6.

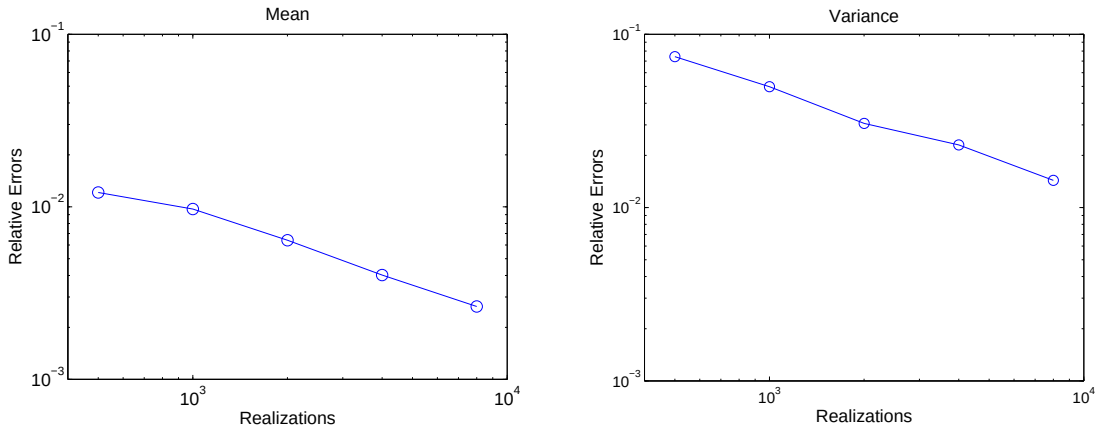

Figure 3.11: Log-log plot of the relative errors of the mean (left plot) and variance (right plot) computed by MC simulation with 500, 1,000, 2,000, 4,000, 8,000 realizations, respectively. The MC method reaches the similar accuracy as the WCE method with 51 coefficients when 8,000 realizations are computed.

show that the sparse WCE truncation can capture the large-scale structure of the random solution in relatively long time intervals with a small number of coefficients. That is a very important and unique feature of the WCE method. Using the sparse WCE method, we will verify numerically the existence of a stationary measure for the additive stochastic Burgers equation (3.1).

We assume that the random forcing in (3.1) has zero mean in space. So $\sigma(x)$ satisfies the condition

$$\int_0^1 \sigma(x)dx = 0. \tag{3.45}$$

Integrating the equation (3.1) in the period $[0,1]$ and using the periodicity, we have

$$\frac{d}{dt}\int_0^1 u(x,t,\omega)dx = \mu u_x\Big|_0^1 - \frac{u^2}{2}\Big|_0^1 - \dot{W}\int_0^1 \sigma(x)dx = 0.$$

So the integral $\int_0^1 u(x,t,\omega)dx$ is conserved in time even though it is a random process. Denote the function space

$$D_c = \Big\{u \in C[0,1],\ u(0) = u(1),\ \int_0^1 u(x)dx = c\Big\}.$$

If the initial condition $u_0(x)$ belongs to the space D_c, then all the later solutions will stay in the same space D_c. In this section, we will focus on the space D_0 only, that is, we only consider initial conditions with zero mean in space. The numerical results for D_c with $c \neq 0$ are very similar.

The solution of the stochastic Burgers equation (3.1) is a Markov process in time. The state of the solution at later time only depends on the current state and the equation itself. It does not depend on the history of the solution. Denote the transition kernel of the Markov chain generated by equation (3.1) as $K^t(u,v)$, where $u(x), v(x) \in D_0$ are two states of the random solution. The transition kernel $K^t(u,v)$ measures the probability that the solution will be in state $v(x)$ at time t when its initial state is $u(x)$. For any measurable subset $A \subset D_0$, denote the transition probability

$$K^t(u,A) = \int_A K^t(u,v)dv.$$

A measure $\pi(du)$ is called a stationary (invariant) measure of the stochastic Burgers equation (3.1) if the following condition is satisfied:

$$\pi(A) = \int K^t(u,\ A)\pi(du).$$

The stationary measure means that if the initial state (initial condition) of the equation (3.1) satisfies the distribution $\pi(u)$, then its solution at later time will always satisfy the same distribution $\pi(u)$. In other words, the distribution of the solution does not change in time and is stationary with respect to the transitional kernel $K^t(u,\ v)$.

It has been proven in [92, 93] that there exists a unique invariant measure for the

stochastic Burgers equation (3.1) when the condition (3.45) is satisfied. In the equation (3.1), the random forcing continuously supplies energy into the system, while the viscous term is responsible for dissipating the energy in the system. The existence of the invariant measure is a direct result of the balance between these two opposite effects. It is crucial that the spatial part of the random forcing satisfies the zero mean condition (3.45). To guarantee the existence of the invariant measure, we should not force the zero mode of the random solution since it is not dissipated by the viscosity.

Even though an invariant measure exists for the equation (3.1), a numerical verification of its existence is quite difficult. Suppose $P(u, x, t)$ is the probability density function (PDF) of the random solution $u(x, t, \omega)$. The PDF $P(u, x, t)$ denotes the probability distribution of the solution u at the point (x, t). If there exists an invariant measure for the equation (3.1), then the PDF $P(u, x, t)$ must have a steady state when t is large. However, it is impossible to compute the PDF function $P(u, x, t)$ directly, since there is no closed form of Fokker-Planck equations for the PDF $P(u, x, t)$, due to the nonlinearity of the problem [23].

The WCE method provides an alternative for studying the invariant measure of the stochastic Burgers equation (3.1) numerically. Denote the WCE of the random solution of (3.1) as

$$u(x,\, t) = \sum_{\alpha} u_\alpha(x,\, t)\, T_\alpha(\xi), \qquad t \le T.$$

The random bases $T_\alpha(\xi)$ are fixed in the time interval $[0, T]$ and do not change in time. Hence the PDF of $u(x, t)$ is uniquely determined by the WCE coefficients $u_\alpha(x, t)$. Theoretically, if all the WCE coefficients $u_\alpha(x, t)$ are known, we should be able to derive the PDF of $u(x, t)$ precisely. In particular, if all the WCE coefficients $u_\alpha(x, t)$ do not change in time, then the PDF of the random solution $u(x, t)$ has a steady state, which is direct evidence of the existence of an invariant measure. So the existence of an invariant measure for equation (3.1) is equivalent to the existence of a stationary solution for its WCE propagator (3.6).

Next we solve the stochastic Burgers equation (3.1) by the sparse WCE method and see whether it can resolve the invariant measure of the solution. We set the spatial part of the forcing as $\sigma(x) = 0.5\cos(2\pi x)$. Obviously, $\sigma(x)$ has zero mean in space. The viscosity

constant is chosen as $\mu = 0.01$. We first consider the initial condition

$$u_0(x) = 0.5\ \cos(4\,\pi\,x),$$

which has zero mean in space. So all the later states of the random solution will stay in the same space D_0. If there exists an invariant measure for the system (3.1), then that measure must be supported in the space D_0.

We solve the stochastic Burgers equation by the WCE method to $T = 6.0$. We keep 12 Gaussian random variables and truncate the WCE at fifth order Wick polynomials. For the sparse truncation, we set the optimal adaptive index as $r = (5, 4, 3, 2, 1, \ldots, 1)$ and reduce the number of Gaussian random variables at higher order polynomials. Specifically, we only include the first six ξ_k in the second and third order Wick polynomials, and the first three ξ_k in the fourth and fifth order Wick polynomials. This is a very sparse WCE truncation for $T = 6.0$. The resulting WCE truncation has only 91 coefficients, as opposed to 6188 coefficients in the simple truncation. We plot the evolution history of the mean and variance of the random solution in Figure 3.12. The mean and variance profiles are plotted for every $\Delta t = 0.1$. There are 61 curves in each plot. Due to the viscosity effects, the mean curve decays monotonously initially. However, the dissipation is quickly slowed down by the constant supply of energy from the random forcing. Soon these two opposite mechanisms balance each other and the mean converges to a steady state (attractor). The variance is initially zero and grows rapidly due to the random forcing. As the variance increases and develops a steep overshoot, the viscose effect becomes dominant and smears the rough structure of the variance. Eventually, the variance also converges to a steady state.

To access the invariance of the mean and variance in time, we compute the L^2 norm of the differences between consecutive time steps. For example, the incremental change of the mean in time is measured by

$$e(n) = \|u_{\mathbf{0}}(\cdot, t_{n+1}) - u_{\mathbf{0}}(\cdot, t_n)\|_2,$$

where $t_n = n\Delta t$ and $\Delta t = 0.1$. The changes of the variance is measured similarly. We plot the incremental changes of the mean and variance in time in Figure 3.13. Though fluctuating quite a bit at the beginning, both the mean and variance change very little

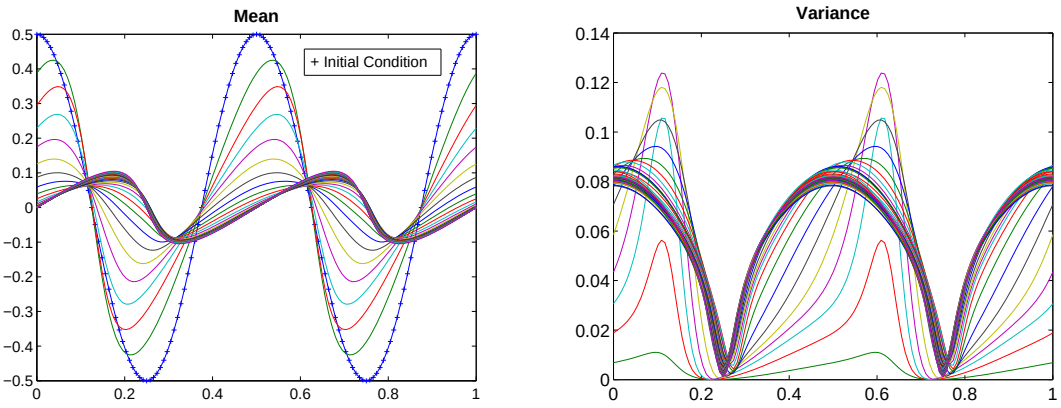

Figure 3.12: Time history of the mean and variance of the stochastic Burgers equation for $t \leq 6.0$, computed by the WCE method with 91 coefficients. The mean decreases monotonously and converges to a steady state very quickly. The variance increases from zero gradually and also converges to a steady state.

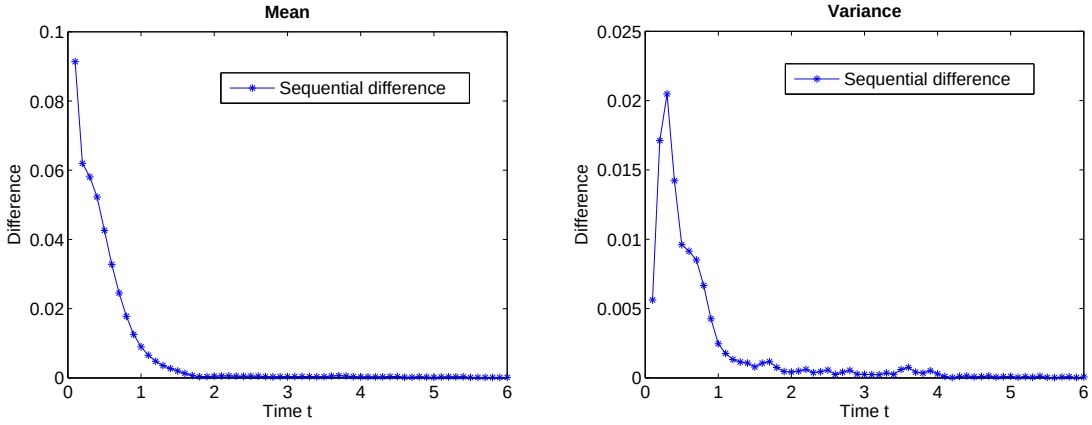

Figure 3.13: Incremental changes of the statistic moments at sequential time steps, measured by $\|u(x,t_{n+1}) - u(x,t_n)\|_2$. The time step size is $\Delta t = 0.1$. The mean and variance fluctuate very little when $t > 2.0$, indicating that all the WCE coefficients are stationary for $t > 2.0$.

after $t > 2.0$, which indicates that all the WCE coefficients are stationary after $t = 2.0$. In fact, we indeed observed in the numerical results that all the WCE coefficients become stationary when $t > 2.0$. As we pointed out before, the existence of a stationary WCE solution is equivalent to the existence of a stationary measure. So the numerical result by the WCE method is a direct evidence that the stochastic Burgers equation (3.1) has a stationary measure. In addition, the sparse WCE method indeed successfully resolved the stationary measure numerically.

Next we consider a different initial condition

$$u_0(x) = 0.5\,\sin(2\,\pi\,x),$$

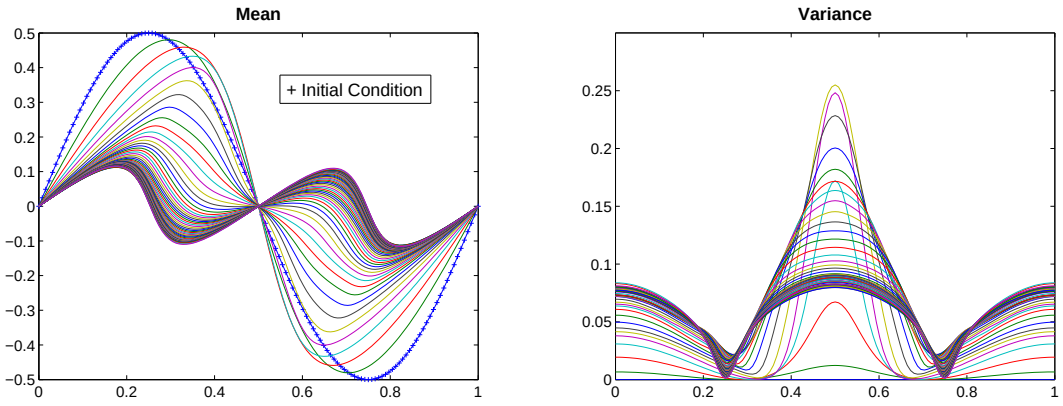

Figure 3.14: Time history of the mean and variance of the stochastic Burgers equation with a different initial condition. The mean and variance converge to the same steady states as in the first example, even though the convergence pattern is quite different.

which is orthogonal to the initial condition used in the previous test. We solve the problem to $T = 6.0$ again and truncate the WCE in the same way as above. We plot the time history of the mean and variance in Figure 3.14. Not surprisingly, the mean and variance converge to the same steady states, even though the convergence pattern is quite different. Figure 3.15 is the comparison of the statistics at $T = 6.0$ for the two different initial conditions. They are indeed very similar to each other. We also tested many other deterministic and random initial conditions in the space D_0 and found that they all converge to the same steady states. These results are consistent with the uniqueness of the invariant measure. The numerical experiments also suggest that the invariant measure is a global attractor in the state space D_0. If this is true, then starting from an arbitrary initial state u_0, the transition kernel $K^t(u_0, \cdot)$ generated by the dynamics of (3.1) will always converge to the unique invariant measure at large t. The global convergence of the invariant measure is closely related to the egordicity of the stochastic dynamics system (3.1).

In the above numerical experiments, we used a very sparse truncation in the WCE solution, with only 91 coefficients for $T = 6.0$. Based on the error estimate (3.10), such a sparse truncation may seem not sufficient. Surprisingly, the sparse WCE method captures the large-scale structures of the random solution quite well, and it indeed converges to the right stationary state. The sparse WCE method may not very accurate in small scales at the transitional stage. However, as the time t increases, the sparse WCE solution still resolves the invariance measure of the random solution correctly.

It is worth pointing out that no invariant measure exists for the stochastic Burgers equa-

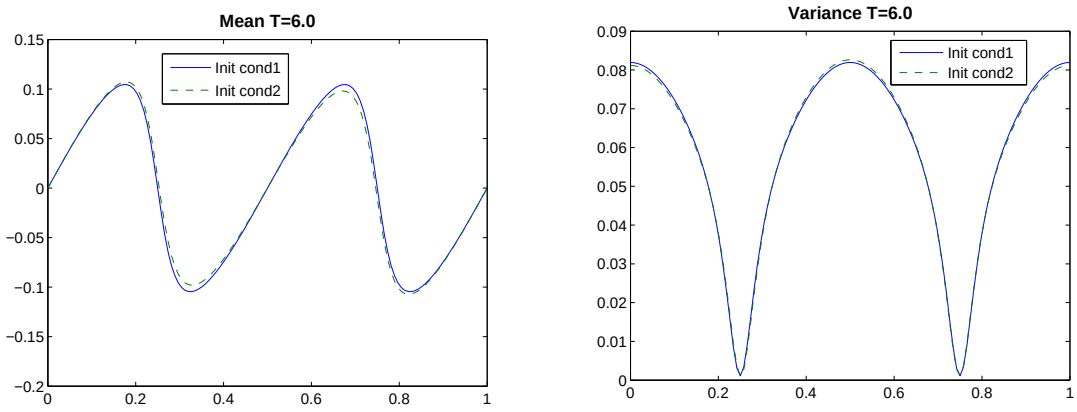

Figure 3.15: The means and variances at $T = 6.0$ from different initial conditions. The random solutions converge to the same steady state, which suggests that the unique invariant measure is a global attractor.

tions with multiplicative forcing (3.25). In the equation (3.25), the random perturbation is along the velocity gradient and it only enhances the convection of the solution. It does not supply energy into the system to balance the dissipation mechanism caused by the viscosity. Hence the stochastic Burgers equation (3.25) is a purely decaying system and no invariant measure can exist.

3.4 Random Forcing with Large Variability

Based on the error estimate (3.10), the accuracy of the WCE truncation also depends on the magnitude of the random forcing. Next, we test how the magnitude of the random forcing affects the performance of the WCE method. In the stochastic Burgers equation (3.1), we set the spatial part of the random forcing as

$$\sigma(x) = A\cos(2\pi x),$$

where A is a constant. The viscosity is chosen as $\mu = 0.02$. We solve the problem to $T = 0.8$ for $A = 1.0,\ 2.0,\ 4.0$, respectively. The numerical results are plotted in Figure 3.16. When the magnitude of the random forcing increases, we need to include more and more WCE coefficients to achieve similar accuracy. When the random forcing is very large, the dynamics of the system is dominated by the random effect. For $A = 4.0$, the variance of the solution is already bigger than its mean. So the random solution will vary dramatically from realization to realization. Theoretically, the increase of the magnitude of the forcing

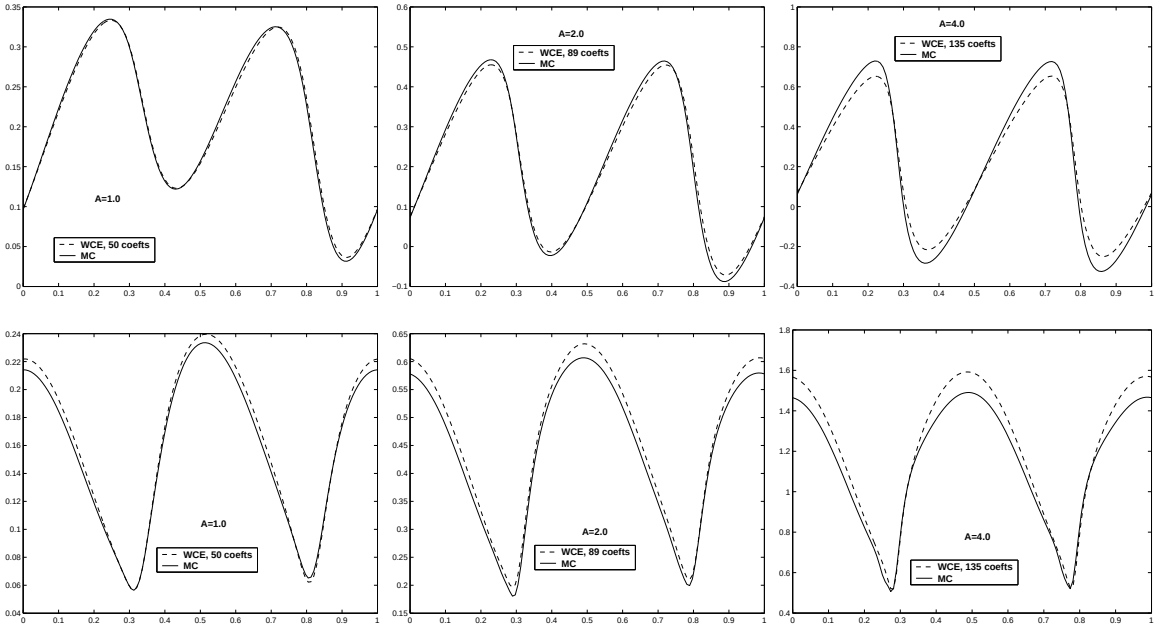

Figure 3.16: Mean (first row) and variance (second row) for different magnitudes of forcing. The first column: $A = 1.0$; the second column: $A = 2.0$; the third column: $A = 4.0$. In all three cases, $10,000$ realizations are computed in MC simulation. When A increases from 1.0 to 4.0, we need more and more WCE bases to achieve similar accuracy. However, the WCE method can approximate the right solution reasonably well when enough coefficients are retained

does not lead to a blowup of the WCE method, as long as the solution still has finite variance. Our numerical experiments indicate that the WCE method is quite stable and does not blow up even when hundreds of the coefficients are included.

The sparse WCE method captures the mean and variance fairly well for $A = 1.0$ and $A = 2.0$. However, for $A = 4.0$ the direct WCE method tends to deviate from the MC simulation near the rough region of the solution. Increasing the order of the WCE truncation will improve the WCE solution, but not very significantly. The reason is that the WCE coefficients at increasingly higher order truncations are smaller and smaller, hence contribute less and less to the WCE solution. Since there are so many of those small coefficients at higher orders, it is impossible to capture their aggregated effects by keeping only a few dozens more of them. In Chapter 6, we will study an alternative computational strategy for the WCE method, where the coarse-scale WCE solution is refined by the MC simulations. By combining the merits of the WCE method and the MC simulation, we can design a more efficient computational strategy for solving SPDEs in long time intervals or with large random forcing. We will revisit this problem in Section 6.3 using the new WCE-MC hybrid method.

Chapter 4

WCE Methods for Stochastic Navier-Stokes Equations

In this section, we apply the WCE method to solve a 2-D stochastic Navier-Stokes (SNS) equation. We consider a temperature distribution convected by a stochastic velocity field $\vec{u}$, where $\vec{u} = (u, v)$ is governed by the Navier-Stokes equation driven by Brownian motions:

$$\begin{cases} \theta_t + \vec{u} \cdot \nabla\theta = \mu\Delta\theta, \\ \vec{u}_t + \vec{u} \cdot \nabla\vec{u} = \nu\Delta\vec{u} - \nabla P + \sigma\dot{\vec{W}}(t), \\ \nabla \cdot \vec{u} = 0, \end{cases} \tag{4.1}$$

where θ is the scaled temperature, and $\vec{W}(t) = (W_1,\ W_2)^T$ is a Brownian motion vector with independent components. The matrix $\sigma(x, y) = \mathrm{diag}(\sigma_1,\ \sigma_2)$ accounts for the spatial dependence of the random forcing. μ and ν are the temperature diffusivity and fluid viscosity, respectively. The temperature θ is convected by the random flow passively, and serves as a visualization media for the fluid particle path. This problem is not only a challenging test for the WCE method, but also an interesting setup for studying the mixing effect induced by the random forcing.

In our numerical experiment, we take the computational domain as $[0, 1]^2$, and assume that θ and $\vec{u}$ are doubly periodic in the domain. We consider the following initial condition for the vorticity $\omega = v_x - u_y$:

$$\omega(x, y, 0) = C - \frac{1}{2\delta} \exp\left(-\frac{I(x)(y - 0.5)^2}{2\delta^2}\right), \tag{4.2}$$

where $I(x) = 1+\epsilon(\cos(4\pi x)-1)$, and C is a constant to make the initial vorticity mean zero: $\int_{[0,1]^2} \omega(x,y,0)dxdy = 0$. The initial vorticity is concentrated in a narrow layer centered at $y = 0.5$. It describes a flat shear layer of characteristic width δ. However, the width is perturbed sinusoidally with amplitude ϵ. If δ goes to zero, the initial vorticity approaches to a perturbed flat vortex sheet.

The initial condition for the temperature is chosen to be

$$\theta(x,y,0) = \begin{cases} H_\delta(y-0.25) & y \le 0.4, \\ 1-2H_\delta(y-0.5) & 0.4 < y < 0.6, \\ -H_\delta(0.75-y) & y \ge 0.6, \end{cases} \tag{4.3}$$

where $H_\delta(x)$ is the mollified Heaviside function

$$H_\delta(x) = \begin{cases} 0 & \text{if} \quad x < -\delta, \\ \dfrac{x+\delta}{2\delta} + \dfrac{\sin(\pi x/\delta)}{2\pi} & \text{if} \quad |x| \le \delta, \\ 1 & \text{if} \quad x > \delta. \end{cases} \tag{4.4}$$

The initial temperature is comprised of four smoothly connected layers. The interfaces between the different layers have the same thickness δ as the initial vorticity. To visualize the particle path of the fluid, one of the temperature interfaces is chosen to coincide with the shear layer.

In this paper we choose $\delta = 0.025$, which gives rise to a very sharp shear layer and temperature interfaces. The perturbation to the initial vortex layer is set to $\epsilon = 0.3$. The initial vorticity and temperature are plotted in Figure 4.1. Due to the Kelvin-Helmholtz instability, the vorticity will roll up and suck in the nearby fluid. As a result, the temperature layers will be convected and mixed up by the rotation of the fluid. The roll-up of a thin shear layer is a well-known challenging problem in computational fluid mechanics [62]. With the extra complexity of the Brownian motion forcing, the stochastic problem (4.1) provides a severe test for the WCE method.

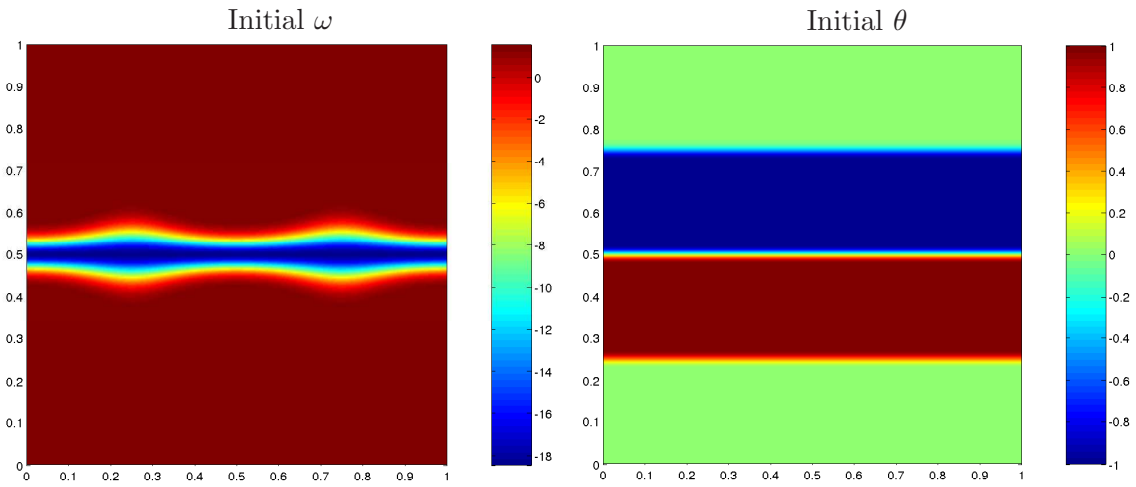

Figure 4.1: Initial vorticity ω and temperature θ

4.1 The WCE method

In this section, we discuss how to solve the stochastic Navier-Stokes equation (4.1) by the WCE method. In the two-dimensional case, the SNS equation (4.1) can be rewritten in the stream function-vorticity formulation. Define the vorticity variable $\omega = v_x - u_y$, and introduce the stream function ψ such that $\vec{u} = (\psi_y, -\psi_x)$, we can rewrite the original system (4.1) as

$$\begin{cases} \theta_t + \vec{u} \cdot \nabla\theta = \mu\Delta\theta, \\ \omega_t + \vec{u} \cdot \nabla\omega = \nu\Delta\omega + (\sigma_2)_x\dot{W}_2 - (\sigma_1)_y\dot{W}_1, \\ -\Delta\psi = \omega, \\ u = \psi_y, \qquad v = -\psi_x. \end{cases} \tag{4.5}$$

We assume that the stream function ψ is also doubly periodic in the domain $[0,1]^2$.

To solve the stochastic system (4.5) by the WCE method, we proceed exactly as we did for the one-dimensional stochastic Burgers equation. We choose an arbitrary orthonormal basis $m_i(t)$, $i = 1, 2, \ldots,$ in $L^2([0,T])$, where $[0,T]$ is the time interval in which we want to solve the problem. Then we project each Brownian motion component $\{W_k(t),\ 0 \le t \le T\}$ onto the L^2 bases:

$$\xi_i^k = \int_0^T m_i(t)dW_k(t), \qquad i = 1, 2, \ldots, \qquad k = 1, 2. \tag{4.6}$$

Let $\alpha = (\alpha_i^k,\ i = 1, 2, \ldots,\ k = 1, 2)$ be a multi-index with nonnegative integer components. Denote the finite multi-index set as

$$\mathcal{J} = \{\alpha,\ |\alpha| = \sum_{i,k} |\alpha_i^k| < \infty\}.$$

For $\alpha \in \mathcal{J}$, define the Wick polynomials by the tensor products

$$T_\alpha(\xi) = \prod_{k=1}^{2} \prod_{i=1}^{\infty} H_{\alpha_i^k}(\xi_i^k). \tag{4.7}$$

Suppose θ, ω, $\vec{u}$ and ψ admit the Wiener Chaos expansions:

$$\theta(x,y,t) = \sum_{\alpha\in\mathcal{J}} \theta_\alpha(x,y,t)\, T_\alpha(\xi), \qquad \omega(x,y,t) = \sum_{\alpha\in\mathcal{J}} \omega_\alpha(x,y,t)\, T_\alpha(\xi),$$

$$\vec{u}(x,y,t) = \sum_{\alpha\in\mathcal{J}} \vec{u}_\alpha(x,y,t)\, T_\alpha(\xi), \qquad \psi(x,y,t) = \sum_{\alpha\in\mathcal{J}} \psi_\alpha(x,y,t)\, T_\alpha(\xi).$$

Similar to 1D stochastic Burgers equation, we can derive the WCE propagator for the system (4.5):

$$\begin{cases} (\theta_\alpha)_t + \mathcal{L}_\alpha(u,\theta)_x + \mathcal{L}_\alpha(v,\theta)_y = \mu\Delta\theta_\alpha, \\ (\omega_\alpha)_t + \mathcal{L}_\alpha(u,\omega)_x + \mathcal{L}_\alpha(v,\omega)_y = \nu\Delta\omega_\alpha \\ \qquad\qquad + (\sigma_2)_x \sum_{j=1}^{\infty} I_{\{\alpha_j^2=1,|\alpha|=1\}} m_j(t) - (\sigma_1)_y \sum_{i=1}^{\infty} I_{\{\alpha_i^1=1,|\alpha|=1\}} m_i(t), \\ -\Delta\psi_\alpha = \omega_\alpha, \\ u_\alpha = (\psi_\alpha)_y, \quad v_\alpha = -(\psi_\alpha)_x. \end{cases} \tag{4.8}$$

In the above formula, the bilinear operator $\mathcal{L}_\alpha(\cdot,\cdot)$ is defined as

$$\mathcal{L}_\alpha(f,g) = \sum_{p\in\mathcal{J}} \sum_{0\le\beta\le\alpha} C(\alpha,\ \beta,\ p)\, f_{\beta+p}\, g_{\alpha-\beta+p} \tag{4.9}$$

for $f = \sum_{\alpha\in\mathcal{J}} f_\alpha\, T_\alpha$ and $g = \sum_{\beta\in\mathcal{J}} g_\beta T_\beta$. The coefficient $C(\alpha,\ \beta,\ p)$ is defined by (2.39).

And $I_{\{\alpha_i^1=1,|\alpha|=1\}}$ is the indicator function

$$I_{\{\alpha_i^1=1,\,|\alpha|=1\}} = \begin{cases} 1 & \alpha_i^1 = 1,\ |\alpha| = 1, \\ \\ 0 & \text{otherwise.} \end{cases}$$

$I_{\{\alpha_i^2=1,|\alpha|=1\}}$ is defined similarly.

Since all the unknowns in (4.5) are doubly periodic in the domain, so are all their WCE coefficients. We solve the WCE propagator (4.8) by pseudo-spectral method in space and the fourth order Adams predictor-corrector scheme [42] in time.

In the following numerical experiment, we set diffusivity as $\mu = 0.0002$, and the viscosity as $\nu = 0.0002$. In the stream function-vorticity formulation (4.5), the spatial parts of the random forcing are taken as

$$(\sigma_1)_y = 0.1\pi \cos(2\pi x)\cos(2\pi y), \qquad (\sigma_2)_x = 0.1\pi \sin(2\pi x)\sin(2\pi y).$$

We solve the problem to $T = 1.0$ by the WCE method. We choose to project each Brownian motion component onto the first four Haar wavelet bases, and truncate the WCE approximation at fourth order Wick polynomials. The resulting system should have 495 coefficients. However, using the sparse WCE truncation, the total number of coefficients can be reduced to 56. We discretize the computational domain by a 256×256 square mesh, and choose $\Delta t = 0.00125$. Then we solve the truncated WCE propagator on the domain $[0,\ 1]^2$ by the Fourier pseudo-spectral method described above. The CPU time is 8.1 hours. We plot the mean and variance of θ and ω in Figure 4.2. Due to the Kelvin-Helmholtz instability, the mean vorticity rolls up and forms a pair of concentrated vortex eyes. The temperature layers are convected and diffused by the random velocity. The sharp edges of the interface are smeared, due to both the diffusivity and the random forcing. The variances of the vorticity and temperature all concentrate along the sharp layers of the interface and are almost zero at the constant regions. In this sense, the random forcing acts as an extra diffusion for the mean equation, even though a closed form of the mean equation is not available.

Another observation is that the third order moment of both the vorticity and temperature are not zero. Actually, they have the same magnitude and similar structure as the respective variances. This shows that the solution of the stochastic Navier-Stokes equation

is far from being a Gaussian process.

To reach similar accuracy as the WCE method, MC simulation needs thousands of realizations, which will take more than two days to compute. For this 2-D stochastic Navier-Stokes equation, it is clear that the WCE method is more efficient than the MC simulation for short time integrations. Since it is not easy to obtain a benchmark solutions by MC simulation or any other method, we will demonstrate the convergence of the WCE method for a special case in the next section.

4.2 Convergence Verification of the WCE Method

To test the convergence of WCE method, we consider a special case of the SNS equation (4.1), where a semi-analytical solution is available. Similar to the 1-D stochastic Burgers equation, we have the following result:

Theorem 4.1 *In system* (4.1)*, if the spatial part* $\sigma(x, y)$ *of the random forcing is a constant matrix, then the stochastic solutions of* (4.1) *can be expressed as*

$$\theta(\vec{x},\, t) = \theta'(\vec{x} - \int_0^t \sigma\, \vec{W}(s)\, ds,\ t), \tag{4.10}$$

$$\vec{u}(\vec{x},\, t) = \vec{u}'(\vec{x} - \int_0^t \sigma\, \vec{W}(s)\, ds,\ t) + \sigma\vec{W}(t), \tag{4.11}$$

$$P(\vec{x},\, t) = P'(\vec{x} - \int_0^t \sigma\, \vec{W}(s)\, ds,\ t),$$

where θ', $\vec{u}'$, P' *are solutions of the corresponding unforced deterministic equations*

$$\begin{cases} \theta'_t + \vec{u}' \cdot \nabla\theta' = \mu\Delta\theta', \\ \vec{u}'_t + \vec{u}' \cdot \nabla\vec{u}' = \nu\Delta\vec{u}' - \nabla P', \\ \nabla \cdot \vec{u}' = 0. \end{cases} \tag{4.12}$$

Theorem 4.1 can be verified easily by direct differentiations. For the random vorticity ω, we can further obtain

$$\omega(\vec{x},\, t) = \omega'\left(\vec{x} - \int_0^t \sigma\, \vec{W}(s)\, ds,\ t\right), \tag{4.13}$$

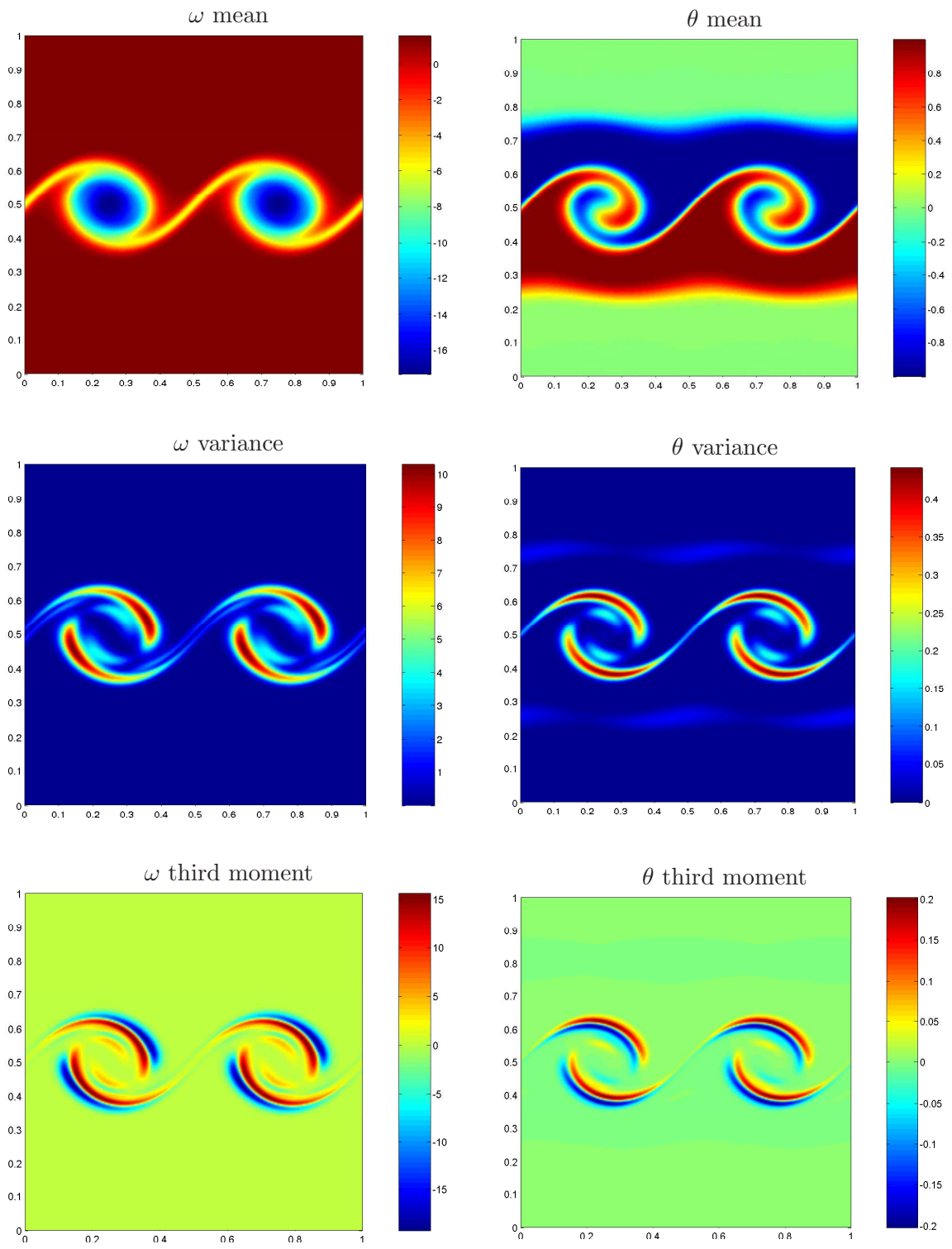

Figure 4.2: Mean, variance, and third order moments of vorticity ω and temperature θ at $T = 1.0$, computed by WCE method with 51 coefficients.

where $\omega' = v'_x - u'_y$ is the deterministic vorticity of the unforced problem (4.12).

