



Swansea University
Prifysgol Abertawe



Swansea University E-Theses

Stochastic finite element modelling of elementary random media.

Li, Chenfeng

How to cite:

Li, Chenfeng (2006) *Stochastic finite element modelling of elementary random media..* thesis, Swansea University.
<http://cronfa.swan.ac.uk/Record/cronfa42770>

Use policy:

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence: copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder. Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

Please link to the metadata record in the Swansea University repository, Cronfa (link given in the citation reference above.)

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>



UNIVERSITY OF WALES SWANSEA
SCHOOL OF ENGINEERING

Stochastic Finite Element Modelling of Elementary Random Media

Chenfeng Li

B.S., M.S. (Tsinghua University, China)

*Thesis submitted to the University of Wales Swansea
in candidature for the degree of Doctor of Philosophy*

September 2006

ProQuest Number: 10807539

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10807539

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346



Declaration and Statements

Declaration

This work has not been previously accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

2006/09/29

Date

Candidate

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. Bibliographies are appended.

2006/09/29

Date

Candidate

Statement 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

2006/09/29

Date

Candidate

Acknowledgements

I would like to express my profound thanks to Prof. D. Roger J. Owen and Dr. Y.T. Feng, my supervisors, for their invaluable guidance, support and encouragement throughout all stages of this research. I would also like to thank Dr. Ian M. Davies for his encouragement and help on stochastic analysis.

I gratefully acknowledge the financial support from EPSRC (grant GR/R87222), which facilitated this research in the early stage (Oct. 2002 – Sep. 2005), and the School of Engineering and the Department of Mathematics in the University of Wales Swansea, which encouraged and supported the final phase of the work (Oct. 2005 – Sep. 2006). My special thanks go to Prof. Owen, Prof. Niels Jacob, Prof. Aubrey Truman and Prof. Nigel P. Weatherill for making this possible.

I would like to thank the Civil & Computational Engineering Centre at the University of Wales Swansea for providing a very supportive and encouraging research environment in which to work.

I am especially grateful to my colleagues Stuart Bates, Wulf G. Dettmer, Jonathan Feldman, Ray Hickey, Mauricio C. Lobao, Miguel R. Paz, Francisco Pires, Andreas Ripple, Prihambodo H. Saksono, Paul Saunders, Satya P. Singh and Z. Zhang, with whom I had many interesting discussions and who made my stay in Swansea very enjoyable.

Last but not least I dedicate this thesis to my parents and my brothers for their constant love and support, without which the completion of this thesis would not have been possible.

Summary

Following a stochastic approach, this thesis presents a numerical framework for elastostatics of random media. Firstly, after a mathematically rigorous investigation of the popular white noise model in an engineering context, the smooth spatial stochastic dependence between material properties is identified as a fundamental feature of practical random media. Based on the recognition of the probabilistic essence of practical random media and driven by engineering simulation requirements, a comprehensive random medium model, namely *elementary random media* (ERM), is consequently defined and its macro-scale properties including stationarity, smoothness and principles for material measurements are systematically explored. Moreover, an explicit representation scheme, namely the Fourier-Karhunen-Loève (F-K-L) representation, is developed for the general elastic tensor of ERM by combining the spectral representation theory of wide-sense stationary stochastic fields and the standard dimensionality reduction technology of principal component analysis. Then, based on the concept of ERM and the F-K-L representation for its random elastic tensor, the stochastic partial differential equations regarding elastostatics of random media are formulated and further discretized, in a similar fashion as for the standard finite element method, to obtain a stochastic system of linear algebraic equations. For the solution of the resulting stochastic linear algebraic system, two different numerical techniques, i.e. the joint diagonalization solution strategy and the directed Monte Carlo simulation strategy, are developed.

Original contributions include the theoretical analysis of practical random medium modelling, establishment of the ERM model and its F-K-L representation, and development of the numerical solvers for the stochastic linear algebraic system. In particular, for computational challenges arising from the proposed framework, two novel numerical algorithms are developed: (a) a quadrature algorithm for multidimensional oscillatory functions, which reduces the computational cost of the F-K-L representation by up to several orders of magnitude; and (b) a Jacobi-like joint diagonalization solution method for relatively small mesh structures, which can effectively solve the associated stochastic linear algebraic system with a large number of random variables.

Table of Contents

Acknowledgements	i
Summary	ii
Notation Symbols and Conventions	vii
1 Introduction	1
1.1 Background and Motivation	1
1.2 Review of Available Techniques	4
1.2.1 Stochastic Partial Differential Equations	4
1.2.2 Stochastic Finite Element Methods	5
1.2.2.1 Overview of Research Related to Stochastic Finite Element Methods	5
1.2.2.2 Major Developments in Stochastic Finite Element Methods	6
1.3 The Aim and Layout of the Thesis	12
1.3.1 The Aim of the thesis	12
1.3.2 Layout of the thesis	13
References	15
2 Basic Mathematical Tools	25
2.1 On Probability Theory	27
2.1.1 Probability Space and Random Variables	27
2.1.2 Statistical Moments and the Characteristic Function	30
2.1.3 Fundamental Inequalities	34
2.1.4 Stochastic Convergence	35
2.1.5 Limit Theorems	38
2.2 On Stochastic Fields	41
2.2.1 General Foundations	41
2.2.2 Gaussian Fields	45
2.2.3 Stationary Stochastic Fields	47
2.3 On Stochastic Analysis	49
2.3.1 White Noise and Brownian Motion	50
2.3.2 Stochastic Ordinary Differential Equations	51
2.3.3 Stochastic Partial Differential Equations	52
References	53
3 Material Modelling – Elementary Random Media	56

3.1	Problem Statements	57
3.2	From White Noise to Elementary Random Media	60
3.2.1	The Integral Functional of White Noise	60
3.2.2	Probabilistic Failings of the White Noise Functional	65
3.2.3	Probabilistic Essence of Random Media	68
3.2.4	Elementary Random Media	69
3.3	Macro-Scale Properties of Elementary Random Media	71
3.3.1	Stationarity	72
3.3.2	Continuity and Differentiability	72
3.3.2.1	Mean-Square Smoothness of Stochastic Fields	73
3.3.2.2	Mean-Square Smoothness of Elementary Random Media	77
3.3.3	Principles of Elementary Random Medium Measurements	77
3.3.3.1	Ergodicity	78
3.3.3.2	Measuring Covariance Functions of Elementary Random Media	80
3.4	Summary and Suggestions for Future Research	81
3.4.1	Summary	81
3.4.2	Suggestions for Future Research	82
3.4.2.1	Making the Ergodicity a Basic Assumption of Elementary Random Media	82
3.4.2.2	A Stochastic Approach for Multi-Scale Simulations	83
	References	87
4	The Fourier-Karhunen-Loève Representation of Elementary Random Media	92
4.1	Overview of Existing Techniques	94
4.2	The Karhunen-Loève Expansion Method	96
4.3	The Fourier-Karhunen-Loève Representation for a Single Elastic Parameter of Elementary Random Media	98
4.3.1	The Fourier-Karhunen-Loève Representation in \mathbb{R}^n	98
4.3.2	The Fourier-Karhunen-Loève Representation in an n -Interval	100
4.3.2.1	The Fourier Expansion of Elementary Random Media	100
4.3.2.2	Principal Component Analysis of Elementary Random Media	103
4.3.3	The Fourier-Karhunen-Loève Representation in an Arbitrary Domain	105
4.4	Computational Issues	105
4.4.1	Computational Techniques	105
4.4.2	The Diagonal Fourier-Karhunen-Loève Approximation Scheme	107
4.4.3	Estimates on the Eigenvalue Decay	109
4.4.4	The Fourier-Karhunen-Loève Algorithm for Scalar Stochastic Fields	111
4.5	Numerical Examples	112
4.5.1	One-Dimensional Examples	112
4.5.1.1	Case I	112
4.5.1.2	Case II	115
4.5.2	A Two-Dimensional Example	118
4.6	The Fourier-Karhunen-Loève Representation for the General Elastic Tensor of Elementary Random Media	120

4.6.1	Background	120
4.6.2	The Solution	120
4.7	Summary and Suggestions for Future Research	124
4.7.1	Summary	124
4.7.2	Suggestions for Future Research	125
Appendix 4A	The Karhunen-Loève Expansion Theorem	127
4A.1	Mercer's Theorem	127
4A.2	The Karhunen-Loève Expansion of Second-Order Stochastic Fields	128
Appendix 4B	The Effective Correlation Length	131
Appendix 4C	The Spectral Representation Theory of Wide-Sense Stationary Stochastic Fields	131
Appendix 4D	Boundary of the Discrete Frequency Domain	133
Appendix 4E	The Covariance Matrix Based on the Real Fourier Basis	134
Appendix 4F	The Exact K-L Expansion in Case I	135
References		136
5	Stochastic Finite Elementary Discretization for Elastostatics of Elementary Random Media	140
5.1	The Stochastic Partial Differential Equation System	140
5.2	The Stochastic System of Linear Algebraic Equations	143
References		146
6	The Joint Diagonalization Solution Strategy for the Stochastic System of Linear Algebraic Equations	147
6.1	Problem Background	148
6.2	Review of Existing Solution Techniques	150
6.2.1	The Monte Carlo Method	150
6.2.2	The Perturbation Method	151
6.2.3	The Neumann Expansion Method	152
6.2.4	The Polynomial Chaos Expansion Method	153
6.3	The Joint Diagonalization Strategy	155
6.3.1	The Formulation	155
6.3.2	A Jacobi-Like Algorithm for the Average Eigenvalue Problem	157
6.3.3	Discussions	161
6.4	Numerical Examples	163
6.4.1	Example 1	163
6.4.2	Example 2	164
6.4.3	Example 3	169
6.5	Summary and Suggestions for Future Research	172
6.5.1	Summary	172
6.5.2	Suggestions for Future Research	173
References		174
7	A Directed Monte Carlo Solution for the Stochastic System of Linear Algebraic Equations	177
7.1	Properties of Stiffness Matrices	179
7.2	Modified Monte Carlo Simulation: One Random Variable	182
7.2.1	Preconditioned Conjugate Gradient Method	183

7.2.2	Exploitation of Spatial Proximity of Monte Carlo Points	185
7.3	Modified Monte Carlo Simulation: Multiple Random Variable	189
7.3.1	Hyper-Spherical Transformation	189
7.3.2	Modified Monte Carlo Simulation in Hyper-Spherical Space	192
7.4	Numerical Experiments	194
7.4.1	Problem Descriptions	195
7.4.2	Performance of the Directed Monte Carlo Method and the Associated Numerical Techniques	197
7.4.2.1	Convergence of Preconditioning and Initial Approximation Schemes	197
7.4.2.2	CPU Time Cost Comparisons	203
7.4.3	Solution Accuracy of the Directed Monte Carlo Method	205
7.5	Summary	207
	References	208
8	Concluding Remarks	210
8.1	Achievements and Conclusions	211
8.1.1	Random Medium Modelling	211
8.1.2	Stochastic Partial Differential Equations	213
8.1.3	The Stochastic System of Linear Algebraic Equations	213
8.1.4	Computer Implementation	214
8.2	Suggestions for Future Research	215
8.2.1	Material Modelling of Random Media	215
8.2.2	Stochastic Partial Differential Equations	216
8.2.3	Numerical Solutions	216
8.2.4	Verification and Application	218
	References	219
	List of Publications	221

Notation Symbols and Conventions

The symbols most frequently used in the text are given below. Any other notation introduced will be defined when required.

- **General conventions**

- I) Scalars are written in regular italic typeface, e.g. a, A, α .
- II) Vector, matrix and tensor quantities are written in bold italic typeface, e.g. $\mathbf{a}, \mathbf{A}, \boldsymbol{\alpha}$.
- III) Equations, figures, examples, algorithms and reference titles are numbered relative to chapters and appendices. Cross-referencing to equations is given in round brackets, e.g. (5.18), while references to bibliography titles are given in rectangular brackets, e.g. [1.8]. Figures, examples and algorithms are referenced using their caption names together with their numbers, e.g. Figure 5.8, Example 2.6 and Algorithm 3.2.
- IV) References are listed at the end of each chapter.

- **Abbreviations**

ERM	Elementary random media
FE	Finite element
FEM	Finite element method
F-K-L	Fourier-Karhunen-Loève
K-L	Karhunen-Loève
ODE	Ordinary differential equations
PDE	Partial differential equations
SFEM	Stochastic finite element

SODE	Stochastic ordinary differential equations
SPDE	Stochastic partial differential equations

• **Some commonly used notations**

\triangleq	Equal to by definition
\square	End of proof
\mathbb{C}	Complex number set
\mathbb{R}	Real number set
\mathbb{Q}	Rational number set
\mathbb{Z}	Integer number set
\mathbb{N}	Natural number set, i.e. $\mathbb{N} \triangleq \{1, 2, 3, \dots\}$
$k!$	Factorial of integer $k \geq 0$, i.e. $k! \triangleq \begin{cases} \prod_{i=1}^k i & k \geq 1 \\ 1 & k = 0 \end{cases}$
$C^m(D)$	The set of all functions defined on D that have up to m -th order continuous derivatives
δ_{ij}	The Kronecker delta, i.e. $\delta_{ij} \triangleq \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$
$\delta(x)$	The Dirac delta function, i.e. $\delta(x) = \begin{cases} 0 & x \neq 0 \\ +\infty & x = 0 \end{cases}$ and $\int_{-\infty}^{+\infty} \delta(x) dx = 1$
inf	Infimum, i.e. the greatest lower bound
sup	Supremum, i.e. the least upper bound
$\mathbf{a} \cdot \mathbf{b}$	Scalar product of vectors \mathbf{a} and \mathbf{b}
$\mathbf{a} : \mathbf{b}$	Double contraction of tensors \mathbf{a} and \mathbf{b}
$\ \mathbf{a}\ _2$	Euclidean norm of vector \mathbf{a} , i.e. $\ \mathbf{a}\ _2 \triangleq \sqrt{\mathbf{a} \cdot \mathbf{a}}$
$\overline{(\bullet)}$	Complex conjugate of (\bullet)
$ a $	Absolute value (or modulus) of a
∇f	Gradient of function f

$\nabla \cdot f$	Divergence of function f
$f * g$	Convolution of functions f and g
a.e.	Almost everywhere
\perp	Orthogonal to (in a Hilbert space)
L^p	L^p space
$\ f\ _p$	The standard L^p norm of function $f(\mathbf{x})$, i.e. $\ f(\mathbf{x})\ _p \triangleq \left(\int_{\mathbb{R}^n} f(\mathbf{x}) ^p d\mathbf{x} \right)^{1/p} \quad p \geq 1$
(Ω, \mathbb{F}, P)	A probability space
Ω	The sure event
ω	The basic event
\mathbb{F}	A σ -algebra of subsets of Ω
\mathbb{B}	The Borel σ -algebra of \mathbb{R}
$P\{\bullet\}$	Probability of $\{\bullet\}$
$E(\bullet)$	Expectation of (\bullet)
$\text{Var}(\bullet)$	Variance of (\bullet)
$\text{Cov}(\bullet, \bullet)$	Covariance of (\bullet, \bullet)
$\Phi(x)$	The standard Gaussian distribution function
$\Gamma(z)$	The gamma function
a.s.	Almost surely
$X_i \xrightarrow{\text{a.s.}} X$	X_i converges to X almost surely (or with probability 1)
$X_i \xrightarrow{\gamma} X$	X_i converges to X in γ -th mean
$X_i \xrightarrow{P} X$	X_i converges to X in probability
$X_i \xrightarrow{d} X$	X_i converges to X weakly (or in distribution)
$X_i \xrightarrow{\text{m.s.}} X$	X_i converges to X in mean square
$X_i \xrightarrow{\text{q.m.}} X$	X_i converges to X in quadratic mean
$\text{l.i.m.}(\bullet)$ $\bullet \rightarrow \bullet$	Limit in mean
A^T	Transpose of matrix/vector A

A^H	Conjugate transpose of matrix/vector A
$\text{tr}(A)$	Trace of tensor/matrix A
$\det(A)$	Determinant of matrix A
$\text{diag}(d)$	The diagonal matrix produced by vector d
$\text{rank}(A)$	Rank of matrix A
$(A)_{ij}$	The entry at the i -th row and j -th column of matrix A
$\text{cond}_2(A)$	The spectral condition number of matrix A

Chapter 1

Introduction

1.1 Background and Motivation

Random media including rocks, soils and concrete etc. are the most common materials on this planet, and they play a very important role in the design of engineering infrastructures. Phenomena involving random media are, however, not only encountered in civil engineering, but also in almost every engineering sector. Although for some materials such as perfect metals, uncertainty factors are often invisible at the macro scale, they become dominant when various unpredictable damage mechanisms have occurred during manufacture, operation and maintenance procedure, or where the micro-scale material structure has to be taken into account. As shown in Figure 1.1, typical engineering problems with respect to random media include risk assessment of rock structures, percolation-flow-field analysis of soils or rocks, life prediction of concrete dams, mechanical analysis of composites with random inclusions, solid-state epitaxy and recrystallization in semiconductors, biomechanical analysis of bones, etc., all of which require complex mathematical predictions. Modern mechanics, especially continuum mechanics, provides comprehensive mathematical models for various engineering systems in terms of partial differential equations. However, due to the intricate stochastic nature of random media, the lack of a proper mathematical model for the associated stochastic fields and the inadequacy of supporting analysis tools for the corresponding Stochastic Partial Differential Equations (SPDEs), the understanding of engineering systems with variable uncertainties is far from being mature.

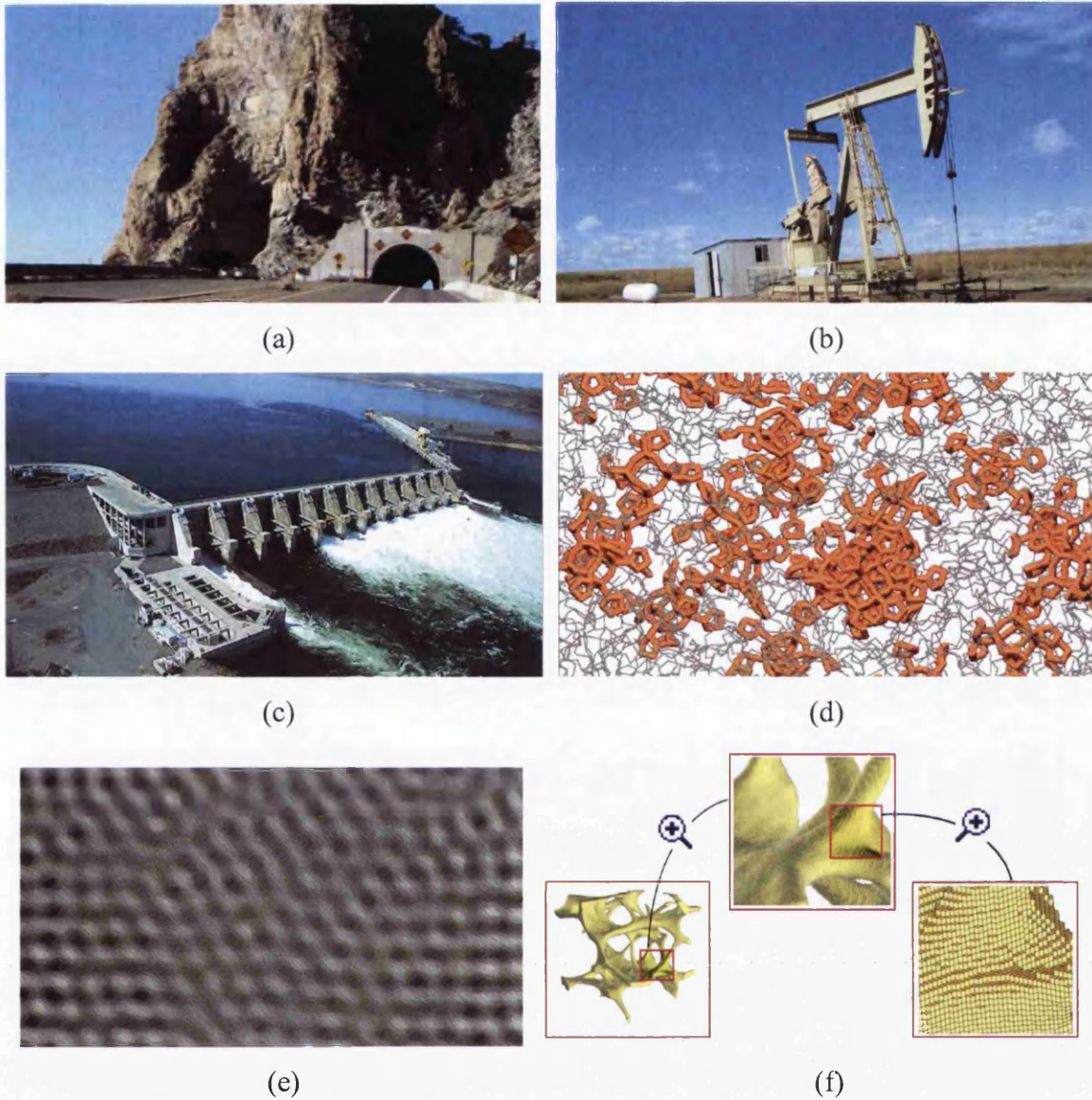


Figure 1.1 Engineering systems with variable uncertainties

To date, the analysis of engineering systems composed of random media still relies largely on empiricism and parametrical (or equivalently enumerative) simulations, which are essentially based on conventional deterministic mathematical models. Consequently, as long as random media exist in a practical engineering problem, conservative safety factors are inevitably applied to the approximate mean-value solution, which results in a significant increase in design, construction and operational costs. The disadvantages of using deterministic mathematical models to study random media are evident: (a) It is often extremely difficult, if not impossible, to provide data for an exhaustive deterministic description of the random media under consideration; (b) Engineering analysis based on a deterministic model can only provide a path-wise solution corresponding to a particular

realization of the engineering system with variable uncertainties; (c) In a practical engineering problem involving random media, a large number of experiments or parametric simulations are required to estimate empirical statistical quantities and as a result, the deterministic methodology may even cease to be useful because of insufficient resources; (d) In some situations such as problems influenced by size-effects, the deterministic mathematical model eliminates the essential stochastic nature of the physical system, and may even lead to wrong conclusions. These unsatisfactory situations that arise in engineering are essentially due to the lack of a versatile probabilistic model for random media and the inadequacy of compatible analysis tools for the corresponding SPDEs. Therefore, it is of crucial importance to explore the mathematical foundation and develop efficient algorithms for practical engineering systems consisting of random media.

In engineering: It is obvious beneficial to know more about how practical engineering systems consisting of random media behave and how their behaviors are influenced by random material properties and by various uncertainty factors. The theoretical model and the computational algorithms developed in this research area will significantly improve the understanding of random media related problems, which in turn will reduce the risk as well as contribute towards optimizing structural performance.

In mechanics: On the one hand, material structures at the micro scale contain various uncertainties; on the other hand, macro-scale mechanical behavior of materials is governed by continuum mechanics, which is a purely deterministic theory. Therefore, continuum mechanics theory can be treated as an average result from a micro-scale probabilistic theory. Consequently, it is expected that the results obtained in this research area could serve as a bridge in material modelling to connect randomness at the micro scale and determinacy at the macro scale. This will obviously be of significant importance in theoretical mechanics.

In mathematics: Since Einstein gave the correct mathematical description of Brownian motion in 1905, extensive results have been obtained in related areas. These include contributions from Langevin, Wiener, Kolmogorov and Itô etc. The study of Brownian motion significantly promoted the developments of probability theory, stochastic fields and stochastic analysis, and the related research eventually became an important and sophisticated area both in mathematics and in physics, and has various

successful applications, for example, in economics and biology. Brownian motion is the movement of a single particle subject to random forces, and Einstein solved this problem by combining elementary probability theory and classical Newtonian particle dynamics. For a continuum, a new theory based on modern stochastic analysis and continuum mechanics can be similarly expected. Hence, the physical problems in this research area could also be of interest in pure mathematical research.

1.2 Review of Available Techniques

Over the last few decades, the theoretical/numerical ingredients of random medium mechanics have been studied independently by both mathematicians and engineers from different points of view, especially in the areas of SPDEs and Stochastic Finite Element Methods (SFEMs) respectively. A number of references will be mentioned in this section. However, for research subjects related to random medium mechanics, the reference list is not complete. It contains the most important items (measured by the times cited) and the items most relevant to the subject discussed in this work. In addition, this technical review does not always address the first but mostly the summarizing concluding references.

1.2.1 Stochastic Partial Differential Equations

The concept of the Itô integral, i.e. the theory of Stochastic Ordinary Differential Equations (SODEs) [1.1], became mature in the 1980s and its successful applications have been observed in a wide range of scientific/technical areas including economics, biology, etc. Promoted by the success of SODEs, the study of SPDEs has developed steadily over the past twenty years, in which major efforts have been focused on expanding Itô's white noise model from one dimension into higher dimensions [1.2]. However, unlike the topology of one dimensional space there is no natural total-order structure equipped with higher dimensional space. Because of this fundamental difference in topology, the white noise approach for SPDEs is still very much ongoing. In addition, some mathematical results obtained in SPDEs are useful for their own interest, but have little to do with specific physical problems. Consequently, engineers feel a need for a justification why these SPDE developments are physically relevant at all. In summary, the present

mathematics has not been able to provide an effective SPDE theory to account for practical engineering systems compounded by the presence of random media. In later parts of this thesis (especially Chapters 2 and 3), the connection between SPDE research and engineering systems with variable uncertainties will be addressed in a more detailed manner after some mathematical preparations.

1.2.2 Stochastic Finite Element Methods

Over the last few decades, Finite Element Methods (FEMs) have become the dominant analysis tool in engineering. In the standard FEM, all the parameters take fixed values. In order to study engineering systems consisting of random media, engineers have replaced some of the parameters in the FEM by random variables and attempted to expand the standard FEM into SFEM [1.15-1.123].

1.2.2.1 Overview of Research Related to Stochastic Finite Element Methods

According to the information from ISI Web of Science [1.3], there are in total 1,805 SCI papers published from 1981 to 2005 which are associated with finite elements and uncertainties. Plotted in Figure 1.2(a) is the histogram of these SCI papers together with their corresponding database keywords. Although many of these papers are not about modelling or algorithms of SFEMs, the histogram 1.2(a) discloses clearly an increasing research interest in the linked area of finite element methods and random phenomena. Among these 1,805 SCI papers, there are 112 papers published in *Computer Methods in Applied Mechanics and Engineering* (CMAME) and *International Journal for Numerical Methods in Engineering* (IJNME), of which the distribution and database keywords are plotted in Figure 1.2(b). Due to the continual exponential increases in computer power and data storage, the research of SFEMs has recently received considerable attention from the computational mechanics community [1.19-1.20]. In order to identify the impact of these 1,805 papers, the histogram of those papers with relatively high citation (cumulatively over twice per year) is plotted in Figure 1.2(c), in which the taller columns denote all the papers regarding finite elements and uncertainties, and the shorter columns denote the papers specifically related to SFEM formulations. Compared with the increasing impact of the former group of papers, the impact of the latter group of papers is so far insignificant, which indicates that the research on SFEM formulations is far from reaching maturity.

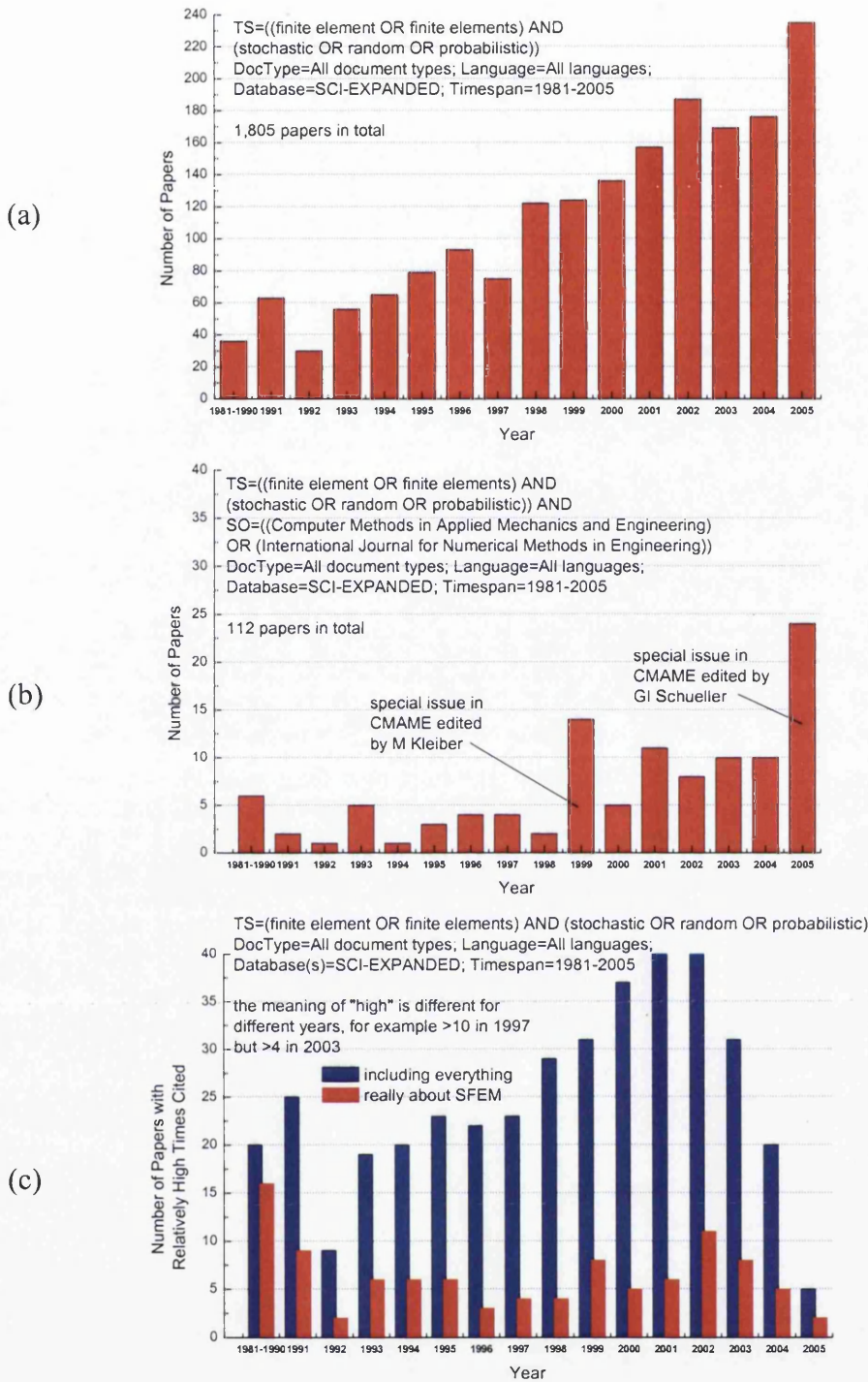


Figure 1.2 Histograms of SCI papers regarding finite elements and uncertainties

1.2.2.2 Major Developments in Stochastic Finite Element Methods

In order to obtain a detailed understanding of major developments in SFEM related research, it was necessary to carefully examine all the important publications in this area, which include the abstracts of the aforementioned 1,805 SCI papers (as shown in

Figure 1.2(a)), the full text of all the SFEM papers published in CMAME and IJNME from 2002 to 2005 (as shown in Figure 1.2(b)) and the full text of all the SFEM papers with relatively high citation (as shown in Figure 1.2(c)). The term SFEM has appeared in the literature since over fifteen years ago. Nevertheless, unlike the standard FEM that holds a clear and uniform Galerkin formulation, there has not been a widely accepted versatile SFEM formulation. To date, the SFEM may be best regarded as some FEM-based numerical technique that deals with certain additive or multiplicative random factors in certain SPDE systems.

During the 1970s and the 1980s, Shinozuka [1.4] started systematically applying the Monte Carlo method and the standard FEM in the reliability analysis of structures with random excitation, random material properties or random geometric configurations [1.21-1.27]. The related research area was soon named stochastic computational mechanics (sometimes also called computational stochastic mechanics) and has seen significant growth since then. There are two central issues in Monte Carlo based SFEM formulations: (a) how to initialize material samples according to the specific random input information (often given in the form of statistical moments); (b) how to reduce the sampling number of Monte Carlo simulations. For the first key issue, the most widely used technique is the trigonometric series approximation method which is based on the spectral representation theory of wide-sense stationary stochastic fields [1.18, 1.28-1.30] (see Chapter 4 for more information). For the second key issue, the most important techniques are the so called FORM (First Order Reliability Method) and SORM (Second Order Reliability Method) [1.31-1.37]. To date, the FORM and SORM are still the most popular approximation techniques for calculating the probability integral encountered in reliability analysis of random structures [1.39], and one of the latest research interests regarding FORM and SORM is applying them to larger engineering structures [1.38]. However, it is also noticed that some engineers [1.40] do not agree with the application of FORM or SORM in practical engineering problems as the approximation may generate seriously biased results.

Other pioneering researchers in stochastic computational mechanics include Vanmarcke [1.5], Nakagiri [1.15], Kiureghian [1.6] and Grigoriu [1.7]. Regarding the early research of SFEMs, it is worth mentioning the summarizing review [1.25] co-authored by Vanmarcke, Shinozuka and Nakagiri et al. In 1983, the international

journal *Structural Safety* (SS) [1.13] was founded by Vanmarcke, and three years later, the international journal *Probabilistic Engineering Mechanics* (PEM) [1.14] was founded by Shinozuka. In 2005, there were 18 papers published on SS and its impact factor was 0.953; for PEM, there were 32 papers published in 2005 and its impact factor was 0.767. In the SCI database 2005, SS and PEM are the only specialized journals for stochastic computational mechanics. The relatively weak impact of these two journals infers that the contemporary SFEMs have not been able to serve the increasing needs of practical engineering systems with variable uncertainties.

In the 1980s and the 1990s, more researchers joined in the research of SFEMs. Among them, Liu [1.8], Spanos [1.9], Ghanem [1.10, 1.16], Kleiber [1.17], Deodatis [1.11] and Schuëller [1.12] are distinguished. Many interesting results, especially a number of non-Monte-Carlo SFEM formulations, were developed during this period. As shown in Table 1.1, most of the major SFEM techniques were developed from 1985 to 1995.

Table 1.1 The most important SFEM developments (1981-2005)

Year ^a	References	Times cited ^b	Outline ^c
1983	E. Vanmarcke M. Grigoriu [1.41]	122	<p>This work is probably the first reasonable SFEM formulation which is not based on Monte Carlo simulations.</p> <p>The key invention in this paper is a local averaging representation scheme for random material properties of beam elements.</p> <p>The method can only be applied to a simple beam and the results are limited to estimates of second-order statistical quantities, i.e. the mean and the covariance.</p>
1986	W.K. Liu et al. [1.42-1.46]	308	<p>The perturbation method.</p> <p>The method commences with a Taylor's expansion of the unknown stochastic field. The Taylor expansion is performed with respect to the random variables that represent the given stochastic field of random media. The unknown coefficients in the expansion are then obtained from grouping like polynomials, whereupon the sum of these like polynomials is set to zero. The perturbation method is computationally more efficient than the direct Monte Carlo method.</p> <p>The failing of this method is a dependence on the random fluctuations being small. Another disadvantage of the perturbation method is that it mainly focuses on the second-order estimate of the response and does not permit higher-order statistical estimates.</p>

Year ^a	References	Times cited ^b	Outline ^c
1987	M.A. Lawrence [1.47]	57	<p>A SFEM formulation based on series expansion and the Galerkin method.</p> <p>This paper considers the equation of the form $Ku=f$. After expanding K, u and f into finite series consisting of random coefficients and deterministic base functions, the equation is solved by a Galerkin approach. The idea of combining series expansion and the Galerkin method in a random context was original and very interesting.</p> <p>The series expansion formulation in this paper is developed in an intuitive manner without paying attention to mathematical rigor, so that some parts of the formulation are suspect.</p>
1988	M. Shinozuka F. Yamazaki G. Deodatis et al. [1.26-1.27]	141	<p>The Neumann expansion method.</p> <p>The stochastic algebraic equation $(K_0+K_\omega)u=f$ is solved by the Neumann expansion of $(K_0+K_\omega)^{-1}$. The advantage of this method is its simple formulation.</p> <p>The computational efficiency of the Neumann expansion depends on the range of random fluctuations. For large-scale random fluctuations, the computational cost of this method could be even more expensive than the direct Monte Carlo method.</p>
1989	R.G. Ghanem P.D. Spanos et al. [1.48-1.60]	336	<p>The polynomial chaos expansion method.</p> <p>This method expands the unknown stochastic field with multiple Hermite polynomials of random variables, i.e. polynomial chaos, and solves the associated unknown coefficients through a Galerkin approach. There is no limit to the scale of random fluctuations and the polynomial chaos expansion can, at least in principle, approximate a functional of Gaussian random variables to any accuracy. Another invention in this work is the application of the Karhunen-Loève expansion in the representation of random material properties.</p> <p>The polynomial chaos expansion method can only be rigorously applied to problems that merely consist of Gaussian random variables. In addition, it does become increasingly difficult to derive and code multiple Hermite polynomials (the basic building blocks of this method) as the number of random variables increases, and the associated computational cost also increases significantly. Indeed, the polynomial chaos expansion has hardly been applied to cases with more than ten random variables.</p>
1991	G. Deodatis M. Shinozuka [1.61-1.62]	76	<p>The weighted integral method.</p> <p>This method is developed for frame structures, and it is based on second-order statistical analysis and a local averaging technique for representing random material properties of beam elements.</p> <p>The results of this method are limited to second-order statistical quantities.</p>

Year ^a	References	Times cited ^b	Outline ^c
1991	P.L. Liu A.D. Kiureghian [1.37]	47	<p>A Monte Carlo SFEM formulation based on FORM and SORM.</p> <p>In order to reduce the number of direct Monte Carlo simulations, the FORM and SORM are employed to approximately calculate the probability integral with respect to the random solution.</p> <p>The accuracy of FORM and SORM may not always satisfy the increasingly more serious requirements to the analysis of engineering systems with variable uncertainties.</p>
1993	C.C. Li A.D. Kiureghian [1.63]	47	<p>This work is not about SFEM formulations, and presents a least-squares approximation method for the representation of random material properties.</p> <p>The criterion of the least-squares approximation within each finite element is the variance of the error between the real stochastic field and the approximate one.</p> <p>In terms of the number of random variables required, the least-squares approximation method is not as efficient as the Karhunen-Loève expansion method, which was first adopted in the polynomial chaos expansion method. Nevertheless, the authors claimed that this method might hold an advantage when the stochastic-field eigenstructure in Karhunen-Loève expansions could not be accurately obtained.</p>
1993	S. Valliappan T.D. Pham S.S. Rao J.P. Sawyer [1.64-1.66]	68	<p>The fuzzy/interval finite element method.</p> <p>The method applies the fuzzy set theory in the context of standard finite element analysis to cope with engineering systems without precise or complete definitions.</p> <p>As the input information of this method is insufficient to properly define an engineering system, the associated results are often vague, imprecise, qualitative and incomplete compared with the solution of a well defined engineering system.</p>
2002	D.B. Xiu G.E. Karniadakis [1.106-1.114]	70	<p>A generalization for the polynomial chaos expansion method.</p> <p>Noting the similarity between the weighted functions in the Askey-scheme of hypergeometric orthogonal polynomials and some typical probability density functions, the polynomial chaos expansion method is generalized simply by replacing the Hermite polynomials with other orthogonal polynomials.</p> <p>This intuitive generalization meets difficulties when multiple random variables are present. In addition, for the single random variable cases, there has not been a proof for the well-posedness or the convergence of the solution scheme.</p>

^a For those works reported in a series of publications, the year is given for the first influential paper;

^b Data from ISI Web of Science, December 2005;

^c More detailed illustration and comparison of some of these SFEMs can be found in Chapters 4 and 6.

In the last decade, a number of interesting results [1.20, 1.67-1.123] were reported

that include the application/generalization of those prototype SFEM techniques (as shown in Table 1.1) in various engineering situations, such as different material parameters (Young's modules, Poisson's ratio and yielding stress etc.), different geometric structures (beams, plates and shells etc.), different physical systems (elastostatics, elastodynamics and thermodynamics etc.) and many more. Among these latest SFEM developments, it is worth mentioning the following work:

- Soize [1.118-1.119] developed a nonparametric SFEM formulation based on the mathematical theory of random matrices.
- In a Galerkin SFEM formulation, Matthies et al. [1.95] applied the Karhunen-Loève expansion in the approximation of the unknown random solution.
- Vandepitte et al. [1.115-1.116] carry on with the developments of the fuzzy/interval finite element method.
- In a Galerkin SFEM formulation, Babuška et al. [1.101] approximated the unknown random solution with so called double orthogonal polynomials.
- Based on the combination of Taylor's series expansions of certain like functions, Xu et al. [1.121] developed a dimension-reduction technique to approximately calculate the probability integral involved in SFEM simulations.
- Li et al. [1.120] proposed an original approach for dynamic response and reliability analysis of stochastic structures. A class of general probability density evolution equation is derived, constructing the relationship between the incremental rate and the realized velocity response.

Nevertheless, compared with the rapid developments between 1986 and 1995, there has hardly been any breakthrough in SFEM formulations since then. Until recently [1.20, 1.67-1.123], three SFEM techniques including the Monte Carlo method, the perturbation method (see e.g. [1.19, 1.42]) and the polynomial chaos expansion method (see e.g. [1.20, 1.48-1.49]) are distinguished.

- To date, the Monte Carlo method appears to be the only versatile method to solve stochastic problems involving non-linearities, dynamic loading, stability effects etc. [1.67-1.82]. However, its computational costs become extremely expensive for large scale problems where a large number of samples have to be computed in order to obtain a rational estimation satisfying the required accuracy.
- For small-scale random variations, the perturbation method is still attractive [1.83-1.92] due to its relatively simple formulation. However, the perturbation

method is limited to the estimation of second-order statistical quantities, and no criteria for convergence have been established in the present context.

- The polynomial chaos expansion method together with its error estimates has recently been put into more general/sophisticated numerical frameworks [1.93-1.105]. Although the initial form of this method uses Hermite polynomials and can only be strictly applied to solve problems consisting of Gaussian random variables, for dynamic problems with a single random variable, different polynomials (e.g. single-variable Jacobi polynomials) have recently been chosen [1.106-1.114], without proof on the suitability, to approximate functions of a non-Gaussian random variable. However, the associated computational costs increase exponentially as the number of random variables grows, which makes the polynomial chaos expansion method impractical for real scale engineering problems in its current form.

In summary, compared with the success of the standard FEM, the SFEM is still in its infancy and many fundamental questions are still outstanding.

1.3 The Aim and Layout of the Thesis

1.3.1 The Aim of the thesis

Following the introduction in Section 1.1, phenomena involved with random media are encountered in many engineering sectors, and the research regarding random medium mechanics is of crucial importance not only in engineering, but also in mechanics and mathematics. However, after the technical review in Section 1.2, it is clear that existing knowledge and methods are not adequate to analyse practical engineering systems composed of random media. On one hand, mathematicians developed SPDE theory from mathematical definitions and assumptions, but before the theory can be applied to engineering systems, the reasonableness of those definitions and assumptions must be verified by the general mechanics theory and practical engineering requirements. On the other hand, engineers modified the traditional FEM to account for engineering systems with variable uncertainties, but as the knowledge of stochastic fields and stochastic analysis is not embedded within the standard solution tools of engineering, some SFEM

techniques have little mathematical foundation and some appear to come with self-contradictions (see Chapter 3 for more details).

The aim of this thesis is to explore the theoretical foundation of random medium mechanics and the possibility of establishing a versatile numerical framework for the analysis of engineering systems composed of random media. The main strategy relies on the combination of mathematical ingredients in probability theory, stochastic fields and stochastic analysis with continuum mechanics and finite element methods. The physical models of random media, the mathematical format of the associated stochastic fields, the governing SPDEs of physical systems and the solution schemes shall all be developed in a synergetic manner. In order to reduce the number of open questions in this challenging subject and make the problem more amenable, attention is restricted to elastostatics of random media. The resulting principles and framework are, however, expected to hold not only for linear static/stationary problems, but also be expandable to certain nonlinear problems and dynamic/transient problems.

1.3.2 Layout of the thesis

The thesis is divided into seven chapters.

Chapter 1: This chapter discusses the background and the motivation which have inspired the research undertaken. A brief but complete technical review on related research is also provided.

Chapter 2: A brief exposition of probability theory, stochastic fields and stochastic analysis is included in this chapter. It provides some conceptual preparations for Chapters 3-7, and in particular initiates the specific approach developed in Chapter 3.

Chapter 3: The possibility of establishing a white noise model (or a generalized white noise model) for practical random media is analyzed in a mathematically rigorous fashion. Based on the fundamental principles of continuum mechanics and the theory of stationary stochastic fields, a comprehensive probabilistic model, namely *elementary random media*, is developed for modelling practical elastic random media, which includes the probabilistic properties of random media, the continuity and differentiability of random media, and principles for the measurements of random material properties.

Chapter 4: In the definition of elementary random media, the random medium is implicitly defined by its statistical moments. However, in the governing SPDE system of an engineering system composed of random media, an explicit representation of the associated stochastic fields of material properties is usually required. Hence, based on the spectral representation theory of wide-sense stationary stochastic fields and the standard dimensionality reduction technology of principal component analysis, an accurate and efficient stochastic-field representation scheme, namely the Fourier-Karnunen-Loève representation, is developed for the general elastic tensor of elementary random media. Several original developments are made in this chapter.

Chapter 5: Based on the Fourier-Karnunen-Loève representation of elementary random media, the governing SPDE system of elastostatics is transformed into a stochastic system of linear algebraic equations, which in turn infers a natural generalization of the standard FEM formulation for deterministic materials to a SFEM formulation for elementary random media.

Chapter 6: This chapter presents a novel solution strategy to the stochastic system of linear algebraic equations $\left(\sum_{i=1}^m \alpha_i A_i\right)x = b$ derived in Chapter 5, in which α_i denote random variables, A_i real symmetric deterministic matrices, b a deterministic/random vector and x the unknown random vector to be solved. Specifically, it is shown that the problem can be approximately treated as an average eigenvalue problem, and can be solved analytically by a deterministic matrix algorithm using a sequence of orthogonal similarity transformations. Once the approximate solution of the random vector x is explicitly obtained, the associated statistical moments and joint probability distributions can be readily calculated.

Chapter 7: Based on Monte Carlo simulations, another solver is developed for the solution of the stochastic system of linear algebraic equations. The basic idea is to direct Monte Carlo samples along straight lines and then utilise their spatial proximity or order to provide high quality initial approximations in order to significantly accelerate the convergence of iterative solvers at each sample.

Chapter 8: Pertinent conclusions are drawn and a broad perspective of the

problem at hand is emphasised.

Certain aspects of this research work have been presented in [1.124-1.129].

References

- [1.1] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publishing Company, Amsterdam, 1981.
- [1.2] H. Holden, B. Øksendal, J. Ubøe and T. Zhang, *Stochastic Partial Differential Equations: A Modelling, White Noise Functional Approach*, Birkhäuser Boston, Boston, 1996.
- [1.3] *ISI Web of Science*, <http://wok.mimas.ac.uk/>.
- [1.4] Prof. M. Shinozuka, Fellow of US National Academy of Engineering, Columbia University, Princeton University and University of California Irvine, <http://shino8.eng.uci.edu/Profile.htm>.
- [1.5] Prof. E. Vanmarcke, Princeton University and Massachusetts Institute of Technology, <http://www.princeton.edu/evm/index.html>.
- [1.6] Prof. A.D. Kiureghian, UC Berkeley, <http://www.ce.berkeley.edu/~adk/>.
- [1.7] Prof. M.D. Grigoriu, Cornell University, <http://www.cee.cornell.edu/fbxk/fcpblctns.cfm?pid=197>.
- [1.8] Prof. W.K. Liu, Northwestern University, http://www.tam.northwestern.edu/wkl/_wkl/hm/contributions.htm.
- [1.9] Prof. P.D. Spanos, Fellow of US National Academy of Engineering, Rice University, <http://www.mems.rice.edu/spanos/>.
- [1.10] Prof. R.G. Ghanem, University of Southern California, <http://venus.usc.edu/users/index.html>.
- [1.11] Prof. G. Deodatis, Columbia University, <http://www.civil.columbia.edu/faculty/deodatis.html>.
- [1.12] Prof. I. Schuëller, Leopold-Franzens University, Austria, <http://www.uibk.ac.at/mechanik/schueller/>.
- [1.13] *Structural Safety*, http://www.elsevier.com/wps/find/journaldescription.cws_home/505664/description#description.
- [1.14] *Probabilistic Engineering Mechanics*,

http://www.elsevier.com/wps/find/journaldescription.cws_home/422923/description#description.

- [1.15] S. Nakagiri and T. Hisada, *An Introduction to Stochastic Finite Element Method: Analysis of Uncertain Structures*, Baifukan, Tokyo, 1985. (In Japanese)
- [1.16] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements – A Spectral Approach*, Revised edition, Dover Publications, New York, 2003.
- [1.17] M. Kleiber and T.D. Hien, *The Stochastic Finite Element Method, Basic Perturbation Technique and Computer Implementation*, John Wiley & Sons, Chichester, 1992.
- [1.18] M. Grigoriu, *Stochastic Calculus: Applications in Science and Engineering*, Birkhauser, 2002.
- [1.19] Edited by M. Kleiber, Special issue on computational stochastic mechanics, *Computer Methods in Applied Mechanics and Engineering*, 168 (1-4) (1999) 1-354.
- [1.20] Edited by G.I. Schuëller, Special issue on computational methods in stochastic mechanics and reliability analysis, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1251-1795.
- [1.21] M. Shinozuka, Simulation of multivariate and multidimensional random processes, *Journal of the Acoustical Society of America*, 49 (1) (1971) 357-367.
- [1.22] M. Shinozuka and Y.K. Wen, Monte-Carlo solution of nonlinear vibrations, *AIAA Journal*, 10 (1) (1972) 37-40.
- [1.23] M. Shinozuka and C.J. Astill, Random eigenvalue problems in structural analysis, *AIAA Journal*, 10 (4) (1972) 456-462.
- [1.24] M. Shinozuka, Monte Carlo solution of structural dynamics, *Computers & Structures*, 2 (5-6) (1972) 855-874.
- [1.25] E. Vanmarcke, M. Shinozuka, S. Nakagiri, G.I. Shuëller and M. Grigoriu, Random-fields and stochastic finite-elements, *Structural Safety*, 3 (3-4) (1986) 143-166.
- [1.26] M. Shinozuka and G. Deodatis, Response variability of stochastic finite element systems, *Journal of Engineering Mechanics – ASCE*, 114 (3) (1988) 499-519.
- [1.27] F. Yamazaki, M. Shinozuka and G. Dasgupta, Neumann expansion for stochastic finite element analysis, *Journal of Engineering Mechanics – ASCE*, 114 (8) (1988) 1335-1354.
- [1.28] M. Shinozuka and G. Deodatis, Simulation of stochastic processes by spectral

- representation, *Applied Mechanics Reviews – ASME*, 44 (4) (1991) 191-204.
- [1.29] M. Grigoriu, A spectral representation based model for Monte Carlo simulation, *Probabilistic Engineering Mechanics*, 15 (4) (2000) 365-370.
- [1.30] J.E. Hurtado and A.H. Barbat, Monte Carlo techniques in computational stochastic mechanics, *Archives of Computational Methods in Engineering*, 5 (1) (1998) 3-29.
- [1.31] A.M. Hasofer and N.C. Lind, An exact and invariant first order reliability format, *Journal of Engineering Mechanics – ASCE*, 100 (1) (1974) 111-121.
- [1.32] B. Fiessler, H.J. Neumann and R. Rackwitz, Quadratic limit states in structural reliability, *Journal of Engineering Mechanics – ASCE*, 105 (4) (1979) 661-676.
- [1.33] M. Hohenbichler and R. Rackwitz, First-order concepts in system reliability, *Structural Safety*, 1 (3) (1983) 177-188.
- [1.34] M. Hohenbichler, S. Gollwitzer, W. Kruse and R. Rackwitz, New light on first- and second- order reliability methods, *Structural Safety*, 4 (4) (1987) 267-284.
- [1.35] L. Tvedt, Distribution of quadratic forms in normal space – application to structural reliability, *Journal of Engineering Mechanics – ASCE*, 116 (6) (1990) 1183-1197.
- [1.36] A.D. Kiureghian and M.D. Stefano, Efficient algorithm for second-order reliability analysis, *Journal of Engineering Mechanics – ASCE*, 117 (12) (1991) 2904-2923.
- [1.37] P.L. Liu and A.D. Kiureghian, Finite element reliability of geometrically nonlinear uncertain structures, *Journal of Engineering Mechanics – ASCE*, 117 (8) (1991) 1806-1825.
- [1.38] K. Imai and D.M. Frangopol, Geometrically nonlinear finite element reliability analysis of structural systems. I: theory, *Computers & Structures*, 77 (6) (2000) 677-691.
- [1.39] R. Rackwitz, Reliability analysis – a review and some perspectives, *Structural Safety*, 23 (4) (2001) 365-395.
- [1.40] P.E.C. Annis, http://www.statisticalengineering.com/limit_state.htm.
- [1.41] E. Vanmarcke and M. Grigoriu, Stochastic finite element analysis of simple beams, *Journal of Engineering Mechanics – ASCE*, 109 (5) (1983) 1203-1214.
- [1.42] W.K. Liu, T. Belytschko and A. Mani, Random field finite elements, *International Journal for Numerical Methods in Engineering*, 23 (10) (1986) 1831-1845.
- [1.43] W.K. Liu, T. Belytschko and A. Mani, Probabilistic finite elements for

- nonlinear structural dynamics, *Computer Methods in Applied Mechanics and Engineering*, 56 (1) (1986) 61-81.
- [1.44] G.H. Besterfield, W.K. Liu, M.A. Lawrence and T. Belytschko, Fatigue crack growth reliability by probabilistic finite elements, *Computer Methods in Applied Mechanics and Engineering*, 86 (3) (1991) 297-320.
- [1.45] W.K. Liu, T. Belytschko and A. Mani, Applications of probabilistic finite-element methods in elastic plastic dynamics, *Journal of Engineering for Industry – ASME*, 109 (1) (1987) 2-8.
- [1.46] G.H. Besterfield, W.K. Liu, M.A. Lawrence and T. Belytschko, Brittle-fracture reliability by probabilistic finite-elements, *Journal of Engineering Mechanics – ASCE*, 116 (3) (1990) 642-659.
- [1.47] M.A. Lawrence, Basis random variables in finite element analysis, *International Journal for Numerical Methods in Engineering*, 24 (10) (1987) 1849-1863.
- [1.48] P.D. Spanos and R.G. Ghanem, Stochastic finite element expansion for random media, *Journal of Engineering Mechanics – ASCE*, 115 (5) (1989) 1035-1053.
- [1.49] R.G. Ghanem and P.D. Spanos, Polynomial chaos in stochastic finite elements, *Journal of Applied Mechanics – ASME*, 57 (1) (1990) 197-202.
- [1.50] R.G. Ghanem and P.D. Spanos, Spectral stochastic finite-element formulation for reliability analysis, *Journal of Engineering Mechanics – ASCE*, 117 (10) (1991) 2351-2372.
- [1.51] R.G. Ghanem and W. Brzakala, Stochastic finite-element analysis of soil layers with random interface, *Journal of Engineering Mechanics – ASCE*, 122 (4) 1996 361-369.
- [1.52] R.G. Ghanem and R.M. Kruger, Numerical solution of spectral stochastic finite element systems, *Computer Methods in Applied Mechanics and Engineering*, 129 (3) (1996) 289-303.
- [1.53] R.G. Ghanem and P.D. Spanos, Spectral techniques for stochastic finite elements, *Archives of Computational Methods in Engineering*, 4 (1) (1997) 63-100.
- [1.54] R.G. Ghanem and S. Dham, Stochastic finite element analysis for multiphase flow in heterogeneous porous media, *Transport in Porous Media*, 32 (3) (1998) 239-262.
- [1.55] R.G. Ghanem, Probabilistic characterization of transport in heterogeneous media, *Computer Methods in Applied Mechanics and Engineering*, 158 (3-4) (1998) 199-220.
- [1.56] 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) (1999) 137-144.