The semi-analytical forms (4.10) and (4.13) state that the stochastic solutions are the deterministic solutions of the corresponding unforced problem compounded with a random perturbation along the flow characteristics. That observation can help us better understand the effects of the random forcing. Formula (4.10) and (4.13) also provide a convenient way to compute the statistical moments of θ and ω numerically. Indeed, for any fixed x, t and integer $n = 1, 2, \ldots$, it is easy to show that

$$E\left[\theta(\vec{x},\, t)\right]^n = \int_{\mathbf{R}^2} \left[\theta'(\vec{x} - \vec{z},\, t)\right]^n \rho(\vec{z},\, t) d\vec{z}, \tag{4.14}$$

$$E\left[\omega(\vec{x},\, t)\right]^n = \int_{\mathbf{R}^2} \left[\omega'(\vec{x} - \vec{z},\, t)\right]^n \rho(\vec{z},\, t) d\vec{z}, \tag{4.15}$$

where

$$\rho(\vec{z},\, t) = \frac{3}{2\pi\sigma_1\sigma_2 t^3} \exp\left(-\frac{3z_1^2}{2\sigma_1^2 t^3} - \frac{3z_2^2}{2\sigma_2^2 t^3}\right).$$

To obtain the benchmark statistics of θ and ω, we can first solve the deterministic equation (4.12) by the pseudo-spectral method, then compute the statistical moments from (4.14) and (4.15) by numerical integrations.

For $\vec{z} = (z_1,\, z_2) \in [0,\, 1]^2$, define

$$\rho_d(z_1, z_2,\, t) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \rho(z_1 + n, z_2 + m,\; t). \tag{4.16}$$

Since both θ' and ω' are doubly periodic in $[0, 1]^2$, formula (4.14) and (4.15) can be rewritten as

$$E\left[\theta(\vec{x},\, t)\right]^n = \int_{[0,\,1]^2} \left[\theta'(\vec{x} - \vec{z},\, t)\right]^n \rho_d(\vec{z}, t)\, d\,\vec{z}. \tag{4.17}$$

$$E\left[\omega(\vec{x},\, t)\right]^n = \int_{[0,\,1]^2} \left[\omega'(\vec{x} - \vec{z},\, t)\right]^n \rho_d(\vec{z}, t)\, d\,\vec{z}. \tag{4.18}$$

Because of the symmetry of the Gaussian kernel ρ, ρ_d is doubly periodic in the unit square $[0, 1]^2$. So the integrals (4.17) and (4.18) are convolutions of two doubly periodic functions, which can be easily computed by FFT. Note that ρ_d is defined by an infinite summation, we need to truncate it at finite terms. Since ρ is a Gaussian kernel and decays exponentially, the truncation can be made very accurate. Under this framework, all the computations are

deterministic. No random number generation is needed. By carefully choosing the deterministic algorithms, we can compute the benchmark statistics of θ and ω very accurately.

When the random forcing is spatially independent, we cannot solve the equation (4.1) by its stream function-vorticity formulation anymore. Since the spatial parts of the random forcing are not mean zero in the physical space, they will induce net flows in the domain. So the velocity $\vec{u}$ would not remain mean zero in space at a later time. That is obvious from formula (4.11): there is a random net flow $\sigma\vec{W}(t)$ in the domain at any instance. As a result, the random stream function $\psi = \int udy - vdx$ is no longer doubly periodic in the computational domain. Generally, we don't know how to prescribe its boundary condition consistently. Instead, we need to work on the primitive formula (4.1) and deal with the incompressibility constraint directly.

In the next numerical experiment, we set the temperature diffusivity as $\mu = 0.0005$, and the fluid viscosity as $\nu = 0.0005$. The constants in the random forcing are chosen as $\sigma_1 = 0.02$, $\sigma_2 = 0.02$. Since the random forcing is spatially independent, its effect will not be damped out in time by viscosity. To avoid excessive smearing induced by the random forcing, we deliberately choose σ_1, σ_2 not much larger than the physical viscosity ν.

We first solve the problem to $T = 1.0$ by the semi-analytical method and obtain the benchmark statistics. To test the performance of the WCE method and the MC simulation, we apply both of them to the same problem and compare their numerical solutions with that of the semi-analytical approach. For the WCE method, we project each Brownian motion component onto the first four Haar wavelet bases, and truncate the WCE propagator at fourth order Wick polynomials. Altogether 56 coefficients are retained in the computation. The maximum mean square error for such a truncation is approximately $O(10^{-2})$, based on the error analysis similar to that in Theorem 5.1. We also solve the problem to $T = 1.0$ by MC simulation with 1,000 realizations. For a fair comparison, the same spatial mesh 256×256 and time step $\Delta t = 0.00125$ are used for all three different methods. The numerical

	CPU Hours	$E(\omega)$	$E(\theta)$	$Var(\omega)$	$Var(\theta)$
WCE	9.3	0.028%	0.044%	1.2%	1.7%
MC	44.7	0.23%	0.36%	8.2%	8.8%

Table 4.1: Relative L_2 error on the domain $[0, 1]^2$ of the WCE method with 56 coefficients and the MC simulation with 1,000 realizations

solutions are compared in Table 4.1. The WCE method with 56 coefficients is more accurate than the MC simulation with 1,000 realizations for all the statistical moments computed. In addition, the WCE method is also several times faster than the MC simulation. To reach the same accuracy as the WCE method, the MC simulation may need much more than 1,000 realizations, which will take even more CPU time. However, the WCE method requires more computer memory than the MC simulation since a coupled PDE system for the WCE coefficients has to be solved simultaneously. For the 1-D and 2-D problems considered in this paper, the memory usage is not a big issue and all the computation can be done on a standard PC. For 3-D stochastic Navier-Stokes equations, we expect that the memory usage will become a bigger concern to the WCE method when many WCE coefficients have to be retained.

Chapter 5

Error Analysis of the WCE Method

Due to the nonlinear coupling in the WCE propagator, it is very difficult to conduct an error analysis of the truncated WCE method in the general setting. In this section, we only establish an error bound for the WCE approximation when the random forcing is independent of the spatial variables. In this case the solutions of the stochastic Navier-Stokes and Burgers equations could be obtained from the respective deterministic counterparts by a stochastic Galilean transformation. Though a special case, the error bound in this particular setting sheds light on the general performance of the WCE method. For simplicity, we will prove the result in the context of the stochastic Burgers equation. For the stochastic Navier-Stokes equation, the proof and results are very similar.

5.1 Convergence Analysis for SPDEs with Additive Forcing

Consider the stochastic Burgers equation

$$\begin{cases} u_t + \frac{1}{2}(u^2)_x = \mu u_{xx} + \sigma(x)\dot{W}(t), \\ u(x,\,0) = u_0(x), \qquad u(0,t) = u(1,t) \end{cases} \qquad (t,x) \in (0,T] \times [0,1], \tag{5.1}$$

where σ is a constant. According to Theorem 3.1, the solution of (5.1) can be expressed as

$$u(x,\,t) = v\left(x - \sigma \int_0^t W(s)ds,\; t\right) + \sigma W(t), \tag{5.2}$$

where $v(x,\, t)$ is the solution of the unforced deterministic Burgers equation

$$\begin{cases} v_t + \frac{1}{2}(v^2)_x = \mu v_{xx}, \\ \\ v(x,0) = u_0(x),\ v(0,t) = v(1,t). \end{cases} \tag{5.3}$$

Suppose the basis functions $m_k(s)$ in the space $L^2([0,t])$ are chosen as the cosine functions (2.29). Denote the truncated index set as

$$\mathcal{J}_{K,N} = \{(\alpha_1, \ldots, \alpha_K);\ |\alpha| \le N\}. \tag{5.4}$$

That is, we keep K number of Gaussian random variables and maximum Nth order Wick polynomials in the truncated WCE. The main result of this chapter is as follows:

Theorem 5.1 (Simple Truncation) *Denote the truncated WCE of the solution of* (5.1) *as*

$$u_{K,N}(x,t) = \sum_{\alpha \in \mathcal{J}_{K,N}} u_\alpha(x,t) T_\alpha, \qquad \textit{where} \qquad u_\alpha(x,t) = E[u(x,t)T_\alpha]. \tag{5.5}$$

Then, the error estimate holds:

$$\max_x \sqrt{E\big|u(x,t) - u_{K,N}(x,t)\big|^2} \le B_{N+1} \frac{\left(\sigma t^{3/2}\right)^{N+1}}{\sqrt{(N+1)!}} + B_1 \sigma \left(\frac{t}{K}\right)^{3/2}, \tag{5.6}$$

where $B_n = \sup_x \left|\frac{\partial^n}{\partial x^n} v(x,t)\right|$ *and* v *is the solution of the deterministic equation* (5.3).

Proof Denote

$$Z(t) = \int_0^t W(s) ds.$$

Due to the expansion (2.30), the Brownian motion process $\{W(s);\ 0 \le s \le t\}$ admits the following expansion:

$$W(s) = \frac{s}{\sqrt{t}}\xi_1 + \sum_{k=2}^{\infty} \xi_k \frac{\sqrt{2t}}{(k-1)\pi} \sin\Big(\frac{(k-1)\pi s}{t}\Big). \tag{5.7}$$

Consequently, we have

$$Z(t) = \frac{\sqrt{t^3}}{2}\xi_1 + \sum_{k=2}^{\infty} a_k \frac{\sqrt{t^3}}{(k-1)^2}\xi_k, \tag{5.8}$$

where $a_k = \frac{\sqrt{2}}{\pi^2}\left[1+(-1)^k\right]$. Denote $Z = Z_K + Z_R$, where

$$Z_K = \frac{\sqrt{t^3}}{2}\xi_1 + \sum_{k=2}^{K} a_k \frac{\sqrt{t^3}}{(k-1)^2}\xi_k,$$

$$Z_R = \sum_{k=K+1}^{\infty} a_k \frac{\sqrt{t^3}}{(k-1)^2}\xi_k, \tag{5.9}$$

Obviously Z_K and Z_R are orthogonal in that $E(Z_K Z_R) = 0$.

Note that $W(t) = \sqrt{t}\,\xi_1$ and the solution (5.2) can be rewritten as

$$u(x,t) = v(x - \sigma Z_K - \sigma Z_R,\ t) + \sigma\sqrt{t}\,\xi_1. \tag{5.10}$$

Expanding $v(x - \sigma Z_K - \sigma Z_R,\ t)$ in Taylor's series, firstly with respect to Z_R and then with respect to Z_K, we arrive at

$$\begin{aligned} u(x,t) &= v(x - \sigma Z_K, t) + I_1 + \sigma\sqrt{t}\,\xi_1 \\ &= v(x,t) + \sum_{n=1}^{N} \frac{(-\sigma Z_K)^n}{n!}\frac{\partial^n v}{\partial x^n}(x,t) + I_2 + I_1 + \sigma\sqrt{t}\,\xi_1, \end{aligned}$$

where

$$I_1(x,t) = -\frac{\partial v}{\partial x}(x - \sigma Z_K - \theta_1, t)\sigma Z_R,$$

$$I_2(x,t) = \frac{(-\sigma Z_K)^{N+1}}{(N+1)!}\frac{\partial^{N+1} v}{\partial x^{N+1}}(x - \theta_2, t)$$

are the Lagrange residuals of the Taylor expansions.

Denote

$$\tilde{u}_{K,N}(x,\ t) = v(x,t) + \sigma\sqrt{t}\,\xi_1 + \sum_{n=1}^{N} \frac{(-\sigma Z_K)^n}{n!}\frac{\partial^n v}{\partial x^n}(x,t). \tag{5.11}$$

Obviously, $\tilde{u}_{K,N}(x,t)$ is a polynomial of $\xi_1, \ldots, \xi_K$ with maximum order N. Since the WCE truncation $u_{K,N}(x,t)$ is a Hermite polynomial expansion, which is an orthogonal (optimal)

projection with respect to the Gaussian measure, it follows that

$$\sqrt{E|u(x,t)-u_{K,N}(x,t)|^2} \le \sqrt{E|u(x,t)-\tilde{u}_{K,N}(x,t)|^2}$$
$$\le B_1\sigma\sqrt{E|Z_R^2|} + B_{N+1}\frac{\sigma^{N+1}}{(N+1)!}\sqrt{E|Z_K^{2N+2}|}, \tag{5.12}$$

where $B_n = \sup_x \left|\frac{\partial^n}{\partial x^n}v(x,t)\right|$. From formula (5.9) we have

$$E|Z_R^2| \le \frac{8t^3}{\pi^4}\sum_{k=K+1}^{\infty}\frac{1}{(k-1)^4} < \frac{8t^3}{\pi^4}\Big(\frac{1}{K^4}+\int_K^{\infty}\frac{dx}{x^4}\Big) < \frac{t^3}{K^3}. \tag{5.13}$$

Note that Z_K is a linear combination of centered Gaussian random variables, so it is a centered Gaussian itself. Due to properties of Gaussian random variables, we have

$$E|Z_K^{2N+2}| = (E|Z_K^2|)^{N+1}(2N+1)!! < (E|Z^2|)^{N+1}(2N+1)!!, \tag{5.14}$$

where $(2N+1)!! = (2N+1)(2N-1)\ldots 1$. It is easy to show that $E|Z^2| = t^3/3$. Substituting (5.13) and (5.14) into (5.12), we obtain

$$\sqrt{E|u(x,t)-u_{K,N}(x,t)|^2} \le B_{N+1}\frac{\sqrt{(2N+1)!!}}{(N+1)!}\sigma^{N+1}\left(\frac{t^3}{3}\right)^{\frac{N+1}{2}} + B_1\sigma\left(\frac{t}{K}\right)^{3/2}$$
$$\le B_{N+1}\frac{\left(\sigma t^{3/2}\right)^{N+1}}{\sqrt{(N+1)!}} + B_1\sigma\left(\frac{t}{K}\right)^{3/2}.$$

Since the right-hand side does not depend on x, taking the maximum in x recovers the estimate (5.6). This completes the proof. □

Remark The first term in the error estimate (5.6) reflects the error of truncating the polynomial order in the Fourier-Hermite expansion of the solution. Making use of the Stirling formula, one can easily check that the asymptotic of this term is given by

$$\frac{\left(\sigma t^{3/2}\right)^{N+1}}{\sqrt{(N+1)!}} = O\Big(\frac{\sigma^2 t^3 e}{N+1}\Big)^{\frac{N+1}{2}}.$$

This confirms our numerical observation that the WCE approximation converges exponentially in terms of the Wick polynomial order N.

The second term in (5.6) stems from the truncation error of the Fourier expansion (5.7).

As we have shown in Theorem 2.1, the Fourier expansion of $W(s)$ converges at the rate

$$E\left[W(s)-\sum_{k=1}^{K}\xi_k\int_0^s m_k(\tau)d\tau ds\right]^2<C\frac{t}{K}.$$

The slow convergence of the second term is a reflection of the irregularity of the Brownian motion forcing.

Both the numerical experiments and Theorem 5.1 indicate that the convergence rate of the WCE method critically depends on the time span $[0,t]$ and the magnitude of the random forcing σ. Clearly, the smaller the time interval $[0,t]$, the faster the convergence. Note that Brownian motion has independent increments. When long time simulations are of interest, one can divide a large time interval into small ones and apply the WCE method repeatedly on those small intervals. If the time subinterval is of size Δt, then by Theorem 5.1 the error of approximation for each step is $O\big(\Delta t\big)^{\frac{3}{2}}$. Hence, the global rate of convergence should be $O\big(\sqrt{\Delta t}\big)$.

A smaller forcing would increase the rate of convergence of the WCE method. In a broad class of SPDEs arising in physics and engineering, the large structures and dominant dynamic effects are captured by the deterministic physical laws, while the unresolved small scales, microscopic effects, and other uncertainties are often modeled as stochastic perturbations with small amplitudes. We expect that the WCE method will be especially effective in applications to this class of problems. In Section 6.4, we will use the WCE method to study the front propagation in a reaction-diffusion equations with the Kolmogorov-Petrovsky-Piskunov (KPP) nonlinearity, where the random forcing is naturally formulated as a small perturbation. The numerical results show that the WCE method is quite effective for this problem.

The proof of Theorem 5.1 is based on the explicit semi-analytical solutions (5.2), which is only true in the case of spatially independent random forcing. In general, the solution of an SPDE is a complicated functional of the Brownian motion path $u=u(x,t;\ W_0^t)$. It does not have an explicit formula as simple as (5.2). However, as we explained in Section 2.3, we can always interpret the stochastic solution $u(x,t;\ W_0^t)$ as a general function of the Gaussian random variables:

$$u(x,t;\ W_0^t)=U(x,t;\ \xi_1,\xi_2,\dots,), \tag{5.15}$$

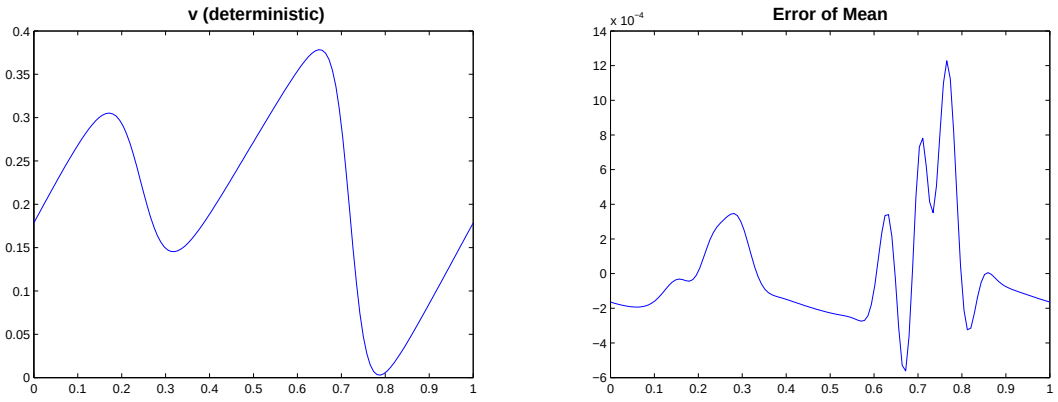

Figure 5.1: The deterministic solution v and the error of the mean computed by WCE methods. The error of WCE solution is indeed much larger near the sharp layers (smoothed shocks) of the deterministic solution v.

where U is an unknown ghost function. Then we can still prove Theorem 5.1 formally based on the unknown function U. In that sense, the error analysis presented in Theorem 5.1 is quite general and has captured the essence of the convergence behavior of the WCE method. In numerical computations, we can still use the special error estimate (5.6) as a guide in choosing the truncation parameters K and N. Numerical results in Chapter 3 and Chapter 4 have shown that truncation decisions based on the error estimate (5.6) work well even in general cases.

Based on Theorem 5.1, the convergence constant of the WCE method depends on the smoothness of the deterministic solution v. Consequently, the error of the WCE method would be larger at the sharp gradient regions of v. Figure 5.1 shows the deterministic solution v and the error of the mean computed by the WCE method. The error of the WCE solution is indeed much larger near the sharp layers of v. This indicates that at large gradient regions more WCE coefficients are needed to achieve the same accuracy.

In the numerical experiments, we have observed that the simple truncation (5.4) is not optimal, and we should use an adaptive truncation strategy to reduce the dimension of the WCE propagator. The following theorem is the rigorous justification for such a consideration.

Theorem 5.2 *For the solution of the stochastic Burgers equation* (5.1)*, its WCE coeffi-*

cients decay asymptotically as

$$u_\alpha = O\left[\frac{1}{\sqrt{\alpha!}}\prod_{k=2}^{\infty}\left(\frac{1}{k-1}\right)^{2\alpha_k}\right]. \tag{5.16}$$

Proof Denote the semi-analytical solution (5.2) as

$$u = V(\xi_1, \xi_2, \dots,). \tag{5.17}$$

Here we have suppressed the variable (x,t) in the expression. From expansion (5.8), the function $V(\xi)$ has the form

$$V(\xi) = v\left(x - \frac{\sigma t^{3/2}}{2}\xi_1 - \sum_{k=2}^{\infty} a_k \frac{\sigma t^{3/2}}{(k-1)^2}\xi_k,\ t\right) + \sigma\sqrt{t}\xi_1.$$

Based on Lemma 2.4 we have

$$\begin{aligned} u_\alpha &= E[u\,T_\alpha(\xi)] = \frac{(-1)^{|\alpha|}}{\sqrt{\alpha!}} E\left[D^\alpha V(\xi)\right] \\ &= \frac{C}{\sqrt{\alpha!}} E\left[\frac{d^{|\alpha|}}{dx^{|\alpha|}} v(x - \sigma Z,\ t)\right] \prod_{k=2}^{\infty}\left(\frac{1}{(k-1)^2}\right)^{\alpha_k} \\ &= O\left[\frac{1}{\sqrt{\alpha!}}\prod_{k=2}^{\infty}\left(\frac{1}{k-1}\right)^{2\alpha_k}\right]. \end{aligned}$$

□

Theorem 5.2 reveals the asymptotic decaying rate of the WCE coefficients u_α in terms of the polynomial order index α. As we emphasized before, the Wiener chaos expansion decays both in the order of the Wick polynomials and the number of retained Gaussian variables. In the asymptotical formula (5.16), the factor $\frac{1}{\sqrt{\alpha!}}$ reflects the decaying rate induced by the polynomial expansion, and the factor $\prod_{k=2}^{\infty}\left(\frac{1}{k-1}\right)^{2\alpha_k}$ reflects the decaying rate by the Gaussian expansion of the Brownian motion. For WCE coefficients u_α with the same order $|\alpha|$, the ones depending on ξ_k with larger subscript k are smaller in magnitude. For example, in the truncation (5.4) the WCE coefficient with $\alpha = (N, 0, \dots, 0)$ has a magnitude of $O\left(\frac{1}{\sqrt{N!}}\right)$, while the coefficient with $\alpha = (0, \dots, 0, N)$ has a magnitude of $O\left(\frac{1}{\sqrt{N!}(K-1)^{2N}}\right)$. So the coefficient $u_{(0,\dots,0,N)}$ is much smaller than the coefficients $u_{(N,0,\dots,0)}$. Instead of using

the same maximum order of Hermite polynomials for every ξ_k, it makes more sense to use lower order Hermite polynomial for ξ_k with larger subscript k. For this purpose, we define an adaptive index

$$r = (r_1, r_2, \ldots, r_K), \qquad \text{with} \quad r_K \leq \ldots \leq r_2 \leq r_1 = N. \tag{5.18}$$

In addition to the simple truncation (5.4), we keep only those Wick polynomials $T_\alpha(\xi)$ such that $\alpha_i \leq r_i$. More precisely, we truncated the WCE expansion by the new index set.

$$\mathcal{J}^r_{K,N} = \{(\alpha_1, \ldots, \alpha_K);\ |\alpha| \leq N,\ \alpha_i \leq r_i\}. \tag{5.19}$$

To determine the optimal adaptive index $r = (r_1, r_2, \ldots, r_K)$, consider the single-mode index

$$e_k = (0, \ldots, 0, r_k, 0, \ldots, 0).$$

We choose r_k so that the all the coefficients $u_{e_k},\ k = 1, 2, \ldots, K$ have similar magnitudes. From Theorem 5.2, the coefficient u_{e_k} has the magnitude

$$u_{e_k} \simeq \frac{1}{\sqrt{r_k!}} \left(\frac{1}{k-1}\right)^{2r_k}.$$

Note that $r_1 = N$ and we always have $u_{e_1} \simeq 1/\sqrt{N!}$. So the r_k is determined by the balance condition $u_{e_k} \simeq u_{e_1}$. Consequently, we have

$$r_1 = N, \qquad \frac{1}{\sqrt{r_k!}} \left(\frac{1}{k-1}\right)^{2r_k} \simeq \frac{1}{\sqrt{N!}}. \tag{5.20}$$

By balancing the magnitudes of the smallest coefficients, we can effectively exclude unimportant coefficients and hence reduce the dimension of the WCE truncations. To estimate the optimal r_k, we just need to solve the equation (5.20) numerically. Since the WCE decays faster in terms of N than in K, in practice we always choose $K > N$ so that the two terms in the error bound (5.6) will balance each other. In real computations, the total order N of the Wick polynomials rarely goes beyond 10. Under this practical assumption, a nearly optimal adaptive index can be set as

$$r = (N,\ N,\ N-1, \ldots, 2,\ 1, \ldots, 1). \tag{5.21}$$

Numerical results in Chapter 3 and Chapter 4 has shown that the simple adaptive index (5.21) is usually suffice in most numerical computations.

In the above adaptive truncations, we excluded all the WCE coefficients with magnitudes less than $O(1/\sqrt{N!})$, which has the same order as the first term in the error estimate (5.6). The extra error introduced by such an adaptive truncation will be no more than $O(1/\sqrt{N!})$. So the adaptive truncation will not change the asymptotics of the first term in the error estimate (5.6). Note that the adaptive truncation deals with how to choose the optimal maximum polynomial order for each Gaussian random variable ξ_k, it does not change the total number of retained Gaussian random variables. Since the second term in the error estimate (5.6) does note depend on the polynomial order, it would not be affected by the adaptive truncation. So the adaptive truncation will not change the asymptotical behavior of the error bound (5.6) for the simple truncation. Based on the above arguments, we have the following conclusion:

Theorem 5.3 (Sparse Truncation) *Denote the sparse WCE solution of* (5.1) *as*

$$u^r_{K,N}(x,t) = \sum_{\alpha \in \mathcal{J}^r_{K,N}} u_\alpha(x,\, t)\, T_\alpha(\xi), \tag{5.22}$$

where $\mathcal{J}^r_{K,N}$ *is defined by* (5.19)*, and* $r = (r_1, r_2, \ldots, r_K)$ *is determined by the balance condition* (5.20)*. Then the sparse WCE truncation* $u^r_{K,N}$ *has the same asymptotic accuracy as the simple truncation* $u_{K,N}$ *defined by* (5.5)*.*

Using the sparse truncation, we can reduce the dimension of the WCE propagator dramatically while keep the same asymptotic convergence rate. For the simple truncation $K = 8$ and $N = 5$, the total number of WCE coefficients would be 1,287. So the resulting WCE propagator will have 1,287 equations, which is almost impossible to solve numerically. However, using the sparse truncation, the total number of WCE coefficients can be reduced to 51. That is a very substantial dimension reduction and the resulting WCE propagator can be solved easily. The numerical results in Chapter 3 and Chapter 4 have shown that computations based on sparse truncations indeed can achieve the same accuracy as the simple truncations.

5.2 Convergence Analysis for SPDEs with Multiplicative Forcing

In this section, we will discuss the convergence properties of WCE methods for solving multiplicative SPDEs. As in the previous section, we will focus on stochastic Burgers equations only. Since the analysis and results are very similar, we only provide the final results without proof. The motivation of the analysis is to help understand the performance of the WCE method and provide guidance for numerical computations.

Consider the multiplicative stochastic Burgers equation

$$\begin{cases} u_t + \left(u + \sigma\dot{W}\right) \circ u_x = \mu u_{xx}, \\ u(x,0) = u_0(x), \quad u(0,t) = u(1,t), \end{cases} \tag{5.23}$$

where σ is a constant. From Theorem 3.2, the solution of (5.23) is given by

$$u(x,\ t) = v\big(x - \sigma W(t),\quad t\big), \tag{5.24}$$

where $v(x,\ t)$ is the solution of the unforced deterministic Burgers equation

$$\begin{cases} v_t + \frac{1}{2}(v^2)_x = \mu v_{xx}, \\ v(x,0) = u_0(x), \quad v(0,t) = v(1,t). \end{cases} \tag{5.25}$$

Similar to the additive stochastic Burgers equation, we have the following result:

Theorem 5.4 (Simple Truncation) *Denote the truncated WCE solution of* (5.23) *as*

$$u_{K,N}(x,t) = \sum_{\alpha\in\mathcal{J}_{K,N}} u_\alpha(x,t)T_\alpha, \qquad u_\alpha(x,t) = E[u(x,t)T_\alpha], \tag{5.26}$$

where the index set $\mathcal{J}_{K,N}$ is defined by (5.4)*. Then, the error estimate holds*

$$\max_x \sqrt{E\big|u(x,t) - u_{K,N}(x,t)\big|^2} \le B_{N+1}\frac{\left(\sigma t^{1/2}\right)^{N+1}}{\sqrt{(N+1)!}} + B_1\sigma\left(\frac{t}{K}\right)^{1/2}, \tag{5.27}$$

where $B_n = \sup_x \left|\frac{\partial^n}{\partial x^n}v(x,t)\right|$ and v is the solution of the deterministic equation (5.25)*.*

The proof of Theorem (5.4) is very similar to that of Theorem (5.1). Comparing the

error estimate 5.27 with 5.6, we notice that the WCE decays slower for multiplicative SPDE than for additive SPDE in terms of the number of retained Gaussian random variables. This explains why multiplicative SPDEs are more difficult to solve than additive SPDEs. In this particular case, the semi-analytical solutions (5.2) for the additive SPDE involves an integral of the Brownian motion, while the semi-analytical solution (5.24) for the multiplicative SPDE only involves the Brownian motion itself. As functionals of the Brownian motion, solution (5.2) has better regularity than solution (5.24) and hence decays faster in terms of the Gaussian expansion of the Brownian motion.

Based on the semi-analytical formula (5.24) for the random solution, we can prove a similar asymptotic formula for its WCE coefficients.

Theorem 5.5 *For the solution of the stochastic Burgers equation* (5.23)*, its WCE coefficients decay asymptotically as*

$$u_\alpha = O\left[\frac{1}{\sqrt{\alpha!}}\prod_{k=2}^{\infty}\left(\frac{1}{k-1}\right)^{\alpha_k}\right]. \tag{5.28}$$

In the asymptotical formula (5.28), the factor $\frac{1}{\sqrt{\alpha!}}$ reflects the decaying rate induced by the polynomial expansion, while the factor $\prod_{k=2}^{\infty}\left(\frac{1}{k-1}\right)^{\alpha_k}$ reflects the decaying rate induced by the Gaussian expansion of the Brownian motion. Comparing formula (5.28) with formula (5.16), it becomes very clear that the WCE coefficients decay slower for the multiplicative SPDE than for the additive SPDE.

As in the case for solving additive SPDEs, we should truncate the polynomial orders for the Wiener chaos expansion adaptively. Instead of using the simple truncation (5.26), we should design a sparse truncation based on the asymptotic decaying rate (5.28) of the WCE coefficients. As before, we set an optimal adaptive index $r = (r_1, r_2, \ldots, r_K)$, where r_k is determined by the balance condition between the WCE coefficients:

$$r_1 = N, \qquad \frac{1}{\sqrt{r_k!}}\left(\frac{1}{k-1}\right)^{r_k} \simeq \frac{1}{\sqrt{N!}}. \tag{5.29}$$

Using the optimal adaptive index r, we can construct a sparse truncation index set $\mathcal{J}^r_{K,N}$ as in (5.19). Based on the same reasons as in additive SPDEs, such constructed sparse truncation will not affect the asymptotic convergence rate of the simple truncation. So we have the following conclusion:

Theorem 5.6 (Sparse Truncation) *Denote the sparse WCE solution of* (5.1) *as*

$$u^r_{K,N}(x,t) = \sum_{\alpha \in \mathcal{J}^r_{K,N}} u_\alpha(x,\, t)\, T_\alpha(\xi), \tag{5.30}$$

where the adaptive index $r = (r_1, r_2, \ldots, r_K)$ *is determined by the balance condition* (5.29)*. Then the sparse WCE truncation* $u^r_{K,N}$ *has the same asymptotic accuracy as the simple truncation* $u_{K,N}$ *defined by* (5.26)*.*

Chapter 6

Long Time Integrations of the WCE Method

The WCE method converges very quickly for short time intervals and small random forcing. However, for long time computations, we need more Gaussian random variables to represent the Brownian motion forcing, and higher order polynomials to captures the nonlinear dynamics of the equation. As a result, the number of WCE coefficients will increase very quickly as the integration time increases. For a WCE truncation with K Gaussian random variables and maximum Nth order polynomials, the total number of Wiener chaos coefficients would be

$$I(K,N) = \binom{K+N}{K} = \frac{(K+N)!}{K!N!}.$$

The number increase very fast as K and N increase. Using the sparse truncation techniques, we can reduce the number of WCE coefficients dramatically without losing much accuracy. Still, it is expected that the direct and naive WCE method will be less efficient for long time integrations.

Similarly, if the magnitude of the random forcing $\sigma(x,t)$ is large, then the WCE may converge much slower. More terms need to be included in the WCE propagator to achieve similar accuracy. The difficulty with large random forcing is equivalent to that with long time integration, since we can always rescale the time variable t to transform one case into the other. Empirically, the L^2 norm of the random forcing

$$E\left[\int_0^T \sigma(x,t)dW(t)\right]^2 = \int_0^T \sigma^2(x,t)dt$$

measures the strength of the random input. Larger T or σ means stronger random input.

Consequently, more WCE bases are required to resolve the randomness in the stochastic solution. To avoid the explosion of dimensions in long time integrations, we need to design a new computational strategy for the WCE method.

In this chapter, we apply the WCE method to a stochastic transport equation, and use it to illustrate the convergence behavior of the WCE method for long time integration. Motivated by the numerical find-outs, we introduce a new computational strategy for long time integration using WCE method. The basic observation is that a sparse WCE solution can capture the coarse-scale variability of the random solution in the probability space, even though it may not be very accurate in small scales. We can first solve the SPDE by a sparse WCE truncation and obtain a coarse-scale WCE solution in the probability space. Then we correct the WCE solution by a small number of MC simulations. To do that, we use the coarse-scale WCE solution as a control variate and subtract it from each MC realization. The coarse-scale WCE solution can reduce the variance of the MC simulation substantially and accelerate the MC convergence by tens to hundreds of times. The combined WCE-MC method is shown to be more efficient than the WCE method or the MC simulation alone in long time integrations. Then we apply the WCE-MC hybrid method to simulate a stochastic reaction-diffusion equation with the Kolmogorov-Petrovsky-Piskunov (KPP) type nonlinearity. The numerical results confirm the quadratic enhancing law of the front propagation speed, as conjectured by Jack Xin in a recent work [76].

6.1 Long Time Integrations of Stochastic Transport Equations

In this section we apply the WCE method to solve a stochastic transport equation on the unit square $\Omega = [0,1]^2$

$$\theta_t + (\mathbf{V} \circ \nabla)\,\theta = \mu \Delta \theta + f, \tag{6.1}$$

where θ is a temperature distribution (or the density of other physical quantities), $\mathbf{V}$ is a given stochastic velocity field, and f is a deterministic forcing. As before, the stochastic integral $(\mathbf{V} \circ \nabla)\,\theta$ is defined in Stratonovich's sense. The linear transport equation (6.1) is widely studied as a prototype model in understanding the anomalous diffusion in turbulence [54, 55, 28, 29]. In the current study, we only use the linear transport equation (6.1) as a

test model for the WCE method, and motivate how to improve the WCE method in long time integrations.

We assume that θ is periodic on the unit square Ω, and consider the following random velocity field

$$\mathbf{V} = (u\dot{W},\ v\dot{W}),$$

where $W(t)$ is a Brownian motion. Hence the transport equation (6.1) can be reformulated as

$$\theta_t + [u\theta_x + v\theta_y] \circ \dot{W} = \mu\Delta\theta + f. \tag{6.2}$$

Denote the spatial part of the velocity field as $\mathbf{b} = \begin{bmatrix} u \\ v \end{bmatrix}$. We assume that $\mathbf{b}$ is a "cellular flow" generated by the stream function

$$\psi(x,y) = A\sin\left(\frac{2\pi x}{\delta}\right)\sin\left(\frac{2\pi y}{\delta}\right),$$

where δ is the characteristic length scale of the flow, and A is a normalization factor. Since

$$\mathbf{b} = \nabla^{\perp}\psi = \begin{bmatrix} -\psi_y \\ \psi_x \end{bmatrix},$$

we have

$$\begin{aligned} u &= -\frac{2\pi A}{\delta}\sin\left(\frac{2\pi x}{\delta}\right)\cos\left(\frac{2\pi y}{\delta}\right), \\ v &= \frac{2\pi A}{\delta}\cos\left(\frac{2\pi x}{\delta}\right)\sin\left(\frac{2\pi y}{\delta}\right). \end{aligned}$$

The level sets of the stream function (streamlines) with $\delta = 0.25$ are plotted in Figure 6.1. The velocity vector is tangent to the streamlines. Obviously, the random velocity field $\mathbf{V}$ is incompressible at any time instance:

$$\nabla\cdot\mathbf{V} = (\nabla\cdot\mathbf{b})\,\dot{W} = \left(\nabla\cdot\nabla^{\perp}\psi\right)\dot{W} = 0.$$

Since the random velocity $\mathbf{V}$ is white-in-time, it flips its directions (the sign) very quickly in time. Such a random velocity field can be regarded as an extreme case of temporally periodic flows with infinitesimal periods.

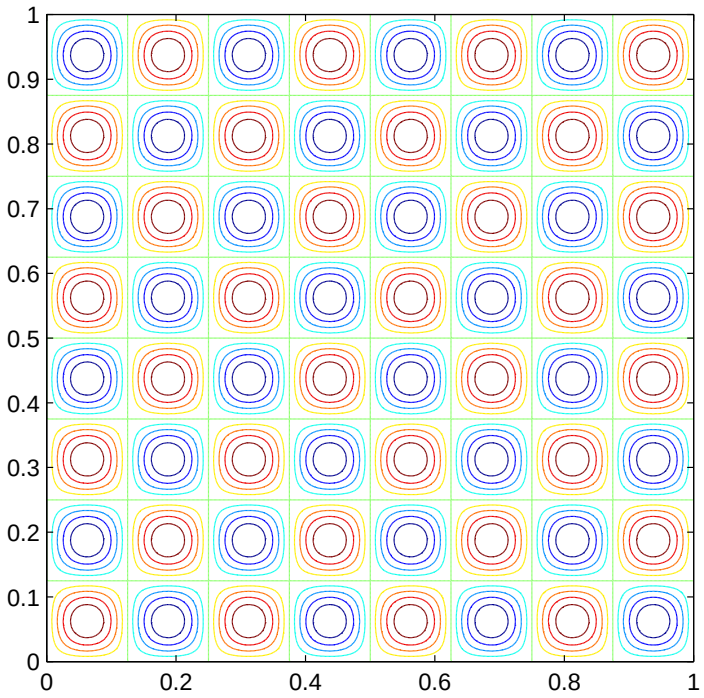

Figure 6.1: The stream function of the "cellular flow" with $\delta = 0.25$.

The stochastic differential equation (6.2) can be rewritten in Ito's sense

$$\theta_t + [(\mathbf{b} \cdot \nabla\theta)] \cdot \dot{W} = \nabla \cdot (\mu\mathbf{I} + \mathbf{D})\nabla\theta + f, \tag{6.3}$$

where

$$\mathbf{D} = \frac{1}{2}\,\mathbf{b} \cdot \mathbf{b}^T = \frac{1}{2}\begin{pmatrix} u^2 & uv \\ vu & v^2 \end{pmatrix}. \tag{6.4}$$

The tensor $\mathbf{D}$ is the convection-enhanced diffusivity tensor, which is only semi-positive definite. Suppose the random solution of (6.3) has the Wiener chaos expansion:

$$\theta(x, y, t) = \sum_{\alpha \in \mathcal{J}} \theta_\alpha(x, y, t) T_\alpha(\xi), \qquad 0 \leqslant t \leqslant T.$$

Similar to the stochastic Burgers equation with multiplicative forcing (see Section 3.2.1),

we can derive the WCE propagator for equation (6.3):

$$(\theta_\alpha)_t + \sum_i m_i(t)\sqrt{\alpha_i}\left[\mathbf{b}\cdot\nabla\theta_{\alpha^-_{(i)}}\right] = \nabla\cdot(\mu\mathbf{I}+\mathbf{D})\nabla\theta_\alpha + fI_{\{\alpha=0\}}, \tag{6.5}$$

where the index $\alpha^-_{(i)}$ is defined by (3.30). In the WCE propagator (6.5), the higher order WCE coefficients depend on the lower order ones, but not vice versa. In particular, the first WCE coefficient $\theta_{\mathbf{0}}$, which is the mean, does not depend on the other WCE coefficients:

$$(\theta_{\mathbf{0}})_t = \nabla\cdot(\mu\mathbf{I}+\mathbf{D})\nabla\theta_{\mathbf{0}} + f. \tag{6.6}$$

So the first equation in the WCE propagator is a closed moment equation for the mean. That is a unique property for the linear transport equation. However, there is no closed moment equation for other order moments.

For highly oscillatory cellular flows (characterized by the small parameter δ), the velocity field $\mathbf{b}$ is very strong, and so is the convection-induced diffusivity. When the diffusivity tensor $\mathbf{D}$ is large in magnitude, the diffusion term $\nabla\cdot\mathbf{D}\nabla\theta_\alpha$ poses serious difficulty for solving the WCE propagator (6.5) numerically. If we discretize the time derivative by explicit schemes such as

$$\frac{\theta^{n+1}_\alpha - \theta^n_\alpha}{k} = \nabla\cdot(\mu\mathbf{I}+\mathbf{D})\nabla\theta^n_\alpha + fI_{\{\alpha=0\}} - \sum_i m_i(t)\sqrt{\alpha_i}\left[\mathbf{b}\cdot\nabla\theta^n_{\alpha^-_{(i)}}\right],$$

and the spatial derivative by central differences, then the CFL condition imposed by the diffusion term would be

$$k = O(h^2), \tag{6.7}$$

where k is the time step and h is the spatial grid size. This constraint on the time step size is very severe for small spatial grid h. To avoid over-small time step sizes, implicit discretizations in time would be necessary. A nature choice is to use explicit schemes for the convection term while implicit schemes for the diffusion term, such as the following one [2]:

$$\frac{\theta^{n+1}_\alpha - \theta^n_\alpha}{k} = \nabla\cdot(\mu\mathbf{I}+\mathbf{D})\nabla\theta^{n+1}_\alpha + fI_{\{\alpha=0\}} - \sum_i m_i(t)\sqrt{\alpha_i}\left[\mathbf{b}\cdot\nabla\theta^n_{\alpha^-_{(i)}}\right].$$

The above time evolution scheme requires solving a potentially ill-conditioned linear system

at every time step

$$[1 - k\nabla \cdot (\mu \mathbf{I} + \mathbf{D})\nabla]\ \theta_\alpha^{n+1} = \theta_\alpha^n + kfI_{\{\alpha=0\}} - k\sum_i m_i(t)\sqrt{\alpha_i}\left[\mathbf{b}\cdot\nabla\theta^n_{\alpha^-_{(i)}}\right].$$

The popular iterative solvers such as GMRES or multi-grid method are not very efficient if μ is relatively small comparing to the magnitude of $\mathbf{D}$.

To avoid the above difficulty, we can add and subtract an artificial diffusion term $\lambda\Delta\theta_\alpha$ in the numerical scheme, where λ is a constant. Then we use the implicit scheme for one term and the explicit scheme for the other. More specifically, we discretize the WCE propagator using the following splitting scheme:

$$\frac{\theta_\alpha^{n+1} - \theta_\alpha^n}{k} = \nabla\cdot(\mu\mathbf{I} + \mathbf{D} - \lambda\mathbf{I})\nabla\theta_\alpha^n + \lambda\Delta\theta_\alpha^{n+1} + fI_{\{\alpha=0\}} - \sum_i m_i(t)\sqrt{\alpha_i}\left[\mathbf{b}\cdot\nabla\theta^n_{\alpha^-_{(i)}}\right]. \quad (6.8)$$

The constant λ is chosen to make $\mu\mathbf{I} + \mathbf{D} - \lambda\mathbf{I}$ a negative definite tensor so that it poses no stability constraint on the time step size. Hence we can choose any λ satisfying

$$\mu + \frac{1}{2}(u^2 + v^2) \leqslant \lambda. \quad (6.9)$$

In the splitting scheme (6.8), we still need to solve a linear system at every time step:

$$(1 - k\lambda\Delta)\theta_\alpha^{n+1} = \theta_\alpha^n + k\left[\nabla\cdot(\mu\mathbf{I} + \mathbf{D} - \lambda\mathbf{I})\nabla\theta_\alpha^n + fI_{\{\alpha=0\}}\right] - k\sum_i m_i(t)\sqrt{\alpha_i}\left[(\mathbf{b}\cdot\nabla\theta^n_{\alpha^-_{(i)}}\right].$$

Since λ is a constant and the problem is periodic, we can use FFT to invert the Laplace operator very easily.

The implicit diffusion term $\lambda\Delta\theta_\alpha^{n+1}$ in (6.8) will help to stabilize the scheme and the CFL condition of the new method is imposed by the convection term only. It can be shown by standard Fourier analysis that the scheme (6.8) is stable if

$$k = O(h). \quad (6.10)$$

The CFL condition (6.10) is much less restrictive than (6.7). For a moderate spatial size $h = O(10^{-2})$, the splitting scheme (6.8) can increase the time step size by two orders and hence reduce the CPU time dramatically.

6.1.1 Full Sparse WCE Solutions

In the numerical experiment, we set $\mu = 0.001$, $\delta = 0.25$ and $A = 0.1\delta/2\pi$. The deterministic forcing is chosen as

$$f(x, y) = 0.5 \cos(8\pi x) \cos(8\pi y).$$

We solve the equation (6.3) to $T = 10.0$ and use the cosine basis functions (2.29) to expand the Brownian motion forcing. To truncate the WCE propagator (6.5) in $[0,\ 10.0]$, we keep ten Gaussian random variables and maximum sixth order Wick polynomials. In the sparse WCE truncation, we set the adaptive index for the Wick polynomial orders as

$$r = (6, 5, 4, 3, 2, 1, 1, 1, 1, 1). \tag{6.11}$$

Moreover, we include only the first six Gaussian random variables in the second and third Wick polynomials, and the first three Gaussian variables in the fourth, fifth, and sixth order Wick polynomials. To implement the sparse truncation, we reduce the length of the adaptive index r as the order of the polynomial increases.

Sparse truncation

1. Set $K = 10$, $N = 6$,
2. Set the adaptive index as $r = (6, 5, 4, 3, 2, 1, 1, 1, 1, 1)$,
   (i) For $2 \leqslant |\alpha| \leqslant 3$, change the adaptive index to $r = (6, 5, 4, 3, 2, 1)$,
   (ii) For $4 \leqslant |\alpha| \leqslant 6$, change the adaptive index to $r = (6, 5, 4)$.

Using the above sparse truncation, we can bring down the total number of WCE coefficients to 118, as opposed to 8,008 in the simple truncation. The truncated WCE propagator can be solved numerically, and it takes approximately the same CPU time as computing 118 MC realizations. Figure 6.2 is the L^2 norm of the computed WCE coefficients, ordered in the same way as in Figure 3.3. The legends *1st*, *2nd*, etc., in the plot indicate the starts of the different order WCE coefficients. The WCE coefficients decay fast as the polynomial order $|\alpha|$ increases. Within the same polynomial order, the WCE coefficients also decay when the Wick basis $T_\alpha(\xi)$ shifts it dependence to ξ_i with bigger subscript i. The L^2 norm of the WCE coefficients is a convenient way to access the convergence of the Wiener chaos

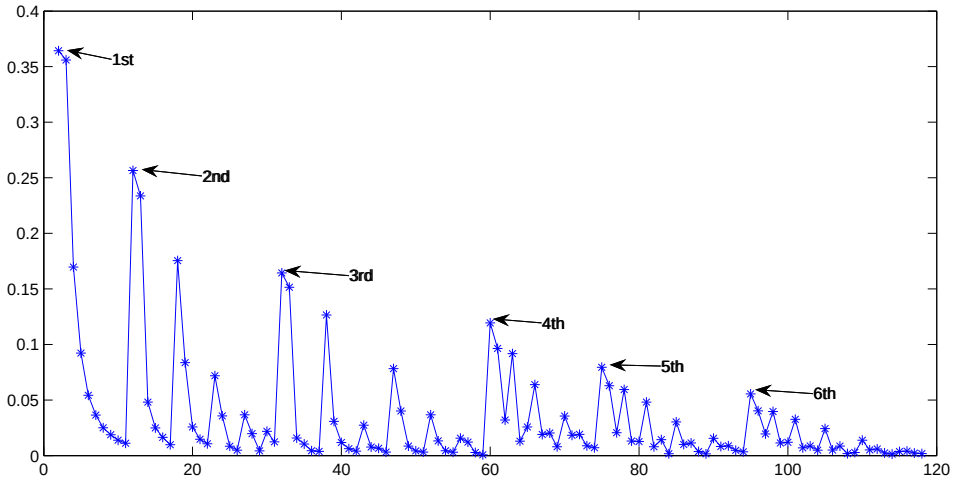

Figure 6.2: L^2 norm of the WCE coefficients for the full sparse truncation (118 coefficients). The legends *1st*, *2nd*, etc., indicate the start of the first, second, etc., polynomial orders of the WCE coefficients. The coefficients decay both in terms of the order of the Wick polynomials and the Gaussian random variables.

expansions.

For comparison purpose, we also solve the equation (6.3) by MC simulations with $50,000$ realizations. The mean and variance computed by the MC simulation are presented in Figure 6.3. We plot the difference between the WCE solutions and the MC solutions in Figure 6.4. Since the "cellular" flow is periodic from cell to cell, we only compare the results in one cell. The mean computed by the WCE method agrees with the MC simulation very well, with a relative error less than 2%. That is because in the WCE method the mean is computed from the closed moment equation (6.6). The difference is actually caused by the error in the MC simulations. However, the relative error for the variance is quite significant, closing to 10% in some local region. And the error is only improved slightly if we keep increasing the number of coefficients in the WCE truncation. The reason is that each individual WCE coefficient captures less and less information of the random solution at the higher order. However, the number of WCE coefficients increases very quickly as the order of truncation increases. For fixed K number of Gaussian random variables, there are

$$\binom{K+n-1}{n} = \frac{(K+n-1)!}{(K-1)!n!}$$

number of Wick polynomials with order n. For $K = 10$, there would be 11,440 number of

Wick polynomials with the order $N = 7$. Though each individual coefficient is very small at this range and can be ignored, their aggregated effect is not completely negligible. Since there are so many of those small coefficients at high orders, it is almost impossible to resolve their aggregated effect by keeping only a few dozens more of them. That is why the sparse WCE method has nearly 10% error in the variance and including more coefficients can only improve the accuracy slightly.

For short time solutions, the above difficulties are much less severe. As we have demonstrated, in short time integrations we only need to keep a small number of Gaussian random variables, and relatively low order Wick polynomials. As the time interval increases, more Gaussian random variables and higher order polynomials are needed, which will increase the number of WCE bases very quickly. To make matters worse, the energy of the solution spreads equally among those high order coefficients and their aggregated effects are very difficult to resolve. That is the main difficulty in applying the WCE method for long time integrations.

6.1.2 Aggressive Sparse WCE Solutions

Since the WCE coefficients at higher order are much less important, we can keep even less WCE coefficients than in the full sparse truncation. At the risk of being less accurate, the resulting coarse-scale WCE solution will still capture most information of the random solution. In the next numerical experiment, we keep only 6 Gaussian random variables and maximum fifth order Wick polynomials in the WCE truncation. We trim the finite WCE further by the adaptive index

$$r = (5, 4, 3, 2, 1, 1),$$

and reduce the number of Gaussian random variables at higher order Wick polynomials. More specifically, we adopt the following aggressive sparse WCE truncation:

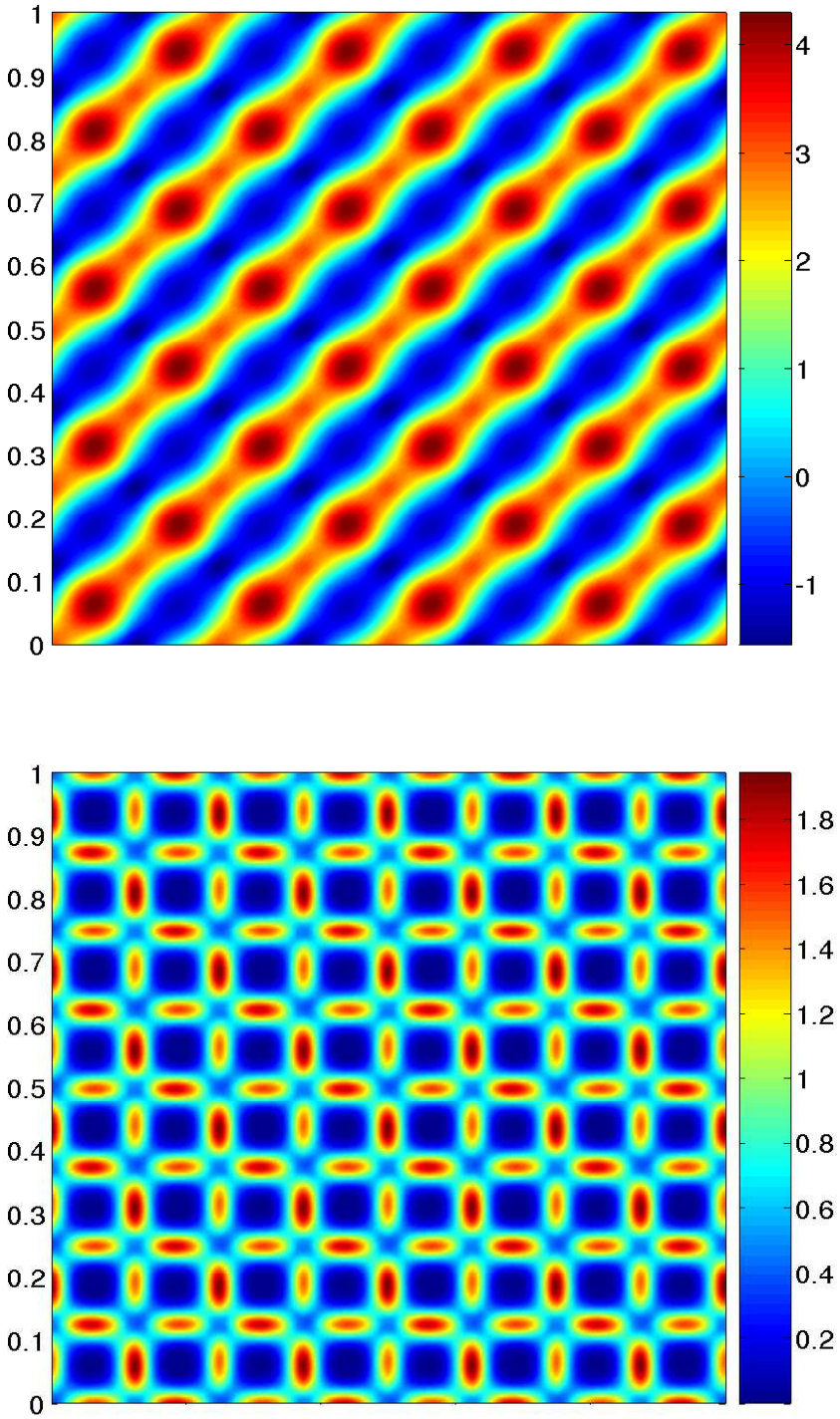

Figure 6.3: Mean (above) and variance (bottom) of θ at $T = 10.0$, computed by MC simulation with $50,000$ realizations. The maximum variance concentrates on the edges of each cell, where the random velocity will convect the particles from cell to cell and enhance the mixing.

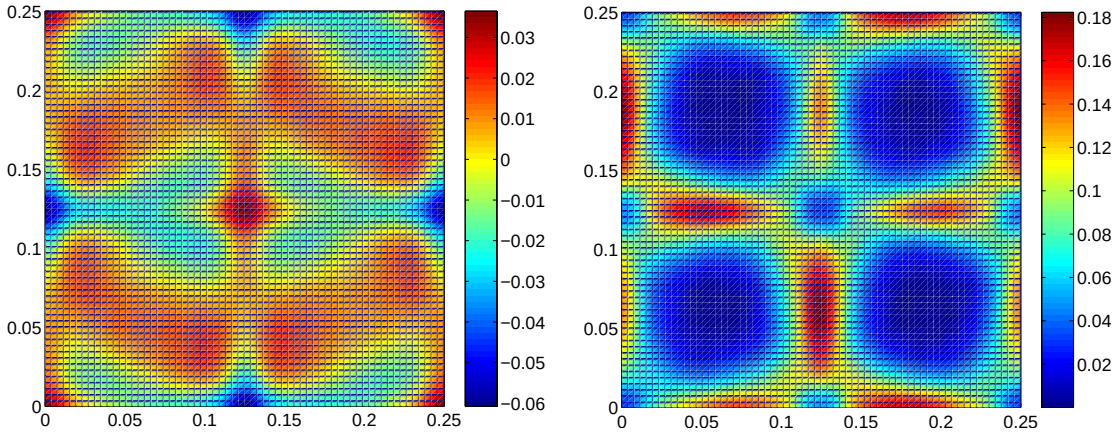

Figure 6.4: Error of the mean (left) and variance (right) of the WCE solution with full sparse truncations (118 coefficients) at $T = 10.0$. The mean has a relative error less than 2%, while the variance has a relative error closing to 10%. Increasing the number of WCE coefficients can only improve the accuracy of the variance slightly.

Aggressive sparse truncation

1. Set $K = 6$, $N = 5$,
2. Set the adaptive index as $r = (5, 4, 3, 2, 1, 1)$,
   (i) For $2 \leqslant |\alpha| \leqslant 3$, change the adaptive index to $r = (5, 4, 3, 2)$,
   (ii) For $4 \leqslant |\alpha| \leqslant 6$, change the adaptive index to $r = (5, 4)$.

With only 37 coefficients, the aggressive sparse truncation is expected to be less accurate than the full sparse truncation. We solve the resulting WCE propagator by the same implicit splitting method (6.8). Figure 6.5 is the L^2 norm of the WCE coefficients for the aggressive sparse truncation. Comparing Figure 6.5 with Figure 6.2, we found that the aggressive sparse truncation shares all the important coefficients with the full sparse truncation. In Figure 6.6 we plot the errors of the mean and variance of the aggressive sparse WCE solution. The error of the mean is the same as in Figure 6.4 since it is solved from the same closed mean equation (6.6). For the variance, the error of the aggressive sparse solution is not much worse than that of the full sparse solution either. That is quite surprising, considering the variance is the square summation of all the WCE coefficients, and much less coefficients are used in the aggressive sparse truncation. In fact, this is a very typical convergence behavior of the WCE solution: The low order WCE coefficients capture most

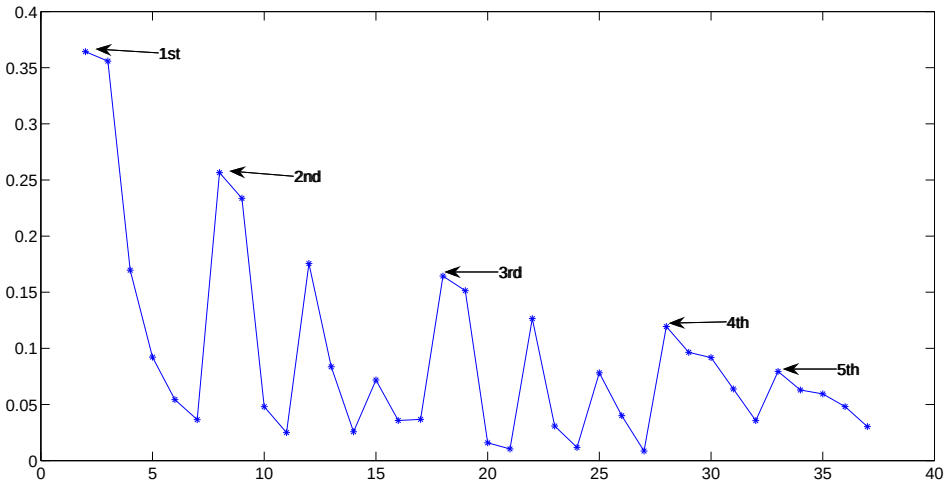

Figure 6.5: L^2 norm of the WCE coefficients with aggressive sparse truncation (37 coefficients). The legends *1st*, *2nd*, etc., indicate the start of the first, second, etc., polynomial orders of the WCE coefficients. The coarse-scale WCE truncation shares all the dominant coefficients with the full WCE truncation.

information of the random solution; the coefficients at increasingly higher order contribute less and less to the WCE solution; the incremental gain by adding more coefficients decreases rapidly as the order of the coefficient increases. That is why the full sparse WCE solution is only slightly better than the aggressive sparse truncation, even though the number of coefficients is tripled. This convergence property poses serious difficulties in computations if we want to resolve all the small scales of the random solution by the WCE method. However, this property of the WCE method can also work to our advantage if we exploit it wisely. Motivated by this observation, we have designed the sparse WCE truncation, which is proven to have the same asymptotical convergence as the simple truncation. Next we will use this same fact to design a long time integration strategy for the WCE method.

6.2 Refining the WCE Solution by MC Simulations

The WCE method is less efficient for long time integration since the required number of coefficients increases dramatically. With an aggressive sparse truncation, the WCE method still captures a large part of the random information of the solution. Adding more coefficients to the WCE truncation will improve the accuracy, but the gain for each added coefficient will decrease quickly as the order of the coefficients increases. To resolve the

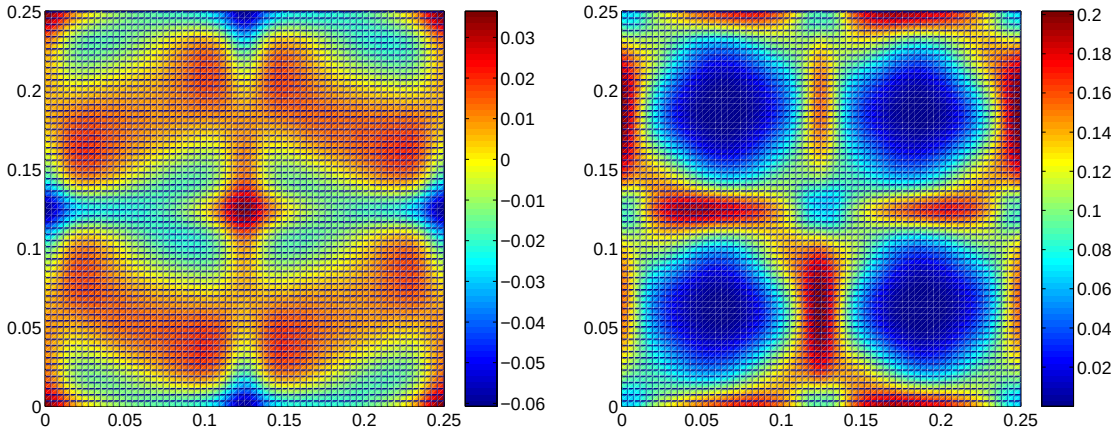

Figure 6.6: Error of the variance by the WCE method with the aggressive sparse truncation (37 coefficients) at $T = 10.0$. With only one-third of the number of coefficients of the full truncation, the error of the aggressive truncation is just slightly worse.

random solution by its WCE to the desired accuracy may require an overwhelmingly large number of coefficients. However, each high order coefficient contributes very little to the WCE solution, even though their aggregated effect is not negligible.

Instead of resolving all the small scales of the random solution by adding more and more WCE coefficients, we can first solve the problem by the WCE method with an aggressive truncation. The sparse WCE solution may not very accurate locally, but it will capture the large structure of the random solution. If we subtract the WCE solution from the true solution, the error will be localized and more or less homogeneous. Then we use MC simulations to correct the error in the WCE solution. Using the coarse WCE solution as a control variate, the variance of the MC estimator is reduced substantially and the correction process can be made very efficient.

6.2.1 Reducing the Variance of MC Simulations Using WCE Solutions

In most applications, the purpose of MC simulation is to estimate the statistical moments of the random solution. Suppose $\theta(\mathbf{x}, t, \omega)$ is the random solution of a SPDE, where ω denotes the random dependence (Brownian motion realizations). Denote the statistical moments of θ as $E[g(\theta)]$, where $g(x) = x^n, n = 1, 2, \ldots$. To simplify the notation, we have suppressed the spatial and temporal variables in the function θ. However, the expectation $E[g(\theta)]$ is always understood as a function of spatial and time variables.

Suppose $\theta(\omega_k),\ k = 1, 2, \ldots N$ are independent realizations of the random solution

computed by MC simulations. Then the expectation $E[g(\theta)]$ can be approximated by the MC ensemble average

$$I_N[g(\theta)] = \frac{1}{N}\sum_{k=1}^{N} g(\theta(\omega_k)). \tag{6.12}$$

Denote the error of the MC estimator (6.12) as

$$\epsilon_g(N) = E[g(\theta)] - I_N[g(\theta)].$$

The error $\epsilon_g(N)$ itself is a random variable. By the Central Limit Theorem [30], the root mean square error (RMSE) of the MC estimator (6.12) decays as

$$\sqrt{E\left[\epsilon_g^2(N)\right]} \simeq \frac{\sigma[g(\theta)]}{\sqrt{N}}, \tag{6.13}$$

where $\sigma^2[g(\theta)]$ is the variance of $g(\theta)$:

$$\sigma^2[g(\theta)] = \int \left[g(\theta) - Eg(\theta)\right]^2 d\omega.$$

So the ensemble average (6.12) converges to $E[g(\theta)]$ at the rate of $N^{-1/2}$ with a proportional constant that is the variance of $g(\theta)$. To reduce the RMSE of (6.12) by a factor of 2, we need to increase the realization number by a factor of 4. This slow convergence rate is an inevitable result of the Central Limit Theorem.

To accelerate the convergence of the MC ensemble average (6.12), we can instead reduce the proportional constant $\sigma[g(\theta)]$, that is, the variance of the MC estimator. Suppose there exists a function $p(\theta)$, which is a good approximation of $g(\theta)$, and whose expectation $E[p(\theta)]$ is known. We can split $E[g(\theta)]$ into two parts

$$E[g(\theta)] = E[p(\theta)] + E[g(\theta) - p(\theta)],$$

and use MC simulation to estimate $E[g(\theta) - p(\theta)]$ only. As a result we have

$$E[g(\theta)] \simeq E[p(\theta)] + \frac{1}{N}\sum_{k=1}^{N}\left[g(\theta(\omega_k)) - p(\theta(\omega_k))\right]. \tag{6.14}$$

The error of the estimation (6.14) is

$$\epsilon_{g-p}(N) = E[g(\theta) - p(\theta)] - \frac{1}{N}\sum_{k=1}^{N}\left[g(\theta(\omega_k)) - p(\theta(\omega_k))\right].$$

Based on the Central Limit Theorem, the RMSE of the estimation (6.14) is

$$\sqrt{E\left[\epsilon_{g-p}^2(N)\right]} \simeq \frac{\sigma_{g-p}}{\sqrt{N}}, \tag{6.15}$$

where σ_{g-p}^2 is the variance of the function $g(\theta) - p(\theta)$:

$$\sigma_{g-p}^2 = \int \left\{ \left[g(\theta) - Eg(\theta)\right] - \left[p(\theta) - Ep(\theta)\right] \right\}^2 d\omega.$$

If $p(\theta) - Ep(\theta)$ is a good approximation of $g(\theta) - Eg(\theta)$ in the mean square sense, then

$$\sigma_{g-p} \ll \sigma_g,$$

and the RMSE (6.15) would be much smaller than the RMSE (6.13). In other words, the ensemble average (6.14) would converge much faster than (6.12). Suppose the error tolerance for the MC estimation is ε, then the number of required realizations is

$$N \simeq \frac{\sigma^2}{\varepsilon^2}.$$

If we can reduce σ by a factor of 2, then the required number of MC realizations can be reduced by a factor of 4. In particular, suppose σ_{g-p} is one-tenth of σ_g, then (6.14) will converge 100 times faster than (6.12). So the variance reduction is quite effective in accelerating the convergence of the MC estimations if we can find a good approximation function $p(\theta)$. Such a function $p(\theta)$ is usually called the control variate of the MC simulation.

To make the above variance reduction work, the control variate $p(\theta)$ should "mimic" the random variability of the estimator $g(\theta)$. More precisely, $p(\theta) - E[p(\theta)]$ should be a good approximation of $g(\theta) - g[p(\theta)]$ in the mean square sense. Note that the mean of $p(\theta)$ does not have to approximate that of $g(\theta)$ very well. The crucial point is that $p(\theta)$ fluctuates around its mean in a similar way as $g(\theta)$ does. Ideally, the realization $p(\theta(\omega_k))$ has to be easily computable for any give Brownian motion path ω_k. Furthermore, the mean $E[p(\theta)]$

should have a closed form. In practice, it is very difficult to find a control variate $p(\theta)$ with all these rare properties, especially when $\theta = \theta(\mathbf{x}, t, \omega)$ is an unknown random process we are looking for.

The WCE solution provides a very effective control variate for variance reduction in MC simulations. Suppose

$$\theta_{\mathrm{WCE}}(\omega) = \sum \theta_\alpha T_\alpha(\xi(\omega)) \tag{6.16}$$

is a truncated WCE solution with a small number of coefficients. As our previous numerical results demonstrate, the coarse-scale WCE solution can capture a major part of the random variability of the solution. By subtracting the WCE solution θ_{WCE} from the true solution, we can estimate $E(\theta)$ by the MC ensemble average

$$E(\theta) \simeq E(\theta_{\mathrm{WCE}}) + \frac{1}{N}\sum_{k=1}^{N} \left[\theta(\omega_k) - \theta_{\mathrm{WCE}}(\omega_k)\right]. \tag{6.17}$$

In the above formula, we estimate the mean $E[\theta]$ by splitting it into two parts. The first term $E[\theta_{\mathrm{WCE}}] = \theta_{\mathbf{0}}$ is the mean computed by the WCE method, and the second term represents the MC correction to the WCE solution. For each given Brownian motion realization ω_k, $\theta(\omega_k)$ is computed by MC simulations. To evaluate the WCE realization $\theta_{\mathrm{WCE}}(\omega_k)$, we first compute the Gaussian random variables ξ_i numerically from the Ito integrals

$$\xi_i(\omega_k) = \int_0^T m_i(t) dW_t(\omega_k) \simeq \sum_n m_i(t_n)\Delta W_n(\omega_k), \qquad i = 1, 2, \ldots,$$

where $\Delta W_n = W(t_{n+1}) - W(t_n)$ is the independent increment of the Brownian motion. Plugging $\xi_i(\omega_k)$ into the WCE formula (6.16), we can easily get the WCE realization $\theta_{\mathrm{WCE}}(\omega_k)$. Note that

$$\sigma^2[\theta - \theta_{\mathrm{WCE}}] \leqslant E[(\theta - \theta_{\mathrm{WCE}})^2],$$

and θ_{WCE} is a fairly good approximation to θ in the mean square norm, $\sigma^2[\theta - \theta_{\mathrm{WCE}}]$ is expected to be much smaller than $\sigma^2[\theta]$. So the ensemble average (6.17) will converge much faster than the direct ensemble average

$$E(\theta) \simeq \frac{1}{N}\sum_{k=1}^{N} \theta(\omega_k). \tag{6.18}$$

In practice, only a few hundreds of MC realizations are needed to obtain a very accurate correction to the WCE solution.

To compute the second order moment $E[\theta^2]$, we can use the same WCE-MC hybrid method

$$E(\theta^2) \simeq E(\theta^2_{\text{WCE}}) + \frac{1}{N}\sum_{k=1}^{N}\big[\theta^2(\omega_k) - \theta^2_{\text{WCE}}(\omega_k)\big]. \tag{6.19}$$

The first term $E(\theta^2_{\text{WCE}}) = \sum \theta^2_\alpha$ is the second order moment estimated by the WCE method, and the second term is the MC correction. Since θ_{WCE} is a good approximation to θ, so is θ^2_{WCE} to θ^2. It is expected that $\sigma^2[\theta^2 - \theta^2_{\text{WCE}}] \ll \sigma^2(\theta^2)$, and the ensemble average (6.19) will converge much faster than the direct ensemble average

$$E(\theta^2) \simeq \frac{1}{N}\sum_{k=1}^{N}\theta^2(\omega_k). \tag{6.20}$$

In the above computational framework, the MC simulation is understood as a correction step to the WCE solution. Alternatively, we can also interpret the WCE solution as a precomputation step to obtain a control variate for MC simulations. This approach is a generalization of the original idea by Chorin [12], where he proposed using Hermite expansions as control variates for MC numerical integrations. To obtain a Hermite expansion of the integrand, MC simulations were employed again to compute the coefficients. Maltz and Hitzl [66] careful analyzed the variance of Chorin's estimator and generalized it to compute high dimensional integrals. Chorin's student Chang [11] applied the same idea in simulating a SODE with constant diffusion, where Chorin's estimator was used at consecutive time steps based on a recursive relation. Chang's approach was particularly designed for constant diffusion problems and cannot be generalized to other SODEs. In our current approach, a Wiener chaos solution is used as a control variate to estimate the statistical moments, which can be regarded as a generalized Chorin estimator. However, a key difference here is that the statistical moments of a random solution are integrals with respect to an unknown measure. Without introducing the Wiener chaos expansion, we cannot transform the measure of the random solution into a multi-dimensional Gaussian measure. Another unique feature of our approach is that the control variate is solved from a deterministic PDE system (WCE propagator), instead of computed by MC computations. When many expansion coefficients need to be computed at the same time, the PDE-based approach will

be much more efficient and accurate than the MC simulations.

6.2.2 The Strongly Convergent Numerical Schemes for MC Simulations

A very important issue in the above variance reduction procedure is that we have to compute the MC realizations $\theta(\omega_k)$ using a strongly convergent numerical scheme. Note that the WCE solution θ_{WCE} is a strong solution to the SPDE (6.3). That is, $\theta_{\text{WCE}}(\omega)$ will approximate the true solution $\theta(\omega)$ for each Brownian motion realization ω. Similarly, the MC solution should also approximate the true solution realization by realization. Approximating the solution by distribution is not enough for the variance reduction purpose. As a result, we have to use strongly convergent numerical schemes in the MC simulation. Particularly, the MC solution should be in the same probability space as the Brownian motion. And we need to compute the MC realization $\theta(\omega_k)$ and the WCE realization $\theta_{\text{WCE}}(\omega_k)$ using the same Brownian motion path ω_k. Please see Appendix B for the difference between strongly convergent and weakly convergent schemes for discretizing SDE.

To discretize the SPDE (6.3), we choose the strongly first order Milstein scheme (B.3), which can be reduced to the following form for equation (6.3):

$$\theta^{n+1} = \theta^n + k\left[\nabla\cdot(\mu\mathbf{I}+\mathbf{D})\nabla\theta^n + f\right] - (\mathbf{b}\cdot\nabla\theta^n)\Delta W_n + \nabla\cdot(\mathbf{D}\nabla\theta^n)\left[(\Delta W_n)^2 - k\right]. \quad (6.21)$$

We can further cancel the term $k\nabla\cdot(\mathbf{D}\nabla\theta^n)$ in the above scheme to get

$$\theta^{n+1} = \theta^n + k\left(\mu\Delta\theta^n + f\right) - (\mathbf{b}\cdot\nabla\theta^n)\Delta W_n + \nabla\cdot(\mathbf{D}\nabla\theta^n)\left[(\Delta W_n)^2\right].$$

If we discretize the spatial derivatives by central difference scheme, then the CFL condition posed by the diffusion term is $k = O(h^2)$ and by the convection term is

$$\Delta W_n \leqslant O(h).$$

Since $\Delta W_n = W(nk+k) - W(nk)$ has the order of $\sqrt{k}$, the CFL condition posed by the convection term is effectively

$$k \leqslant O(h^2).$$

So in the numerical scheme (6.21), the convection term poses similar CFL condition as the

diffusion term.

6.2.3 Numerical Results of the WCE-MC Hybrid Method

Next we test the efficiency of the WCE-MC hybrid method on the linear transport equation (6.3). We first solve the problem to $T = 10.0$ by the aggressive sparse WCE method as in Section 6.1.2. This coarse-scale WCE solution captures a major part of the random variability of the solution. Though the mean is quite accurate, the variance has a maximum error of 10%. To refine the numerical results by the WCE method, we compute 100 MC realizations and correct the variance by the ensemble averages (6.19). The CPU time for this WCE-MC hybrid method is approximately the same as the MC simulation with 200 realizations. Figure 6.7 is the error of the variance after the MC correction. Comparing with Figure 6.6, the accuracy of the variance is improved significantly. The maximum error of the variance is reduced from 0.2 to 0.03, and the relative error is less than 2%. So the MC correction step is indeed very efficient since only 100 realizations are computed.

We also plot the error of the mean by the WCE-MC method in Figure 6.7, even thought the mean of the WCE solution is quite accurate and needs no MC corrections. Alternatively, we can regard the WCE solution solely as a control variate for the MC simulation. As we emphasized before, how well the mean of the control variate approximates that of the estimator is irrelevant for the variance reduction purpose. The only thing that matters is the control variate fluctuates around its own mean in a similar way as the estimation target, i.e., the control variate captures the random variability of the estimator. Figure 6.7 demonstrates that the WCE solution is indeed a very effective control variate for the MC simulation. With only 100 MC simulations, both the mean and variance reach a relative error within 2% when the WCE solution is used as a control variate. To reach the same accuracy, the direct MC simulation need to compute at least 10,000 realizations. So the WCE-MC hybrid method is more than 50 times faster than the MC method in this case. To achieve the same accuracy by the direct WCE method, an overwhelming number of coefficients may be needed, and the resulting WCE propagator quickly becomes unmanageable numerically. So the WCE-MC hybrid method has significant advantages over the MC method or the direct WCE method in long time integrations.

The substantial acceleration in the WCE-MC convergence is due to the reduced proportional constant in the convergence rate $O\big(N^{-1/2}\big)$. While estimating the mean $E(\theta)$ by the

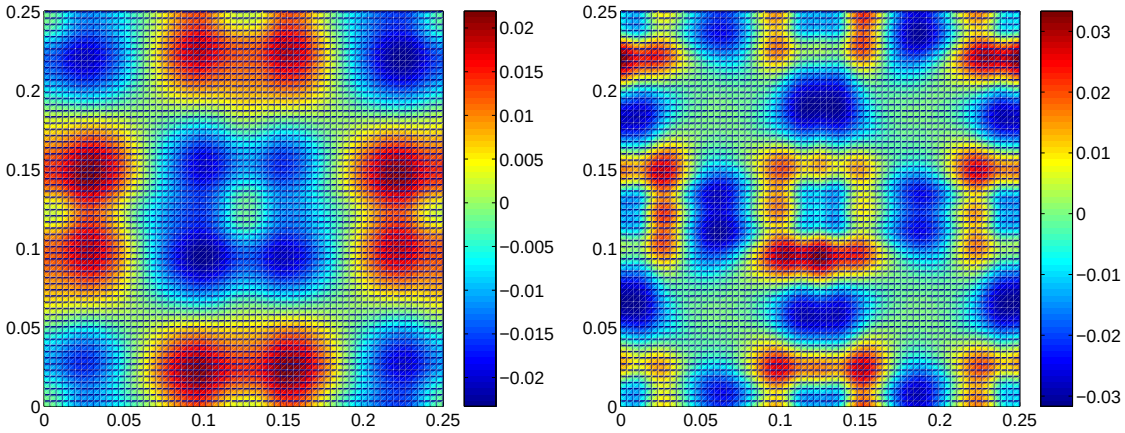

Figure 6.7: Errors of the mean and variance computed by the WCE-MC hybrid method with 37 WCE coefficients and 100 MC realizations. The WCE-MC hybrid converge more than 50 times faster than the MC simulation.

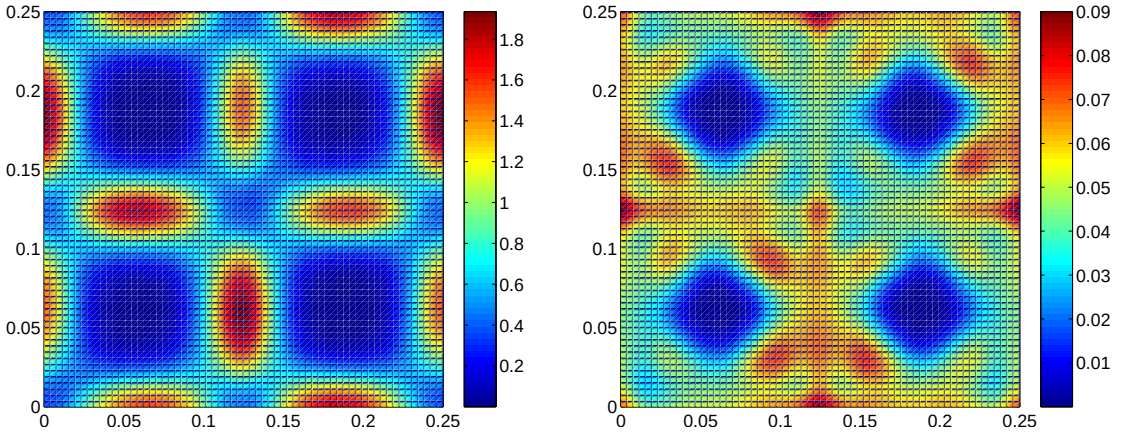

Figure 6.8: The convergence constants for estimating $E(\theta)$ by different ensemble averages: The left plot is the convergence constant $\sigma^2[\theta]$ for the direct MC ensemble average; the right plot is the convergence constant $\sigma^2[\theta - \theta_{\mathrm{WCE}}]$ for the WCE-MC ensemble average. Since $\sigma^2[\theta - \theta_{\mathrm{WCE}}]$ is less than one-twentieth of $\sigma^2[\theta]$, the WCE-MC ensemble average converges much faster than the direct MC ensemble average.

direct MC method (6.18) and the WCE-MC hybrid method (6.17), the constants in front of the convergence rates are $\sigma[\theta]$ and $\sigma[\theta - \theta_{\mathrm{WCE}}]$, respectively. We plot the constant $\sigma^2[\theta]$ and $\sigma^2[\theta - \theta_{\mathrm{WCE}}]$ for these two different ensemble averages in Figure 6.8. As we can see, the constant $\sigma^2[\theta - \theta_{\mathrm{WCE}}]$ is less than one-twentieth of $\sigma^2[\theta]$. That is why the MC estimator (6.17) which uses the WCE solution as the control variate converges much faster than the direct MC ensemble average (6.18).