- [1.57] R.G. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, *Computer Methods in Applied Mechanics and Engineering*, 168 (1-4) (1999) 19-34 (Included in Ref. [1.19]).
- [1.58] R.G. Ghanem, Stochastic finite elements with multiple random non-Gaussian properties, *Journal of Engineering Mechanics – ASCE*, 125 (1) 1999 26-40.
- [1.59] R.G. Ghanem, The nonlinear Gaussian spectrum of log-normal stochastic processes and variables, *Journal of Applied Mechanics - ASME*, 66 (4) (1999) 964-973.
- [1.60] S. Sakamoto and R.G. Ghanem, Polynomial chaos decomposition for the simulation of non-Gaussian nonstationary stochastic processes, *Journal of Engineering Mechanics – ASCE*, 128 (2) (2002) 190-201.
- [1.61] G. Deodatis, Weighted integral method I: stochastic stiffness matrix, *Journal of Engineering Mechanics – ASCE*, 117 (8) (1991) 1851-1864.
- [1.62] G. Deodatis and M. Shinozuka, Weighted integral method. II: response variability and reliability, *Journal of Engineering Mechanics – ASCE*, 117 (8) (1991) 1865-1877.
- [1.63] C.C. Li and A.D. Kiureghian, Optimal discretization of random fields, *Journal of Engineering Mechanics – ASCE*, 119 (6) (1993) 1136-1154.
- [1.64] S. Valliappan and T.D. Pham, Fuzzy finite element analysis of a foundation on an elastic soil medium, *International Journal for Numerical and Analytical Methods in Geomechanics*, 17 (11) (1993) 771-798.
- [1.65] S. Valliappan and T.D. Pham, Elastoplastic finite-element analysis with fuzzy parameters, *International Journal for Numerical Methods in Engineering*, 38 (4) (1995) 531-548.
- [1.66] S.S. Rao and J.P. Sawyer, Fuzzy finite element approach for the analysis of imprecisely defined systems, *AIAA Journal*, 33 (12) (1995) 2364-2370.
- [1.67] M.E. Cruz and A.T. Patera, A parallel Monte-Carlo finite-element - procedure for the analysis of multicomponent random-media, *International Journal for Numerical Methods in Engineering*, 38 (7) (1995) 1087-1121.
- [1.68] M. Papadrakakis and V. Papadopoulos, Robust and efficient methods for stochastic finite element analysis using Monte Carlo simulation, *Computer Methods in Applied Mechanics and Engineering*, 134 (3-4) (1996) 325-340.
- [1.69] G.I. Schuëller and H.J. Pradlwarter, On the stochastic response of nonlinear FE models, *Archive of Applied Mechanics*, 69 (9-10) (1999) 765-784.
- [1.70] I. Elishakoff and Y.J. Ren, The bird's eye view on finite element method for structures with large stochastic variations, *Computer Methods in Applied*

- Mechanics and Engineering*, 168 (1-4) (1999) 51-61 (Included in Ref. [1.19]).
- [1.71] S. Adhikari and C.S. Manohar, Dynamic analysis of framed structures with statistical uncertainties, *International Journal for Numerical Methods in Engineering*, 44 (8) (1999) 1157-1178.
- [1.72] G.I. Schuëller, Computational stochastic mechanics – recent advances, *Computers & Structures*, 79 (22-25) (2001) 2225-2234.
- [1.73] N. Impollonia and A. Sofi, A response surface approach for the static analysis of stochastic structures with geometrical nonlinearities, *Computer Methods in Applied Mechanics and Engineering*, 192 (37-38) (2003) 4109-4129.
- [1.74] C. Proppe, H.J. Pradlwarter and G.I. Schuëller, Equivalent linearization and Monte Carlo simulation in stochastic dynamics, *Probabilistic Engineering Mechanics*, 18 (1) (2003) 1-15.
- [1.75] G. Stefanou and M. Papadrakakis, Stochastic finite element analysis of shells with combined random material and geometric properties, *Computer Methods in Applied Mechanics and Engineering*, 193 (1-2) (2004) 139-160.
- [1.76] J. Guggenberger and H. Grundmann, Stochastic response of large FEM models with hysteretic behaviour in beam elements, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1739-1756 (Included in Ref. [1.20]).
- [1.77] F.J. Ferrante and L.L. Graham-Brady, Stochastic simulation of non-Gaussian/non-stationary properties in a functionally graded plate, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1675-1692 (Included in Ref. [1.20]).
- [1.78] H.J. Pradlwarter, M.F. Pellissetti, C.A. Schenk, G.I. Schuëller, A. Kreis, S. Fransen, A. Calvi and M. Klein, Realistic and efficient reliability estimation for aerospace structures, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1597-1617 (Included in Ref. [1.20]).
- [1.79] D.C. Charmpis and M. Papadrakakis, Improving the computational efficiency in finite element analysis of shells with uncertain properties, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1447-1478 (Included in Ref. [1.20]).
- [1.80] J. Ching, S.K. Liu and J.L. Beck, Reliability estimation for dynamical systems subject to stochastic excitation using subset simulation with splitting, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1557-1579 (Included in Ref. [1.20]).
- [1.81] L. Katafygiotis and S.H. Cheung, A two-stage subset simulation-based approach for calculating the reliability of inelastic structural systems subjected to Gaussian random excitations, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1581-1595 (Included in Ref. [1.20]).

- [1.82] V. Papadopoulos and M. Papadrakakis, The effect of material and thickness variability on the buckling load of shells with random initial imperfections, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1405-1426 (Included in Ref. [1.20]).
- [1.83] T.D. Hien and M. Kleiber, Finite-element analysis based on stochastic Hamilton variational principle, *Computers & Structures*, 37 (6) (1990) 893-902.
- [1.84] T.D. Hien and M. Kleiber, Stochastic design sensitivity in structural dynamics, *International Journal for Numerical Methods in Engineering*, 32 (6) (1991) 1247-1265.
- [1.85] I. Elishakoff, Y.J. Ren and M. Shinozuka, Improved finite element method for stochastic problems, *Chaos, Solitons and Fractals*, 5 (5) (1995) 833-846.
- [1.86] M. Kamiński and M. Kleiber, Stochastic structural interface defects in fiber composites, *International Journal of Solids and Structures*, 33 (20-22) (1996) 3035-3956.
- [1.87] T.D. Hien and M. Kleiber, Stochastic finite element modelling in linear transient heat transfer, *Computer Methods in Applied Mechanics and Engineering*, 144 (1-2) (1997) 111-124.
- [1.88] M. Kamiński and M. Kleiber, Numerical homogenization of N-component composite including stochastic interface defects, *International Journal for Numerical Methods in Engineering*, 47 (5) (2000) 1001-1027.
- [1.89] S. Rahman and B.N. Rao, A perturbation method for stochastic meshless analysis in elastostatics, *International Journal for Numerical Methods in Engineering*, 50 (8) (2001) 1969-1991.
- [1.90] B.V. Nieuwenhof and J.P. Coyette, Modal approaches for the stochastic finite element analysis of structures with material and geometric uncertainties, *Computer Methods in Applied Mechanics and Engineering*, 192 (33-34) (2003) 3705-3729.
- [1.91] H.C. Noh, A formulation for stochastic finite element analysis of plate structures with uncertain Poisson's ratio, *Computer Methods in Applied Mechanics and Engineering*, 193 (42-44) (2004) 4857-4873.
- [1.92] I. Doltsinis, Z. Kang and G. Cheng, Robust design of non-linear structures using optimization methods, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1779-1795 (Included in Ref. [1.20]).
- [1.93] H.G. Matthies, C.E. Brenner and C.G. Bucher, Uncertainties in probabilistic numerical analysis of structures and solids – stochastic finite elements, *Structural Safety*, 19 (3) (1997) 283-336.
- [1.94] H.G. Matthies and C.G. Bucher, Finite elements for stochastic media problems, *Computer Methods in Applied Mechanics and Engineering*, 168 (1-4) (1999)

3-17 (Included in Ref. [1.19]).

- [1.95] H.G. Matthies and A. Keese, Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1295-1331 (Included in Ref. [1.20]).
- [1.96] M. Anders and M. Hori, Stochastic finite element method for elasto-plastic body, *International Journal for Numerical Methods in Engineering*, 46 (11) (1999) 1897-1916.
- [1.97] M. Anders and M. Hori, Three-dimensional stochastic finite method for elasto-plastic bodies, *International Journal for Numerical Methods in Engineering*, 51 (4) (2001) 449-478.
- [1.98] O.P.L. Maître, O.M. Knio, H.N. Najm and R.G. Ghanem, A stochastic projection method for fluid flow I. basic formulation, *Journal of Computational Physics*, 173 (2) (2001) 481-511.
- [1.99] P. Frauenfelder, C. Schwab and R.A. Todor, Finite elements for elliptic problems with stochastic coefficients, *Computer Methods in Applied Mechanics and Engineering*, 194 (2-5) (2005) 205-228.
- [1.100] M.K. Deb, I.M. Babuška and J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, *Computer Methods in Applied Mechanics and Engineering*, 190 (48) (2001) 6359-6372.
- [1.101] I.M. Babuška, R. Tempone and G.E. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1251-1294 (Included in Ref. [1.20]).
- [1.102] I.M. Babuška, P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, *Computer Methods in Applied Mechanics and Engineering*, 191 (37-38) (2002) 4093-4122.
- [1.103] I.M. Babuška, R. Tempone and G.E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, *SIAM Journal on Numerical Analysis*, 42 (2) (2004) 800-825.
- [1.104] I.M. Babuška and K.M. Liu, On solving stochastic initial-value differential equations, *Mathematical Models and Methods in Applied Sciences*, 13 (5) (2003) 715-745.
- [1.105] D.B. Chung, M.A. Gutiérrez and R. Borst, Object-oriented stochastic finite element analysis of fibre metal laminates, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1427-1446 (Included in Ref. [1.20]).
- [1.106] D.B. Xiu and G.E. Karniadakis, The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations, *SIAM Journal on Scientific Computing*, 24

(2) (2002) 619-644.

- [1.107] D.B. Xiu and G.E. Karniadakis, Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos, *Computer Methods in Applied Mechanics and Engineering*, 191 (43) 4927-4948, 2002.
- [1.108] D.B. Xiu and G.E. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, *Journal of Computational Physics*, 187 (1) (2003) 137-167.
- [1.109] D.B. Xiu and G.E. Karniadakis, A new stochastic approach to transient heat conduction modeling with uncertainty, *International Journal of Heat and Mass Transfer*, 46 (24) (2003) 4681-4693.
- [1.110] D. Lucor, D.B. Xiu and C.H. Su, Predictability and uncertainty in CFD, *International Journal for Numerical Methods in Engineering*, 43 (5) (2003) 483-505.
- [1.111] X.L. Wang, D.B. Xiu and G.E. Karniadakis, Stochastic solutions for the two-dimensional advection-diffusion equation, *SIAM Journal on Scientific Computing*, 26 (2) (2004) 578-590.
- [1.112] D.B. Xiu and G.E. Karniadakis, Supersensitivity due to uncertain boundary conditions, *International Journal for Numerical Methods in Engineering*, 61 (12) (2004) 2114-2138.
- [1.113] D.B. Xiu and I.G. Kevrekidis, Equation-free, multiscale computation for unsteady random diffusion, *Multiscale Modeling & Simulation*, 4 (3) (2005) 915-935.
- [1.114] D.B. Xiu, I.G. Kevrekidis and R.G. Ghanem, An equation-free, multiscale approach to uncertainty quantification, *Computing in Science & Engineering*, 7 (3) (2005) 16-23.
- [1.115] D. Monens and D. Vandepitte, An interval finite element approach for the calculation of envelope frequency response functions, *International Journal for Numerical Methods in Engineering*, 61 (14) (2004) 2480-2507.
- [1.116] D. Moens and D. Vandepitte, A survey of non-probabilistic uncertainty treatment in finite element analysis, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1527-1555 (Included in Ref. [1.20]).
- [1.117] R. Popescu, G. Deodatis and J.H. Prevost, Simulation of homogeneous nonGaussian stochastic vector fields, *Probabilistic Engineering Mechanics*, 13 (1) (1998) 1-13.
- [1.118] C. Soize, A nonparametric model of random uncertainties for reduced matrix models in structural dynamics, *Probabilistic Engineering Mechanics*, 15 (3) (2000) 277-294.
- [1.119] C. Soize, Random matrix theory for modeling uncertainties in computational

- mechanics, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1331-1366 (Included in Ref. [1.20]).
- [1.120] J. Li and J.B. Chen, Dynamic response and reliability analysis of structures with uncertain parameters, *International Journal for Numerical Methods in Engineering*, 62 (2) (2004) 289-315.
- [1.121] H. Xu and S. Rahman, A generalized dimension-reduction method for multidimensional integration in stochastic mechanics, *International Journal for Numerical Methods in Engineering*, 61 (12) (2004) 1992-2019.
- [1.122] X.F. Xu and L. Graham-Brady, A stochastic computational method for evaluation of global local behavior of random elastic media, *Computer Methods in Applied Mechanics and Engineering*, 194 (42-44) (2005) 4362-4385.
- [1.123] P.D. Spanos, J. Tezcan and P. Tratskas, Stochastic processes evolutionary spectrum estimation via harmonic wavelets, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1367-1383 (Included in Ref. [1.20]).
- [1.124] C.F. Li, Y.T. Feng and D.R.J. Owen, Continuum mechanics and finite elements in stationary stochastic fields: the formulation, *The Third M.I.T. Conference on Computational Fluid and Solid Mechanics*, Cambridge, MA, USA, June 14-17, 2005.
- [1.125] Y.T. Feng, C.F. Li, D.R.J. Owen, Continuum mechanics and finite elements in stationary stochastic fields: computational issues, *The Third M.I.T. Conference on Computational Fluid and Solid Mechanics*, Cambridge, MA, USA, June 14-17, 2005.
- [1.126] C.F. Li, Y.T. Feng, D.R.J. Owen and I.M. Davies, Fourier representation of random media fields in stochastic finite element modelling, *Engineering Computations*, 23 (7) (2006) 794-817.
- [1.127] C.F. Li, Y.T. Feng and D.R.J. Owen, Explicit solution to stochastic system of linear algebraic equations $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)x = b$, *Computer Methods in Applied Mechanics and Engineering*, 195 (44-47) (2006) 6560-6576.
- [1.128] C.F. Li, Y.T. Feng, D.R.J. Owen and I.M. Davies, A simple semi-analytic solution for Karhunen-Loève expansions of stationary stochastic fields, *The Seventh World Congress on Computational Mechanics*, Los Angeles, California, USA, July 16-22, 2006.
- [1.129] Y.T. Feng, C.F. Li and D.R.J. Owen, Joint diagonalization solution for the stochastic system of linear algebraic equations $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)x = b$, *The Seventh World Congress on Computational Mechanics*, Los Angeles, California, USA, July 16-22, 2006.

Chapter 2

Basic Mathematical Tools

Stochastic mathematics, including probability theory, the theory of stochastic fields[†] and stochastic analysis etc., studies quantitative relations in various random phenomena, and has become one of the major branches in mathematics. This chapter contains a brief excursion through the foundations of stochastic mathematics, which provides some basic but necessary preparations for stochastic modelling and analysis of random media, and also initiates the specific approach to be developed in Chapters 3-7. Since they have been included mainly for the sake of completeness, and contain quite standard results, proofs are not given. The detailed and complete description of these mathematical results can be found in the references listed at the end of this chapter. Specifically, good basic treatments of probability theory, the theory of stochastic fields and stochastic analysis are included respectively in Refs. [2.9-2.14], [2.15-2.20] and [2.21-2.30].

The history of *probability theory* goes back as early as the 17th century, when two French mathematicians, Pascal [2.1] and Fermat [2.1], worked on some probabilistic problems arising from gambling, in which the basic principles of probability theory were formulated for the first time. In the 18th century, the major contributors were Bernoulli [2.1], who proposed and proved the first limit theorem in probability theory (a weak law of large numbers), and De Moivre [2.1]. The most important contributions in the 19th

[†] Throughout this thesis, a stochastic field can be defined in an n -dimensional ($n \in \mathbb{N}$) domain and in order to avoid possible confusions, the term stochastic process (which is probably more popular in mathematics textbooks) is only used for one-dimensional stochastic fields.

century were made by Laplace [2.1], who introduced a host of new ideas and mathematical techniques, and proved the second limit theorem in probability theory (an early form of the central limit theorem which was first proposed by De Moivre). However, the modern understanding of probability theory is due to Kolmogorov [2.1], who in 1933 [2.2] provided a measure-theoretic foundation for probability which is now treated as axiomatic. Similar to many other branches of mathematics, the development of probability theory has been stimulated by the variety of its applications. Conversely, each advance in the theory has enlarged the scope of its influence. Mathematical statistics is one important branch of applied probability; other applications occur in such widely different fields as genetics, psychology, economics, and engineering.

The study of *stochastic fields*, by definition, is the study of random functions defined over some Euclidean space. In 1905 [2.3], Einstein [2.1] discovered the quantitative description of Brownian motion, which turned out to be the first milestone in the research of stochastic fields. In 1923 [2.4], Wiener [2.1] presented the mathematical definition of Brownian motion, i.e. the Wiener process. However, the concept of probability was neither precise nor comprehensive at that time, and was insufficient for the study of stochastic fields. Research on the general theory of stochastic fields started in the 1930s. In 1931 [2.5], Kolmogorov laid the foundation of the theory of Markov processes; in 1934 [2.6], Khinchin [2.1] established the theoretical basis for stationary stochastic fields. In 1953 [2.7], Doob [2.1] systematically and rigorously summarized the theory of stochastic processes including much of his own development of martingale theory.

Stochastic analysis is one of the youngest branches in stochastic mathematics, and it studies the integration and differentiation of functionals of stochastic fields. The foundation of stochastic analysis was established in the 1940s by Itô [2.1], who developed the theory of stochastic (ordinary) differential equations (SODEs), i.e. Itô integral. The theory of SODEs deals with the integration and differentiation of functionals of Wiener processes and its applications in such area as economics have been extremely successful [2.8]. Stimulated by the success of the Itô integral, continuous efforts have been made in the past two decades to extend the SODE theory into SPDEs. Nevertheless, the research of SPDEs is still very much on going and many fundamental questions are still outstanding.

Remark: Throughout this chapter, a vector is always assumed as a row vector

unless otherwise stated. In addition, some basic results in the classic measure theory [2.31-2.32] and functional analysis [2.33-2.35] are accepted and used without additional explanation.

2.1 On Probability Theory

2.1.1 Probability Space and Random Variables

Let Ω be a set with generic element ω , \mathbb{F} be a σ -algebra of subsets of Ω , P be a measure defined on \mathbb{F} . In measure theory, (Ω, \mathbb{F}) is termed a measurable space while (Ω, \mathbb{F}, P) is termed a measure space. In probability theory, (Ω, \mathbb{F}, P) is termed a *probability space* if P further satisfies

$$0 \leq P\{A\} \leq 1 \quad (\forall A \in \mathbb{F}) \quad \text{and} \quad P\{\Omega\} = 1. \quad (2.1)$$

The elements of \mathbb{F} will generally be referred to as *events*. In particular, Ω is called the *sure event* and ω the *basic event*.

Let \mathbb{R}^m ($m \in \mathbb{N}$) denote the Euclidean m -space, and \mathbb{B}^m the Borel σ -algebra on \mathbb{R}^m . Then, $(\mathbb{R}^m, \mathbb{B}^m)$ forms a measurable space. Now, let X be a measurable mapping from (Ω, \mathbb{F}) into $(\mathbb{R}^m, \mathbb{B}^m)$, so that for every Borel set $B \in \mathbb{B}^m$, the set

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \quad (2.2)$$

is an element of \mathbb{F} . Then X is called an \mathbb{R}^m -valued *random variable*. For each ω , $X(\omega)$ is a vector of length m , and its coordinates is often written as $X_1(\omega)$, $X_2(\omega)$, ..., $X_m(\omega)$. Since each X_j ($j=1,2,\dots,m$) is a measurable mapping from (Ω, \mathbb{F}) to (\mathbb{R}, \mathbb{B}) , it is also a well-defined random variable. When $m=1$, X is said to be *real valued*, and when $m=2$ X can be considered as *complex valued* with obvious conventions.

For the sake of simplicity, the term *random vector* is not used in this chapter, and instead the term random variable is used in all occasions, including real-valued X and

\mathbb{R}^m -valued X . In addition, without special statement, a random variable is always real valued or \mathbb{R}^m valued. Nevertheless, it is straightforward to generalize some of the mathematical results addressed in this chapter into complex-valued or \mathbb{C}^m -valued random variables.

The measurability of X induces a probability measure F_X (or simply F) on \mathbb{B}^m , defined by

$$F_X(B) \triangleq P\{X^{-1}(B)\} \quad (2.3)$$

for every $B \in \mathbb{B}^m$. To this measure there corresponds a point function, which is also denoted by F_X (or simply F), defined by

$$F_X(\mathbf{x}) \triangleq F_X((-\infty, x_1] \times (-\infty, x_2] \times \cdots \times (-\infty, x_m]). \quad (2.4)$$

This function is called the *distribution function* of the random variable X . Two random variables X and Y are said to be *identically distributed* if they have the same distribution functions, i.e. $F_X \equiv F_Y$. Using λ_m (or simply λ) to denote the Lebesgue measure on \mathbb{R}^m and assuming that F_X is absolutely continuous, there exists a non-negative Borel function $f_X: \mathbb{R}^m \rightarrow \mathbb{R}$ for which

$$F_X(B) = \int_B f_X(\mathbf{x}) d\lambda(\mathbf{x}) \quad (2.5)$$

for any $B \in \mathbb{B}^m$. This function is known as the *probability density function* of the random variable X .

If X is an \mathbb{R}^m -valued random variable and $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a Borel function (so that $g^{-1}(B) \in \mathbb{B}^m$ for every $B \in \mathbb{B}^n$), then $Y = g(X(\omega))$ is a measurable mapping from (Ω, \mathbb{F}) to $(\mathbb{R}^n, \mathbb{B}^n)$, and so is an \mathbb{R}^n -valued random variable. Its corresponding probability measure F_Y is defined by

$$F_Y(B) = F_X(g^{-1}(B)) = P\{X^{-1}(g^{-1}(B))\} \quad (2.6)$$

for any $B \in \mathbb{B}^n$.

Since the notation

$$P\{\omega: X(\omega) \in B\} \quad (2.7)$$

is unduly clumsy, the simpler, but less exact, notation

$$P\{X \in B\} \quad (2.8)$$

is generally used instead.

In probability theory, the interesting properties of random variables are probabilistic, so that there is no difference between two variables that are equal with probability one. Hence, random variables X and Y are called *equivalent* and considered to be indistinguishable if

$$P\{X = Y\} = 1. \quad (2.9)$$

An important special case is that, a random variable X is said to be *degenerate* at X_d if $P\{X = X_d\} = 1$ and X_d is a constant. In this light, deterministic variables can be regarded as a special class of random variables.

Note that any \mathbb{R}^m -valued random variable X can in fact be treated as a collection of m real-valued variables, by considering the m coordinates of X separately. When considering X in this viewpoint the function $F(\mathbf{x})$ is called the *joint distribution function* of these variables, and questions of dependence, independence, and conditional probability arise. Two \mathbb{R}^m - and \mathbb{R}^n -valued random variables X and Y , defined on the same probability space (Ω, \mathbb{F}, P) , are called *independent* if

$$P\{X \in B_1, Y \in B_2\} = P\{X \in B_1\} P\{Y \in B_2\} \quad (2.10)$$

for all $B_1 \in \mathbb{B}^m$ and $B_2 \in \mathbb{B}^n$. The *conditional probability* of $X \in B_1$ given $Y \in B_2$ is defined by

$$P\{X \in B_1 | Y \in B_2\} \triangleq \frac{P\{X \in B_1, Y \in B_2\}}{P\{Y \in B_2\}} \quad (2.11)$$

for all pairs of Borel sets $B_1 \in \mathbb{B}^m$ and $B_2 \in \mathbb{B}^n$ satisfying $P\{Y \in B_2\} > 0$. This leads naturally to the concept of *conditional distribution function* of X given Y , defined, in measure form, by

$$F_{X|Y}(B_1 | B_2) \triangleq \frac{\int_{B_1} \int_{B_2} dF_{(X,Y)}(\mathbf{x}, y)}{\int_{B_2} dF_Y(y)} \quad (2.12)$$

where $F_{(X,Y)}$ is the distribution function for the \mathbb{R}^{m+n} -valued random variable (X, Y) .

If all the distribution functions are absolutely continuous, the *conditional probability density function* of X given Y can be defined as

$$f_{X|Y}(\mathbf{x}, y) \triangleq \frac{f_{(X,Y)}(\mathbf{x}, y)}{f_Y(y)}. \quad (2.13)$$

Finally, two events (elements of \mathbb{F}) A and B are called *independent events* if

$$P\{A \cap B\} = P\{A\}P\{B\}. \quad (2.14)$$

The *conditional probability* of A given B is defined by

$$P\{A|B\} \triangleq \frac{P\{A \cap B\}}{P\{B\}}. \quad (2.15)$$

whenever $P\{B\} > 0$. It is left undefined, or given an arbitrary value, when $P\{B\} = 0$.

2.1.2 Statistical Moments and the Characteristic Function

If X is an \mathbb{R}^m -valued random variable and $\mathbf{g}: \mathbb{R}^m \rightarrow \mathbb{R}^n$ a Borel function, then the *expectation* of $\mathbf{g}(X)$ is defined as the Lebesgue-Stieltjes integral

$$E(\mathbf{g}(X)) \triangleq \int_{\Omega} \mathbf{g}(X(\omega)) dP(\omega), \quad (2.16)$$

provided that the integral exists. Because of the measurability of $X(\omega)$, this is equal to the Riemann-Stieltjes integral given by

$$E(\mathbf{g}(X)) = \int_{\mathbb{R}^m} \mathbf{g}(\mathbf{x}) dF_X(\mathbf{x}). \quad (2.17)$$

Since $\mathbf{g}(X) = (g_1(X), g_2(X), \dots, g_n(X))$ is \mathbb{R}^n valued, so is $E(\mathbf{g}(X))$. Noting this, it is easy to see that $E(\mathbf{g}(X))$ exists if $\int_{\mathbb{R}^m} |g_i(\mathbf{x})| dF_X(\mathbf{x}) < +\infty$ for $i = 1, 2, \dots, n$.

If X is a real-valued random variable and $E(|X|) < +\infty$, X is said to be *integrable*. If $E(|X|^2) < +\infty$, X is said to be *square integrable*. The simple expectation

$$\mu_X = E(X) \quad (2.18)$$

is termed the *mean* of X , and X is called *centred* if $\mu_X = 0$. The quantity

$$\sigma_X^2 = E((X - \mu_X)^2) = \text{Var}(X) \quad (2.19)$$

is termed the *variance* of X , while $\sigma_X \geq 0$ is known as the *standard deviation*. More generally, for any $j \in \mathbb{N}$, the expectations $E(X^j)$, $E((X - \mu_X)^j)$, $E(|X|^j)$ and $E(|X - \mu_X|^j)$ are termed respectively the (j -th order) *moment*, *central moment*, *absolute moment* and *central absolute moment* of X .

The concept of statistical moments can also be defined for \mathbb{R}^m -valued random variables. For example, the product expectation

$$\sigma_{XY}^2 = E((X - \mu_X)(Y - \mu_Y)) = \text{Cov}(X, Y) \quad (2.20)$$

involving two real-valued random variables X and Y is known as *covariance* (i.e. second order central moment), while

$$\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y} \quad (2.21)$$

is called the *correlation coefficient* of X and Y . Note when $\text{Cov}(X, Y) = 0$, X and Y are said to be *uncorrelated* to each other.

With respect to the expectation of real-valued random variables, there are the following properties:

- Consider the constant c as a special random variable, then $E(c) = c$.
- For any constants a, b and random variables X, Y ,

$$E(aX + bY) = aE(X) + bE(Y). \quad (2.22)$$

- If $X \geq 0$, then $E(X) \geq 0$.
- If random variables X and Y are independent to each other, then

$$E(XY) = E(X)E(Y). \quad (2.23)$$

With respect to the variance of real-valued random variables, the following properties hold

- $\text{Var}(X) = E(X^2) - (E(X))^2$.
- The variance of a random variable X equals zero if, and only if, there exists a constant c such that

$$P\{X = c\} = 1. \quad (2.24)$$

- For any constant a ,

$$\text{Var}(aX) = a^2 \text{Var}(X). \quad (2.25)$$

- For any constant a ,

$$\text{Var}(X) \leq E((X - a)^2), \quad (2.26)$$

and the equal sign in the above inequality holds if and only if $a = E(X)$.

- If random variables X and Y are uncorrelated to each other, then

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y). \quad (2.27)$$

Let X be a real-valued random variable with distribution function $F(x)$. The function

$$\varphi(t) = E(e^{\sqrt{-1}tx}) = \int_{\Omega} e^{\sqrt{-1}tX(\omega)} dP(\omega) = \int_{-\infty}^{+\infty} e^{\sqrt{-1}tx} dF(x) \quad (2.28)$$

defined for all real numbers t , is called the *characteristic function* of X or $F(x)$. If $F(x)$ is absolutely continuous such that $f(x)$, the probability density function of X , exists, then

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{\sqrt{-1}tx} dF(x) = \int_{-\infty}^{+\infty} e^{\sqrt{-1}tx} f(x) dx. \quad (2.29)$$

That is, differing by a constant scalar $(2\pi)^{-1}$, $\varphi(t)$ is the inverse Fourier transform of $f(x)$, and $f(x)$ can be expressed in terms of the Fourier transform of $\varphi(t)$, i.e.

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\sqrt{-1}tx} \varphi(t) dt. \quad (2.30)$$

With respect to the characteristic function $\varphi(t)$, there are the following properties:

- (a) $\varphi(t)$ is well defined for all $t \in \mathbb{R}$ and satisfies $\varphi(0) = 1$, $|\varphi(t)| \leq 1$; (b) $\varphi(t)$ is uniformly continuous in \mathbb{R} ; (c) $\varphi(t)$ is *non-negative definite* (See also Section 2.2.3), that is,

$$\sum_{i=1}^k \sum_{j=1}^k \varphi(t_i - t_j) z_i \bar{z}_j \geq 0 \quad (2.31)$$

holds for any positive integer k , any set of real numbers t_1, t_2, \dots, t_k and any set of complex numbers z_1, z_2, \dots, z_k . Conversely, for any function $g(t)$ satisfying

these three conditions (a, b, c), there exists a random variable whose characteristic function is $g(t)$.

- The characteristic function $\varphi(t)$ and the distribution function $F(x)$ are uniquely determined by each other. Furthermore, if the probability density function $f(x)$ exists, then $F(x)$, $f(x)$ and $\varphi(t)$ are equivalent in defining the associated random variable X .
- If X_1, X_2, \dots, X_k are independent random variables with characteristic functions $\varphi_{X_1}(t), \varphi_{X_2}(t), \dots, \varphi_{X_k}(t)$, then the characteristic function of $X_1 + X_2 + \dots + X_k$ can be obtained as

$$\varphi_{X_1+X_2+\dots+X_k}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)\cdots\varphi_{X_k}(t). \quad (2.32)$$

- For any constants a and b , the characteristic function of $aX + b$ is

$$\varphi_{aX+b}(t) = e^{\sqrt{-1}bt} \varphi_X(at). \quad (2.33)$$

- If, for some integer $k > 0$, the moment $E(X^k)$ is finite, then $\varphi(t)$ is differentiable k times and

$$\varphi^{(i)}(0) = (\sqrt{-1})^i E(X^i) \quad 0 \leq i \leq k. \quad (2.34)$$

If moments $E(X^i)$ for all orders are finite, then the Taylor expansion

$$\varphi(t) = \sum_{i=0}^{+\infty} E(X^i) \frac{(\sqrt{-1}t)^i}{i!} \quad (2.35)$$

holds for all t for which the right-hand side converges, which is known from calculus to be equivalent to $|t| < 1/\varepsilon$, where $\varepsilon = \limsup_{i \rightarrow +\infty} \left(|E(X^i)|/i! \right)^{1/i}$.

More generally, the characteristic function $\varphi(t)$ of an \mathbb{R}^m -valued random variable X with distribution function $F(x)$ is defined by

$$\varphi(t) \triangleq E\left(e^{\sqrt{-1}t \cdot X}\right) = \int_{\mathbb{R}^m} e^{\sqrt{-1}t \cdot x} dF(x) \quad t \in \mathbb{R}^m. \quad (2.36)$$

Similar to the real-valued case, when the corresponding probability density function $f(x)$ exists, $\varphi(t)$ and $f(x)$ are related to each other via Fourier transforms, i.e.

$$\varphi(t) = \int_{\mathbb{R}^m} e^{\sqrt{-1}t \cdot x} f(x) dx, \quad (2.37)$$

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-\sqrt{-1}t \cdot x} \varphi(t) dt. \quad (2.38)$$

2.1.3 Fundamental Inequalities

There are a number of inequalities that play important roles in probability theory, and they are outlined below for references.

Theorem 2.1 (Chebyshev's inequality) Let X be an integrable random variable, then the inequality

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{\text{Var}(X)}{\varepsilon^2} \quad (2.39)$$

holds for any $\varepsilon > 0$.

Theorem 2.2 (Kolmogorov's inequality) Let X_1, X_2, \dots, X_k be k independent random variables that satisfy $E(X_i) = 0$ and $\text{Var}(X_i) = \sigma_i^2 < +\infty$ for all $i = 1, 2, \dots, k$. Then, the inequality

$$P\left\{\max_{1 \leq i \leq k} |X_1 + X_2 + \dots + X_i| \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2} (\sigma_1^2 + \sigma_2^2 + \dots + \sigma_k^2) \quad (2.40)$$

holds for any $\varepsilon > 0$.

Theorem 2.3 (Hölder's inequality) Let X and Y be two random variables, and $p > 1$, $q > 1$ such that $1/p + 1/q = 1$. Then

$$E(|XY|) \leq \left(E(|X|^p)\right)^{1/p} \left(E(|Y|^q)\right)^{1/q}. \quad (2.41)$$

If $\left(E(|X|^p)\right)^{1/p} < +\infty$ and $\left(E(|Y|^q)\right)^{1/q} < +\infty$, then the equal sign in the above inequality holds if and only if $P\{X = 0\} = 1$ or $P\{Y = 0\} = 1$ or there exists a constant $c > 0$ such that $P\{|X|^p = c|Y|^q\} = 1$. When $p = q = 2$, this is known as the *Cauchy-Schwartz inequality*.

Theorem 2.4 Let X_1, X_2, \dots, X_k be k random variables. Then

$$E(|X_1 X_2 \dots X_k|) \leq \prod_{i=1}^k \left(E(|X_i|^k)\right)^{1/k}. \quad (2.42)$$

Theorem 2.5 (Jensen's inequality) Let X be a random variable and g a convex function in \mathbb{R} . That is, for any $\lambda \in [0,1]$ and $u < v \in \mathbb{R}$,

$$g(u + \lambda(v-u)) \leq g(u) + \lambda(g(v) - g(u)). \quad (2.43)$$

Then,

$$g(E(X)) \leq E(g(X)), \quad (2.44)$$

if both $E(|X|)$ and $E(|g(X)|)$ are finite.

Theorem 2.6 (Moment inequality) Let X be a random variable and $0 < s < t$, then

$$\left(E(|X|^s)\right)^{1/s} \leq \left(E(|X|^t)\right)^{1/t}. \quad (2.45)$$

When $E(|X|^t) < +\infty$, the equal sign in the above inequality holds if, and only if, there exists a constant c such that $P\{X = c\} = 1$.

Theorem 2.7 (Minkowski inequality) Let X and Y be two random variables, and $p > 0$. If $E(|X|^p + |Y|^p) < +\infty$, then

$$\text{I) } p \geq 1 \quad \left(E(|X+Y|^p)\right)^{1/p} \leq \left(E(|X|^p)\right)^{1/p} + \left(E(|Y|^p)\right)^{1/p}, \quad (2.46a)$$

in which the equal sign holds if and only if: (a) when $p > 1$, $P\{X = 0\} = 1$ or $P\{Y = 0\} = 1$ or there exists a constant c such that $P\{X = cY\} = 1$; or (b) when $p = 1$, $P\{XY \geq 0\} = 1$.

$$\text{II) } 0 < p < 1 \quad E(|X+Y|^p) \leq E(|X|^p) + E(|Y|^p), \quad (2.46b)$$

in which the equal sign holds if and only if $P\{XY = 0\} = 1$.

2.1.4 Stochastic Convergence

In calculus, the concept of convergence is the mathematical basis of continuity, differentiability and integrability of functions. Similarly, the concepts of stochastic convergence play a fundamental role in stochastic mathematics. There are four basic modes of convergence for a sequence $\{X_i\}$ of \mathbb{R}^m -valued random variables.

I) X_i is said to converge to X with probability one, or almost surely (a.s.), if

$$P\left\{\lim_{i \rightarrow +\infty} \|X_i - X\|_2 = 0\right\} = 1, \quad (2.47)$$

or equivalently, if there exists a set $N \subset \Omega$ such that $P\{N\} = 0$ and

$$\lim_{i \rightarrow +\infty} \|X_i(\omega) - X(\omega)\|_2 = 0 \quad (2.48)$$

holds for every $\omega \notin N$.

II) X_i is said to converge to X in γ -th mean ($\gamma \geq 1$) if

$$E\left(\|X_i - X\|_2^\gamma\right) \rightarrow 0 \quad \text{as } i \rightarrow +\infty. \quad (2.49)$$

III) X_i is said to converge to X in probability if

$$P\left\{\|X_i - X\|_2 > \varepsilon\right\} \rightarrow 0 \quad \text{as } i \rightarrow +\infty \quad (2.50)$$

holds for every $\varepsilon > 0$.

IV) X_i is said to converge to X in distribution, or weakly, if their distribution functions satisfy

$$F_{X_i}(\mathbf{x}) \rightarrow F_X(\mathbf{x}) \quad \text{as } i \rightarrow +\infty \quad (2.51)$$

at every continuity point \mathbf{x} of $F_X(\mathbf{x})$.

These four modes of convergence are denoted symbolically by

$$\begin{array}{ll} \text{I) } X_i \xrightarrow{\text{a.s.}} X, & \text{II) } X_i \xrightarrow{\gamma} X, \\ \text{III) } X_i \xrightarrow{P} X, & \text{IV) } X_i \xrightarrow{d} X. \end{array} \quad (2.52)$$

Convergence in γ -th mean when $\gamma = 2$ is of particular importance, and is known as *mean square convergence*, which is often denoted by

$$X_i \xrightarrow{\text{m.s.}} X, \quad X_i \xrightarrow{\text{q.m.}} X \quad \text{or} \quad X = \underset{i \rightarrow +\infty}{\text{l.i.m.}} X_i, \quad (2.53)$$

where l.i.m. is read as *limit in mean*.

In each of the first three cases there is a corresponding mode of *mutual convergence* of a sequence $\{X_i\}$ which is often useful. A sequence of $\{X_i\}$ is said to converge mutually almost surely, mutually in γ -th mean, or mutually in probability, if (2.54a), (2.54b) and (2.54c), respectively, holds:

$$\sup_{j>i} \|X_j - X_i\|_2 \xrightarrow{\text{a.s.}} 0 \quad \text{as } i \rightarrow +\infty, \quad (2.54a)$$

$$\sup_{j>i} E\left(\|X_j - X_i\|_2^\gamma\right) \rightarrow 0 \quad \text{as } i \rightarrow +\infty, \quad (2.54b)$$

$$\sup_{j>i} P\left\{\|X_j - X_i\|_2 > \varepsilon\right\} \rightarrow 0 \quad \text{as } i \rightarrow +\infty, \quad \forall \varepsilon > 0. \quad (2.54c)$$

In each case, the mutual convergence of a sequence of random variables in a particular mode implies the existence of a limit random variable to which the sequence converges in that mode, and vice versa.

Let $\gamma_1 \geq \gamma_2 \geq 1$, the relation among various modes of stochastic convergence can be represented as the following diagram, in which $A \Rightarrow B$ denotes “ A leads to B ”.

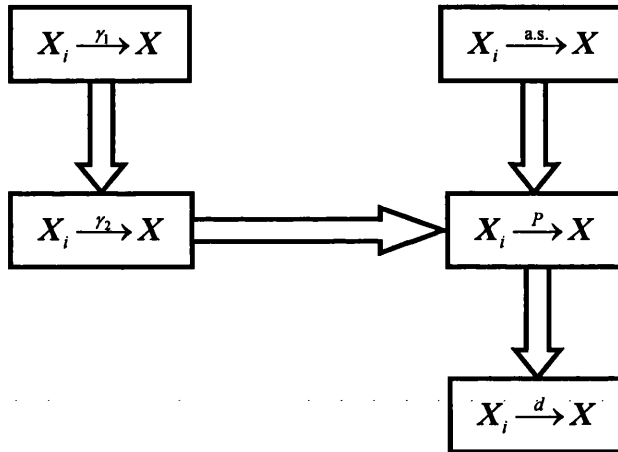


Figure 2.1 The relation of stochastic convergence

The following theorems summarize a number of results relating the various modes of stochastic convergence.

Theorem 2.8 Let $\{X_i\}$ be a sequence of random variables:

- I) If $X_i \xrightarrow{P} X$ then there is a subsequence $\{X_{i_k}\}$ of $\{X_i\}$ such that $X_{i_k} \xrightarrow{\text{a.s.}} X$ as $k \rightarrow +\infty$.
- II) If $X_i \xrightarrow{d} X$ and X is degenerate, then $X_i \xrightarrow{P} X$.
- III) $X_i \xrightarrow{d} X$ if, and only if, $\varphi_{X_i}(t) \rightarrow \varphi_X(t)$ for every t .

Theorem 2.9 (Lévy's monotone convergence theorem) Let $\{X_i\}$ be a

sequence of non-negative random variables for which $X_i(\omega) \leq X_{i+1}(\omega)$ for all i and ω , and suppose there exists a random variable X such that $X_i \xrightarrow{\text{a.s.}} X$. Then

$$\lim_{i \rightarrow +\infty} E(X_i) = E(X). \quad (2.55)$$

Theorem 2.10 (Fatou's lemma) Let $\{X_i\}$ be a sequence of random variables, and suppose there exists an integrable random variable X such that $X_i(\omega) \geq X(\omega)$ for all i and ω . Then

$$\liminf_{i \rightarrow +\infty} E(X_j) \geq E\left(\liminf_{j \geq i} X_j\right). \quad (2.56)$$

Theorem 2.11 (Lebesgue's dominated convergence theorem) Let $\{X_i\}$ be a sequence of random variables and X a random variable such that $X_i \xrightarrow{\text{a.s.}} X$. Suppose that there exists an integrable random variable Y for which $|X_i(\omega)| \leq Y(\omega)$ for all i and ω . Then

$$\lim_{i \rightarrow +\infty} E(X_i) = E(X). \quad (2.57)$$

Theorem 2.12 (The Borel-Contelli lemma) Let $\{A_i\}$ be a sequence of events.

If $\sum_{i=1}^{+\infty} P\{A_i\} < 1$, then the probability of an infinite number of the A_i occurring is zero, i.e.

$$P\left\{\bigcap_{i=1}^{+\infty} \bigcup_{j=i}^{+\infty} A_j\right\} = 0. \quad (2.58)$$

Theorem 2.13 (Kolmogorov's zero-one law) Let $\{A_i\}$ be a sequence of independent events. Corresponding to the convergence or divergence of $\sum_{i=1}^{+\infty} P\{A_i\}$, the probability of an infinite number of the A_i occurring is zero or one, respectively.

2.1.5 Limit Theorems

Until about sixty years ago, limit theorems have been the central problems in

probability theory for more than two centuries. There are two basic types of limit theorems: (a) *Laws of large numbers* that determine whether or not the arithmetic average of a sequence of random variables degenerate to a constant; (b) *Central limit theorems* that determine whether or not the partial sum of a sequence of random variables asymptotically follows a Gaussian distribution.

Laws of large numbers are established with respect to different modes of stochastic convergence. In particular, there are *weak laws of large numbers*, which correspond to convergence in probability, and *strong laws of large numbers*, which corresponds to almost sure convergence.

Theorem 2.14 (Bernoulli's weak law of large numbers) In k repeated independent trials of a random experiment, let p denote the probability that the event A occurs in any given trial, v_k the total number of times that A occurs in k trials. Then,

$$\lim_{k \rightarrow +\infty} P \left\{ \left| \frac{v_k}{k} - p \right| \geq \varepsilon \right\} = 0 \quad (2.59)$$

holds for any $\varepsilon > 0$.

Theorem 2.15 (due to Chebyshev) Let $\{X_i\}$ be a sequence of uncorrelated random variables. If there exists a constant c such that

$$\text{Var}(X_i) \leq c \quad (2.60)$$

for all X_i , then

$$\lim_{k \rightarrow +\infty} P \left\{ \left| \frac{S_k - E(S_k)}{k} \right| \geq \varepsilon \right\} = 0 \quad (2.61)$$

holds for all $\varepsilon > 0$, where $S_k = \sum_{i=1}^k X_i$.

Theorem 2.16 (due to Kolmogorov) Let $\{X_i\}$ be a sequence of independent and identically distributed random variables. Then,

$$P \left\{ \lim_{k \rightarrow +\infty} \frac{1}{k} \sum_{i=1}^k X_i = \mu \right\} = 1 \quad (2.62)$$

holds if and only if $E(X_i)$ exists and $E(X_i) = \mu$.

Theorem 2.17 (Kolmogorov's strong law of large numbers) Let $\{X_i\}$ be a sequence of independent random variables satisfying

$$\sum_{i=1}^{+\infty} \frac{\text{Var}(X_i)}{i^2} < +\infty, \quad (2.63)$$

then

$$P \left\{ \lim_{k \rightarrow +\infty} \frac{S_k - E(S_k)}{k} = 0 \right\} = 1 \quad (2.64)$$

where $S_k = \sum_{i=1}^k X_i$.

In science and engineering, many practical problems are influenced by a number of independent random factors. When the effect of each random factor is small, the overall effect is often observed following a Gaussian distribution. It is the central limit theorem that demonstrates this intuitive observation in a mathematically rigorous manner.

Theorem 2.18 (due to De Moivre and Laplace) Assume random variable v_k to follow the binomial distribution $B(k, p)$, i.e.

$$P\{v_k = i\} = \binom{k}{i} p^i q^{k-i} \quad (2.65)$$

where $i \leq k \in \mathbb{N}$, $0 < p < 1$ and $q = 1 - p$. Then

$$\lim_{k \rightarrow +\infty} P \left\{ a < \frac{v_k - kp}{\sqrt{kpq}} \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \quad (2.66)$$

holds for any $a < b \in \mathbb{R}$.

Theorem 2.19 (due to Lindeberg and Lévy) Let $\{X_i\}$ be a sequence of independent and identically distributed random variables. If $\text{Var}(X_i)$ exists, then

$$\lim_{k \rightarrow +\infty} P \left\{ \frac{S_k - k\mu}{\sqrt{k\sigma^2}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy \quad (2.67)$$

where $S_k = \sum_{i=1}^k X_i$, $\mu = E(X_i)$ and $\sigma^2 = \text{Var}(X_i)$.

Theorem 2.20 (due to Lindeberg and Feller) Let $\{X_i\}$ be a sequence of

independent random variables with distribution functions $F_i(x)$, $i \in \mathbb{N}$; let $a_i = E(X_i)$,

$b_i^2 = \text{Var}(X_i)$ and $B_k^2 = \sum_{i=1}^k b_i^2$. If the Lindeberg condition holds, i.e.

$$\lim_{k \rightarrow +\infty} \frac{1}{B_k^2} \sum_{i=1}^k \int_{|x-a_i| > \varepsilon B_k} (x-a_i)^2 dF_i(x) = 0 \quad \forall \varepsilon > 0, \quad (2.68)$$

then

$$\lim_{k \rightarrow +\infty} P \left\{ \frac{1}{B_k} \sum_{i=1}^k (X_i - a_i) \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy \quad \forall x \in \mathbb{R} \quad (2.69)$$

and the Feller condition

$$\lim_{k \rightarrow +\infty} \frac{1}{B_k^2} \max_{1 \leq i \leq k} \sigma_i^2 = 0 \quad (2.70)$$

hold; and vice versa.

2.2 On Stochastic Fields

Simply speaking, a stochastic field is just a random function defined over some index set. Hence, the research of stochastic fields covers an extremely wide area, since any question that can be asked about deterministic functions can be expected for their stochastic counterparts. In the study of stochastic fields, the greatest success is achieved in a special class of stochastic fields, namely Markov processes, which play a very important role in many random problems (especially those involved with time evolution). However, the Markov process has little to do with this research regarding random media, which will become clear in the latter parts of this thesis (e.g. Section 2.3 and Chapter 3). Consequently, the theory of Markov processes is not included in this section.

2.2.1 General Foundations

There are two virtually different approaches to defining stochastic fields. One is essentially a measure-theoretic approach and leads ultimately to a probabilistic setting, while the other starts probabilistically and can be eventually placed in a measure-theoretic framework.

The measure-theoretic approach is considered first. Let $G^{n,m}$ denote the set of all finite \mathbb{R}^m -valued functions in \mathbb{R}^n , $m, n \in \mathbb{N}$, and $\mathbb{G}^{n,m}$ the σ -algebra containing all sets of the form $\left\{ g \in G^{n,m} : g(t) \in \bigcup_{i=1}^k B_i, t \in \mathbb{R}^n \right\}$ where $k > 0$ is an arbitrary integer and B_i are half-open intervals in \mathbb{R}^m . Then, similar to the definition of random variables, a *stochastic field* can be defined as a measurable mapping from (Ω, \mathbb{F}) into $(G^{n,m}, \mathbb{G}^{n,m})$. The notation $X(t, \omega)$ is used to denote an \mathbb{R}^m -valued n -dimensional stochastic field, where $X \in \mathbb{R}^m$, $t \in \mathbb{R}^n$ and $\omega \in \Omega$. As long as there is no danger of confusion, t , ω , or both could be suppressed.

Let t_1, t_2, \dots, t_k be k arbitrary points in \mathbb{R}^n . Then, $X(t_1), X(t_2), \dots, X(t_k)$ are a sequence of \mathbb{R}^m -valued random variables. Given the existence of the probability measure P on \mathbb{F} , the following probability measure

$$F_{(t_1, t_2, \dots, t_k)}(B) \triangleq P\{(X(t_1), X(t_2), \dots, X(t_k)) \in B\} \quad (2.71)$$

is well defined for every Borel set $B \in \mathbb{B}^{km}$; or equivalently, $F_{(t_1, t_2, \dots, t_k)}(x_1, x_2, \dots, x_k)$, the joint distribution function of $X(t_1), X(t_2), \dots, X(t_k)$, is well defined by the probability measure P on \mathbb{F} . The collection of these joint distribution functions $\left\{ F_{(t_1, t_2, \dots, t_k)}(x_1, x_2, \dots, x_k) : t_1, t_2, \dots, t_k \in \mathbb{R}^n, k \in \mathbb{N} \right\}$ is known as the family of *finite-dimensional distributions* for the stochastic field $X(t, \omega)$. In general, it is the finite-dimensional distributions that are dealt with in the study of a stochastic field. Hence, it is desirable that the P measure of all Borel sets in the σ -algebra $\mathbb{G}^{n,m}$ could also be uniquely defined by the finite-dimensional distributions. In order to make this property hold, it is necessary to assume some other condition such as separability of $X(t, \omega)$, which shall become clear later in this subsection.

The second definition of an \mathbb{R}^m -valued n -dimensional *stochastic field*, which might be more natural from a modelling viewpoint, is to consider it as a collection of random variables $\{X(t) \in \mathbb{R}^m : t \in \mathbb{R}^n\}$ together with a collection of joint distribution

functions $\left\{F_{(t_1, t_2, \dots, t_k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) : t_1, t_2, \dots, t_k \in \mathbb{R}^n, k \in \mathbb{N}\right\}$. A natural question to ask is whether or not one can always find a stochastic field $X(t, \omega)$, according to the first definition, which possesses these $F_{(t_1, t_2, \dots, t_k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ as its finite-dimensional distributions. This is answered by the following theorem.

Theorem 2.21 (Kolmogorov's Consistency Principle) For any $k \in \mathbb{N}$ and any points $t_1, t_2, \dots, t_k \in \mathbb{R}^n$, let $F_{(t_1, t_2, \dots, t_k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{R}^m$, denote a joint distribution function. Then, there exist a probability space (Ω, \mathbb{F}, P) and a corresponding stochastic field $X(t, \omega)$ ($X \in \mathbb{R}^m$, $t \in \mathbb{R}^n$ and $\omega \in \Omega$) such that the finite-dimensional distribution of $X(t, \omega)$ with respect to t_1, t_2, \dots, t_k coincides with $F_{(t_1, t_2, \dots, t_k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$, if and only if

I) For any permutation $(\lambda_1, \lambda_2, \dots, \lambda_k)$ of $(1, 2, \dots, k)$, the relation

$$F_{(t_1, t_2, \dots, t_k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = F_{(t_{\lambda_1}, t_{\lambda_2}, \dots, t_{\lambda_k})}(\mathbf{x}_{\lambda_1}, \mathbf{x}_{\lambda_2}, \dots, \mathbf{x}_{\lambda_k}) \quad (2.72)$$

holds for any $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{R}^m$.

II) Writing $F_{(t_1, t_2, \dots, t_k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ as a probability measure $F_{(t_1, t_2, \dots, t_k)}$, the relation

$$F_{(t_1, t_2, \dots, t_{k+j})}(B \times \mathbb{B}^{jm}) = F_{(t_1, t_2, \dots, t_k)}(B) \quad (2.73)$$

holds for every $j, k \in \mathbb{N}$ and $B \in \mathbb{B}^{km}$.

It is clear that not every family of joint distribution functions (or equivalently probability measures) corresponds to a stochastic field. However, only those families that do are generally interesting to researchers.

In probability theory, two random variables X and Y are called equivalent if $P\{X=Y\}=1$. This implies that, for all intents and purposes, these two variables are indistinguishable. Similarly, two stochastic fields $X(t)$ and $Y(t)$ are said to be *equivalent*, if

$$P\{X(t)=Y(t)\}=1 \quad \text{for every } t \in \mathbb{R}^n. \quad (2.74)$$

Two equivalent stochastic fields generate equivalent probability measures on $\mathbb{G}^{n,m}$ but

are not necessarily indistinguishable in every sense. For example, let $n = m = 1$, $\Omega = [0,1]$, and define the probability as

$$P\{B\} = \int_{B \cap [0,1]} d\lambda(x) \quad \text{for every Borel set } B \in \mathbb{B} \quad (2.75)$$

where λ is the Lebesgue measure on \mathbb{R} . Consider the following two stochastic fields,

$$X(t, \omega) = 0 \quad \text{for all } t \text{ and } \omega, \quad (2.76a)$$

$$Y(t, \omega) = \begin{cases} 0 & t \neq \omega \\ 1 & t = \omega \end{cases}. \quad (2.76b)$$

Then, X and Y are clearly equivalent, but

$$P\{X(t) \text{ is continuous for } t \in [0,1]\} = 1, \quad (2.77a)$$

$$P\{Y(t) \text{ is continuous for } t \in [0,1]\} = 0. \quad (2.77b)$$

Thus, certain probabilistic properties of these two stochastic fields are quite different.

For a given $\omega_0 \in \Omega$, $X(t, \omega_0)$ is simply a deterministic \mathbb{R}^m -valued function on \mathbb{R}^n , which is termed a *realization* (or a *sample path*) of the stochastic field X . Thus the difference between the equivalent stochastic processes X and Y defined in (2.73) could be said to be a difference in *sample path behaviour*. It infers that the behaviour of a sample path $X(t, \omega_0)$ is not necessary determined by the finite-dimensional distributions of the stochastic field $X(t, \omega)$. In order to overcome this uncomfortable situation, Doob [2.7] demonstrated that finite-dimensional distributions in fact do determine sample path properties by requiring that sample paths are essentially determined by their values on an everywhere dense, countable set of points in the parameter space \mathbb{R}^n . Specifically, a stochastic field $X(t, \omega)$, $t \in \mathbb{R}^n$, is said to be *separable* if there exists a countable set $D \subset \mathbb{R}^n$ and a fixed event N with $P\{N\} = 0$, such that for any closed interval $B \in \mathbb{R}^m$ and open interval $I \subset \mathbb{R}^n$ the two sets

$$\{\omega : X(t, \omega) \in B, t \in I\} \quad \text{and} \quad \{\omega : X(t, \omega) \in B, t \in I \cap D\} \quad (2.78)$$

differ by a subset of N . It is easy to see from this definition that the process $Y(t, \omega)$ defined by (2.76b) is not separable. In the remainder of this thesis, it is assumed, without further statement, that all stochastic fields are separable. This is not a serious assumption, since it can be proven that to every stochastic field there corresponds an equivalent

separable stochastic field.

Let $X(t) = (X_1(t), X_2(t), \dots, X_m(t))$ denote an \mathbb{R}^m -valued n -dimensional stochastic field, the vector-valued function

$$\mu(t) = E(X(t)) \quad (2.79)$$

is called the *expectation function* of $X(t)$, and the matrix-valued function

$$R(s, t) = \text{Cov}(X(s), X(t)) = E\left(\left(X(s) - \mu(s)\right)^T \left(X(t) - \mu(t)\right)\right) \quad (2.80)$$

is called the *covariance function*. Another important and widely used concept is the so called *second-order stochastic field*. Specifically, $X(t)$ is said to be a second-order stochastic field if $\sum_{i=1}^m E(|X_i(t)|^2) < +\infty$ holds for all $t \in \mathbb{R}^n$.

The next two subsections, as indicated by their subtitles, introduce two special classes of stochastic fields, namely Gaussian fields and stationary stochastic fields, which will play an important role in Chapters 3 and 4.

2.2.2 Gaussian Fields

Denoted by $X \sim N(\mu, \sigma^2)$, a real-valued random variable X is said to be *Gaussian* (or *normally distributed*) if its mean $\mu = E(X)$ and variance $\sigma^2 = E(|X - \mu|^2)$ are both finite, and its distribution function is given by

$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy. \quad (2.81)$$

Hence, the probability density function of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (2.82)$$

and the characteristic function is

$$\varphi_X(t) = e^{\sqrt{-1}\mu t - \frac{1}{2}\sigma^2 t^2}. \quad (2.83)$$

The case $\mu = 0$, $\sigma^2 = 1$ is rather special, and is termed a *standard Gaussian/normal distribution*. The distribution function of a standard Gaussian random variable is usually

denoted by $\Phi(x)$ so that

$$\Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy. \quad (2.84)$$

Denoted by $X \sim N(\mu, \Sigma)$, an \mathbb{R}^m -valued random variable X is said to be (*multivariate*) *Gaussian* if its probability density function is given by

$$f_X(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m \det(\Sigma)}} e^{-\frac{1}{2}(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu)^T} \quad (2.85)$$

where $\mu = E(X)$ is the m -dimensional mean vector and $\Sigma = E((X - \mu)^T (X - \mu))$ is the $m \times m$ non-negative definite *covariance matrix*. Then, it is relatively straightforward to check from the definition that the corresponding characteristic function is

$$\varphi_X(\mathbf{t}) = e^{\sqrt{-1}\mu^T \mathbf{t} - \frac{1}{2}\mathbf{t}^T \Sigma \mathbf{t}}. \quad (2.86)$$

The class of Gaussian random variables has a number of advantageous properties. In particular, let $X = (X_1, X_2, \dots, X_m) \sim N(\mu, \Sigma)$, we have:

- I) Probabilistic properties of X are completely determined by its first- and second-order moments.
- II) X_1, X_2, \dots, X_m are mutually independent if, and only if, Σ is a diagonal matrix.
- III) Let A be a $m \times n$ matrix ($m \geq n \in \mathbb{N}$) with $\text{rank}(A) = n$, then

$$Y = XA \sim N(\mu A, A^T \Sigma A). \quad (2.87)$$

Due to the above properties, such notations as *Wiener chaos expansions*[†] and *Wick products*[†] (see e.g. [2.20, 2.30]), which are based on Gaussian random variables, are often put into a framework of Hilbert space. Specifically, a *Gaussian linear space* is a real linear space of random variables, defined on some probability space (Ω, \mathbb{F}, P) , such that each variable in the space is centred Gaussian. Obviously, a Gaussian linear space is a linear

[†] The Wiener chaos expansion (or equivalently polynomial chaos expansion) is the theoretical foundation of the polynomial chaos expansion method, a very popular SFEM formulation initiated by R.G. Ghanem and P.D. Spanos (see Table 1.1). However, the research discussed in this thesis follows a completely different approach and has nothing to do with Wiener chaos expansions. Hence, the introduction to the Wiener chaos expansion and the associated Wick product is not included in this subsection.

subspace of $L^2(\Omega, \mathbb{F}, P)$, therefore it shares the same norm and inner product of the L^2 -space. A *Gaussian Hilbert space* is a Gaussian linear space which is complete, i.e. a closed subspace of $L^2(\Omega, \mathbb{F}, P)$ consisting of centred Gaussian random variables.

We can now define a *Gaussian stochastic field* (or simply a *Gaussian field*) to be a stochastic field possessing finite-dimensional distributions all of which are multivariate Gaussian. It is then clear that, all the finite-dimensional distributions of an \mathbb{R}^m -valued n -dimensional Gaussian field $X(t)$, and hence the probability measures they induce on $\mathbb{G}^{n,m}$, are completely determined by the expectation function $E(X(t))$ and the covariance function $\text{Cov}(X(s), X(t))$.

2.2.3 Stationary Stochastic Fields

Although the major interest of this thesis lies in \mathbb{R}^m -valued stochastic fields, it is mathematically more convenient to consider the stationary stochastic field in the more general context of \mathbb{C}^m -valued fields. Hence, unless otherwise stated, it is assumed in this subsection that a stochastic field $X(t) = (X_1(t), X_2(t), \dots, X_m(t))$ always takes values in the complex space \mathbb{C}^m , while t lies in \mathbb{R}^n . It is also assumed $E\left(\sum_{i=1}^m |X_i(t)|^2\right) < +\infty$ for all $t \in \mathbb{R}^n$, i.e. $X(t)$ is a second-order stochastic field.

With these assumptions, the *expectation function* of $X(t)$ is defined as

$$\mu(t) = E(X(t)), \quad (2.88)$$

and it is a deterministic function from \mathbb{R}^n to \mathbb{C}^m . The *covariance function* $R(s, t): \mathbb{R}^{2n} \rightarrow \mathbb{C}^{m \times m}$ is defined by

$$R(s, t) = \text{Cov}(X(s), X(t)) = E\left(\left(X(s) - \mu(s)\right)^T \overline{\left(X(t) - \mu(t)\right)}\right) \quad (2.89)$$

where the bar denotes the complex conjugate. When $X(t)$ is \mathbb{R}^m valued, the complex conjugate has no effect, so that the above definition is equivalent to that given in (2.80) for \mathbb{R}^m -valued stochastic fields.

From (2.89), it is easy to see that the matrix-valued covariance function satisfies

$$\mathbf{R}(\mathbf{s}, \mathbf{t}) = \overline{\mathbf{R}(\mathbf{t}, \mathbf{s})} \quad (2.90)$$

for all $\mathbf{s}, \mathbf{t} \in \mathbb{R}^n$. Let $R_{ij}(\mathbf{s}, \mathbf{t})$, $i, j \in \{1, 2, \dots, m\}$, denote the entry located at the i -th row and the j -th column of matrix $\mathbf{R}(\mathbf{s}, \mathbf{t})$. Then, from (2.89) and due to the Cauchy-Schwartz inequality (see e.g. Theorem 2.3), $R_{ij}(\mathbf{s}, \mathbf{t})$ is finite for all $\mathbf{s}, \mathbf{t} \in \mathbb{R}^n$. Specifically,

$$\begin{aligned} R_{ij}(\mathbf{s}, \mathbf{t}) &= E\left(\left(X_i(\mathbf{s}) - \mu_i(\mathbf{s})\right)\overline{\left(X_j(\mathbf{t}) - \mu_j(\mathbf{t})\right)}\right) \\ &\leq E\left(\left|X_i(\mathbf{s}) - \mu_i(\mathbf{s})\right|^2\right)^{1/2} E\left(\left|X_j(\mathbf{t}) - \mu_j(\mathbf{t})\right|^2\right)^{1/2} \quad \forall \mathbf{s}, \mathbf{t} \in \mathbb{R}^n. \quad (2.91) \\ &< +\infty \end{aligned}$$

Furthermore, the diagonal entry $R_{ii}(\mathbf{s}, \mathbf{t})$ is a *non-negative definite* function on \mathbb{R}^{2n} .

That is, the Hermitian form $\sum_{j_1=1}^k \sum_{j_2=1}^k R_{ii}(\mathbf{t}_{j_1}, \mathbf{t}_{j_2}) z_{j_1} \overline{z_{j_2}}$ is always real and non-negative for any collection of points $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_k$ in \mathbb{R}^n , and any complex numbers z_1, z_2, \dots, z_k . This is because,

$$\begin{aligned} &\sum_{j_1=1}^k \sum_{j_2=1}^k R_{ii}(\mathbf{t}_{j_1}, \mathbf{t}_{j_2}) z_{j_1} \overline{z_{j_2}} \\ &= \sum_{j_1=1}^k \sum_{j_2=1}^k E\left(\left(X_i(\mathbf{t}_{j_1}) - \mu_i(\mathbf{t}_{j_1})\right)\overline{\left(X_i(\mathbf{t}_{j_2}) - \mu_i(\mathbf{t}_{j_2})\right)}\right) z_{j_1} \overline{z_{j_2}} \quad \forall k \in \mathbb{N}. \quad (2.92) \\ &= E\left(\left|\sum_{j=1}^k z_j \left(X_i(\mathbf{t}_j) - \mu_i(\mathbf{t}_j)\right)\right|^2\right) \geq 0 \end{aligned}$$

The property of non-negative definiteness actually characterizes covariance functions, so that given any function $\mu(\mathbf{t}): \mathbb{R}^n \rightarrow \mathbb{C}$ and a non-negative definite $R(\mathbf{s}, \mathbf{t}): \mathbb{R}^{2n} \rightarrow \mathbb{C}$ it is always possible to construct a complex-valued n -dimensional stochastic field for which $\mu(\mathbf{t})$ and $R(\mathbf{s}, \mathbf{t})$ are the expectation and covariance functions, respectively. When $\mu(\mathbf{t})$ and $R(\mathbf{s}, \mathbf{t})$ are both real valued, this field can even be taken to be Gaussian, since the multivariate Gaussian finite-dimensional distributions generated by $\mu(\mathbf{t})$ and $R(\mathbf{s}, \mathbf{t})$ are easily seen to satisfy the Kolmogorov consistency principle (Theorem 2.21).

A \mathbb{C}^m -valued n -dimensional stochastic field $X(\mathbf{t})$ is said to be *strictly stationary* if its finite-dimensional distributions are invariant under translations in the

parameter \mathbf{t} . That is, for any set of points $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_k$ and $\boldsymbol{\tau}$ in \mathbb{R}^n , the joint distribution of the random variables $X(\mathbf{t}_1), X(\mathbf{t}_2), \dots, X(\mathbf{t}_k)$ should be the same as for the variables $X(\mathbf{t}_1 + \boldsymbol{\tau}), X(\mathbf{t}_2 + \boldsymbol{\tau}), \dots, X(\mathbf{t}_k + \boldsymbol{\tau})$. Two immediate consequences of this property are: a) the expectation function $\boldsymbol{\mu}(\mathbf{t})$ is identically equal to a constant; b) the covariance function $\mathbf{R}(\mathbf{s}, \mathbf{t})$ is forced to be a function of the difference $\mathbf{s} - \mathbf{t}$ only.

Very often, when dealing with stochastic fields, it is not necessary to impose the rather restrictive condition of strictly stationarity, and it is sufficient to demand only those two consequences just noted. Hence, the question of weak stationarity arises. A \mathbb{C}^m -valued n -dimensional stochastic field $X(\mathbf{t})$ is called *wide-sense stationary* (or *weakly stationary*), if

$$E(X(\mathbf{t})) = \boldsymbol{\mu} \text{ is a constant,} \quad (2.93)$$

and

$$\text{Cov}(X(\mathbf{s}), X(\mathbf{t})) = \mathbf{R}(\mathbf{s} - \mathbf{t}) \text{ is a function of } \mathbf{s} - \mathbf{t} \text{ only.} \quad (2.94)$$

A strictly stationary stochastic field is clearly wide-sense stationary, but in general the reverse is not true. However, it is true for \mathbb{R}^m -valued Gaussian fields.

An important special class of wide-sense stationary stochastic fields is the so called *isotropic stochastic field*. Specifically, a \mathbb{C}^m -valued n -dimensional wide-sense stationary stochastic field $X(\mathbf{t})$ is called *isotropic*, if its covariance function $\mathbf{R}(\mathbf{s} - \mathbf{t})$ is a function of $\|\mathbf{s} - \mathbf{t}\|_2$ only; i.e. for any two points $\mathbf{s}, \mathbf{t} \in \mathbb{R}^n$, the covariance of $X(\mathbf{s})$ and $X(\mathbf{t})$ depends only on the Euclidian distance between \mathbf{s} and \mathbf{t} .

Finally, it is highlighted that the theory of stationary stochastic field is quite rich and powerful, of which detailed expositions can be found in [2.17-2.19].

2.3 On Stochastic Analysis

Stochastic analysis is the generalization of the classic calculus to the random context, and it studies differentiation and integration of stochastic fields. Similar to the

structure of calculus, stochastic analysis may be divided into two parts, namely the SODE theory and the SPDE theory, that deal with one-dimensional stochastic processes and higher-dimensional stochastic fields, respectively. Hence, this section is organized as follows. Section 2.3.1 introduces two fundamental concepts in stochastic analysis, i.e. white noise and Brownian motion. Then, Sections 2.3.2 and 2.3.3 bring up respectively SODEs and SPDEs in a way that leads to the specific modelling strategy which will be introduced in Chapter 3.

2.3.1 White Noise and Brownian Motion

Let $X(t)$ denote an \mathbb{R}^n -valued n -dimensional wide-sense stationary stochastic field. Then, $X(t)$ is called *white noise* if $E(X(t)) = \mathbf{0}$, $\forall t \in \mathbb{R}^n$, and random variables $X(t_1), X(t_2), \dots, X(t_k)$ are independent and identically distributed for any set of points t_1, t_2, \dots, t_k in \mathbb{R}^n . In particular, when white noise $X(t)$ is a Gaussian field, it is termed *Gaussian/normal white noise*. It can be proven that the white noise field is discontinuous almost everywhere.

In 1828 the Scottish botanist Brown observed that pollen grains suspended in liquid performed an irregular motion. The motion was later explained by the random collisions with the molecules of the liquid. To describe the motion mathematically it is natural to use the concept of a stochastic process $B(t, \omega)$, interpreted as the position at time t of the pollen grain ω . The solution of $B(t, \omega)$ is largely due to Einstein (1905) and Wiener (1923).

In mathematics, the term *Brownian motion* (or *Wiener process*) describes a real-valued stochastic field $B(t, \omega)$, $t \geq 0$, that satisfies:

- I) For every $\omega \in \Omega$, $B(t, \omega) \in C^0[0, +\infty)$; i.e. $B(t, \omega)$, as a function of t , is continuous in $[0, +\infty)$.
- II) $B(0, \omega) = 0$ and

$$P\{B(t_i, \omega) - B(t_{i-1}, \omega) \leq x_i, i = 1, 2, \dots, k\} = \prod_{i=1}^k \Phi\left(\frac{x_i}{\sqrt{t_i - t_{i-1}}}\right) \quad (2.95)$$

holds for any $k \in \mathbb{N}$, any $0 = t_0 < t_1 < \dots < t_k < +\infty$, and any $x_1, x_2, \dots, x_k \in \mathbb{R}$.

It can be shown that the Brownian motion defined above represents a coordinate of $B(t, \omega)$, the position of pollen particles in liquid.

An equivalent mathematical definition of the Brownian motion $B(t, \omega)$, $t \geq 0$, is:

- I) For every $\omega \in \Omega$, $B(t, \omega) \in C^0[0, +\infty)$.
- II) For any $k \in \mathbb{N}$ and any $t_1, t_2, \dots, t_k \in [0, +\infty)$, random variables $B(t_1, \omega)$, $B(t_2, \omega), \dots, B(t_k, \omega)$ follow a multivariate Gaussian distribution.
- III) For any $t \in [0, +\infty)$

$$E(B(t, \omega)) = 0, \quad (2.96)$$

and for any $s, t \in [0, +\infty)$

$$E(B(s, \omega)B(t, \omega)) = \min(s, t). \quad (2.97)$$

2.3.2 Stochastic Ordinary Differential Equations

A general SODE has the form

$$\frac{dX}{dt} = u(t, X) + v(t, X) \cdot r(t, X, \omega) \quad t \in \mathbb{R}, \quad (2.98)$$

where X is the random unknown and functions u , v and r are all given; specifically, $u(t, X)$ and $v(t, X)$ are formally deterministic, and $r(t, X, \omega)$ is a random function. It is very difficult to develop a versatile SODE theory capable to deal with a general random term $r(t, X, \omega)$ in the above equation. A useful and comprehensive theory is, however, achieved with respect to a special class of random functions $r(t, X, \omega)$, i.e. white noise.

In this special case, Eq. (2.98) is usually written as

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot w_t \quad t \geq 0, \quad (2.99)$$

where $X_t = X(t, \omega)$ is the random unknown, $b(t, X_t)$ and $\sigma(t, X_t)$ are formally deterministic functions, and w_t represents a Gaussian white noise process.

As a white noise process is discontinuous almost everywhere, it is inconvenient to

directly deal with Eq. (2.99). Hence, its integral form

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s, \quad (2.100)$$

is considered instead, in which $B_s = B(s, \omega)$ denotes the Brownian motion. Note that the third term $\int_0^t \sigma(s, X_s) dB_s$ in the right-hand side of Eq. (2.100) could not be defined in the Riemann-Stieltjes sense because the variation of the Brownian motion is too large. Hence, a new integral formulation, namely *Itô integral*, is developed for $\int_0^t \sigma(s, X_s) dB_s$.

In mathematics, the theory developed from (2.100) is often termed *Itô calculus*, and it has become a mature and powerful tool in many applications containing large random jumps with time evolution. Although the Itô calculus considers only a special case of Eq. (2.98), the theory and its applications have been so successful that, the term SODE (or simply SDE) actually infers the Itô integral unless otherwise stated. A good introduction to Itô calculus is provided in [2.26] and more mathematically-rigorous treatments can be found in [2.21-2.25].

2.3.3 Stochastic Partial Differential Equations

Due to mathematical difficulties, the successful SODE theory (i.e. Itô calculus) is achieved for the special equation (2.99) driven by white noise rather than the general equation (2.98) containing an arbitrary random function. It can be naturally expected that, the challenge to develop a comprehensive theory for general SPDEs could be even more significant. Hence, instead of targeting general SPDEs, the SPDE research community have been mainly engaged in extending Itô's white noise model to higher-dimensional stochastic fields. As the concept of white noise holds in any dimensionality, Eq. (2.99), which in the SODE theory is usually regarded as a shorthand version of the formal integral equation (2.100), is often taken as the starting point in SPDE research (see e.g. [2.29-2.30]). However, this generalization is not straightforward.

For one-dimensional stochastic processes, there is a natural "total order" (i.e. the sequential order of real numbers) that represents the evolution of the process. The Itô integral is based on the concept of Brownian motion, and its development significantly benefits from the martingale theory and the theory of Markov processes, which are

essentially associated with the topological feature of total order.

Unfortunately, there is no such structure of total order in higher-dimensional Euclidean space. Consequently, neither the martingale theory nor the theory of Markov processes can be easily generalized to high-dimensional stochastic fields. This fundamental topological difference also complicates the generalization of Brownian motion to higher-dimensional cases, e.g. two-dimensional Brownian sheet. Specifically, the two *equivalent* definitions of Brownian motion addressed in Section 2.3.1 lead to completely *different* concepts of Brownian sheet.

In summary, current mathematics has not been able to provide an effective SPDE theory for the study of complicated random medium problems.

References

- [2.1] <http://www-history.mcs.st-andrews.ac.uk/BiogIndex.html>.
- [2.2] A.N. Kolmogorov, *Foundations of the Theory of Probability*, Springer, Berlin, 1933. (In German)
- [2.3] A. Einstein, On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat, *Annalen der Physik*, 17 (1905) 549-560. (In German)
- [2.4] N. Wiener, Differential space, *Journal of Mathematical Physics*, 2 (1923) 131-174.
- [2.5] A.N. Kolmogorov, On analytic methods in probability theory, *Mathematische Annalen*, 104 (1931) 415-458. (In German)
- [2.6] A.Y. Khinchin, Correlation theory of stationary stochastic processes, *Mathematische Annalen*, 109 (1934) 604-615. (In German)
- [2.7] J.L. Doob, *Stochastic Processes*, Wiley, New York, 1953.
- [2.8] P. Samuelson (1970) and R. Merton and M. Scholes (1997), awarded *Nobel Prizes* for their applications and developments of the SODE theory in economics, <http://nobelprize.org/>.
- [2.9] R.G. Wang, *An Introduction to Probability Theory*, Peking University Press, Beijing, 1994. (In Chinese)

- [2.10] S.H. Cheng, *Advanced Probability Theory*, Peking University Press, Beijing, 1996. (In Chinese)
- [2.11] M. Loève, *Probability Theory*, 4th Edition, Vols. I and II, Springer-Verlag, New York, 1977.
- [2.12] O. Kallenberg, *Foundations of Modern Probability*, Springer, New York, 1997.
- [2.13] J. Galambos, *Advanced Probability Theory*, Marcel Dekker Inc., New York, 1988.
- [2.14] K.L. Chung, *A Course in Probability Theory*, 2nd Edition, Academic Press, New York, 2000.
- [2.15] M.P. Qian and G.L. Gong, *Applied Stochastic Processes*, Peking University Press, Beijing, 1998. (In Chinese)
- [2.16] M.P. Qian and G.L. Gong, *Theory of Stochastic Processes*, 2nd Edition, Peking University Press, Beijing, 2000. (In Chinese)
- [2.17] A.M. Yaglom and R.A. Silverman, *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1962.
- [2.18] H. Cramér and M.R. Leadbetter, *Stationary and Related Stochastic Processes*, John Wiley & Sons, New York, 1966.
- [2.19] R.J. Adler, *The Geometry of Random Fields*, John Wiley & Sons, New York, 1981.
- [2.20] S. Janson, *Gaussian Hilbert Spaces*, Cambridge University Press, Cambridge, UK, 1997.
- [2.21] G.L. Gong, *An Introduction to Stochastic Differential Equations*, 2nd Edition, Peking University Press, Beijing, 2000. (In Chinese)
- [2.22] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales – Volume 1 Foundations*, 2nd Edition, Cambridge University Press, Cambridge, UK, 2000.
- [2.23] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales – Volume 2 Itô Calculus*, John Wiley & Sons, New York, 1987.
- [2.24] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publishing Company, Amsterdam, 1981.
- [2.25] Z. Schuss, *Theory and Applications of Stochastic Differential Equations*, John Wiley & Sons, New York, 1980.
- [2.26] B.K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5th Edition, Springer, Berlin, 1998.
- [2.27] G.D. Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*,

Cambridge University Press, Cambridge, UK, 1992.

- [2.28] P. Malliavin, *Stochastic Analysis*, Springer, New York, 1997.
- [2.29] T.H. Hida, H.H. Kuo, J. Potthoff and L. Streit, *White Noise: An Infinite Dimensional Calculus*, Springer, New York, 1993.
- [2.30] H. Holden, B. Øksendal, J. Ubøe and T. Zhang, *Stochastic Partial Differential Equations: A Modelling, White Noise Functional Approach*, Birkhäuser Boston, Boston, 1996.
- [2.31] M.Q. Zhou, *Functions of Real Variables*, 2nd Edition, Peking University Press, Beijing, 1998. (In Chinese)
- [2.32] H.L. Royden, *Real Analysis*, 3rd Edition, Prentice Hall, Englewood Cliffs, New Jersey, 1988.
- [2.33] G.Q. Zhang and Y.Q. Lin, *Lecture Note on Functional Analysis*, Vol. I, Peking University Press, Beijing, 1998. (In Chinese)
- [2.34] G.Q. Zhang and M.Z. Guo, *Lecture Note on Functional Analysis*, Vol. II, Peking University Press, Beijing, 1998. (In Chinese)
- [2.35] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1991.

Chapter 3

Material Modelling – Elementary

Random Media

As in the simulation of deterministic engineering systems, one of the key issues in solving an engineering system composed of random media is to develop an appropriate material model for the random media under consideration. Furthermore, in order to achieve a general SFEM framework, the random material model needs to be: (a) simple and robust so that it can be readily adopted in different equation systems; (b) comprehensive and versatile so that it can be employed to model a variety of practical random media.

Following the technical review in Chapter 1 and from a mathematical viewpoint, it might be said that there has been at least three classes of approaches in random material modelling, which are based on the fuzzy set theory [3.1-3.5], the theory of random matrices [3.6-3.7] and the probability theory [3.8-3.35], respectively. Although the first two classes of models might be useful in some applications, it appears that the third class of models have been more widely accepted in the SFEM community. This is simply because the mathematical foundation of probability (see e.g. Chapter 2) is much more solid than those of fuzzy sets [3.36-3.37] and random matrices [3.38]. Nevertheless, in most SFEM formulations (e.g. [3.8-3.35]), the material models of random media are developed in an intuitive manner, which are neither based on theoretical deduction nor based on experimental induction. As a result, some of these material models are even

ambiguous in concept and contain self-contradiction[†].

Some basic questions one may ask about a random material model are: (a) How to define the random medium in term of mathematical concepts? (b) From the viewpoint of physics, why does the practical random medium follow the specific mathematical description? (c) What properties does the random medium have? It appears that these fundamental questions have not been systematically answered according to the current state of SFEM research. Hence, following a probabilistic approach and driven by physical requirements of practical random media, this chapter attempts to develop, in a more mathematically rigorous fashion, a robust and comprehensive random material model. It has not been possible to make any experimental investigation during the course of this thesis, and efforts are made to answer the above questions from theoretical deduction.

3.1 Problem Statements

In the classic theory of continuum mechanics, material parameters take fixed values and engineering systems are governed by various PDE systems determined by their physical nature. In order to perform a successful analysis for an engineering structure, all the material properties have to be known in advance, so that the structural responses are uniquely determined by specific boundary conditions. For instance, it is well known that the governing equations and associated boundary conditions for elastostatics are

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma}(\mathbf{x}) + \mathbf{f}(\mathbf{x}) = 0 \\ \boldsymbol{\varepsilon}(\mathbf{x}) = \frac{1}{2}(\mathbf{u}(\mathbf{x})\nabla + \nabla\mathbf{u}(\mathbf{x})) \\ \boldsymbol{\sigma}(\mathbf{x}) = \mathbf{C}(\mathbf{x}) : \boldsymbol{\varepsilon}(\mathbf{x}) \end{cases} \quad \begin{cases} \mathbf{u}(\mathbf{x}) = \tilde{\mathbf{u}}(\mathbf{x}) & \text{on } \partial D_u \\ \boldsymbol{\sigma}(\mathbf{x}) = \tilde{\boldsymbol{\sigma}}(\mathbf{x}) & \text{on } \partial D_\sigma \end{cases} \quad (3.1)$$

$\mathbf{x} \in D$

where $\boldsymbol{\sigma}(\mathbf{x})$ and $\tilde{\boldsymbol{\sigma}}(\mathbf{x})$ denote the stress tensors, $\boldsymbol{\varepsilon}(\mathbf{x})$ the strain tensor, $\mathbf{u}(\mathbf{x})$ and $\tilde{\mathbf{u}}(\mathbf{x})$ the displacement vectors, $\mathbf{f}(\mathbf{x})$ the load vector, and $\mathbf{C}(\mathbf{x})$ the elastic tensor; $D \subset \mathbb{R}^n$ is the material domain, ∂D_u the displacement boundary of D , and ∂D_σ the stress boundary of D .

[†] More details can be found in Section 3.3.

Many heterogeneous materials (e.g. concrete, composite materials with random inclusions, soil and rocks etc.) encountered in civil engineering, material science, geomechanics and hydrogeology etc. have structures with a non-deterministic disorder and are often termed as *random media* in engineering. Although engineering systems consisting of random media conform to the same principles of mechanics, it should be emphasized that the phenomena involved in random media are far more complex than those arising when homogeneous materials are present. This is not only because of the intricate stochastic nature of material properties, but also because the material structure of random media often contains discontinuities and multiple length scales. The resulting formidable mathematical and computational challenges cannot be met merely by employing larger computational platforms, but require fundamentally new mathematical insights and algorithmic developments. A better understanding and prediction of engineering systems consisting of random media requires a modelling approach based on a combination of probabilistic concepts with methods of mechanics.

In the past few decades [3.1-3.35], different mathematical models have been developed to describe the irregular variation of material properties through the random medium, and most of these material models are based on stochastic fields. That is, when random media are concerned, the elastic tensor \mathbf{C} in the simple elastostatic equations (3.1) is not only a function over the medium domain $D \subset \mathbb{R}^n$, but is also a function over a probability space (Ω, \mathbb{F}, P) such that

$$\mathbf{C} = \mathbf{C}(\mathbf{x}, \omega) \quad \mathbf{x} \in D, \omega \in \Omega. \quad (3.2)$$

A simple example of material properties of random media is shown in Figure 3.1, in which $g(\mathbf{x}) = g(\mathbf{x}, \omega_0)$ denotes a particular realization of the Young's modulus of a random medium plate.

In the stochastic-field model, material properties of random media are regarded as some random functions defined in Euclidean space instead of as a sequence of individual random variables. Due to the well established analysis tools in calculus, this field-based strategy makes it easier to deal with various complicated interactions through the random medium.

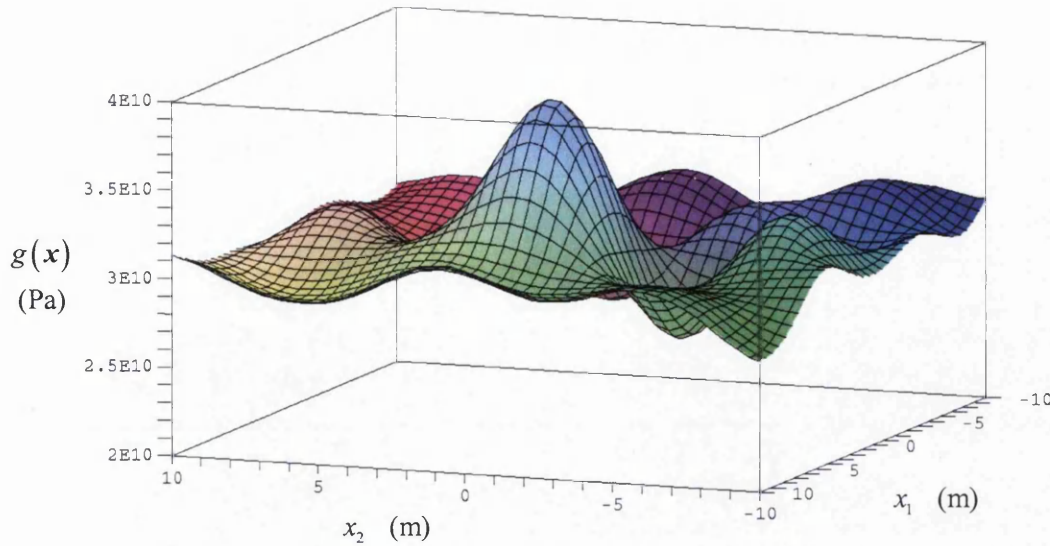


Figure 3.1 Young's modulus of a random medium plate

Consequently, for random medium structures, the governing equations and associated boundary conditions of elastostatics become a SPDE system as follows

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, \omega) + \mathbf{f}(\mathbf{x}, \omega) = 0 \\ \boldsymbol{\epsilon}(\mathbf{x}, \omega) = \frac{1}{2} (\mathbf{u}(\mathbf{x}, \omega) \nabla + \nabla \mathbf{u}(\mathbf{x}, \omega)) \\ \boldsymbol{\sigma}(\mathbf{x}, \omega) = \mathbf{C}(\mathbf{x}, \omega) : \boldsymbol{\epsilon}(\mathbf{x}, \omega) \end{cases} \quad \begin{cases} \mathbf{u}(\mathbf{x}, \omega) = \tilde{\mathbf{u}}(\mathbf{x}, \omega) & \text{on } \partial D_u \\ \boldsymbol{\sigma}(\mathbf{x}, \omega) = \tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega) & \text{on } \partial D_\sigma \end{cases} \quad (3.3)$$

$\mathbf{x} \in D, \omega \in \Omega$

in which all the mathematical quantities are stochastic fields and share the same physical meanings as their deterministic counterparts in (3.1). Except for the possible random boundary conditions $\tilde{\mathbf{u}}(\mathbf{x}, \omega)$ and $\tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega)$, which can be treated separately, the randomness in the above SPDEs is introduced by the spatially irregular variation of material properties. Therefore, the probability distributions of the stochastic fields $\mathbf{C}(\mathbf{x}, \omega)$ and $\mathbf{f}(\mathbf{x}, \omega)$ determine the probabilistic properties of the random solutions $\mathbf{u}(\mathbf{x}, \omega)$, $\boldsymbol{\epsilon}(\mathbf{x}, \omega)$ and $\boldsymbol{\sigma}(\mathbf{x}, \omega)$. This relationship reflects the stochastic mechanism of many typical random medium systems regardless of their physical nature. Hence, in solving engineering systems with variable uncertainties, e.g. Eq. (3.3), it is of crucial importance to identify and analyse the common features of the stochastic fields describing random material properties, e.g. $\mathbf{C}(\mathbf{x}, \omega)$ and $\mathbf{f}(\mathbf{x}, \omega)$.

Remark: As stated in Section 2.1.4, there are four basic modes of stochastic

convergence, which lead to different modes of stochastic continuity and stochastic differentiability. Hence, an essential question arising from the SPDE system (3.3) is that, in which sense are the associated stochastic fields including $C(\mathbf{x}, \omega)$, $\mathbf{u}(\mathbf{x}, \omega)$, $\varepsilon(\mathbf{x}, \omega)$ and $\sigma(\mathbf{x}, \omega)$ differentiable? Naturally, an appropriate mode of stochastic continuity and stochastic differentiability has to satisfy the requirements both in engineering and in mathematics. For practical random media, this question will be rigorously answered later in this chapter. For the moment, it is assumed that all the stochastic fields involved are path-wise differentiable, which corresponds to almost sure convergence, so that Eq. (3.3) could at least be valid in mathematics.

In the remainder of this chapter, a specific stochastic-field model will be developed to define material properties of practical random media. In order to simplify the notation, the discussions will mainly focus on scalar fields. However, the resulting concepts and principles generally hold for tensor fields without changes. Unless otherwise stated, all the random variables take real values.

3.2 From White Noise to Elementary Random Media

In order to develop a mathematically rigorous material model for random media, this section starts with the popular white noise model. Although the white noise approach does not succeed in the end, it helps to identify the engineering requirements and the probabilistic essence of practical random media, which in turn leads to the definition of a comprehensive mathematical model, namely elementary random media.

3.2.1 The Integral Functional of White Noise

It is reviewed in Section 2.3 that, one of the most important concepts in stochastic analysis [3.39-3.42] is white noise. The white noise model not only plays a fundamental role in Itô calculus [3.39], but is also exceedingly popular in the current research of SPDEs [3.40]. Due to the available mathematical tools in stochastic analysis as well as the extremely irregular variation of material properties through a random medium, an easy way, perhaps also the most mathematically attractive approach, to tackle practical random

media is by referring to white noise. However, it can be proven that a white noise field is discontinuous almost everywhere (see e.g. [3.39] or Section 3.3.2). This hampers the direct application of the white noise model in physical systems that usually have a continuity nature at the macro scale.

From a mathematical point of view, a natural way to bypass the above obstacle of discontinuity is to consider, instead, the integral functional of white noise. Indeed, the integral strategy has been adopted in the development of the SODE theory [3.39] and in the attempt of extending the SODE theory to the SPDE case [3.40]. The integral functional approach can be regarded as an average solution, in which the original field at any fixed point is approximated by its average value in the neighbourhood of the fixed point. From the point of view of physics, this is equivalent to investigating the white noise field in a larger length scale. The advantages of this strategy based on integral functionals are obvious:

- I) With certain local averaging techniques, a better continuity/differentiability in a larger scale may be achieved for the white noise model, which in turn makes the existing PDE tools applicable to this study.
- II) The “average” solution could be expected to converge to the “exact” solution when the selected view-window tends to a single view-point.

Let $w(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{R}^n$, $\omega \in \Omega$ denote a real-valued n -dimensional white noise field. That is, for any two points $\mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n$, the associated random variables $w(\mathbf{x}_1, \omega)$ and $w(\mathbf{x}_2, \omega)$ are independent and identically distributed. Without loss of generality, the expectation function of $w(\mathbf{x}, \omega)$ is

$$E(w(\mathbf{x}, \omega)) = 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad (3.4)$$

and the covariance function of $w(\mathbf{x}, \omega)$ is

$$\text{Cov}(w(\mathbf{x}_1, \omega), w(\mathbf{x}_2, \omega)) = \begin{cases} 0 & \mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n \\ \sigma^2 & \mathbf{x}_1 = \mathbf{x}_2 \in \mathbb{R}^n \end{cases} \quad (3.5)$$

where σ^2 (see Section 3.4.4.2 for more information) is the variance of the white noise $w(\mathbf{x}, \omega)$.

If $w(\mathbf{x}, \omega)$ is measurable[†] in $\mathbb{R}^n \times \Omega$, its locally averaging field can be written, in a general form, as the integral functional

$$W(\mathbf{x}, \omega) = (w * k)(\mathbf{x}, \omega) = \int_{\mathbb{R}^n} w(\mathbf{y}, \omega) k(\mathbf{x} - \mathbf{y}) d\mathbf{y} \quad (3.6)$$

where $*$ denotes the convolution operator, and the convolution kernel $k(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$, serves as a weight function to control the effective averaging domain, i.e. the view-window. Let

$$\delta(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \neq \mathbf{0} \\ +\infty & \mathbf{x} = \mathbf{0} \end{cases} \quad \int_{\mathbb{R}^n} \delta(\mathbf{x}) d\mathbf{x} = 1 \quad (3.7)$$

denote the Dirac delta function in \mathbb{R}^n . It is then evident that, when the view-window in (3.6) degenerates into a single view-point, i.e.

$$k(\mathbf{x}) = \delta(\mathbf{x}), \quad (3.8)$$

the locally averaging field will become identical with the original white noise field, i.e.

$$W(\mathbf{x}, \omega) = w(\mathbf{x}, \omega). \quad (3.9)$$

This verifies the second intuitive expectation of the integral functional of white noise. However, to verify the first intuitive expectation of the white noise functional, i.e. the smoothness of $W(\mathbf{x}, \omega)$, it is necessary to introduce the following lemma about the convolution in the right-hand side of (3.6).

Given a non-negative integer m , let $C^m(\mathbb{R}^n)$ denote the set of all functions defined on \mathbb{R}^n that have up to the m -th order continuous derivatives. Furthermore, let $C_0^m(\mathbb{R}^n)$ denote all the functions in $C^m(\mathbb{R}^n)$ that have compact support (the support of function $f(\mathbf{x})$ $\mathbf{x} \in \mathbb{R}^n$ is the closed set of $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \neq 0\}$). When there is no danger of confusion, symbols $C^m(\mathbb{R}^n)$ and $C_0^m(\mathbb{R}^n)$ can be abbreviated as C^m and C_0^m , respectively. For example, C^∞ denote all the infinitely differentiable functions on \mathbb{R}^n , and C_0^∞ denote all the functions in C^∞ that have compact support.

[†] In fact, it will be proven in Section 3.4.4.2 that there does not exist a measurable white noise field in the normal sense. Nevertheless, the formalistic derivation addressed here is helpful for disclosing the probabilistic essence of practical random media, which in turn leads to the definition of the so called elementary random media. This logic will eventually become clear at the end of this chapter.

Lemma 3.1 Let $1 \leq p \leq +\infty$ denote a real number and $m \geq 0$ an integer. If $f(\mathbf{x}) \in L^p(\mathbb{R}^n)$ and $g(\mathbf{x}) \in C_0^m(\mathbb{R}^n)$, then $f * g \in C^m(\mathbb{R}^n)$ and

$$\left(\frac{\partial^{\alpha_1 + \dots + \alpha_n}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} (f * g) \right) (\mathbf{x}) = \left(f * \frac{\partial^{\alpha_1 + \dots + \alpha_n}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} g \right) (\mathbf{x}) \quad (3.10)$$

where α_k ($k=1, 2, \dots, n$) are non-negative integers that satisfy $\sum_{i=1}^n \alpha_i \leq m$.

Proof:

First, it is proven that, if the function $g(\mathbf{x})$ is continuous in \mathbb{R}^n and with compact support, then $f * g$ is continuous. Consider the absolute difference

$$\begin{aligned} & |(f * g)(\mathbf{x} + \Delta \mathbf{x}) - (f * g)(\mathbf{x})| \\ &= \left| \int_{\mathbb{R}^n} f(\mathbf{y}) g(\mathbf{x} + \Delta \mathbf{x} - \mathbf{y}) d\mathbf{y} - \int_{\mathbb{R}^n} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) d\mathbf{y} \right| \\ &= \left| \int_{\mathbb{R}^n} f(\mathbf{x} - \mathbf{t}) [g(\mathbf{t} + \Delta \mathbf{x}) - g(\mathbf{t})] d\mathbf{t} \right| = \Delta I \geq 0 \end{aligned} \quad (3.11)$$

in three cases corresponding respectively to $1 < p < +\infty$, $p = 1$ and $p = +\infty$:

I) If $1 < p < +\infty$, then there exists $1 < q < +\infty$ such that $1/p + 1/q = 1$. According to Hölder's inequality (see e.g. the probabilistic version in Section 2.1.3 or the general form in [3.43-3.44]), we have

$$\Delta I \leq \|f\|_p \left(\int_{\mathbb{R}^n} |g(\mathbf{t} + \Delta \mathbf{x}) - g(\mathbf{t})|^q d\mathbf{t} \right)^{1/q}. \quad (3.12)$$

As $g(\mathbf{x})$ is continuous and with compact support, it is uniformly continuous.

Therefore,

$$\left(\int_{\mathbb{R}^n} |g(\mathbf{t} + \Delta \mathbf{x}) - g(\mathbf{t})|^q d\mathbf{t} \right)^{1/q} \rightarrow 0 \quad \text{as } \Delta \mathbf{x} \rightarrow \mathbf{0}. \quad (3.13)$$

II) If $p = 1$, according to the dominated convergence theorem (see e.g. the probabilistic version in Section 2.1.4 or the general form in [3.43-3.44]), we have

$$\Delta I \leq \int_{\mathbb{R}^n} |f(\mathbf{x} - \mathbf{t}) [g(\mathbf{t} + \Delta \mathbf{x}) - g(\mathbf{t})]| d\mathbf{t} \rightarrow 0 \quad \text{as } \Delta \mathbf{x} \rightarrow \mathbf{0}. \quad (3.14)$$

III) If $p = +\infty$, we have

$$\Delta I \leq \|f\|_\infty \int_{\mathbb{R}^n} |g(\mathbf{t} + \Delta \mathbf{x}) - g(\mathbf{t})| d\mathbf{t} \rightarrow 0 \quad \text{as } \Delta \mathbf{x} \rightarrow \mathbf{0}. \quad (3.15)$$

Summarizing the above three cases, it is concluded that $\Delta I \rightarrow 0$ as $\Delta \mathbf{x} \rightarrow \mathbf{0}$. Hence, $f * g$ is continuous.

Secondly, consider $g \in C_0^m$ where $m \geq 1$. For a fixed integer j , $1 \leq j \leq n$, let

$$\Delta \mathbf{x} = \left(\underbrace{0, \dots, 0}_{1 \dots j-1}, \Delta x_j, \underbrace{0, \dots, 0}_{j+1 \dots n} \right) \quad \Delta x_j \neq 0. \quad (3.16)$$

Then, according to the mean value theorem, we have

$$\begin{aligned} & \frac{(f * g)(\mathbf{x} + \Delta \mathbf{x}) - (f * g)(\mathbf{x})}{\Delta x_j} - \left(f * \frac{\partial g}{\partial x_j} \right)(\mathbf{x}) \\ &= \int_{\mathbb{R}^n} f(\mathbf{t}) \left[\frac{g(\mathbf{x} - \mathbf{t} + \Delta \mathbf{x}) - g(\mathbf{x} - \mathbf{t})}{\Delta x_j} \right] d\mathbf{t} - \left(f * \frac{\partial g}{\partial x_j} \right)(\mathbf{x}) \\ &= \int_{\mathbb{R}^n} f(\mathbf{t}) \left[\frac{\partial g}{\partial x_j}(\mathbf{x} - \mathbf{t} + \boldsymbol{\xi}) - \frac{\partial g}{\partial x_j}(\mathbf{x} - \mathbf{t}) \right] d\mathbf{t} \end{aligned} \quad (3.17)$$

in which

$$\boldsymbol{\xi} = \left(\underbrace{0, \dots, 0}_{1 \dots j-1}, \xi_j, \underbrace{0, \dots, 0}_{j+1 \dots n} \right) \quad 0 \leq |\xi_j| \leq |\Delta x_j|. \quad (3.18)$$

Since $g(\mathbf{x})$ is continuously differentiable and has compact support, following a similar analysis process as from (3.11) to (3.15), it can be verified that the right-hand side of (3.17)

converges to zero as $\Delta x_j \rightarrow 0$. That is, $\frac{\partial}{\partial x_j}(f * g)(\mathbf{x})$ exists and equals $\left(f * \frac{\partial g}{\partial x_j} \right)(\mathbf{x})$.

Note from the first part of this proof that $f * g$ is continuous, it can be concluded that the theorem is true for $m = 1$. When $m \geq 2$, the theorem can be similarly proven by repeatedly applying the above proof procedure. \square

Consequently, according to Lemma 3.1, the path-wise differentiability of $w(\mathbf{x}, \omega)$ can be improved to any order by the integral functional (3.6), and the differentiability of $W(\mathbf{x}, \omega)$ depends on the smoothness of the weight function $k(\mathbf{x})$.

With the above discussions, it appears that, by choosing an appropriate kernel $k(\mathbf{x})$ to control the view-window in accordance with practical requirements of

engineering systems with variable uncertainties, the white noise functional $W(\mathbf{x}, \omega)$ may be able to provide an effective material model for random media. It will be very interesting if this is true since a large portion of existing mathematical results in stochastic analysis will then become available to this study. On the other hand, an ideal stochastic-field model of random media should not only be smooth to meet, at the macro scale, the continuity/differentiability requirements of a general physical system, but also be sufficiently flexible, in the sense of probability, to capture the stochastic nature of random media in various situations. This is the other aspect that needs to be checked in order to evaluate the potential of the white noise approach in random medium modelling.

3.2.2 Probabilistic Failings of the White Noise Functional

In accordance with Eqs. (3.4) and (3.6), the expectation function of $W(\mathbf{x}, \omega)$ can be readily obtained as

$$E(W(\mathbf{x}, \omega)) = \int_{\mathbb{R}^n} E(w(\mathbf{y}, \omega))k(\mathbf{x} - \mathbf{y})d\mathbf{y} = 0 \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (3.19)$$

Moreover, following Eqs. (3.4 – 3.6), the covariance function of $W(\mathbf{x}, \omega)$ is

$$\begin{aligned} & \text{Cov}(W(\mathbf{x}_1, \omega), W(\mathbf{x}_2, \omega)) \\ &= E\left(\left(W(\mathbf{x}_1, \omega) - E(W(\mathbf{x}_1, \omega))\right)\left(W(\mathbf{x}_2, \omega) - E(W(\mathbf{x}_2, \omega))\right)\right) \\ &= E\left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} w(\mathbf{y}_1, \omega)w(\mathbf{y}_2, \omega)k(\mathbf{x}_1 - \mathbf{y}_1)k(\mathbf{x}_2 - \mathbf{y}_2)d\mathbf{y}_1d\mathbf{y}_2\right) \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n. \quad (3.20) \\ &= \int_{\mathbb{R}^{2n}} \text{Cov}(w(\mathbf{y}_1, \omega), w(\mathbf{y}_2, \omega))k(\mathbf{x}_1 - \mathbf{y}_1)k(\mathbf{x}_2 - \mathbf{y}_2)d\mathbf{y}_1d\mathbf{y}_2 \\ &= \int_{\{(\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{R}^{2n} : \mathbf{y}_1 = \mathbf{y}_2\}} \sigma^2 k(\mathbf{x}_1 - \mathbf{y}_1)k(\mathbf{x}_2 - \mathbf{y}_2)d\mathbf{y}_1d\mathbf{y}_2 \end{aligned}$$

Note that $\{(\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{R}^{2n} : \mathbf{y}_1 = \mathbf{y}_2\}$ is a set of measure zero (see Section 3.4.4.2 for more information) in \mathbb{R}^{2n} , therefore the right-hand side of the above equation equals zero, i.e.

$$\text{Cov}(W(\mathbf{x}_1, \omega), W(\mathbf{x}_2, \omega)) = 0 \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n. \quad (3.21)$$

Thus, the first- and second- order statistical moments of the white noise functional $W(\mathbf{x}, \omega)$ are both zero. Indeed, much more than this is true. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_j$ be a set of fixed points in \mathbb{R}^n . Then, we have

$$\begin{aligned}
 & E\left(W(\mathbf{x}_1, \omega)W(\mathbf{x}_2, \omega)\cdots W(\mathbf{x}_j, \omega)\right) \\
 &= E\left(\int_{\mathbb{R}^n} (w(y_1, \omega)\cdots w(y_j, \omega))(k(\mathbf{x}_1 - y_1)\cdots k(\mathbf{x}_j - y_j)) dy_1 \cdots dy_j\right), \quad (3.22) \\
 &= \int_{\mathbb{R}^n} M(y_1, \dots, y_j)(k(\mathbf{x}_1 - y_1)\cdots k(\mathbf{x}_j - y_j)) dy_1 \cdots dy_j
 \end{aligned}$$

in which

$$M(y_1, \dots, y_j) = E(w(y_1, \omega)\cdots w(y_j, \omega)) \quad (3.23)$$

is a function defined on \mathbb{R}^{jn} . Following the independent assumption of white noise,

$$M(y_1, \dots, y_j) = \begin{cases} E(w^j(y_1)) & y_1 = y_2 = \cdots = y_j \\ 0 & \text{otherwise} \end{cases}. \quad (3.24)$$

Thus, $M(y_1, \dots, y_j)$ is identically equal to zero on \mathbb{R}^{jn} except for a set of measure zero.

Hence,

$$E(W(\mathbf{x}_1, \omega)W(\mathbf{x}_2, \omega)\cdots W(\mathbf{x}_j, \omega)) = 0 \quad \forall j \in \mathbb{N}, \forall \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_j \in \mathbb{R}^n. \quad (3.25)$$

That is, the j -th order statistical moment of $W(\mathbf{x}, \omega)$ is zero.

For a fixed point $\mathbf{x}_0 \in \mathbb{R}^n$, let $W(\mathbf{x}_0, \omega)$ be the corresponding random variable taken from the white noise functional $W(\mathbf{x}, \omega)$. It is well known that the probabilistic property of the random variable $W(\mathbf{x}_0, \omega)$ is completely determined by its characteristic function $\varphi(\mathbf{x}_0, t)$, which can be expanded into a Taylor series (see e.g. Section 2.1.2)

$$\varphi(\mathbf{x}_0, t) = \sum_{i=0}^{+\infty} E(W^i(\mathbf{x}_0, \omega)) \frac{(\sqrt{-1}t)^i}{i!} \quad (3.26)$$

where

$$E(W^0(\mathbf{x}_0, \omega)) \equiv 1. \quad (3.27)$$

Following immediately from (3.25), we have

$$E(W^i(\mathbf{x}_0, \omega)) = 0 \quad \forall i \in \mathbb{N}. \quad (3.28)$$

Substituting (3.27) and (3.28) into (3.26) yields

$$\varphi(\mathbf{x}_0, t) \equiv 1 \quad \forall t \in \mathbb{R}. \quad (3.29)$$

Note $E(W(\mathbf{x}_0, \omega)) = 0$, hence the above characteristic function infers that the random variable $W(\mathbf{x}_0, \omega)$ degenerates to a constant zero. As \mathbf{x}_0 is arbitrary, the stochastic field $W(\mathbf{x}, \omega)$ degenerates to a constant deterministic field, i.e.

$$W(\mathbf{x}, \omega) \equiv 0 \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (3.30)$$

After all the discussions in Section 3.2.1, the above result is a bit surprising since following Eq. (3.30) there are only two possibilities:

- I) The white noise approach is basically unsuitable for random medium modelling.
- II) It is sufficient to model random media, at least some types of random media, with the conventional deterministic approach.