Similarly, while estimating the second order moment $E(\theta^2)$ by the direct MC method (6.20) and the WCE-MC hybrid method (6.19), the constants in front of the convergence

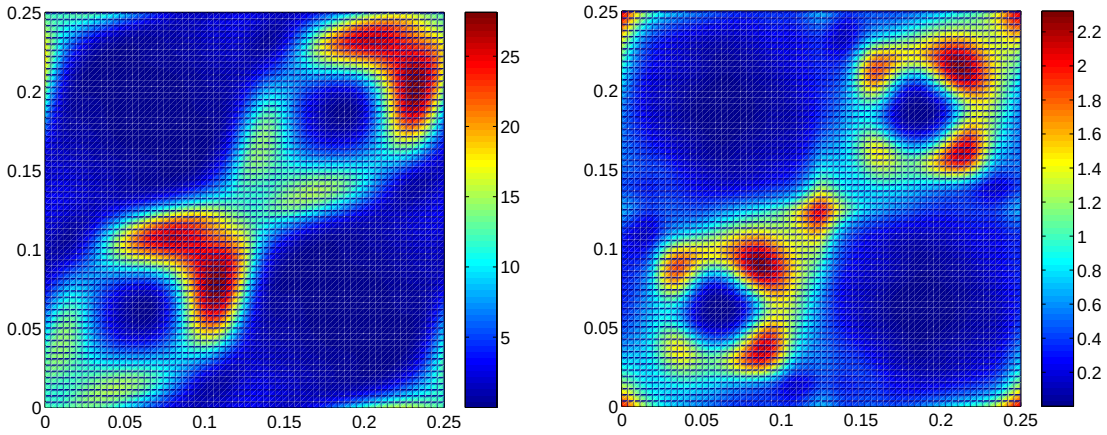

Figure 6.9: The convergence constants for estimating $E(\theta^2)$ by different ensemble averages: The left plot is the convergence constant $\sigma^2[\theta^2]$ for the direct MC ensemble average; the right plot is the convergence constant $\sigma^2[\theta^2 - \theta^2_{\rm WCE}]$ for the WCE-MC ensemble average. Since $\sigma^2[\theta^2 - \theta^2_{\rm WCE}]$ is less than one-tenth of $\sigma^2[\theta^2]$, the WCE-MC ensemble average converges much faster than the direct MC ensemble average.

rates are $\sigma[\theta^2]$ and $\sigma[\theta^2 - \theta^2_{\rm WCE}]$, respectively. In Figure 6.9, we plot the constant $\sigma^2[\theta]$ and $\sigma^2[\theta - \theta_{\rm WCE}]$ for these two different ensemble averages. The constant $\sigma^2[\theta^2 - \theta^2_{\rm WCE}]$ is less then one tenth of $\sigma^2[\theta^2]$. That is why the WCE-MC ensemble average (6.19) converges much faster than the direct MC ensemble average (6.20).

In summary, the WCE-MC hybrid method can solve SPDEs in much longer time intervals than the direct WCE method can. By using the WCE solution as a control variate, the WCE-MC hybrid method converges more than 50 times faster than the direct MC simulations. To achieve the same accuracy, the direct WCE method needs an overwhelming number of coefficients. The resulting WCE propagator will have too many equations to be numerically solvable. To obtain an approximate control variate by the WCE method, only a very sparse WCE truncation will suffice. So the explosion of dimensions can be partly avoided in long time computations. Comparing with the WCE method or the MC method alone, the WCE-MC hybrid method combines the merits of both methods and has greater flexibility.

However, the WCE-MC hybrid method is still sensitive to the length of the time intervals. For arbitrarily long time integrations, the error in the sparse WCE solution may be very large, and hence the performance of the WCE-MC will not be much better than that of MC simulations. Designing more efficient WCE-type numerical methods that are not sensitive to the integration time is still a challenging and open question. Nevertheless, the WCE-MC hybrid method can extend the application of WCE method to much longer time

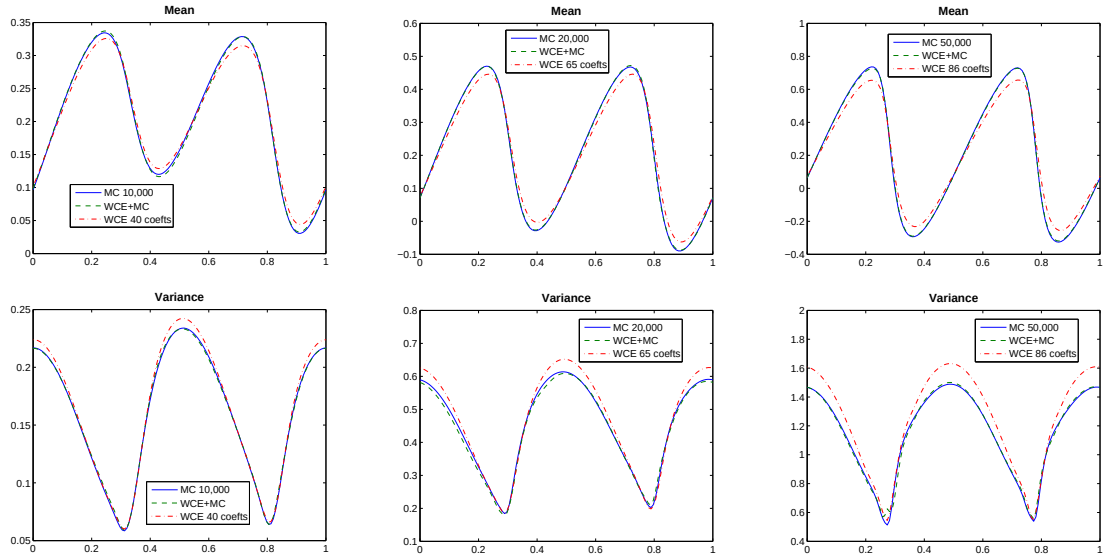

Figure 6.10: Mean (first row) and variance (second row) for random forcing with different magnitudes. The first column: $A = 1.0$; the second column: $A = 2.0$; the third column: $A = 4.0$. We correct the WCE results with 200 MC simulations in all three cases. The WCE-MC hybrid method can achieve the same accuracy as the MC simulations with only a fraction of the computational costs.

intervals, which already enables us to study some interesting SPDEs numerically.

6.3 Stochastic Burgers Equations with Large variability

In Section 3.4, we studied how the variability of the random forcing affects the convergence of the WCE method. We found that when the magnitude of the random forcing increases, the WCE method needs to include more and more coefficients to achieve similar accuracy. For random forcing with very large magnitude, the error of the WCE method is quite significant in the rough region of the solution. Nevertheless, the WCE solutions still capture the large structures of the random solution. In this section, we will test the WCE-MC hybrid method on the same stochastic Burgers equation (3.1) with large random forcing. Our numerical results demonstrate that the WCE-MC hybrid method can improve the accuracy of the WCE method for nonlinear problems with large random forcing as well.

We consider the same setup as in Section 3.4, where the spatial part of the random forcing is

$$\sigma(x) = A\cos(2\pi x),$$

with $A = 1.0,\ 2.0,\ 4.0$, respectively. To apply the WCE-MC hybrid method, we first solve

the problem by the WCE method with an aggressive sparse truncation. Though the sparse WCE solutions are not very accurate in the rough regions, they successfully capture the large structure of the random solution. To correct the unresolved small scales in the WCE results, we compute 200 MC realizations and compute their ensemble averages using the WCE solutions as the control variate. For comparison, we also solve the problems by MC simulations with 10,000, 20,000 and 50,000 realizations, respectively. For stronger random forcing, the random solution has much larger variance, hence we have to average more MC realizations to estimate its statistic moments. We compare the numerical results of WCE-MC with those of MC simulation in Figure 6.10. Since the coarse-scale WCE solutions already capture most of the variability of the random solution, the MC simulation with the WCE solution as the control variate provides very effective correction to the WCE solution. As we can see, the numerical results by the WCE-MC method agree very well with those by the MC simulations. However the WCE-MC hybrid method is 50 to 100 times more efficient than the MC simulations.

6.4 Front Propagation Speeds of Stochastic Reaction-Diffusion Equations

Front propagation in heterogeneous fluid flows appears in many scientific disciplines, such as combustion of premixed flames, reactive pollutant transports in porous media, etc. (see the review [98] and the reference therein). It is known that the large-scale front speed can be enhanced due to the presence of multi-scales in the fluid flows. A fundamental issue is how to characterize and compute the large time front speed, an upscaled quantity that depends on the statistics of the random flows in a highly nontrivial manner. Analysis of prototype models has been an efficient way to improve our understanding of front propagation with complex structures.

In this section, we consider the front propagation through a random shear flow in a two-dimensional channel $D = R \times \Omega$, where $R = (-\infty, +\infty)$ and Ω is an interval. The model equation is a reaction-diffusion (R-D) equation

$$u_t = \mu \Delta u + \delta\, V \cdot \nabla u + f(u), \tag{6.22}$$

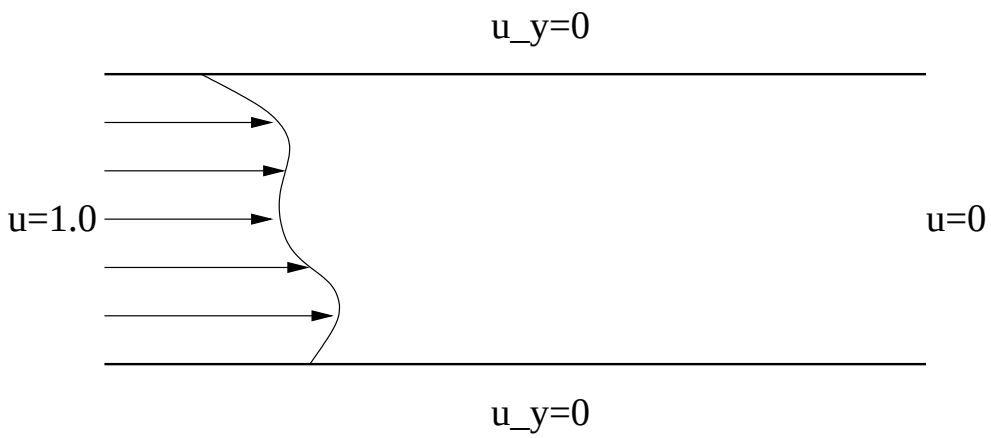

Figure 6.11: Illustrations of the front propagation for R-D equations

where $V = (\xi(y,t),\ 0)$ is a given shear flow, and $f(u) = u(1-u)$ is the Kolmogorov-Petrovsky-Piskunov (KPP) nonlinear reaction term. The small parameter δ measures the strength of the shear flow. We assume that

$$u(-\infty, y, t) = 1, \qquad u(+\infty, y, t) = 0,$$

and zero Neumann boundary condition $\partial u/\partial y$=0 at $\partial\Omega$. For the combustion problem, $u = 1$ stands for the hot burned region, and $u = 0$ represents the cold reactant region. The burning front is moving along the x-direction from the left to the right, and we are interested in how the front propagation speed depends on the shear strength δ in long time. The setup of this problem is illustrated in Figure 6.11

If the shear field V is deterministic and periodic in both space and time, the speed $c^*(\delta)$ of the front obeys the quadratic enhancement law [74]:

$$c^*(\delta) = c_0(1 + \alpha\,\delta^2 + h.o.t), \qquad \delta \ll 1, \tag{6.23}$$

where α is a positive constant depending only on ξ, c_0 is the pure R-D front speed with $\delta = 0$, and h.o.t is short for higher order terms.

However, the situation becomes more complicated if the flow field V is random, in which case the front speed $c^*(\delta, \omega)$ is a random variable. Suppose $\xi = \xi(y)$ is a stationary random process in space (does not depends on time) with zero ensemble mean, then the quadratic law (6.23) is still valid for the ensemble averaged speed $E[c^*(\delta, \omega)]$ [75]. If $\xi = \xi(y, t)$ is a

white-in-time, stationary, and ergodic Gaussian process, it was proved [99] that the front speed $c^*(\delta, \omega)$ is almost surely a constant $c^*(\delta)$ provided that the channel is the whole two-dimensional space. And the quadratic law (6.23) is shown to be true for the speed limit $c^*(\delta)$. For a bounded channel with a finite cross section, Xin et al. [76] derived an exact formula as well as upper bounds for the front speed $c^*(\delta, \omega)$. Based on a variational analysis, their results reveal that the random front speed $c^*(\delta, \omega)$ converges to a deterministic constant $c^*(\delta)$ almost surely when t is sufficiently large. However, no precise asymptotical relation as (6.23) has been found for the steady front speed $c^*(\delta)$. Their numerical simulations with an Ornstein-Uhlenbeck process suggest that the front speed still obeys a quadratic law. Due to the computational intensity [76], no numerical result is available yet for the interesting case where the shear flow is white-in-time and homogeneous in space.

For the white ergodic Gaussian shear flow, the almost surely convergence of the random front speed $c^*(\delta, \omega)$ is due to the ergodicity and self-averaging of the fluid flow in the y-direction or in the time direction. With the ergodicity assumption, we can replace the spatial or time average by the ensemble average. Instead of computing a single realization for infinitely long time ($T = 30,000$ was used in [76]), we can compute the mean front speed for many realizations in a relatively short time. The ensemble average generally yields more accurate and stable numerical results than a single realization. For this purpose, we apply the WCE-MC hybrid method to solve the R-D equation with a white-in-time and homogeneous shear flow. Our numerical results confirm that the front speed indeed obeys the quadratic enhancement law (6.23).

6.4.1 Problem Setup

We consider a random shear flow $\xi(y, t)$ which is mean zero, white-in-time and homogeneous in space, with the covariance function given by

$$E[\xi(y_1, t_1)\xi(y_2, t_2)] = \Gamma(|y_1 - y_2|)\, \delta(t_1 - t_2). \tag{6.24}$$

Obviously, the spatial covariance function $\Gamma(|y_1 - y_2|)$ is symmetric and semi-positive definite. We assume the cross section of the channel is $\Omega = [0, 1]$. Consider the eigenvalue problem:

$$\int_0^1 \Gamma(|y_1 - y_2|)\, \psi_k(y_2) dy_2 = \lambda_k \psi_k(y_1). \tag{6.25}$$

Suppose the eigenvalues are sorted in descending order by magnitude: $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant 0$. For convenience, we assume that the eigenfunctions $\psi_k(y)$ are normalized so that $\|\psi_k\|_{L^2} = 1$. It can be shown that $\{\psi_k;\ k = 1, 2, \ldots\}$ form an orthonormal basis in $L^2[\Omega]$. Furthermore, the spatial covariance function has the decomposition

$$\Gamma(|y_1 - y_2|) = \sum_{k=1}^{\infty} \lambda_k \, \psi_k(y_1)\psi_k(y_2), \tag{6.26}$$

and the Gaussian process $\xi(y, t)$ can be decomposed as

$$\xi(y, t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\psi_k(y) \, \dot{W}_k(t), \tag{6.27}$$

where $W_k(t)$ are independent Brownian motions. The expansion (6.27) is a special case of the Karhunen-Loeve expansion. Please see Section 7.1 for a detailed discussion about Karhunen-Loeve expansion. It is easy to verify that the random process given by formula (6.27) is indeed a stationary Gaussian process and has (6.24) as its covariance function. Since $\xi(y, t)$ is white-in-time, the random convection term in the R-D equation (6.22) should be proposed in Stratonovich's sense. Consequently, the R-D equation (6.22) can be rewritten as

$$u_t = \mu\Delta u + \delta \sum_{k=1}^{\infty} \sqrt{\lambda_k}\psi_k(y) \, \dot{W}_k(t) \circ u_x + f(u), \qquad (x, y) \in R \times [0, 1]. \tag{6.28}$$

From the relations between Stratonovich's integrals and Ito's integrals, we have

$$\begin{aligned} \delta \sum_{k=1}^{\infty} \sqrt{\lambda_k}\psi_k(y) \, \dot{W}_k(t) \circ u_x &= \delta \sum_{k=1}^{\infty} \sqrt{\lambda_k}\psi_k(y) \, \dot{W}_k(t) \cdot u_x + \frac{\delta^2}{2} \sum_{k=1}^{\infty} \lambda_k \, \psi_k^2(y) u_{xx} \\ &= \delta \sum_{k=1}^{\infty} \sqrt{\lambda_k}\psi_k(y) \, \dot{W}_k(t) \cdot u_x + \frac{\delta^2}{2}\Gamma(0) \, u_{xx}. \end{aligned}$$

So we can rewrite the SPDE (6.28) in Ito's form:

$$u_t = \mu\Delta u + \frac{1}{2}\delta^2 \, \Gamma(0) \, u_{xx} + \delta \sum_{k=1}^{\infty} \sqrt{\lambda_k}\psi_k(y) \, \dot{W}_k(t) \cdot u_x + u(1 - u). \tag{6.29}$$

We are interested in the average front speed of the equation (6.29) at the regime $\delta \ll 1$.

We consider the shear flow $\xi(y,t)$ with the spatial covariance function

$$\Gamma(|y_1 - y_2|) = \cos(2k\pi(y_1 - y_2)), \tag{6.30}$$

where k is a given integer. Obviously, the covariance function (6.30) is symmetric and semi-positive definite. The eigenvalue problem (6.25) has only two nonzero eigenvalues: $\lambda_1 = \lambda_2 = \frac{1}{2}$, and

$$(\lambda_1, \psi_1) = \left(\frac{1}{2},\ \sqrt{2}\cos(2k\pi y)\right), \qquad (\lambda_2, \psi_2) = \left(\frac{1}{2},\ \sqrt{2}\sin(2k\pi y)\right).$$

All the other eigenvalues are zero $\lambda_k = 0, k \geqslant 3$. As a result, the Gaussian process $\xi(y,t)$ has a simple representation

$$\xi(y,t) = \cos(2k\pi y)\dot{W}_1(t) + \sin(2k\pi y)\dot{W}_2(t). \tag{6.31}$$

The random process given by (6.31) is a spatially mean zero, white-in-time, and stationary Gaussian process. For the covariance function (6.30), we have $\Gamma(0) = R(y,y) = 1$. Hence the R-D equation (6.29) can be simplified as

$$u_t = \mu\Delta u + \delta\left[\cos(2k\pi y)\dot{W}_1(t) + \sin(2k\pi y)\,\dot{W}_2(t)\right]\cdot u_x + \frac{\delta^2}{2}u_{xx} + u(1-u). \tag{6.32}$$

This is the final model equation we will be working with.

6.4.2 Wiener Chaos Propagator for the R-D Equation

Next we derive the WCE propagator for the equation (6.32) in a finite time interval $[0,T]$. Suppose $\{m_i(t);\ i = 1,2,\ldots\}$ is an orthonormal basis in the Hilbert space $L^2[0,T]$. Define the independent Gaussian random variables

$$\xi_i^k = \int_0^T m_i(t)dW_k(t), \qquad i = 1,2,\ldots, \quad k = 1,2.$$

Let $\alpha = (\alpha_i^k;\ i = 1,2,\ldots,\ k = 1,2)$ be a multi-index with nonnegative integer components. Denote the finite multi-index set as $\mathcal{J} = \left\{\alpha;\ |\alpha| = \sum_{i,k}\alpha_i^k < \infty\right\}$. For $\alpha \in \mathcal{J}$, define the

Wick polynomial

$$T_\alpha(\xi) = \prod_{k=1}^{2}\prod_{i=1}^{\infty} H_{\alpha_i^k}(\xi_i^k).$$

Denote $\mathcal{F}_t$ as the σ-algebra generated by the Brownian motions $\{W_1(s), W_2(s); 0 \leqslant s \leqslant t\}$. Define $T_\alpha(t) = E(T_\alpha(\xi) \mid \mathcal{F}_t)$, then $T_\alpha(t)$ is a martingale and satisfies the differential equation:

$$dT_a(t) = \left(\sum_i m_i(t)\sqrt{\alpha_i^1}\, T_{\alpha^-_{(1,i)}}(t)\right) dW_1(t) + \left(\sum_i m_i(t)\sqrt{\alpha_i^2}\, T_{\alpha^-_{(2,i)}}(t)\right) dW_2(t). \quad (6.33)$$

where

$$\alpha^-_{(k,i)}(l,j) = \begin{cases} \alpha_i^k - 1 & \text{if } l = k, j = i, \\ \alpha_j^l & \text{else}. \end{cases}$$

The proof of (6.33) is very similar to Theorem 2.5 with one Brownian motion.

Suppose the solution of the SPDE (6.32) has the following Wiener chaos expansion

$$u(x,y,t) = \sum_{\alpha\in\mathcal{J}} u_\alpha(x,y,t)T_\alpha(\xi), \qquad 0 \leqslant t \leqslant T. \quad (6.34)$$

Similar to the multiplicative stochastic Burgers equation in Section 3.2.1, we can obtain the WCE propagator for the equation (6.32):

$$\begin{aligned}(u_\alpha)_t &= \mu\Delta u_\alpha + \frac{\delta^2}{2}(u_\alpha)_{xx} + u_\alpha - \sum_{p\in\mathcal{J}}\sum_{0\leqslant\beta\leqslant\alpha} C(\alpha,\beta,p)u_{\alpha-\beta+p}u_{\beta+p} \\ &\quad +\delta\cos(2k\pi y)\sum_i m_i(t)\sqrt{\alpha_i^1}\frac{\partial}{\partial x}\left(u_{\alpha^-_{(1,i)}}\right) + \delta\sin(2k\pi y)\sum_i m_i(t)\sqrt{\alpha_i^2}\frac{\partial}{\partial x}\left(u_{\alpha^-_{(2,i)}}\right).\end{aligned}$$

By solving the WCE propagator numerically, we can compute the ensemble average of the front speeds very easily.

6.4.3 Numerical Simulations

In the numerical simulations, we consider a finite channel with $D = [0, L]\times[0, 1]$. We choose a front-like (step function) initial condition

$$u(x,y,0) = 1 \quad \text{for} \quad 0 \leqslant x \leqslant 1, \qquad u(x,y,0) = 0 \quad \text{for} \quad 1 < x \leqslant L. \quad (6.35)$$

On the opening ends of the channel, u satisfies the Dirichlet boundary condition

$$u(0, y, t) = 1, \qquad u(L, y, t) = 0. \tag{6.36}$$

On the upper and bottom solid boundary, the zero Neumann boundary condition is specified.

For a given parameter δ, we solve the WCE system numerically and compute the integral

$$I(t) = \int_0^L \int_0^1 E[u(x, y, t, \omega)] dy dx = \int_D u_{\mathbf{0}}(x, y, t) dy dx, \tag{6.37}$$

where $u_{\mathbf{0}}$ is the first WCE coefficient. $I(t)$ is the average area of the burned region behind the flame front. Since the area of the cross section of the channel is one, $I(t)$ is also the average location of the front along the x-direction. We compute the integral (6.37) numerically by composite Simpson rule. Then we approximate the average front speed $c^*(\delta)$ as the optimal slope for $I(t)$ at the time interval $[t_1, T]$:

$$c^*(\delta) = \arg\min_b \int_{t_1}^T |I(t) - bt - a|^2 dt. \tag{6.38}$$

We will solve the R-D equation to a sufficiently long time T so that the front location $I(t)$ is already well stabilized at the time interval $[t_1, T]$.

In the numerical experiment, we set $\mu = 0.02$ and $L = 10.0$. By some trial computations, we found that the front speed becomes stable after $t = 4.0$ (dimensionless). We solve the WCE propagator to $T = 8.0$ and fit the front speed by (6.38) on the time interval $[5.0, 8.0]$. In the WCE method, we represent each Brownian motion by eight cosine functions, and truncate the Wick polynomials at maximum third order. Using the sparse WCE truncation, the resulting WCE propagator has only 61 coefficients. Note that the nonlinearity in the R-D equation only involves the function itself, but not its derivatives. So the nonlinearity in the R-D equations is relatively weaker than that of the stochastic Burgers or stochastic Navier-Stokes equations. Furthermore, the random shear flow has small magnitude in the regime $\delta \ll 1.0$. As we have shown before, small random forcing will greatly accelerate the convergence of the WCE method. Due to these reasons, the above moderate order of WCE truncation is fairly accurate in computing the mean front. We further refine the mean of the WCE solution by 50 MC simulations. The MC simulation provides minor corrections to the WCE solution near the sharp front regions. To simulate the same problem by MC

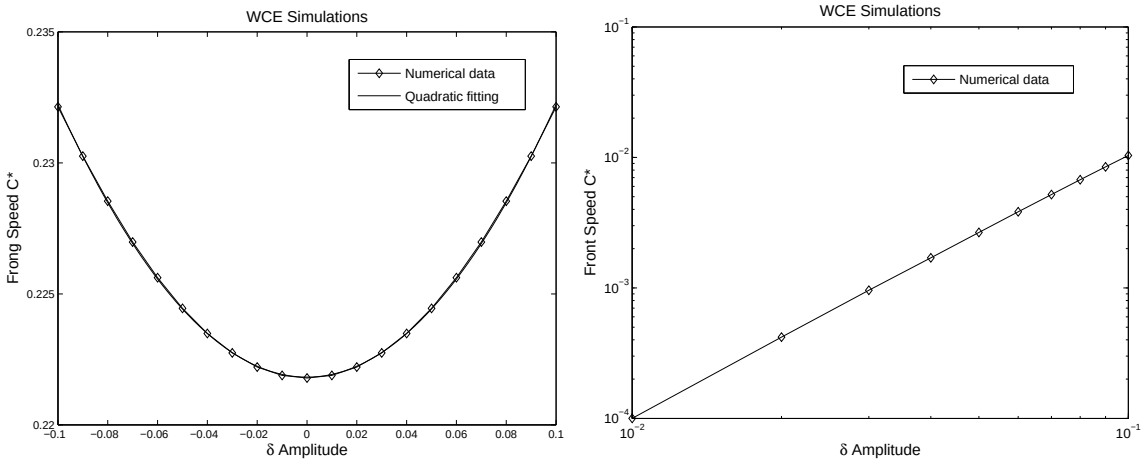

Figure 6.12: Ensemble average front speed computed by the WCE-MC hybrid method. In the left plot, the front speed $c^*(\delta)$ can be fitted by a parabola $c^*(\delta) \simeq c_0(1 + r\delta^2)$ almost perfectly, with $c_0 = 0.22183$ and r =4.6901. The right plot is the log-log plot of $c^*(\delta) - c_0$ against δ. The curve can be fitted by a linear function with a slop $\nu = 2.012$. So the front speed $c^*(\delta)$ indeed obeys the quadratic enhancement law.

simulations alone, thousands of MC realizations need to be computed for each value of δ. So the WCE-MC hybrid method is scores of times more efficient than the simple MC simulations.

To find out how $c^*(\delta)$ scales as a function of δ, we solve the R-D equation by the WCE-MC hybrid method repeatedly with different δ. The parameter δ takes value in the range $[-0.1,\ 0.1]$ with an increment of 0.01. We plot the average front speed $c^*(\delta)$ as a function of δ in Figure 6.12. The curve can be fitted by a quadratic polynomial very accurately, with

$$c^*(\delta) \simeq 0.22183 + 1.0404\ \delta^2 = 0.22183(1 + 4.6901\ \delta^2).$$

In Figure 6.12, we also plot $c^*(\delta)$ as a function of δ in log-log scale. We found that the exponent ν in the relation $c^* - c_0 \simeq O(\delta^\nu)$ is 2.012. So the numerical results confirm that $c^*(\delta)$ indeed satisfies the quadratic enhancement law (6.23).

The Wiener chaos expansion also provides a useful tool for studying the stochastic R-D problem (6.22) theoretically. Using the WCE method, we can reduce the stochastic R-D equation into a deterministic PDE system (WCE propagator). The WCE propagator is equivalent to the original stochastic R-D equation and resembles a deterministic R-D equation system with spatial-temporal periodic shear flows. Since the front propagation problems have been well studied in deterministic periodic cases, we can translate those

techniques and results to the WCE system, and use it to analyze the properties of the original stochastic equation. This approach is quite different from the variational principles used by Xin et al. [75, 76], where they choose to attack the stochastic R-D equation directly. It is expected the WCE perspective can bring new tools to the research of the stochastic reaction-diffusion problems.

Chapter 7

WCE Methods and Upscaling for Stochastic Elliptic Equations

In simulating fluid flows in porous media, people often need to consider the elliptic problem

$$\nabla \cdot k \nabla P = 0, \tag{7.1}$$

where P is the pressure, and $k(x)$ is the permeability field of the subsurface media. Due to lack of data, it is almost impossible to obtain a detailed description of $k(x)$. To model the uncertainty in describing the permeability field, people usually introduce a certain level of random variability to $k(x)$ and assume that $k(x)$ is a stochastic process. As a result, the solution $P(x)$ is also a random field, which reflects how the uncertainty in the input will propagator and affect the output of the model.

Since $k(x)$ is modeled as a random process, (7.1) becomes a stochastic elliptic equation. In an original work, Ghanem and Spanos [39] developed a new method, called stochastic finite element method, for solving the equation (7.1) numerically. Based on Karhunen-Loeve expansions, they first expanded the random input as a series of random Gaussian variables. Then the random solution was represented as a Fourier-Hermite expansion of those Gaussian random variables. Once the governing equations of the expansion coefficients are solved numerically, they can quantify the uncertainly in the random solution (the response) based on the deterministic coefficients and the random Hermite bases. This novel approach for solving stochastic elliptic equations was further studied by many others, notably, Babuska et al. [5, 4, 3], Schwab et al. [91, 32], Keese et al. [67, 51], Roman et al. [89], etc.

An important assumption made in previous work [5, 4, 91] for the stochastic elliptic equation (7.1) is that the random coefficient $k(x)$ satisfies the strong elliptic condition.

That is, there exist two constants k_{min} and k_{max} such that

$$0 < k_{min} \leq k(x, \omega) \leq k_{max} < \infty. \tag{7.2}$$

However, the assumption (7.2) is usually not true for the widely used Gaussian models, where $k(x, \omega)$ is not uniformly bounded. How to model the permeability field $k(x)$ is still a challenging problem to be addressed. In previous work [32, 5, 4, 3], the authors proposed truncating the Karhunen-Loeve expansion of the permeability field $k(x, \omega)$ first, and obtain a finite dimensional approximation $k_M(x, \omega)$. Then they approximate the original equation by the following one:

$$\nabla \cdot k_M(x, \omega) \nabla P = 0.$$

A potential problem with this approach is that the truncated permeability $k_M(x, \omega)$ may violate the elliptic assumption (7.2) if the Karhunen-Loeve expansion of k is not uniformly convergent. However, the uniform convergence of the Karhunen-Loeve expansion is available only for very special cases.

In this chapter, we will study the stochastic elliptic equations (7.1) in the context of reservoir modeling. We will follow Ghanem's stochastic finite element approach, but re-derive the WCE system for the stochastic elliptic equation using the formulae developed in Chapter 2. A new ingredient of our derivation is an explicit and exact formula for the stiffness coefficients. Based on this explicit formula, we find that even though the permeability field is represented by an infinite series, only finite terms will contribute to each stiffness coefficients. So there is no need to truncate the permeability field.

In the following, we first introduce the Karhunen-Loeve expansion for a stochastic process, which is the base for solving the stochastic elliptic problem (7.1) by WCE method. Then we derive an elliptic system for the WCE coefficients and analyze its property. At the end we will present an upscaling formulation for the WCE elliptic system. Multi-scale finite element methods and upscaling for scaler elliptic equations have been studied by many authors [19, 20, 26, 21, 47]. However, the upscaling of elliptic systems is relatively less studied. Combining the WCE method with the multiscale upscaling is a novel practice. In Chapter 8, we will apply the upscaled WCE method to the uncertainty quantification in subsurface modeling.

7.1 Karhunen-Loeve Expansion

Suppose $Y(x,\omega)$, $x \in \Omega$, is a random process with finite second order moment

$$\int_\Omega E[Y^2(x,\omega)]dx < \infty.$$

For simplicity, we assume that $E[Y] = 0$. For a given orthonormal basis $\{\psi_k\}$ in $L^2(\Omega)$, we can expand $Y(x,\omega)$ as a generalized Fourier series

$$Y(x,\omega) = \sum_{k=1}^{\infty} Y_k(\omega)\psi_k(x), \tag{7.3}$$

where

$$Y_k(\omega) = \int_\Omega Y(x,\,\omega)\,\psi_k(x)dx, \qquad k = 1, 2, \ldots$$

are random variables with zero means. We look for a special basis $\{\phi_k\}$ that makes Y_k uncorrelated: $E(Y_iY_j) = 0$ for all $i \neq j$. Denote the covariance function of $Y(x,\omega)$ as $R(x,y) = E\left[Y(x,\omega)\,Y(y,\omega)\right]$. Then the special basis functions $\{\phi_k\}$ should satisfy

$$E[Y_iY_j] = \int_\Omega \phi_i(x)dx \int_\Omega R(x,y)\phi_j(y)dy = 0, \qquad i \neq j.$$

Since $\{\phi_k\}$ is complete and orthonormal in $L^2(\Omega)$, it follows that $\phi_k(x)$ are eigenfunctions of $R(x,y)$:

$$\int_\Omega R(x,y)\phi_j(y)dy = \lambda_j\phi_j(x), \qquad j = 1, 2, \ldots, \tag{7.4}$$

where $\lambda_j = E[Y_j^2] > 0$. If we choose the basis functions $\phi_k(x)$ as the solutions of the eigen problem (7.4), then the random variables $Y_k(\omega)$ are indeed uncorrelated. Denote $\theta_k = Y_k/\sqrt{\lambda_k}$, we have the following expansion:

$$Y(x,\omega) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\,\theta_k(\omega)\,\phi_k(x), \tag{7.5}$$

where θ_k satisfy $E(\theta_k) = 0$ and $E(\theta_i\theta_j) = \delta_{ij}$. If $Y(x,\omega)$ is a Gaussian process, then $Y_k(\omega)$ are independent Gaussian random variables. The expansion (7.5) is called the Karhunen-Loeve expansion (KLE) of the stochastic process $Y(x,\omega)$. The KLE resembles the eigenvalue decomposition of a symmetric semi-positive definite matrix. For finite discrete processes,

the KLE reduces to the principal component decomposition.

The KLE (7.5) represents the stochastic process as a series of uncorrelated random variables. The L^2 basis functions $\phi_k(x)$ are deterministic and resolve the spatial dependence of the random process. The KLE converges to the random process $Y(x,\omega)$ in the mean square sense

$$\lim_{N\to\infty}\int_\Omega E|Y(x,\omega)-Y_N(x,\omega)|^2\,dx=0, \tag{7.6}$$

where

$$Y_N=\sum_{k=1}^{N}\sqrt{\lambda_k}\,\theta_k\,\phi_k \tag{7.7}$$

is a finite term KLE. For the Fourier expansion (7.3) with an arbitrary basis $\{\psi_k\}$, it can be shown that

$$\int_\Omega E\big|Y(x,\omega)-Y_N(x,\omega)\big|^2\,dx=\min\int_\Omega E\big|Y(x,\omega)-\sum_{k=1}^{N}Y_k\psi_k(x)\big|^2\,dx. \tag{7.8}$$

So among all the Fourier expansions (7.3), the KLE with the basis function $\phi_k(x)$ determined by the eigen problem (7.4) has the fastest convergence rate.

The convergence rate of the KLE only depends on the smoothness of the covariance function, but not on the specific probability structure of the process being expanded. The KLE converges fast if the covariance function $R(x,y)$ is very smooth. In this case, we only need to keep the leading order terms (quantified by the magnitude of λ_k) in the finite KLE and still capture most of the energy of the stochastic process $Y(x,\omega)$. For the finite KLE (7.7), we define its energy ratio as

$$e(N):=\frac{\int_\Omega E|Y_N|^2dx}{\int_\Omega E|Y|^2dx}=\frac{\sum_{k=1}^{N}\lambda_k}{\sum_{k=1}^{\infty}\lambda_k}. \tag{7.9}$$

If the eigenvalues $\lambda_k, k=1,2,\ldots,$ decay very fast, then the finite term KLE would be good approximations of the stochastic process. Figure 7.1 shows the eigenvalues of KLE with a Gaussian covariance function

$$R(x,y)=\sigma^2\exp\Big(-\frac{|x-y|^2}{2L^2}\Big), \tag{7.10}$$

where $\sigma=1.0$ and $L=0.1$ is the correlation length. Since the Gaussian covariance function is infinitely smooth, the eigenvalues decay exponentially fast [32]. The finite KLE

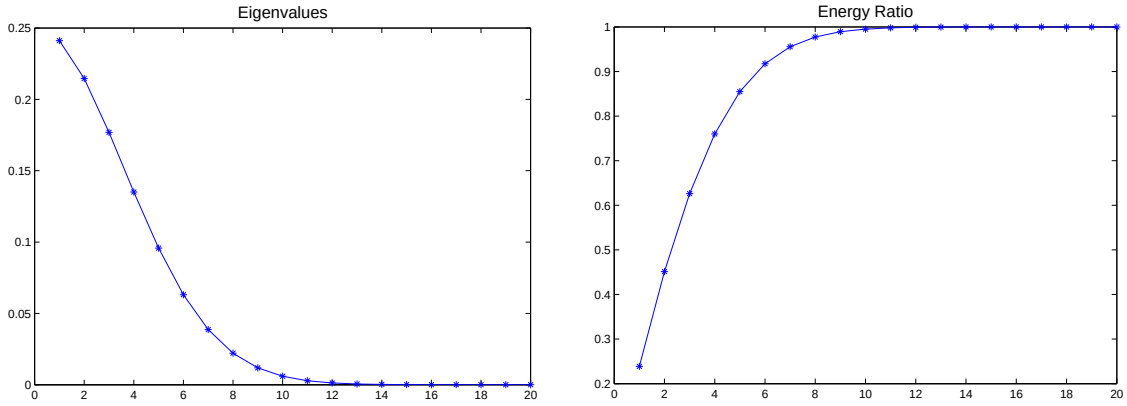

Figure 7.1: The first 20 eigenvalues of the KLE with a Gaussian covariance function. The right plot is the energy ratio (7.9) of the finite KLEs with different number of terms. The eigenvalues decay very fast and the KLE with ten terms captures 99% of the total energy of the random process.

can capture 99% of the total energy of the random process by keeping only ten terms.

If the covariance function is less smooth, then the KLE will decay slower. Consider the exponential covariance function

$$R(x,y) = \sigma^2 \exp\Big(-\frac{|x-y|}{L}\Big), \tag{7.11}$$

where $\sigma = 1.0$, and $L = 0.1$ is the correlation length. It has been shown that the eigenvalues of the exponential covariance function (7.11) only decay algebraically, since (7.11) is piecewise smooth. Figure 7.2 is the first 40 eigenvalues of the KLE. The eigenvalues indeed decay much slower. With 20 terms, the finite KLE only captures 90% of the total energy of the random process.

The convergence rate of the eigenvalues also depends on the correlation length L of the covariance function. The smaller the correlation length, the slower the convergence rate. Figure 7.3 shows the eigenvalues of the same exponential covariance function (7.11) with a larger correlation length $L = 0.4$. Comparing with Figure (7.2), we notice that the eigenvalues indeed decay faster. The finite KLE captures 95% of the total energy of the random process with only ten terms.