From a physical/engineering point of view, when a random medium is being investigated from different length scales, the stochastic nature of the random medium should not change dramatically and the behaviours of the random medium system must remain compatible. Hence, if a white noise model in a way reflects the stochastic nature of practical random media, the results obtained from investigating the white noise field in a larger length scale should be similar to those obtained from the investigation based on a smaller length scale, in particular when the two length scales are of little difference. However, it is noted that the specific shape (e.g. narrow or wide) of the view-window, which is controlled by the weight function $k(\mathbf{x})$ in the local average (3.6), has no effect in the above analysis from (3.19) to (3.30). That is, no matter how small the observing scale (or equivalently the locally averaging domain) is, the corresponding white noise functional $W(\mathbf{x}, \omega)$ always completely loses the randomness of the white noise field $w(\mathbf{x}, \omega)$. This contradicts the common sense engineering requirement. Consequently, although “mathematically sound” [3.39-3.42], the starting point of the white noise approach is not quite right for representing material properties of random media.

Proposition 3.2: Due to the lack of sufficient capability in capturing the stochastic nature of practical random media, the white noise functional $W(\mathbf{x}, \omega)$ is inadequate for modelling random media in engineering systems. In particular, from a physical point of view, the white noise model is equivalent to a deterministic model in a practical engineering system with variable uncertainties.

3.2.3 Probabilistic Essence of Random Media

The failings of the white noise approach indicate that, at least one of the two basic assumptions of white noise, i.e. independent and identical distribution, contradicts the stochastic nature of random media in practical engineering systems.

In order to locate the defect of the white noise model and disclose the probabilistic essence of practical random media, it is suggested to consider a special stochastic field $b(i, \omega)$ $i \in \mathbb{N}, \omega \in \Omega$ with $E(b(i, \omega)) = 0 \quad \forall i \in \mathbb{N}$. The stochastic field $b(i, \omega)$ is essentially a sequence of centred random variables. According to Kolmogorov's strong law of large numbers (see e.g. Section 2.1.5), if random variables $b(i, \omega)$ are independent and satisfy

$$\sum_{i=1}^{+\infty} \frac{\text{Var}(b(i, \omega))}{i^2} < +\infty, \quad (3.31)$$

then the following limit relation holds

$$P\left(\lim_{k \rightarrow +\infty} B_k(i, \omega) = 0\right) = 1 \quad \forall i \in \mathbb{N}, \quad (3.32)$$

in which

$$B_k(i, \omega) = \frac{1}{k} \sum_{j=1}^k b(i+j, \omega) \quad (3.33)$$

denotes the local average of $b(i, \omega)$. That is, under the technical condition (3.31), the local average of the independent random variables $b(i, \omega)$ converges almost surely to a constant zero as the upper bound k approaches infinity. Comparing the local average $B_k(i, \omega)$ with the white noise functional $W(x, \omega)$, it is observed that the crucial common factor that results in the loss of randomness in these two cases is the “independent assumption”. Hence, it may be said that the “independent distribution” assumption of $w(x, \omega)$ contradicts the stochastic nature of practical random media, which in turn makes the white noise approach unsuitable for this study. Consequently, the following insight is taken as the basis for random medium modelling in this thesis.

Proposition 3.3: The stochastic dependence among different points through the medium is an essential characteristic of many practical random media in physical systems.

3.2.4 Elementary Random Media

It is implied in Proposition 3.3 that an effective random material model has to be based on a dependent stochastic field, which, as a mathematical concept, is much too general, on its own, and is difficult to be tackled in a quantitative manner. Hence, in order to make the problem treatable in mathematics, additional assumptions must be made regarding the dependent stochastic field according to specific features of the random media under consideration.

Naturally, as addressed in Section 3.2.1, a key issue that needs to be taken into account in random medium modelling is the continuity and differentiability of the stochastic field involved, since it will determine whether or not, and furthermore in which sense, the SPDE (3.3) holds. In addition, due to the insufficiency of experimental data, only the first- and second- order statistical moments of random material properties are available in many practical engineering problems. Hence, it will be beneficial for applications if the random material model does not require higher order statistical moments for input settings.

Driven by the above modelling requirements and based on the existing mathematical tools reviewed in Chapter 2, a random material model is defined as follows.

Definition 3.4 (Elementary random media) Let $m \geq 0 \in \mathbb{Z}$. Random medium $D \subset \mathbb{R}^n$ is termed a C^m elementary random medium (ERM), if its material property is represented by a second-order stochastic field $a(\mathbf{x}, \omega)$ $\mathbf{x} \in D, \omega \in \Omega$ that satisfies:

- I) The expectation function of $a(\mathbf{x}, \omega)$ has up to the m -th order continuous derivatives in D , i.e.

$$E(a(\mathbf{x}, \omega)) = a_o(\mathbf{x}) \quad (3.34)$$

where $a_o(\mathbf{x}) \in C^m(D)$.

- II) The covariance function of $a(\mathbf{x}, \omega)$ can be written as

$$\text{Cov}(a(\mathbf{x}_1, \omega), a(\mathbf{x}_2, \omega)) \equiv R(\mathbf{x}_1 - \mathbf{x}_2) = R(\boldsymbol{\tau}) \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D \quad (3.35)$$

such that

$$\int_{\mathbb{R}^n} R(\boldsymbol{\tau}) d\boldsymbol{\tau} = \int_{\mathbb{R}^n} |R(\boldsymbol{\tau})| d\boldsymbol{\tau} < +\infty. \quad (3.36)$$

III) The covariance function $R(\boldsymbol{\tau})$ is C^{2m} continuous at $\boldsymbol{\tau} = \mathbf{0}$, i.e. it has up to the $2m$ -th order continuous derivatives at the origin.

IV) With respect to functions $a_o(\mathbf{x})$ and $R(\boldsymbol{\tau})$, the following relation holds

$$R(\mathbf{0}) \leq \kappa |a_o(\mathbf{x})| \quad \forall \mathbf{x} \in D \quad (3.37)$$

where $\kappa > 0$ is a fixed constant depending on the physical nature of $a(\mathbf{x}, \omega)$.

V) The stochastic field $a(\mathbf{x}, \omega)$ is a Gaussian field.

In particular, the random medium D defined above is referred to as a *generalized ERM* if the condition V) is not required.

Remark: For the sake of simplicity, the material property of ERM is represented by the scalar $a(\mathbf{x}, \omega)$ in the above definition. However, the ERM model holds for the general case in which the material properties are represented by a general elastic tensor $C(\mathbf{x}, \omega)$ and conditions I)-V) are accordingly defined with respect to each scalar entry of $C(\mathbf{x}, \omega)$.

With respect to the above mathematical definition, the following engineering insights of ERM are highlighted:

- I) The material property of an ERM is defined by its expectation function and covariance function. The finite assumption of these two statistical moments is purely technical, which is always met in practical engineering problems.
- II) The spatial stochastic dependence of an ERM is described by the associated covariance function $R(\boldsymbol{\tau}) = R(\mathbf{x}_1 - \mathbf{x}_2)$, which implies that the pair-wise dependence is invariant under translations in parameter \mathbf{x} . Furthermore, as shown in (3.36), the pair-wise dependence has a short-range nature such that $R(\boldsymbol{\tau})$ vanishes as $\|\boldsymbol{\tau}\|_2$, i.e. the distance between two points, approaches infinity.
- III) As the definition of ERM is based on second order stochastic fields, the mean square convergence is taken as the mathematical foundation to define the continuity/differentiability in the SPDE system with respect to ERM. It should be noted that the mean square continuity (or differentiability) does not imply the path-wise continuity (or differentiability). Hence, the material property of a

particular ERM sample may not be smooth. However, the corresponding SPDEs hold in the context of probability, and specifically in the mean square sense, all the standard differential/integral operators in calculus can be similarly performed.

- IV) In order to ensure the mean square continuity/differentiability of the stochastic field $a(\mathbf{x}, \omega)$, it is required that its statistical moments $a_o(\mathbf{x})$ and $R(\boldsymbol{\tau})$ satisfy certain smoothness conditions. The underline mathematics will be addressed in detail in Section 3.3.
- V) In the ERM model, the material property at any individual point $\mathbf{x}_0 \in D$ is described by a random variable $a(\mathbf{x}_0, \omega)$, which is determined by its first two statistical moments $a_o(\mathbf{x}_0)$ and $R(\mathbf{0})$. Note in continuum mechanics that the value of a specific material parameter is often restricted in a certain range determined by its physical nature. Hence, in order to satisfy, in a probabilistic sense, the physical requirement on the value range of the material property concerned, the factor κ is employed in (3.37) to control the variation of $a(\mathbf{x}, \omega)$.
- VI) Determined by the central limit theorem (see e.g. Section 2.1.5), a Gaussian distribution is often encountered in practical engineering problems when the observation regards the total effect of a large number of small random factors. Consequently, it is expected that Gaussian fields provide a reasonably good approximation of many practical random media. In addition, a Gaussian field is completely defined by its first- and second- order statistical moments, and due to its linearity, it can be put into the context of Hilbert spaces (see e.g. Section 2.2.2). Hence, having “Gaussian field” as a basic assumption of ERM makes it possible to achieve explicit solutions in some typical random medium problems. However, in a serious engineering problem, whether or not the specific random material property can be treated as a Gaussian field must be examined against experimental data. In fact, the Gaussian assumption of ERM is not essential, which will become clear later in this thesis.

3.3 Macro-Scale Properties of Elementary Random Media

In the last section, the random material model ERM is defined in terms of mathematical concepts. This section explores its properties including the stationarity, the smoothness and principles for material measurements. These macro-scale properties will have strong influences on the theoretical and algorithmic developments of ERM as well as the associated applications in engineering practice.

Following Definition 3.4, let

$$a(\mathbf{x}, \omega) = a_o(\mathbf{x}) + a_s(\mathbf{x}, \omega) \quad (3.38)$$

where the mean $a_o(\mathbf{x})$ represents the deterministic component of $a(\mathbf{x}, \omega)$, and the fluctuation field $a_s(\mathbf{x}, \omega)$ the random component. It is then straightforward to obtain from (3.34) and (3.35) that

$$E(a_s(\mathbf{x}, \omega)) = 0 \quad \forall \mathbf{x} \in D, \quad (3.39)$$

$$\begin{aligned} & \text{Cov}(a_s(\mathbf{x}_1, \omega), a_s(\mathbf{x}_2, \omega)) \\ &= \text{Cov}(a(\mathbf{x}_1, \omega), a(\mathbf{x}_2, \omega)) \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D. \\ &= R(\boldsymbol{\tau}) \end{aligned} \quad (3.40)$$

That is, the fluctuation field $a_s(\mathbf{x}, \omega)$ has mean zero and the same covariance function as the original stochastic field $a(\mathbf{x}, \omega)$. As it is usually simpler to study a centred stochastic field, the discussion in this section will mainly focus on the properties of $a_s(\mathbf{x}, \omega)$. However, all the results obtained for $a_s(\mathbf{x}, \omega)$ can be readily translated into the conclusions corresponding to $a(\mathbf{x}, \omega)$, which only differs by a smooth deterministic function $a_o(\mathbf{x})$.

3.3.1 Stationarity

It is shown in (3.39, 3.40) that both the expectation function and the covariance function of $a_s(\mathbf{x}, \omega)$ are invariant under coordinate translations. Thus, the fluctuation

field $a_s(\mathbf{x}, \omega)$ is wide-sense stationary. As $a(\mathbf{x}, \omega)$ is a Gaussian field for ERM, it can be further concluded that $a_s(\mathbf{x}, \omega)$ is a strictly stationary stochastic field whose finite dimensional distributions are invariant under coordinate translations. The wide-sense stationarity plays an important role in many aspects of ERM. It will soon become clear that in many ways, the theoretical/algorithmic developments of ERM in this thesis is about wide-sense stationary stochastic fields (see e.g. Section 2.2.3).

3.3.2 Continuity and Differentiability

As pointed out in the beginning of this chapter, the smoothness of stochastic fields is an essential issue in random medium modelling. However, it appears that this fundamental aspect has not received adequate attention in the development of SFEMs, since some elementary and fatal inconsistencies have occurred in the literature. For example, isotropic stochastic fields with such covariance functions as $c_1(1 - c_2\|\boldsymbol{\tau}\|_2)$ and $c_1e^{-\|\boldsymbol{\tau}\|_2/c_2}$ have been employed in many SFEM formulations [3.13, 3.15, 3.19-3.24, 3.26, 3.28, 3.32] to describe random material properties, some of which are noticed in the recent work [3.29]. These stochastic fields often lead to simple theoretical treatments, and sequentially provide efficient numerical algorithms. However, it can be proven that these stochastic fields are non-differentiable almost everywhere and inevitably result in defective, if not meaningless, SPDE systems of random media.

In this thesis, the continuity and differentiability regarding ERM are defined based on the mean-square convergence. As addressed in Section 3.2.4, the mean-square continuity/differentiability does not necessarily imply path-wise smoothness, and theoretically speaking, this provides the ERM model some freedom to accommodate the discontinuities in practical random medium samples. More importantly, in the mean square sense, the SPDE system can be similarly treated with traditional PDE tools.

3.3.2.1 Mean-Square Smoothness of Stochastic Fields

The mathematical research regarding mean-square smoothness of stochastic fields can be traced back over half a century ago to Khinchin [3.45], Kolmogorov [3.46] and Wiener [3.47]. The continuity/differentiability of one-dimensional wide-sense stationary

stochastic processes was first systematically and rigorously addressed by Yaglom [3.48], and higher dimensional cases can be similarly treated.

Theorem 3.5 A centred second-order stochastic field $b(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$ is continuous in the mean-square sense at $\mathbf{x} = \mathbf{x}_0$, if and only if the covariance function $\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)) = R(\mathbf{x}_1, \mathbf{x}_2)$ is continuous at the point $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_0, \mathbf{x}_0)$.

Proof:

I) The “if” part: This follows immediately from the relation

$$\begin{aligned} & E\left(\left|b(\mathbf{x}_0 + \Delta\mathbf{x}, \omega) - b(\mathbf{x}_0, \omega)\right|^2\right) \\ &= R(\mathbf{x}_0 + \Delta\mathbf{x}, \mathbf{x}_0 + \Delta\mathbf{x}) - R(\mathbf{x}_0 + \Delta\mathbf{x}, \mathbf{x}_0) - R(\mathbf{x}_0, \mathbf{x}_0 + \Delta\mathbf{x}) + R(\mathbf{x}_0, \mathbf{x}_0) \end{aligned} \quad (3.41)$$

where $\Delta\mathbf{x} \in \mathbb{R}^n$.

II) The “only if” part: Let $\Delta\mathbf{x}_1, \Delta\mathbf{x}_2 \in \mathbb{R}^n$, consider the following absolute difference

$$\begin{aligned} & \left| R(\mathbf{x}_0 + \Delta\mathbf{x}_1, \mathbf{x}_0 + \Delta\mathbf{x}_2) - R(\mathbf{x}_0, \mathbf{x}_0) \right| \\ &= \left| E\left(\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)\right)\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)\right)\right) + \right. \\ & \quad \left. E\left(\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)\right)b(\mathbf{x}_0, \omega)\right) + \right. \\ & \quad \left. E\left(\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)\right)b(\mathbf{x}_0, \omega)\right) \right| \\ &\leq E\left(\left|\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)\right)\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)\right)\right|\right) + \\ & \quad E\left(\left|\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)\right)b(\mathbf{x}_0, \omega)\right|\right) + \\ & \quad E\left(\left|\left(b(\mathbf{x}_0 + \Delta\mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)\right)b(\mathbf{x}_0, \omega)\right|\right) \\ &= \Delta I \end{aligned} \quad (3.42)$$

According to Hölder’s inequality (see e.g. Section 2.1.3), we have

$$\begin{aligned} \Delta I &\leq \left(E\left(\left|b(\mathbf{x}_0 + \Delta\mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)\right|^2\right)\right)^{1/2} \left(E\left(\left|b(\mathbf{x}_0 + \Delta\mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)\right|^2\right)\right)^{1/2} + \\ & \quad \left(E\left(\left|b(\mathbf{x}_0 + \Delta\mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)\right|^2\right)\right)^{1/2} \left(E\left(\left|b(\mathbf{x}_0, \omega)\right|^2\right)\right)^{1/2} + \\ & \quad \left(E\left(\left|b(\mathbf{x}_0 + \Delta\mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)\right|^2\right)\right)^{1/2} \left(E\left(\left|b(\mathbf{x}_0, \omega)\right|^2\right)\right)^{1/2} \end{aligned} \quad (3.43)$$

Hence,

$$\left| R(\mathbf{x}_0 + \Delta\mathbf{x}_1, \mathbf{x}_0 + \Delta\mathbf{x}_2) - R(\mathbf{x}_0, \mathbf{x}_0) \right| \rightarrow 0 \quad \text{as } \Delta\mathbf{x}_1, \Delta\mathbf{x}_2 \rightarrow \mathbf{0}. \quad (3.44)$$

The theorem is proven. \square

A trivial application of the above theorem is to analyse the continuity of the white noise field $w(\mathbf{x}, \omega)$ in Section 3.2.1. From (3.4, 3.5), it is straightforward to conclude that $w(\mathbf{x}, \omega)$ is discontinuous everywhere.

As the point \mathbf{x}_0 in Theorem 3.5 is arbitrary, the following theorem about “everywhere continuity” holds.

Theorem 3.6 A centred second-order stochastic field $b(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$ is everywhere continuous in mean square, if and only if its covariance function $R(\mathbf{x}_1, \mathbf{x}_2)$ is continuous at every diagonal point $\mathbf{x}_1 = \mathbf{x}_2$.

Before moving to the differentiability of stochastic fields, it is necessary to introduce the following convergence criterion of random variables.

Lemma 3.7 (due to Loève) A sequence of random variables X_1, X_2, \dots converges to X in mean square, if and only if

$$\lim_{i, j \rightarrow +\infty} E(X_i X_j) = c \quad (3.45)$$

where c is a finite constant.

Proof:

I) The “if” part is due to

$$\begin{aligned} & \lim_{i, j \rightarrow +\infty} E(|X_i - X_j|^2) \\ &= \lim_{i, j \rightarrow +\infty} (E(X_i^2) - 2E(X_i X_j) + E(X_j^2)) \\ &= c - 2c + c \\ &= 0 \end{aligned} \quad (3.46)$$

II) The “only if” part follows immediately

$$E(X_i X_j) \rightarrow E(X^2) = c \quad \text{as } i, j \rightarrow +\infty. \quad (3.47)$$

The lemma is proven. \square

With respect to the point-wise mean-square differentiability of stochastic fields, there is the following theorem.

Theorem 3.8 A centred second-order stochastic field $b(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$ is differentiable in mean square at $\mathbf{x} = \mathbf{x}_0$, if and only if its covariance function $R(\mathbf{x}_1, \mathbf{x}_2)$ has finite second-order derivatives $\frac{\partial^2 R(\mathbf{x}_1, \mathbf{x}_2)}{\partial x_{1i} \partial x_{2i}}$ $i = 1, 2, \dots, n$ at the point $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_0, \mathbf{x}_0)$.

Proof:

For a fixed i , $1 \leq i \leq n$, let

$$\Delta \mathbf{x} = \left(\underbrace{0, \dots, 0}_{1 \dots i-1}, \Delta x_i, \underbrace{0, \dots, 0}_{i+1 \dots n} \right) \quad \Delta x_i \neq 0. \quad (3.48)$$

Thus, with respect to the i -th coordinate of point \mathbf{x} , the mean-square partial derivative of $b(\mathbf{x}, \omega)$ is

$$\frac{\partial b(\mathbf{x}, \omega)}{\partial x_i} \triangleq \text{l.i.m.}_{\Delta x_i \rightarrow 0} \frac{b(\mathbf{x} + \Delta \mathbf{x}, \omega) - b(\mathbf{x}, \omega)}{\Delta x_i} \quad (3.49)$$

where l.i.m. denotes “limit in mean”. Let

$$\begin{cases} \Delta \mathbf{x}_1 = \left(\underbrace{0, \dots, 0}_{1 \dots i-1}, \Delta x_{1i}, \underbrace{0, \dots, 0}_{i+1 \dots n} \right) & \Delta x_{1i} \neq 0 \\ \Delta \mathbf{x}_2 = \left(\underbrace{0, \dots, 0}_{1 \dots i-1}, \Delta x_{2i}, \underbrace{0, \dots, 0}_{i+1 \dots n} \right) & \Delta x_{2i} \neq 0 \end{cases}, \quad (3.50)$$

we have

$$\begin{aligned} & E \left(\frac{b(\mathbf{x}_0 + \Delta \mathbf{x}_1, \omega) - b(\mathbf{x}_0, \omega)}{\Delta x_{1i}} \cdot \frac{b(\mathbf{x}_0 + \Delta \mathbf{x}_2, \omega) - b(\mathbf{x}_0, \omega)}{\Delta x_{2i}} \right) \\ &= \frac{R(\mathbf{x}_0 + \Delta \mathbf{x}_1, \mathbf{x}_0 + \Delta \mathbf{x}_2) - R(\mathbf{x}_0 + \Delta \mathbf{x}_1, \mathbf{x}_0) - R(\mathbf{x}_0, \mathbf{x}_0 + \Delta \mathbf{x}_2) + R(\mathbf{x}_0, \mathbf{x}_0)}{\Delta x_{1i} \Delta x_{2i}}. \end{aligned} \quad (3.51)$$

Hence, according to Lemma 3.7, the stochastic field $b(\mathbf{x}, \omega)$ has a mean-square partial derivative $\frac{\partial b(\mathbf{x}, \omega)}{\partial x_i}$ at $\mathbf{x} = \mathbf{x}_0$, if and only if the second derivative $\frac{\partial^2 R(\mathbf{x}_1, \mathbf{x}_2)}{\partial x_{1i} \partial x_{2i}}$ exists at the point $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_0, \mathbf{x}_0)$. Note the integer i is arbitrary in $\{1, 2, \dots, n\}$, the theorem is proven. \square

Remark: According to Theorem 3.8, isotropic stochastic fields with covariance functions of the forms

$$R(\mathbf{x}_1, \mathbf{x}_2) = c_1(1 - c_2 \|\mathbf{x}_1 - \mathbf{x}_2\|_2) \quad \text{and} \quad R(\mathbf{x}_1, \mathbf{x}_2) = c_1 e^{-\|\mathbf{x}_1 - \mathbf{x}_2\|_2 / c_2} \quad (3.52)$$

are non-differentiable everywhere. Hence, these stochastic fields are unsuitable for representing material properties in the SPDE system (3.3), in which the elastic tensor $C(\mathbf{x}, \omega)$ needs to be differentiable.

From Theorem 3.8, it is straightforward to obtain the following theorem about everywhere differentiability of stochastic fields.

Theorem 3.9 A centred second-order stochastic field $b(\mathbf{x}, \omega) \quad \mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$ is everywhere differentiable in mean square, if and only if its covariance function $R(\mathbf{x}_1, \mathbf{x}_2)$ has finite second order derivatives $\frac{\partial^2 R(\mathbf{x}_1, \mathbf{x}_2)}{\partial x_{1i} \partial x_{2i}} \quad i = 1, 2, \dots, n$ at every diagonal point $\mathbf{x}_1 = \mathbf{x}_2$.

Theorems 3.5-3.6 and 3.8-3.9 address the criteria of the continuity and the first-order differentiability of stochastic fields. In a similar way, sufficient and necessary conditions regarding the higher order differentiability of stochastic fields can be established. In particular, if the stochastic field in consideration is wide-sense stationary, it can be proven that the local properties (e.g. continuity and differentiability) of the stochastic field are completely determined by the behaviour of its covariance function in the neighbourhood of the origin. This is summarised by the following theorem.

Theorem 3.10 (Smoothness of wide-sense stationary stochastic fields) Let m denote a non-negative integer and $b(\mathbf{x}, \omega) \quad \mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$ a centred second-order stochastic field with covariance function $\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)) \equiv R(\mathbf{x}_1 - \mathbf{x}_2) = R(\boldsymbol{\tau})$.

The m -th order mean-square partial derivatives $\frac{\partial^m b(\mathbf{x}, \omega)}{\partial x_{i_1} \cdots \partial x_{i_m}} \quad 1 \leq i_1, \dots, i_m \leq n$ exist

everywhere, if and only if the $2m$ -th order partial derivatives $\frac{\partial^{2m} R(\boldsymbol{\tau})}{\partial^2 \tau_{i_1} \cdots \partial^2 \tau_{i_m}}$ exists at the

point $\boldsymbol{\tau} = \mathbf{0}$.

3.3.2.2 Mean-Square Smoothness of Elementary Random Media

Consider the stochastic field $a(\mathbf{x}, \omega)$ of a C^m ERM D . Following Definition 3.4, Eq. (3.38) and Theorem 3.10, it can be concluded that both $a_o(\mathbf{x})$ and $a_s(\mathbf{x}, \omega)$ have up to the m -th order continuous derivatives in mean square. Consequently, $a(\mathbf{x}, \omega)$ has the C^m mean-square continuity on D .

3.3.3 Principles of Elementary Random Medium

Measurements

The inputs of the ERM model are the expectation function $a_o(\mathbf{x})$ and the covariance function $R(\boldsymbol{\tau})$. In general, experimental measurements of a large number of material samples are required to construct these two statistical moments according to their definitions (3.34) and (3.35). In order to reduce the associated experimental costs and improve the applicability of the ERM model, it is important to be able to accurately construct $a_o(\mathbf{x})$ and $R(\boldsymbol{\tau})$ from as few material samples as possible. In many engineering problems, $a_o(\mathbf{x})$, the deterministic component of $a(\mathbf{x}, \omega)$, can be estimated in accordance with existing knowledge and measuring techniques of deterministic material models. Consequently, the major challenge in the random material measurement is to obtain $R(\boldsymbol{\tau})$, which is essentially determined by $a_s(\mathbf{x}, \omega)$, the random component of $a(\mathbf{x}, \omega)$. The solution of this problem is attributed to another property of ERM, i.e. *ergodicity*.

3.3.3.1 Ergodicity

The mathematical background of ergodicity is outlined below, and detailed explanations and proofs can be found in [3.49-3.52].

Recall from Section 2.2.1 that $X(t, \omega)$ denotes an \mathbb{R}^m -valued n -dimensional stochastic field, and $(G^{n,m}, \mathbb{G}^{n,m}, P)$ the corresponding probability space. That is, $G^{n,m}$ denotes the set of all finite \mathbb{R}^m -valued functions on \mathbb{R}^n , $\mathbb{G}^{n,m}$ denotes the σ -algebra

containing all sets of the form $\left\{ \mathbf{g} \in G^{n,m} : \mathbf{g}(\mathbf{t}) \in \bigcup_{i=1}^k B_i, \mathbf{t} \in \mathbb{R}^n \right\}$ where $k > 0$ is an arbitrary integer and B_i are half-open intervals in \mathbb{R}^m , and P is the probability measure uniquely defined on all sets of $G^{n,m}$ by the finite-dimensional distributions of $X(\mathbf{t}, \omega)$. With this choice of probability space, the basic events, ω , are now individual sample functions $\mathbf{g}(\mathbf{t}) \in \mathbb{R}^m$, $\mathbf{t} \in \mathbb{R}^n$. For a fixed vector $\boldsymbol{\tau} \in \mathbb{R}^n$, the *shift transformation* of $\mathbf{g}(\mathbf{t})$ is defined as

$$T_{\boldsymbol{\tau}} \circ \mathbf{g} \triangleq \mathbf{g}(\mathbf{t} + \boldsymbol{\tau}). \quad (3.53)$$

Similarly, each $T_{\boldsymbol{\tau}}$ takes any set $A \in G^{n,m}$ into the set

$$T_{\boldsymbol{\tau}} \circ A \triangleq \{ \mathbf{g}(\mathbf{t} + \boldsymbol{\tau}) : \mathbf{g}(\mathbf{t}) \in A \} \quad (3.54)$$

that consists of all the functions of A shifted by $\boldsymbol{\tau}$.

A set $A \in G^{n,m}$ is called an *invariant set* of $X(\mathbf{t}, \omega)$ if, for every $\boldsymbol{\tau}$, the sets A and $T_{\boldsymbol{\tau}} \circ A$ differ, at most, by a set of P -measure zero. That is, there exist two sets B_1 and B_2 , both of which have measure zero, for which

$$A \cup B_1 = (T_{\boldsymbol{\tau}} \circ A) \cup B_2. \quad (3.55)$$

Definition 3.11 (Ergodicity) An \mathbb{R}^m -valued n -dimensional strictly stationary stochastic field $X(\mathbf{t}, \omega)$ is said to be ergodic if the σ -algebra of invariants sets only contains sets of probability zero or one.

The importance of the notion of ergodicity lies in the following theorems.

Theorem 3.12 (Ergodic theorem) Let $X(\mathbf{t}, \omega) = (X_1(\mathbf{t}, \omega), \dots, X_m(\mathbf{t}, \omega))$ be an \mathbb{R}^m -valued n -dimensional, strictly stationary and ergodic stochastic field; let $S(r)$ denote the sphere in \mathbb{R}^n of radius $r > 0$, and $V(r)$ the volume of $S(r)$. If

$$E(|X_i(\mathbf{t}, \omega)|) < +\infty \quad i = 1, 2, \dots, m \quad (3.56)$$

and with probability one, the \mathbb{R}^m -valued Riemann integral

$$\int_{S(r)} X(\mathbf{t}, \omega) dt \quad (3.57)$$

exists for every $r > 0$, then

$$\frac{1}{V(r)} \int_{S(r)} X(t) dt \xrightarrow{\text{a.s.}} E(X(0)) \quad \text{as } r \rightarrow +\infty. \quad (3.58)$$

Theorem 3.13 Let $X(t, \omega)$ be an \mathbb{R}^m -valued n -dimensional stochastic field. If $X(t, \omega)$ is strictly stationary and ergodic, then the stochastic field generated by a shift transformation of $X(t, \omega)$ is also strictly stationary and ergodic.

In general, it is not easy to provide a simple condition which ensures ergodicity. For Gaussian fields, however, this can be done.

Theorem 3.14 Let $X(t, \omega)$ be a real-valued n -dimensional strictly stationary Gaussian field with covariance function $R(\tau)$. Then, $X(t, \omega)$ is ergodic if

$$R(\tau) \rightarrow 0 \quad \text{as } \|\tau\|_2 \rightarrow +\infty. \quad (3.59)$$

3.3.3.2 Measuring Covariance Functions of Elementary Random Media

With the above mathematical preparations, the ERM measurement problem can now be solved.

Let $R(\tau)$ denote the ERM covariance function to be statistically estimated. Following Eqs. (3.39-3.40) and at a fixed point $\tau_0 \in \mathbb{R}^n$, the value of $R(\tau)$ is

$$R(\tau_0) = E(\alpha(\omega)) \quad (3.60)$$

where

$$\alpha(\omega) = a_s(\tau_0, \omega) a_s(0, \omega). \quad (3.61)$$

That is, $R(\tau_0)$ is the expectation of the random variable $\alpha(\omega)$, which is defined on sample paths of the fluctuation field $a_s(x, \omega)$. By the definition (3.60), the estimation of $R(\tau_0)$ requires arithmetically averaging a number of samples of $\alpha(\omega)$, each of which corresponds to a different material sample of the random medium under consideration. However, measuring a large number of material samples is inconvenient, and sometimes even impossible, in engineering practice. It is therefore beneficial to be able to obtain

$R(\boldsymbol{\tau}_0)$ from a small number of material samples, and ideally from one sample.

As shown in Section 3.3.1, $a_s(\mathbf{x}, \omega)$ is a strictly stationary Gaussian field. Note in (3.36) that $R(\boldsymbol{\tau})$ has a short range nature such that $R(\boldsymbol{\tau}) \rightarrow 0$ as $\|\boldsymbol{\tau}\|_2 \rightarrow +\infty$. Hence, according to Theorem 3.14, $a_s(\mathbf{x}, \omega)$ is ergodic. Furthermore, according to Theorem 3.13, the following stochastic field

$$A(\mathbf{x}, \omega) \triangleq a_s(\boldsymbol{\tau}_0 + \mathbf{x}, \omega) a_s(\mathbf{x}, \omega) \quad \mathbf{x} \in \mathbb{R}^n \quad (3.62)$$

is strictly stationary and ergodic. Consequently, following the ergodic theorem 3.12, we have

$$\begin{aligned} R(\boldsymbol{\tau}_0) &= E(A(\omega)) = \lim_{r \rightarrow +\infty} \left(\frac{1}{V(r)} \int_{S(r)} A(\mathbf{x}, \omega) d\mathbf{x} \right) \\ &= \lim_{r \rightarrow +\infty} \left(\frac{1}{V(r)} \int_{S(r)} (a(\boldsymbol{\tau}_0 + \mathbf{x}, \omega) - a_o(\boldsymbol{\tau}_0 + \mathbf{x})) (a(\mathbf{x}, \omega) - a_o(\mathbf{x})) d\mathbf{x} \right) \end{aligned} \quad (3.63)$$

where $S(r)$ and $V(r)$ denote the n -sphere and its volume, respectively. Since there is no probability integral involved in the right-hand side of the above equation, $R(\boldsymbol{\tau}_0)$ can be accordingly obtained from the measurements of just one material sample. As the point $\boldsymbol{\tau}_0$ is arbitrary in (3.63), the covariance function $R(\boldsymbol{\tau})$ can be obtained from the measurements of one material sample.

3.4 Summary and Suggestions for Future Research

3.4.1 Summary

A comprehensive and robust random material model, namely elementary random media, is developed in this chapter. First, in the formalistic analysis of the white noise approach, it is observed that the assumption of independent distribution is the main deficiency that results in the failure of the white noise model. Therefore, the stochastic dependence is recognized as an essential characteristic of practical random media. Next, based on this initiative and driven by the modelling requirements both in engineering

practice and in mathematics, the ERM model is defined. Finally, a number of macro-scale properties of ERM are addressed, in which the stationarity is fundamental. Although there are five mathematical assumptions in the definition of ERM, the ERM model is not very restrictive for engineering applications. Indeed, most random media in the SFEM literature [3.1-3.35] could be described by the (generalized) ERM model.

It should be noted that the idea of using wide-sense stationary stochastic fields to describe random material properties is not new. Instead, the major contributions in this chapter are:

- The contradiction-analysis strategy of random medium modelling, which, as a whole, is original. The modelling strategy not only leads to the definition of ERM but also indicates a stochastic approach for multi-scale simulations. This will be explained in the next subsection.
- The systematic combination of random medium modelling and the theory of wide-sense stationary stochastic fields. In this process, a serious error regarding smoothness of stochastic fields, which has been occurring in a variety of SFEM formulations for two decades, is highlighted, and consequently the correct mode and corresponding conditions for the continuity and differentiability of practical random media are defined. More effort and consequently more benefits along this strategic route will be addressed in Chapter 4.
- The establishment of the ERM model, in which the random material properties are represented by a general elastic tensor. As the discussion and analysis in this chapter do not involve interactions between different random material parameters, they are demonstrated by virtue of scalar stochastic fields to simplify the notation. However, it is important in random medium modelling to take into account the possible interactions between different material properties, and the complete treatment of these random interactions is a fundamental feature of the ERM model, which will be explained in detail at the end of Chapter 4.

3.4.2 Suggestions for Future Research

3.4.2.1 *Making the Ergodicity a Basic Assumption of Elementary Random Media*

It is seen in Section 3.3.3 that the ergodicity plays a crucial role in reducing the experimental work in random medium measurements. This aspect is of fundamental importance in random medium modelling, since the significance of a random material model is limited if its input parameters can not be easily obtained from experimental measurements of practical random medium samples.

However, it is generally difficult to provide a simple condition to ensure the ergodicity. In the ERM model, the ergodicity is ensured by the properties of wide-sense stationarity and Gaussian distribution. These two conditions are sufficient but not necessary, and they do generate unnecessary restrictions for practical applications of the ERM model. An alternative choice is to remove the Gaussian assumption and make the ergodicity a basic assumption of ERM instead. Specifically, as shown in (3.63), a weak ergodicity regarding second-order statistical moments would be sufficient. This alternative definition will certainly improve the coverage of the ERM model. Nevertheless, it will be shown in Chapter 4 that the Gaussian field assumption makes it easier to obtain an explicit representation of ERM. This is the main reason that the Gaussian assumption is kept in Definition 3.4.

Hence, further investigations are required in order to achieve a more general ERM model. A key issue in developing the new ERM model is how to obtain the explicit representation of ERM by using the ergodic rather than the Gaussian assumption. This viewpoint will become clearer in Chapter 4.

3.4.2.2 *A Stochastic Approach for Multi-Scale Simulations*

At the macro scale, the classic continuum mechanics, which is purely deterministic, provides an effective mathematical description for the behaviours of many solid materials. At the micro scale, various uncertain factors inevitably exist in the material structure, whose behaviours are therefore better described by a stochastic theory of mechanics. Naturally, it can be expected that there exists a general framework of multi-scale

simulations, in which the macro-scale continuum mechanics serves as an average theory of the micro-scale stochastic mechanics. This intuitive viewpoint is widely accepted in the research of multi-scale simulations, and specifically the averaging technique is often termed *homogenization* in the literature.

As indicated by the question marks in Figure 3.2, there are two key issues in the development of the framework of multi-scale simulations that involve micro-scale uncertainties:

- I) Developing a stochastic mechanics theory to describe the behaviours of micro-scale materials.
- II) Constructing a bridge to connect the stochastic mechanics at the micro scale and the classic continuum mechanics at the macro scale.

Figure 3.2 also indicates a theoretical structure of multi-scale simulations, in which the framework is divided into two parts connected together by one bridge. This simple structure has been widely used in the current research of multi-scale simulations.

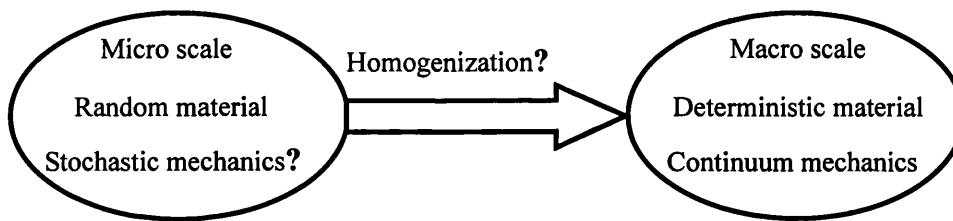


Figure 3.2 A simple framework of multi-scale simulations

The random-medium modelling discussions in this chapter infer a possibility to tackle the above two problems by describing the material properties at the micro scale with stochastic fields, as it is done in the ERM model, and describing the material properties at the macro scale with deterministic variables, as is done in classic continuum mechanics. It is then expected that the average effects of the stochastic-field material model could be approximated by the conventional deterministic material model. However, in general, the spatial average (or integral functional) of a stochastic field is again a stochastic field, and the resulting field does not automatically degenerate to deterministic variables. This contradiction makes it difficult to develop a homogenization technique for the framework shown in Figure 3.2.

Note in Sections 3.2.2 and 3.2.3 that, due to the independent distribution

assumption of white noise, the white noise functional degenerates to a constant field so that the stochastic-field model based on white noise is equivalent to a deterministic material model. This fact indicates that, there may be an independent middle state existing between the random material structure at the micro scale and the deterministic material structure at the macro scale. Given the existence of such an independent middle state, the framework of multi-scale simulations is drawn in Figure 3.3. There are two homogenization operations in this new framework; the first homogenization transforms the stochastic-field material model (e.g. the ERM model) into a generalized white noise model for which only the independent assumption is retained, and the second homogenization transforms the generalized white noise model into the deterministic material model in classic continuum mechanics.

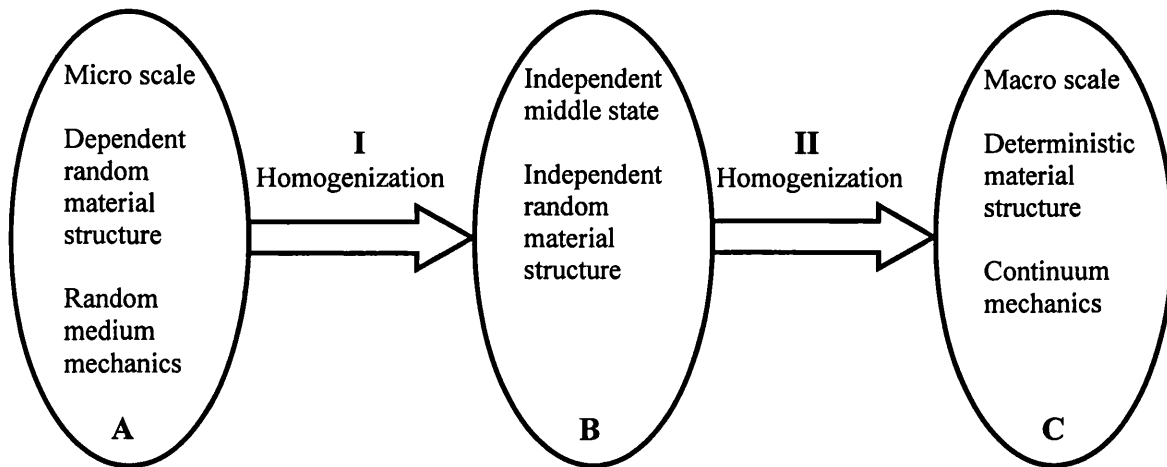


Figure 3.3 A stochastic approach for multi-scale simulations

In order to obtain a more detailed picture of the independent middle state B shown in Figure 3.3, it is necessary to carefully investigate the measurability of white noise. The following analysis on measurability will also clarify the footnote highlighted in the beginning of Section 3.2.1.

Recall Section 3.2.1 that $w(x, \omega)$ $x \in \mathbb{R}^n, \omega \in \Omega$ denotes a white noise field. Suppose $w(x, \omega)$ is measurable on $B \times \Omega$ where $B \subset \mathbb{R}^n$ is a Lebesgue measurable set. If ΔB is any Lebesgue measurable subset of B , it is trivial to obtain from Hölder's inequality (see e.g. 2.1.3) that

$$E\left(\int_{\Delta B} \int_{\Delta B} |w(x_1, \omega) w(x_2, \omega)| dx_1 dx_2\right) < +\infty. \quad (3.64)$$

Then, following Fubini's theorem (see e.g. [3.43-3.44]) and the condition (3.5), we have

$$\begin{aligned}
 & E\left(\left(\int_{\Delta B} w(\mathbf{x}, \omega) d\mathbf{x}\right)^2\right) \\
 &= E\left(\int_{\Delta B} \int_{\Delta B} w(\mathbf{x}_1, \omega) w(\mathbf{x}_2, \omega) d\mathbf{x}_1 d\mathbf{x}_2\right). \\
 &= \int_{\Delta B} \int_{\Delta B} E(w(\mathbf{x}_1, \omega) w(\mathbf{x}_2, \omega)) d\mathbf{x}_1 d\mathbf{x}_2 \\
 &= 0
 \end{aligned} \tag{3.65}$$

From the above equation,

$$\int_{\Delta B} w(\mathbf{x}, \omega) d\mathbf{x} = 0 \quad \text{for } \omega \notin N_{\Delta B} \tag{3.66}$$

where $N_{\Delta B} \subset \Omega$ denotes an event of probability zero, i.e.

$$P(N_{\Delta B}) = 0. \tag{3.67}$$

As the subset ΔB in (3.64-3.67) is arbitrary, the following relation

$$w(\mathbf{x}, \omega) = 0 \tag{3.68}$$

holds almost surely for all $\mathbf{x} \in B$ except possibly for a set of Lebesgue measure zero.

Hence,

$$E\left(\int_B w(\mathbf{x}, \omega)^2 d\mathbf{x}\right) = 0. \tag{3.69}$$

However, according to Eq. (3.5) and following Fubini's theorem, we have

$$E\left(\int_B w(\mathbf{x}, \omega)^2 d\mathbf{x}\right) = \int_B E(w(\mathbf{x}, \omega)^2) d\mathbf{x} = V(B), \tag{3.70}$$

where $V(B)$ is the Lebesgue measure of B . Consequently, it is concluded from (3.69-3.70) that

$$V(B) = 0. \tag{3.71}$$

That is, the white noise field $w(\mathbf{x}, \omega)$ is measurable on $B \times \Omega$ only if the Lebesgue measure of B is zero.

It is important to have a useful mathematical model of a "white noise" field. A natural choice of B is the countable set in \mathbb{R}^n , then the stochastic field $w(\mathbf{x}, \omega)$ $\mathbf{x} \in B, \omega \in \Omega$ becomes a family of discrete, independent and identically distributed random variables. This white noise model has been widely used in engineering, such as time series analysis [3.53]. Another way to define a white noise field is to describe its covariance by the Delta function, i.e.

$$\text{Cov}(w(\mathbf{x}_1, \omega), w(\mathbf{x}_2, \omega)) = \sigma^2 \delta(\mathbf{x}_1 - \mathbf{x}_2). \quad (3.72)$$

Compared with the engineering white noise field defined in (3.5), the above definition is preferred in mathematics. Substitution (3.72) into (3.20) yields

$$\begin{aligned} & \text{Cov}(W(\mathbf{x}_1, \omega), W(\mathbf{x}_2, \omega)) \\ &= \int_{\mathbb{R}^{2n}} \text{Cov}(w(y_1, \omega), w(y_2, \omega)) k(\mathbf{x}_1 - y_1) k(\mathbf{x}_2 - y_2) dy_1 dy_2 \\ &= \int_{\mathbb{R}^n} \sigma^2 k(\mathbf{x}_1 - \mathbf{y}) k(\mathbf{x}_2 - \mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^n} \sigma^2 k(\mathbf{x}_1 - \mathbf{x}_2 - \mathbf{z}) k(-\mathbf{z}) d\mathbf{z} \\ &= R(\mathbf{x}_1 - \mathbf{x}_2) \end{aligned} \quad (3.73)$$

which again indicates stationarity, one of the basic macro-scale properties of ERM.

With the above discussions, the stochastic approach of multi-scale simulations shown in Figure 3.3 can now be explained in more detail. Unlike the simple framework shown in Figure 3.2, the stochastic approach shown in Figure 3.3 divides the simulations into three scales (or stages) A, B and C. Two homogenization procedures are required to connect the scales A – B and the scales B – C, respectively. Consequently, at the scales A, B and C, the material under multi-scale investigation is described by three different mathematical models.

- I) At the micro scale A, material properties are modelled by stochastic fields (or random variables). The ERM model, depending on the specific micro-scale material structure, might be a choice for this purpose.
- II) At the independent state B, material properties are modelled by a countable set of independent random variables. Note, following the discussion in Section 3.2.3, these random variables do not necessarily have identical distribution.
- III) At the macro scale C, material properties are modelled by deterministic fields (or variables). It is in this scale that the material behaviours are described by classic continuum mechanics.

The original goal of this PhD research was to develop a general stochastic approach for multi-scale simulations, which, from a theoretical viewpoint, was expected to be closed, and from a numerical viewpoint, was expected to be versatile. However, it turned out that the research was too ambitious to be finished during the course of this thesis. Consequently, multi-scale considerations were completely removed from the

research schedule, and it was decided to concentrate on random phenomena in one scale, which in the end becomes the focus of this thesis. Following the strategic discussion and the mathematical analysis in this subsection, much more research work, especially in the aspect of physics, need to be done in order to fulfil the framework shown in Figure 3.3.

References

- [3.1] S. Valliappan and T.D. Pham, Fuzzy finite element analysis of a foundation on an elastic soil medium, *International Journal for Numerical and Analytical Methods in Geomechanics*, 17 (11) (1993) 771-798.
- [3.2] S. Valliappan and T.D. Pham, Elastoplastic finite-element analysis with fuzzy parameters, *International Journal for Numerical Methods in Engineering*, 38 (4) (1995) 531-548.
- [3.3] S.S. Rao and J.P. Sawyer, Fuzzy finite element approach for the analysis of imprecisely defined systems, *AIAA Journal*, 33 (12) (1995) 2364-2370.
- [3.4] D. Moens and D. Vandepitte, An interval finite element approach for the calculation of envelope frequency response functions, *International Journal for Numerical Methods in Engineering*, 61 (14) (2004) 2480-2507.
- [3.5] D. Moens and D. Vandepitte, A survey of non-probabilistic uncertainty treatment in finite element analysis, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1527-1555 (Included in Ref. [3.29]).
- [3.6] C. Soize, A nonparametric model of random uncertainties for reduced matrix models in structural dynamics, *Probabilistic Engineering Mechanics*, 15 (3) (2000) 277-294.
- [3.7] C. Soize, Random matrix theory for modeling uncertainties in computational mechanics, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1331-1366 (Included in Ref. [3.29]).
- [3.8] M. Shinozuka, Simulation of multivariate and multidimensional random processes, *Journal of the Acoustical Society of America*, 49 (1) (1971) 357-367.
- [3.9] M. Shinozuka and Y.K. Wen, Monte-Carlo solution of nonlinear vibrations, *AIAA Journal*, 10 (1) (1972) 37-40.
- [3.10] M. Shinozuka and C.J. Astill, Random eigenvalue problems in structural analysis, *AIAA Journal*, 10 (4) (1972) 456-462.
- [3.11] M. Shinozuka, Monte Carlo solution of structural dynamics, *Computers &*

Structures, 2 (5-6) (1972) 855-874.

- [3.12] S. Nakagiri and T. Hisada, *An Introduction to Stochastic Finite Element Method: Analysis of Uncertain Structures*, Baifukan, Tokyo, 1985. (In Japanese)
- [3.13] E. Vanmarcke and M. Grigoriu, Stochastic finite element analysis of simple beams, *Journal of Engineering Mechanics – ASCE*, 109 (5) (1983) 1203-1214.
- [3.14] P.L. Liu and A.D. Kiureghian, Finite element reliability of geometrically nonlinear uncertain structures, *Journal of Engineering Mechanics – ASCE*, 117 (8) (1991) 1806-1825.
- [3.15] M. Shinozuka and G. Deodatis, Response variability of stochastic finite element systems, *Journal of Engineering Mechanics – ASCE*, 114 (3) (1988) 499-519.
- [3.16] F. Yamazaki, M. Shinozuka and G. Dasgupta, Neumann expansion for stochastic finite element analysis, *Journal of Engineering Mechanics – ASCE*, 114 (8) (1988) 1335-1354.
- [3.17] G. Deodatis, Weighted integral method I: stochastic stiffness matrix, *Journal of Engineering Mechanics – ASCE*, 117 (8) (1991) 1851-1864.
- [3.18] G. Deodatis and M. Shinozuka, Weighted integral method. II: response variability and reliability, *Journal of Engineering Mechanics – ASCE*, 117 (8) (1991) 1865-1877.
- [3.19] W.K. Liu, T. Belytschko and A. Mani, Random field finite elements, *International Journal for Numerical Methods in Engineering*, 23 (10) (1986) 1831-1845.
- [3.20] W.K. Liu, T. Belytschko and A. Mani, Probabilistic finite elements for nonlinear structural dynamics, *Computer Methods in Applied Mechanics and Engineering*, 56 (1) (1986) 61-81.
- [3.21] M. Kleiber and T.D. Hien, *The Stochastic Finite Element Method, Basic Perturbation Technique and Computer Implementation*, John Wiley & Sons, Chichester, 1992.
- [3.22] P.D. Spanos and R.G. Ghanem, Stochastic finite element expansion for random media, *Journal of Engineering Mechanics – ASCE*, 115 (5) (1989) 1035-1053.
- [3.23] R.G. Ghanem and P.D. Spanos, Polynomial chaos in stochastic finite elements, *Journal of Applied Mechanics – ASME*, 57 (1) (1990) 197-202.
- [3.24] R.G. Ghanem and P.D. Spanos, Spectral stochastic finite-element formulation for reliability analysis, *Journal of Engineering Mechanics – ASCE*, 117 (10) (1991) 2351-2372.
- [3.25] R.G. Ghanem, Ingredients for a general purpose stochastic finite elements

- implementation, *Computer Methods in Applied Mechanics and Engineering*, 168 (1-4) (1999) 19-34 (Included in Ref. [3.28]).
- [3.26] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements – A Spectral Approach*, Revised edition, Dover Publications, New York, 2003.
- [3.27] M. Grigoriu, *Stochastic Calculus: Applications in Science and Engineering*, Birkhauser, 2002.
- [3.28] Edited by M. Kleiber, Special issue on computational stochastic mechanics, *Computer Methods in Applied Mechanics and Engineering*, 168 (1-4) (1999) 1-354.
- [3.29] Edited by G.I. Schuëller, Special issue on computational methods in stochastic mechanics and reliability analysis, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1251-1795.
- [3.30] B.V. Nieuwenhof and J.P. Coyette, Modal approaches for the stochastic finite element analysis of structures with material and geometric uncertainties, *Computer Methods in Applied Mechanics and Engineering*, 192 (33-34) (2003) 3705-3729.
- [3.31] I. Doltsinis, Z. Kang and G. Cheng, Robust design of non-linear structures using optimization methods, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1779-1795 (Included in Ref. [3.29]).
- [3.32] M.K. Deb, I.M. Babuška and J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, *Computer Methods in Applied Mechanics and Engineering*, 190 (48) (2001) 6359-6372.
- [3.33] D.B. Xiu and G.E. Karniadakis, The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations, *SIAM Journal on Scientific Computing*, 24 (2) (2002) 619-644.
- [3.34] G. Stefanou and M. Papadrakakis, Stochastic finite element analysis of shells with combined random material and geometric properties, *Computer Methods in Applied Mechanics and Engineering*, 193 (1-2) (2004) 139-160.
- [3.35] H.J. Pradlwarter, M.F. Pellissetti, C.A. Schenk, G.I. Schuëller, A. Kreis, S. Fransen, A. Calvi and M. Klein, Realistic and efficient reliability estimation for aerospace structures, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1597-1617 (Included in Ref. [3.29]).
- [3.36] Prof. L.A. Zadeh, <http://www.cs.berkeley.edu/~zadeh/awfeho.html>.
- [3.37] Website of Wikipedia, http://en.wikipedia.org/wiki/Fuzzy_set.
- [3.38] M.L. Mehta, *Random Matrices – Revised and Enlarged Second Edition*, Academic Press, New York, 1991.
- [3.39] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion*

Processes, North-Holland Publishing Company, Amsterdam, 1981.

- [3.40] H. Holden, B. Øksendal, J. Ubøe and T. Zhang, *Stochastic Partial Differential Equations: A Modelling, White Noise Functional Approach*, Birkhäuser Boston, Boston, 1996.
- [3.41] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales – Volume 1 Foundations*, 2nd Edition, Cambridge University Press, Cambridge, UK, 2000.
- [3.42] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales – Volume 2 Itô Calculus*, John Wiley & Sons, New York, 1987.
- [3.43] M.Q. Zhou, *Functions of Real Variables*, 2nd Edition, Peking University Press, Beijing, 1998. (In Chinese)
- [3.44] H.L. Royden, *Real Analysis*, 3rd Edition, Prentice Hall, Englewood Cliffs, New Jersey, 1988.
- [3.45] A.Y. Khinchin, Correlation theory of stationary stochastic processes, *Mathematische Annalen*, 109 (1934) 604-615. (In German)
- [3.46] A.N. Kolmogorov, Stationary sequences in Hilbert space, *Bulletin of Moscow State University*, 2 (6) (1941) 1-40. (In Russian)
- [3.47] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, M.I.T. Technology Press and John Wiley & Sons, New York, 1950.
- [3.48] A.M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall, Englewood Cliffs, 1962.
- [3.49] M.P. Qian and G.L. Gong, *Applied Stochastic Processes*, Peking University Press, Beijing, 1998. (In Chinese)
- [3.50] H. Cramér and M.R. Leadbetter, *Stationary and Related Stochastic Processes*, John Wiley & Sons, New York, 1966.
- [3.51] R.J. Adler, *The Geometry of Random Fields*, John Wiley & Sons, New York, 1981.
- [3.52] P. Walters, *An Introduction to Ergodic Theory*, Springer, New York, 2000.
- [3.53] M.B. Priestley, *Spectral Analysis and Time Series Vol. 1 Univariate Series*, Academic Press, London, 1981.

Chapter 4

The Fourier-Karhunen-Loève Representation of Elementary Random Media

The ERM model introduced in the last chapter is defined by its first- and second-order statistical moments. That is, for an ERM $D \subset \mathbb{R}^n$, the random elastic tensor $C(\mathbf{x}, \omega)$ in Eq. (3.3) is implicitly given by

$$E(C(\mathbf{x}, \omega)) = C_o(\mathbf{x}) \quad \forall \mathbf{x} \in D \quad (4.1)$$

and

$$\begin{aligned} & \text{Cov}(C(\mathbf{x}_1, \omega), C(\mathbf{x}_2, \omega)) \\ & \triangleq \text{Cov}(c(\mathbf{x}_1, \omega), c(\mathbf{x}_2, \omega)) \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D, \\ & = R(\mathbf{x}_1 - \mathbf{x}_2) \\ & = R(\boldsymbol{\tau}) \end{aligned} \quad (4.2)$$

in which $C_o(\mathbf{x})$ is a tensor-valued function whose scalar elements are the expectations of the corresponding scalar entries in $C(\mathbf{x}, \omega)$; $c(\mathbf{x}, \omega)$ is a vector constructed from the scalar entries of $C(\mathbf{x}, \omega)$ via a one-to-one mapping; and the covariance function $R(\boldsymbol{\tau})$, where $\boldsymbol{\tau} = \mathbf{x}_1 - \mathbf{x}_2$, is matrix valued.

However, there is no probability integral directly involved in the SPDE system

(3.3), and in order to perform the associated differential operators, an explicit representation of $C(\mathbf{x}, \omega)$ is required, instead. An ideal form of the explicit representation is

$$C(\mathbf{x}, \omega) = \text{l.i.m.}_{k \rightarrow +\infty} \sum_{i=0}^k z_i(\omega) C_i(\mathbf{x}), \quad (4.3)$$

where $z_i(\omega)$ are a sequence of scalar-valued random variables, and $C_i(\mathbf{x})$ a sequence of tensor-valued deterministic functions. The advantage of the series (4.3) is evident since, following the separation of the random and deterministic parts of $C(\mathbf{x}, \omega)$, traditional PDE tools become immediately applicable to the SPDE (3.3).

Hence, the aim of this chapter is to develop for ERM an effective representation scheme in the form (4.3). This representation problem is not a unique issue for the ERM model. In fact, the first two statistical moments, due to their simplicity, have been widely used in various SFEM formulations [4.1-4.18] to define random material properties, and consequently the associated problem of stochastic-field representation has always been a key issue in the SFEM research. The differences between the aim of this chapter and the related research in the literature are:

- The emphasis in this chapter is on the general elastic tensor $C(\mathbf{x}, \omega)$, in which multiple random scalars are present and consequently various possible interactions between these different random material parameters need to be taken into consideration. However, to date, existing techniques in the SFEM literature are mainly about individual elastic parameters and in many cases [4.1-4.18], just one major random material property (e.g. Young's modulus or Poisson's ratio) is considered.
- This chapter attempts to develop a general analytical solution for the expansion (4.3) so that the resulting stochastic-field representation scheme can be sufficiently accurate and efficient both for numerical simulations of ERM and for theoretical studies of the associated SPDE system (3.3).

The remainder of this chapter is organized as follows.

- I) Section 4.1 briefly reviews existing stochastic-field representation techniques, among which the Karhunen-Loève expansion method, due to its special importance, is separately addressed in Section 4.2.

- II) In Section 4.3, based on the spectral representation theory of wide-sense stationary stochastic fields and the standard dimensionality reduction technology of principal component analysis, the so called Fourier-Karhunen-Loève representation scheme is derived for a single elastic parameter of ERM, first in an infinite domain, then in a regular domain and finally in an arbitrary domain.
- III) Section 4.4 discusses the computational issues, which also provides a theoretical platform to analyze and further reveal respectively the eigenvalue-decay properties and the harmonic essence of ERM.
- IV) Two examples are employed in Section 4.5 to investigate the performance of this new stochastic-field representation scheme.
- V) Based on the developments in Sections 4.3-4.5, the Fourier-Karhunen-Loève representation for the general elastic tensor of ERM is formulated in Section 4.6, in which the interactions between different random scalars are fully taken into account.
- VI) The chapter finishes in Section 4.7 with a summary of the main features of the Fourier-Karhunen-Loève representation and some suggestions for future research along this direction.

4.1 Overview of Existing Techniques

In SFEM research, a key issue for the analysis of random medium problems is to rationally describe the irregular variation of material properties through the medium, so that an appropriate stochastic finite element formulation can be established for the problem under consideration. A simple example of a material parameter of random media is shown in Figure 3.1.

Over the last few decades, a number of stochastic-field representation methods have been developed for the description of random material properties. Most of these methods are essentially developed for scalar stochastic fields, and they mainly include:

- The middle point method [4.1, 4.2]. This method discretizes the definition domain of the stochastic field with a FE (finite element) mesh, and in each element, the stochastic field is simply approximated by the random variable at the element

centre. Then, these central random variables are determined through a Cholesky decomposition of their covariance matrix, which is constructed by directly sampling the given covariance function.

- The local averaging method [4.3]. After discretizing the definition domain of the stochastic field with a FE mesh, the stochastic field in each element is represented by a random variable, which is calculated from a spatial average of the local stochastic field within the element. Due to the complexity of the specific averaging operation, this method is only applicable to simple beam elements.
- The shape function method [4.4-4.6]. This method also discretizes the definition domain of the stochastic field with a FE mesh. However, the stochastic field in each element is interpolated by nodal random variables and the associated shape functions, and the covariance between each pair of nodal random variables are simply enforced to take the exact value from the given covariance function.
- The least-squares discretization method [4.7]. In this method, the stochastic field is approximated in the same discretization format as in the shape function method. However, the associated nodal random variables are determined via an optimization procedure, in which the approximation accuracy within each element is measured by the variance of the error between the exact stochastic field and the interpolated stochastic field.
- The trigonometric series approximation method [4.8-4.11]. No FE mesh is required in this method. The stochastic field is approximated by a general trigonometric series with random coefficients, which is obtained by directly discretizing the spectral representation of the wide-sense stationary stochastic field under consideration. More details regarding this method can be found in Appendix 4C.
- The Karhunen-Loève expansion method [4.12-4.17]. The method, as indicated by its name, is based on K-L (Karhunen-Loève) expansions [4.19] of second-order stochastic fields, which was first introduced into the SFEM research by Ghanem and Spanos [4.12], and since then has been widely used to describe random material properties [4.13-4.17]. To date, this method appears to be the most significant step forward for this stochastic-field representation problem. The K-L expansion is also part of the theoretical foundation of the F-K-L (Fourier-Karhunen-Loève) representation scheme to be developed in this chapter, and therefore the K-L expansion method will be explained with more details in the

next section.

It is observed in the above technical review that, FE meshes, as a discretization format, have been widely used in the solution of this stochastic-field representation problem. That is, for those random media whose material properties are essentially described by a scalar stochastic field, the deterministic basis functions in (4.3) have often been expressed in terms of piecewise polynomials. The advantage of representing random material properties by virtue of piecewise polynomials (or equivalently FE meshes) is evident. Specifically, the random material properties can be described with the same mesh employed in solving the SPDE system, which in turn makes it easier to deal with random media in a similar framework as the standard finite element method.

4.2 The Karhunen-Loève Expansion Method

In the current SFEM research, the *K-L expansion method* is probably the most popular method for describing random material properties. Assume that the material property of a random medium $D \in \mathbb{R}^n$ is described by a second-order stochastic field $b(\mathbf{x}, \omega)$ $\mathbf{x} \in D, \omega \in \Omega$, which is given by its first- and second- order statistical moments. Then, according to the K-L expansion theorem (see e.g. Appendix 4A), $b(\mathbf{x}, \omega)$ can be expressed as

$$b(\mathbf{x}, \omega) = E(b(\mathbf{x}, \omega)) + \sum_{i=1}^{+\infty} \sqrt{\lambda_i} \xi_i(\omega) \psi_i(\mathbf{x}) \quad (4.4)$$

in which the random variables $\xi_i(\omega)$ satisfy

$$E(\xi_i(\omega)) = 0 \quad \text{and} \quad E(\xi_i(\omega) \overline{\xi_j(\omega)}) = \delta_{ij}, \quad (4.5)$$

where δ_{ij} denotes the Kronecker delta; and the deterministic λ_i and $\psi_i(\mathbf{x})$ are eigenvalues and eigenfunctions of the characteristic equation

$$\int_D \text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)) \psi(\mathbf{x}_1) d\mathbf{x}_1 = \lambda \psi(\mathbf{x}_2) \quad (4.6)$$

such that $\lambda_i > 0$ and

$$\int_D \psi_i(\mathbf{x}) \overline{\psi_j(\mathbf{x})} d\mathbf{x} = \delta_{ij}. \quad (4.7)$$

In a numerical treatment, Eq. (4.4) is truncated into a finite series after sorting λ_i in a descending order, and the number of terms of the finite series can be determined by the required accuracy using the following trace relation

$$\sum_{i=1}^{+\infty} \lambda_i = \int_D \text{Cov}(b(\mathbf{x}, \omega), b(\mathbf{x}, \omega)) d\mathbf{x}. \quad (4.8)$$

It is highlighted that in the K-L expansion (4.4), the probability distribution of $\xi_i(\omega)$ remains undetermined unless $b(\mathbf{x}, \omega)$ is a Gaussian field, for which $\xi_i(\omega)$ are Gaussian random variables [4.19]. The first two statistical moments of $b(\mathbf{x}, \omega)$ are insufficient to fully define a general second-order stochastic field, and additional information must be provided in order to determine the probability distribution of $\xi_i(\omega)$.

In the SFEM literature, integral equation (4.6) is usually solved by the standard Galerkin finite element method (see e.g. [4.13, 4.17]), and the same FE mesh for the solution of the partial differential equations is also adopted. This leads to a generalized algebraic eigenvalue problem

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{B}\mathbf{v}, \quad (4.9)$$

in which \mathbf{A} is a full matrix (real symmetric and non-negative definite), \mathbf{B} a sparse matrix (real symmetric and positive definite), and λ and \mathbf{v} denote respectively the unknown eigenvalue and eigenvector. The associated computational costs mainly include two parts: constructing the full matrix \mathbf{A} and solving for the corresponding generalized eigenproblem. These operations can be very challenging because the dimensionality of the full matrix \mathbf{A} equals the total number of FE nodes and a large number of elements are required to achieve a reasonable K-L solution for practical random media [4.17]. Another deficiency in the above finite element based K-L representation scheme is the relatively poor approximation to the eigenfunctions $\psi_i(\mathbf{x})$ [4.13]. This is because there is no direct error control for the eigenfunctions although the error of the eigenvalues can be partially controlled via (4.8). These computational difficulties make it hard to apply the “K-L expansion method[†]” to describe practical random media, of which the ratio between the material dimension and the effective correlation length (see Appendix 4B for the exact

[†] In this thesis, the term “K-L expansion” is different from the term “K-L expansion method”. Specifically, the former denotes the mathematical result in the K-L expansion theorem (see e.g. Appendix 4A), while the latter represents the corresponding FE-mesh based stochastic-field representation method for random material properties.

definition) is often greater than 10.

4.3 The Fourier-Karhunen-Loève Representation for a Single Elastic Parameter of Elementary Random Media

In mathematics, the eigenpair λ_i and $\psi_i(\mathbf{x})$ in the K-L expansion (4.4) are essentially the spectrum of the linear integral operator

$$(Tf)(\mathbf{x}_2) \triangleq \int_D \text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)) f(\mathbf{x}_1) d\mathbf{x}_1, \quad (4.10)$$

where $f(\mathbf{x}) \in L^2(\mathbb{R}^n)$. It is usually difficult to accurately obtain the spectrum of T with respect to a general second-order stochastic field $b(\mathbf{x}, \omega)$. However, due to the harmonic essence of wide-sense stationary stochastic fields [4.24-4.25], it is possible to establish an effective representation scheme for ERM with little computational costs.

In order to simplify the notation and avoid possible confusions, the F-K-L representation of ERM is first established for a special case via three progressive steps in this section, where the random material property is described by a scalar stochastic field; and the general case, in which the random material properties are described by a tensor stochastic field, will be accordingly formulated in Section 4.6.

4.3.1 The Fourier-Karhunen-Loève Representation in \mathbb{R}^n

Consider an infinite n -dimensional ERM with its random material property represented by $a(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$. Recall from Chapter 3 that $a_o(\mathbf{x})$, $a_s(\mathbf{x}, \omega)$ and $R(\boldsymbol{\tau}) = R(\mathbf{x}_1 - \mathbf{x}_2)$ denote respectively the expectation function, the fluctuation field and the covariance function of $a(\mathbf{x}, \omega)$. Then, in accordance with

- I) The continuity and stationarity of $a_s(\mathbf{x}, \omega)$, and the spectral representation theory of wide-sense stationary stochastic fields (see Appendix 4C)
- II) The boundedness and integrability of $R(\boldsymbol{\tau})$, and the existence criterion of Fourier

transforms (see e.g. [4.27-4.28])

the following spectral decompositions hold

$$a(\mathbf{x}, \omega) = a_0(\mathbf{x}) + \int_{\mathbb{R}^n} e^{\sqrt{-1}\mathbf{x}\cdot\mathbf{y}} dZ(\mathbf{y}, \omega), \quad (4.11)$$

$$R(\boldsymbol{\tau}) = \int_{\mathbb{R}^n} e^{\sqrt{-1}\boldsymbol{\tau}\cdot\mathbf{y}} dF(\mathbf{y}) = \int_{\mathbb{R}^n} f(\mathbf{y}) e^{\sqrt{-1}\boldsymbol{\tau}\cdot\mathbf{y}} d\mathbf{y}, \quad (4.12)$$

where $F(\mathbf{y})$ is a bounded, real-valued function satisfying

$$\int_A dF(\mathbf{y}) \geq 0 \quad (4.13)$$

for all measurable $A \subset \mathbb{R}^n$; $f(\mathbf{y}) \geq 0$ vanishes at infinity, i.e.

$$\lim_{\|\mathbf{y}\|_2 \rightarrow +\infty} f(\mathbf{y}) = 0; \quad (4.14)$$

and $Z(\mathbf{y}, \omega)$, satisfying

$$\begin{aligned} Z(-\infty, \omega) &= 0 & E(Z(\mathbf{y}, \omega)) &= 0 \\ E(|Z(\mathbf{y}, \omega)|^2) &= F(\mathbf{y}) & E(|dZ(\mathbf{y}, \omega)|^2) &= dF(\mathbf{y}) = f(\mathbf{y}) d\mathbf{y} \end{aligned} \quad (4.15)$$

is a complex-valued stochastic field with orthogonal increments, i.e.

$$E\left(Z(\mathbf{F}_1^n, \omega) \overline{Z(\mathbf{F}_2^n, \omega)}\right) = 0 \quad (4.16)$$

for any pair of disjoint n -intervals $\mathbf{F}_1^n, \mathbf{F}_2^n \subset \mathbb{R}^n$. In Eqs. (4.11, 4.12), both $a(\mathbf{x}, \omega)$ and $R(\boldsymbol{\tau})$ are expressed in the frequency space and in the form of the inverse Fourier transform. In particular, $F(\mathbf{y})$ is termed the *spectral distribution function* of $a(\mathbf{x}, \omega)$ and $R(\boldsymbol{\tau})$, and $f(\mathbf{y})$ is the associated *spectral density function* that can be readily obtained via the Fourier transform

$$f(\mathbf{y}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} R(\boldsymbol{\tau}) e^{-\sqrt{-1}\boldsymbol{\tau}\cdot\mathbf{y}} d\boldsymbol{\tau}. \quad (4.17)$$

Eq. (4.11) is termed the *F-K-L representation* of ERM in \mathbb{R}^n . Applying Eq. (4.17) to the following integration of $R(\mathbf{x}_1 - \mathbf{x}_2) e^{-\sqrt{-1}\mathbf{x}_1\cdot\mathbf{y}}$ yields

$$\begin{aligned} & \int_{\mathbb{R}^n} R(\mathbf{x}_1 - \mathbf{x}_2) e^{-\sqrt{-1}\mathbf{x}_1\cdot\mathbf{y}} d\mathbf{x}_1 \\ &= e^{-\sqrt{-1}\mathbf{x}_2\cdot\mathbf{y}} \int_{\mathbb{R}^n} R(\mathbf{x}_1 - \mathbf{x}_2) e^{-\sqrt{-1}(\mathbf{x}_1 - \mathbf{x}_2)\cdot\mathbf{y}} d\mathbf{x}_1. \\ &= (2\pi)^n f(\mathbf{y}) e^{-\sqrt{-1}\mathbf{x}_2\cdot\mathbf{y}} \end{aligned} \quad (4.18)$$

The above characteristic equation is satisfied for all $y \in \mathbb{R}^n$. More importantly, by comparing (4.18) with (4.6), it is clear that, without solving Eq. (4.6), the expressions (4.11, 4.12) provide an explicit solution for the K-L expansion of $a(x, \omega)$ $x \in \mathbb{R}^n$ in terms of Fourier integrals. Specifically, $dZ(y, \omega)$ and $e^{\sqrt{-1}x \cdot y}$ in (4.11) correspond respectively to $\sqrt{\lambda_i} \xi_i(\omega)$ and $\psi_i(x)$ in (4.4), and the continual spectrum indexed by y corresponds to the discrete spectrum indexed by i .

4.3.2 The Fourier-Karhunen-Loève Representation in an n -Interval

4.3.2.1 The Fourier Expansion of Elementary Random Media

Let

$$\mathbf{T}^n \triangleq \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : -t_k \leq x_k \leq t_k, t_k > 0 \quad (k = 1, \dots, n)\} \quad (4.19)$$

denote an n -interval with volume

$$V_{\mathbf{T}^n} = 2^n \prod_{k=1}^n t_k, \quad (4.20)$$

and restrict the definition domain of $a(x, \omega)$ in \mathbf{T}^n . Note that the Fourier integral and series are essentially equivalent except that the former is defined in \mathbb{R}^n but the latter is defined in an n -interval (see e.g. [4.27-4.28]). The inverse Fourier transform (4.11) of $a(x, \omega)$ $x \in \mathbb{R}^n$ infers the following complex Fourier expansion of $a(x, \omega)$ $x \in \mathbf{T}^n$

$$a(x, \omega) = a_o(x) + \sum_{m_1=-\infty}^{+\infty} \dots \sum_{m_n=-\infty}^{+\infty} \Delta Z(m, \omega) e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}, \quad (4.21)$$

where $m = (m_1, \dots, m_n) \in \mathbb{Z}^n$;

$$\theta_k = \frac{\pi}{t_k} \quad k = 1, \dots, n; \quad (4.22)$$

$e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}$ are the Fourier basis defined in \mathbf{T}^n ; and

$$\Delta Z(m, \omega) = \frac{1}{V_{\mathbf{T}^n}} \int_{\mathbf{T}^n} (a(x, \omega) - a_o(x)) e^{-\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} dx \quad (4.23)$$

are complex-valued random Fourier coefficients. The convergence and well-posedness of (4.21) are ensured by the completeness and orthogonality of the Fourier basis (see e.g. [4.27-4.29]). It is then obvious that

$$E(\Delta Z(\mathbf{m}, \omega)) = 0. \tag{4.24}$$

Although the stochastic field $Z(\mathbf{y}, \omega)$ in (4.11) is with orthogonal increments, the random sequence $\Delta Z(\mathbf{m}, \omega)$ in (4.21) is not necessarily (and usually is not) composed of orthogonal random variables. This is because of the simplification from the Fourier integrals to the Fourier series; the frequency of the trigonometric functions in (4.21) is preset by \mathbf{m} and consequently random coefficients $\Delta Z(\mathbf{m}, \omega)$ must be adapted for the convergence of the infinite Fourier series, for which the strictly orthogonal property of $Z(\mathbf{y}, \omega)$ is inevitably destroyed. Hence, Eq. (4.21) is not the K-L expansion of $a(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbf{T}^n$. However, it can be expected that $\Delta Z(\mathbf{m}, \omega)$ are nearly orthogonal to each other such that

$$E\left(Z(\mathbf{m}_1, \omega) \overline{Z(\mathbf{m}_2, \omega)}\right) \approx 0 \quad \forall \mathbf{m}_1 \neq \mathbf{m}_2, \tag{4.25}$$

therefore $\Delta Z(\mathbf{m}, \omega)$ and $e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}$ provide a good initial estimate of the K-L eigenstructure of $a(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbf{T}^n$. This intuitive insight will be further exploited in Section 4.4.1.

In a numerical formulation, the convergent Fourier series (4.21) is truncated such that

$$\begin{aligned} a(\mathbf{x}, \omega) &\approx a_o(\mathbf{x}) + \sum_{m_1=-M_1}^{M_1} \cdots \sum_{m_n=-M_n}^{M_n} \Delta Z(\mathbf{m}, \omega) e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} \\ &= a_o(\mathbf{x}) + (\Delta Z(\omega))^T \mathbf{e}(\mathbf{x}) \\ &= a_N(\mathbf{x}, \omega) \end{aligned} \tag{4.26}$$

where $\mathbf{M} = (M_1, \dots, M_n) \in \mathbb{N}^n$;

$$N = \prod_{k=1}^n (2M_k + 1); \tag{4.27}$$

and vectors $\Delta Z(\omega)$ and $\mathbf{e}(\mathbf{x})$ are respectively constructed from $\sqrt{V_{\mathbf{T}^n}} \Delta Z(\mathbf{m}, \omega)$ and



$e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} / \sqrt{V_{\mathbf{T}^n}}$. The constant coefficient $\sqrt{V_{\mathbf{T}^n}}$ is chosen to normalise the Fourier basis, and the entries in $\Delta \mathbf{Z}(\omega)$ and $e(\mathbf{x})$ are organized in an ascending order of the frequency number m . In the Fourier approximation (4.26), the truncation error $a(\mathbf{x}, \omega) - a_N(\mathbf{x}, \omega)$ depends on the discrete frequency domain

$$\mathbf{M}^n \triangleq \{m \in \mathbb{Z}^n : -M_k \leq m_k \leq M_k \quad (k=1, \dots, n)\} \quad (4.28)$$

with boundary M . It is then desirable to have an explicit error control on M according to the required accuracy, and this can be achieved with the assistance of the spectral representation of $R(\boldsymbol{\tau})$.

Specifically, let

$$\mathbf{F}^n \triangleq \{y \in \mathbb{R}^n \mid -f_k \leq y_k \leq f_k, f_k > 0 \quad (k=1, \dots, n)\} \quad (4.29)$$

denote the smallest domain in the frequency space of $R(\boldsymbol{\tau})$ such that

$$\frac{\int_{\mathbf{F}^n} f(y) dy}{\int_{\mathbb{R}^n} f(y) dy} = \frac{\int_{\mathbf{F}^n} f(y) dy}{R(\mathbf{0})} \geq \mu \quad 0 < \mu < 1 \quad (4.30)$$

where constant μ is the required accuracy for approximating the spectrum identified by $f(y)$. The boundary M is then determined by

$$M_k = \left\lceil \frac{f_k}{\theta_k} \right\rceil = \left\lceil \frac{f_k t_k}{\pi} \right\rceil \quad k=1, \dots, n \quad (4.31)$$

where operator $\lceil x \rceil$ gives the minimum integer that is not less than x . The derivation is explained in Appendix 4D.

Regarding the above Fourier expansion of ERM, the following two remarks are in order.

- I) It is well known that the Fourier transforms of many typical functions can be obtained analytically, therefore the spectral density function $f(y)$ (4.17) can be exactly obtained for many typical ERM encountered in practice. If this is not the case, the standard FFT can be applied to numerically calculate $f(y)$.
- II) According to the well known Heisenberg inequality [4.27-4.28] in Fourier analysis,

a wider $R(\boldsymbol{\tau})$, which approximately corresponds to a larger effective correlation length, results in a narrower $f(\boldsymbol{y})$, which approximately corresponds to a smaller major frequency domain \mathbf{F}^n ; and vice versa. Consequently, from (4.31), \mathbf{M}^n expands with the growth of the random medium domain, but normally shrinks as the effective correlation length increases; and vice versa.

4.3.2.2 Principal Component Analysis of Elementary Random Media

In (4.26), in terms of the fixed Fourier vector $\boldsymbol{e}(\boldsymbol{x})$, $a_N(\boldsymbol{x}, \omega) - a_o(\boldsymbol{x})$ is represented by the random vector $\Delta\mathbf{Z}(\omega)$ whose entries $\Delta Z(\boldsymbol{m}, \omega)$ are not strictly orthogonal to each other. However, as discussed in the previous subsection, $\Delta\mathbf{Z}(\boldsymbol{m}, \omega)$ and $\boldsymbol{e}^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}$ do provide a good initial estimate of the K-L eigenstructure of $a(\boldsymbol{x}, \omega)$ $\boldsymbol{x} \in \mathbf{T}^n$. Hence, due to the orthogonality of the Fourier basis, the standard dimensionality reduction technique of PCA (principal component analysis) can be performed to orthogonalise $\Delta\mathbf{Z}(\boldsymbol{m}, \omega)$ and improve this approximate K-L solution. The PCA operation for $\Delta\mathbf{Z}(\omega)$ is outlined below, and its theoretical background can be found in standard textbooks (e.g. [4.26]).

Let

$$\mathbf{G} = E\left(\Delta\mathbf{Z}(\omega)(\Delta\mathbf{Z}(\omega))^H\right) \quad (4.32)$$

denote the covariance matrix of $\Delta\mathbf{Z}(\omega)$, $\boldsymbol{e}_i(\boldsymbol{x})$ the i -th term of $\boldsymbol{e}(\boldsymbol{x})$ and $\Delta Z_i(\omega)$ the i -th term of $\Delta\mathbf{Z}(\omega)$. Then the matrix \mathbf{G} can be constructed from

$$\begin{aligned} (\mathbf{G})_{ij} &= \text{Cov}(\Delta Z_i(\omega), \Delta Z_j(\omega)) \\ &= \int_{\mathbf{T}^n} \int_{\mathbf{T}^n} R(\boldsymbol{x}_1 - \boldsymbol{x}_2) \overline{\boldsymbol{e}_i(\boldsymbol{x}_1)} \boldsymbol{e}_j(\boldsymbol{x}_2) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \end{aligned} \quad (4.33)$$

where $(\mathbf{G})_{ij}$ is the entry at the i -th row and j -th column of \mathbf{G} . Due to the symmetric and non-negative definite properties of $R(\boldsymbol{x}_1 - \boldsymbol{x}_2)$ (see e.g. [4.19, 4.24-4.25] or Section 2.2.3), \mathbf{G} is a non-negative definite Hermitian matrix (see Section 4.4.1 for more details), therefore there exists a unitary matrix $\mathbf{Q}^{-1} = \mathbf{Q}^H$ such that

$$\mathbf{Q}^H \mathbf{G} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N), \quad (4.34)$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$. Let

$$\boldsymbol{\psi}(\mathbf{x}) = (\psi_1(\mathbf{x}), \dots, \psi_N(\mathbf{x}))^T = \mathbf{Q}^H \mathbf{e}(\mathbf{x}), \quad (4.35)$$

$$\boldsymbol{\xi}(\omega) = (\xi_1(\omega), \dots, \xi_N(\omega))^T = \text{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_N}}\right) \mathbf{Q}^H \Delta \mathbf{Z}(\omega), \quad (4.36)$$

then it follows from Eq. (4.26) that

$$\begin{aligned} & a_N(\mathbf{x}, \omega) \\ &= a_o(\mathbf{x}) + \sum_{i=1}^N \sqrt{\lambda_i} \xi_i(\omega) \psi_i(\mathbf{x}) \\ &\approx a_o(\mathbf{x}) + \sum_{i=1}^{N^*} \sqrt{\lambda_i} \xi_i(\omega) \psi_i(\mathbf{x}) \\ &= a_{N^*}(\mathbf{x}, \omega) \end{aligned} \quad (4.37)$$

in which N^* is the smallest integer such that

$$\frac{\sum_{i=1}^{N^*} \lambda_i}{\sum_{i=1}^N \lambda_i} \geq \mu^* \quad 0 < \mu^* \leq 1, \quad (4.38)$$

where μ^* is the required accuracy for approximating the total variance of $a_N(\mathbf{x}, \omega)$.

Eq. (4.37) together with (4.33-4.36) is termed the *F-K-L representation* of ERM in n -intervals. Due to the completeness and orthogonality of the Fourier basis and the PCA procedure, $a_{N^*}(\mathbf{x}, \omega)$ provides a series solution for the K-L expansion of $a(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbf{T}^n$. As $a(\mathbf{x}, \omega)$ is a Gaussian field in the ERM model, it is trivial to see that $\xi_i(\omega)$ in (4.37) are stochastically independent standard Gaussian random variables. The truncation error $a(\mathbf{x}, \omega) - a_{N^*}(\mathbf{x}, \omega)$ is explicitly controlled by (4.30-4.31) and (4.38), and specifically μ controls the error of the spectrum (i.e. eigenvalues and eigenfunctions) of $a(\mathbf{x}, \omega)$ and μ^* controls the error of the total variance of $a(\mathbf{x}, \omega)$.

4.3.3 The Fourier-Karhunen-Loève Representation in an Arbitrary Domain

Consider an ERM $D \subset \mathbb{R}^n$ whose random material property is described by $a(\mathbf{x}, \omega)$ $\mathbf{x} \in D, \omega \in \Omega$. Letting $\widetilde{\mathbf{T}}^n$ be the minimal n -interval such that $D \subset \widetilde{\mathbf{T}}^n$, there exists the F-K-L representation for $a_{N^*}(\mathbf{x}, \omega)$ $\mathbf{x} \in \widetilde{\mathbf{T}}^n$. Due to the convergence property of the K-L expansion (see Appendix 4A), it is clear that for any point $\mathbf{x}_0 \in D \subset \widetilde{\mathbf{T}}^n$, $a_{N^*}(\mathbf{x}_0, \omega)$ converges to $a(\mathbf{x}_0, \omega)$ in a mean-square sense. Consequently, with this simple expansion of D into $\widetilde{\mathbf{T}}^n$, an explicit representation scheme is achieved for $a(\mathbf{x}, \omega)$ $\mathbf{x} \in D$, and this is called the *F-K-L representation* defined in an arbitrary domain. However, it should be noted that the F-K-L representation scheme provides a series solution for the K-L expansion of an ERM if and only if the ERM is defined in the whole space or an n -interval.

4.4 Computational Issues

4.4.1 Computational Techniques

Computational costs of the F-K-L representation scheme include two parts: constructing the $N \times N$ covariance matrix \mathbf{G} and solving the associated standard algebraic eigenvalue problem to obtain λ_i and \mathbf{Q} .

Following (4.33),

$$(\mathbf{G})_{ij} = \frac{1}{V_{\mathbf{T}^n}} \int_{\mathbf{T}^n} \int_{\mathbf{T}^n} R(\mathbf{x}_1 - \mathbf{x}_2) e^{-\sqrt{-1} \sum_{k=1}^n m_{ik} \theta_k x_{1k}} e^{\sqrt{-1} \sum_{k=1}^n m_{jk} \theta_k x_{2k}} dx_1 dx_2. \quad (4.39)$$

The substitution of (4.12) into (4.39) followed by a trivial calculation yields

$$(\mathbf{G})_{ij} = \frac{2^n}{V_{\mathbf{T}^n}} \int_{\mathbb{R}^n} f(\mathbf{y}) \prod_{k=1}^n \frac{\cos\left(\left(m_{jk} - m_{ik}\right)\pi\right) - \cos\left(2y_k t_k - \left(m_{ik} + m_{jk}\right)\pi\right)}{\left(y_k - m_{ik} \frac{\pi}{t_k}\right) \left(y_k - m_{jk} \frac{\pi}{t_k}\right)} dy \quad (4.40)$$

Substituting (4.17) into (4.40) yields

$$(\mathbf{G})_{ij} = \frac{1}{\pi^n V_{\mathbf{T}^n}} \int_{\mathbb{R}^n} R(\boldsymbol{\tau}) \left(\prod_{k=1}^n G_k \right) d\boldsymbol{\tau}, \quad (4.41)$$

in which

$$G_k = \int_{\mathbb{R}} \frac{\cos\left(\left(m_{jk} - m_{ik}\right)\pi\right) - \cos\left(2y_k t_k - \left(m_{ik} + m_{jk}\right)\pi\right)}{\left(y_k - m_{ik} \frac{\pi}{t_k}\right)\left(y_k - m_{jk} \frac{\pi}{t_k}\right)} e^{-\sqrt{-1}y_k \tau_k} dy_k. \quad (4.42)$$

After explicitly calculating the integral (4.42), which is a standard Fourier transform, the expression (4.41) can be simplified as

$$(\mathbf{G})_{ij} = \int_0^{2t_1} \cdots \int_0^{2t_n} \tilde{R}(\boldsymbol{\tau}) \left(\prod_{k=1}^n A_{ijk}(\tau_k) B_{ijk}(\tau_k) \right) d\boldsymbol{\tau}, \quad (4.43)$$

where

$$\tilde{R}(\boldsymbol{\tau}) = \frac{1}{2^n} \sum_{k_1=1}^2 \cdots \sum_{k_n=1}^2 R\left((-1)^{k_1} \tau_1, \dots, (-1)^{k_n} \tau_n\right) \quad (4.44)$$

denotes the axis-symmetric component of $R(\boldsymbol{\tau})$; and

$$A_{ijk}(\tau_k) = \begin{cases} 2 - \frac{\tau_k}{t_k} & m_{ik} = m_{jk} \\ \frac{(-1)^{m_{ik} + m_{jk}}}{(m_{ik} - m_{jk})\pi} & m_{ik} \neq m_{jk} \end{cases}, \quad (4.45)$$

$$B_{ijk}(\tau_k) = \begin{cases} \cos\left(m_{ik} \frac{\pi}{t_k} \tau_k\right) & m_{ik} = m_{jk} \\ \sin\left(m_{jk} \frac{\pi}{t_k} \tau_k\right) - \sin\left(m_{ik} \frac{\pi}{t_k} \tau_k\right) & m_{ik} \neq m_{jk} \end{cases}. \quad (4.46)$$

From (4.43-4.46), it is observed that the matrix \mathbf{G} is real and symmetric. Hence, the non-negative definite Hermitian matrix \mathbf{G} is specifically real-symmetric and non-negative definite. The concise expressions of (4.43-4.46) will not only make efficient computational techniques applicable to this problem but also provide a good starting point to explore the properties of finite ERM.

Note in (4.43-4.46) that, all the entries in \mathbf{G} are essentially Fourier coefficients of $\tilde{R}(\boldsymbol{\tau}) \prod_{k=1}^n A_{ijk}(\tau_k)$; and up to a constant scalar determined by i and j , there are only 2^n

($n=1,2$ or 3 denotes the dimensionality of the stochastic field $a(\mathbf{x},\omega)$) possible forms of $\tilde{R}(\boldsymbol{\tau})\prod_{k=1}^n A_{ijk}(\tau_k)$. Hence, the covariance matrix can be efficiently constructed by using the FFT [4.30], and the computational cost is $O(2^n N \log N)$. It is worth mentioning that the right-hand side of (4.43) may be analytically obtained for some typical covariance functions encountered in practice.

In addition, there are two observations:

- I) The integral kernel $A_{ijk}(\tau_k)$ is much larger in a diagonal entry of \mathbf{G} than in an off-diagonal entry; with the decreasing rate $1/\prod_{k=1}^n |m_{ik} - m_{jk}|$, it becomes increasingly smaller when the associated entry moves away from the diagonal.
- II) In practice, the function $\tilde{R}(\boldsymbol{\tau})$ usually exhibits its largest value at $\boldsymbol{\tau}=\mathbf{0}$ and decreases as $\|\boldsymbol{\tau}\|_2$ increases. Therefore, the major non-zero domain of $\tilde{R}(\boldsymbol{\tau})$ is in the vicinity of the origin, in which $B_{ijk}(\tau_k)$ is much larger in a diagonal entry than in an off-diagonal entry depending on the decay rate of $\tilde{R}(\boldsymbol{\tau})$.

Consequently, following these two observations, it appears that the covariance matrix \mathbf{G} is diagonally dominant. Furthermore, \mathbf{G} can be transformed into a block diagonal matrix containing two diagonally-dominant blocks, and the dimensionality of each block matrix is approximately half of the dimensionality of \mathbf{G} (see Appendix 4E for details). Consequently, the algebraic eigenstructure of \mathbf{G} can be readily obtained by using existing eigensolvers with moderate computational costs.

4.4.2 The Diagonal Fourier-Karhunen-Loève Approximation Scheme

The diagonally-dominant property of \mathbf{G} infers that $\Delta Z_i(\omega)$ are nearly orthogonal, which implies that $\Delta Z_i(\omega)$ and $e_i(\mathbf{x})$ provide a good initial estimate to the K-L eigenstructure of $a(\mathbf{x},\omega)$ $\mathbf{x} \in \mathbf{T}^n$. This insight makes it interesting to investigate the following stochastic field constructed only from the diagonal entries of \mathbf{G}

$$a_d(\mathbf{x}, \omega) \triangleq a_o(\mathbf{x}) + \sum_{i=1}^{+\infty} \Delta \tilde{Z}_i(\omega) e_i(\mathbf{x}), \quad (4.47)$$

where

$$E(\Delta \tilde{Z}_i(\omega)) = 0 \quad (4.48)$$

and

$$E(\Delta \tilde{Z}_i(\omega) \overline{\Delta \tilde{Z}_j(\omega)}) = \delta_{ij} E(\Delta Z_i(\omega) \overline{\Delta Z_j(\omega)}). \quad (4.49)$$

Utilising (4.43-4.46), the covariance function of $a_d(\mathbf{x}, \omega)$ can be obtained as

$$\begin{aligned} & \text{Cov}(a_d(\mathbf{x}_1, \omega), a_d(\mathbf{x}_2, \omega)) \\ &= \frac{1}{V_{\mathbf{T}^n}} \sum_{m_1=-\infty}^{+\infty} \cdots \sum_{m_n=-\infty}^{+\infty} \left(\int_0^{2t_1} \cdots \int_0^{2t_n} \tilde{R}(\boldsymbol{\tau}) \left(\prod_{k=1}^n \left(2 - \frac{\tau_k}{t_k} \right) \cos \left(m_k \frac{\pi}{t_k} \tau_k \right) \right) d\boldsymbol{\tau} \right) e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k (x_{1k} - x_{2k})}. \quad (4.50) \\ &= R_d(\mathbf{x}_1 - \mathbf{x}_2) = R_d(\boldsymbol{\tau}) \end{aligned}$$

According to (4.47, 4.50), the stochastic field $a_d(\mathbf{x}, \omega)$ has the same expectation function as $a(\mathbf{x}, \omega)$, and the stochastic field $a_d(\mathbf{x}, \omega) - a_o(\mathbf{x})$ is wide-sense stationary. Naturally, it is expected that $R_d(\boldsymbol{\tau})$ approximates $R(\boldsymbol{\tau})$ in a certain sense such that $a_d(\mathbf{x}, \omega)$ can provide a reasonable estimate of $a(\mathbf{x}, \omega)$.

According to the Fourier expansion in (4.50), $R_d(\boldsymbol{\tau})$ is essentially the axis-symmetric component of $\tilde{R}(\boldsymbol{\tau}) \prod_{k=1}^n \left(2 - \frac{\tau_k}{t_k} \right)$. In addition, $R_d(\boldsymbol{\tau})$ has a period of \mathbf{T}^n , therefore it is only necessary to consider the behaviour of $R_d(\boldsymbol{\tau})$ in one half of \mathbf{T}^n . That is,

$$\begin{aligned} & R_d(\boldsymbol{\tau}) \\ &= \frac{1}{2^n} \sum_{i_1=1}^2 \cdots \sum_{i_n=1}^2 \tilde{R}(t_1 + (-1)^{i_1} (t_1 - \tau_1), \dots, t_n + (-1)^{i_n} (t_n - \tau_n)) \prod_{k=1}^n \frac{t_k - (-1)^{i_k} (t_k - \tau_k)}{t_k}. \quad (4.51) \\ & \quad 0 \leq \tau_k \leq t_k \quad k = 1, \dots, n \end{aligned}$$

For the sake of simplicity, consider $R_d(\boldsymbol{\tau})$ in 1D cases, i.e. $n=1$, then

$$R_d(\tau) = \frac{1}{2} \left(\tilde{R}(\tau) \frac{2t - \tau}{t} + \tilde{R}(2t - \tau) \frac{\tau}{t} \right) \quad 0 \leq \tau \leq t. \quad (4.52)$$

As for 1D wide-sense stationary stochastic field

$$R(\boldsymbol{\tau}) \equiv R(-\boldsymbol{\tau}) \Rightarrow \tilde{R}(\boldsymbol{\tau}) \equiv R(\boldsymbol{\tau}), \quad (4.53)$$

the following relation holds

$$R_d(\boldsymbol{\tau}) - R(\boldsymbol{\tau}) = \frac{\boldsymbol{\tau}}{2t} (R(2t - \boldsymbol{\tau}) - R(\boldsymbol{\tau})) \quad 0 \leq \boldsymbol{\tau} \leq t. \quad (4.54)$$

Note that $R(\boldsymbol{\tau})$ in practice is monotone decreasing from the origin. Hence, for a sufficiently large t , we have

$$\begin{cases} R_d(\boldsymbol{\tau}) - R(\boldsymbol{\tau}) = 0 & \boldsymbol{\tau} = 0, t \\ R_d(\boldsymbol{\tau}) - R(\boldsymbol{\tau}) \approx 0 & 0 < \boldsymbol{\tau} < t \end{cases} \quad (4.55)$$

Consequently, for $\boldsymbol{\tau} \in [-t, t]$, the stochastic field $a_d(\boldsymbol{x}, \boldsymbol{\omega})$ provides a good estimate of $a(\boldsymbol{x}, \boldsymbol{\omega})$ in terms of the first two statistical moments. For $a(\boldsymbol{x}, \boldsymbol{\omega})$ of a higher dimensional ERM, it can be proven via a similar procedure that for $\boldsymbol{\tau} \in \mathbf{T}^n$, $a_d(\boldsymbol{x}, \boldsymbol{\omega})$ provides a good estimate of $a(\boldsymbol{x}, \boldsymbol{\omega})$ as long as \mathbf{T}^n is sufficiently large and $a(\boldsymbol{x}, \boldsymbol{\omega})$ is isotropic or orthotropic. The above analysis requires $a(\boldsymbol{x}, \boldsymbol{\omega})$ to be isotropic or orthotropic in higher dimensional cases. This is because, although $R(\boldsymbol{\tau}) \equiv R(-\boldsymbol{\tau})$ holds for any real-valued 2D/3D wide-sense stationary stochastic fields, $\tilde{R}(\boldsymbol{\tau}) = R(\boldsymbol{\tau})$ holds only for isotropic or orthotropic 2D/3D wide-sense stationary stochastic fields.

The advantage of the diagonal F-K-L approximation (4.47) is its inexpensive computational cost as no equation solving is involved. It should also be noted that (4.47) does not provide the K-L solution of $a(\boldsymbol{x}, \boldsymbol{\omega}) \quad \boldsymbol{x} \in \mathbf{T}^n$. For large random media with small effective correlation length, the full F-K-L representation scheme requires the solution of a large algebraic eigenvalue problem, which may be beyond available computer storage (especially in 3D cases). For those “unsolvable” situations, the diagonal F-K-L approximation scheme may provide an acceptable alternative.

4.4.3 Estimates on the Eigenvalue Decay

In a SFEM formulation, the size of the final algebraic equation system or the number of necessary parametric simulations is largely determined by the total number of random variables included in the random medium system. Generally, the more random variables in the system, the more computing power is required to solve the associated

SPDEs. Hence, in the analysis of ERM systems, it is desirable to have an approximate prediction on the eigenvalue decay of the F-K-L representation, which in turn determines the total number of random variables required by the description of random material properties. Let λ_i/λ_1 denote the eigenvalue decay rate. In the F-K-L representation scheme, λ_i/λ_1 can be readily predicted by two different methods.

In the first method, λ_i/λ_1 is estimated by using the spectral density function $f(y)$. Specifically, through a similar procedure to the derivation of (4.30, 4.31),

$$\frac{\lambda_i}{\lambda_1} \approx \frac{f\left(m_{i1} \frac{\pi}{t_1}, \dots, m_{in} \frac{\pi}{t_n}\right)}{f(\mathbf{0})} \quad i \geq 1. \quad (4.56)$$

From the above prediction, it is clear that the larger the random medium domain \mathbf{T}^n the slower the eigenvalues decay; and vice versa. In addition, due to Eq. (4.17) and the Heisenberg inequality [4.27, 4.28], a smaller effective correlation length normally corresponds to a wider spectral density function $f(y)$, and therefore results in slower eigenvalue decay; and vice versa.

In the second predicting method, due to the diagonally-dominant property of \mathbf{G} , λ_i/λ_1 can be estimated by using the diagonal entries. Specifically, according to (4.43-4.46),

$$\frac{\lambda_i}{\lambda_1} \approx \frac{\int_0^{2t_1} \dots \int_0^{2t_n} \tilde{R}(\boldsymbol{\tau}) \left(\prod_{k=1}^n \left(2 - \frac{\tau_k}{t_k} \right) \cos \left(m_{ik} \frac{\pi}{t_k} \tau_k \right) \right) d\boldsymbol{\tau}}{\int_0^{2t_1} \dots \int_0^{2t_n} \tilde{R}(\boldsymbol{\tau}) \prod_{k=1}^n \left(2 - \frac{\tau_k}{t_k} \right) d\boldsymbol{\tau}} \quad i \geq 1. \quad (4.57)$$

In the above prediction, the eigenvalues are estimated by the Fourier coefficients of the axis-symmetric component of $\tilde{R}(\boldsymbol{\tau}) \prod_{k=1}^n \left(2 - \frac{\tau_k}{t_k} \right)$. Hence, the eigenvalue decay rate λ_i/λ_1 is approximately the decay rate of some Fourier coefficients. Let $c_j(g)$ $j = 0, \pm 1, \pm 2, \dots$ be the Fourier coefficients of a function $g(x)$ defined in an interval. It is well known in the classic theory of Fourier series [4.27-4.29] that:

D) If g has bounded variation, then

$$|c_j(g)| \leq V(g)/|4j| \quad (4.58)$$

where $V(g)$ is the total variation of g ;

II) If the p -th derivative of g is absolutely continuous, then

$$c_j(g) = o\left(1/|j|^{p+1}\right) \quad \text{as } j \rightarrow +\infty. \quad (4.59)$$

Consequently, in 1D cases, λ_i/λ_1 is at least proportional to $1/i$ (for sufficiently large i) as long as the corresponding condition of the bounded variation is satisfied; moreover, the smoother the covariance function the faster the eigenvalues decay. The generalization of these conclusions to higher dimensional cases with separable $\tilde{R}(\boldsymbol{\tau})$ is immediate. For general higher dimensional cases, although the analysis on Fourier coefficients is not straightforward, it can be expected that essentially similar estimates hold for λ_i/λ_1 .

4.4.4 The Fourier-Karhunen-Loève Algorithm for Scalar Stochastic Fields

For an ERM described by $a(\mathbf{x}, \omega)$ $\mathbf{x} \in D$, given the expectation function $a_o(\mathbf{x})$, the covariance function $R(\boldsymbol{\tau})$ and the error-control parameters μ and μ^* , the corresponding F-K-L representation can be obtained with the following algorithm:

Algorithm 4.1 The F-K-L representation for a single elastic parameter of ERM

1. Find the minimum n -interval $\tilde{\mathbf{T}}^n \subset \mathbb{R}^n$ such that $D \subset \tilde{\mathbf{T}}^n$.
2. Determine the boundary of \mathbf{M}^n according to μ and Eqs. (4.17, 4.30-4.31).
3. Form the covariance matrix \mathbf{G} according to Eqs. (4.43-4.46).
4. Solve the standard eigenvalue problem with respect to \mathbf{G} .
5. Construct the F-K-L representation according to μ^* and Eqs. (4.35, 4.37-4.38).

It is worth mentioning that for an irregular domain D , the above F-K-L representation constructed via the n -interval $\tilde{\mathbf{T}}^n$ is not equivalent to performing an exact K-L expansion of $a(\mathbf{x}, \omega)$ $\mathbf{x} \in D$. Hence, in terms of the number of random variables contained in the

series expression, the F-K-L representation may be less economical than the exact K-L expansion depending on the specific domain D .

4.5 Numerical Examples

In this section, the overall performance of the F-K-L representation scheme for a single elastic parameter of ERM is examined through two examples. Without loss of generality, it is assumed that $a_o(x) \equiv 0$ in these examples.

4.5.1 One-Dimensional Examples

Two stochastic fields, which are both widely used in the description of random material properties, are considered.

- Case I: $a_{1D}^1(x, \omega)$ $x \in [-t, t]$ with covariance function $e^{-|t|/a_1}$ $a_1 > 0$;
- Case II: $a_{1D}^2(x, \omega)$ $x \in [-t, t]$ with covariance function e^{-x^2/a_2} $a_2 > 0$.

4.5.1.1 Case I

Although $a_{1D}^1(x, \omega)$ has been widely used in the SFEM literature (e.g. [4.3, 4.4, 4.6-4.7, 4.12-4.13, 4.15-4.18]), this thesis does not advocate its use in describing material properties of random media due to the discussions regarding continuity and differentiability in Chapter 3. The SOLE objective of choosing $a_{1D}^1(x, \omega)$ is to investigate the accuracy of the proposed F-K-L representation scheme.

Assuming $t = 65$ and $a_1 = 1.25$, the effective correlation length of $a_{1D}^1(x, \omega)$ is 5.75. The K-L expansion of $a_{1D}^1(x, \omega)$ is respectively obtained with the exact solution (see Appendix 4F), the finite element method (see e.g. [4.13, 4.17]) and the proposed F-K-L representation scheme. Figure 4.1 plots $\sum_{i=1}^N \lambda_i / \sum_{i=1}^{+\infty} \lambda_i$ against N , in which

$\sum_{i=1}^{+\infty} \lambda_i = 130$ and λ_i are calculated from the exact solution. It is observed that the K-L

expansion requires the first $N = 100$ eigenpairs to approximate 80% of the total variance of $a_{1D}^1(x, \omega)$, i.e. $\sum_{i=1}^N \lambda_i / \sum_{i=1}^{+\infty} \lambda_i = 0.8$. Indeed, the number of terms contained in a reasonable K-L expansion of a practical random medium can easily reach a few hundred, even in 1D problems. This is because the ratio of the random medium size over the effective correlation length can be even larger than that in this simple example. Hence, in order to achieve a reasonable K-L expansion of the stochastic field concerned, it is important to accurately obtain all the required eigenvalues and eigenfunctions instead of only the first few eigenpairs.