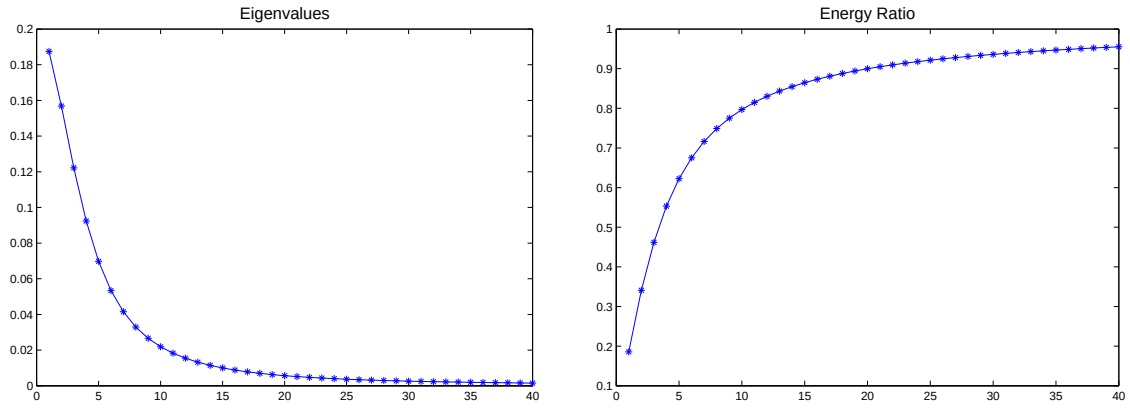

Figure 7.2: The first 40 eigenvalues of the KLE with the exponential covariance function with $L = 0.1$. The right plot is the energy ratio of the finite KLE with different number of terms. The KLE with 20 terms only captures 90% of the total energy of the random process. The KLE converges much slower for the exponential covariance function.

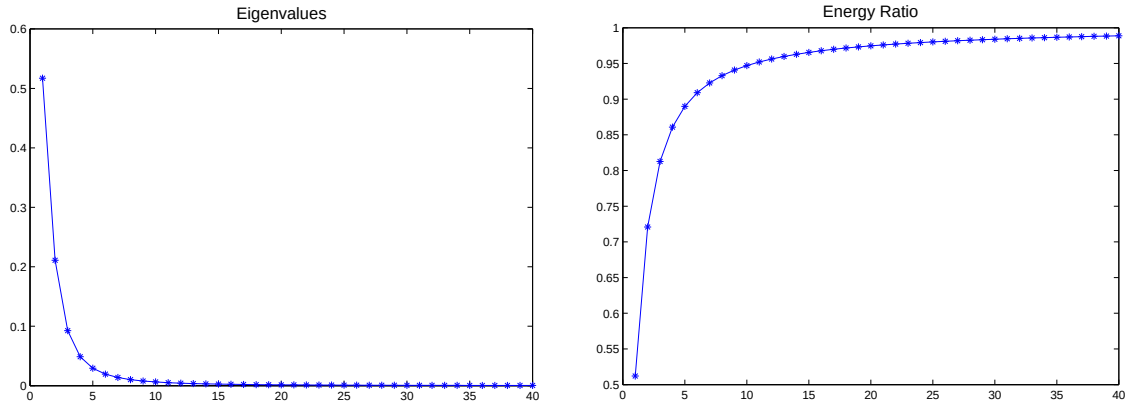

Figure 7.3: The first 40 eigenvalues of the exponential covariance function with $L = 0.4$; The right plot is the energy ratio of the finite KLE with different number of terms. For the same type of covariance function, the KLE converges much faster with larger correlation length.

7.2 WCE Methods for Stochastic Elliptic Equations

Consider the stochastic elliptic equation

$$\begin{cases} \nabla_x \cdot [k(x,\omega)\nabla_x P(x,w)] = q(x), \quad x \in D. \\ P(x,\omega) = 0 \quad \text{on} \quad \partial D, \end{cases} \tag{7.12}$$

where P is the pressure, k is the permeability field, and $q(x)$ is the (deterministic) source term. The random permeability field $k(x,\omega)$ is assumed to be

$$k(x,\omega) = k_0(x) + \delta\, e^{Y(x,\omega)}, \tag{7.13}$$

where $k_0(x)$ is a deterministic function, $\delta > 0$ is a constant, and $Y(x, w)$ is a homogeneous isotropic Gaussian field. If we think of $k_0(x)$ as the reference permeability field, then $\delta\, e^{Y(x,\omega)}$ is the random variability to model the uncertainty in the description of $k_0(x)$. We assume that $k_0(x)$ satisfies the uniform lower bound

$$0 < a_1 \leqslant k_0(x), \qquad \text{for } x \in D.$$

This is a reasonable assumption because the permeability field is always positive. Since $\delta > 0$, it follows immediately that

$$0 < a_1 \leqslant k_0(x) \leqslant k(x, \omega), \qquad \text{for} \quad x \in D,\ \omega \in \Omega. \tag{7.14}$$

So the random permeability field $k(x, \omega)$ is uniformly bounded from below.

Define the Hilbert space

$$H_0^1(D, k) = \left\{ v(x, \omega) :\ v|_{\partial D} = 0,\ E \int_D k(x, \omega) |\nabla_x v(x, \omega)|^2 d\, x\ < +\infty \right\}. \tag{7.15}$$

The inner product on the Hilbert space $H_0^1(D, k)$ is defined as

$$a(u, v) = E \int_D k(x, \omega) \nabla_x u(x, \omega) \nabla_x v(x, \omega) d\, x\,.$$

Denote

$$(q, v) = E \int_D q(x)\, v(x, \omega) d\, x\,.$$

The weak form of the problem (7.12) is proposed as: Find $P(x, \omega) \in H_0^1(D, k)$ such that

$$a(P, v) = (q, v), \qquad \forall\ v \in H_0^1(D, k). \tag{7.16}$$

We have the following result:

Theorem 7.1 *The elliptic problem (7.12) has a unique weak solution in the space $H_0^1(D, k)$.*

Proof We first prove that (q, v) is a bounded functional on the Hilbert space $H_0^1(D, k)$. From the Cauchy-Schwartz inequality,

$$(q, v)^2 \leqslant \int_D q^2(x)\, d\, x \int_D E[v^2(x, \omega)]\, d\, x\,.$$

Since $k(x,\omega)$ is bounded uniformly from below, we have

$$a_1 \int_D E|\nabla_x v(x,\omega)|^2 d\,x \leqslant E \int_D k(x,\omega)|\nabla_x v(x,\omega)|^2 d\,x. \tag{7.17}$$

Combining the Poincare inequality and (7.17), we can show that

$$\int_D E[v^2(x,\omega)] d\,x \lesssim \int_D E|\nabla_x v(x,\omega)|^2 d\,x \lesssim E \int_D k(x,\omega)|\nabla_x v(x,\omega)|^2 d\,x\,.$$

So (q,v) is indeed a bounded functional on the Hilbert space $H_0^1(D,k)$. Based on the Riesz lemma, there exists a unique $P \in H_0^1(D,k)$ such that $a(P,v) = (q,v)$. □

If the permeability $k(x,\omega)$ satisfies the strong elliptic condition (7.2), then a unique weak solution exists for (7.12) in the Hilbert space

$$H_0^1(D) = \left\{ v(x,\omega) :\ u|_{\partial D} = 0,\ E \int_D |\nabla v(x,\omega)|^2 d\,x\ < +\infty \right\}.$$

Since $Y(x,\omega)$ is a Gaussian process, the permeability (7.13) is not uniformly bounded from above. For any constant k_{max}, there is always a positive probability that $k(x,\omega) > k_{max}$. That is why the weak solution of (7.12) has to be proposed in the new space $H_0^1(D,k)$, which depends on the property of $k(x,\omega)$.

Since $Y(x,\omega)$ is a Gaussian process, its distribution is determined uniquely by its mean and covariance function. For simplicity, we assume $E[Y(x,\omega)] = 0$. Denote the covariance function of $Y(x,\omega)$ as $R(x,y) = E[Y(x)Y(y)]$. The typical covariance function used for $Y(x,\omega)$ is the Gaussian covariance function

$$R(x,y) = \sigma^2 \exp\left[-\frac{(x_1-y_1)^2}{2L_1^2} - \frac{(x_2-y_2)^2}{2L_2^2} \right], \tag{7.18}$$

or the exponential covariance function

$$R(x,y) = \sigma^2 \exp\left[-\frac{|x_1-y_1|}{L_1} - \frac{|x_2-y_2|}{L_2} \right]. \tag{7.19}$$

In either case, we always have

$$E[Y(x)]^2 = R(x,x) = \sigma^2.$$

So $Y(x,\omega)$ determined by (7.18) or (7.19) is indeed homogeneous in space.

Denote the Karhunen-Loeve expansion of $Y(x,\omega)$ as

$$Y(x,\omega) = \sum_{i=1}^{\infty} \sqrt{\mu_i}\ \theta_i(\omega)\psi_i(x), \tag{7.20}$$

where $\psi_i(x)$ and μ_i are eigenfunctions and eigenvalues of the integral equation

$$\int_D R(x,y)\,\psi_i(y)\,dy = \mu_i\,\psi_i(x).$$

Since $Y(x,\omega)$ is a Gaussian process, so θ_i are independent Gaussian random variables. Denote $\theta = (\theta_1, \theta_2, \ldots)$, we can define the Wick polynomials of θ as before

$$T_\alpha(\theta) = \prod_{i=1}^{\infty} H_{\alpha_i}(\theta_i),$$

where $\alpha = (\alpha_1, \alpha_2, \ldots)$ is a finite multi-index with integer component. Denote $\mathcal{J}$ as the set of all the finite multi-indices α. Since $k(x,\omega) = k_0(x) + \delta\, e^{Y(x,\omega)}$ has finite variance, we can expand the permeability field $k(x,\omega)$ into a Fourier-Hermite series of θ :

$$k(x,\omega) = \sum_{\alpha\in\mathcal{J}} k_\alpha(x) T_\alpha(\theta), \qquad k_\alpha(x) = E[k(x)T_\alpha(\theta)]. \tag{7.21}$$

With $Y(x,\omega)$ given by the KLE (7.20), we can derive an exact formula for the WCE coefficient $k_\alpha(x)$.

Lemma 7.1 *Suppose $k(x,\omega) = k_0(x) + \delta\, e^{Y(x,\omega)}$ has the WCE expansion (7.21), then the WCE coefficients $k_\alpha(x)$ are given by the formulae*

$$k_{\mathbf{0}}(x) = k_0(x) + \delta e^{\sigma^2/2}$$

and

$$k_\alpha(x) = \delta\, e^{\sigma^2/2} \prod_{i=1}^{\infty} \frac{\left[\sqrt{\mu_i}\psi_i(x)\right]^{\alpha_i}}{\sqrt{\alpha_i!}}, \qquad \text{for} \quad \alpha \neq \mathbf{0},$$

Proof For any fixed x, we have

$$\begin{aligned} E[e^{Y(x,\omega)}T_\alpha(\theta)] &= \prod_{i=1}^{\infty}\left(\int_{-\infty}^{+\infty} e^{\sqrt{\mu_i}\psi_i\theta_i}H_{\alpha_i}(\theta_i)\frac{1}{\sqrt{2\,\pi}}\,e^{-\frac{\theta_i^2}{2}}d\theta_i\right) \\ &= \prod_{i=1}^{\infty} e^{\frac{1}{2}\mu_i\psi_i^2}\int_{-\infty}^{+\infty} H_{\alpha_i}(\theta_i)\frac{1}{\sqrt{2\,\pi}}\,e^{-\frac{(\theta_i-\sqrt{\mu_i}\psi_i)^2}{2}}d\theta_i \\ &= \prod_{i=1}^{\infty} e^{\frac{1}{2}\mu_i\psi_i^2}E[H_{\alpha_i}(\theta_i+\sqrt{\mu_i}\psi_i)]. \end{aligned}$$

Note that $\sigma^2 = E(Y^2) = \sum \mu_i\psi_i^2$. Using Lemma 2.3, we have

$$E[e^{Y(x,\omega)}T_\alpha(\theta)] = e^{\sigma^2/2}\prod_{i=1}^{\infty}\frac{(\sqrt{\mu_i}\psi_i)^{\alpha_i}}{\sqrt{\alpha_i!}}.$$

Consequently, we have

$$k_{\mathbf{0}}(x) = E[k(x,\omega)] = k_0(x) + E[\delta\, e^{Y(x,\omega)}] = k_0(x) + \delta e^{\sigma^2/2},$$

and for $\alpha \neq \mathbf{0}$

$$k_\alpha(x) = E[\delta\, e^{Y(x,\omega)}T_\alpha(\theta)] = \delta\, e^{\sigma^2/2}\prod_{i=1}^{\infty}\frac{(\sqrt{\mu_i}\psi_i)^{\alpha_i}}{\sqrt{\alpha_i!}}$$

□

Lemma 7.1 provides an exact and explicit formula for the WCE coefficients of the permeability field. Once the KLE of $Y(x,\omega)$ is available, we can easily compute the WCE coefficients $k_\alpha(x)$. We will see later that the availability of an exact formula for k_α is very important in obtaining a well-posed WCE elliptic system of the problem (7.12).

Since the permeability field $k(x,\omega)$ depends on the Gaussian random variables $\theta_i, i = 1,2,\ldots$, so does the random solution $P(x,\omega)$. As we have shown, the random solution P belongs to the Hilbert space $H_0^1(D,k)$. Based on the uniform lower bound of $k(x,\omega)$ and the Poincare inequality, we can easily show that $E[P^2(x,\omega)] < \infty$ for almost all $x \in D$ (see the proof of Theorem 7.1). So $P(x,\omega)$ has a finite variance and admits a Wiener chaos expansion

$$P(x,\omega) = \sum_{\alpha\in\mathcal{J}} P_\alpha(x)\,T_\alpha(\theta). \tag{7.22}$$

Multiply both sides of (7.12) by $T_\alpha(\theta)$ and take expectations, and we have

$$\begin{aligned} E[q\, T_\alpha] &= \nabla \cdot E\left[(k\nabla P)\, T_\alpha\right] \\ &= \nabla \cdot E\left[k\Big(\sum_{\gamma\in\mathcal{J}} \nabla P_\gamma\, T_\gamma\Big) T_\alpha\right] \\ &= \nabla \cdot \sum_{\gamma\in\mathcal{J}} E[k\, T_\alpha\, T_\gamma]\nabla P_\gamma. \end{aligned}$$

Denote

$$A_{\alpha,\gamma}(x) = E[k\ T_\alpha\, T_\gamma].$$

Expanding the product of the Wick polynomial $T_\alpha\, T_\gamma$ by formula (2.33) and using the orthogonality of T_α, we get

$$\begin{aligned} A_{\alpha,\gamma}(x) &= \sum_{p\le\alpha\wedge\gamma} B(\alpha,\gamma,p)\ E\left(k\, T_{\alpha+\gamma-2p}\right) \\ &= \sum_{p\le\alpha\wedge\gamma} B(\alpha,\gamma,p)\ k_{\alpha+\gamma-2p}(x). \end{aligned}$$

Therefore, we have the following result:

Theorem 7.2 *The WCE coefficients $P_\alpha(x)$ of the expansion* (7.22) *satisfy the elliptic system*

$$\nabla \cdot \left(\sum_{\gamma\in\mathcal{J}} A_{\alpha,\gamma}(x)\nabla P_\gamma(x)\right) = qI_{\{\alpha=0\}}, \qquad \alpha\in\mathcal{J}, \tag{7.23}$$

where

$$A_{\alpha,\gamma}(x) = \sum_{p\le\alpha\wedge\gamma} B(\alpha,\gamma,p)\ k_{\alpha+\gamma-2p}(x). \tag{7.24}$$

The elliptic system (7.23) has infinite number of equations. For numerical purposes, we need to truncate it to finite terms. However, the truncation should be done very carefully. For the stochastic elliptic equation (7.12), it is critical that the permeability $k(x,\omega)$ satisfy the uniform lower bound condition (7.14). Otherwise, the equation (7.12) can be ill-posed and no solution would exist. When truncating the WCE system (7.23), we have to preserve the lower bound condition (7.14) for elliptic system (7.23). In the literature [3, 4, 32], the

authors assume a Karhunen-Loeve expansion for the permeability directly:

$$k(x,\omega) = k_0(x) + \sum_{i=1}^{\infty} \sqrt{\mu_i}\,\theta_i \psi_i(x). \tag{7.25}$$

Then they truncate the KLE of $k(x,\omega)$ and obtain a finite dimensional permeability field $k_M(x,\omega)$. At the end, they propose to solve the following approximate elliptic equation:

$$\nabla_x \cdot [k_M(x,\omega)\nabla_x P(x,w)] = q(x). \tag{7.26}$$

Since $k_M(x,\omega)$ has only a finite number of random variables, they can apply the polynomial chaos expansion to solve it numerically.

A problem with the above approach is that the truncated permeability $k_M(x,\omega)$ may no longer satisfy the elliptic condition (7.2), even though the original permeability $k(x,\omega)$ does. For $k_M(x,\omega)$ to satisfy a condition similar to (7.2), the KLE of $k(x,\omega)$ has to converge uniformly both in x and in ω variables so that the truncation error can be controlled. In lots of practical cases, the random variables θ_k are not uniformly bounded, such as the Gaussian random variables or exponential random variables, and the KLE does not converge uniformly. For any constant $k_{\min}$, there is always a positive probability that $k_M(x,\omega) < k_{\min}$. It has been shown by Babuska et al. [5] that even condition (7.2) is only violated with a very small but positive probability, system (7.26) can still be ill-posed. For those cases, the approach adopted by [32, 4, 3] is not readily applicable.

If we look at the infinite elliptic system (7.23), it becomes clear that there is no need to truncate the permeability field at all. We only need to truncate the WCE of the random solution $P(x,\omega)$. Denote

$$P_{K,N}(x,\omega) = \sum_{\mathcal{J}_{K,N}} k_\alpha(x) T_\alpha(\theta_i),$$

where $\mathcal{J}_M$ is a finite (truncated) subset of the index set $\mathcal{J}$. As a result, the full elliptic system (7.23) is reduced to the finite system

$$\nabla \cdot \left(\sum_{\gamma \in \mathcal{J}_M} A_{\alpha,\gamma}(x) \nabla P_\gamma(x) \right) = q I_{\{\alpha=0\}}, \qquad \alpha \in \mathcal{J}_M. \tag{7.27}$$

However, to get the finite system (7.27) we never need to truncate the expansion of $k(x,\omega)$. For any given index pair (α,γ), the formula for $A_{\alpha,\gamma}(x)$ is still $A_{\alpha,\gamma}(x) = E[kT_\alpha T_\gamma]$, instead

of $A_{\alpha,\gamma}(x) = E[k_M T_\alpha T_\gamma]$. Using the analytical formula for products of Wick polynomials, we can obtain an explicit formula (7.24) for $A_{\alpha,\gamma}(x)$. Even though $k(x,\omega)$ is an infinite series, the formula for $A_{\alpha,\gamma}(x)$ is a finite summation and is easily computable from the coefficients k_α. The summation in (7.24) will include all the eligible terms even if the index $\alpha+\gamma-2p \notin \mathcal{J}_M$. That is the precise difference between truncating and not truncating the permeability $k(x,\omega)$.

Both Matthies et al. [67] and Roman et al. [89] pointed out that there is no need to truncate the WCE expansion (7.21) or the KLE expansion (7.25) for the permeability field. A new ingredient of the current work is an analytical and explicit formula (7.24) for the stiffness coefficients $A_{\alpha,\beta}$. The formula (7.24) clearly illustrates that even though $k(x,\omega)$ is an infinite series, the stiffness coefficients $A_{\alpha,\gamma}$ are still computable, hence no truncation of the permeability is needed.

While deriving the finite elliptic system (7.12), it is important that we have an exact and easily computable formula (Lemma 7.1) for the WCE coefficients $k_\alpha(x)$. That will avoid any other forms of truncations of the permeability field. In the papers [48, 104], the authors used Taylor expansion of $e^{Y(x,\theta)}$ to estimate the WCE coefficients $k_\alpha(x)$. When they truncated the Taylor expansion, they essentially truncated the permeability field, which will inevitably destroy the lower bound condition (7.14).

7.2.1 Property of the Truncated Elliptic Systems

It is important that the truncated WCE elliptic system (7.27) preserves the properties of the original elliptic equation. Note that $A_{\alpha\beta}(x) = E\,(k\,T_\alpha\,T_\beta)$, it is obvious that

Lemma 7.2 *The elliptic system (7.27) is symmetric:*

$$A_{\alpha,\beta}(x) = A_{\beta,\alpha}(x).$$

Since the original stochastic elliptic equation is positive definite (uniformly bounded from below), we expect that the elliptic system (7.27) is also positive definite. Indeed we have

Lemma 7.3 *For any $u_\alpha \in R^d$, $\alpha \in \mathcal{J}_M$, we have*

$$0 < a_1|u|^2 \leqslant \sum_{\alpha,\beta\in\mathcal{J}_M} u_\alpha^T A_{\alpha,\beta}(x) u_\beta, \qquad x \in D,$$

where $|u|^2 = \sum_{\alpha\in\mathcal{J}_M} |u_\alpha|^2$ *and* a_1 *is the same lower bound (7.14) for* $k(x,\omega)$*. Therefore, the elliptic system (7.27) is symmetric and positive definite.*

Proof From the formula of $A_{\alpha,\beta}(x)$ we have

$$\begin{aligned}
\sum_{\alpha,\beta} u_\alpha^T\, A_{\alpha,\beta}(x)\, u_\beta &= \sum_{\alpha,\beta} u_\alpha^T E\,(k\, T_\alpha\, T_\beta)\, u_\beta \\
&= E\Big[\Big(\sum_\alpha u_\alpha\, T_\alpha\Big)^T k\Big(\sum_\beta u_\beta\, T_\beta\Big)\Big] \\
&\geqslant a_1\, E\Big[\Big(\sum_\alpha u_\alpha\, T_\alpha\Big)^T \Big(\sum_\beta u_\beta\, T_\beta\Big)\Big] \\
&= a_1\, |u|^2
\end{aligned}$$

□

As we can see, the truncated elliptic system (7.27) preserves the lower bound (7.14) since we did not truncate the permeability field. So the truncated elliptic system (7.27) is well-posed.

7.3 Upscaling of the WCE Elliptic System

Since the permeability field $k(x,\omega)$ usually has multi-scale structures, so do the stiffness coefficients $A_{\alpha,\beta}(x)$. In practice, it is impossible and unnecessary to resolve all the scales in the solution. In stead, we can capture the net effects of the small scales by deriving an upscaled equation. In this section, we will generalize the upscaling formulation for a single elliptic equation to the WCE elliptic system (7.27).

Consider the elliptic system

$$\frac{\partial}{\partial x_i}\left[A_{\alpha\beta}^{ij}\left(x,\frac{x}{\varepsilon}\right)\frac{\partial}{\partial x_j}P_\beta(x)\right] = 0. \tag{7.28}$$

In the above formula, the summation over the repeated indices is assumed. Denote $y = x/\varepsilon$. We assume that $x \in \Omega$ and $y \in Y$, and $A_{\alpha\beta}(x,y)$ is periodic in y. Expand the solutions in multiscales:

$$P_\beta = P_\beta^0\left(x,\frac{x}{\varepsilon}\right) + \varepsilon P_\beta^1\left(x,\frac{x}{\varepsilon}\right) + \varepsilon^2 P_\beta^2\left(x,\frac{x}{\varepsilon}\right) + \ldots. \tag{7.29}$$

Note that for any function $f\left(x, \frac{x}{\varepsilon}\right)$,

$$\frac{\partial}{\partial x} f\left(x, \frac{x}{\varepsilon}\right) = \frac{\partial}{\partial x} f(x,y) + \frac{1}{\varepsilon}\frac{\partial}{\partial y} f(x,y).$$

Applying the chain rule to (7.28), we get

$$\left[\frac{\partial}{\partial x_i} + \frac{1}{\varepsilon}\frac{\partial}{\partial y_i}\right]\left\{A^{ij}_{\alpha\beta}(x,y)\left[\frac{\partial}{\partial x_j} + \frac{1}{\varepsilon}\frac{\partial}{\partial y_j}\right] P_\beta(x,y)\right\} = 0.$$

Substitute the multiscale expansion (7.29) of P_β into the above equation, we get

$$\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial x_j}P^0_\beta(x,y)\right) + \frac{1}{\varepsilon}\frac{\partial}{\partial y_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial x_j}P^0_\beta(x,y)\right) +$$

$$\frac{1}{\varepsilon}\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^0_\beta(x,y)\right) + \frac{1}{\varepsilon^2}\frac{\partial}{\partial y_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^0_\beta(x,y)\right) +$$

$$\varepsilon\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial x_j}P^1_\beta(x,y)\right) + \frac{\partial}{\partial y_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial x_j}P^1_\beta(x,y)\right) +$$

$$\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^1_\beta(x,y)\right) + \frac{1}{\varepsilon}\frac{\partial}{\partial y_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^1_\beta(x,y)\right) +$$

$$\varepsilon^2\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial x_j}P^2_\beta(x,y)\right) + \varepsilon\frac{\partial}{\partial y_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial x_j}P^2_\beta(x,y)\right) +$$

$$\varepsilon\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^2_\beta(x,y)\right) + \frac{\partial}{\partial y_i}\left(A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^2_\beta(x,y)\right) + \cdots = 0.$$

Collect all the ε^{-2} terms we get

$$\frac{\partial}{\partial y_i}\left[A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^0_\beta(x,y)\right] = 0.$$

Multiply both sides by $P^0_\alpha(x,y)$ and integration by part. From the periodicity in Y and Lemma 7.3, we have

$$0 = \int_Y A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^0_\beta(x,y)\frac{\partial}{\partial y_i}P^0_\alpha(x,y)\,dy \geq a_1\int_Y \sum_\alpha \left|\frac{\partial}{\partial y_i}P^0_\alpha(x,y)\right|^2 dy.$$

As a result, we have

$$P^0_\beta(x,y) = P^0_\beta(x). \tag{7.30}$$

Without Lemma 7.3, we cannot reach such a conclusion. So it is indeed crucial to preserve the positiveness for the truncated elliptic system.

By collecting the ε^{-1} terms, we get

$$\frac{\partial}{\partial y_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial x_j}P_\beta^0(x,y)\right]+\frac{\partial}{\partial x_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial y_j}P_\beta^0(x,y)\right]$$
$$+\frac{\partial}{\partial y_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial y_j}P_\beta^1(x,y)\right]=0.$$

Because of (7.30), the second term in the above formula is zero and we can separate the first term. Consequently, we have

$$\frac{\partial}{\partial y_i}A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial x_j}P_\beta^0(x)+\frac{\partial}{\partial y_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial y_j}P_\beta^1(x,y)\right]=0.$$

Assume that

$$P_\beta^1(x,y)=N_{\beta\gamma}^j(x,y)\frac{\partial}{\partial x_j}P_\gamma^0(x), \tag{7.31}$$

then $N_{\beta\gamma}^j$ satisfies the following equation:

$$\frac{\partial}{\partial y_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial y_j}N_{\beta\gamma}^k(x,y)\right]=-\frac{\partial}{\partial y_i}A_{\alpha\beta}^{ik}(x,y). \tag{7.32}$$

By collecting the ε^0 terms, we have

$$\frac{\partial}{\partial x_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial x_j}P_\beta^0(x)\right]+\frac{\partial}{\partial x_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial y_j}P_\beta^1(x,y)\right]$$
$$+\frac{\partial}{\partial y_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial x_j}P_\beta^1(x,y)\right]+\frac{\partial}{\partial y_i}\left[A_{\alpha\beta}^{ij}(x,y)\frac{\partial}{\partial y_j}P_\beta^2(x,y)\right]=0.$$

Let $\langle\cdot\rangle$ be the spatial average over Y, and note that

$$\langle\nabla_y\cdot f\rangle=\int_{\partial Y}f(x,y)\cdot n\,dl=0$$

if $f(x,y)$ is periodic in y. If we average the above equation over the period Y, the last two

terms will disappear. Plugging in formula (7.31), we have

$$\begin{aligned}0 =&\frac{\partial}{\partial x_i}\left[\left\langle A^{ij}_{\alpha\beta}(x,y)\right\rangle \frac{\partial}{\partial x_j}P^0_\beta(x)\right]+\frac{\partial}{\partial x_i}\left\langle A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}P^1_\beta(x,y)\right\rangle\\ =&\frac{\partial}{\partial x_i}\left[\left\langle A^{ij}_{\alpha\beta}(x,y)\right\rangle \frac{\partial}{\partial x_j}P^0_\beta(x)\right]+\frac{\partial}{\partial x_i}\left\langle A^{ij}_{\alpha\beta}(x,y)\frac{\partial}{\partial y_j}N^k_{\beta\gamma}(x,y)\frac{\partial}{\partial x_k}P^0_\gamma(x)\right\rangle.\end{aligned}$$

Denote

$$\left(A^{ij}_{\alpha\beta}\right)^\star=\left\langle A^{ij}_{\alpha\beta}(x,y)\right\rangle+\left\langle A^{ik}_{\alpha\gamma}(x,y)\frac{\partial}{\partial y_k}N^j_{\gamma\beta}(x,y)\right\rangle, \tag{7.33}$$

then the previous equation for the ε^0 terms is reduced to

$$\frac{\partial}{\partial x_i}\left[\left(A^{ij}_{\alpha\beta}\right)^\star\frac{\partial}{\partial x_j}P^0_\beta(x)\right]=0. \tag{7.34}$$

Equation (7.34) is the upscaled equation for the leading order terms P^0_β, with $\left(A^{ij}_{\alpha\beta}\right)^\star$ as the effective stiffness coefficients.

In order to compute the effective coefficients $\left(A^{ij}_{\alpha\beta}\right)^\star$, we must first find $N^k_{\beta\gamma}(x,y)$. In practice, the coefficients $A^{ij}_{\alpha\beta}(x)$ have no explicit separation of scales, and they are not periodic either. To apply the above upscaling formulations, we partition the computational domain Ω into coarse blocks $\Omega=\bigcup_{k=1}^K\Omega_k$. Then we solve the cell problem (7.32) on each coarse block Ω_k:

$$\frac{\partial}{\partial x_i}\left[A^{ij}_{\alpha\beta}(x)\frac{\partial}{\partial x_j}N^k_{\beta\gamma}(x)+A^{ik}_{\alpha\beta}(x)\right]=0,\qquad x\in\Omega_k.$$

If we write $A^{ik}_{\alpha\beta}=A^{ij}_{\alpha\gamma}\delta_{\gamma\beta}\delta_{jk}$ then we can simplify the above equation as

$$\frac{\partial}{\partial x_i}\left[A^{ij}_{\alpha\beta}\left(\frac{\partial}{\partial x_j}N^k_{\beta\gamma}+\delta_{\beta\gamma}\delta_{jk}\right)\right]=\frac{\partial}{\partial x_i}\left[A^{ij}_{\alpha\beta}\frac{\partial}{\partial x_j}\left(N^k_{\beta\gamma}(x,y)+\delta_{\beta\gamma}x_k\right)\right]=0.$$

Therefore, we only need to solve the following local problems

$$\frac{\partial}{\partial x_i}\left(A^{ij}_{\alpha\beta}\frac{\partial}{\partial x_j}\varphi_{\beta\gamma}\right)=0,\qquad x\in\Omega_k \tag{7.35}$$

with certain boundary conditions. Once $\varphi_{\beta\gamma}$ is available, we can compute the effective

stiffness coefficients on each block by the local average

$$\left(A^{ij}_{\alpha\beta}\right)^{\star}_{\Omega_k} = \left\langle A^{ij}_{\alpha\beta}(x)\,\frac{\partial}{\partial x_j}\varphi_{\beta\gamma}\right\rangle_{\Omega_k}. \tag{7.36}$$

Chapter 8

Uncertainty Quantification in Reservoir Modeling

8.1 Introduction

Uncertainties in the detailed description of reservoir lithofacies, porosity, and permeability are major contributors to uncertainty in reservoir performance forecasting. Making decisions in reservoir management requires efficient ways for quantifying the uncertainty. Large uncertainties in reservoirs can greatly affect the production and decision making on well drilling. Thus, quantifying and reducing the uncertainty is an important problem in subsurface modeling. Additional dynamic data, such as the production data, can be used to improve the predictions and reduce the uncertainty. To predict future reservoir performance, the reservoir properties, such as porosity and permeability, need to be conditioned to dynamic data. In general it is difficult to calculate this posterior probability distribution because the process of predicting flow and transport in petroleum reservoirs is nonlinear. Instead, people estimate this probability distribution from the outcomes of flow predictions for a large number of realizations of the reservoir. It is essential that the permeability (and porosity) realizations adequately reflect the uncertainty in the reservoir properties, i.e., we correctly sample this probability distribution.

The prediction of permeability fields based on dynamic data is a challenging problem because permeability fields are typically defined on a large number of grid blocks. The Markov chain Monte Carlo (MCMC) method and its modifications have been used previously to sample the posterior distribution of the permeability field. Oliver et al. [78, 79] proposed the randomized maximum likelihood (RML) method, which generates uncondi-

tional realizations of the production and permeability data and then solves a deterministic gradient-based inverse problem. The solution of this minimization problem is taken as a proposal and accepted with probability 1 because the rigorous acceptance probability is very difficult to estimate. In addition to the need of solving a gradient-based inverse problem, this method does not guarantee a proper sampling of the posterior distribution. Developing efficient and rigorous MCMC calculations with high acceptance rates remains a challenging problem.

In this chapter, we study a two-stage MCMC method for sampling the permeability fields, where coarse-scale models of the fluid flows are used to increase the acceptance rate of the MCMC simulations. In the two-stage MCMC method, we first test each MCMC proposal by a coarse-scale model. If the proposal is accepted at the first-stage, then a fine-scale simulation will be performed to determine the acceptance probability of the proposal. The first stage of the MCMC method filters unlikely proposals and avoids expensive fine-scale computations for them. Compared with the regular MCMC method, the two-stage MCMC method generates a modified Markov chain by incorporating the coarse-scale information of the problem. Under some technical assumptions, we can prove that the modified Markov chain is ergodic and converges to the correct posterior distribution. The validity of the assumptions for our application is discussed in the paper. Our numerical results show that the two-stage MCMC method can be ten times more efficient than the regular MCMC method. We would like to note that two-stage MCMC algorithms have been used previously (e.g., [14, 44, 58, 80]) in different situations.

In the second part, we employ the Langevin algorithms to generate proposals for the two-stage MCMC methods. Langevin algorithms provide efficient sampling techniques because they use the gradient information of the target distributions. In fact, the popular RML sampler designed by Oliver et al. [78, 79] can be regarded a variation of the Langevin algorithm. However, the direct Langevin algorithm is very expensive because the computation of the distribution gradients requires solving the nonlinear PDE on fine scales many times. Alternatively, we can compute the Langevin proposal using the coarse gradient of the target distribution. Using the upscaled WCE method, we can represent the random solution as a stochastic spectral expansion and obtain an explicit formula for the coarse-scale target distribution. Then we can generate the Langevin proposals easily using this explicit coarse-scale distributions, without the need for solving the nonlinear PDE repeatedly. It

is important that we have derived an upscaling formulation for the WCE elliptic equations so that it can be solved efficiently on the coarse-grid. The Langevin algorithms can be combined with the two-stage MCMC method, and hence provide a very efficient way for sampling the permeability fields.

8.2 The Model Equations and Problem Setting

We consider two-phase flows in a domain Ω under the assumption that the displacement is dominated by viscous effects. We neglect the effects of gravity, compressibility, and capillary pressure. The two phases are referred to as water (aqueous phase) and oil (nonaqueous phase liquid), designated by subscripts w and o, respectively. We write Darcy's law, with all quantities dimensionless, for each phase as follows:

$$\boldsymbol{v}_j = -\frac{k_{rj}(S)}{\mu_j} k \cdot \nabla p, \tag{8.1}$$

where $\boldsymbol{v}_j$, $j = w, o$, is the phase velocity, k is the permeability tensor, k_{rj} is the relative permeability of the phase j, S is the water saturation (volume fraction), and p is the pressure. In this work, a single set of relative permeability curves is used and k is taken to be a diagonal tensor. Combining Darcy's law with a statement of conservation of mass allows us to express the governing equations in terms of the so-called pressure and saturation equations:

$$\nabla \cdot (\lambda(S) k \nabla p) = q, \tag{8.2}$$

$$\frac{\partial S}{\partial t} + \boldsymbol{v} \cdot \nabla f(S) = -q_w, \tag{8.3}$$

where $\lambda(S)$ is the total mobility, q and q_w are the source terms, $\boldsymbol{v}$ is the total velocity, and $f(S)$ is the flux function, which are respectively given by

$$\lambda(S) = \frac{k_{rw}(S)}{\mu_w} + \frac{k_{ro}(S)}{\mu_o}, \tag{8.4}$$

$$\boldsymbol{v} = \boldsymbol{v}_w + \boldsymbol{v}_o = -\lambda(S) k \nabla p, \tag{8.5}$$

$$f(S) = \frac{k_{rw}(S)/\mu_w}{k_{rw}(S)/\mu_w + k_{ro}(S)/\mu_o}. \tag{8.6}$$

The above description is referred to as the fine model of the two-phase flow problem. For the single-phase flow, we have $\lambda(S) = 1$ and $f(S) = S$. Throughout, the porosity is assumed to be constant.

The dynamical data of the two-phase flow problem include the fractional flow curve, defined as

$$F(t) = 1 - \frac{\int_{\partial\Omega^{out}} f(S)\, v_n \, dl}{\int_{\partial\Omega^{out}} v_n \, dl}, \tag{8.7}$$

where $\partial\Omega^{out}$ is the outflow boundary and $v_n = \boldsymbol{v} \cdot \boldsymbol{n}$ is the normal velocity on the boundary. The fractional flow $F(t)$ (denoted simply by F thereafter) is the fraction of oil in the produced fluid and is observable from the production data. Based on geological surveys, the permeability field $k(x)$ is usually known at some sparse location $k(x_i) = k_i$, $i = 1, 2, \ldots, N$. The problem under consideration is: *How do we sample the permeability fields $k(x)$ based on the fractional flow F and the static data k_k?*. Since the fractional flow is an integrated response, the map from the permeability field to the fractional flow is not one-to-one. Hence this problem is ill-posed in the sense that there exist many different permeability realizations for the given production data.

From the probabilistic point of view, this problem can be regarded as sampling the permeability field conditioning on the fractional flow data with measurement errors. Consequently, our goal is to sample from the conditional distribution $P(k|F)$, where k is the fine-scale permeability field and F is the fractional flow curve measured from the production data. Using the Bayes formula we can write

$$P(k|F) \propto P(F|k)P(k). \tag{8.8}$$

In the above formula, $P(k)$ is the prior distribution of the permeability field, which is assumed to be known. The prior distribution $P(k)$ will also incorporate the information of the permeability field at the sparse locations. The likelihood function $P(F|k)$ denotes the conditional probability that the outcome of the measurement is F when the true permeability is k.

In practice, the measured fractional flow F contains measurement errors. Denote the fractional flow for a given k as F_k, which can be computed by solving the model equation (8.1)–(8.3) on the fine-grid. The computed F_k will contain a modeling error as well as a numerical error. In this paper, we assume that the combined errors from the measurement, modeling, and numerics satisfy a Gaussian distribution. That is, the likelihood function $P(F|k)$ takes the form

$$P(F|k) \propto \exp\Big(-\frac{\|F - F_k\|^2}{\sigma_f^2}\Big), \tag{8.9}$$

where F is the observed fractional flow, F_k is the fractional flow computed by solving the model equations (8.1)–(8.3) on the fine-grid for a given k, and σ_f is the precision associated with the measurement F and the numerical solution F_k. Since both F and F_k are functions of t, $\|F - F_k\|^2$ denotes the L_2 norm

$$\|F - F_k\|^2 = \int_0^T [F(t) - F_k(t)]^2 \, dt.$$

It is worth noting that the method discussed in this paper does not depend on the specific form of the error functions. A more general error model can be used in the simulations. We would like to emphasize that different permeability fields may produce the same fractional flow curve. Thus, the likelihood distribution $P(F|k)$ is a multi-modal function of k with multiple local maxima.

Denote the posterior distribution as

$$\pi(k) = P(k|F) \propto \exp\Big(-\frac{\|F - F_k\|^2}{\sigma_f^2}\Big) P(k). \tag{8.10}$$

Sampling from the distribution $\pi(k)$ can be accomplished by using the Markov chain Monte Carlo (MCMC) method. The main idea of MCMC method is to generate a Markov chain with $\pi(k)$ as its stationary distribution. A key step to this approach is to construct the desired transition kernel for the Markov chain. In this paper, we use the Metropolis-Hasting

algorithm. Suppose $q(y|x)$ is a general transitional probability distribution, which is easy to sample and has an explicit form. The Metropolis-Hasting MCMC algorithm (see, e.g., [88]) consists of the following steps.

Algorithm I (Metropolis-Hasting MCMC [88])

- Step 1. At state k_n generate k from $q(k|k_n)$.
- Step 2. Accept k as a sample with probability

$$p(k_n, k) = \min\left(1,\ \frac{q(k_n|k)\pi(k)}{q(k|k_n)\pi(k_n)}\right), \tag{8.11}$$

  i.e., take $k_{n+1} = k$ with probability $p(k_n, k)$, and $k_{n+1} = k_n$ with probability $1 - p(k_n, k)$.