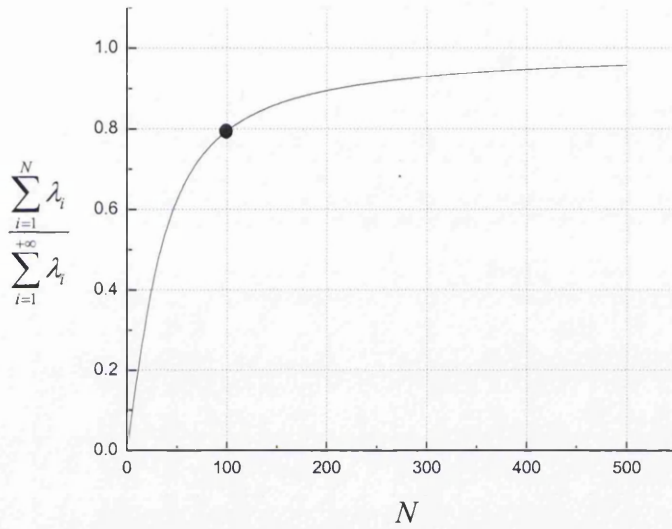


Figure 4.1 Asymptotic property of eigenvalues (Case I)

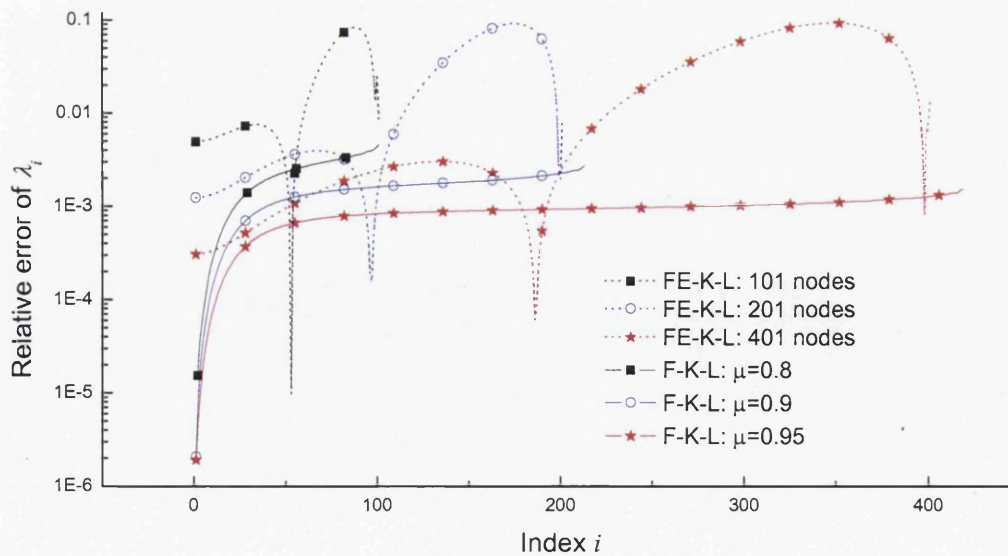


Figure 4.2 Relative errors of eigenvalues (Case I)

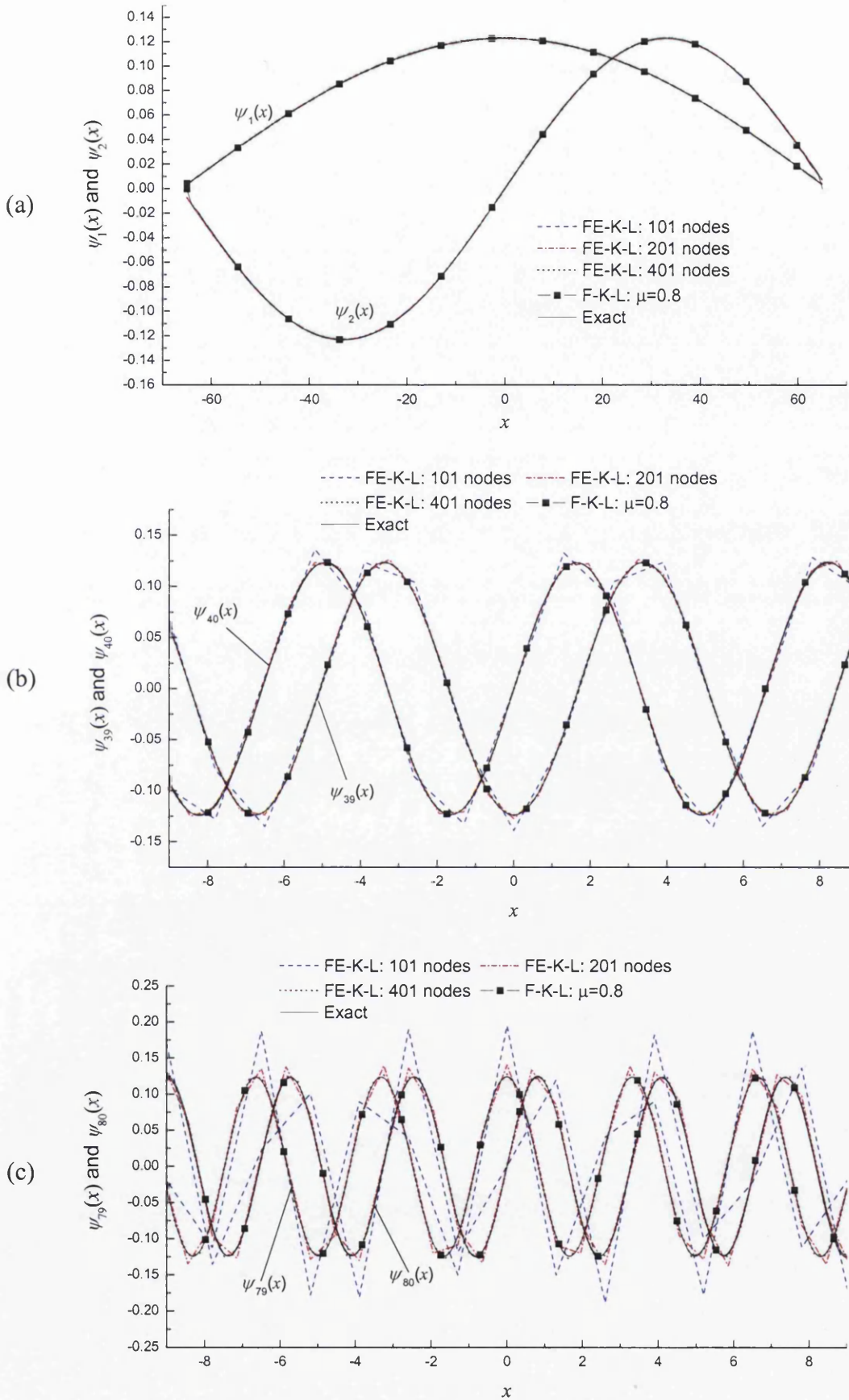


Figure 4.3 Comparison of eigenfunctions (Case I)

The relative errors of K-L eigenvalues are compared in Figure 4.2, in which the curves labelled F-K-L are obtained with the F-K-L representation scheme and the curves labelled FE-K-L are obtained with the K-L expansion method based on FE meshes. It is observed that the eigenvalue accuracy of the F-K-L representation scheme is much higher than that of the K-L expansion method. Specifically, in the F-K-L scheme, the approximation error is explicitly controlled by μ and all the eigenvalues are obtained with good accuracy; however, in the FE-K-L scheme, only half of the eigenvalues are obtained with reasonable accuracy.

Figure 4.3 compares some eigenfunctions of $a_{1D}^1(x, \omega)$, and as expected the F-K-L scheme shows a significant advantage over the K-L expansion method based on FE meshes. Although the accuracy of the K-L expansion method can be improved by employing more elements, it provides relatively poor results for higher order eigenfunctions. In the F-K-L solution with $\mu = 0.8$, the dimensionality of the covariance matrix \mathbf{G} is 100; however, for the eigenfunctions shown in Figure 4.3, there is no visible difference between the associated F-K-L solution and the exact solution.

4.5.1.2 Case II

The aim of this particular example is to investigate the influences of the random medium size and the effective correlation length on the F-K-L representation scheme. Letting $a_2 = 2$, the effective correlation length of $a_{1D}^2(x, \omega)$ is 3.0. Consider $a_{1D}^2(x, \omega)$ defined in $[-1.5r, 1.5r]$, where $r \in \{5, 10, 15, 20\}$ denotes the ratio of the interval size over the effective correlation length.

The K-L expansion of $a_{1D}^2(x, \omega)$ is obtained with the F-K-L representation scheme with $\mu = 0.95$. The diagonally dominant property of \mathbf{G} is verified in Figure 4.4, where the matrix elements are linearly mapped into pixels whose greyscales represent the values of the corresponding entries. It is observed that the larger the ratio r the more diagonally-dominant \mathbf{G} is.

Figure 4.5 plots the eigenvalue decay rate λ_i/λ_1 , in which scatters \blacksquare \square \blacktriangle \triangle are directly calculated from the eigenvalues, and the solid and dashed curves are predicted by

using respectively (4.56) and (4.57). As expected, the larger the ratio r the slower the eigenvalues decay. Both the spectral density function prediction and the diagonal entry prediction provide good estimates of the eigenvalue decay rate.

To examine the diagonal F-K-L approximations of $a_{1D}^2(x, \omega)$, Figure 4.6 compares their covariance functions with the exact covariance function $e^{-\tau^2/2}$, and reasonable agreements are achieved in $\tau \in [-1.5r, 1.5r]$.

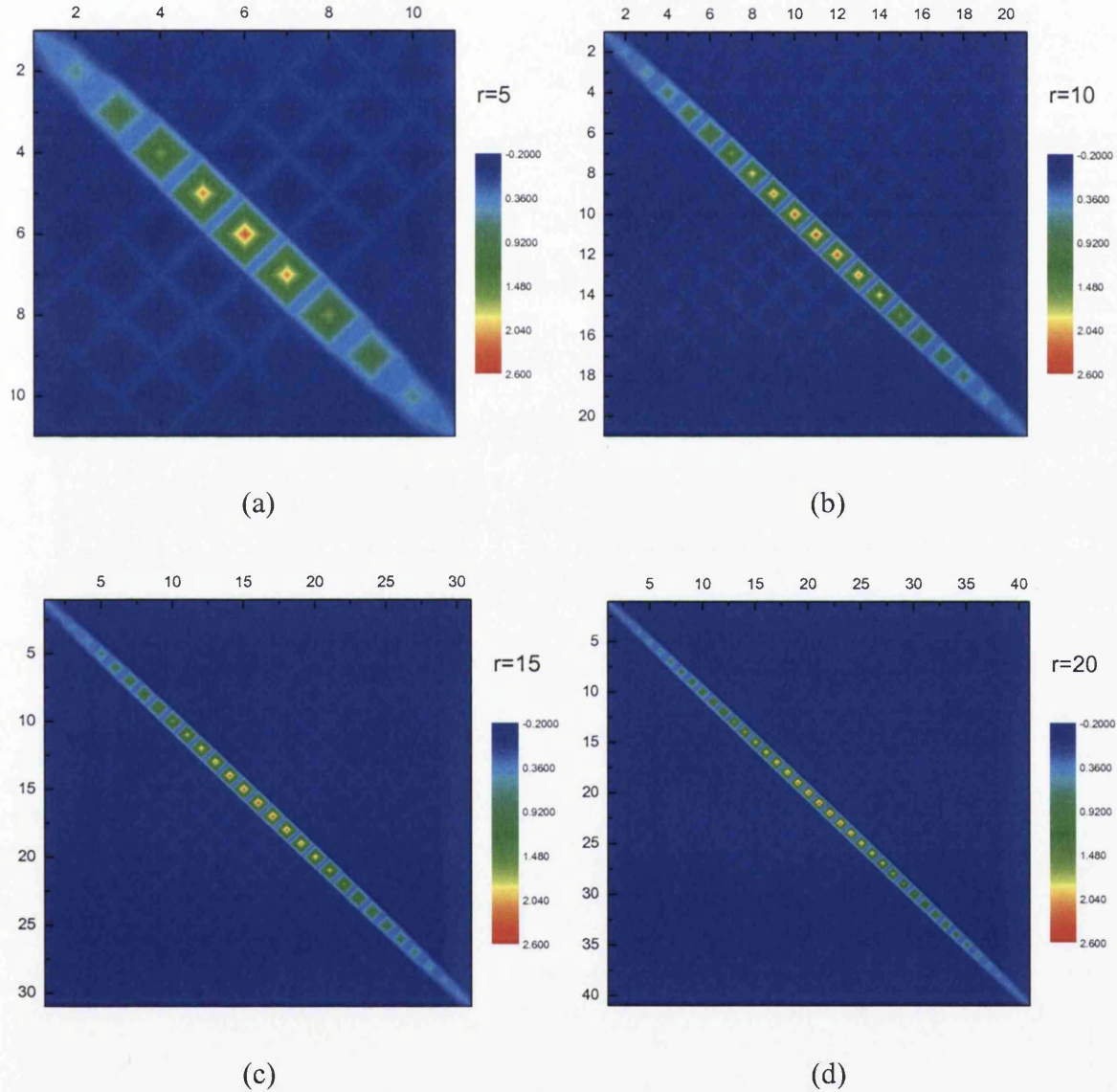


Figure 4.4 The covariance matrix G (Case II)

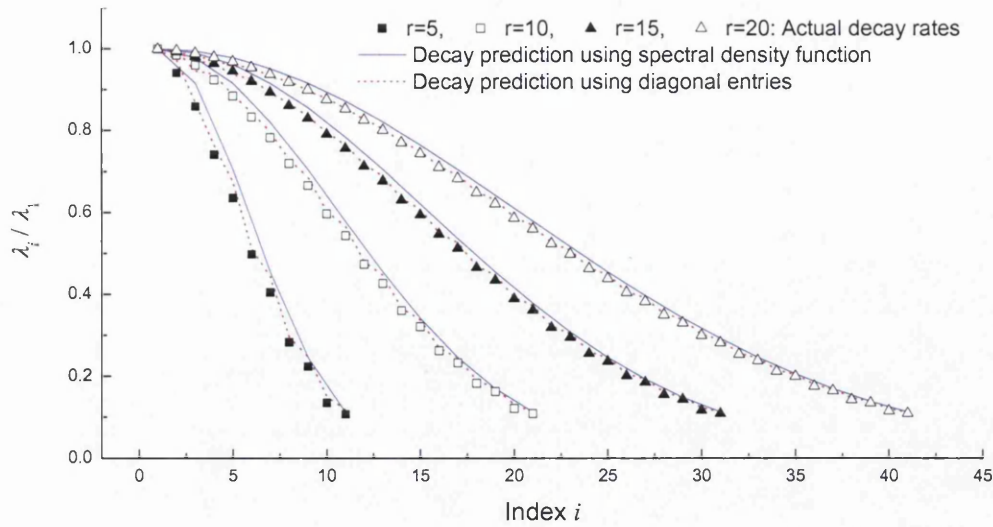


Figure 4.5 Eigenvalue decay rates (Case II)

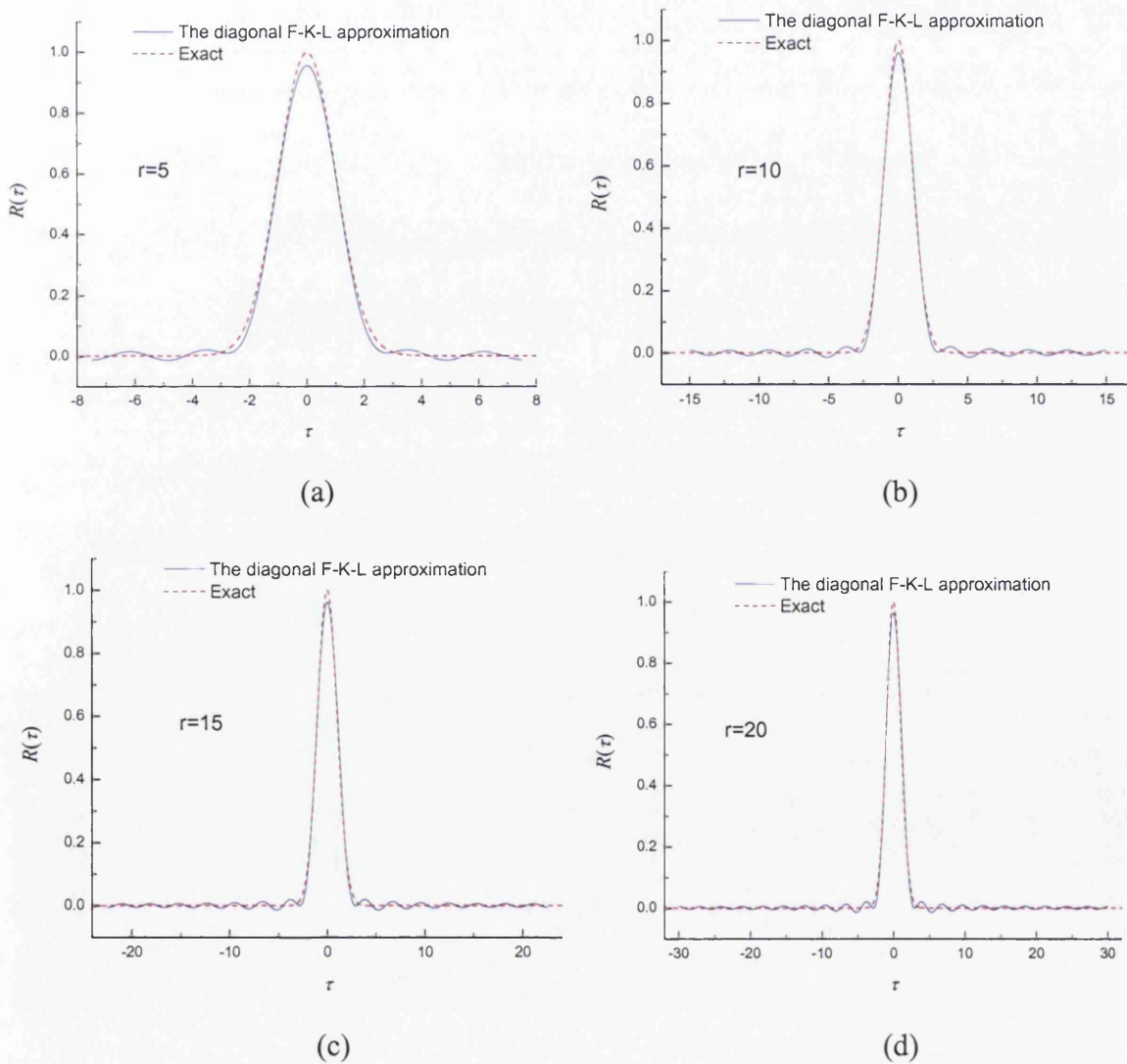


Figure 4.6 Accuracy of the diagonal F-K-L approximation scheme (Case II)

4.5.2 A Two-Dimensional Example

This example considers the stochastic field $a_{2D}(\mathbf{x}, \omega)$ $\mathbf{x} \in [-15, 15] \times [-15, 15]$ with covariance function $e^{-(\tau_1^2/3 + \tau_2^2/4)^{2/3}}$ whose effective correlation length is 6.3.

As the covariance function is nonseparable, there is no exact K-L solution available for $a_{2D}(\mathbf{x}, \omega)$. Hence, only the F-K-L representation is constructed for this general stochastic field. By setting $\mu = 0.95$ in the F-K-L representation scheme, there are around 2,800 K-L eigenpairs obtained for $a_{2D}(\mathbf{x}, \omega)$. In a PC system with an Intel Xeon 2.4 GHz processor and 1.0 GB DDR memory, it takes 23.6 seconds to construct the diagonally-dominant covariance matrix \mathbf{G} , and with the standard MATLAB eigensolver it takes 216.2 seconds to solve for the entire algebraic eigenstructure of \mathbf{G} .

Figures 4.7 and 4.8 plot respectively the eigenvalues and eigenfunctions, and Figure 4.9 shows a particular realization of $a_{2D}(\mathbf{x}, \omega)$, which is generated from the F-K-L representation obtained with $\mu^* = 0.8$.

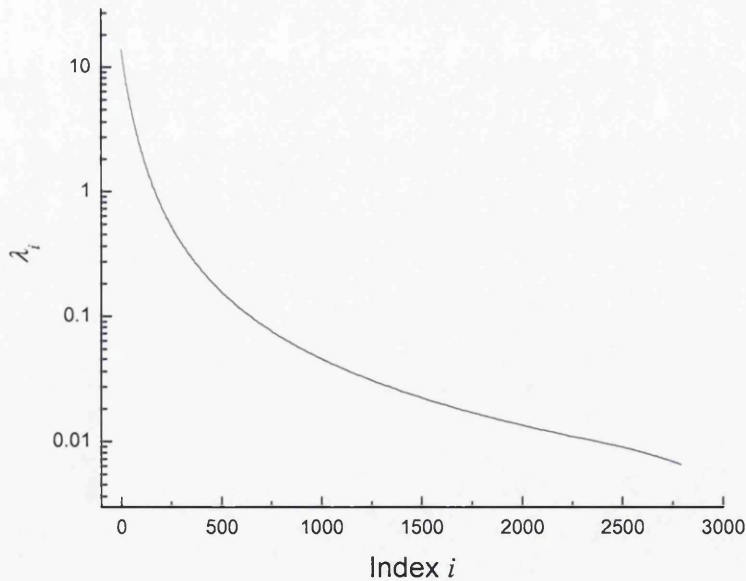


Figure 4.7 The eigenvalues of $a_{2D}(\mathbf{x}, \omega)$

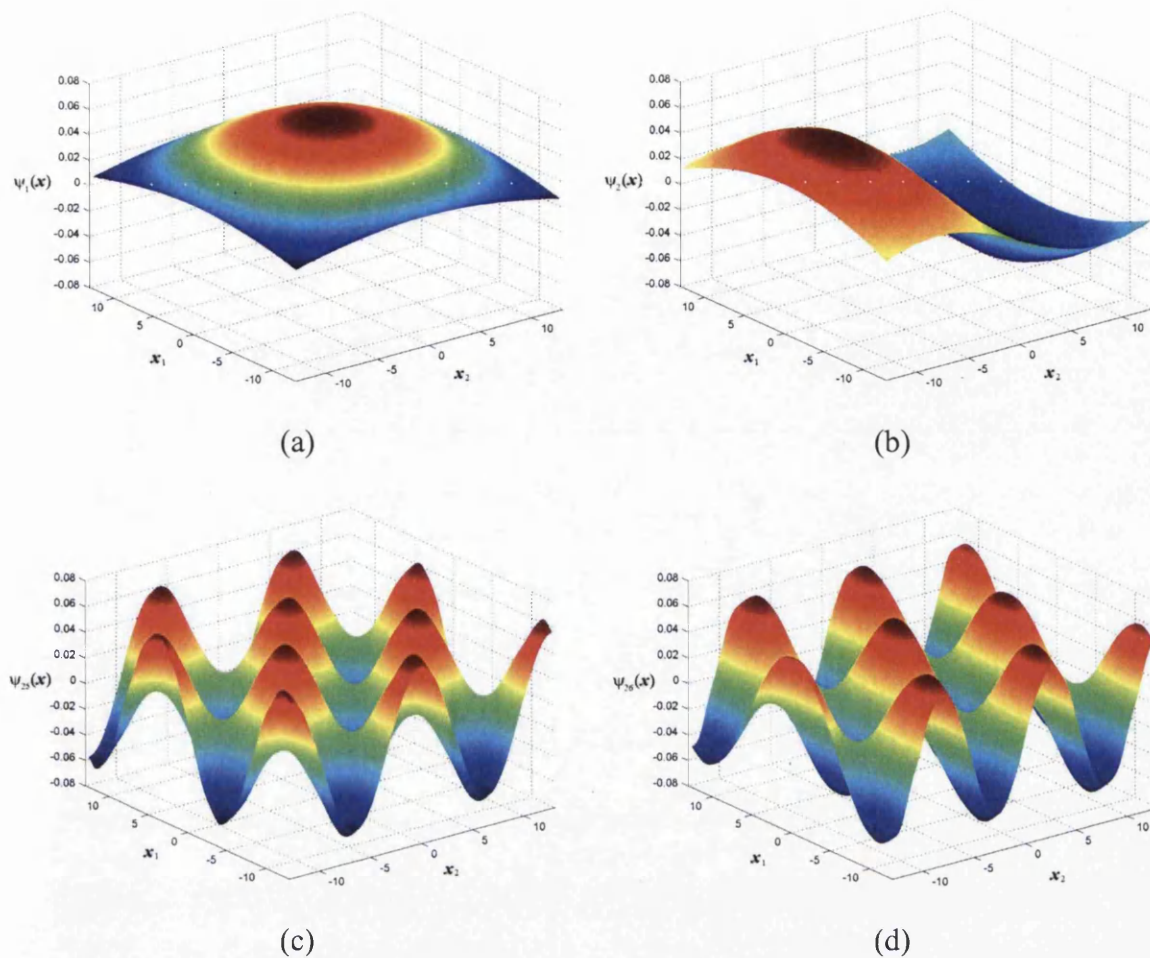


Figure 4.8 The eigenfunctions of $a_{2D}(\mathbf{x}, \omega)$

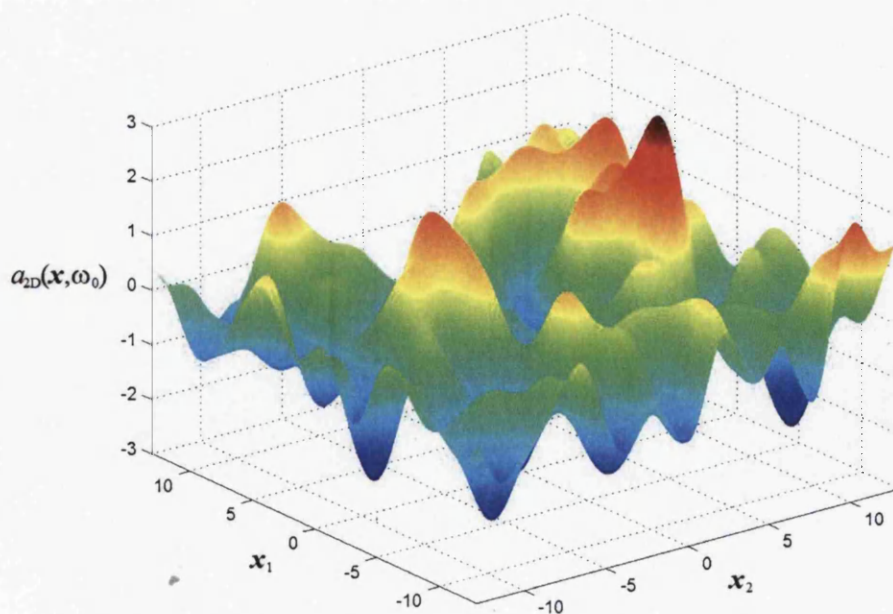


Figure 4.9 A particular realization of $a_{2D}(\mathbf{x}, \omega)$

4.6 The Fourier-Karhunen-Loève Representation for the General Elastic Tensor of Elementary Random Media

4.6.1 Background

In Sections 4.3-4.5, the F-K-L representation is developed for a single elastic parameter of ERM. However, this is rarely the exact situation in reality since a practical random medium system usually contains multiple random material parameters. Unfortunately, it is not straightforward to apply the previously developed F-K-L representation scheme to describe multiple random material properties, unless these properties are linearly dependent to each other, for which the multivariate stochastic-field representation problem can be transformed into a simple problem with respect to a single scalar stochastic field.

On the other hand, up to now [4.1-4.18], it appears that there has not been any general technique reported to effectively represent multiple random material properties. More specifically, as pointed out in [4.18], the random Young's modulus is mainly dealt with in SFEM research, and very few works are dedicated to dealing with a random Poisson's ratio.

Hence, the aim of this section is to establish the F-K-L representation scheme for the general elastic tensor of ERM, which fully takes into account the interactions between different random elastic parameters.

4.6.2 The Solution

Recall from the introduction in the beginning of this chapter that, $C(x, \omega)$ $x \in D, \omega \in \Omega$ denotes the elastic tensor of an ERM $D \subset \mathbb{R}^n$, which is defined by its expectation function $C_o(x, \omega)$ (4.1) and covariance function $R(\tau)$ (4.2). The aim is to obtain the F-K-L representation of $C(x, \omega)$ in the form of (4.3).

Let

$$\mathbf{c}(\mathbf{x}, \omega) = (c_1(\mathbf{x}, \omega), c_2(\mathbf{x}, \omega), \dots, c_p(\mathbf{x}, \omega))^T, \quad (4.60)$$

denote the vector constructed from the scalar entries of $\mathbf{C}(\mathbf{x}, \omega)$ via the one-to-one mapping

$$\mathbf{C}(\mathbf{x}, \omega) \leftrightarrow \mathbf{c}(\mathbf{x}, \omega). \quad (4.61)$$

The expectation function of $\mathbf{c}(\mathbf{x}, \omega)$ is

$$\mathbf{c}_o(\mathbf{x}) = (c_{o,1}(\mathbf{x}), c_{o,2}(\mathbf{x}), \dots, c_{o,p}(\mathbf{x}))^T, \quad (4.62)$$

which due to the mapping (4.61) is known from $\mathbf{C}_o(\mathbf{x}, \omega)$. Denoted by the matrix-valued function $\mathbf{R}(\boldsymbol{\tau})$ (4.2), the covariance functions of $\mathbf{c}(\mathbf{x}, \omega)$ and $\mathbf{C}(\mathbf{x}, \omega)$ are the same.

It is simpler in notation to derive the F-K-L representation for the vector $\mathbf{c}(\mathbf{x}, \omega)$, and due to the one-to-one mapping (4.61), this is equivalent to investigating the tensor $\mathbf{C}(\mathbf{x}, \omega)$.

Let \mathbf{T}^n defined as (4.19) denote the minimum n -interval such that $D \subset \mathbf{T}^n$. Then, following the discussions in Section 4.3.2.1, each entry of $\mathbf{c}(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbf{T}^n$ can be approximated by a finite Fourier series as

$$\begin{aligned} c_i(\mathbf{x}, \omega) &\approx c_{o,i}(\mathbf{x}) + \sum_{m_1=-M_{i,1}}^{M_{i,1}} \dots \sum_{m_n=-M_{i,n}}^{M_{i,n}} \Delta Z_i(\mathbf{m}, \omega) e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} \quad i=1,2,\dots,p, \quad (4.63) \\ &= c_{o,i}(\mathbf{x}) + (\Delta \mathbf{Z}_i(\omega))^T \mathbf{e}_i(\mathbf{x}) \end{aligned}$$

in which $e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}$ are the Fourier basis functions defined on \mathbf{T}^n , $\Delta Z_i(\mathbf{m}, \omega)$ are the Fourier coefficients given by

$$\Delta Z_i(\mathbf{m}, \omega) = \frac{1}{V_{\mathbf{T}^n}} \int_{\mathbf{T}^n} (c_i(\mathbf{x}, \omega) - c_{o,i}(\mathbf{x})) e^{-\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} d\mathbf{x}; \quad (4.64)$$

the discrete spectrum boundaries $M_i = (M_{i,1}, \dots, M_{i,n})$ can be similarly determined as (4.30-4.31) according to the given error control parameter μ ; and the vectors

$$\Delta \mathbf{Z}_i(\omega) = (\Delta Z_{i,1}(\omega), \Delta Z_{i,2}(\omega), \dots, \Delta Z_{i,N_i}(\omega))^T \quad (4.65)$$

and

$$\mathbf{e}_i(\mathbf{x}) = (e_{i,1}(\mathbf{x}), e_{i,2}(\mathbf{x}), \dots, e_{i,N_i}(\mathbf{x}))^T \quad (4.66)$$

are respectively constructed from $\sqrt{V_{T^n}} \Delta Z_i(\mathbf{m}, \omega)$ and $e^{\frac{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}{\sqrt{V_{T^n}}}}$.

Consider the following row vectors

$$\Delta \mathbf{Z}(\omega) \triangleq \left((\Delta \mathbf{Z}_1(\omega))^T, (\Delta \mathbf{Z}_2(\omega))^T, \dots, (\Delta \mathbf{Z}_p(\omega))^T \right)^T, \quad (4.67)$$

$$\mathbf{e}(\mathbf{x}) \triangleq \left((\mathbf{e}_1(\mathbf{x}))^T, (\mathbf{e}_2(\mathbf{x}))^T, \dots, (\mathbf{e}_p(\mathbf{x}))^T \right)^T, \quad (4.68)$$

whose dimensionality is

$$N = \sum_{i=1}^p N_i. \quad (4.69)$$

It is clear that $\Delta \mathbf{Z}(\omega)$ is the projection coordinate of $\mathbf{c}(\mathbf{x}, \omega) - \mathbf{c}_o(\mathbf{x})$ on the Fourier axis $\mathbf{e}(\mathbf{x})$. Naturally, the dimensionality of $\Delta \mathbf{Z}(\omega)$ can be reduced by PCA.

Specifically, the covariance matrix of $\Delta \mathbf{Z}(\omega)$ has the following form

$$\mathbf{G} = E \left(\Delta \mathbf{Z}(\omega) (\Delta \mathbf{Z}(\omega))^H \right) = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} & \cdots & \mathbf{G}_{1p} \\ \mathbf{G}_{21} & \mathbf{G}_{22} & \cdots & \mathbf{G}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_{p1} & \mathbf{G}_{p2} & \cdots & \mathbf{G}_{pp} \end{pmatrix}, \quad (4.70)$$

where

$$\mathbf{G}_{ij} = E \left(\Delta \mathbf{Z}_i(\omega) (\Delta \mathbf{Z}_j(\omega))^H \right) \quad \text{for every } i, j = 1, 2, \dots, p \quad (4.71)$$

is again a covariance matrix, of which the dimensionality is $N_i \times N_j$. Following Eq. (4.64), each such covariance matrix \mathbf{G}_{ij} is defined as,

$$\begin{aligned} (\mathbf{G}_{ij})_{kl} &= \text{Cov}(\Delta Z_{i,k}(\omega), \Delta Z_{j,l}(\omega)) & 1 \leq k \leq N_i \\ &= \int_{T^n} \int_{T^n} R_{ij}(\mathbf{x}_1 - \mathbf{x}_2) \overline{e_{i,k}(\mathbf{x}_1)} e_{j,l}(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 & 1 \leq l \leq N_j \end{aligned} \quad (4.72)$$

in which $(\mathbf{G}_{ij})_{kl}$ is the entry located at the k -th row and the l -th column of \mathbf{G}_{ij} , and

$$R_{ij}(\mathbf{x}_1 - \mathbf{x}_2) = \text{Cov}(c_i(\mathbf{x}_1, \omega), c_j(\mathbf{x}_2, \omega)) \quad (4.73)$$

is the entry located at the i -th row and j -th column of the matrix-valued covariance

function $R(\boldsymbol{\tau})$.

As the covariance matrix of the random vector $\Delta\mathbf{Z}(\omega)$, \mathbf{G} is a non-negative definite Hermitian matrix[†]. Hence, there exists a unitary matrix $\mathbf{Q}^{-1} = \mathbf{Q}^H$ such that

$$\mathbf{Q}^H \mathbf{G} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N), \quad (4.74)$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$. Consequently, the random vector $\Delta\mathbf{Z}(\omega)$ can be written as

$$\begin{aligned} \Delta\mathbf{Z}(\omega) &= \left(\Delta Z_{1,1}(\omega), \dots, \Delta Z_{1,N_1}(\omega), \dots, \Delta Z_{p,1}(\omega), \dots, \Delta Z_{p,N_p}(\omega) \right)^T \\ &= \mathbf{Q} \left(\sqrt{\lambda_1} \xi_1(\omega), \sqrt{\lambda_2} \xi_2(\omega), \dots, \sqrt{\lambda_N} \xi_N(\omega) \right)^T \end{aligned} \quad (4.75)$$

in which $\xi_i(\omega)$ $i = 1, 2, \dots, N$ are stochastically independent standard Gaussian random variables. Let N^* be the smallest integer such that

$$\frac{\sum_{i=1}^{N^*} \lambda_i}{\sum_{i=1}^N \lambda_i} \geq \mu^* \quad 0 < \mu^* \leq 1, \quad (4.76)$$

where μ^* is the required accuracy for approximating the total variance of $c(\mathbf{x}, \omega)$. Then, from (4.75), the dimensionality of $\Delta\mathbf{Z}(\omega)$ can be reduced to N^* such that each entry in $\Delta\mathbf{Z}(\omega)$ has the following form

$$\Delta Z_{i,k}(\omega) \approx \sum_{q=1}^{N^*} Q_{i,k,q} \sqrt{\lambda_q} \xi_q(\omega) \quad \begin{array}{l} i = 1, 2, \dots, p \\ k = 1, 2, \dots, N_i \end{array} \quad (4.77)$$

in which $Q_{i,k,q}$ are corresponding constant entries in \mathbf{Q} .

Substituting (4.77) into (4.63) yields

$$\begin{aligned} c_i(\mathbf{x}, \omega) &\approx c_{o,i}(\mathbf{x}) + \sum_{k=1}^{N_i} \Delta Z_{i,k}(\omega) e_{i,k}(\mathbf{x}) \\ &\approx c_{o,i}(\mathbf{x}) + \sum_{k=1}^{N_i} \sum_{q=1}^{N^*} Q_{i,k,q} \sqrt{\lambda_q} \xi_q(\omega) e_{i,k}(\mathbf{x}) \quad i = 1, 2, \dots, p, \\ &= c_{o,i}(\mathbf{x}) + \sum_{q=1}^{N^*} \sqrt{\lambda_q} \xi_q(\omega) \psi_{i,q}(\mathbf{x}) \end{aligned} \quad (4.78)$$

where

[†] In fact, following a similar analysis as Section 4.4.1, it can be proven that the covariance matrix \mathbf{G} defined in (4.70) is real-symmetric and non-negative definite.

$$\psi_{i,q}(\mathbf{x}) = \sum_{k=1}^{N_i} Q_{i,k,q} e_{i,k}(\mathbf{x}). \quad (4.79)$$

It should be noted that, unlike the scalar case discussed in Sections 4.2-4.5, the basis functions $\psi_{i,q}(\mathbf{x})$ defined above are not necessarily orthogonal to each other. Due to the mapping (4.61), Eq. (4.78) indicates that the elastic tensor $\mathbf{C}(\mathbf{x}, \omega)$ can be represented as

$$\mathbf{C}(\mathbf{x}, \omega) \approx \mathbf{C}_o(\mathbf{x}) + \sum_{q=1}^{N^*} \sqrt{\lambda_q} \xi_q(\omega) \mathbf{C}_q(\mathbf{x}), \quad (4.80)$$

in which $\mathbf{C}_q(\mathbf{x})$ is a tensor-valued function constructed from $\psi_{i,q}(\mathbf{x})$ (4.79).

Eq. (4.80) is termed the *F-K-L representation* for the general elastic tensor of ERM. The associated F-K-L algorithm is formally the same as Algorithm 4.1, but the calculation at each algorithmic step should follow the corresponding formulation developed in this section. For the sake of simplicity, the dependent scalar entries in the fourth-order tensor $\mathbf{C}(\mathbf{x}, \omega)$ are not distinguished in the above developments. However, it is well known in elasticity that there are maximum 21 (rather than 81) independent entries in the elastic tensor, and for the simplest cases there are only two independent entries. Consequently, to be computationally efficient in practical random medium modelling, it is only necessary to construct the vector $\mathbf{c}(\mathbf{x}, \omega)$ with those independent non-zero scalar entries in $\mathbf{C}(\mathbf{x}, \omega)$, from which the F-K-L representation for the whole elastic tensor can be readily constructed.

4.7 Summary and Suggestions for Future Research

4.7.1 Summary

In this chapter, an explicit stochastic-field representation scheme, namely the F-K-L representation, is established to describe the random material properties of the ERM model. The main features of the proposed scheme are listed as follows:

- The F-K-L representation scheme is developed for the general elastic tensor of ERM, which makes the ERM model applicable to any elastic constitutive relation.
- In contrast to various FE-mesh based stochastic-field representation schemes, the

F-K-L representation scheme is completely meshfree and also independent of the detailed shape of the random structure under consideration.

- When there is only one random material parameter involved in the ERM model, the F-K-L representation provides a semi-analytical solution for the K-L expansion of the corresponding stochastic field. Comparing with the widely used K-L expansion method based on FE meshes, the F-K-L representation scheme is not only more accurate but is also computationally more efficient. This is due to the harmonic essence of wide-sense stationary stochastic fields.
- The F-K-L representation is achieved with a prior error control. Specifically, the approximation error is explicitly controlled by two parameters, which control respectively the error of the spectrum of stochastic fields and the error of the total variance.
- In the F-K-L representation scheme, the covariance matrix \mathbf{G} can be obtained semi-analytically by using Fourier transforms. The associated theoretical formulation of \mathbf{G} not only significantly improves the computational efficiency of the F-K-L solution, but also results in some interesting conclusions regarding ERM, which include the diagonally-dominant property of \mathbf{G} , the diagonal F-K-L approximation scheme and the eigenvalue decay prediction.

Developed in Section 4.4.1, the specific quadrature method for multi-dimensional oscillatory functions is the other contribution made in this chapter. With this novel quadrature technique, the computational cost of constructing the matrix \mathbf{G} is significantly reduced (up to several orders of magnitude). It is clearly possible to apply this technique for the quadrature of more general multi-dimensional oscillatory functions, and this will be addressed in [4.33].

Parts of the results in this chapter have already been reported in [4.36-4.38].

4.7.2 Suggestions for Future Research

Regarding the F-K-L representation scheme presented in this chapter, the following aspects need further investigation.

- It is worth mentioning that for an ERM described by $a(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{T}^n$, some eigenfunctions obtained by the F-K-L representation scheme do not converge to the

exact eigenfunctions near $\partial\mathbf{T}^n$, i.e. the boundary of \mathbf{T}^n . This is due to the Gibbs phenomenon [4.29] where some eigenfunctions of $a(\mathbf{x}, \omega)$ have discontinuities on $\partial\mathbf{T}^n$. This adverse phenomenon can be overcome by adjusting the Fourier summation with the Lanczos sigma factor [4.31] or slightly extending the computational domain of \mathbf{T}^n .

- The success of the F-K-L representation is essentially due to the harmonic essence of wide-sense stationary stochastic fields. Hence, within the same framework, it is interesting to investigate further applications of the theory of harmonic analysis, especially wavelet theory, in the stochastic-field representation scheme. The wavelet theory has been firmly established as a very powerful transformation tool, alternative to Fourier analysis, for many problems and predominately for signal/image processing and data analysis. The most attractive feature of wavelets is their dual localisation in both spatial and frequency domains. It is expected that at least two benefits can be gained from the wavelet representation: (a) A faster convergence than the Fourier representation is expected to be achieved, leading to a smaller number of basis (wavelet) functions needed; and (b) A successful wavelet representation will provide a natural route to considering multi-scale effects in random media. In this phase of work, several different sets of wavelet functions, with different local compactness and smoothness, will be investigated to identify the most suitable set for the problem considered.
- In Section 4.6, the F-K-L representation is established for the elastic tensor of ERM, however, for this general case, its numerical performance has not been fully examined. A possible failing may arise in the F-K-L representation scheme in that, different elastic parameters are equally considered in the covariance matrix \mathbf{G} . Specifically, the random values of some elastic parameters may differ in order of magnitude, and when this difference is very large the matrix \mathbf{G} may become ill conditioned, which in turn makes the solution difficult. This potential problem can be solved by replacing in the F-K-L algorithm the covariance matrix \mathbf{G} with the correlation matrix \mathbf{H} . In contrast to Eq. (4.72), each entry in \mathbf{H} is a normalized correlation coefficient of the form

$$\left(\mathbf{H}_{ij}\right)_{kl} = \frac{\text{Cov}(\Delta Z_{i,k}(\omega), \Delta Z_{j,l}(\omega))}{\sqrt{\text{Var}(\Delta Z_{i,k}(\omega))} \sqrt{\text{Var}(\Delta Z_{j,l}(\omega))}} \in [0,1] \quad \begin{matrix} 1 \leq k \leq N_i \\ 1 \leq l \leq N_j \end{matrix} \quad (4.81)$$

Other technical defects might also exist in the F-K-L representation of the general elastic tensor which may not become apparent until a comprehensive numerical examination is completed.

- Recall from Section 3.4.2.1 that a more general definition of the ERM model can be achieved by replacing the Gaussian assumption with an assumption of ergodicity (3.62). It is seen in Sections 4.3-4.6 that the only use of the ERM Gaussian assumption in the F-K-L representation scheme is to determine the probability distribution of the random sequence $\xi_i(\omega)$ (4.37, 4.80). Consequently, with respect to the F-K-L representation, two key questions arise from the alternative ERM definition: (a) Without the Gaussian assumption, whether or not it is possible to properly define $\xi_i(\omega)$ from the ergodicity assumption; (b) If it is possible to do so, then how to approximately determine the joint probability distribution of random variables $\xi_i(\omega)$.

The ERM model and its F-K-L representation will be reported in [4.34, 4.35] together with some of the above research which is currently under development.

Appendix 4A The Karhunen-Loève Expansion Theorem

This appendix explains the mathematical background of K-L expansions.

4A.1 Mercer's Theorem

Definition 4A.1 Let $D \subset \mathbb{R}^n$ denote a Lebesgue measurable set. A complex-valued function $R(\mathbf{x}_1, \mathbf{x}_2)$ $\mathbf{x}_1, \mathbf{x}_2 \in D$ is said to be symmetric and non-negative definite in D if

$$R(\mathbf{x}_1, \mathbf{x}_2) = \overline{R(\mathbf{x}_2, \mathbf{x}_1)} \quad \forall (\mathbf{x}_1, \mathbf{x}_2) \in D \times D \quad (4A.1)$$

and

$$\sum_{i=1}^k \sum_{j=1}^k R(\mathbf{x}_i, \mathbf{x}_j) z_i \overline{z_j} \geq 0 \quad (4A.2)$$

for all finite sequences of points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ in D and all choices of complex numbers

z_1, z_2, \dots, z_k .

Theorem 4A.2 (Mercer's theorem) Let D denote a compact set in \mathbb{R}^n , and $R(\mathbf{x}_1, \mathbf{x}_2)$ $\mathbf{x}_1, \mathbf{x}_2 \in D$ denote a complex-valued function. If $R(\mathbf{x}_1, \mathbf{x}_2)$ is continuous, symmetric and non-negative definite in D , then there exists a set of functions $\{\psi_i(\mathbf{x})\}_{i \in \mathbb{N}}$ in $L^2(D)$ satisfying

$$\int_D \psi_i(\mathbf{x}) \overline{\psi_j(\mathbf{x})} d\mathbf{x} = \delta_{ij} \quad (4A.3)$$

and the characteristic equation

$$\int_D R(\mathbf{x}_1, \mathbf{x}_2) \psi(\mathbf{x}_1) d\mathbf{x}_1 = \lambda \psi(\mathbf{x}_2) \quad (4A.4)$$

where the eigenvalues $\{\lambda_i\}_{i \in \mathbb{N}}$ are non-negative and the eigenfunctions $\{\psi_i(\mathbf{x})\}_{i \in \mathbb{N}}$ corresponding to those non-zero eigenvalues are continuous in D . Furthermore, the integral kernel $R(\mathbf{x}_1, \mathbf{x}_2)$ has the following representation

$$R(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=1}^{+\infty} \lambda_i \psi_i(\mathbf{x}_1) \overline{\psi_i(\mathbf{x}_2)} \quad (4A.5)$$

whose convergence is absolute and uniform. In particular,

$$\text{Tr}(R(\mathbf{x}_1, \mathbf{x}_2)) \triangleq \int_D R(\mathbf{x}, \mathbf{x}) d\mathbf{x} = \sum_{i=1}^{+\infty} \lambda_i \quad (4A.6)$$

is termed the trace of $R(\mathbf{x}_1, \mathbf{x}_2)$.

Mercer's theorem is one of the most important theoretical tools in the theory of integral equations, and its proof can be found in [4.21-4.23].

4A.2 The Karhunen-Loève Expansion of Second-Order Stochastic Fields

In the Hilbert space theory of stochastic fields, the K-L expansion is a technique that can be used to simplify a stochastic field; more formally it is a transformation that chooses a new set of basis functions for the stochastic field such that the greatest variance by any projection of the stochastic field lies on the first basis-function axis (called the first principal component); the second greatest variance on the second axis, and so on. Assuming a zero mean for the complex-valued stochastic field $b(\mathbf{x}, \omega)$ $\mathbf{x} \in D, \omega \in \Omega$,

the first principal component $\psi_1(\mathbf{x})$ of $b(\mathbf{x}, \omega)$ is defined as :

$$\psi_1(\mathbf{x}) = \arg \max_{\int_D |\psi(\mathbf{x})|^2 d\mathbf{x} = 1} E \left(\left| \int_D b(\mathbf{x}, \omega) \overline{\psi(\mathbf{x})} d\mathbf{x} \right|^2 \right) \quad (4A.7)$$

where $\arg \max$ stands for the argument of the maximum. When the first $k-1$ components are defined, the k -th component can be found by subtracting the first $k-1$ principal components from $b(\mathbf{x}, \omega)$:

$$\hat{b}_{k-1}(\mathbf{x}, \omega) = b(\mathbf{x}, \omega) - \sum_{i=1}^{k-1} \psi_i(\mathbf{x}) \int_D b(\mathbf{x}, \omega) \overline{\psi_i(\mathbf{x})} d\mathbf{x}, \quad (4A.8)$$

and then by treating this as the new stochastic field to find its principal component

$$\psi_k(\mathbf{x}) = \arg \max_{\int_D |\psi(\mathbf{x})|^2 d\mathbf{x} = 1} E \left(\left| \int_D \hat{b}_{k-1}(\mathbf{x}, \omega) \overline{\psi(\mathbf{x})} d\mathbf{x} \right|^2 \right). \quad (4A.9)$$

The stochastic field $b(\mathbf{x}, \omega)$ then has the following representation

$$b(\mathbf{x}, \omega) = \sum_{i=1}^{+\infty} z_i(\omega) \psi_i(\mathbf{x}) \quad (4A.10)$$

where $z_i(\omega)$ are random coordinates with respect to $\psi_i(\mathbf{x})$ such that

$$z_i(\omega) = \int_D b(\mathbf{x}, \omega) \overline{\psi_i(\mathbf{x})} d\mathbf{x}. \quad (4A.11)$$

If the stochastic field $b(\mathbf{x}, \omega)$ is mean-square continuous, it is then trivial to verify against Definition 4A.1 that the covariance function $\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega))$ is continuous, symmetric and non-negative definite (see e.g. Section 2.2.3). Consequently, based on Mercer's theorem, a simpler formulation can be developed to obtain the principal components $\psi_i(\mathbf{x})$ defined in (4A.9), the random coordinates $z_i(\omega)$ defined in (4A.11) and the expansion (4A.10). This is known as the *K-L expansion theorem*.

Theorem 4A.3 (Karhunen-Loève expansion) Let D denote a compact set in \mathbb{R}^n , and $b(\mathbf{x}, \omega)$ $\mathbf{x} \in D, \omega \in \Omega$ a centred complex-valued second-order stochastic field. If $b(\mathbf{x}, \omega)$ is mean-square continuous, then it can be expressed as

$$b(\mathbf{x}, \omega) = \sum_{i=1}^{+\infty} z_i(\omega) \psi_i(\mathbf{x}) \quad (4A.12)$$

where the convergence is in mean square and is absolute and uniform; $z_i(\omega)$ are a

sequence of mutually uncorrelated complex-valued random variables satisfying

$$E\left(z_i(\omega)\overline{z_j(\omega)}\right) = \delta_{ij}\lambda_i; \quad (4A.13)$$

and λ_i and $\psi_i(\mathbf{x})$ are the positive eigenvalues and the corresponding orthonormal eigen-functions generated by the covariance function of $b(\mathbf{x}, \omega)$ according to Mercer's theorem. In particular, if the stochastic field $b(\mathbf{x}, \omega)$ is Gaussian, then $z_i(\omega)$ are stochastically independent Gaussian random variables and the convergence of (4A.12) becomes almost surely, absolute and uniform.

The proof of Theorem 4A.3 can be found in [4.19, 4.20]. From a mathematical point of view, the K-L expansion is essentially equivalent to principal component analysis [4.26], a well known technique for data compression and signal reconstruction in the context of pattern recognition and signal processing. The difference between the K-L expansion and principal component analysis is that, the former deals with continuous stochastic fields while the latter focuses mainly on discrete random sequences. Due to this mathematical similarity, principal component analysis is sometimes referred to as K-L transforms.

Since series (4A.12) is of absolute convergence, the sequence of $z_i(\omega)\psi_i(\mathbf{x})$ appearing in (4A.12) is arbitrary, and for the most efficient approximation, it can be chosen as the descending order of the corresponding eigenvalues $\{\lambda_i\}_{i \in \mathbb{N}}$. The number of terms needed in the approximate expansion, i.e. the partial sum of infinite series (4A.12), is determined by the trace relation (4A.6) and the required accuracy. For example, if the trace of a given covariance function is

$$\text{Tr}\left(\text{Cov}\left(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)\right)\right) = 10 \quad (4A.14)$$

and the associated eigenvalues are obtained as

$$\{\lambda_1, \lambda_2, \lambda_3, \dots\} = \{4, 3, 1.5, \dots\}, \quad (4A.15)$$

then an approximation of 70% accuracy can be readily achieved by truncating series (4A.12) from the second term, i.e.

$$b(\mathbf{x}, \omega) \approx z_1(\omega)\psi_1(\mathbf{x}) + z_2(\omega)\psi_2(\mathbf{x}) \quad (4A.16)$$

where $\psi_1(\mathbf{x})$ and $\psi_2(\mathbf{x})$ are the first two orthonormal eigenfunctions of the given

covariance function, and random variables $z_1(\omega)$ and $z_2(\omega)$ satisfy

$$\begin{aligned} E(|z_1(\omega)|^2) &= \lambda_1 = 4 & E(|z_2(\omega)|^2) &= \lambda_2 = 3 \\ E(z_1(\omega)\overline{z_2(\omega)}) &= 0 & E(\overline{z_1(\omega)}z_2(\omega)) &= 0 \end{aligned} \quad (4A.17)$$

Eq. (4A.16) is called a 70% approximation for $b(\mathbf{x}, \omega)$ due to the following calculation

$$\frac{\lambda_1 + \lambda_2}{\text{Tr}(\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)))} = \frac{4+3}{10} = 70\%. \quad (4A.18)$$

Appendix 4B The Effective Correlation Length

Consider a second-order stochastic field $b(\mathbf{x}, \omega)$ $\mathbf{x} \in D, \omega \in \Omega$ with the covariance function $\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega))$. In this thesis, the *effective correlation length* of $b(\mathbf{x}, \omega)$ is defined as the minimum value τ_c such that

$$\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_2, \omega)) \leq \frac{\text{Cov}(b(\mathbf{x}_1, \omega), b(\mathbf{x}_1, \omega))}{100} \quad (4B.1)$$

holds for all points $\mathbf{x}_1, \mathbf{x}_2 \in D$ satisfying

$$\|\mathbf{x}_1 - \mathbf{x}_2\|_2 > \tau_c. \quad (4B.2)$$

Appendix 4C The Spectral Representation Theory of Wide-Sense Stationary Stochastic Fields

The spectral representation theory of wide-sense stationary stochastic fields is outlined below, and detailed explanations and proofs can be found in [4.24, 4.25] and the references therein.

Theorem 4C.1 A continuous function $R(\boldsymbol{\tau})$ from \mathbb{R}^n to the complex plane is non-negative definite[†] if and only if it can be represented in the form

[†] See e.g. Section 2.1.2 for the exact definition.

$$R(\boldsymbol{\tau}) = \int_{\mathbb{R}^n} e^{\sqrt{-1}\boldsymbol{\tau} \cdot \mathbf{y}} dF(\mathbf{y}), \quad (4C.1)$$

where $F(\mathbf{y})$ is a bounded, real-valued function satisfying

$$\int_A dF(\mathbf{y}) \geq 0 \quad (4C.2)$$

for all Lebesgue measurable set $A \subset \mathbb{R}^n$.

The function $F(\mathbf{y})$ in the above theorem is only specified up to an additive constant. When $F(\mathbf{y})$ is fixed by demanding that

$$F(-\infty) = F(-\infty, \dots, -\infty) = 0 \quad (4C.3)$$

(thus implying that $F(+\infty) = F(+\infty, \dots, +\infty) = R(\mathbf{0})$), Eq. (4C.1) is called the spectral representation of $R(\boldsymbol{\tau})$, and the function $F(\mathbf{y})$ is called the spectral distribution function of $R(\boldsymbol{\tau})$.

Note in Section 2.2.3 that every covariance function of wide-sense stationary stochastic fields is non-negative definite. Hence, due to Theorem 4C.1, each such covariance function has a spectral representation of the form (4C.1), which further indicates a possibility to establish an explicit representation for the corresponding stochastic field. This leads to the following *spectral representation theorem* of wide-sense stationary stochastic fields.