Starting with an arbitrary initial permeability sample k_0, the MCMC algorithm generates a Markov chain $\{k_n\}$. At each iteration, the probability of moving from state k_n to a next state k is $q(k|k_n)p(k_n, k)$, so the transition kernel for the Markov chain $\{k_n\}$ is

$$K(k_n, k) = p(k_n, k)q(k|k_n) + \Big(1 - \int p(k_n, k)q(k|k_n)dk\Big)\delta_{k_n}(k).$$

Using the explicit formula of the transition kernel, it is not difficult to prove that the target distribution $\pi(k)$ is indeed the stationary distribution of the Markov chain $\{k_n\}$. As a result, we can take k_n as samples of the distribution $\pi(k)$ after the chain reaches steady state.

8.3 Preconditioning MCMC Simulations Using Coarse-scale Models

8.3.1 The Preconditioned MCMC method

In the above Metropolis-Hasting MCMC algorithm, the major computational cost is to compute F_k in the target distribution $\pi(k)$, which involves solving the coupled nonlinear PDE system (8.1)–(8.3) on the fine-grid. Generally, the MCMC method requires thousands of iterations before it converges to the steady state. To quantify the uncertainty of the

permeability field accurately, one also needs to generate a large number of different samples. However, the acceptance rate of the direct MCMC method is very low, due to the large dimensionality of the permeability field. The algorithm needs to test many proposals to accept only a few permeability samples. Most of the CPU time is spent on simulating the rejected samples. That makes the direct (full) MCMC simulations prohibitively expensive in practice.

One way to improve the direct MCMC method is to increase its acceptance rate by modifying the proposal distribution $q(k|k_n)$. In this work, we propose an algorithm in which the proposal distribution $q(k|k_n)$ is adapted to the target distribution using the coarse-scale model. Instead of testing each proposal by fine-scale computations directly, the algorithm first tests the proposal by the coarse-scale model. This is achieved by comparing the fractional flow curves on the coarse grid first. If the proposal is accepted by the coarse-scale test, then a full fine-scale computation will be conducted and the proposal will be further tested as in the direct MCMC method. Otherwise, the proposal will be rejected by the coarse-scale test and a new proposal will be generated from $q(k|k_n)$. The coarse-scale test filters the unacceptable proposals and avoids the expensive fine-scale tests for those proposals. The filtering process essentially modifies the proposal distribution $q(k|k_n)$ by incorporating the coarse-scale information of the problem. That is why the modified method is called a **preconditioned MCMC method**.

Recall that the fine-scale target distribution is given by (8.10). We approximate the distribution $\pi(k)$ on the coarse scale by

$$\pi^*(k) \propto \exp\Big(-\frac{\|F - F_k^*\|^2}{\sigma_c^2}\Big) P(k), \tag{8.12}$$

where F_k^* is the fractional flow computed by solving a coarse-scale model of (8.1)–(8.3) for the given k, and σ_c is the precision associated with the coarse-scale model. The proposed coarse-scale model consists of upscaling the pressure equation (8.2) to obtain the velocity field on the coarse-grid, and then using it in (8.3) to resolve the saturation on the coarse-grid. The pressure equation is upscaled using the multiscale finite element method. The details of the method are presented in Appendix C. Since no subgrid modeling is performed for the saturation equation, this upscaling procedure introduces errors. In Figure 8.1, we present a comparison of the typical fractional flows computed by fine- and coarse-scale models.

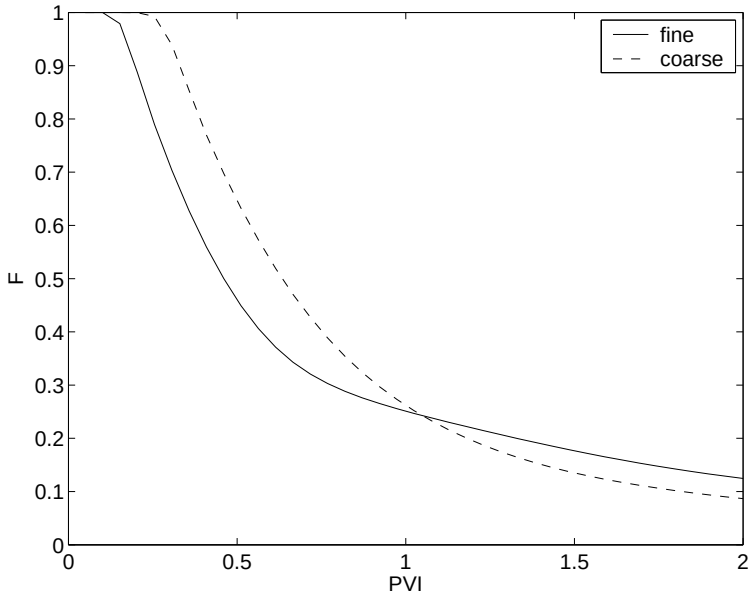

Figure 8.1: Typical fine- and coarse-scale fractional flows.

The fractional flows are plotted against the dimensionless time "pore volume injected" (PVI). The pore volume injected (PVI) at time T is defined as $\frac{1}{V_p}\int_0^T q_t(\tau)d\tau$, where q_t is the combined flow rates of water and oil at the production edge, and V_p is the total pore volume of the system. PVI provides the dimensionless time for the flow displacement.

Using the coarse-scale distribution $\pi^*(k)$ as a filter, the preconditioned MCMC can be described as follows:

Algorithm II (preconditioned MCMC)

- Step 1. At k_n, generate a trial proposal k' from distribution $q(k'|k_n)$.
- Step 2. Take the real proposal as

$$k = \begin{cases} k' & \text{with probability } g(k_n, k'), \\ k_n & \text{with probability } 1 - g(k_n, k'), \end{cases}$$

where

$$g(k_n, k') = \min\left(1, \frac{q(k_n|k')\pi^*(k')}{q(k'|k_n)\pi^*(k_n)}\right). \tag{8.13}$$

Therefore, the final proposal k is generated from the effective instrumental distribution

$$Q(k|k_n) = g(k_n, k)q(k|k_n) + \Big(1 - \int g(k_n, k)q(k|k_n)dk\Big)\delta_{k_n}(k). \tag{8.14}$$

- Step 3. Accept k as a sample with probability

$$\rho(k_n, k) = \min\left(1, \frac{Q(k_n|k)\pi(k)}{Q(k|k_n)\pi(k_n)}\right), \tag{8.15}$$

  i.e., $k_{n+1} = k$ with probability $\rho(k_n, k)$, and $k_{n+1} = k_n$ with probability $1 - \rho(k_n, k)$.

In the above algorithm, if the trial proposal k' is rejected by the coarse-scale test (Step 2), k_n will be passed to the fine-scale test as the proposal. Since $\rho(k_n, k_n) \equiv 1$, no further (fine-scale) computation is needed. Thus, the expensive fine-scale computations can be avoided for those proposals that are unlikely to be accepted. In comparison, the regular MCMC method requires a fine-scale simulation for every proposal k, even though most of the proposals will be rejected at the end.

It is worth noting that there is no need to compute $Q(k|k_n)$ and $Q(k_n|k)$ in (8.15) by the integral formula (8.14). The acceptance probability (8.15) can be simplified as

$$\rho(k_n, k) = \min\left(1, \frac{\pi(k)\pi^*(k_n)}{\pi(k_n)\pi^*(k)}\right). \tag{8.16}$$

In fact, (8.16) is obviously true for $k = k_n$ since $\rho(k_n, k_n) \equiv 1$. For $k \neq k_n$,

$$\begin{aligned} Q(k_n|k) = g(k, k_n)q(k_n|k) &= \frac{1}{\pi^*(k)} \min\Big(q(k_n|k)\pi^*(k),\ q(k|k_n)\pi^*(k_n)\Big) \\ &= \frac{q(k|k_n)\pi^*(k_n)}{\pi^*(k)} g(k_n, k) = \frac{\pi^*(k_n)}{\pi^*(k)} Q(k|k_n). \end{aligned}$$

Substituting the above formula into (8.15), we immediately get (8.16).

Since the computation of the coarse-scale solution is very cheap, Step 2 of the preconditioned MCMC method can be implemented very fast to decide whether or not to run the fine-scale simulation. The second step of the algorithm serves as a filter that avoids unnecessary fine-scale runs for the rejected samples. It is possible that the coarse-scale test may reject an individual sample that will otherwise have a (small) probability to be accepted in the fine-scale test. However, that doesn't play a crucial role, since we are only interested

in the statistical property of the samples. As we will show later that the preconditioned MCMC algorithm converges under some mild assumptions.

We would like to note that the Gaussian error model for the coarse-scale distribution $\pi^*(k)$ is not very accurate. We only use it in the filtering stage to decide whether or not to run the fine-scale simulations. The choice of the coarse-scale precision parameter σ_c is important for increasing the acceptance rate. If σ_c is too large, then too many proposals can pass the coarse-scale tests and the filtering stage will become less effective. If σ_c is too small, then eligible proposals may be incorrectly filtered out, which will result in biased sampling. Our numerical results show that the acceptance rate is optimal when σ_c is of the same order as σ_f. The optimal value of σ_c can be estimated based on the correlation between $\|F - F_k\|$ and $\|F - F_k^*\|$ (cf. Figure 8.2).

Based on the Gaussian precision models (8.10) and (8.12), the acceptance probability (8.16) has the form

$$\rho(k_n, k) = \min\left(1, \frac{\pi(k)\pi^*(k_n)}{\pi(k_n)\pi^*(k)}\right) = \min\left(1, \frac{\exp\left(-\frac{E_k - E_{k_n}}{\sigma_f^2}\right)}{\exp\left(-\frac{E_k^* - E_{k_n}^*}{\sigma_c^2}\right)}\right), \tag{8.17}$$

where

$$E_k = \|F - F_k\|^2, \qquad E_k^* = \|F - F_k^*\|^2.$$

If E_k^* is strongly correlated with E_k, then the acceptance probability (8.17) could be close to 1 for certain choices of σ_c. Hence a high acceptance rate can be achieved at Step 3 of the preconditioned MCMC method. To demonstrate that E_k^* is indeed strongly correlated with E_k, we simulate E_k and E_k^* for many different permeability samples k, and plot E_k against E_k^* in Figure 8.2. We find that the correlation coefficient between E_k^* and E_k is approximately 0.9. If the correlation between E_k and E_k^* is strong, we can write

$$E_k \simeq \alpha E_k^* + \beta.$$

Substituting this into (8.17) and choosing $\sigma_c^2 = \sigma_f^2/\alpha$, we can obtain the acceptance rate close to 1 in Step 3. In practice, however, one does not know *a priori* the correlation constant α. The approximate value of α can be estimated by *a priori* numerical simulations where E_k and E_k^* are computed for a number of permeability samples.

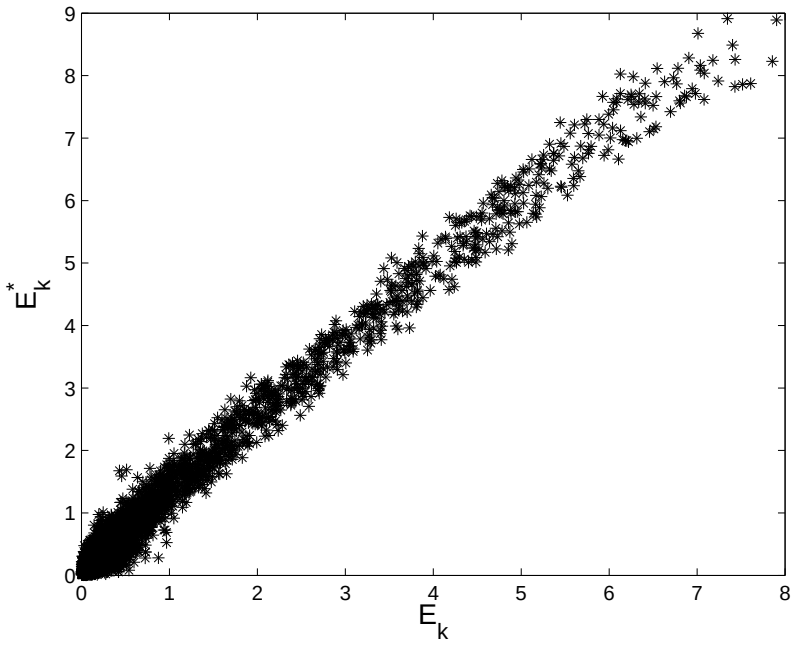

Figure 8.2: Cross-plot between the fine fractional error $E_k = \|F - F_k\|^2$ and the coarse fractional error $E_k^* = \|F - F_k^*\|^2$. These two quantities are strongly correlated.

The preconditioned MCMC method employs multiscale finite element methods in the preconditioning step. If a proposal is accepted by the coarse-scale test (Step 2), one can use the precomputed multiscale basis functions to reconstruct the velocity field on the fine-scale. Then the transport equation can be solved on the fine-grid coupled with the coarse-grid pressure equation [26, 49, 50, 1]. This approach provides an accurate approximation to the production data on the fine-grid and can be used to replace the fine-scale computation in the second-stage (step 3). In this procedure, the basis functions are not updated in time, or updated only in a few coarse blocks. Thus the fine-scale computation in the second-stage of the preconditioned MCMC method (step 3) can also be implemented fast. Since the basis functions from the first-stage are reused for the fine-scale computation, this combined multiscale approach can be very efficient for our sampling problem.

8.3.2 Analysis of the Preconditioned MCMC Method

Next we will analyze the preconditioned MCMC method in more detail. Denote

$$\begin{aligned} \mathcal{E} &= \big\{k;\ \pi(k) > 0\big\}, \\ \mathcal{E}^* &= \big\{k;\ \pi^*(k) > 0\big\}, \\ \mathcal{D} &= \big\{k;\ q(k|k_n) > 0 \text{ for some } k_n \in \mathcal{E}\big\}. \end{aligned} \tag{8.18}$$

The set $\mathcal{E}$ is the support of the posterior (target) distribution $\pi(k)$. $\mathcal{E}$ contains all the permeability field k that has a positive probability of being accepted as a sample. Similarly, $\mathcal{E}^*$ is the support of the coarse-scale distribution $\pi^*(k)$, which contains all the k acceptable by the coarse-scale test. $\mathcal{D}$ is the set of all the proposals which can be generated by the instrumental distribution $q(k|k_n)$. For the preconditioned MCMC method to work properly, the conditions $\mathcal{E} \subseteq \mathcal{D}$ and $\mathcal{E} \subseteq \mathcal{E}^*$ must hold (up to a zero measure set) simultaneously. If one of these conditions is not true, say, $\mathcal{E} \not\subseteq \mathcal{E}^*$, then there will exist a subset $A \subset (\mathcal{E} \setminus \mathcal{E}^*)$ such that

$$\pi(A) = \int_A \pi(k)dk > 0 \qquad \text{and} \qquad \pi^*(A) = \int_A \pi^*(k)dk = 0,$$

which means no element of A can pass the coarse-scale test and A will never be visited by the Markov chain $\{k_n\}$. Thus, $\pi(k)$ cannot be sampled properly.

For most practical proposals $q(k|k_n)$, such as the random walk samplers and independent samplers, the conditions $\mathcal{E}, \mathcal{E}^* \subset \mathcal{D}$ can be naturally satisfied. By choosing the parameter σ_c in $\pi^*(k)$ properly, the condition $\mathcal{E} \subset \mathcal{E}^*$ can also be satisfied (see the discussion below). As a result, we have $\mathcal{E} \subset \mathcal{E}^* \subset \mathcal{D}$. In this case, $\mathcal{E}^*$ is identical to the support of the effective proposal $Q(k|k_n)$:

$$\mathcal{E}^* = \big\{k;\ Q(k|k_n) > 0 \text{ for some } k_n \in \mathcal{E}\big\}.$$

Next we will discuss the stability property of the preconditioned MCMC method. We shall show that the preconditioned MCMC method shares the same convergence property as the regular MCMC method. Denote by K the transition kernel of the Markov chain $\{k_n\}$ generated by the preconditioned MCMC method. Since its effective proposal is $Q(k|k_n)$ as

defined by (8.14), we get

$$K(k_n, k) = \rho(k_n, k)Q(k|k_n) \qquad \text{for } k \neq k_n, \tag{8.19}$$

$$K(k_n, \{k_n\}) = 1 - \int_{k \neq k_n} \rho(k_n, k)Q(k|k_n)dk. \tag{8.20}$$

That is, the transition kernel $K(k_n, \cdot)$ is continuous when $k \neq k_n$ and has a positive probability for the event $\{k = k_n\}$.

As in the regular MCMC method, it is easy to show that $K(k_n, k)$ satisfies the detailed balance condition

$$\pi(k_n)K(k_n, k) = \pi(k)K(k, k_n) \tag{8.21}$$

for any $k, k_n \in \mathcal{E}$. In fact, the equality (8.21) is obviously true when $k = k_n$. If $k \neq k_n$, then from (8.19) we have

$$\pi(k_n)K(k_n, k) = \pi(k_n)\rho(k_n, k)Q(k|k_n) = \min\Big(Q(k|k_n)\pi(k_n),\ Q(k_n|k)\pi(k)\Big)$$

$$= \min\left(\frac{Q(k|k_n)\pi(k_n)}{Q(k_n|k)\pi(k)},\ 1\right) Q(k_n|k)\pi(k) = \rho(k, k_n)Q(k_n|k)\pi(k) = \pi(k)K(k, k_n).$$

So the detailed balance condition (8.21) is always satisfied. Using (8.21), we can easily show that $\pi(A) = \int K(k, A)dk$ for any $A \in \mathcal{B}(\mathcal{E})$, where $\mathcal{B}(\mathcal{E})$ denotes all the measurable subsets of $\mathcal{E}$. Thus, $\pi(k)$ is indeed the stationary distribution of the transition kernel $K(k_n, k)$.

In the regular MCMC method, the proposal $q(k|k_n)$ is usually chosen to satisfy

$$q(k|k_n) > 0 \text{ for any } (k_n, k) \in \mathcal{E} \times \mathcal{E}, \tag{8.22}$$

which guarantees that the resulting regular MCMC method is irreducible. The similar statement is true for the preconditioned MCMC method.

Lemma 8.1 *If the proposal distribution $q(k|k_n)$ satisfies* (8.22) *and $\mathcal{E} \subset \mathcal{E}^*$ holds, then the chain $\{k_n\}$ generated by the preconditioned MCMC method is strongly π-irreducible.*

Proof According to the definition of strong irreducibility, we only need to show that

$K(k_n, A) > 0$ for all $k_n \in \mathcal{E}$ and any measurable set $A \subset \mathcal{E}$ with $\pi(A) > 0$. Note that

$$\begin{aligned} K(k_n, A) &\geq \int_{A \setminus k_n} K(k_n, k)dk = \int_{A \setminus k_n} \rho(k_n, k)Q(k_n, k)dk \\ &= \int_{A \setminus k_n} \rho(k_n, k)g(k_n, k)q(k|k_n)dk. \end{aligned}$$

In the above inequality, the equal sign holds when $k_n \notin A$. Since $\pi(A) = \int_A \pi(k)dk > 0$, it follows that $m(A) = m(A \setminus k_n) > 0$, where $m(A)$ is the Lebesgue measure. Since $A \subset \mathcal{E}$ and $\mathcal{E} \subset \mathcal{E}^*$, both $\rho(k_n, k)$ and $g(k_n, k)$ are positive for $k \in A$. Combining the positivity assumption (8.22), we can easily conclude that $K(k_n, A) > 0$. □

Most practical proposal distributions, such as random walk samplers or independent samplers, satisfy the positivity condition (8.22). Thus condition (8.22) poses only a mild restriction in practice. As we will see later, the proposals used in our numerical experiment naturally satisfy the condition (8.22).

Based on the stability property of Markov chains [88, 69], the following convergence result is readily available:

Theorem 8.1 *[88] Suppose* (8.22) *is true and* $\mathcal{E} \subset \mathcal{E}^*$ *holds, then the preconditioned Markov chain* $\{k_n\}$ *is ergodic: For any function* $h(k)$*,*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} h(k_n) = \int h(k)\pi(k)dk. \tag{8.23}$$

If the chain $\{k_n\}$ *is also aperiodic, then the distribution of* k_n *converges to* $\pi(k)$ *in the total variation norm*

$$\lim_{n \to \infty} \sup_{A \in \mathcal{B}(\mathcal{E})} \left| K^n(k_0, A) - \pi(A) \right| = 0 \tag{8.24}$$

for any initial state k_0*.*

To get convergence property (8.24), we need to show that the Markov chain $\{k_n\}$ generated by the preconditioned MCMC method is indeed aperiodic. Recall that a simple sufficient condition for aperiodicity is that $K(k_n, \{k_n\}) > 0$ for some $k_n \in \mathcal{E}$. In other words, the event $\{k_{n+1} = k_n\}$ happens with a positive probability in the preconditioned

MCMC method. From the definition (8.20), we have

$$K(k_n, \{k_n\}) = 1 - \int_{k \neq k_n} \rho(k_n, k) Q(k|k_n) dk = 1 - \int_{k \neq k_n} \rho(k_n, k) g(k_n, k) q(k|k_n) dk.$$

Consequently, $K(k_n, \{k_n\}) \equiv 0$ requires $g(k_n, k) = 1$ and $\rho(k_n, k) = 1$ for almost all $k \in \mathcal{D}$, which means that all the proposals generated by $q(k|k_n)$ are correct samples of distributions $\pi(k)$ and $\pi^*(k)$. This is obviously not true in practice. Thus, the practical preconditioned MCMC method is always aperiodic and converges to the target distribution $\pi(k)$ in the sense of (8.24).

Next we discuss the necessary condition $\mathcal{E} \subseteq \mathcal{E}^*$, which is essential to guarantee the convergence of the preconditioned MCMC method. Due to the Gaussian form of the posterior distribution, $\pi(k)$ and $\pi^*(k)$ do not have a compact support and the domain $\mathcal{E}$ (or $\mathcal{E}^*$) is the whole space spanned by all k. However, if the precision parameters σ_f and σ_c are relatively small, then $\pi(k)$ and $\pi^*(k)$ are very close to zero for most proposals. From the numerical point of view, the proposal k is very unlikely to be accepted if $\pi(k)$ or $\pi^*(k)$ is close to zero. Consequently, the support of the distributions should be interpreted as $\mathcal{E} = \{k; \pi(k) > \delta\}$ and $\mathcal{E}^* = \{k; \pi^*(k) > \delta\}$, where δ is a small positive number.

If $k \in \mathcal{E}$, then $\pi(k) > \delta$ and $\|F_k - F\|^2/\sigma_f^2$ is not very large. To make $k \in \mathcal{E}^*$, $\|F_k^* - F\|^2/\sigma_c^2$ should not be very large either. If $\|F_k^* - F\|^2$ is bounded by $\|F_k - F\|^2$ up to a multiplicative constant, then the condition $\mathcal{E} \subseteq \mathcal{E}^*$ can be satisfied by choosing the parameter σ_c properly. For most upscaled models, the coarse-scale quantity is indeed bounded by the corresponding fine-scale quantity. For example, the upscaled velocity $\boldsymbol{v}^*$ in the saturation equation is obtained by averaging the fine-scale velocity $\boldsymbol{v}$ over the coarse-grid blocks

$$\boldsymbol{v}^*(x) = \sum_i \left(\frac{1}{|\Omega_i|} \int_{\Omega_i} \boldsymbol{v}(y) dy \right) 1_{\Omega_i}(x),$$

where $\Omega_i \subset \Omega$ are the coarse blocks. It is easy to show that

$$\begin{aligned} \|\boldsymbol{v}^*\|^2_{L^2(\Omega)} &= \sum_i \frac{1}{|\Omega_i|} \Big(\int_{\Omega_i} \boldsymbol{v}(y) dy \Big)^2 \\ &\leq \sum_i \frac{1}{|\Omega_i|} \Big(\int_{\Omega_i} 1(y)^2 dy \Big) \Big(\int_{\Omega_i} \boldsymbol{v}^2(y) dy \Big) = \|\boldsymbol{v}\|^2_{L^2(\Omega)}. \end{aligned} \tag{8.25}$$

Thus, the coarse-scale velocity is bounded by the corresponding fine-scale one. We would

like to remark that for some nonlinear averaging operators, one can also show that the coarse-scale quantities are bounded by the corresponding fine-scale quantities. One of the examples is the homogenization operator for linear elliptic equations.

In general, it is difficult to carry out such estimates for fractional flows. However, coarse-scale fractional flows can be interpreted as some type of average of the corresponding fine-scale fractional flows. Indeed, the fine-scale fractional flow curve can be regarded as the travel times along the characteristics of the particles that start at the inlet. The coarse-scale fractional flow, on the other hand, represents an average of these travel times over characteristics within the coarse domain. In general, the estimation similar to (8.25) does not hold for fractional flow curves, as our next counter-example shows. For simplicity, we present the counter-example for the single-phase flow in porous media with four layers. This example can be easily generalized. Denote by t_i, $i = 1, 2, 3, 4$ the breakthrough times for the layers. Consider two fine-scale (with four layers) permeability fields with breakthrough times $t_1 = T_1$, $t_2 = T_2$, $t_3 = T_1$, $t_4 = T_2$ and $t_1 = T_1$, $t_2 = T_1$, $t_3 = T_2$, $t_4 = T_2$, respectively. These two fine-scale permeability fields will give the same fractional flows, since the times of the flights are the same up to a permutation. Now we consider the upscaling of these two fine-scale permeability fields to two-layered media. Upscaling is equivalent to averaging the breakthrough times. Consequently, the breakthrough times for the corresponding upscaled models are $t_1^* = 0.5(T_1 + T_2)$, $t_2^* = 0.5(T_1 + T_2)$, and $t_1^* = 0.5(T_1 + T_1) = T_1$, $t_2^* = 0.5(T_2 + T_2) = T_2$, respectively. Thus, the coarse-scale models give different fractional flows, even though the fractional flows are identical for the fine-scale models. However, this type of counter-example can be avoided in practice, because the near-well values of the permeability are known, and consequently, permutation of the layers can be avoided.

8.3.3 Numerical Results

In this section we discuss the implementation details of the preconditioned MCMC method and present the numerical results. Suppose the permeability field $k(\boldsymbol{x})$, where $\boldsymbol{x} = (x, z)$, is defined on the unit square $\Omega = [0, 1]^2$. We assume that the permeability field $k(\boldsymbol{x})$ is a log-normal process and its covariance function is known. The observed data include the fractional flow curve F and the values of the permeability at sparse locations. We discretize the domain Ω by a rectangular mesh and the permeability field k is represented by a matrix (thus k is a high dimensional vector). For the boundary conditions, we assume that $p = 1$,

$S = 1$ on the boundary $x = 0$, and $p = 0$ on the boundary $x = 1$. We assume no flow conditions on the lateral boundaries $z = 0$ and $z = 1$. We will consider both single-phase and two-phase flow displacements in the simulation.

Using the Karhunen-Loeve expansion [59, 97], the permeability field can be expanded in terms of an optimal L^2 basis. By truncating the expansion we can represent the permeability matrix by a small number of random parameters. To impose the hard constraints (the values of the permeability at prescribed locations), we will sample a linear subspace of the random parameter space, which will automatically yield permeability fields satisfying the hard constrains.

Denote $Y(\boldsymbol{x}, \omega) = \log[k(\boldsymbol{x}, \omega)]$, where the random element ω is included to remind us that k is modeled as a random field. Suppose $Y(\boldsymbol{x}, \omega)$ is a second order stochastic process. Without loss of generality, we assume that $E[Y(\boldsymbol{x}, \omega)] = 0$. Denote the covariance function of Y as $R(\boldsymbol{x}, \boldsymbol{y}) = E\left[Y(\boldsymbol{x}, \omega)Y(\boldsymbol{y}, \omega)\right]$. Suppose $Y(\boldsymbol{x})$ has the KLE

$$Y(\boldsymbol{x}, \omega) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \theta_k(\omega)\, \phi_k(\boldsymbol{x}), \tag{8.26}$$

where ϕ_k and λ_k satisfy

$$\int_{\Omega} R(\boldsymbol{x}, \boldsymbol{y})\phi_k(\boldsymbol{y})d\boldsymbol{y} = \lambda_k \phi_k(\boldsymbol{x}), \qquad k = 1, 2, \ldots. \tag{8.27}$$

We assume that the eigenvalues λ_k are ordered in descending order $\lambda_1 \geq \lambda_2 \geq \ldots$.

Suppose the permeability field $k(\boldsymbol{x}, \omega)$ is a log-normal homogeneous stochastic process. Then $Y(\boldsymbol{x}, \omega)$ is a Gaussian process and θ_k are independent standard Gaussian random variables. We assume that the covariance function of $Y(\boldsymbol{x}, \omega)$ has the form

$$R(\boldsymbol{x}, \boldsymbol{y}) = \sigma^2 \exp\Big(-\frac{|x_1 - y_1|^2}{2L_1^2} - \frac{|x_2 - y_2|^2}{2L_2^2}\Big), \tag{8.28}$$

where L_1 and L_2 are the correlation lengths in each dimension, and $\sigma^2 = E(Y^2)$ is a constant. In our first example, we set $L_1 = 0.2$, $L_2 = 0.2$ and $\sigma^2 = 2$. We first solve the eigenvalue problem (8.27) numerically and obtain the eigenpairs $\{\lambda_k, \phi_k\}$. In Figure 8.3 the first 50 eigenvalues are plotted. As we can see, the eigenvalues of the KLE decay very fast. It has been shown in [32] that the eigenvalues decay exponentially fast for the covariance function (8.28). Therefore, only a small number of terms need to be retained

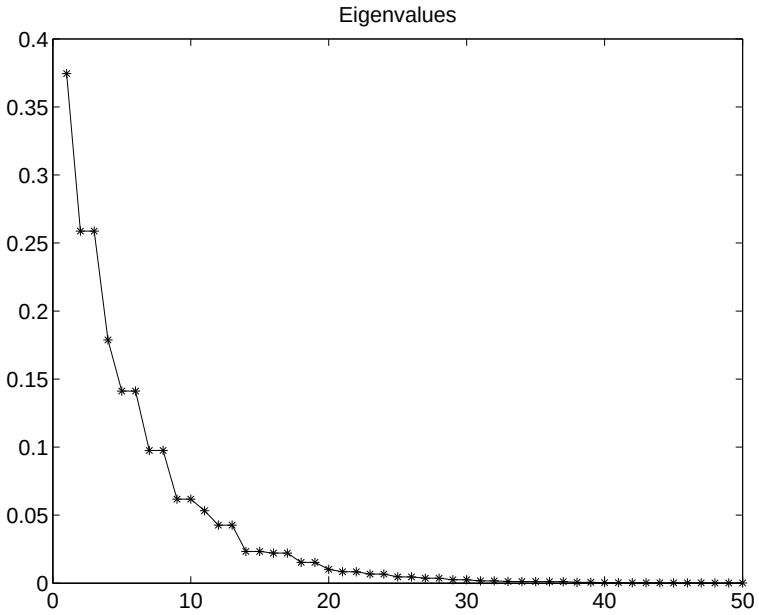

Figure 8.3: Eigenvalues of the KLE for the Gaussian covariance with $L_1 = L_2 = 0.2$. The eigenvalues decay very fast.

in the truncated expansion (8.26). We can sample $Y(\boldsymbol{x}, \omega)$ easily from the truncated KLE (8.26) by generating independent Gaussian random variables θ_k.

For a different covariance function such as

$$R(\boldsymbol{x}, \boldsymbol{y}) = \sigma^2 \exp\Big(-\frac{|x_1 - y_1|}{L_1} - \frac{|x_2 - y_2|}{L_2}\Big),$$

the eigenvalues of the integral equation (8.27) may decay slowly (only algebraically [32]). To achieve the same accuracy, more terms should be retained in the truncated expansion (8.26), which will increase the dimension of the parameter space to represent the permeability. As a result, sampling the permeability from the distribution will be more expensive for both the direct MCMC method and the preconditioned MCMC method. However, small parameter space does not favor the preconditioned MCMC method and the preconditioning technique is applicable independent of the problem dimension. For permeability with higher dimensional parameters, the acceptance rates of the direct MCMC method will be even lower. Consequently, the preconditioned MCMC method will be preferable since its filtering procedure can increase the acceptance rates dramatically. Note that if the permeability field is not a log-normal process, then θ_k in the expansion (8.26) are not necessarily Gaussian random variables. However, we can still sample the permeability field from the truncated

expansion (8.26) by sampling the random variables θ_k.

In the numerical experiments, we first generate a reference permeability field by the full KLE (8.26) and compute the corresponding fractional flow. To propose permeability fields from the prior (unconditioned) distribution, we keep 20 terms in the KLE. Suppose the permeability field is known at eight distinct points. This condition is imposed by setting

$$\sum_{k=1}^{20} \sqrt{\lambda_k}\theta_k\phi_k(\boldsymbol{x}_j) = \alpha_j, \tag{8.29}$$

where α_j $(j = 1, \ldots, 8)$ are prescribed constants. For simplicity, we set $\alpha_j = 0$ for all $j = 1, \ldots, 8$. In the simulations we propose 12 θ_i and calculate the rest of θ_i by solving the linear system (8.29). In all the simulations, we test 50,000 proposals and iterate the Markov chain 50,000 times. Because the direct MCMC computations are expensive, we do not select the large model problems, and only consider 40×40 and 60×60 fine-scale models. However, the preconditioned MCMC method is applicable independent of the size of the permeability field.

We have considered two types of instrumental proposal distributions $q(k|k_n)$: the independent sampler and the random walk sampler. In the case of independent sampler, the proposal distribution $q(k|k_n)$ is chosen to be independent of k_n and equal to the prior (unconditioned) distribution. In the random walk sampler, the proposal distribution depends on the previous value of the permeability field and is given by $k = k_n + \epsilon_n$, where ϵ_n is a random perturbation with prescribed distribution. If the variance of ϵ_n is chosen to be very large, then the random walk sampler becomes similar to the independent sampler. Although the random walk sampler allows us to accept more realizations, it often gets stuck in the neighborhood of a local maximum of the target distribution. For both proposal distributions, we have observed consistently more than ten times of increase in the acceptance rate when the preconditioned MCMC method is used. In this thesis we only report the results using the independent samplers.

For the first set of numerical tests, we use 40×40 fine-scale permeability field and 10×10 coarse-scale models. The permeability field is assumed to be log-normal, with $L_1 = L_2 = 0.2$ and $\sigma^2 = 2$ for the covariance function (7.10). In Figure 8.4, the acceptance rates are plotted against different coarse-scale precisions σ_c. Here the acceptance rate refers to the ratio between the number of accepted permeability samples and the number of fine-

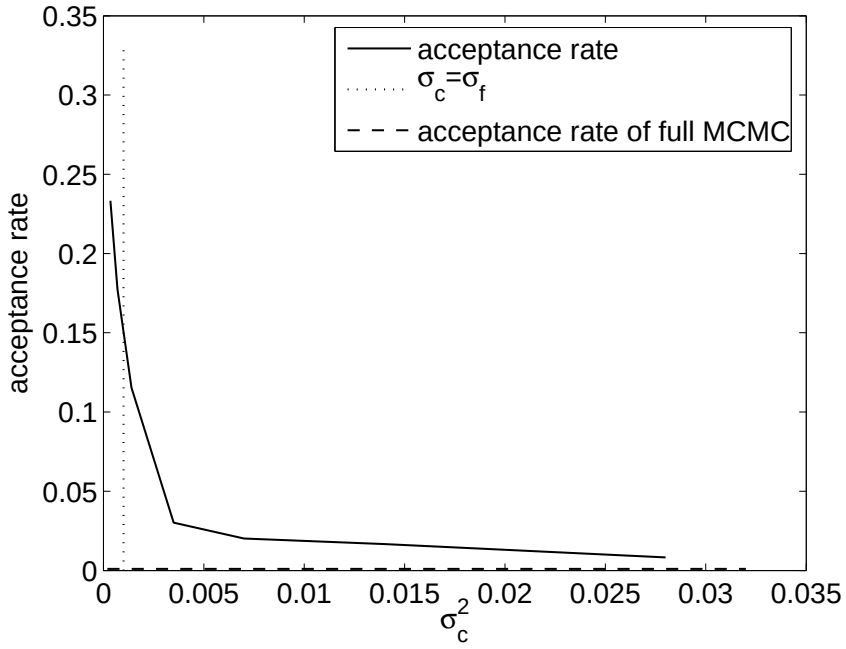

Figure 8.4: Acceptance rate of the preconditioned MCMC with different coarse-scale precisions. Single-phase flow and $\sigma_f^2 = 0.001$.

scale simulations that are performed. The acceptance rate for the direct (full) MCMC is plotted using dashed line, and it is equal to 0.001. The vertical doted line marks the coarse-scale precision $\sigma_c = \sigma_f$. If σ_c is very small, then the total number of accepted realizations is also small, even though the acceptance rate is higher. We have found that if σ_c is of the same order as σ_f then the preconditioned MCMC method accepts almost the same number of proposals as the direct MCMC, but requires only 10% of the fine-scale runs. Note that as σ_c increases the acceptance rate decreases and reaches the acceptance rate of full MCMC. Indeed, if σ_c is very large, then all the proposals will be accepted by the coarse-scale test, and there is no gain in preconditioning. In general, one can estimate the optimal σ_c based on a limited number of simulations, prior to the full simulations as described above.

In Figure 8.5 we plot the fractional flows of the accepted permeability realizations. The bold solid line is the reference fractional flow curve and the dotted lines are the fractional flows of accepted samples. As we can see, the fractional flows of accepted realizations are very close to the observed fractional flow, because the precision is taken to be $\sigma_f^2 = 0.001$. In Figure 8.6, we plot the fractional flow error $E_k = \|F - F_k\|^2$ of the accepted samples for both the direct and preconditioned MCMC methods. We observe that the errors of both of the Markov chains converge to a steady state within 20 accepted iterations (corresponds

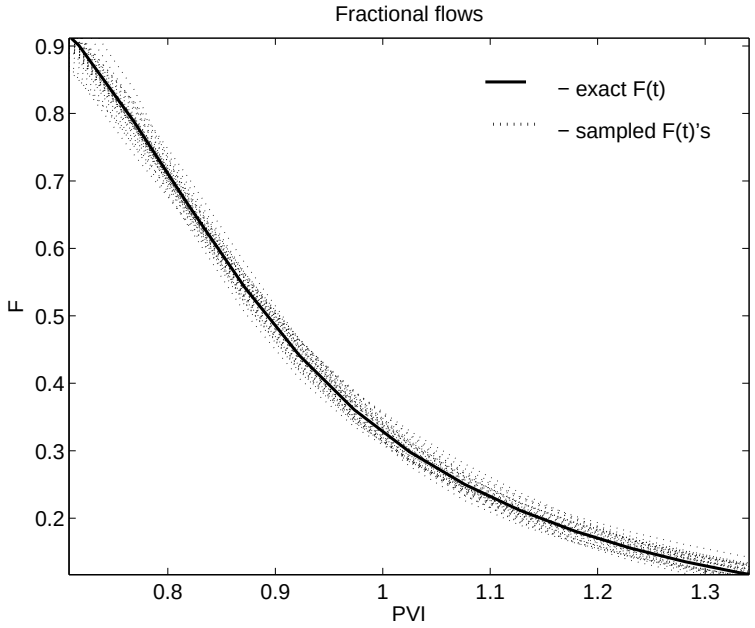

Figure 8.5: The reference fractional flow (solid line) and fractional flows of accepted permeability samples (dotted lines). The fractional flow of the sampled permeability fields are very close to the reference fractional flow.

to 20,000 proposals). Note that we can assess the convergence of the MCMC methods based on the fractional flow errors. This is a reasonable indicator for the convergence and is frequently used in practice. Given the convergence result of the MCMC method, a longer chain can be easily generated when it is needed.

We present the last five accepted permeability realizations generated by the preconditioned MCMC method in Figure 8.7. The first plot is the reference (true) permeability field and the others are the last five accepted permeability realizations. Some of these realizations closely resemble the true permeability field. And they are all eligible samples because their fractional flows are in good agreement with the reference (true) fractional flow. One can use these permeability samples for the uncertainty quantifications.