Theorem 4C.2 (The spectral representation theorem) For every centred, mean-square continuous, wide-sense stationary and complex-valued stochastic field $a(\mathbf{x}, \omega)$ $\mathbf{x} \in \mathbb{R}^n, \omega \in \Omega$, there exists a complex-valued stochastic field $Z(\mathbf{y}, \omega)$ $\mathbf{y} \in \mathbb{R}^n, \omega \in \Omega$ with orthogonal increments such that for each $\mathbf{x} \in \mathbb{R}^n$, $a(\mathbf{x}, \omega)$ has the following representation in the mean-square sense:

$$a(\mathbf{x}, \omega) = \int_{\mathbb{R}^n} e^{\sqrt{-1}\mathbf{x} \cdot \mathbf{y}} dZ(\mathbf{y}, \omega). \quad (4C.4)$$

The stochastic field $Z(\mathbf{y}, \omega)$ is defined up to an additive constant. If this is fixed by setting

$$Z(-\infty, \omega) = 0, \quad (4C.5)$$

we have

$$E(Z(\mathbf{y}, \omega)) = 0, \quad E(|Z(\mathbf{y}, \omega)|^2) = F(\mathbf{y}) \quad \text{and} \quad E(|Z(\mathbf{F}^n)|^2) = F(\mathbf{F}^n), \quad (4C.6)$$

where \mathbf{F}^n is any n -interval in \mathbb{R}^n and $F(\mathbf{y})$ is determined by Eq. (4C.1), in which

$$R(\boldsymbol{\tau}) = R(\mathbf{x}_1 - \mathbf{x}_2) = \text{Cov}(a(\mathbf{x}_1, \omega), a(\mathbf{x}_2, \omega)). \quad (4C.7)$$

It should be noted that the spectral representation (4C.4) is valid only for wide-sense stationary stochastic field on \mathbb{R}^n , while the K-L expansion illustrated in Appendix 4A holds for all second-order stochastic fields.

Clearly, by approximating the infinite integrations in (4C.4, 4C.1) with finite summations, the stochastic field $a(\mathbf{x}, \omega)$ defined on \mathbb{R}^n can be expressed as a finite series of general trigonometric functions. For instance, a real-valued wide-sense stationary stochastic field $h(x, \omega)$ $x \in \mathbb{R}$ with mean zero and spectral distribution function $F_h(y)$ can be approximated by

$$h(x, \omega) \approx \sum_{i=1}^q \sigma_i (A_i(\omega) \cos(y_i x) + B_i(\omega) \sin(y_i x)), \quad (4C.8)$$

where $A_i(\omega)$ and $B_i(\omega)$ are uncorrelated random variables with mean zero and variance one, and

$$\sigma_i^2 = \int_{y_k - \Delta y/2}^{y_k + \Delta y/2} dF_h(y) \quad \Delta y > 0. \quad (4C.9)$$

Since no equation solving is involved, this trigonometric series approximation scheme and its improvements have long been applied in SFEMs (mainly in Monte Carlo simulations [4.8-4.11] to generate material samples) in an intuitive manner. However, in order to ensure the convergence of such a general trigonometric series, the integration domain in (4C.4, 4C.1) has to be truncated at a sufficiently large domain in the frequency space (i.e. $q \rightarrow +\infty$) while the summation steps have to be set sufficiently small (i.e. $\Delta y \rightarrow 0$), which in turn makes this simple scheme inefficient for practical uses.

Appendix 4D Boundary of the Discrete Frequency

Domain

Following (4.26), the covariance function of $a_N(\mathbf{x}, \omega)$ is

$$R_N(\mathbf{x}_1, \mathbf{x}_2) = E \left(\left(\sum_{m_1=-M_1}^{M_1} \cdots \sum_{m_n=-M_n}^{M_n} \Delta Z(\mathbf{m}, \omega) e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_{1k}} \right) \overline{\left(\sum_{m_1=-M_1}^{M_1} \cdots \sum_{m_n=-M_n}^{M_n} \Delta Z(\mathbf{m}, \omega) e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_{2k}} \right)} \right), \quad (4D.1)$$

whose right-hand side term with the highest frequency is

$$E \left(\Delta Z(\mathbf{M}, \omega) e^{\sqrt{-1} \sum_{k=1}^n M_k \theta_k x_{1k}} \overline{\Delta Z(\mathbf{M}, \omega) e^{\sqrt{-1} \sum_{k=1}^n M_k \theta_k x_{2k}}} \right) = E \left(\Delta Z(\mathbf{M}, \omega) \overline{\Delta Z(\mathbf{M}, \omega)} \right) e^{\sqrt{-1} \sum_{k=1}^n M_k \theta_k (x_{1k} - x_{2k})} \quad (4D.2)$$

Hence, the boundary of the discrete frequency domain of $R_N(\mathbf{x}_1, \mathbf{x}_2)$ is $(M_1 \theta_1, \dots, M_n \theta_n)$. This should be compatible with the boundary of \mathbb{F}^n , the continual frequency domain of $R(\mathbf{x}_1, \mathbf{x}_2)$. Consequently,

$$(M_k - 1) \theta_k < f_k \leq M_k \theta_k \quad k = 1, \dots, n, \quad (4D.3)$$

which leads to the result in (4.31).

Appendix 4E The Covariance Matrix Based on the Real Fourier Basis

In (4.33, 4.39, 4.43), \mathbf{G} is constructed with respect to the complex Fourier basis

$$\frac{1}{\sqrt{V_{\mathbb{T}^n}}} e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}, \quad (4E.1)$$

and equivalently the covariance matrix can be constructed with respect to the real Fourier basis

$$\frac{1}{\sqrt{V_{\Gamma^n}}/2} \left\{ \frac{1}{\sqrt{2}}, \cos \sum_{k=1}^n m_k \theta_k x_k, \sin \sum_{k=1}^n m_k \theta_k x_k \right\}. \quad (4E.2)$$

Noting that

$$\cos \left(\sum_{k=1}^n m_k \theta_k x_k \right) = \frac{e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} + e^{-\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}}{2} \quad (4E.3)$$

and

$$\sin \left(\sum_{k=1}^n m_k \theta_k x_k \right) = \frac{e^{\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k} - e^{-\sqrt{-1} \sum_{k=1}^n m_k \theta_k x_k}}{2\sqrt{-1}}, \quad (4E.4)$$

the covariance matrix based on the real Fourier basis can be obtained from \mathbf{G} via an orthogonal transformation.

In the covariance matrix based on the real Fourier basis, consider the entry

$$\frac{2}{V_{\Gamma^n}} \int_{\Gamma^n} \int_{\Gamma^n} R(\mathbf{x}_1 - \mathbf{x}_2) \left(\cos \sum_{k=1}^n m_{i_k} \theta_k x_{1k} \right) \left(\sin \sum_{k=1}^n m_{j_k} \theta_k x_{2k} \right) dx_1 dx_2. \quad (4E.5)$$

The above expression is obviously real-valued, and also according to (4E.3-4E.4, 4.39, 4.43), its real part is zero. Hence, all the entries of the form (4E.5) are zero. This decoupling between sine terms and cosine terms infers that the covariance matrix based on the real Fourier basis can be organized as a block diagonal matrix with two blocks: one block is constructed from the sine terms and the other block is constructed from the cosine terms together with the constant term. From (4E.3-4E.4, 4.43-4.46), it is trivial to conclude that both of these two blocks are diagonally dominant.

No matter how the covariance matrix is constructed with respect to the complex Fourier basis or the real Fourier basis, the final results of the K-L eigenstructure λ_i and $\psi_i(\mathbf{x})$ are the same. However, by using the above decoupling transformation on \mathbf{G} , the problem size of the associated algebraic eigenvalue problem can be reduced by half, which in turn significantly improves the computational efficiency of solving for λ_i and \mathbf{Q} (especially for large-scale covariance matrices \mathbf{G}).

Appendix 4F The Exact K-L Expansion in Case I

For the simple stochastic field $a_{1D}^1(x, \omega)$ $x \in [-t, t]$ considered in Case I, the eigenpairs λ_i and $\psi_i(x)$ in the K-L expansion can be analytically obtained as follows (see e.g. [4.13, 4.32] for detailed derivation):

$$\lambda_i = \frac{2a_1}{1 + a_1^2 \omega_i^2} \quad i = 1, 2, 3, \dots \quad (4F.1)$$

and

$$\psi_i(x) = \begin{cases} \frac{\cos(\omega_i x)}{\sqrt{t + \frac{\sin(2\omega_i t)}{2\omega_i}}} & i = 1, 3, 5, \dots \\ \frac{\sin(\omega_i x)}{\sqrt{t - \frac{\sin(2\omega_i t)}{2\omega_i}}} & i = 2, 4, 6, \dots \end{cases}, \quad (4F.2)$$

where $\omega_1 \leq \omega_2 \leq \omega_3 \leq \dots$ are given by the following transcendental equations

$$1 - a_1 \omega \tan(\omega t) = 0 \quad \text{or} \quad a_1 \omega + \tan(\omega t) = 0. \quad (4F.3)$$

References

- [4.1] M. Shinozuka, Monte Carlo solution of structural dynamics, *Computers & Structures*, 2 (5-6) (1972) 855-874.
- [4.2] F. Yamazaki, M. Shinozuka and G. Dasgupta, Neumann expansion for stochastic finite element analysis, *Journal of Engineering Mechanics - ASCE*, 114 (8) (1988) 1335-1354.
- [4.3] E.H. Vanmarcke and M. Grigoriu, Stochastic finite element analysis of simple beams, *Journal of Engineering Mechanics - ASCE*, 109 (5) (1983) 1203-1214.
- [4.4] W.K. Liu, T. Belytschko and A. Mani, Random field finite-elements. *International Journal for Numerical Methods in Engineering*, 23 (10) (1986) 1831-1845.
- [4.5] M. Kamiński and M. Kleiber, Stochastic structural interface defects in fiber composites, *International Journal of Solids and Structures*, 33 (20-22) (1996) 3035-3056.

- [4.6] T.D. Hien and M. Kleiber, Stochastic finite element modelling in linear transient heat transfer, *Computer Methods in Applied Mechanics and Engineering*, 144 (1-2) (1997) 111-124.
- [4.7] C.C. Li and A. Derkiureghian, Optimal discretization of random fields, *Journal of Engineering Mechanics – ASCE*, 119 (6) (1993) 1136-1154.
- [4.8] M. Shinozuka and G. Deodatis, Simulation of stochastic processes by spectral representation, *Applied Mechanics Reviews – ASME*, 44 (4) (1991) 191-204.
- [4.9] J.E. Hurtado and A.H. Barbat, Monte Carlo techniques in computational stochastic mechanics, *Archives of Computational Methods in Engineering: State of the Art Reviews*, 5 (1) (1998) 3-30.
- [4.10] M. Grigoriu, A spectral representation based model for Monte Carlo simulation, *Probabilistic Engineering Mechanics*, 15 (2000) 365-370.
- [4.11] M. Grigoriu, *Stochastic Calculus: Applications in Science and Engineering*, Birkhauser, 2002.
- [4.12] P.D. Spanos and R.G. Ghanem, Stochastic finite element expansion for random media, *Journal of Engineering Mechanics – ASCE*, 115 (5) (1989) 1035-1053.
- [4.13] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Revised edition, Dover Publications, New York, 2003.
- [4.14] R.G. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, *Computer Methods in Applied Mechanics and Engineering*, 168 (1999) 19-34.
- [4.15] H.G. Matthies and C. Bucher, Finite elements for stochastic media problems, *Computer Methods in Applied Mechanics and Engineering*, 168 (1999) (1-4) 3-17.
- [4.16] I.M. Babuška and P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, *Computer Methods in Applied Mechanics and Engineering*, 191 (37-38) (2002) 4093-4122.
- [4.17] P. Frauenfelder, C. Schwab and R.A. Todor, Finite elements for elliptic problems with stochastic coefficients, *Computer Methods in Applied Mechanics and Engineering*, 194 (2-5) (2005) 205-228.
- [4.18] H.C. Noh, A formulation for stochastic finite element analysis of plate structures with uncertain Poisson's ratio, *Computer Methods in Applied Mechanics and Engineering*, 193 (45-47) (2004) 4857-4873.
- [4.19] M. Loève, *Probability Theory*, 4th Edition, Vol. II, Springer-Verlag, New York, 1977.
- [4.20] B.B. Ash, *Information Theory*, Dover Publications, New York, 1990.

- [4.21] A.C. Zaanen, *Linear Analysis: Measure and Integral, Banach and Hilbert Space, Linear Integral Equations*, North Holland Publishing Co., Amsterdam, 1964.
- [4.22] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I, Interscience, New York, 1953.
- [4.23] B. Simon, *Functional Integration and Quantum Physics*, Academic Press, New York, 1979.
- [4.24] A.M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, Prentice-hall Inc., New Jersey, U.S., 1962.
- [4.25] R.J. Adler, *The Geometry of Random Fields*, John Wiley & Sons, New York, 1981.
- [4.26] I.T. Jolliffe, *Principal Component Analysis*, 2nd Edition, New York, 2002.
- [4.27] W.J. Pan, *Fourier Analysis and its Applications*, Peking University Press, Beijing, 2000. (In Chinese)
- [4.28] H. Dym and H.P. McKean, *Fourier Series and Integrals*, Academic Press, New York, 1972.
- [4.29] R.E. Edwards, *Fourier Series: A Modern Introduction*, Vol. 1, 2nd Edition, Springer-Verlag, New York, 1979.
- [4.30] W.H. Press, S.A. Teukolsky and W.T. Vetterling et al., *Numerical Recipes in C: the Art of Scientific Computing*, 2nd Edition, Cambridge University Press, Cambridge, 1995.
- [4.31] R.W. Hamming, *Numerical Methods for Scientists and Engineers*, 2nd Edition, Dover, New York, 1987.
- [4.32] H.L. Vantrees, *Detection, Estimation and Modulation Theory*, Part 1, Wiley, New York, 1968.
- [4.33] C.F. Li, Y.T. Feng and D.R.J. Owen, On the quadrature of certain multi-dimensional oscillatory functions, *Computer Physics Communications*, In preparation.
- [4.34] C.F. Li, Y.T. Feng and D.R.J. Owen, Elastostatics of elementary random media: part I – theory and algorithms, *Journal of the Mechanics and Physics of Solids*, In preparation.
- [4.35] C.F. Li, Y.T. Feng and D.R.J. Owen, Elastostatics of elementary random media: part II – verification and application, *Journal of the Mechanics and Physics of Solids*, In preparation.
- [4.36] C.F. Li, Y.T. Feng, D.R.J. Owen and I.M. Davies, Fourier representation of random media fields in stochastic finite element modelling, *Engineering*

Computations, 23 (7) (2006) 794-817.

- [4.37] C.F. Li, Y.T. Feng, D.R.J. Owen and I.M. Davies, A simple semi-analytic solution for Karhunen-Loève expansions of stationary stochastic fields, *The Seventh World Congress on Computational Mechanics*, Los Angeles, California, USA, July 16-22, 2006.
- [4.38] C.F. Li, Y.T. Feng, D.R.J. Owen, D.F. Li and I.M. Davies, A Fourier-Karhunen-Loève representation scheme for wide-sense stationary stochastic fields, *International Journal for Numerical Methods in Engineering*, Submitted.

Chapter 5

Stochastic Finite Element Discretization for Elastostatics of Elementary Random Media

Following the ERM model established in Chapter 3 and its F-K-L representation scheme developed in Chapter 4, this chapter first sets up, in a mathematically rigorous fashion, a closed form SPDE system for elastostatics of ERM. Then, based on standard finite element technology, the SPDE system is discretized to achieve a stochastic system of linear algebraic equations.

5.1 The Stochastic Partial Differential Equation System

Recall Section 3.1 that, the governing equation system for elastostatics of random media has the following form

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, \omega) + \rho(\mathbf{x}, \omega) \mathbf{g} = 0 \\ \boldsymbol{\varepsilon}(\mathbf{x}, \omega) = \frac{1}{2}(\mathbf{u}(\mathbf{x}, \omega) \nabla + \nabla \mathbf{u}(\mathbf{x}, \omega)) \\ \boldsymbol{\sigma}(\mathbf{x}, \omega) = \mathbf{C}(\mathbf{x}, \omega) : \boldsymbol{\varepsilon}(\mathbf{x}, \omega) \end{cases} \quad \begin{cases} \mathbf{u}(\mathbf{x}, \omega) = \tilde{\mathbf{u}}(\mathbf{x}, \omega) & \text{on } \partial D_u \\ \boldsymbol{\sigma}(\mathbf{x}, \omega) = \tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega) & \text{on } \partial D_\sigma \end{cases} \quad (5.1)$$

$\mathbf{x} \in D, \omega \in \Omega$

in which D denotes the random medium domain, ∂D_u the displacement boundary,

∂D_σ , the stress boundary, and stochastic fields $\tilde{\mathbf{u}}(\mathbf{x}, \omega)$ and $\tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega)$ the corresponding boundary conditions; the unknown stochastic fields $\boldsymbol{\sigma}(\mathbf{x}, \omega)$, $\boldsymbol{\varepsilon}(\mathbf{x}, \omega)$ and $\mathbf{u}(\mathbf{x}, \omega)$ denote respectively the stress tensor, the strain tensor and the displacement vector; vector \mathbf{g} is the gravitational acceleration, and the given stochastic fields $\rho(\mathbf{x}, \omega)$ and $\mathbf{C}(\mathbf{x}, \omega)$ denote the density and the elastic tensor, respectively.

For a C^m ERM, the random material properties $\rho(\mathbf{x}, \omega)$ and $\mathbf{C}(\mathbf{x}, \omega)$ are defined by their first- and second- order statistical moments. That is, the density $\rho(\mathbf{x}, \omega)$ is given by

$$E(\rho(\mathbf{x}, \omega)) = \rho_o(\mathbf{x}) \quad \forall \mathbf{x} \in D \quad (5.2)$$

and

$$\text{Cov}(\rho(\mathbf{x}_1, \omega), \rho(\mathbf{x}_2, \omega)) = R_\rho(\boldsymbol{\tau}) \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D \quad (5.3)$$

where $\boldsymbol{\tau} = \mathbf{x}_1 - \mathbf{x}_2$; and the elastic tensor $\mathbf{C}(\mathbf{x}, \omega)$ is given by

$$E(\mathbf{C}(\mathbf{x}, \omega)) = \mathbf{C}_o(\mathbf{x}) \quad \forall \mathbf{x} \in D \quad (5.4)$$

and

$$\text{Cov}(\mathbf{C}(\mathbf{x}_1, \omega), \mathbf{C}(\mathbf{x}_2, \omega)) = \mathbf{R}_c(\boldsymbol{\tau}) \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D, \quad (5.5)$$

in which $\mathbf{C}_o(\mathbf{x})$ is a tensor whose scalar entries are the expectation functions of the corresponding scalar entries in $\mathbf{C}(\mathbf{x}, \omega)$, and $\mathbf{R}_c(\boldsymbol{\tau})$ is a matrix whose elements are the covariance functions of the corresponding scalar entries in $\mathbf{C}(\mathbf{x}, \omega)$.

In an equation system regarding C^m ERM, the differential/integral operators are all defined based on the mean square convergence, and the smoothness of the corresponding ERM is specified by the integer m . As only the first order derivatives are involved in Eq. (5.1), it is required that $m=1$ such that the density $\rho(\mathbf{x}, \omega)$ and the elastic tensor $\mathbf{C}(\mathbf{x}, \omega)$ of the associated ERM have at least C^1 continuity. Consequently, in the mean square sense, $\rho_o(\mathbf{x})$ and every scalar entry in $\mathbf{C}_o(\mathbf{x})$ are required to be at least C^1 continuous with respect to \mathbf{x} , and $R_\rho(\boldsymbol{\tau})$ and every scalar entry in $\mathbf{R}_c(\boldsymbol{\tau})$ are required to be at least C^2 continuous with respect to $\boldsymbol{\tau}$.

According to the F-K-L representation of ERM and Eqs. (5.2-5.3), the density $\rho(\mathbf{x}, \omega)$ can be expressed as

$$\rho(\mathbf{x}, \omega) \approx \rho_o(\mathbf{x}) + \sum_{i=1}^{N_\rho^*} \sqrt{\lambda_i^\rho} \xi_i^\rho(\omega) \rho_i(\mathbf{x}) \quad (5.6)$$

where $\lambda_i^\rho > 0$ are constants, $\xi_i^\rho(\omega)$ are stochastically independent standard Gaussian random variables, $\rho_i(\mathbf{x})$ are a sequence of orthonormal deterministic functions, and the integer N_ρ^* is explicitly determined by the error-control parameters in the F-K-L representation scheme. Similarly, from Eqs. (5.4-5.5), the elastic tensor $\mathbf{C}(\mathbf{x}, \omega)$ can be expressed as

$$\mathbf{C}(\mathbf{x}, \omega) \approx \mathbf{C}_o(\mathbf{x}) + \sum_{i=1}^{N_C^*} \sqrt{\lambda_i^C} \xi_i^C(\omega) \mathbf{C}_i(\mathbf{x}) \quad (5.7)$$

where $\lambda_i^C > 0$ are constants, $\xi_i^C(\omega)$ are a stochastically independent standard Gaussian random sequence, $\mathbf{C}_i(\mathbf{x})$ are a sequence of tensor-valued deterministic functions, and the integer N_C^* is explicitly determined by the error-control parameters in the F-K-L representation scheme.

In Eq. (5.1), the displacement and stress boundary conditions $\tilde{\mathbf{u}}(\mathbf{x}, \omega)$ and $\tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega)$ are also allowed to be random. However, the possible randomness from these boundary conditions can be treated separately since it has no direct effect on the associated differential operators. Hence, the definition method of $\tilde{\mathbf{u}}(\mathbf{x}, \omega)$ and $\tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega)$ are not specified in the ERM model, and they can be defined either by their statistical moments or by any other appropriate methods.

Remark: To simplify the notation, the linear dependent scalar entries in $\mathbf{C}(\mathbf{x}, \omega)$ are not distinguished in Eqs. (5.4-5.5). However, in reality, it is totally unnecessary to provide every entry of functions $\mathbf{C}_o(\mathbf{x})$ and $\mathbf{R}_C(\boldsymbol{\tau})$, which are fully defined by the first two statistical moments of those linear independent scalar entries in the elastic tensor. For example, it is well known that there are only two linear independent elastic parameters for an isotropic material; hence, the first- and second- order statistical moments of these two elastic parameters are sufficient to fully define the elastic tensor of an isotropic ERM and

achieve the corresponding F-K-L representation in the form (5.7). In addition, $\rho(\mathbf{x}, \omega)$ and $\mathbf{C}(\mathbf{x}, \omega)$ are assumed stochastically independent in Eqs. (5.2-5.5), and this leads to different random sequences $\sqrt{\lambda_i^\rho} \xi_i^\rho(\omega)$ and $\sqrt{\lambda_i^C} \xi_i^C(\omega)$ in Eqs. (5.6-5.7). However, if the stochastic dependence between the density and the elastic tensor needs to be taken into account, they should be defined via the same vector as in Section 4.6.

5.2 The Stochastic System of Linear Algebraic Equations

Eqs. (5.1-5.7) form the complete governing equation system for elastostatics of ERM. In some situations where the geometric configurations and boundary conditions are both relatively simple, it is possible to obtain the analytical solution for this SPDE system. Although interesting, the analytical solution approach is not the emphasis of this thesis; and instead via a stochastic finite element approach, this work concentrates on obtaining the numerical solutions in general situations.

It is clear that the SPDE system (5.1-5.7) is a generalization of the conventional PDE system (3.1) that describes the elastostatics of deterministic materials. Naturally, the stochastic finite element discretization for the SPDE system can be expected to be achieved by generalizing the finite element discretization for the PDE system. The well-established finite element method has become a standard analysis tool in engineering, and its details can be found in various textbooks such as [5.1-5.3]. Hence, the following stochastic finite element formulation will be derived without addressing those details that are similar to or essentially the same as the standard finite element method.

First, the random medium D is discretized with a finite element mesh. It should be emphasised that the size of each element must be sufficiently small to satisfy the requirements from both the stress gradient consideration and the stochastic field variability. Specifically, (a) a finer mesh is needed in a location where the stress gradient is more significant; (b) the element size must also be examined in comparison with the effective correlation length of the ERM.

Then, from Eq. (5.1) and following the standard finite element discretization procedure, the total potential of the ERM D can be written as

$$\begin{aligned} \Pi_p = & \sum_e (\mathbf{u}^e(\omega))^T \left(\int_{D^e} \frac{1}{2} (\mathbf{B}(\mathbf{x}))^T \mathbf{D}(\mathbf{x}, \omega) \mathbf{B}(\mathbf{x}) dV \right) \mathbf{u}^e(\omega) \\ & - \sum_e (\mathbf{u}^e(\omega))^T \int_{D^e} (\mathbf{N}(\mathbf{x}))^T \mathbf{g} \rho(\mathbf{x}, \omega) dV \\ & - \sum_e (\mathbf{u}^e(\omega))^T \int_{\partial D^e} (\mathbf{N}(\mathbf{x}))^T \tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega) dS \end{aligned} \quad (5.8)$$

where \sum_e represents the summation over all the elements, $\mathbf{u}^e(\omega)$ denotes the nodal displacement vector organized with respect to each element, $\mathbf{B}(\mathbf{x})$ the strain matrix, $\mathbf{D}(\mathbf{x}, \omega)$ the elastic matrix, and $\mathbf{N}(\mathbf{x})$ the shape function matrix. Due to Eq. (5.7), the elastic matrix can be written as

$$\mathbf{D}(\mathbf{x}, \omega) = \mathbf{D}_o(\mathbf{x}) + \sum_{i=1}^{N_C^*} \sqrt{\lambda_i^C} \xi_i^C(\omega) \mathbf{D}_i(\mathbf{x}), \quad (5.9)$$

where $\mathbf{D}_o(\mathbf{x})$ and $\mathbf{D}_i(\mathbf{x})$ are the elastic matrices constructed respectively from the elastic tensors $\mathbf{C}_o(\mathbf{x})$ and $\mathbf{C}_i(\mathbf{x})$. Substituting (5.6) and (5.9) into (5.8) yields

$$\begin{aligned} \Pi_p = & \sum_e (\mathbf{u}^e(\omega))^T \frac{1}{2} \left(\mathbf{K}_o^e + \sum_i^{N_C^*} \xi_i^C(\omega) \mathbf{K}_i^e \right) \mathbf{u}^e(\omega) \\ & - \sum_e (\mathbf{u}^e(\omega))^T \left(\mathbf{P}_o^e + \sum_i^{N_p^*} \xi_i^p(\omega) \mathbf{P}_i^e \right) - \sum_e (\mathbf{u}^e(\omega))^T \tilde{\mathbf{P}}^e(\omega) \end{aligned} \quad (5.10)$$

in which the element stiffness matrices \mathbf{K}_o^e and \mathbf{K}_i^e , the element volume-force vectors \mathbf{P}_o^e and \mathbf{P}_i^e , and the element external-load vector $\tilde{\mathbf{P}}^e(\omega)$ are given below

$$\mathbf{K}_o^e = \int_{D^e} (\mathbf{B}(\mathbf{x}))^T \mathbf{D}_o(\mathbf{x}) \mathbf{B}(\mathbf{x}) dV, \quad (5.11)$$

$$\mathbf{K}_i^e = \sqrt{\lambda_i^C} \int_{D^e} (\mathbf{B}(\mathbf{x}))^T \mathbf{D}_i(\mathbf{x}) \mathbf{B}(\mathbf{x}) dV \quad i = 1, 2, \dots, N_C^*; \quad (5.12)$$

$$\mathbf{P}_o^e = \int_{D^e} (\mathbf{N}(\mathbf{x}))^T \mathbf{g} \rho_o(\mathbf{x}) dV, \quad (5.13)$$

$$\mathbf{P}_i^e = \sqrt{\lambda_i^p} \int_{D^e} (\mathbf{N}(\mathbf{x}))^T \mathbf{g} \rho_i(\mathbf{x}) dV \quad i = 1, 2, \dots, N_p^*; \quad (5.14)$$

$$\tilde{\mathbf{P}}^e(\omega) = \int_{\partial D^e} (\mathbf{N}(\mathbf{x}))^T \tilde{\boldsymbol{\sigma}}(\mathbf{x}, \omega) dS. \quad (5.15)$$

The total potential Π_p is minimized by the real displacement solution. Hence, taking variation with respect to the displacement on both sides of Eq. (5.10), the following stochastic system of linear algebraic equations holds

$$\left(\mathbf{K}_o + \sum_{i=1}^{N_c^*} \xi_i^C(\omega) \mathbf{K}_i \right) \mathbf{U}(\omega) = \left(\mathbf{P}_o + \sum_{i=1}^{N_\rho^*} \xi_i^\rho(\omega) \mathbf{P}_i \right) + \tilde{\mathbf{P}}(\omega), \quad (5.16)$$

where \mathbf{K}_o and \mathbf{K}_i are the global stiffness matrices assembled respectively from \mathbf{K}_o^e and \mathbf{K}_i^e ; \mathbf{P}_o and \mathbf{P}_i are the global volume-force vectors assembled respectively from \mathbf{P}_o^e and \mathbf{P}_i^e ; $\tilde{\mathbf{P}}(\omega)$ is the global external-load vector assembled from $\tilde{\mathbf{P}}^e(\omega)$; and $\mathbf{U}(\omega)$ is the unknown nodal displacement vector. Following a similar analysis as in the standard finite element formulation, it can be proven that the matrix \mathbf{K}_o is real, symmetric and non-negative definite, the matrices \mathbf{K}_i are all real and symmetric, and the random matrix sum $\mathbf{K}_o + \sum_{i=1}^{N_c^*} \xi_i^C(\omega) \mathbf{K}_i$ is, in the context of probability, real, symmetric and non-negative definite.

As in the standard finite element method, the displacement boundary conditions $\tilde{\mathbf{u}}(\mathbf{x}, \omega)$ of the SPDE system (5.1) can be directly introduced into the stochastic linear algebraic system (5.16). The details are listed as follows.

- To set

$$(\mathbf{U}(\omega))_j \equiv 0 \quad (5.17)$$

where $(\mathbf{U}(\omega))_j$ is the j -th entry of $\mathbf{U}(\omega)$, the stiffness matrices and the nodal load vectors need to be modified as

$$(\mathbf{K}_o)_{jk} = (\mathbf{K}_o)_{kj} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}, \quad (5.18)$$

$$(\mathbf{K}_i)_{jk} = (\mathbf{K}_i)_{kj} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \quad i = 1, 2, \dots, N_c^*, \quad (5.19)$$

$$(\mathbf{P}_o)_j = 0, \quad (5.20)$$

$$(\mathbf{P}_i)_j = 0 \quad i = 1, 2, \dots, N_\rho^*, \quad (5.21)$$

$$\left(\tilde{\mathbf{P}}(\omega)\right)_j \equiv 0, \quad (5.22)$$

where $(\mathbf{K}_o)_{jk}$ and $(\mathbf{K}_i)_{jk}$ denote the entry at the j -th row and k -th column of \mathbf{K}_o and \mathbf{K}_i , respectively; $(\mathbf{P}_o)_j$, $(\mathbf{P}_i)_j$ and $\left(\tilde{\mathbf{P}}(\omega)\right)_j$ denote the j -th entry of \mathbf{P}_o , \mathbf{P}_i and $\tilde{\mathbf{P}}(\omega)$, respectively.

- To set

$$\left(\mathbf{U}(\omega)\right)_j = U_j(\omega) \neq 0, \quad (5.23)$$

the change to the stiffness matrices and the nodal load vectors should be made as follows,

$$\left(\mathbf{K}_o\right)_{ij} = \kappa; \quad (5.24)$$

$$\left(\mathbf{P}_o\right)_j = \kappa U_j(\omega); \quad (5.25)$$

where $\kappa > 0$ is a sufficiently large number.

Once the proper displacement boundary conditions are applied to Eq. (5.16), the random stiffness matrix sum in the left-hand side of the equation becomes almost surely real, symmetric and positive definite such that

$$P \left\{ \det \left(\mathbf{K}_o + \sum_{i=1}^{N_c^*} \xi_i^C(\omega) \mathbf{K}_i \right) \neq 0 \right\} = 1. \quad (5.26)$$

Consequently, the stochastic system of linear algebraic equations (5.16) is always well defined and the random displacement solution $\mathbf{U}(\omega)$ exists with probability one.

References

- [5.1] O.C. Zienkiewicz, R.L. Taylor and J.Z. Zhu, *The Finite Element Method: Its Basis and Fundamentals*, 6th Edition, Butterworth-Heinemann, Oxford, 2005.
- [5.2] O.C. Zienkiewicz and R.L. Taylor, *The Finite Element Method for Solid and Structural Mechanics*, 6th Edition, Butterworth-Heinemann, Oxford, 2005.
- [5.3] O.C. Zienkiewicz, R.L. Taylor and P. Nithiarasu, *The Finite Element Method for Fluid Dynamics*, 6th Edition, Butterworth-Heinemann, Oxford, 2005.

Chapter 6

The Joint Diagonalization Solution Strategy for the Stochastic System of Linear Algebraic Equations

For the solution of the stochastic system of linear algebraic equations obtained in Chapter 5, a novel solution strategy, namely the joint diagonalization method, is developed in this chapter. The form of the stochastic linear algebraic system (5.16) is not new in SFEM research, and in particular for simple elastostatic problems [6.1-6.2, 6.4-6.7, 6.9-6.10, 6.12-6.13] where the Young's modulus is the only random material parameter and also for simple steady-state heat conduction problems [6.2-6.3, 6.8] where the only random material property is the thermal conductivity, essentially similar equations have long been noticed in the literature. Consequently, a number of methods have been developed for the solution of these specific stochastic linear algebraic equations.

In order to highlight the difference between the new solution method and various existing techniques and also in order to make the discussions regarding solution strategies stand in a more general context, it is decided not to use the notation in (5.16) and instead present all the formulations by virtue of the equation $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m) \mathbf{x} = \mathbf{b}$, in which α_i ($i=1, \dots, m$) denote random variables, A_i ($i=1, \dots, m$) real symmetric deterministic matrices, \mathbf{b} a deterministic/random vector and \mathbf{x} the unknown random vector to be solved.

The remainder of this chapter is organized as follows. First, the general background of the stochastic system of linear algebraic equations, including its link with the conventional linear algebraic equations arising from the standard finite element analysis, is outlined in Section 6.1. Then, existing solution techniques for the stochastic linear algebraic system are briefly reviewed in Section 6.2. Next, the joint diagonalization method is explained in detail in Section 6.3, including its operation process, algorithm properties and performance analysis. In Section 6.4, three numerical examples are employed to investigate in detail the performance of the new method. The chapter concludes in Section 6.5 with a summary of the main features and limitations of the new solution strategy, and some suggestions for future research.

Throughout this chapter, a vector (e.g. \mathbf{x}) always indicates a column vector unless otherwise stated, and the entry located at the i -th row and j -th column of a matrix \mathbf{A} is represented by $(\mathbf{A})_{ij}$.

6.1 Problem Background

The classical linear finite element equations have the following generic form

$$\mathbf{Ax} = \mathbf{b} \tag{6.1}$$

where the physical meanings of \mathbf{A} , \mathbf{x} and \mathbf{b} relate to the system under consideration. For example, in a simple problem of elastostatics \mathbf{A} denotes the elastic stiffness matrix, \mathbf{x} the unknown nodal displacement vector and \mathbf{b} the nodal load vector; in a steady-state heat conduction problem, they denote the heat conduction matrix, the unknown nodal temperature vector and the nodal temperature-load vector respectively.

In reality, due to various uncertainties, the deterministic physical model described by Eq. (6.1) is seldom valid. Consequently, an appropriate safety factor is often adopted to amend the result for a practical engineering problem. Numerous successful applications have shown the power of this deterministic approach. Conversely, there exist a number of situations where deterministic models have failed to give satisfactory solutions, and these include problems in rocks, soil and ground water where the stochastic uncertainty becomes dominant. Consequently, a more generalized solution technique, namely the

SFEM, is needed to model these intricate random physical systems. The study of SFEMs has been in progress for several decades and many different techniques have been presented to formulate uncertainties in various applications [6.1-6.14, 6.17]. However, compared to the success of the deterministic FEM, the SFEM is still in its infancy and many fundamental questions are still outstanding. Due to the continual exponential increase in computer power and data storage, the formulation of the SFEM has recently received considerable attention from the computational mechanics community [6.17] and, generally, for many engineering applications, the final linear SFEM equations take the following form

$$(\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_m A_m) \mathbf{x} = \mathbf{b} , \quad (6.2)$$

where the deterministic $n \times n$ real symmetric matrices A_i ($i=1, \dots, m$), unknown real random vector \mathbf{x} and deterministic real vector \mathbf{b} have essentially the same physical meanings as their primary counterparts in Eq. (6.1); real scalars α_i ($i=1, \dots, m$) denote a series of dimensionless random factors that capture various intrinsic uncertainties in the system. The advantage of Eq. (6.2) is that the physical problem with a stochastic nature is decoupled into deterministic and stochastic parts in a simple regular fashion. Typically, different SFEM formulations define α_i , A_i and \mathbf{b} differently but give rise to the same stochastic system of linear algebraic equations as shown above [6.1-6.11]. It should be noted that, Eq. (6.2) is not always explicitly addressed in the literature [6.1-6.5], since there is no significant advantage of transforming early SFEM techniques into the form of (6.2).

The aim of this chapter is to introduce a new numerical procedure for the solution of the stochastic system of linear algebraic equations. Existing techniques applied to the solution of Eq. (6.2) mainly include the Monte Carlo method [6.12-6.13], the perturbation method [6.1-6.3], the Neumann expansion method [6.4-6.5] and the polynomial chaos expansion method [6.6-6.11]. However, none of these schemes have promoted the SFEM to such a mature level as the well established FEM. For the development of a solver for Eq. (6.2), these methods will be considered in the next section.

Remarks: For simplicity, uncertainties arising from the right-hand side of Eq. (6.2) have at this stage been temporarily ignored. However, as will be shown in Section 6.3, it is trivial to relax this unnecessary restriction.

6.2 Review of Existing Solution Techniques

Until recently, there have been four major techniques employed to solve the stochastic system of linear algebraic equations (6.2): the Monte Carlo method, the perturbation method, the Neumann expansion method and the polynomial chaos expansion method. In order to avoid possible confusion in terminology, it should be noted that sometimes the same titles are also employed in the literature to indicate the entire schemes corresponding to their SFEM formulations. Although the following discussion will only explore these methods from the viewpoint of a solver for Eq. (6.2), the principles hold respectively for their associated SFEM schemes.

6.2.1 The Monte Carlo Method

First, N sets of samples of the random variables α_i ($i=1, \dots, m$) are generated. For each sample path $\alpha_i = a_{ij}$, where the right-hand side sample values are distinguished by sample index $j=1, \dots, N$, Eq. (6.2) becomes a standard deterministic system of linear algebraic equations:

$$\mathbf{B}_j \mathbf{x}_j = \mathbf{b} \quad (6.3)$$

with

$$\mathbf{B}_j = \sum_{i=1}^m a_{ij} \mathbf{A}_i. \quad (6.4)$$

Solutions of Eq. (6.3), i.e.

$$\mathbf{x}_j = \mathbf{B}_j^{-1} \mathbf{b} \quad (j=1, \dots, N), \quad (6.5)$$

form a sample set for the random solution \mathbf{x} , which can then be employed to calculate the associated empirical joint distribution or empirical statistical quantities.

The Monte Carlo method is simple and perhaps the most versatile method used to date, but typically the computational costs are extremely high, especially for large scale problems where a large number of samples have to be computed in order to obtain a rational estimation satisfying the required accuracy. Improvements [6.12-6.13] have been made to overcome this computational difficulty.

6.2.2 The Perturbation Method

In a perturbation scheme, all the random variables α_i ($i=1, \dots, m$) are first decomposed into a deterministic part and a random part, i.e.

$$\alpha_i = E(\alpha_i) + \alpha_i^s \quad (i=1, \dots, m), \quad (6.6)$$

where $E(\alpha_i)$ and α_i^s denote respectively the mean value and the centred deviation random variable. As a result, Eq. (6.2) is then transformed into

$$(A_0 + \alpha_1^s A_1 + \alpha_2^s A_2 + \dots + \alpha_m^s A_m) \mathbf{x} = \mathbf{b}, \quad (6.7)$$

where

$$A_0 = \sum_{i=1}^m E(\alpha_i) A_i \quad (6.8)$$

is the deterministic mean of sum $\sum_{i=1}^m \alpha_i A_i$. Meanwhile, the random solution \mathbf{x} is approximated by Taylor's series

$$\mathbf{x} \approx \mathbf{x} \Big|_{\alpha_i^s=0 \ (i=1, \dots, m)} + \sum_{i=1}^m \frac{\partial \mathbf{x}}{\partial \alpha_i^s} \alpha_i^s = \mathbf{x}_o + \sum_{i=1}^m \mathbf{c}_i \alpha_i^s \quad (6.9)$$

where the origin \mathbf{x}_o and coefficients \mathbf{c}_i ($i=1, \dots, m$) are both unknown deterministic vectors. Next, substituting (6.9) into (6.7) yields

$$(A_0 + \alpha_1^s A_1 + \dots + \alpha_m^s A_m) (\mathbf{x}_o + \mathbf{c}_1 \alpha_1^s + \dots + \mathbf{c}_m \alpha_m^s) = \mathbf{b} \quad (6.10a)$$

and

$$A_0 \mathbf{x}_o + \sum_{i=1}^m (A_0 \mathbf{c}_i + A_i \mathbf{x}_o) \alpha_i^s + \sum_{i=1}^m \sum_{j=1}^m A_i \mathbf{c}_j \alpha_i^s \alpha_j^s = \mathbf{b}. \quad (6.10b)$$

Inspired by the idea of perturbation in deterministic problems, some researchers [6.1-6.3] believe that the following equalities resulting respectively from the zero-order "random perturbation" and the first-order "random perturbation" hold, provided that the scale of random fluctuations α_i^s is sufficiently small (e.g. less than ten percent [6.1]):

Zero order

$$A_0 \mathbf{x}_o = \mathbf{b}, \quad (6.11a)$$

First order

$$A_0 \mathbf{c}_i = -A_i \mathbf{x}_o \quad (i=1, \dots, m), \quad (6.11b)$$

from which x_0 and c_i can be solved and then employed in (6.9) to provide an approximation to the random solution x in terms of the first-order Taylor's series.

In the literature, the aforementioned procedure is called the first-order perturbation method and similarly, the second-order perturbation method follows immediately from expanding the first-order Taylor series (6.9) to the second order. Applications of higher-order perturbations are, however, rare due to the increasingly high complexity of analytic derivations as well as computational costs. Although a number of numerical experiments with small random fluctuations have been reported to show good agreements between the perturbation method and the Monte Carlo method, no criteria for convergence have been established in the present context [6.3].

6.2.3 The Neumann Expansion Method

The first step of the Neumann expansion method [6.4-6.5] also consists of centralizing the random variables α_i ($i=1, \dots, m$) to obtain Eq. (6.7) whose left-hand side coefficients are further treated as a sum of a deterministic matrix and a stochastic matrix, i.e.

$$(A_0 + \Delta A)x = b \quad (6.12)$$

with the stochastic matrix ΔA defined as

$$\Delta A = \alpha_1^s A_1 + \alpha_2^s A_2 + \dots + \alpha_m^s A_m. \quad (6.13)$$

The solution of Eq. (6.12) yields

$$x = (I + A_0^{-1} \Delta A)^{-1} A_0^{-1} b. \quad (6.14)$$

The term $(I + A_0^{-1} \Delta A)^{-1}$ can be expressed in a Neumann series expansion giving

$$x = (I - B + B^2 - B^3 + \dots) A_0^{-1} b \quad (6.15)$$

with

$$B = A_0^{-1} \Delta A. \quad (6.16)$$

The random solution vector can now be represented by the following series

$$x = x_0 - Bx_0 + B^2 x_0 - B^3 x_0 + \dots \quad (6.17)$$

with

$$x_0 = A_0^{-1} b. \quad (6.18)$$

It is well known that the Neumann expansion shown in Eq. (6.15) converges if the spectral radius of the stochastic matrix B is always less than 1. This is not a serious restriction in engineering problems, as in most cases the random variations are smaller than the mean value. The most significant feature of this approach is that the inverse of matrix A_0 is required only once for all samples and, at least in principle, the statistical moments of the solution x in (6.17) can be obtained analytically by recognising that there is no inverse operation on ΔA in (6.13). Nevertheless, the computational costs of the Neumann expansion method increase with the number of terms required in Eq. (6.17) to achieve a given accuracy and therefore for problems with large random fluctuations, the method loses its advantage and could become even more expensive than the direct Monte Carlo method. Note that improvements [6.12] have been made by combining the Neumann expansion method and the preconditioned conjugate gradient method.

6.2.4 The Polynomial Chaos Expansion Method

In the theory of probability and stochastic processes, the polynomial chaos expansion [6.16] was originally developed by Wiener [6.18], Cameron, Martin [6.19] and Itô [6.20] more than fifty years ago. Concisely, a function $f(\beta_1, \beta_2, \dots, \beta_q)$ of a standard Gaussian random vector $\beta = (\beta_1, \beta_2, \dots, \beta_q)^T$ has a unique representation, i.e. the so called polynomial chaos expansion

$$f(\beta_1, \beta_2, \dots, \beta_q) = \sum_{i=1}^{\infty} h_i H_i(\beta_1, \beta_2, \dots, \beta_q), \quad (6.19)$$

where h_i are constant coefficients and $H_i(\beta_1, \beta_2, \dots, \beta_q)$ are orthonormal multivariate Hermite polynomials [6.23] with the weight function $e^{-\frac{1}{2}\beta^T\beta}$. Hence, if α_i ($i=1, \dots, m$) in Eq. (6.2) are mutually independent Gaussian random variables the associated solution x then has the following polynomial chaos expansion

$$x \approx \sum_i^M h_i H_i(\alpha_1, \alpha_2, \dots, \alpha_m) \quad (6.20)$$

with $H_i(\alpha_1, \alpha_2, \dots, \alpha_m)$ ($i=1, \dots, M$) generated by the weight function $e^{-\frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \dots + \alpha_m^2)}$.

As shown below, the unknown vector-valued coefficients h_i ($i=1, \dots, M$) can be solved

from Eq. (6.2) through a Galerkin approach whose shape functions are provided by these multivariate Hermite polynomials. Substituting (6.20) into (6.2) yields

$$\left(\sum_{i=1}^m \alpha_i A_i \right) \left(\sum_{j=1}^M h_j H_j(\alpha_1, \alpha_2, \dots, \alpha_m) \right) - \mathbf{b} \approx \mathbf{0}. \quad (6.21)$$

Multiplying by $H_k(\alpha_1, \alpha_2, \dots, \alpha_m)$ on both sides, Eq. (6.21) becomes

$$\left(\left(\sum_{i=1}^m \alpha_i A_i \right) \left(\sum_{j=1}^M h_j H_j \right) - \mathbf{b} \right) H_k \approx \mathbf{0} \quad (k=1, 2, \dots, M). \quad (6.22)$$

As these shape functions are random, the integral weak form of Eq. (6.21) is then obtained by enforcing the expectation to equal zero instead of the Riemann integration of the left-hand sides of Eq. (6.22), i.e.

$$E \left(\left(\left(\sum_{i=1}^m \alpha_i A_i \right) \left(\sum_{j=1}^M h_j H_j \right) - \mathbf{b} \right) H_k \right) = \mathbf{0} \quad (k=1, 2, \dots, M), \quad (6.23a)$$

$$\sum_{i=1}^m \sum_{j=1}^M E(\alpha_i H_j H_k) A_i h_j = E(H_k) \mathbf{b} \quad (k=1, 2, \dots, M). \quad (6.23b)$$

Therefore, the solution \mathbf{x} can be readily achieved in terms of the polynomial chaos expansion (6.20) after solving the n -vector-valued coefficients \mathbf{h}_j ($j=1, \dots, M$) from the above deterministic equation system of dimensionality Mn .

Due to the early theoretical work by Wiener et al. [6.18-6.20], the polynomial chaos expansion method [6.6-6.10] has a rigorous mathematical foundation, including its suitability and convergence. However, this method can only be strictly applied to solve equations consisting of Gaussian random variables, though for dynamic problems with a single random variable, different polynomials (e.g. single-variable Jacobi polynomials) [6.11] have recently been chosen, without proof of the suitability, to approximate functions of a non-Gaussian random variable. Furthermore, the complexity in terms of both the derivation of multivariate Hermite polynomials, i.e. the basic building blocks of this method, and the associated computational costs does increase exponentially as the number of random variables grows. This is because the total number of multivariate Hermite polynomials required in the solution for Eq. (6.2) is given by

$$M = \frac{(m+r)!}{m!r!}, \quad (6.24)$$

where r denotes the highest order of polynomials employed in the approximate

polynomial chaos expansion (6.20). An example is presented in Ref. [6.7] where a small stochastic equation system with six Gaussian random variables is considered. It was necessary to use up to fourth-order polynomials to satisfy the required accuracy, and consequently there are $\frac{(6+4)!}{6!4!} = 210$ six-variable Hermite polynomials in the fourth-order primary function set $\left\{H_j^{(0,1,\dots,4)}(\alpha_1, \alpha_2, \dots, \alpha_6)\right\}_{j=1}^{210}$. Therefore, the computational cost becomes a serious issue even for a rather small scale problem.

6.3 The Joint Diagonalization Strategy

It should be noted that the term “random matrix” has been avoided in this chapter since the discussions here have nothing to do with the existing theory of random matrices [6.15], which has recently been employed to develop a non-parametric SFEM formulation [6.14]. Furthermore, Eq. (6.2) is termed a stochastic system of linear algebraic equations compared to a standard deterministic system of linear algebraic equations (6.1). After the critical review of the existing methods, this section provides a pure algebraic treatment to a new solution strategy – termed the joint diagonalization - for the stochastic linear algebraic system (6.2).

6.3.1 The Formulation

The solution of Eq. (6.2) is well defined if the left-hand side sum of the matrices is non-singular almost surely, i.e.

$$P\left(|\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m| \neq 0\right) = 1 \quad (6.25)$$

The objective here is to invert the matrix sum $\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m$ to obtain an explicit solution of x in terms of random variables α_i ($i = 1, \dots, m$).

Assume that there exists an invertible matrix P to simultaneously diagonalize all the matrices A_i ($i = 1, \dots, m$) such that

$$P^{-1} A_i P = A_i = \text{diag}(\lambda_{i1}, \lambda_{i2}, \dots, \lambda_{in}) \quad (i = 1, \dots, m), \quad (6.26)$$

where λ_j ($j=1, \dots, n$) are eigenvalues of the $n \times n$ real symmetric matrix A . Eq. (6.2)

can then be transformed into:

$$P(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)P^{-1}x = b. \quad (6.27)$$

The solution x is given by

$$x = PA^{-1}P^{-1}b \quad (6.28)$$

with

$$A^{-1} = (\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)^{-1} \\ = \text{diag} \left(\frac{1}{\sum_{i=1}^m \alpha_i \lambda_{i1}}, \frac{1}{\sum_{i=1}^m \alpha_i \lambda_{i2}}, \dots, \frac{1}{\sum_{i=1}^m \alpha_i \lambda_{in}} \right). \quad (6.29)$$

Letting

$$P^{-1}b = (d_1, d_2, \dots, d_n)^T, \quad (6.30)$$

expression (6.28) becomes

$$x = P \begin{pmatrix} \left(\sum_{i=1}^m \alpha_i \lambda_{i1} \right)^{-1} & & & \\ & \left(\sum_{i=1}^m \alpha_i \lambda_{i2} \right)^{-1} & & \\ & & \ddots & \\ & & & \left(\sum_{i=1}^m \alpha_i \lambda_{in} \right)^{-1} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} \\ = P \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \begin{pmatrix} \left(\sum_{i=1}^m \alpha_i \lambda_{i1} \right)^{-1} \\ \left(\sum_{i=1}^m \alpha_i \lambda_{i2} \right)^{-1} \\ \vdots \\ \left(\sum_{i=1}^m \alpha_i \lambda_{in} \right)^{-1} \end{pmatrix} \quad (6.31)$$

Hence, in a more concise form, the solution is

$$x = D \left(\frac{1}{\sum_{i=1}^m \alpha_i \lambda_{i1}}, \frac{1}{\sum_{i=1}^m \alpha_i \lambda_{i2}}, \dots, \frac{1}{\sum_{i=1}^m \alpha_i \lambda_{in}} \right)^T \quad (6.32)$$

where

$$\mathbf{D} = \mathbf{P} \text{diag}(d_1, d_2, \dots, d_n). \quad (6.33)$$

Eq. (6.32), in which the coefficients \mathbf{D} and λ_{ij} ($i=1, \dots, m; j=1, \dots, n$) are constants determined completely by matrices \mathbf{A}_i ($i=1, \dots, m$) and vector \mathbf{b} , gives an explicit solution to Eq. (6.2). As a result, the associated joint probability distribution and statistical moments (e.g. expectation and covariance) can be readily computed. A key issue of the proposed strategy is therefore to obtain the transform matrix \mathbf{P} and the corresponding eigenvalues λ_{ij} ($i=1, \dots, m; j=1, \dots, n$), which is essentially an average eigenvalue problem and will be treated in detail in the next subsection.

6.3.2 A Jacobi-Like Algorithm for the Average Eigenvalue Problem

The eigenvalue problem of a single matrix is well understood in linear algebra. It is well known that a $n \times n$ real symmetric matrix \mathbf{A} can always be transformed into a real diagonal matrix $\mathbf{\Lambda}$ through an orthogonal similarity transformation, i.e.

$$\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} = \mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \quad (6.34)$$

where $\lambda_i \in \mathbb{R}$ ($i=1, \dots, n$) and $\mathbf{Q}^{-1} = \mathbf{Q}^T$. There are various numerical algorithms to obtain $\mathbf{\Lambda}$ and \mathbf{Q} , and among them the classical Jacobi method [6.21-6.22] diagonalizes \mathbf{A} by vanishing its off-diagonal elements through a sequence of Givens similarity transformations. Although Jacobi's original consideration is for a single matrix, his idea is rather general and the classical algorithm can be readily modified to accommodate multiple real symmetric matrices, as described below.

Denoted by \mathbf{G} , the orthogonal Givens matrix corresponding to a rotation angle θ is

$$(\mathbf{A}_k^*)_{pq} = (\mathbf{A}_k^*)_{qp} = \frac{1}{2} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right) \sin 2\theta + (\mathbf{A}_k)_{pq} \cos 2\theta, \quad (6.38f)$$

from which it is easy to verify that

$$(\mathbf{A}_k^*)_{pp}^2 + (\mathbf{A}_k^*)_{qq}^2 + 2(\mathbf{A}_k^*)_{pq}^2 = (\mathbf{A}_k)_{pp}^2 + (\mathbf{A}_k)_{qq}^2 + 2(\mathbf{A}_k)_{pq}^2 \quad (6.39)$$

The Frobenius norm of a matrix remains invariant under orthogonal similarity transformations [6.21-6.22], therefore the following equalities hold

$$\begin{aligned} \text{off}(\mathbf{A}_k^*) &= \|\mathbf{A}_k^*\|_F^2 - \sum_{i=1}^n (\mathbf{A}_k^*)_{ii}^2 \\ &= \|\mathbf{A}_k\|_F^2 - \sum_{\substack{i=1 \\ i \neq p, q}}^n (\mathbf{A}_k)_{ii}^2 - \left((\mathbf{A}_k^*)_{pp}^2 + (\mathbf{A}_k^*)_{qq}^2 \right), \\ &= \text{off}(\mathbf{A}_k) - 2(\mathbf{A}_k)_{pq}^2 + 2(\mathbf{A}_k^*)_{pq}^2 \end{aligned} \quad (6.40)$$

$$\sum_{k=1}^m \text{off}(\mathbf{A}_k^*) = \sum_{k=1}^m \text{off}(\mathbf{A}_k) - \sum_{k=1}^m 2(\mathbf{A}_k)_{pq}^2 + \sum_{k=1}^m 2(\mathbf{A}_k^*)_{pq}^2. \quad (6.41)$$

Hence, the minimization of $\sum_{k=1}^m \text{off}(\mathbf{A}_k^*)$ is equivalent to minimizing $\sum_{k=1}^m 2(\mathbf{A}_k^*)_{pq}^2$.