For the next numerical examples, we consider the anisotropic permeability case with $L_1 = 0.4$, $L_2 = 0.05$, and $\sigma^2 = 2$. We consider a 60×60 fine grid. As in the previous example, we use eight conditioning points and truncate the KLE expansion of the permeability field with 20 terms to maintain a sufficient accuracy. In Figure 8.8, we plot the acceptance rates for 6×6 and 10×10 coarse-scale models with different coarse-scale precision σ_c^2. The acceptance rate for the direct (full) MCMC is 0.0008 and it is designated by the dashed line. The acceptance rates is increased by more than ten times in the preconditioned MCMC

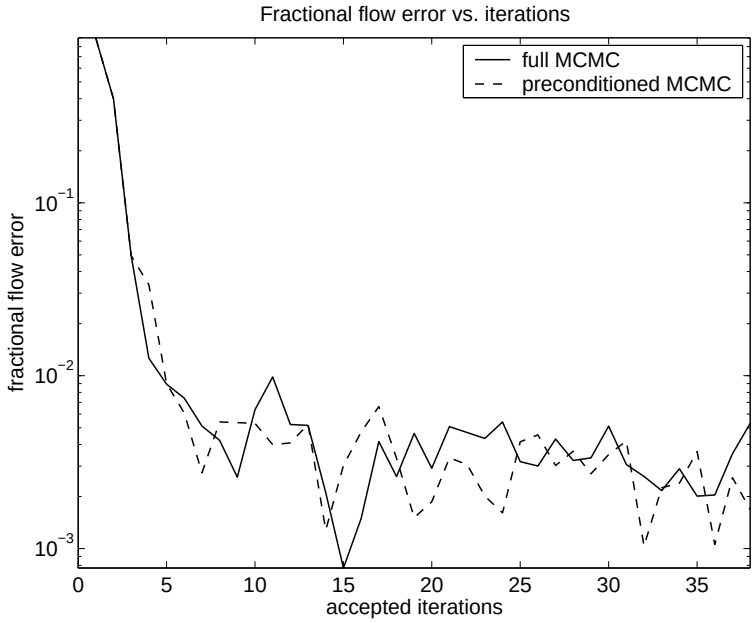

Figure 8.6: Fractional flow errors for the accepted permeability samples

method when σ_c is slightly larger than σ_f (the vertical doted line marks the choice $\sigma_c = \sigma_f$). We also observe higher acceptance rate for 10×10 coarse-scale model than for 6×6 coarse-scale model. This is because 10×10 coarse-scale model provides more accurate predictions of the fine-scale results compared to the 6×6 coarse-scale model. As in the previous cases, when the σ_c is slightly larger than σ_f, the preconditioned MCMC method can accept the same number of samples as the underlying full MCMC but performs only 10 percent of the fine-scale simulations. Moreover, both the direct (full) MCMC and the preconditioned MCMC methods converge to the steady state within 20 accepted iterations, which indicates that both chains have the similar convergence properties. In Figure 8.9, we plot the last five accepted samples of the permeability field generated by the preconditioned MCMC method using 6×6 coarse-scale model. Some of these samples closely resemble the reference (true) permeability field.

Our next set of numerical experiments are for the two-phase flow simulations. We have observed very similar rate increases for two-phase flow simulations, and thus restrict our numerical results to only a few examples. We consider $\mu_w/\mu_o = 5.0$ and $k_{rw}(S) = S^2$, $k_{ro}(S) = (1-S)^2$. Typically, the upscaling error for two-phase flows is very similar to that of the single-phase flows. We consider 40×40 fine-scale log-normal permeability field with $L_1 = L_2 = 0.2$ and 10×10 coarse-scale models. In Figure 8.10, the acceptance rate with

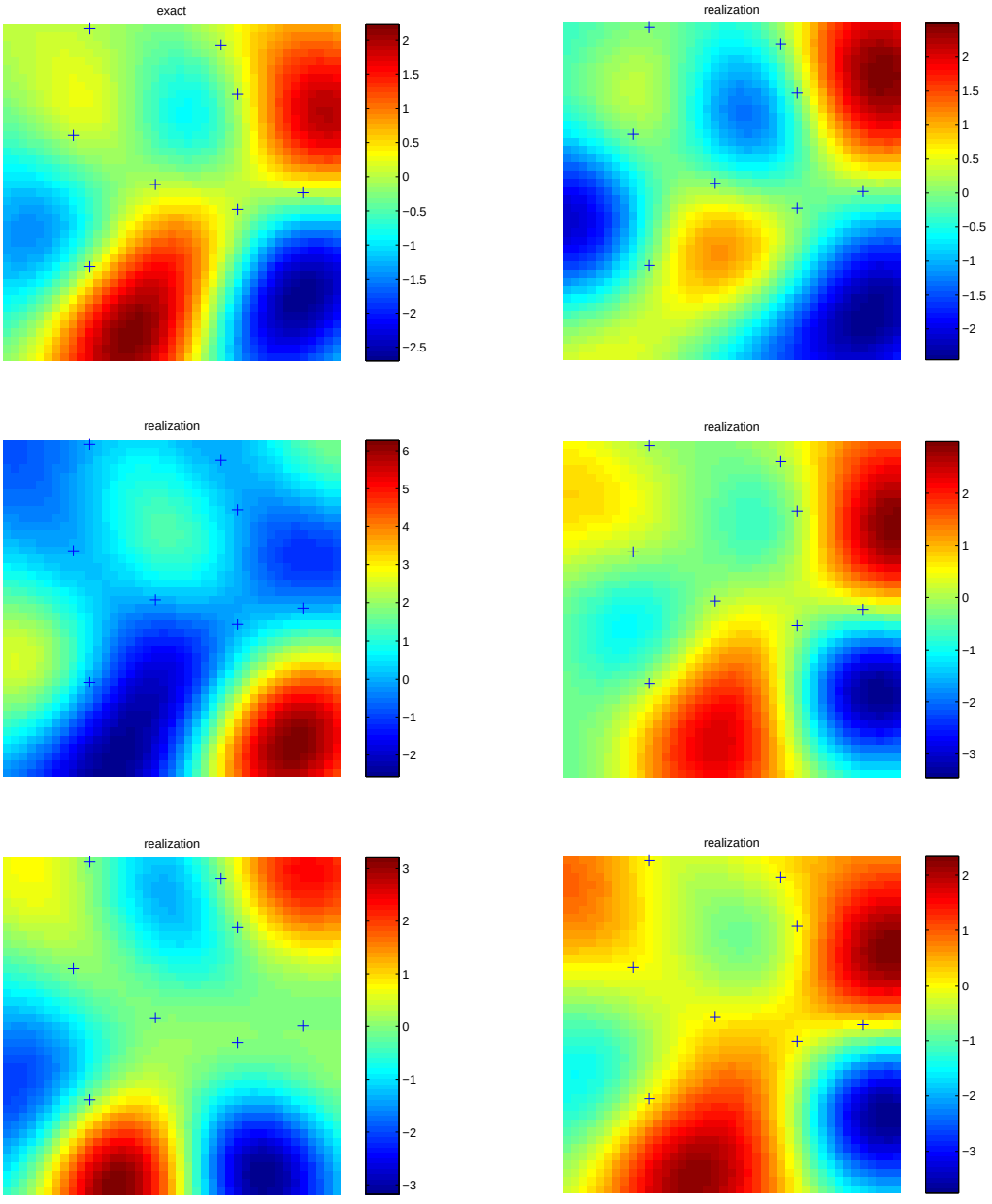

Figure 8.7: The last five accepted realizations of the log permeability field. The "+" sign marks the locations of the hard data. Some of these realizations closely resemble the exact permeability field.

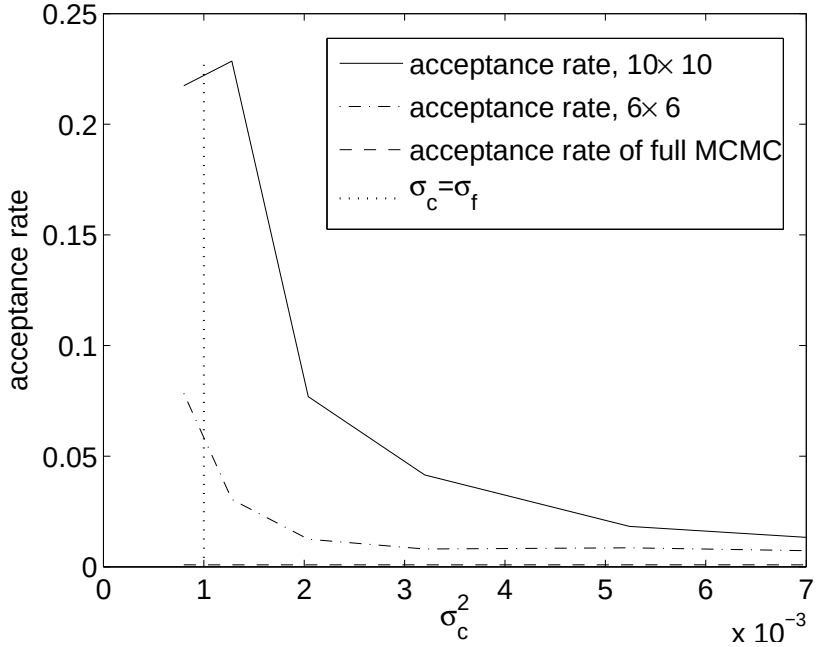

Figure 8.8: Acceptance rate of the preconditioned MCMC method with 6×6 and 10×10 coarse-scale models. Anisotropic single-phase flow and $\sigma_f^2 = 0.001$.

fine-scale precision $\sigma_f^2 = 0.001$ is plotted. As in the case of the single-phase flow simulations, we observe more than ten times increase in the acceptance rate. The preconditioned MCMC method accepts the same amount of samples as in the full MCMC with less than 10% of the fine-scale runs. To study the relative convergence of the preconditioned MCMC method, we plot the fractional flow error for both full and preconditioned MCMC simulations in Figure 8.11. It can be seen from this figure that both the full and preconditioned MCMC methods reach the steady state within 20 accepted iterations. This indicates that the direct and preconditioned MCMC methods have similar convergence properties. The typical samples for the two-phase flow are very similar to those for the single-phase flow, and we do not present them here.

As we mentioned earlier, the full MCMC method and the preconditioned MCMC method accept approximately the same amount of samples for a fixed number of tested proposals. Denote N as the total number of tested proposals, then the direct MCMC method requires exactly N number of fine-scale simulations. Suppose $M < N$ is the number of fine-scale simulations conducted in the preconditioned MCMC method. Denote t_f and t_c as the CPU times for a single fine-scale and coarse-scale forward simulation. Then the computational costs for the direct MCMC method and the preconditioned MCMC method would be Nt_f

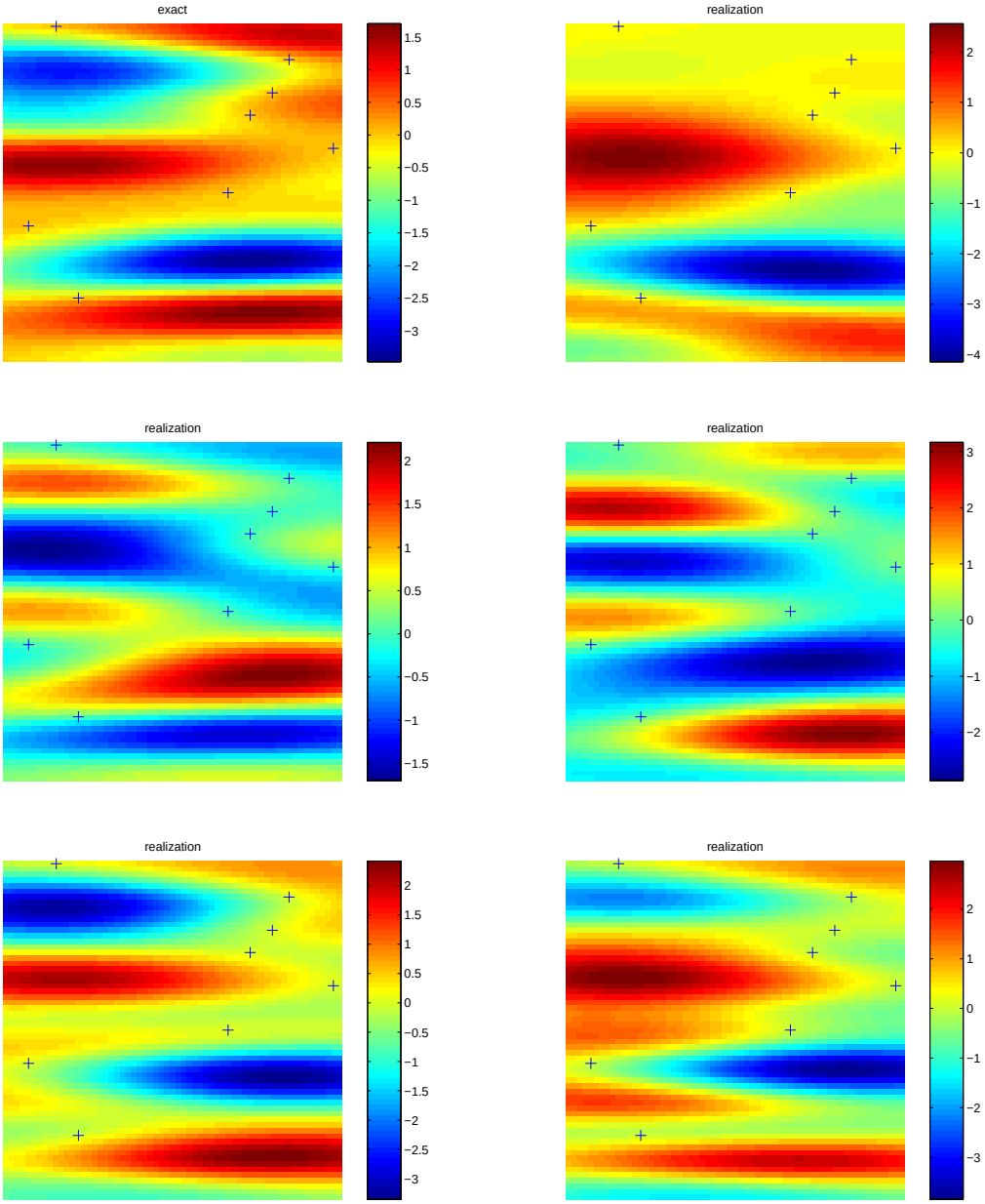

Figure 8.9: The last five accepted realizations of log permeability fields for the anisotropic single-phase flow. The "+" sign marks the locations of the hard data. Some of these samples closely resemble the exact permeability field.

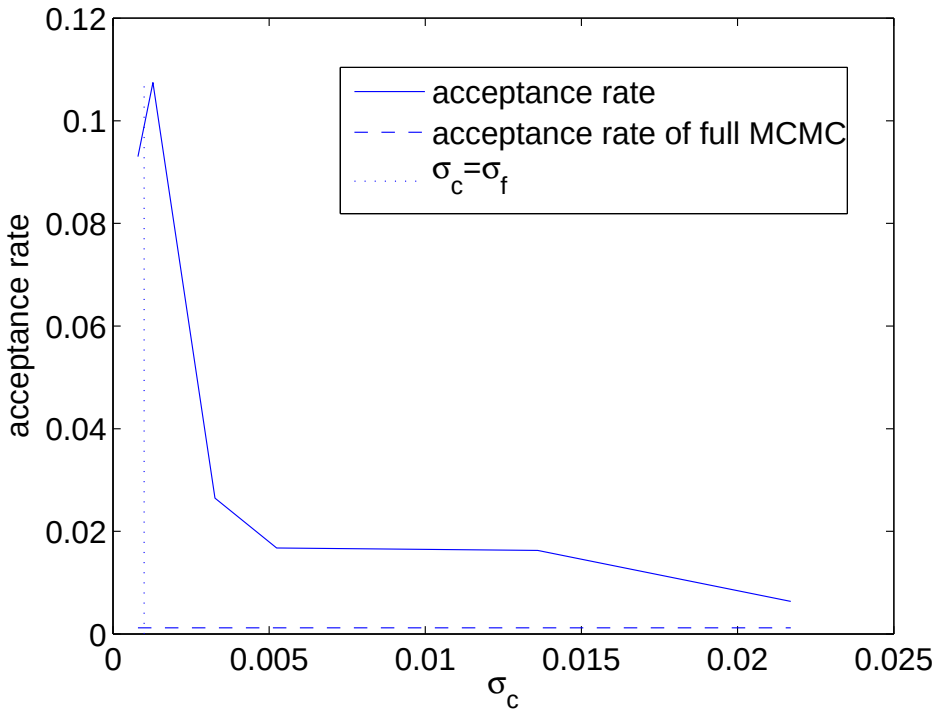

Figure 8.10: Acceptance rate of the preconditioned MCMC method with different coarse-scale precisions. Two-phase flow and $\sigma_f^2 = 0.001$.

and $Nt_c + Mt_f$ respectively. Therefore, the CPU cost for the preconditioned MCMC method is only $\frac{t_c}{t_f} + \frac{M}{N}$ of that of the direct MCMC method. The coarse-scale computational cost t_c is usually much smaller than the fine-scale computational cost t_f. Suppose the fine-scale model is upscaled five times in each direction. Then solving the pressure equation at each time step is about 25 times faster on the coarse grid than on the fine grid. Moreover, the saturation equation is also solved on the *coarse grid* and with *larger time steps*. This makes the overall coarse-scale computations of the two-phase flow equation at least 25 times faster than the fine-scale computations, i.e., $t_c \approx 0.04 t_f$. If the acceptance rate is increased by more than ten times in the preconditioned MCMC method, as in our numerical experiments, then $\frac{M}{N} < 0.1$, and the overall CPU cost of the preconditioned MCMC method would be only 10% of the CPU costs of the direct MCMC method. Note that using very coarse-scale models (fewer coarse blocks) reduces t_c but increases the fine-scale run ratio $\frac{M}{N}$. On the other hand, using finer coarse-scale models reduces the ratio $\frac{M}{N}$ but increases t_c. Consequently, a somewhat moderate coarsening (five to ten times coarsening in each direction for large-scale fine models) can provide an optimal choice in the preconditioning of the MCMC simulations.

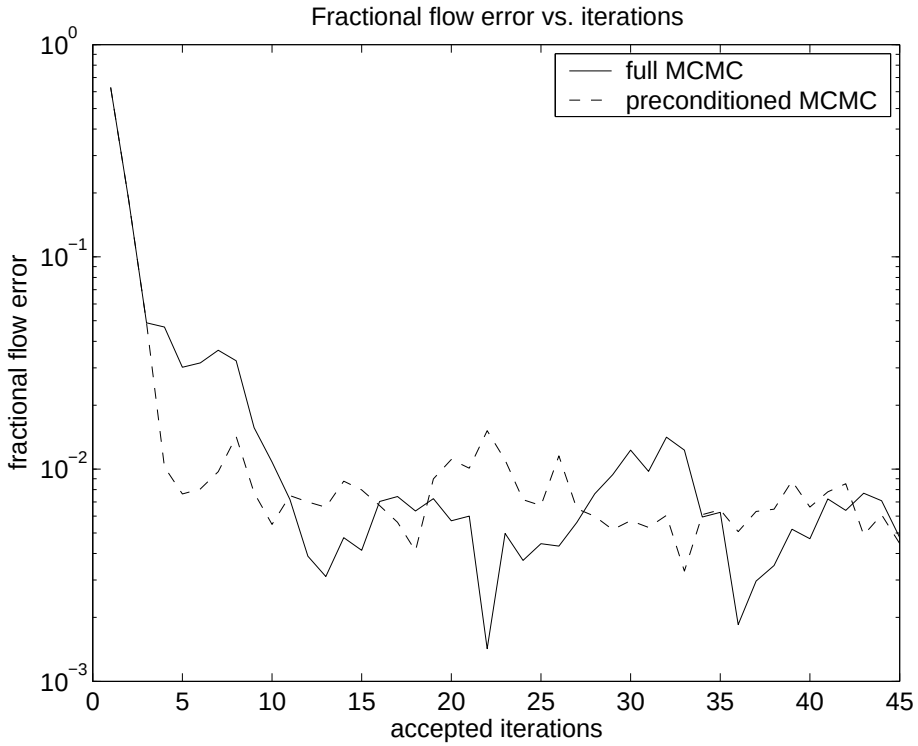

Figure 8.11: Fractional flow errors for the accepted permeability samples in two phase-flow.

8.4 WCE Methods and Coarse Gradient Langevin Algorithms

8.4.1 Proposal Distributions Based on Langevin Diffusions

In the above numerical method, we used the independent sampler as the proposal distribution for the preconditioned MCMC method. Another important type of proposal distribution can be derived from the Langevin diffusion, as proposed by Grenander and Miller [41]. The Langevin diffusion is defined by the stochastic differential equation

$$dk_\tau = \frac{1}{2}\nabla \log \pi(k_\tau)d\tau + dW_\tau, \tag{8.30}$$

where W_τ is the standard Brownian motion vector with independent components. Denote the probability density distribution (pdf) of k_τ at time τ as

$$p(k,\tau) = P\Big(k_\tau = k \,\Big|\, k_{\{\tau=0\}} = k_0\Big),$$

then the pdf $p(k,\tau)$ satisfies the Fokker-Planck equation

$$\begin{cases} \frac{\partial}{\partial\tau}p(k,\tau) + \frac{1}{2}\nabla\cdot\Big(\nabla\log\pi(k)\,p(k,\tau)\Big) = \frac{1}{2}\Delta p(k,\tau), \\ p(k,0) = \delta(k-k_0). \end{cases}$$

The above equation is a heat equation with time independent coefficients. Its solution converges to a steady state when τ goes to infinity. In other words, the Markov chain k_τ determined by (8.30) has a stationary distribution when τ is large. Denote

$$P(k) = \lim_{\tau\to\infty} p(k,\tau),$$

then $P(k)$ satisfies the equation

$$\nabla\cdot\Big(\nabla\log\pi(k)\,P(k)\Big) = \Delta P(k). \tag{8.31}$$

It is easy to check that $P(k) = \pi(k)$ is the solution of the elliptic equation (8.31). So $\pi(k)$ is the stationary distribution of the random solution k_τ of the diffusion equation (8.30).

Suppose we discretize the equation (8.30) by a numerical scheme, such as

$$k_{n+1} = k_n + \frac{\Delta\tau}{2}\nabla\log\pi(k_n) + \sqrt{\Delta\tau}\,\epsilon_n,$$

where ϵ_n are independent standard normal distributions. Since $\pi(k)$ is the stationary distribution of k_τ, it may be tempting to take k_n as samples from the distribution $\pi(k)$ for large n. However, the discrete solution k_n can have quite different asymptotic behavior from the continuous diffusion process k_τ [88]. In general, the discrete solution k_n does not necessarily have $\pi(k)$ as its stationary distribution. Instead of taking k_n as samples directly, we use them as test proposals for the MCMC algorithm. The samples will be further tested and corrected by the Metropolis acceptance-rejection step (8.11). Consequently, we can choose the proposal generator $q(Y|k_n)$ in Algorithm I as

$$Y = k_n + \frac{\Delta\tau}{2}\nabla\log\pi(k_n) + \sqrt{\Delta\tau}\,\epsilon_n. \tag{8.32}$$

Since ϵ_n are independent Gaussian vectors, the transition distribution of the proposal gen-

erator (8.32) is

$$
\begin{aligned}
q(Y|k_n) &\propto \exp\left(-\frac{\|Y - k_n - \frac{\Delta\tau}{2}\nabla \log \pi(k_n)\|^2}{2\Delta\tau}\right), \\
q(k_n|Y) &\propto \exp\left(-\frac{\|k_n - Y - \frac{\Delta\tau}{2}\nabla \log \pi(Y)\|^2}{2\Delta\tau}\right).
\end{aligned}
\tag{8.33}
$$

The proposal (8.32) can be regarded as a problem-adapted random walk. The gradient of the target distribution is included to move the samples to regions with larger probability. The use of the gradient information in inverse problems for subsurface characterization is not new. In an original work, Oliver et al [78, 79] developed the randomized maximum likelihood method, which uses the gradient information of the target distribution. This approach uses unconditional realizations of the production and permeability data and solves a deterministic gradient-based minimization problem. The solution of this minimization problem is taken as a proposal and is accepted with probability 1, since the acceptance probability is very difficult to estimate. In addition to the need of solving a gradient-based inverse problem, this method does not guarantee a proper sampling of the posterior distribution. Though the Langevin formula (8.32) shares the same spirit with the randomized maximum likelihood method, it is more efficient and rigorous, and one can compute the acceptance probability easily. The Langevin algorithms also allow the MCMC method to achieve high acceptance rates.

However, the Langevin algorithms require computing the gradients of the target distribution $\pi(k)$, which is very expensive. As we emphasized before, to valuate the distribution $\pi(k)$ for a given k, we need to solve the nonlinear PDE system (8.1)–(8.3) numerically. To compute the gradient of $\pi(k)$ numerically, we need to solve the nonlinear PDE system (8.1)–(8.3) multi times. Following the idea of the preconditioned MCMC method where coarse-scale models are used, we propose to compute the gradients of the distribution $\pi(k)$ based on its coarse-scale models.

To model $\pi(k)$ on the coarse-scale, we define a coarse-grid map F_k^* between the permeability field k and the fractional flow F. The coarse scale map F_k^* can be constructed explicitly from the upscaled WCE solution of the PDE system (8.1)-(8.3). We will discuss the details of constructing F_k^* in the next section. Based on the coarse-scale fractional flow

F_k^*, the target distribution $\pi(k)$ can be approximated on the coarse-scale as

$$\pi^*(k) \propto \exp\left(-\frac{\|F - F_k^*\|^2}{\sigma_c^2}\right) P(k), \tag{8.34}$$

where σ_c is the precision of the coarse-scale models. Then we generate the Langevin samples (8.32) using the gradient of the coarse-scale target distribution

$$Y = k_n + \frac{\Delta\tau}{2}\nabla \log \pi^*(k_n) + \sqrt{\Delta\tau}\,\epsilon_n. \tag{8.35}$$

The transition distribution of the coarse-grid proposal (8.35) is

$$\begin{aligned} q^*(Y|k_n) &\propto \exp\left(-\frac{\|Y - k_n - \frac{\Delta\tau}{2}\nabla \log \pi^*(k_n)\|^2}{2\Delta\tau}\right), \\ q^*(k_n|Y) &\propto \exp\left(-\frac{\|k_n - Y - \frac{\Delta\tau}{2}\nabla \log \pi^*(Y)\|^2}{2\Delta\tau}\right). \end{aligned} \tag{8.36}$$

Since the coarse-scale distribution $\pi^*(k_n)$ can be constructed explicitly from the WCE solution, computing its gradient is straightforward. There is no need to solve the nonlinear PDE system (8.1)–(8.3) repeatedly. The coarse-scale distribution $\pi^*(k)$ serves as a regularization of the original fine-scale distribution $\pi(k)$. By replacing the fine-scale gradient with the coarse-scale gradient, we can reduce the computational cost dramatically but still direct the proposals to regions with larger probabilities.

Because of the discretization errors, the proposals generated by the Langevin algorithms are not exact samples from its stationary distribution. We still need to test them by the Metropolis acceptance-rejection step. Using the coarse-scale Langevin algorithm as the proposal distribution for the preconditioned MCMC method, we get the following revised MCMC algorithm:

Algorithm III (preconditioned coarse-gradient Langevin algorithm)

- Step 1. At k_n, generate a trial proposal Y from the coarse Langevin algorithm (8.35).
- Step 2. Take the proposal k as

$$k = \begin{cases} Y \text{ with probability } g(k_n, Y), \\ k_n \text{ with probability } 1 - g(k_n, Y), \end{cases}$$

where

$$g(k_n, Y) = \min\left(1, \frac{q^*(k_n|Y)\pi^*(Y)}{q^*(Y|k_n)\pi^*(k_n)}\right).$$

Therefore, the proposal k is generated from the effective instrumental distribution

$$Q(k|k_n) = g(k_n, k)q^*(k|k_n) + \left(1 - \int g(k_n, k)q^*(k|k_n)dk\right)\delta_{k_n}(k). \tag{8.37}$$

- Step 3. Accept k as a sample with probability

$$\rho(k_n, k) = \min\left(1, \frac{Q(k_n|k)\pi(k)}{Q(k|k_n)\pi(k_n)}\right), \tag{8.38}$$

i.e., $k_{n+1} = k$ with probability $\rho(k_n, k)$, and $k_{n+1} = k_n$ with probability $1 - \rho(k_n, k)$.

Using the Karhunen-Loeve expansion (8.26), the permeability field k is represented as Hermite polynomial expansions of the Gaussian random variables θ. By changing the variables from k to θ, we can reduce the dimension of the permeability dramatically. To sample the permeability k, we only need to sample the random parameters θ. Consequently, we can rewrite the posterior distribution $\pi^*(k)$ in the θ parameter space and apply Algorithm III to sample θ. For this purpose, denote

$$\Pi^*(\theta) = \pi^*(k(\theta)),$$

then we can generate samples of θ using the following Langevin algorithm:

$$d\theta_s = \frac{1}{2}\nabla \log \Pi^*(\theta_s) + d\,W_s. \tag{8.39}$$

For each sample of θ, we can easily get the corresponding sample for the permeability field $k(x)$. In the next section, we construct the coarse-scale distribution $\Pi^*(\theta)$ explicitly using the upscaled WCE method. With the explicit formula for $\Pi^*(\theta)$, we can generate samples for θ by the Langevin algorithm (8.39) very easily, without the need of solving any nonlinear PDEs.

8.4.2 Constructing $\Pi^*(\theta)$ Explicitly by the Upscaled WCE Method

To construct the coarse-scale distribution $\Pi^*(\theta)$, we only need to find out the nonlinear map between θ and the coarse-scale fractional flow $F^* = F^*(\theta)$. To do that, we first solve the flow problem (8.1)–(8.3) by the WCE method and represent all the random solutions as spectral expansions of θ. Then the fractional flow map $F^*(\theta)$ can be constructed explicitly based on its definition (8.7).

Next we will discuss how to solve the nonlinear PDE system (8.1)–(8.3) by the WCE method. For the ease of presentation, we will focus on the single phase flows only. The same procedure can be applied to two-phase flows in a very similar way, though the formulae are more complicated. More specifically, we consider the following PDE system

$$\begin{cases} \nabla \cdot (k \nabla P) = q, \\ v = -k \nabla P, \\ \frac{\partial S}{\partial t} + \nabla \cdot (vS) = 0. \end{cases} \tag{8.40}$$

The domain and boundary conditions are the same as in Section 8.3.3.

Since the permeability field k is a function of the Gaussian random variables θ, so are all the random solutions of equation (8.40). Denote their WCEs as

$$\begin{aligned} S(x,t) &= \sum_{\alpha \in \mathcal{J}} S_\alpha(x,t) T_\alpha(\theta), \\ v(x,t) &= \sum_{\alpha \in \mathcal{J}} S_\alpha(x,t) T_\alpha(\theta), \\ P(x,t) &= \sum_{\alpha \in \mathcal{J}} P_\alpha(x,t) T_\alpha(\theta). \end{aligned}$$

The WCE coefficients satisfy the following PDE system:

$$\begin{cases} \nabla \cdot \Big(\sum_{\gamma \in \mathcal{J}} A_{\alpha,\gamma}(x) \nabla P_\gamma(x) \Big) = q I_{\{\alpha=0\}}, \\ v_\alpha = - \sum_{\gamma \in \mathcal{J}} A_{\alpha,\gamma}(x) \nabla P_\gamma(x) \ , \\ \dfrac{\partial S_\alpha}{\partial t} + \sum_{p \in \mathcal{J}} \sum_{0 \le \beta \le \alpha} C(\alpha,\ \beta,\ p)\ \nabla \cdot (v_{\alpha-\beta+p}\ S_{\beta+p}) = 0, \end{cases} \tag{8.41}$$

where the coefficient $A_{\alpha,\gamma}(x)$ is defined by formula (7.24) and $C(\alpha, \beta, p)$ is defined by (2.39).

To solve the WCE system (8.41) numerically, we first solve the elliptic system by the upscaled WCE method (see Chapter 7). Then we construct the coarse-scale velocity field from the pressure solution, and use it resolve the coarse-scale saturation from the transport equation. This procedure is very similar to the upscaling of the deterministic single-phase flow problem. Once the WCE coefficients of the random solutions are available, we can construct the coarse-scale fractional flow map explicitly

$$F^*(\theta,t) = 1 - \frac{\int_{\partial\Omega^{out}} S\, v_n dl}{\int_{\partial\Omega^{out}} v_n dl} = 1 - \frac{\sum_\alpha f_\alpha(t)\, T_\alpha(\theta)}{\sum_\alpha g_\alpha(t)\, T_\alpha(\theta)}, \tag{8.42}$$

where

$$f_\alpha(t) = \sum_p \sum_{\beta \leqslant \alpha} \int_{\partial\Omega^{out}} C(\alpha,\beta,p) S_{\alpha-\beta+p} v_{\beta+p} \cdot n dl,$$

and

$$g_\alpha(t) = \int_{\partial\Omega^{out}} v_\alpha \cdot n dl.$$

With the coarse-scale fractional flow map $F^*(\theta,t)$ available explicitly, the coarse-scale distribution $\Pi^*(\theta)$ becomes explicit also. We can immediately generate permeability proposals based on the coarse-scale Langevin algorithm (8.39), and apply Algorithm III for the dynamic data integration. Since the formula for $\Pi^*(\theta)$ is available explicitly, we don't need to solve the nonlinear PDE system (8.1)–(8.3) repeatedly using the coarse-scale model based on the multiscale finite element. Instead, the upscaled WCE system (8.41) is solve only once. So the Algorithm III combines the merits of the Langevin algorithm and upscaled WCE method with that of the two-stage MCMC sampling. In a recent [18] work, we did extensive numerical experiments for the coarse-gradient Langevin algorithm combined with the two-stage MCMC method, but without using the upscaled WCE method as the coarse-scale model. Instead the regular coarse-scale model based on the multiscale finite element was used. That intermediate version of Algorithm III was shown to be ten times more efficient than the regular MCMC method. In the future, we will incorporate the upscaled WCE solver into the algorithm and avoid the repeated simulations of the coarse-model in the first stage. It is expected the upscaled WCE solver will reduce the computational cost further and accelerate the two-stage MCMC sampling.

Chapter 9

Concluding Discussions

In this thesis we explored a numerical method based on Wiener chaos expansions for solving SPDEs driving by Brownian motion forcing. Using the exact expansions for products of WCE bases, we were able to derive explicit Wiener chaos propagators for a wide range of SPDEs arising in fluid mechanics. We also derived analytical formulae for high order statistical moments in terms of the Wiener chaos coefficients. Based on the asymptotic decaying rate of Wiener chaos coefficients, we designed a sparse truncation strategy for the WCE propagator, which can substantially reduce the dimension of the problem while retaining the same asymptotic convergence rate. By building a generic and robust numerical solver, we can handle nonlinear WCE propagators with hundreds of PDEs in our computations. We applied the WCE method to a wide range of applications and carefully studied its performance. Our numerical results convincingly demonstrate that the WCE-based method is more efficient and accurate than MC simulation for short to moderate time solutions. To better understand the performance of the WCE method, we conducted a rigorous error analysis for a model equation, where a semi-analytical solution can be constructed. The analysis reveals the important convergence properties of the WCE method and provides guidance for the numerical computations.

Though quite efficient and accurate in short time intervals, the WCE method has serious difficulties for long time integrations. When the SPDEs are solved in longer time intervals, the required number of WCE coefficients will increase very quickly. If truncated at a limited order, the WCE solutions will have difficulty resolving the small structures of the random solutions. To solve this problem partially, we proposed a new computational strategy where MC simulations are used to correct the unresolved small scales in the Wiener chaos solutions. Since it is very easy to construct realizations from the Wiener chaos solution, we can subtract

the Wiener chaos solution from each MC realization, and hence substantially reduce the variance of the MC simulation. The numerical experiments demonstrate that this WCE-MC hybrid method can handle SPDEs in much longer time intervals or with larger random forcing than the direct WCE method can. The WCE-MC hybrid method is also shown to be more efficient than the Wiener chaos method or the MC simulation alone in relatively long time intervals.

Using the new numerical methods we had designed, we verified numerically the existence of a unique stationary measure for a stochastic Burgers equation. For the reaction-diffusion equation in random shear flows, our numerical results demonstrate that the front propagation speed obeys the quadratic enhancement law, which confirms the conjecture by Jack Xin et al. [76] about the precise asymptotical relations of the front speed in white-in-time shear flows.

In the second part of the thesis, we revisited the stochastic finite element method for solving stochastic elliptic equations, and resolved a few technical difficulties in this approach. We further derived an upscaling formulation for the elliptic system of the Wiener chaos coefficients, and applied it for uncertainty quantification in subsurface modeling. Uncertainty quantification is a very important research topic in petroleum engineering. An effective way to reduce the uncertainties in the reservoir simulation is to sample the reservoir properties conditioned on dynamical production data. For this purpose, we have developed a two-stage MCMC method to sample the reservoir permeability. The new method employs multiscale finite element method as the coarse-scale model and uses it to filter unacceptable permeability proposals. The coarse-scale model can effectively reduce the fine-scale computational costs and increase the acceptance rate of the MCMC sampling. We applied the new method to both single-phase flows and two-phase flows. Numerical results demonstrate that the two-stage MCMC method is ten times more efficient than the regular MCMC sampling method.

9.1 Future Work

The WCE-MC method can extend the application of the WCE method to much longer time intervals. However, the WCE-MC method is still sensitive to the size of the time interval. In an arbitrarily long time interval, the sparse WCE solution can only resolve

limited information of the random solution, hence the WCE-MC will not be much more efficient than the MC simulation combined with other acceleration techniques. Designing efficient numerical methods based on Wiener chaos expansion remains to be an open and challenging research problem.

In our current study, we choose to apply the WCE method to a single large time intervals. Since Brownian motion forcing has independent increments, it would be more desirable to divide the large time interval into small ones, and apply the WCE method repeatedly on those small intervals. The difficulty with this approach is that we have to deal with random initial conditions with rapidly increasing dimensions. At each small time interval, the initial condition is the final solution in the previous time interval, so it is already represented as a Wiener chaos expansion of the historical randomness. The dimension of the random initial condition will increase very quickly. How to compress efficiently the random initial conditions is the key for this approach to work. Initially, we tried to compress the initial conditions by its Karhunen-Loeve expansion (7.5). As we have pointed out, the Karhunen-Loeve expansion is an optimal expansion and hence an ideal candidate for the compression purpose. However, the initial conditions on the subsequent small time intervals are non-Gaussian processes. As a result, the uncorrelated random variable θ_k in the KLE of the initial conditions are not Gaussian variables. In fact, their distributions are not known explicitly, and we don't know how to incorporate them in the framework of the Wiener chaos expansion. In other words, it is very difficult to evolve the initial randomness correctly if we compress them by their KLEs. How to effectively reduce the dimension of the random initial condition is the most difficult part in designing a practical localized WCE method. Much more research efforts are required to make any progress along this direction.

Using the WCE method, we can reduce the stochastic reaction-diffusion equation (6.22) into a deterministic system, which opens a door for analyzing the asymptotical behavior of the stochastic system. Since the front propagation problems have been well studied in deterministic cases, we can naturally translate those techniques and results to the deterministic WCE system, and use it to analyze the properties of the original stochastic systems. This approach is quite different from the variational principles used by Xin et al. [75, 76], and is expected to be able to bring new perspectives to the research of the stochastic reaction-diffusion problems.

Appendix A

Stochastic Integrals

In this appendix we will give an introduction to the stochastic integrals, both in Ito's sense and Strotonovich's sense. The presentation of this appendix follows the textbook by Øksendal [77], and the note by Evans [27].

Suppose W_t is a one-dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, P)$. Since W_t is nowhere differentiable, the stochastic differential equation

$$dX_t = a(X_t, t)dt + b(X_t, t)\, dW_t$$

is always interpreted as an integral equation

$$X_t = X_0 + \int_0^t a(X_s, s)ds + \int_0^t b(X_s, s)dW_s. \tag{A.1}$$

To study the integral equation (A.1), we must first define the stochastic integral

$$\int_0^T f(t, \omega)\, dW_t \tag{A.2}$$

for a wide class of stochastic process $f(t, \omega)$. Since W_t has unbounded total variation, the stochastic integral (A.2) will behave quite differently from the deterministic integrals.

Denote $\mathcal{F}_t$ as the σ-algebra generated by the Brownian motion $\{W_s(\omega);\ s \leqslant t\}$. $\{\mathcal{F}_t\}_{t\geqslant 0}$ is an increasing family of σ-algebra in the probability space $(\Omega, \mathcal{F}, P)$. A random process $f(t, \omega)$ is called $\mathcal{F}_t$-adapted if for each $t \geqslant 0$ the random variable $f(t, \omega)$ is $\mathcal{F}_t$-measurable. We first describe the class of functions for which the stochastic integral (A.2) will be defined.

Definition A.1 *Let $\mathcal{V}(0,T)$ be a class of stochastic process*

$$f(t,\omega) : [0,\infty) \times \Omega \to R$$

such that

(i) $(t,\omega) \to f(t,\omega)$ *is* $\mathcal{B} \times \mathcal{F}$*-measurable, where* $\mathcal{B}$ *is the Borel* σ*-algebra on* $[0,\infty)$*,*

(ii) $f(t,\omega)$ *is* $\mathcal{F}_t$*-adapted,*

(iii) $E\int_0^T f^2(t,\omega)dt < \infty$*.*

Suppose $0 = t_0 < t_1 < \ldots < t_N = T$ is a partition of the time interval $[0,T]$. Denote $\Delta = \max\{\Delta t_n\}$, where $\Delta t_n = t_{n+1} - t_n$. For any $f(t,\omega) \in \mathcal{V}(0,T)$, the *Ito integral* of (A.2) is defined as

$$\int_0^T f(t,\omega) \cdot dW_t = \lim_{\Delta \to 0} \sum_{n=0}^{N-1} f(t_n,\omega)\,(W_{n+1} - W_n)\,, \tag{A.3}$$

where $W_n = W(t_n,\omega)$, and the *Stratonovich integral* of (A.2) is defined as

$$\int_0^T f(t,\omega) \circ dW_t = \lim_{\Delta \to 0} \sum_{n=0}^{N-1} f(t_{n+1/2},\omega)\,(W_{n+1} - W_n)\,, \tag{A.4}$$

where $t_{n+1/2} = (t_n + t_{n+1})/2$. Both (A.3) and (A.4) converge in the mean square sense in the probability space $(\Omega, \mathcal{F}, P)$.

The Ito integral and Stratonovich integral are particular cases of the regular Riemann-Stieltjes integrals

$$\int_0^T f(t)dt = \lim_{\Delta \to 0} \sum_{n=0}^{N-1} f(t_n^*)\,(t_{n+1} - t_n),$$

where the representative point $t_n^* \in [t_n, t_{n+1}]$ is arbitrary. In the Ito integral we take $t_n^* = t_n$, while in the Stratonovich integral we take $t_n^* = (t_n + t_{n+1})/2$. Due to the unbounded variation of W_t, the different choices of t_n^* indeed result in quite different stochastic integrals. Next we will describe the differences and connections between Ito integrals and Stratonovich integrals.

Definition A.2 *A stochastic process* $\{M_t\}_{t \geqslant 0}$ *on* $(\Omega, \mathcal{F}, P)$ *is called a martingale with respect to the increasing* σ*-algebra family* $\{\mathcal{F}_t\}_{t \geqslant 0}$ *if*

(i) M_t *is* $\mathcal{F}_t$*-measurable for all* t*,*

(ii) $E|M_t| < \infty$ *for all* t,

(iii) $E(M_s|\mathcal{F}_t) = M_t$ *for all* $s \geqslant t$.

The Ito integral (A.3) has the following properties:

Theorem A.1

(i) The Ito integral (A.3) is mean zero: $E\left[\int_0^T f(t,\omega)\cdot dW_t\right] = 0$,

(ii) $F(t,\omega) = \int_0^t f(s,\omega)\cdot dW_s$ *is a martingale with respect to* $\{\mathcal{F}_t\}_{t\geqslant 0}$,

(iii) The Ito integral (A.3) satisfies the isometry property

$$E\left[\int_0^T f(t,\omega)\cdot dW_t\right]^2 = E\int_0^T f^2(t,\omega)dt. \tag{A.5}$$

In general, the above properties are not shared by the Stratonovich integrals (A.4). That is, the Stratonovich integral is not mean zero, not a martingale, and does not satisfy the isometry relation (A.5). Because of the nice properties of the Ito integrals, they are usually easier to handle than Stratonovich integrals. For this reason, it is important that we known how to transform one integral form to the other. Suppose the stochastic integral equation (A.1) is proposed in Stratonovich's sense:

$$X_t = X_0 + \int_0^t a(X_s,s)ds + \int_0^t b(X_s,s)\circ dW_s, \tag{A.6}$$

then it is equivalent to the following integral equation in Ito's sense

$$X_t = X_0 + \int_0^t a(X_s,s)ds + \frac{1}{2}\int_0^t b_x b(X_s,s)dt + \int_0^t b(X_s,s)\cdot dW_s. \tag{A.7}$$

Using the above relation, we can easily transform the Stratonovich integral into the Ito integral, and vice versa. If $b = b(t)$ does not depend on X_t, then the Stratonovich integral is the same as the Ito integral.

Due to the unbounded total variation of the Brownian motion, the calculus rules of the stochastic integrals are quite different from the deterministic ones. In particular, stochastic integrals have different differential chain rule, which is described by the Ito formula.

Theorem A.2 (the one-dimensional Ito formula)

Let X_t be a stochastic process satisfying the stochastic differential equation

$$d\,X_t = a\,dt + b\,d\,W_t.$$

Let $g(x,t)$ be a function with continuous derivatives $\frac{\partial g}{\partial t}$, $\frac{\partial g}{\partial x}$ and $\frac{\partial^2 g}{\partial x^2}$. Then the stochastic process

$$Y_t = g(X_t, t)$$

satisfies the differential equation

$$dY_t = \frac{\partial g}{\partial t}(X_t,t)dt + \frac{\partial g}{\partial x}(X_t,t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t,t)\cdot(dX_t)^2, \tag{A.8}$$

where $(dX_t)^2 = dX_t \cdot dX_t$ is computed according to the rules

$$dt\cdot dt = dt\cdot dW_t = dW_t\cdot dt = 0, \qquad dW_t\cdot dW_t = dt.$$

The Ito formula (A.8) describes the differential chain rule for functions of stochastic processes. Comparing with the deterministic chain rule, the stochastic chain rule (A.8) has an extra term

$$\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t,t)\cdot(dX_t)^2 = \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t,t)\cdot\, b^2(dW_t)^2.$$

Heuristically, the "derivative" of the stochastic process Y_t is the leading order term in the Taylor expansion:

$$\begin{aligned} Y(t+\Delta t) - Y(t) &= \frac{\partial g}{\partial t}(X_t,t)\Delta t + \frac{\partial g}{\partial x}(X_t,t)\Delta X_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t,t)\cdot\, b^2(\Delta W_t)^2 + O(\Delta t\Delta W + \Delta t^2) \\ &\simeq A\Delta t + O(\Delta t\Delta W + \Delta t^2). \end{aligned}$$

Since the Brownian motion increment $\Delta W_t = W(t+\Delta t) - W(t)$ has the Holder property

$$E[\Delta W_t]^2 = \Delta t,$$

we need to include the term $(\Delta W_t)^2$ in the leading order term $A\Delta t$. That is why the stochastic chain rule is different from the deterministic chain rule.

Appendix B

Monte Carlo Simulations of Stochastic Differential Equations

In this appendix, we will briefly discuss how to solve a general stochastic differential equation (SDE)

$$dX_t = a(t, X_t)dt + b(t, X_t) \cdot dW(t). \tag{B.1}$$

by Monte Carlo simulations. For more details, please see the comprehensive references [53, 84].

The idea of Monte Carlo simulation is very simple: sample the random forcing and solve the equation (B.1) realization by realization. For each given sample path (realization) of the Brownian motion, the equation (B.1) becomes deterministic. It may be tempting to solve the resulting equations by the well-developed deterministic numerical schemes. However, the Brownian motion trajectory has unbounded variation and is nowhere differentiable. Due to the unique differential rules (Ito formula) of stochastic integrals, the numerical schemes for discretizing B.1) are quite different from the ones for smooth deterministic differential equations. In this appendix we will briefly discuss the subtle nature of time discretizations for SDEs and present a few popular numerical schemes for solving SDEs. We will also introduce several useful variance reduction techniques in MC simulations.

B.1 Strong and Weak Convergence

Suppose

$$0 = t_0 < t_1 < t_2 < \ldots < t_N = T$$

is a uniform partition of the time interval $[0, T]$, where $t_n = nh$ and $h = \frac{T}{N}$. Suppose $X_h(t)$ is a discrete solution to (B.1). Then $X_h(t)$ should be an approximation to the true solution $X(t)$ in some sense. Motivated by the needs of applications, there are two major approximation criteria:

(a) approximate the same path of the random solution for each Brownian motion realization, or

(b) approximate the distributions of the random solution.

The first approximation criterion is useful when we need to study individual realizations of the random solution. The second criteria is useful if we are only interested in the statistical property of the random solution. Accordingly, the numerical schemes for solving SDEs can be roughly divided into two categories: strongly convergent schemes and weakly convergent schemes.

Definition B.1 *Strong convergence: We say that the numerical solution X_h converges to the solution X_t strongly with order $\gamma > 0$ if there exists a finite constant K such that*

$$E(|X_t - X_h|) \leq Kh^{\gamma}.$$

Strong convergence emphasizes that the numerical solution approximates the exact solution realization by realization. However, approximating the solution realization by realization is not necessary in many practical situations. For example, if only the statistic moments of the random solution are the interests, approximating the distribution of the solution will be suffice. That is why the weak convergence is considered:

Definition B.2 *Weak convergence: Suppose that all the statistical moments of X_t exist. We say that X_h converges to X_t weakly with order β if for any polynomial $g(x)$ there exists a finite constant K_g such that*

$$\big|E[g(X_t)] - E[g(X_h)]\big| \leq K_g h^{\beta}.$$

Weak convergence emphasizes that the numerical solution has approximate probability distribution as the true solution. Weak convergence leaves much more freedom in designing numerical schemes for SDEs. To approximate the distribution of X, the numerical solution

X_h can even in a different probability space than X. Take $g(x) = x^n$, we immediately conclude that all the statistical moments of the numerical solution will converge under the weak convergence criterion. That is why weak convergence is useful in practice.

To see the difference between strong convergence and weak convergence, let us consider the Euler discretization of (B.1):

$$X_{n+1} = X_n + a(t_n, X_n)h + b(t_n, X_n)\Delta W_n, \tag{B.2}$$

where $\Delta W_n = W(t_{n+1}) - W(t_n)$. If b depends on X_t, it is known that the Euler scheme (B.2) converges strongly with order $\gamma = \frac{1}{2}$, and weakly with order $\beta = 1$. That is quite different from the deterministic Euler scheme, which always has first order accuracy. The reason is that ΔW_n has the order of $(\Delta t)^{1/2}$ only. To achieve first order strong convergence in (B.2), we also need to include the terms depending on $(\Delta W_n)^2$. To modify the Euler scheme, Milstein [73] proposed the following strongly first order numerical scheme:

$$X_{n+1} = X_n + a(t_n, X_n)h + b(t_n, X_n)\Delta W_n + \frac{1}{2} b\, b_x(t, X)\{(\Delta W_n)^2 - h\}. \tag{B.3}$$

Comparing with the Euler scheme (B.2), the Milstein scheme has an extra term that involves $(\Delta W_n)^2$. That is why it can achieve strongly first order accuracy. For additive forcing where b does not depend on X_t, the Milstein scheme (B.3) reduces to the Euler scheme. Hence for additive random forcing, the Euler scheme converges with first order for both weak and strong convergence criteria.

Generally, the strongly convergent numerical schemes are much more complicated than the weakly convergent ones. Strongly convergent numerical schemes with orders higher than 1 are all unpleasantly complicated. In this thesis, we will use the Milstein scheme only when strong convergence is needed. For the weak convergence, we use the popular weakly second order Runge-Kutta method [53] throughout this paper:

$$\begin{aligned} X_{n+1} = X_n \quad &+ \quad \frac{1}{2}\left[a(t_n, X_n) + a(t_{n+1}, \hat{X})\right] h \\ &+ \quad \frac{1}{4}\left[2b(t_n, X_n) + b(t_{n+1}, X^+) + b(t_{n+1}, X^-)\right] \Delta W_n \\ &+ \quad \frac{1}{4}\left[b(t_{n+1}, X^+) - b(t_{n+1}, X^-)\right] \left\{(\Delta W_n)^2 - h\right\} h^{-1/2}, \end{aligned} \tag{B.4}$$

where

$$\hat{X} = X_n + a(t_n, X_n)h + b(t_n, X_n)\Delta W_n$$

and

$$X^{\pm} = X_n + a(t_n, X_n)h \pm b(t_n, X_n)\sqrt{h}.$$

The above numerical scheme has a weakly second order accuracy. Comparing with the first order Euler scheme, the second order scheme involves the extra terms $\Delta t \Delta W_n$ and $(\Delta W_n)^2 - \Delta t$, which have the order of $(\Delta t)^{\frac{3}{2}}$. That is why the above numerical scheme can achieve weakly second order accuracy $O(\Delta t)^2$ in the time discretization. For stochastic differential equations with additive forcing, those extra terms will disappear and the numerical scheme (B.4) is reduced to

$$X_{n+1} = X_n + \frac{1}{2}\left[a(t_n, X_n) + a(t_{n+1}, \hat{X})\right]h + \frac{1}{2}[b(t_n) + b(t_{n+1})]\Delta W_n.$$

B.2 Variance Reduction of MC Simulations

In lots of applications, the purpose of MC simulation is to estimate the expectation $E[g(X_t)]$, where $g(x)$ is a deterministic function. For example, we can set $g(x) = x^n$ if we want to compute the statistic moments of the random solution. The MC estimation of the expectation takes the form

$$E[g(X_t)] \simeq \frac{1}{N}\sum_{k=1}^{N} g(X_t(\omega_k)), \tag{B.5}$$

where $X_t(\omega_k), k = 1, 2, \ldots N$ are independent realizations of the random solution, and N is the total number of realizations. Denote the error of the estimation as

$$\epsilon_g(N) = E[g(X_t)] - \frac{1}{N}\sum_{k=1}^{N} g(X_t(\omega_k)).$$

The error of the estimator is a random variable itself. By the Central Limit Theorem [30], the error decays as

$$\sqrt{E(\epsilon_g^2)} \simeq \frac{\sigma[g(X_t)]}{\sqrt{N}},$$

where $\sigma^2[g(X)]$ is the variance

$$\sigma^2[g(X)] = \int \big[g(X) - E\, g(X)\big]^2 d\omega.$$

The error of the MC estimator decays rather slowly, in the order of $O(N^{-1/2})$. If we want to reduce the error by a factor of 2, we need to increase the realization number by a factor of 4. Figure B.1 is a typical convergence behavior of MC simulations with different realization numbers. The MC estimator converges faster at the beginning but slows down quickly. Eventually it reaches a flat region where the gain by increasing the number of realizations becomes smaller and smaller.

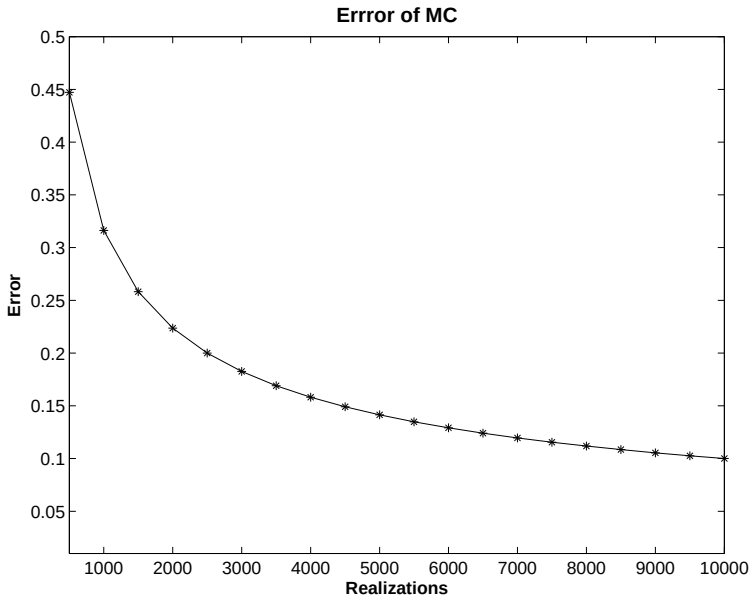

Figure B.1: Asymptotic convergence rate of MC simulations with different realization numbers. The MC method converges faster at the beginning but slows down quickly. Eventually it reaches a flat region where incremental computational costs only improve the estimation very slightly.

There are two typical ways to accelerate the convergence of the MC estimation: increase the order of the convergence, or reduce the proportional constant $\sigma[g(X)]$. For the first approach, a modification of the regular MC method, called quasi Monte Carlo method [9], can increase the convergence order of MC from $O(N^{-1/2})$ to $O\left((\log N)^k N^{-1}\right)$. This is a nearly optimal convergence rate for MC estimations since only N number of realizations are computed. The idea of quasi MC method is to generate deterministic (rather than random or pseudo-random) sequences, which are correlated (rather than independent) to provide greater uniformity. quasi MC is designed specifically for numerical integration purpose and

is not directly applicable for simulating solutions of stochastic differential equations.

Another way to accelerate the MC method is to reduce the variance of the estimator $\sigma[g(X)]$. Suppose the error tolerance of the MC estimation is ε, then the number of required MC realizations to achieve such an accuracy would be

$$N \simeq \frac{\sigma^2[g(X)]}{\varepsilon^2}.$$

For a fixed error ε, if we can reduce $\sigma[g(X)]$ by a factor of 2, then the required number of realization can be reduced by a factor of 4. So variance reduction is a quite effective way to improve the convergence of the MC method. Next we will introduce two popular variance reduction methods which can be applied to simulate statistics of SDEs.

B.2.1 Antithetic Variables

Antithetic variable method is one of the simplest and most widely used variance reduction techniques. The method works as follows: for each sample path ω_k of the Brownian motion, also use the sample path $-\omega_k$. The resulting MC estimator is

$$E[g(X)] \simeq \frac{1}{2N}\sum_{n=1}^{N}\Big\{g(X(\omega_k)) + g(X(-\omega_k))\Big\}.$$

For example, suppose $\omega_k = (\Delta W_1, \Delta W_2, \dots \Delta W_L)$ is a discrete sample of the Brownian motion increments, then we also include $-\omega_k = (-\Delta W_1, -\Delta W_2, \dots, -\Delta W_L)$ as a sample of the Brownian motion. The use of antithetic variables is motivated by an expansion for small values of the variance. Consider, for example, the expectation $E[g(X)]$ in which X is an $N(0, \sigma^2)$ random variable with small σ. Set $X = \sigma\hat{X}$. The Taylor expansion of $g(X)$ for small σ is

$$g(X) = g(0) + g(0)\sigma\hat{X} + O(\sigma^2).$$

Since the distribution of $\hat{X}$ is symmetric about 0, the average $E(\hat{X})$ is 0. In a standard MC estimation of $E[g(X)]$, this term does not cancel exactly, so that the MC error is proportional to σ. With antithetic variables, on the other hand, the linear term cancel exactly, so that the remaining error is proportional to σ^2.

The antithetic variable method is very easy to apply and can be combined with other variance reduction method, such as the control variate method. However, for nonlinear

SDE, the random solution X is by no means a Gaussian random variable or symmetric. The acceleration provided by antithetic variables is quite limited, as our numerical experiments with nonlinear SPDEs have demonstrated.

B.2.2 Control Variates

The idea of control variate is to use a function $p(X)$, which has similar variability as $g(X)$ and has known expectation $E[p(X)]$. Then the expectation $E[g(X)]$ can be rewritten as

$$E[g(X)] = E[p(X)] + E[g(X) - p(X)]\,.$$

Since $E[p(X)]$ is already known, the MC method is employed to compute $E[g(X) - p(X)]$ only:

$$E[g(X) - p(X)] \simeq \frac{1}{N}\sum_{k=1}^{N}\big\{g(X(\omega_k)) - p(X(\omega_k))\big\}. \tag{B.6}$$

The error of the above MC estimator has the form

$$\sqrt{E(\epsilon_{g-p}^2)} \simeq \frac{\sigma_{g-p}}{\sqrt{N}},$$

where σ_{g-p}^2 is the variance of the random function $g(X) - p(X)$:

$$\sigma_{g-p}^2 = \int \Big\{[g(X) - E\,g(X)] - [p(X) - E\,p(X)]\Big\}^2 d\omega. \tag{B.7}$$

If $p(X)$ is a good approximation of $g(X)$ in the mean square sense and hence

$$\sigma_{g-p} \ll \sigma_g,$$

then the ensemble average (B.6) will converge much faster than the ensemble average (B.5). For example, if σ_{g-p} is only one-tenth of σ_g, then the convergence of (B.6) can be accelerated 100 times. Note that to make the control variate work efficiently, the mean of $p(X)$ does not have to approximate the mean of $g(X)$. As we can see from (B.7), the crucial point is that $p(X)$ fluctuate around its mean in a similar way as $g(X)$ does.

There exist many other variance reduction techniques for MC simulations, such as importance sampling, matching moments, stratification, etc. However, those methods are more suitable for MC integrations than for simulating stochastic differential equations. For

more details, please see [9, 84] and the references therein.

Appendix C

Coarse-scale Models Using Multiscale Finite Element Methods

In this appendix, we will describe the details of the coarse-scale model we used in Section 8.3 for preconditioning MCMC simulations. As we noted earlier, this model is similar to single-phase upscaling methods. However, instead of coarsening the absolute permeability, we use pre-computed multiscale finite element basis functions. The key idea of the method is the construction of basis functions on the coarse grids such that these basis functions capture the small scale information on each of these coarse grids. The method that we use follows its finite element counterpart presented in [47]. The basis functions are constructed from the solution of the leading order homogeneous elliptic equation on each coarse element with some specified boundary conditions. Thus if we consider a coarse element K which has d vertices, the local basis functions $\phi^i, i = 1, \cdots, d$ satisfy the following elliptic problem:

$$\begin{aligned} -\nabla \cdot (k \cdot \nabla \phi^i) &= 0 \quad \text{in } K \\ \phi^i &= g^i \quad \text{on } \partial K, \end{aligned} \tag{C.1}$$

for some function g^i defined on the boundary of the coarse element K. Hou et al. [47] have demonstrated that a careful choice of boundary condition would guarantee the performance of the basis functions to incorporate the local information and hence improve the accuracy of the method. The function g^i for each i varies linearly along ∂K. Thus, for example, in the case of a constant diagonal tensor the solution of (C.1) would be a standard linear/bilinear basis function. We note that, as usual, we require $\phi^i(\xi_j) = \delta_{ij}$. Finally, a nodal basis function associated with the vertex ξ in the domain Ω is constructed from the combination of the local basis functions that share this ξ and are zero elsewhere. These nodal basis

functions are denoted by $\{\psi_\xi\}_{\xi \in Z_h^0}$.

Having described the basis functions, we denote the space of our approximate pressure solution as V^h, which is spanned by the basis functions $\{\psi_\xi\}_{\xi \in Z_h^0}$. Now we may formulate the finite dimensional problem corresponding to the finite volume element formulation of (8.2). A statement of mass conservation on a control volume V_ξ is formed from (8.2), where now the approximate solution is written as a linear combination of the basis functions. Assembly of this conservation statement for all control volumes would give the corresponding linear system of equations that can be solved accordingly. The resulting linear system has incorporated the fine-scale information through the involvement of the nodal basis functions on the approximate solution. To be specific, the problem now is to seek $p^h \in V^h$ with $p^h = \sum_{\xi \in Z_h^0} p_\xi \psi_\xi$ such that

$$\int_{\partial V_\xi} \lambda(S) k \cdot \nabla p^h \cdot \vec{n}\, dl = \int_{V_\xi} f\, dA, \tag{C.2}$$

for every control volume $V_\xi \subset \Omega$. Here $\vec{n}$ defines the unit normal vector on the boundary of the control volume, ∂V_ξ, and S is the fine scale saturation field at this point. We note that, concerning the basis functions, a vertex-centered finite volume difference is used to solve (C.1), and a harmonic average is employed to approximate the permeability k at the edges of fine control volumes.

Furthermore, the pressure solution may be used to compute the total velocity field at the coarse-scale level, denoted by $\overline{v} = (\overline{v}_x, \overline{v}_z)$ via (8.5). In general, the following equations are used to compute the velocities in horizontal and vertical directions, respectively:

$$\overline{v}_x = -\frac{1}{h_z} \sum_{\xi \in Z_h^0} p_\xi \left(\int_E \lambda(S) k_x \frac{\partial \psi_\xi}{\partial x} dz \right), \tag{C.3}$$

$$\overline{v}_z = -\frac{1}{h_x} \sum_{\xi \in Z_h^0} p_\xi \left(\int_E \lambda(S) k_z \frac{\partial \psi_\xi}{\partial z} dx \right), \tag{C.4}$$

where E is the edge of V_ξ. Furthermore, for the control volumes V_ξ adjacent to the Dirichlet boundary (which are half control volumes), we can derive the velocity approximation using the conservation statement derived from (8.2) on V_ξ. One of the terms involved is the integration along part of the Dirichlet boundary, while the rest of the three terms are known from the adjacent internal control volumes calculations. The analysis of the two-scale finite

volume method can be found in [40].

As for the upscaling of the saturation equation, we only use the coarse scale velocity to update the saturation field on the coarse-grid, i.e.,

$$\frac{\partial \overline{S}}{\partial t} + \overline{v} \cdot \nabla f(\overline{S}) = 0, \tag{C.5}$$

where $\overline{S}$ denotes the saturation on the coarse-grid. In this case the upscaling of the saturation equation does not take into account subgrid effects. This kind of upscaling techniques in conjunction with the upscaling of absolute permeability are commonly used in applications (see e.g. [20, 21, 22]). The difference of our approach is that the coupling of the small scales is performed through the finite volume element formulation of the pressure equation.

Bibliography

[1] J. Aarnes. On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation. *SIAM Multiscale Model. Simul.*, 2:421–439, 2004.

[2] U. M. Ascher, S. J. Ruuth, and B. Wetton. Implicit-explicit methods for time-dependent partial differential equations. *SIAM J. Numer. Anal.*, 32:797–823, 1995.

[3] I. Babuska and P. Chatzipantelidis. On solving elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 191(37-38):4093–4122, 2002.

[4] I. Babuska, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825, 2004.

[5] I. Babuska, R. Tempone, and G. E. Zouraris. Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation. *Comput. Methods Appl. Mech. Engrg.*, 194:1251–1294, 2005.

[6] J. Bec, U. Frisch, and K. Khanin. Kicked Burgers turbulence. *J. Fluid Mech.*, 418:239–267, 2000.

[7] A. Bensoussan and R. Temam. Equations stochastiques du type Navier-Stokes. *J. Func. Anal.*, 13:195–222, 1973.

[8] J. P. Boyd. *Chebyshev and Fourier Spectral Methods*. Dover Publication, New York, 2000.

[9] R. E. Caflisch. Monte Carlo and Quasi-Monte Carlo methods. *Acta Numerica*, pages 1–49, 1998.

[10] R. H. Cameron and W. T. Martin. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Ann. Math.*, 48:385–392, 1947.

[11] C. C. Chang. Numerical solution of stochastic differential equations with constant diffusion coefficients. *Math. Comput.*, 49(180):523–542, 1987.

[12] A. J. Chorin. Hermite expansion in Monte-Carlo simulations. *J. Comput. Phys.*, 8:472–482, 1971.

[13] A. J. Chorin. Gaussian fields and random flow. *J. Fluid Mech.*, 63:21–32, 1974.

[14] A. Christen and C. Fox. MCMC using an approximation. Technical report, Department of Mathematics, The University of Auckland, New Zealand, 2004.

[15] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Wiley-Interscience, New York, 1953.

[16] S. C. Crow and G. H. Canavan. Relationship between a Wiener-Hermite expansion and an energy cascade. *J. Fluid Mech.*, 41:387–403, 1970.

[17] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.

[18] P. Dostert, Y. Efendiev, T. Hou, and W. Luo. Coarse-gradient Langevin algorithms for dynamic data integration and uncertainty quantification. *accepted by J. Comput. Phys.*, 2006.

[19] L. J. Durlofsky. Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. *Water Resour. Res.*, 27:699–708, 1991.

[20] L. J. Durlofsky. Coarse scale models of two phase flow in heterogeneous reservoirs: Volume averaged equations and their relationship to the existing upscaling techniques. *Computational Geosciences*, 2:73–92, 1998.

[21] L. J. Durlofsky, R. A. Behrens, R. C. Jones, and A. Bernath. Scale up of heterogeneous three dimensional reservoir descriptions. *SPE paper*, page 30709, 1996.

[22] L. J. Durlofsky, R. C. Jones, and W. J. Milliken. A nonuniform coarsening approach for the scale up of displacement processes in heterogeneous media. *Advances in Water Resources*, 20:335–347, 1997.

[23] W. E. and E Vanden Eijnden. Statistical theory for the stochastic Burgers equation in the inviscid limit. *Comm. Pure Appl. Math.*, 53:852–901, 2000.

[24] W. E., K. Khanin, A. Mazel, and Y. Sinai. Probability distribution functions for the random forced Burgers equation. *Phys. Rev. Lett.*, 78:1904–1907, 1997.

[25] W. E., K. Khanin, A. Mazel, and Y. Sinai. Invariant measures for Burgers equation with stochastic forcing. *Ann. Math.*, 151:877–960, 2000.

[26] Y. Efendiev, V. Ginting, T. Hou, and R. Ewing. Accurate multiscale finite element methods for two-phase flow simulations. *Submitted*, 2005.

[27] L. C. Evans. *An Introduction to Stochastic Differential Equations.* Lecture Notes, Department of Math, UC Berkeley.

[28] A. Fannjiang and G. Papanicolaou. Convection enhanced diffusion for periodic flows. *SIAM J. Appl. Math.*, 54(2):333–408, 1994.

[29] A. Fannjiang and G. Papanicolaou. Convection-enhanced diffusion for random flows. *J. Statist. Phys.*, 88:1033–1076, 1997.

[30] W. Feller. *An Introduction to Probability Theory and Its Application*, volume 1. John Wiley & Sons, third edition, 1968.

[31] F. Flandoli and F. Gozzi. Kolmogorov equation associated to a stochastic Navier-Stokes equation. *J. Funct. Anal.*, 160:312–336, 1998.

[32] P. Frauenfelder, C. Schwab, and R. A. Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Method Appl. Mech. Engrg.*, 194:205–228, 2005.

[33] R. Ghanem. The nonlinear Gaussian spectrum of log-normal stochastic processes and variables. *J. Appl. Mech. Trans.*, 66(4):964–973, 1999.

[34] R. Ghanem and S. Dham. Stochastic finite element analysis for multiphase flow in heterogeneous porous media. *Transp. Porous Media*, 32(3):239–262, 1998.

[35] R. G. Ghanem. Probablistic characterization of transport in heterogeneous media. *Comput. Methods Appl. Mech. Engrg.*, 158:199–220, 1998.

[36] R. G. Ghanem. Stochastic finite element with multiple random non-Gaussian properties. *J. Engrg. Mech.*, 125(1):26–40, 1998.

[37] R. G. Ghanem. Ingredients for a general purpose stochastic finite element formulation. *Comput. Methods Appl. Mech. Engrg.*, 168:19–34, 1999.

[38] R. G. Ghanem and J. Red-Horse. Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach. *PHYSICA D*, 133(1-4):137–144, 1999.

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

[40] V. Ginting. Analysis of two-scale finite volume element method for elliptic problem. *J. Numer. Math.*, 12:119–142, 2004.

[41] U. Grenander and M. I. Miller. Representations of knowledge in complex systems (with discussion). *J. R. Statist. Soc. B*, 56:549–603, 1994.

[42] W. D. Henshaw, H.-O. Kreiss, and L. G. Reyna. Smallest scale estimates for the Navier-Stokes equations for incompressible fluids. *Arch. Rational Mech. Anal.*, 112:21–44, 1990.

[43] T. Hida, H. H. Kuo, J. Potthoff, and L. Streit. *White Noise*. Kluwer Academic Publishers, Dordrecht, 1993.

[44] D. Higdon, H. Lee, and Z. Bi. A Bayesian approach to characterizing uncertainty in inverse problems using coarse and fine-scale information. *IEEE Transactions on Signal Processing*, 50:388–399, 2002.

[45] H. Holden, B. Øksendal, J. Uboe, and T. Zhang. *Stochastic Partial Differential Equations. A Modeling, White Noise Functional Approach*. Birkhäuser, Boston, 1996.

[46] T. Y. Hou, J. S. Lowengrub, and M. J. Shelley. Removing the stiffness from interfacial flows with surface tension. *J. Comput. Phys.*, 114:312–338, 1994.

[47] T. Y. Hou and X. H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134:169–189, 1997.

[48] M. Jardak, C. H. Su, and G. E. Karniadakis. Spectral polynomial chaos solutions of the stochastic advection equation. *J. Sci. Comput.*, 17:319–338, 2002.

[49] P. Jenny, S. H. Lee, and H. Tchelepi. Multi-scale finite volume method for elliptic problems in subsurface flow simulation. *J. Comput. Phys.*, 187:47–67, 2003.

[50] P. Jenny, S. H. Lee, and H. Tchelepi. Adaptive multi-scale finite volume method for multi-phase flow and transport in porous media. *SIAM Multiscale Model. Simul.*, 3:30–64, 2005.

[51] A. Keese. A review of recent developements in the numerical solution of stochastic partial differential equations (stochastic finite elements). Technical report, Institute of Scientific Computing, Technical University Braunschweig, Brunswick, Germany, October, 2003.

[52] J. Kevorkian. *Partial Differential Equations, Analytical Solution Techinques.* Springer, 2000.

[53] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, 1992.

[54] R. H. Kraichnan. Small-scale structure of a scalar field convected by turbulence. *Phys. Fluids*, 11:945–953, 1968.

[55] R. H. Kraichnan. Anomalous scaling of a randomly advected passive scalar. *Phys. Rev. Lett.*, 72:1016–1019, 1994.

[56] R. H. Kraichnan. Note on forced Burgers turbulence. *Phys. Fluid*, 11:3738–3742, 1999.

[57] R. Kress. *Linear Integral Equations.* Springer, second edition, 1999.

[58] J. S. Liu. *Monte Carlo Strategies in Scientific Computing.* Springer, New-York, 2001.

[59] M. Loeve. *Probability Theory.* Springer, Berlin, fourth edition, 1977.

[60] S. Lototsky, R. Mikulevicius, and B. Rozovskii. Nonlinear filtering revisited: a spectral approach. *SIAM J. Control Optim.*, 35:435–461, 1997.

[61] S. V. Lototsky. Nonlinear filtering of diffusion processes in correlated noise: analysis by separation of variables. *Appl. Math. Optim.*, 47(2):167–194, 2003.

[62] J. S. Lowengrub, M. J. Shelley, and B. Merriman. High-order and efficient methods for the vorticity formulation of the Euler equations. *SIAM J. Sci. Comput.*, 14:1107–1142, 1993.

[63] Z. Lu and D. Zhang. A comparative study on uncertainty quantification for flow in randomly heterogeneous media using Monte Carlo simulations and conventional and KL-based moment-equation approaches. *SIAM J. Sci. Comput.*, 26(2), 2004.

[64] O. P. Le Maitre, O. M. Knio, H. N. Najm, and R. G. Ghanem. A stochastic projection method for fluid flow, I. basic formulation. *J. Comput. Phys.*, 173:481–511, 2001.

[65] O. P. Le Maitre, M. T. Reagan, H. N. Najm, R. G. Ghanem, and O. M. Knio. A stochastic projection method for fluid flow, II. random process. *J. Comput. Phys.*, 181:9–44, 2002.

[66] F. H. Maltz and D. L. Hitzl. Variance reduction in Monte Carlo computations using multi-dimensional Hermite polynomials. *J. Comput. Phys.*, 32:345–376, 1979.

[67] H. G. Matthies and A. Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194:1295–1331, 2005.

[68] W. C. Meecham and A. Siegel. Wiener-Hermite expansion in model turbulence at large Reynolds number. *Phys. Fluids*, 7:1178, 1964.

[69] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1996.

[70] R. Mikulevicius and B. Rozovskii. Linear parabolic stochastic PDEs and Wiener chaos. *SIAM J. Math. Anal.*, 29:452–480, 1998.

[71] R. Mikulevicius and B. Rozovskii. Stochastic Navier-Stokes equations for turbulence flow. *SIAM J. Math. Anal.*, 35:1250–1310, 2004.

[72] R. Mikulevicius and B. L. Rozovskii. Stochastic Navier-Stokes equations: Propagation of chaos and statistical moments. In J. L. Menaldi, E. Rofmann, and A. Sulem, editors,

Optimal Control and Partial Differential Equations, pages 258–267, Amsterdam, 2001. IOS Press. In honor of Professor Alain Bensoussan's 60th Birthday.

[73] G. N. Milstein. Approximate integration of stochastic differential equations. *Theory Prob. Appl.*, 19:557–562, 1974.

[74] J. Nolen and J. Xin. Reaction-diffusion front speeds in spatially-temporally periodic shear flows. *Multiscale Model. Simul.*, 1:554–570, 2003.

[75] J. Nolen and J. Xin. A variational principle based study of KPP minimal front speeds in random shears. *Nonlinearity*, 18:1655–1675, 2005.

[76] J. Nolen and J. Xin. Variational principle of KPP front speeds in temporally random shear flows with applications. *preprint*, 2005.

[77] B. Øksendal. *Stochastic Differential Equations, An Introduction with Applicatins.* Springer, fifth edition, 2000.

[78] D. Oliver, L. Cunha, and A. Reynolds. Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology*, 29, 1997.

[79] D. Oliver, N. He, and A. Reynolds. Conditioning permeability fields to pressure data. In *5th European conference on the mathematics of oil recovery*, Leoben, Austria, 1996.

[80] H. Omre and O. P. Lodoen. Improved production forecasts and history matching using approximate fluid flow simulators. *SPE Journal*, pages 339–351, 2004.

[81] S. A. Orszag and L. R. Bissonnette. Dynamical properties of truncated Wiener-Hermite expansions. *Phys. Fluids*, 10:2603–2613, 1967.

[82] G. Papanicolaou. Wave propagation in an one-dimensional random medium. *SIAM J. Appl. Math.*, 21:13–18, 1971.

[83] G. Papanicolaou. Diffusion in random media. In J. B. Keller, D. McLaughlin, and G. Papanicolaou, editors, *Surveys in Applied Mathematics*, pages 205–255. Plenum Press, 1995.

[84] E. Platen. An introduction to numerical methods for stochastic differential equations. *Acta Numerica*, pages 197–246, 1999.

[85] G. Da Prato and A. Debussche. Ergodicity for the 3D stochastic Navier-Stokes equations. *J. Math. Pure Appl.*, 82:877–947, 2003.

[86] G. Da Prato, A. Debussche, and R. Temam. Stochastic Burgers' equation. *Nonlinear Differential Equations Appl.*, 1:389–402, 1994.

[87] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, second edition, 1999.

[88] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer-Verlag, New York, 1999.

[89] L. J. Roman and M. Sarkis. Stochastic Galerkin method for elliptic SPDEs: A white noise approach. *Discrete and Continuous Dynamical Systems (accepted)*, 2005.

[90] S. Sakamoto and R. Ghanem. Polynomial chaos decomposition for the simulation of non-Gaussian non-stationary stochastic processes. *J. Eng. Mech.*, 128:190–201, 2002.

[91] C. Schwab and R.-A. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.*, 95(4):707–734, 2003.

[92] Y. Sinai. Two results concerning asymptotic behavior of solutions of the Burgers equation with force. *J. Stat. Phys.*, 64:1–12, 1992.

[93] Y. Sinai. Burgers system driven by a periodic stochastic flow. In *Ito's Stochastic Calculus and Probability Theory*, pages 347–355. Springer-Verlag, 1996.

[94] G. Szegö. *Orthogonal Polynomials*. Amer. Math. Soc., fourth edition, 1975.

[95] N. Wiener. The homogeneous chaos. *Amer. J. Math.*, 60:897–936, 1938.

[96] N. Wiener. *Nonlinear Problems in Random Theory*. MIT Press, Cambridge, 1958.

[97] E. Wong. *Stochastic Processes in Information and Dynamical System*. McGraw-Hill, 1979.

[98] J. Xin. Front propagation in heterogeneous media. *SIAM Rev.*, 42:161–230, 2000.

[99] J. Xin. KPP front speeds in random shears and the parabolic Anderson problem. *Methods Appl. Anal.*, 10:191–198, 2003.

[100] D. Xiu and G. E. Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Engrg.*, 191(43):4927–4948, 2002.

[101] D. Xiu and G. E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phy.*, 187:137–167, 2003.

[102] D. Xiu and G. E. Karnidakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24:619–644, 2002.

[103] D. Xiu, D. Lucor, C. H. Su, and G. E. Karniadakis. Stochastic modeling of flow-structure interactions using generalized polynomial chaos. *J. Fluid Engrg.*, 124:51–59, 2002.

[104] D. Zhang and Z. Lu. An efficient, high-order perburbation approach for flow in random porous media via Karhunen-Loeve and polynomial expansion. *J. Comput. Phy.*, 194:137–167, 2003.