According to expressions (6.38), we have

$$\begin{aligned} \sum_{k=1}^m 2(\mathbf{A}_k^*)_{pq}^2 &= \sum_{k=1}^m 2 \left(\frac{1}{2} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right) \sin 2\theta + (\mathbf{A}_k)_{pq} \cos 2\theta \right)^2 \\ &= \sum_{k=1}^m \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \end{pmatrix}^T \begin{pmatrix} 2(\mathbf{A}_k)_{pq}^2 & (\mathbf{A}_k)_{pq} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right) \\ (\mathbf{A}_k)_{pq} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right) & \frac{1}{2} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right)^2 \end{pmatrix} \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \end{pmatrix} \\ &= (\cos 2\theta \quad \sin 2\theta) \mathbf{J} \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \end{pmatrix} \end{aligned} \quad (6.42)$$

where

$$\mathbf{J} = \sum_{k=1}^m \begin{pmatrix} 2(\mathbf{A}_k)_{pq}^2 & (\mathbf{A}_k)_{pq} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right) \\ (\mathbf{A}_k)_{pq} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right) & \frac{1}{2} \left((\mathbf{A}_k)_{qq} - (\mathbf{A}_k)_{pp} \right)^2 \end{pmatrix} \quad (6.43)$$

The left-hand side of equality (6.42) is a quadratic sum and the right-hand side is a real quadratic form. By noticing the fact that $(\cos 2\theta \quad \sin 2\theta)^T$ is a unit vector due to the trigonometric identity $\cos^2 2\theta + \sin^2 2\theta \equiv 1$ and is able to represent any unit vector in a

plane, the 2×2 real symmetric matrix J defined in (6.43) is concluded to be nonnegative definite since its corresponding quadratic form is always nonnegative. Let $e_1^{(j)}$ and $e_2^{(j)}$ denote the unit eigenvectors (different up to a sign coefficient -1) of J , and $\lambda_1^{(j)} \geq \lambda_2^{(j)} \geq 0$ the eigenvalues, the range of (6.42) is known from the theory of quadratic form, i.e.

$$\lambda_2^{(j)} \leq \sum_{k=1}^m 2(A_k^*)_{pq}^2 = (\cos 2\theta \quad \sin 2\theta) J \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \end{pmatrix} \leq \lambda_1^{(j)} \quad (6.44)$$

whose maximum and minimum are reached when $(\cos 2\theta \quad \sin 2\theta)^T$ is equal to $e_1^{(j)}$ and $e_2^{(j)}$ respectively. Hence, $\sum_{k=1}^m 2(A_k^*)_{pq}^2$ is minimized by setting $(\cos 2\theta \quad \sin 2\theta)^T$ equal to the unit eigenvector corresponding to the smaller eigenvalue of the 2×2 matrix J . Without loss of generality, $\cos 2\theta$ can be assumed always nonnegative. Therefore, the optimal Givens rotation angle θ_{opt} to minimize $\sum_{k=1}^m 2(A_k^*)_{pq}^2$ is uniquely determined by

$$(\cos 2\theta_{opt} \quad \sin 2\theta_{opt})^T = e_2^{(j)} \quad (\cos 2\theta_{opt} \geq 0), \quad (6.45)$$

from which the corresponding Givens matrix follows immediately.

Finally, the classical Jacobi algorithm [6.21-6.22] is modified as follows to accommodate multiple real symmetric matrices:

I) Sweep in turn all the entries of matrices A_k ($k=1, \dots, m$) and find an entry (p, q) $p \neq q$ such that

$$\sum_{k=1}^m (A_k)_{pq}^2 \neq 0. \quad (6.46)$$

II) For every entry (p, q) satisfying the above condition, compute the optimal Givens rotation angle θ_{opt} according to Eq. (6.45) and form the corresponding Givens matrix $G(p, q, \theta_{opt})$.

III) Apply Givens similarity transformation $G(p, q, \theta_{opt}) A_k G^{-1}(p, q, \theta_{opt})$ to all the matrices A_k ($k=1, \dots, m$) respectively and update these matrices into

A_k^* ($k=1, \dots, m$). (Note: only the p -th and q -th rows/columns in these matrices need to be updated.)

IV) Repeat the above procedure until the process converges.

As the Givens matrix corresponding to $\theta = 0$ is an identity matrix, we have

$$\sum_{k=1}^m 2(A_k^*)^2_{pq} \Big|_{\theta=\theta_{opt}} \leq \sum_{k=1}^m 2(A_k^*)^2_{pq} \Big|_{\theta=0} = \sum_{k=1}^m 2(A_k)_{pq}^2. \quad (6.47)$$

Substituting (6.47) into (6.41) yields

$$0 \leq \sum_{k=1}^m \text{off}(A_k^*) \Big|_{\theta=\theta_{opt}} \leq \sum_{k=1}^m \text{off}(A_k), \quad (6.48)$$

which indicates that $\sum_{k=1}^m \text{off}(A_k)$ is monotonously decreasing in this iterative procedure.

Therefore, the convergence to an average eigenstructure is guaranteed by the proposed Jacobi-like joint diagonalization algorithm. Assuming the above procedure has been performed K times with Givens matrices G_1, G_2, \dots, G_K respectively, the transform matrix P in (6.26) is then given by

$$P = G_1^{-1} G_2^{-1} \dots G_K^{-1} = G_1^T G_2^T \dots G_K^T, \quad (6.49)$$

and the corresponding eigenvalues λ_{ij} ($i=1, \dots, m$ $j=1, \dots, n$) are those diagonal entries in the final matrices A_k^* ($k=1, \dots, m$).

6.3.3 Discussions

- For the stochastic linear algebraic system (6.2) with one random variable, i.e.

$$\alpha_1 A_1 x = b, \quad (6.50)$$

the Jacobi-like algorithm for multiple real symmetric matrices reduces to the classical Jacobi algorithm for a single real symmetric matrix and the proposed joint diagonalization solution strategy gives the exact solution of x in this simple case, i.e.

$$\begin{aligned} x &= P A_1^{-1} P^{-1} b \\ &= P \text{diag} \left(\frac{1}{\alpha_1 \lambda_{11}}, \frac{1}{\alpha_1 \lambda_{12}}, \dots, \frac{1}{\alpha_1 \lambda_{1n}} \right) P^{-1} b = \frac{1}{\alpha_1} A_1^{-1} b. \end{aligned} \quad (6.51)$$

However, existing methods, such as the Monte Carlo method, the perturbation method, the Neumann expansion method and the polynomial chaos method, lack the above feature. In addition, it can be seen from Eq. (6.50) and (6.1), that the solution for the deterministic equation system (6.1) can be regarded as a special case of the proposed solver for the more general stochastic equation system (6.2).

- For the stochastic system of linear algebraic system (6.2) with more than one random variable, i.e. $m \geq 2$, the joint diagonalization can only be approximately achieved in this manner unless all the real symmetric matrices A_i ($i=1, \dots, m$) share exactly the same eigenstructure. Consequently, the approximate result is essentially an average eigenstructure that minimizes all the off-diagonal entries measured by $\sum_{k=1}^m \text{off}(A_k)$. It should be noted that in a practical problem, the approximate similarity among matrices A_i ($i=1, \dots, m$) is not only determined by the stochastic field of the physical problem under consideration but is also influenced by the method employed to construct these matrices, which is directly related to the choice of random variables α_i ($i=1, \dots, m$). The effectiveness and efficiency of the proposed approach depend on the degree of the eigenstructure similarity of the matrices involved. The applicability and limitations of the approach will be explored further in the next section.
- As shown in the explicit solution (6.32), the performance of the proposed solution strategy is not influenced by either the range or the type of random variations. In the previous sections, the right-hand side vector \mathbf{b} of Eq. (6.2) has been assumed deterministic, however, this restriction can be readily removed in the present approach, since there is essentially no intermediate operation required on \mathbf{b} until the final solution \mathbf{x} is calculated for the given random variables α_i ($i=1, \dots, m$).
- The major computational cost of the proposed approach is the Jacobi-like joint diagonalization procedure which is obviously proportional to m , the total number of matrices, as illustrated in the last section. This implies that the algorithm can be easily parallelized and the total computational cost is proportional to the total number of random variables in the system.

6.4 Numerical Examples

Three examples are employed to provide a numerical assessment of the overall performance of the proposed joint diagonalization method. All the numerical tests are conducted on a PC system with an Intel Xeon 2.4 GHz processor and 1.0 GB DDR memory.

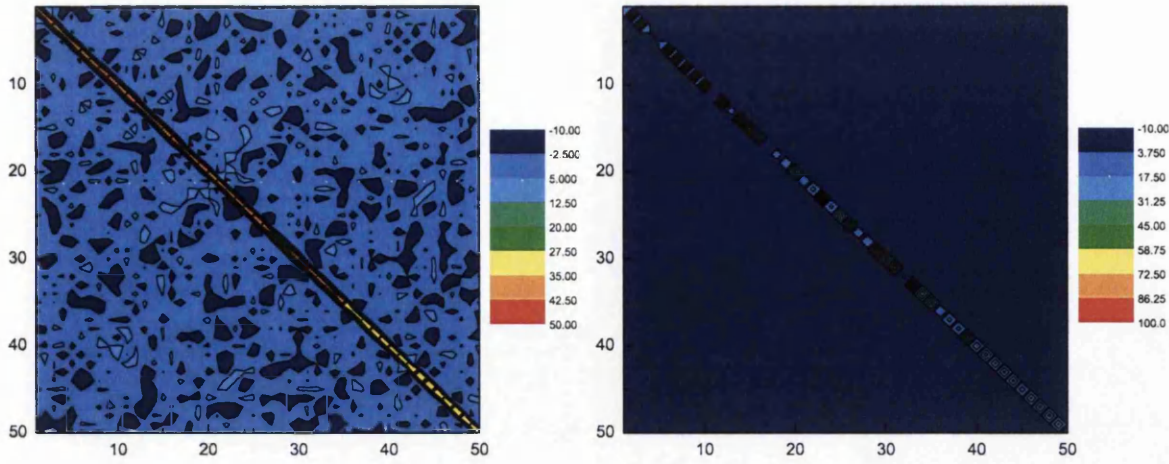
6.4.1 Example 1

This example is an artificially designed problem, aiming to demonstrate the effectiveness of the proposed procedure for matrices with similar eigenstructures. The matrices A_i ($i=1, \dots, m$) are generated by

$$A_i = Q(A_i + \Delta_i)Q^T \quad (i=1, \dots, m), \quad (6.52)$$

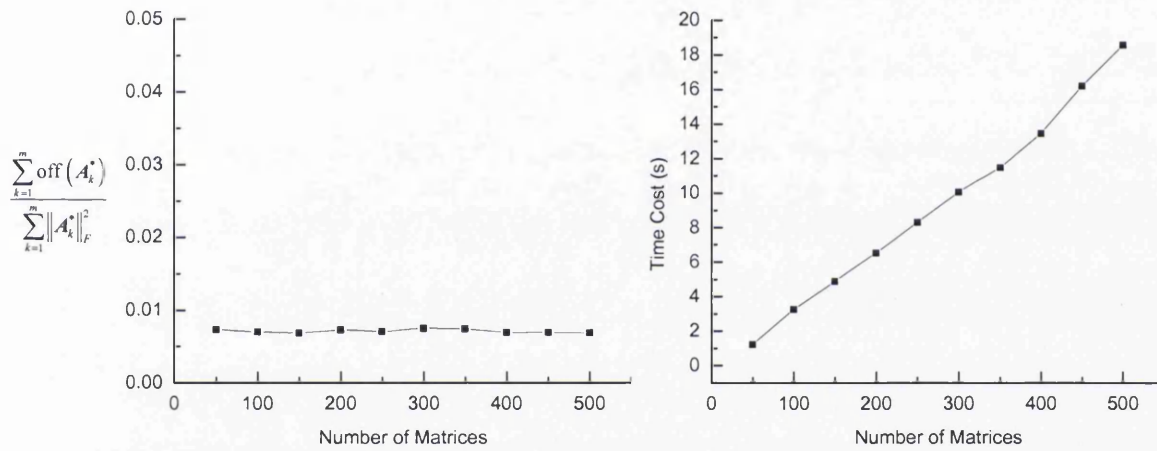
where matrix Q is a given orthogonal matrix that remains the same for all the matrices A_i ; and matrices A_i and Δ_i , representing the diagonal and the off-diagonal entries of the average eigenstructure respectively, are randomly generated matrices that differ with index i . In particular, matrices A_i are diagonal matrices whose diagonal entries are all positive; matrices Δ_i are real symmetric matrices whose diagonal entries are all zero and the absolute values of off-diagonal entries are relatively small (about 10% in this example) compared with the corresponding diagonal entry in matrix A_i .

A typical matrix A_i and its transformed counterpart $P^{-1}A_iP$ are respectively illustrated in Figures 6.1(a) and 6.1(b), where the entries in the matrix are linearly mapped into image pixels whose colours represent the values of the corresponding matrix entries. It can be seen that the Jacobi-like joint diagonalization algorithm significantly reduces the magnitudes of the off-diagonal entries when these matrices share approximately the same eigenstructure. As observed in Section 6.3.2, the computational cost of the proposed Jacobi-like joint diagonalization algorithm is proportional to the total number of matrices. This relationship is verified in Figure 6.2(b), where the number of matrices m ranges from 50 to 500 and approximately the same convergence level is achieved as shown in Figure 6.2(a).



(a) An example of matrix A_i (b) An example of matrix $P^{-1}A_iP$

Figure 6.1 Contour plots of matrices A_i and $P^{-1}A_iP$



(a) Convergence control level (b) Time cost of the Jacobi-like joint diagonalization

Figure 6.2 Computational cost of the Jacobi-like joint diagonalization algorithm

6.4.2 Example 2

As shown in Figure 6.3, the second example considers an elastic plane stress problem of an isotropic ERM D whose material properties are completely defined by its Young's modulus $E(\mathbf{x},\omega)$ and Poisson's ratio $\nu(\mathbf{x},\omega)$. Furthermore, for the sake of simplicity, it is assumed that the Poisson's ratio takes a constant value and the only random material property of D is the Young's modulus. Specifically,

$$\nu(\mathbf{x},\omega) = 0.3 \quad \text{a.s. for } \forall \mathbf{x} \in D, \tag{6.53}$$

$$E(E(\mathbf{x}, \omega)) = 3.0 \times 10^{10} \text{ Pa} \quad \forall \mathbf{x} \in D, \quad (6.54a)$$

$$\begin{aligned} & \text{Cov}(E(\mathbf{x}_1, \omega), E(\mathbf{x}_2, \omega)) \\ &= 6.561 \times 10^{19} e^{-\frac{(x_2-x_1)^2+(y_2-y_1)^2}{2.0^2}} \text{ Pa}^2 \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D, \end{aligned} \quad (6.54b)$$

Following the SFEM formulation described in Chapter 5, a stochastic system of linear algebraic equations is obtained, which consists of $m = 26$ real symmetric matrices associated with a constant parameter $\alpha_1 \equiv 1$ and mutually independent standard Gaussian random variables α_i ($i = 2, \dots, 26$).

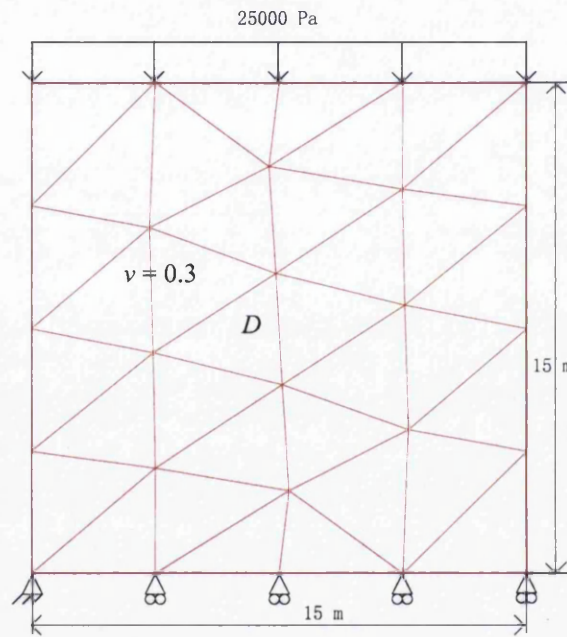


Figure 6.3 Structural illustration of Example 2

To demonstrate the effectiveness of the joint diagonalization procedure, the images of matrices A_1 and A_2 , and their transformed counterparts $A_1^* = P^{-1}A_1P$ and $A_2^* = P^{-1}A_2P$ are respectively depicted in Figures 6.4 and 6.5. It can be seen from the figures that the total value of the off-diagonal entries, i.e. $\sum_{k=1}^m \text{off}(A_k)$, is significantly reduced by the Jacobi-like joint diagonalization. From the resulting average eigenstructure, the ratio of the off-diagonal entries to the Frobenius norm is 9.655×10^{-4} , i.e.

$$\sum_{k=1}^m \text{off}(A_k^*) : \sum_{k=1}^m \|A_k^*\|_F^2 = 9.655 \times 10^{-4} : 1. \quad (6.55)$$

The corresponding convergence history is plotted in Figure 6.6.

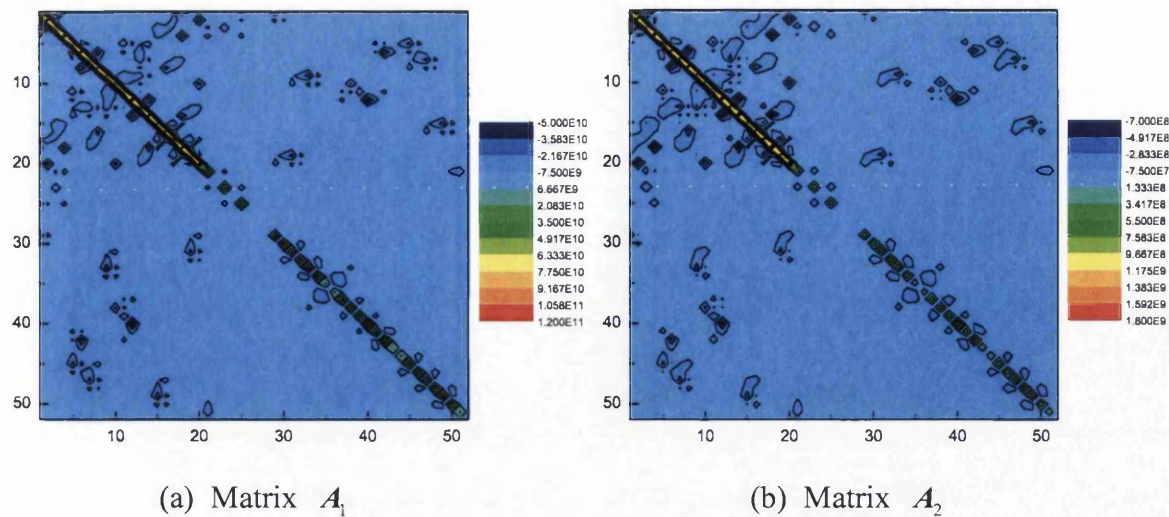


Figure 6.4 Contour plots of matrices A_1 and A_2

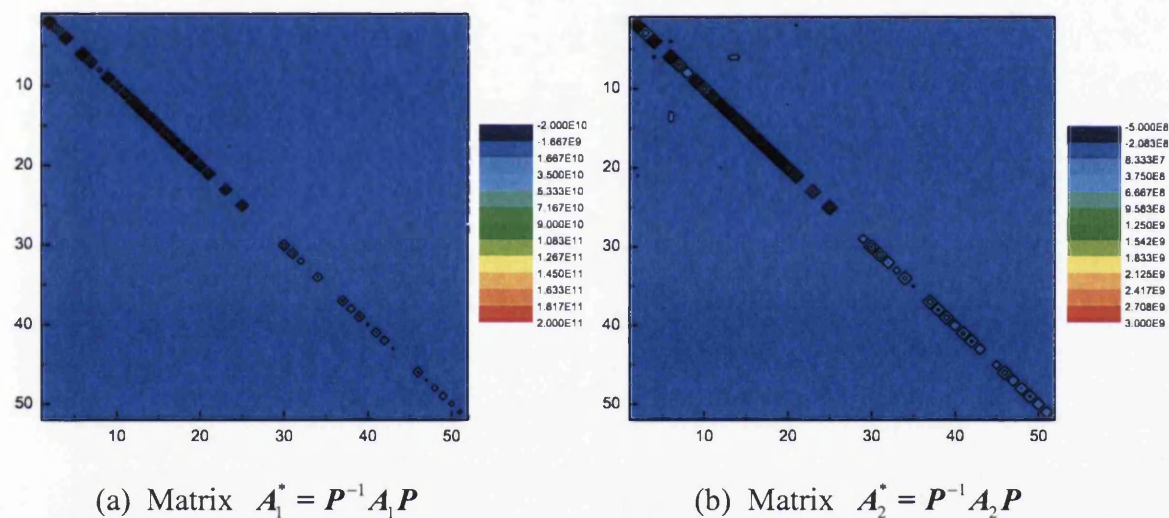


Figure 6.5 Contour plots of matrices $A_1^* = P^{-1}A_1P$ and $A_2^* = P^{-1}A_2P$

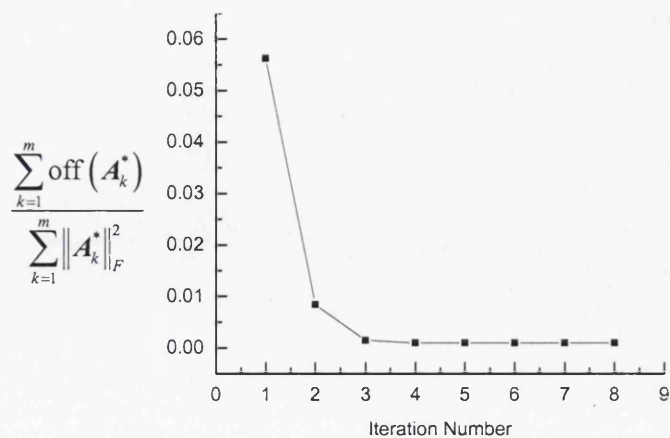


Figure 6.6 Convergence history of the Jacobi-like joint diagonalization

Due to the number of random variables used in this example and the fluctuation range of these random variables, only the Monte Carlo method and the Neumann expansion method (up to the sixth order expansion in Eq. (6.17)) are implemented for comparison with the current joint diagonalization method. For one sample path of α_i ($i=1, \dots, 26$), the comparison of the solutions x obtained respectively with the Monte Carlo method, the Neumann expansion method and the joint diagonalization method is illustrated in Figure 6.7. It can be seen that the proposed method obtains a good path-wise solution (strong solution) to the stochastic system of linear algebraic equations (6.2). Figure 6.8 shows the comparison for eight sample paths and the current method also gives a very good statistical solution (weak solution) to the stochastic linear algebraic system (6.2). In order to obtain good empirical statistics of the 52 unknown random variables (i.e. the nodal displacements) used in this example, 500 path-wise solutions are calculated with the three different methods. The corresponding CPU time costs of 500 solutions are recorded in Table 6.1. It can be seen that the joint diagonalization solution strategy exhibits the best performance in terms of efficiency. However, as the joint diagonalization algorithm proposed in this work is based on the classical Jacobi algorithm, this advantage may disappear for large-scale matrices.

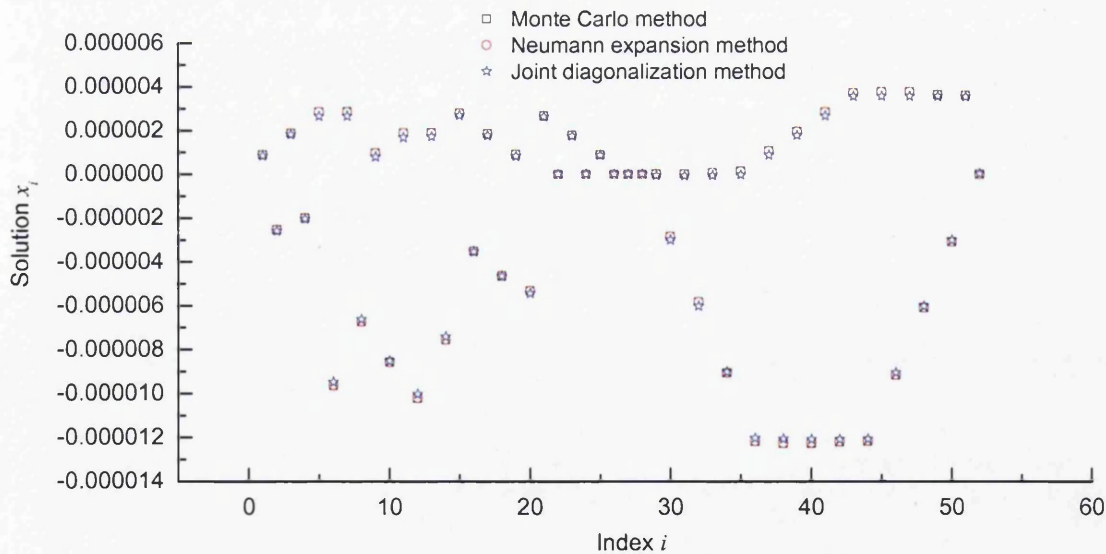


Figure 6.7 Solutions obtained with Monte Carlo, Neumann expansion and joint diagonalization methods (one sample path)

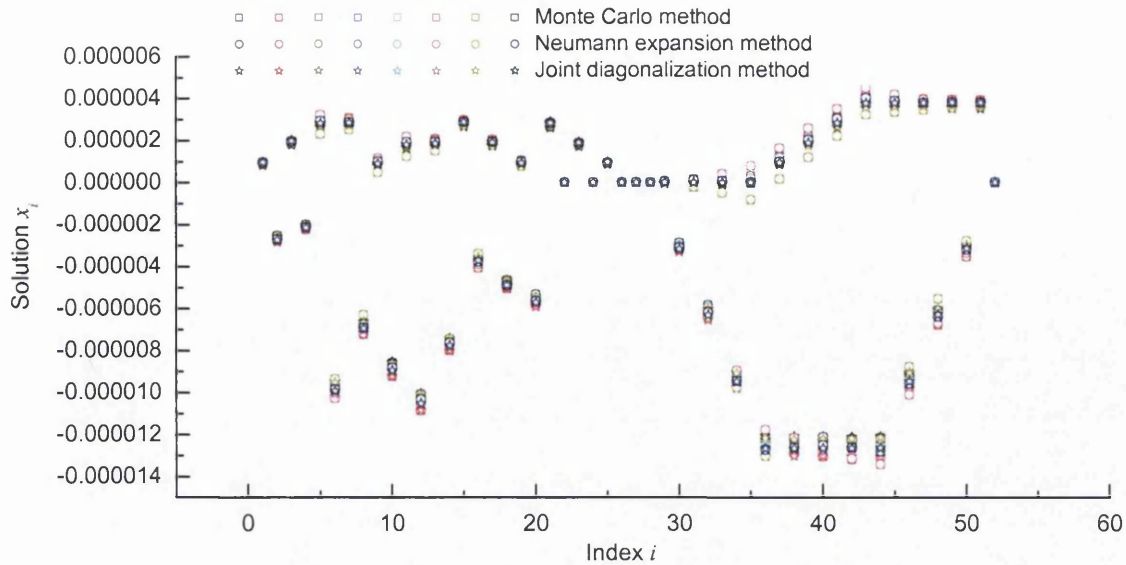
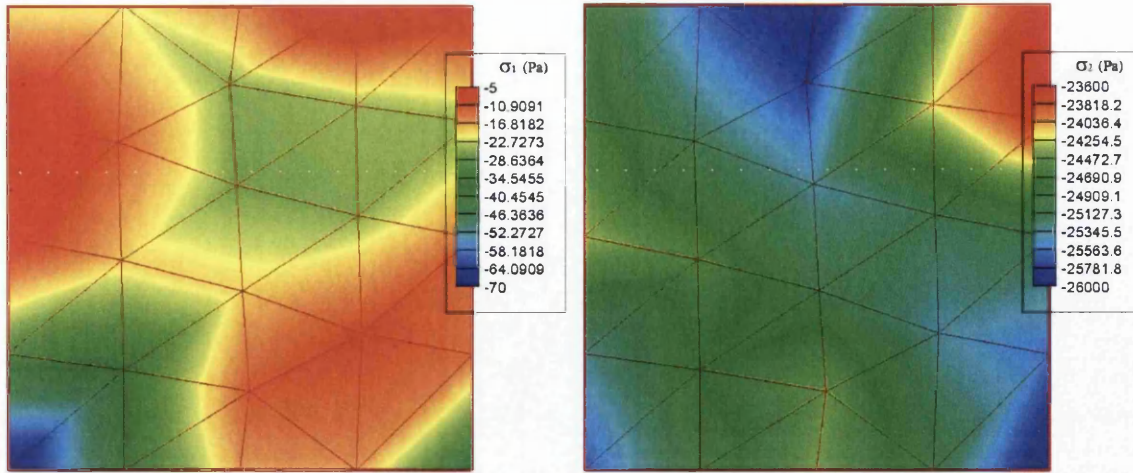


Figure 6.8 Solutions obtained with Monte Carlo, Neumann expansion and joint diagonalization methods (eight sample paths)

Table 6.1 CPU time costs of 500 solutions

	Monte Carlo method	Neumann expansion method	Joint diagonalization method
CPU time cost (s)	0.427	0.926	0.343

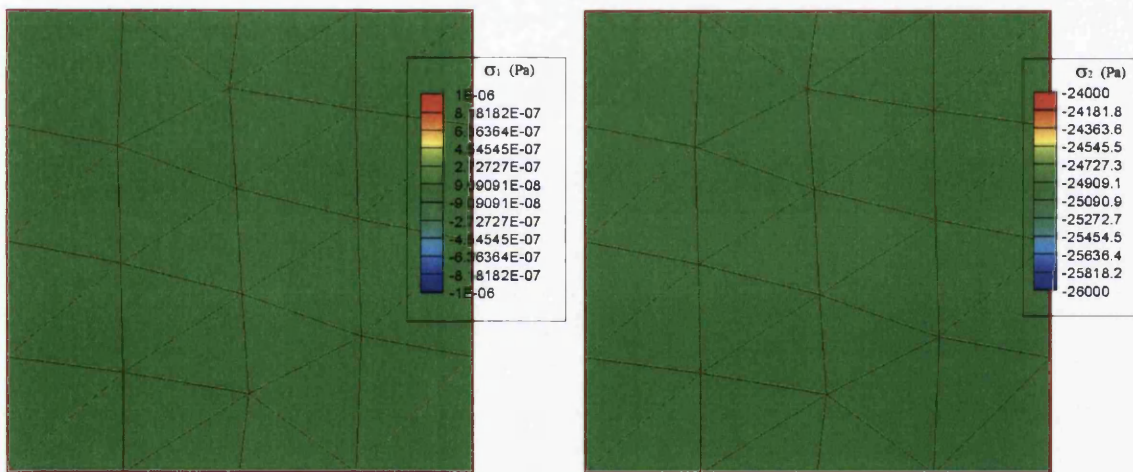
Finally, for one sample path of the random medium, the contour plots of principal stresses are shown in Figure 6.9. The deterministic counterpart of this simple example, in which the Poisson's ratio remains the same and the Young's modulus takes a fixed value $E(\mathbf{x}, \omega) \equiv 3.0 \times 10^{10}$ Pa, is also analysed using the standard finite element method with the same mesh structure, and the principal stresses are illustrated in Figure 6.10 for comparison. As shown in Figure 6.10(a), there is no horizontal stress distribution in this simple problem when the material properties are constant. However, due to the spatial variation of material properties, a horizontal stress distribution is observed in Figure 6.9(a). It can be seen in Figure 6.9(b) and Figure 6.10(b) that, although the principal stress level in the stochastic case is very close to the deterministic case, there is a visible random variation of the principal stresses resulting from the property variation throughout the random medium, as expected.



(a) Principal stress σ_1

(b) Principal stress σ_2

Figure 6.9 Principal stress distribution of the ERM (for one sample path)



(a) Principal stress σ_1

(b) Principal stress σ_2

Figure 6.10 Principal stress distribution of the corresponding deterministic medium

6.4.3 Example 3

A simplified tunnel model as shown in Figure 6.11 is considered in this example.

The associated ERM D is described by

Young's modulus

$$E(E(\mathbf{x}, \omega)) = 3.0 \times 10^{10} \text{ Pa} \quad \forall \mathbf{x} \in D, \quad (6.56a)$$

$$\begin{aligned} & \text{Cov}(E(\mathbf{x}_1, \omega), E(\mathbf{x}_2, \omega)) \\ &= 3.8025 \times 10^{20} e^{\frac{(x_2-x_1)^2+(y_2-y_1)^2}{2.0^2}} \text{ Pa}^2 \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in D; \end{aligned} \quad (6.56b)$$

Poisson's ratio

$$v(\mathbf{x}, \omega) = 0.3 \quad \text{a.s. for } \forall \mathbf{x} \in D. \quad (6.57)$$

The stochastic system of linear algebraic equations is formed through a similar procedure as used in the previous example. After joint diagonalization, the ratio of the off-diagonal entries to the Frobenius norm is 6.024×10^{-3} , and the corresponding convergence history is shown in Figure 6.12.

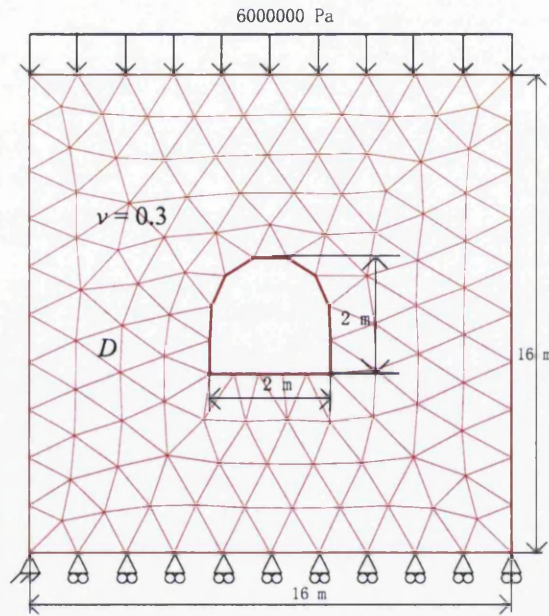


Figure 6.11 Illustration of Example 3

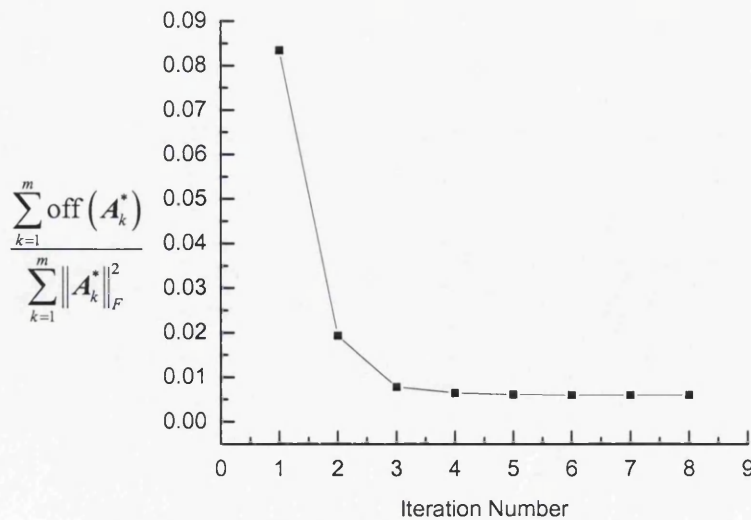


Figure 6.12 Convergence history of the Jacobi-like joint diagonalization

A corresponding deterministic model, sharing the same geometric configuration and the same Poisson's ratio, but setting Young's modulus to a fixed value 3.0×10^{10} Pa, is also analysed using the standard finite element method. Principal stress distributions of the stochastic model and the deterministic model are respectively shown in Figures 6.13 and 6.14, from which it can be observed that the stress distribution of the stochastic model is not only influenced by the model structure but also by the random material property variation throughout the medium.

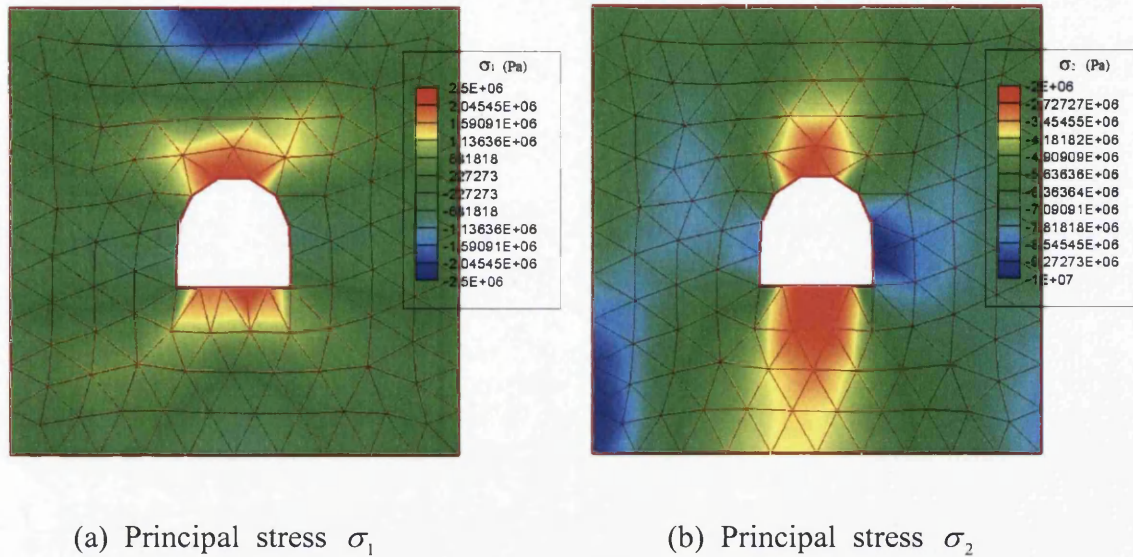


Figure 6.13 Principal stress distribution of a simplified tunnel model (for one sample path)

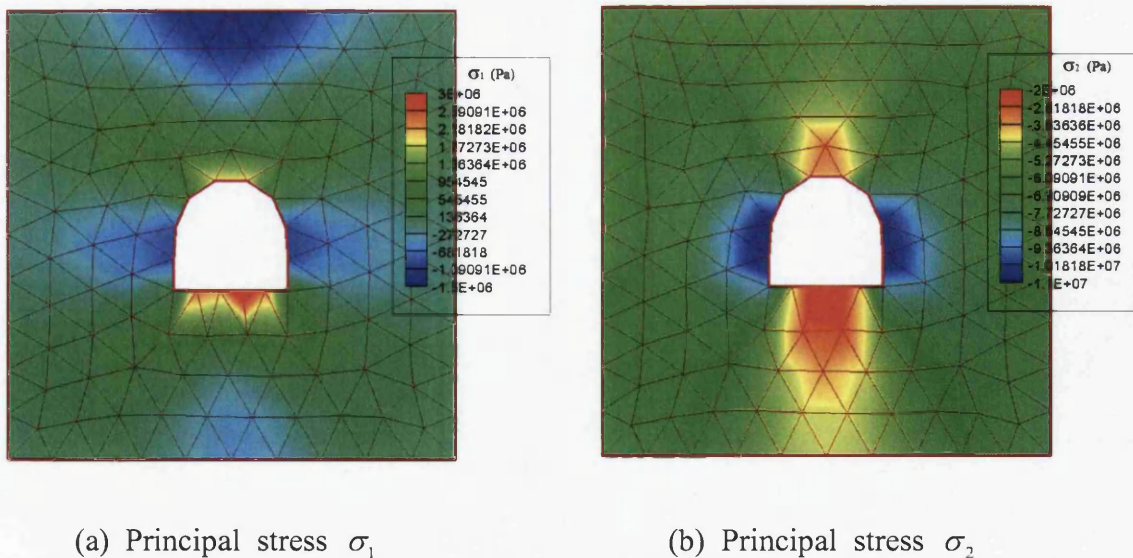


Figure 6.14 Principal stress distribution of the corresponding deterministic model

6.5 Summary and Suggestions for Future Research

6.5.1 Summary

This paper presents a novel strategy for solving the stochastic system of linear algebraic equations (6.2) arising from the SFEM formulation in Chapter 5. Firstly, the solution strategy simultaneously diagonalizes all the matrices in the system to obtain an average eigenstructure. The stochastic linear algebraic system is then decoupled by joint diagonalization, and its approximate solution is explicitly obtained by inverting the resulting diagonal stochastic matrix and performing the corresponding similarity transformation. Once the approximate solution is obtained in an explicit form, it is trivial to calculate the associated statistical moments and joint probability distributions. For the joint diagonalization, the classical Jacobi method has been modified for use with multiple symmetric matrices, while preserving the fundamental properties of the original version including the convergence and the explicit solution to the optimal Givens rotation angle. The computational cost of this Jacobi-like joint diagonalization algorithm is proportional to the total number of matrices in the system. This infers that it can be easily parallelized. For the solution of the stochastic system of linear algebraic equations, using the proposed approach, there is no restriction regarding either the range or the type of random variations in consideration.

Even though the presented strategy gives an explicit solution to Eq. (6.2) in a closed form, it is not advocated to use the Jacobi-like joint diagonalization algorithm for large-scale matrices. Indeed for a single matrix the classical Jacobi algorithm is not the most efficient method and does become extremely slow in dealing with a larger matrix. In this chapter, the joint diagonalization of multiple matrices is achieved through a similarity transformation by using an orthogonal matrix, thus its performance depends on the approximate similarity of the matrix family, which is not only determined by the stochastic field associated with the physical problem in consideration but is also strongly influenced by the method used to construct these matrices. These are the major limitations of the proposed Jacobi-like joint diagonalization algorithm.

It is well known that for a deterministic system of linear algebraic equations (6.1), there exist various numerical algorithms developed for different types of matrices and

different solution requirements, which are all explicitly or implicitly based on inverting the matrix under consideration. For a more general stochastic system of linear algebraic equations (6.2), it can be similarly expected that there will be different numerical algorithms based on the joint diagonalization of multiple matrices, which essentially give an approximate inverse of the matrix family as shown in the previous sections.

Parts of the results in this chapter have already been reported in [6.24-6.25].

6.5.2 Suggestions for Future Research

Regarding the joint diagonalization solution strategy proposed in this chapter, the following aspects need further investigation.

- In Sections 6.4.2 and 6.4.3, it is assumed that the Young's modulus is the only random material property of the ERM under consideration. This is because the full numerical investigation for the F-K-L representation scheme of the general elastic tensor of ERM has not been completed in Chapter 4 and consequently considering ERM with multiple random material properties may introduce unnecessary numerical errors to the stochastic linear algebraic system (6.2). In principle, the joint diagonalization solution strategy for Eq. (6.2) holds no matter how many independent random material parameters are actually contained in the ERM elastic tensor. Nevertheless, a thorough numerical investigation is important for the new approach and it should be performed with carefully designed examples when the examination of the F-K-L representation scheme is completed.
- As shown in Section 6.3, the performance of the Jacobi-like joint diagonalization solution depends on the similarity between those real symmetric matrices in the stochastic linear algebraic system. Hence, in order to obtain an error estimate for the approximate solution, it is necessary to further analyse the algebraic properties of the matrices involved, especially the properties of their average eigenstructure. The semi-analytic F-K-L representation addressed in Chapter 4 and the simple SFEM discretization described in Chapter 5 provide a good start to tackle this algebraic problem.
- The proposed Jacobi-like joint diagonalization method may provide an alternative way to solve the stochastic system of linear algebraic equations in small-scale cases. However, it should be noted that neither the similarity transformation nor the

orthogonal matrix is necessary and essential in this new solution strategy. For example, it is trivial to prove that for two real symmetric matrices A and B , if at least one of these is positive definite, then there exists an invertible matrix C , which is not necessarily (and usually is not) orthogonal, to simultaneously diagonalize both matrices through a congruent transformation such that $C^T A C$ and $C^T B C$ are real diagonal matrices. Indeed, it can be shown that the solution of Eq. (6.2) is mainly determined by the average eigenstructure corresponding to the smallest eigenvalues of matrices A_i , which can be approximately obtained by an explicit transform of individual eigenstructures of A_i . Hence, for the stochastic linear algebraic system with large-scale matrices, an improved algorithm which combines the current procedure with dimension-reduction techniques is very promising.

The joint diagonalization solution strategy together with some of the above research that is currently under development will be reported in [6.26-6.27].

References

- [6.1] W.K. Liu, T. Belytschko and A. Mani, Random field finite-elements, *International Journal for Numerical Methods in Engineering*, 23 (10) (1986) 1831-1845.
- [6.2] M. Kleiber and T.D. Hien, *The Stochastic Finite Element Method – Basic Perturbation Technique and Computer Implementation*, John Wiley & Sons, Chichester, 1992.
- [6.3] T.D. Hien and M. Kleiber, Stochastic finite element modelling in linear transient heat transfer, *Computer Methods in Applied Mechanics and Engineering*, 144 (1-2) (1997) 111-124.
- [6.4] F. Yamazaki, M. Shinozuka and G. Dasgupta, Neumann expansion for stochastic finite element analysis, *Journal of Engineering Mechanics – ASCE*, 114 (8) (1988) 1335-1354.
- [6.5] M. Shinozuka and G. Deodatis, Response variability of stochastic finite element systems, *Journal of Engineering Mechanics – ASCE*, 114 (3) (1988) 499-519.
- [6.6] R.G. Ghanem and P.D. Spanos, Spectral stochastic finite-element formulation

- for reliability-analysis, *Journal of Engineering Mechanics – ASCE*, 117 (10) (1990) 2351-2372.
- [6.7] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements – A Spectral Approach*, Revised edition, Dover Publications, New York, 2003.
- [6.8] M.K. Deb, I.M. Babuška and J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, *Computer Methods in Applied Mechanics and Engineering*, 190 (48) (2001) 6359-6372.
- [6.9] R.G. Ghanem and R.M. Kruger, Numerical solution of spectral stochastic finite element systems, *Computer Methods in Applied Mechanics and Engineering*, 129 (1996) 289-303.
- [6.10] M.F. Pellissetti and R.G. Ghanem, Iterative solution of systems of linear equations arising in the context of stochastic finite elements, *Advances in Engineering Software*, 31 (2000) 607-616.
- [6.11] D. Xiu and G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM Journal on Scientific Computing*, 24 (2) (2002) 619-644.
- [6.12] M. Papadrakakis M and V. Papadopoulos, Robust and efficient methods for stochastic finite element analysis using Monte Carlo simulation, *Computer Methods in Applied Mechanics and Engineering*, 134 (3-4) (1996) 325-340.
- [6.13] D.C. Charnpis and M. Papadrakakis, Improving the computational efficiency in finite element analysis of shells with uncertain properties, *Computer Methods in Applied Mechanics and Engineering*, 194 (2005) 1447-1478. (Included in Ref. [6.17])
- [6.14] C. Soize, Random matrix theory for modeling uncertainties in computational mechanics, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1333-1366. (Included in Ref. [6.17])
- [6.15] M.L. Mehta, *Random Matrices – Revised and Enlarged Second Edition*, Academic Press, New York, 1991.
- [6.16] H. Holden, B. Øksendal and Jan Ubøe et al., *Stochastic Partial Differential Equations – A Modelling White Noise Functional Approach*, Birkhäuser, Boston, 1996.
- [6.17] Edited by G.I. Schuëller, Special issue on computational methods in stochastic mechanics and reliability analysis, *Computer Methods in Applied Mechanics and Engineering*, 194 (12-16) (2005) 1251-1795.
- [6.18] N. Wiener, The homogeneous chaos, *American Journal of Mathematics*, 60 (40) (1938) 897-936.
- [6.19] R.H. Cameron and W.T. Martin, The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals, *Annals of Mathematics*,

48 (2) (1947) 385-392.

- [6.20] K. Itô, Multiple Wiener integral, *Journal of the Mathematical Society of Japan*, 3 (1) (1951) 157-169.
- [6.21] Z. Guan and Y.F. Lu, The Fundamental of Numerical Analysis, *Higher Education Press*, Beijing, 1998. (In Chinese)
- [6.22] C.E. Fröberg, *Introduction to Numerical Analysis*, Second edition, Addison-Wesley Publishing Company, Reading, 1970.
- [6.23] Y. Xu and C.F. Dunkl, *Orthogonal Polynomials of Several Variables (Encyclopedia of Mathematics and its Applications)*, Cambridge University Press, Cambridge, 2001.
- [6.24] C.F. Li, Y.T. Feng and D.R.J. Owen, Explicit solution to stochastic system of linear algebraic equations $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)x = b$, *Computer Methods in Applied Mechanics and Engineering*, 195 (44-47) (2006) 6560-6576.
- [6.25] Y.T. Feng, C.F. Li and D.R.J. Owen, Joint diagonalization solution for the stochastic system of linear algebraic equations $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)x = b$, *The Seventh World Congress on Computational Mechanics*, Los Angeles, California, USA, July 16-22, 2006.
- [6.26] C.F. Li, Y.T. Feng and D.R.J. Owen, Elastostatics of elementary random media: part I – theory and algorithms, *Journal of the Mechanics and Physics of Solids*, In preparation.
- [6.27] C.F. Li, Y.T. Feng and D.R.J. Owen, Elastostatics of elementary random media: part II – verification and application, *Journal of the Mechanics and Physics of Solids*, In preparation.

Chapter 7

A Directed Monte Carlo Solution for the Stochastic System of Linear Algebraic Equations

The main objective of this chapter is to propose a Monte Carlo based method for the solution of the stochastic system of linear algebraic equations (5.16) obtained in Chapter 5. To simplify the notation, Eq. (5.16) is rewritten as follows:

$$(K_o + K(\xi))u(\xi) = b \quad (7.1)$$

where

$$K(\xi) = \xi_1 K_1 + \xi_2 K_2 + \cdots + \xi_m K_m \quad \xi = (\xi_1, \xi_2, \dots, \xi_m), \quad (7.2)$$

in which K_o and K_i ($i=1,2,\dots,m$) are $N \times N$ deterministic matrices; ξ_i are m mutually independent standard Gaussian random variables; b is a deterministic[†] external load vector; and $u(\xi)$ is the random displacement vector to be sought.

The development of effective solution strategies for Eq. (7.1) to obtain various statistical properties of the solution $u(\xi)$ and the associated results, such as strain and stress, is one of the central issues in SFEM and becomes increasingly more important when larger scale problems with many random variables need to be considered. As reviewed in Chapter 6, various solution approaches have been proposed over the past two

[†] For the sake of simplicity, randomness in the external load vector is not considered in this chapter.

decades. Among all existing solution methods, Monte Carlo (MC) simulations are considered to be the most versatile approach and are in fact almost always used at some stage in most stochastic solution procedures. The main disadvantage of Monte Carlo simulations, however, is the intensive computational cost involved, particularly for large scale problems with many random variables, since a large number of samples are normally required to achieve a reasonable solution accuracy and a new system of equations needs to be solved at each MC sample.

In addition to those schemes traditionally proposed to improve the sampling strategies of Monte Carlo methods, including importance sampling, stratified sampling, recursive stratified sampling and adaptive Monte Carlo and especially VERGAS [7.2] (see [7.3] for a brief review on the topic), some progress has also been made in SFEM where the focus is on the solution of Eq. (7.1) by employing iterative approaches, such as Preconditioned Conjugate Gradients (PCG), so that large scale problems could be tackled. Recent developments in this aspect can be found, e.g. in [7.4-7.7]. One example is the recent work of [7.8] which proposes an explicit stochastic Incomplete Cholesky (IC) preconditioner for $K_o + K(\xi)$ based on the polynomial chaos expansion concept (see e.g. Ref. [7.1] or Section 6.2.4) to achieve a fast solution convergence of PCG at each Monte Carlo sampling point. It has to be pointed out however that in contrast to the solution of the deterministic system of linear algebraic equations, existing stochastic solution procedures are generally far from computationally adequate to handle large scale problems with many random variables, and therefore new and more advanced solution techniques are urgently needed in order to greatly improve the modelling capability of SFEM for practical applications.

The key feature of Monte Carlo samples is the randomness and unpredictability in the sequence of their spatial positions, in the space formed by the random variables, except for their statistical property. Therefore it appears necessary to solve a separate system of equations for (7.1) at each Monte Carlo sampling point of ξ . Note that this is always the case when Eq. (7.1) is solved by a direct solver. However, after all the Monte Carlo samples are generated, many of them may be closely positioned spatially, particularly in the high probability regions and when the number of the sampling points is large. This spatial proximity of Monte Carlo samples suggests that if an iterative solver is employed, the solution obtained at one or more MC sample may be used to provide a good initial

approximation for their immediate neighbours which could result in a significant reduction in the number of iterations required at these points. It is this observation that will be exploited in this chapter in order to improve the computational efficiency of MC simulations for the solution of Eq. (7.1).

The remainder of this chapter is organized as follows. In Section 7.1, the joint eigen properties of the stochastic matrix $\mathbf{K}(\xi)$ and the deterministic matrix \mathbf{K}_o will be briefly discussed to identify a possible very effective preconditioning scheme to be used later in PCG. For the solution of Eq. (7.1), a modified Monte Carlo approach, termed the directed Monte Carlo method and based on the utilisation of the spatial proximity of MC sampling points, will be developed first for equations with one random variable in Section 7.2. Two integrated numerical techniques essential for the success of the new method, including preconditioning and initial approximation prediction, are discussed in detail. The extension of the directed Monte Carlo approach to equations containing multiple random variables is undertaken by the adoption of a general hyper-spherical concept in Section 7.3. Numerical experiments are conducted in Section 7.4 to assess the performance of the proposed solution strategy and associated numerical techniques in terms of computational costs and solution accuracy. Finally, some conclusions regarding the proposed method are made in Section 7.5.

7.1 Properties of Stiffness Matrices

In addition to the statistical behaviour of the random vector ξ in Eq. (7.1), the relationship between \mathbf{K}_o and $\mathbf{K}(\xi) = \sum_{i=1}^m \xi_i \mathbf{K}_i$ also plays a paramount role in the solution of the stochastic linear algebraic system. Particularly, gaining a fundamental understanding of their underlying relationship will shed light onto the development of effective solution strategies for Eq. (7.1). This important issue is briefly addressed here.

Recall from Chapter 5 that the matrix \mathbf{K}_o is real symmetric and positive definite while the matrices \mathbf{K}_i ($i=1,2,\dots,m$) are real symmetric but not necessarily (and usually are not) positive definite. Hence, \mathbf{K}_o permits a standard \mathbf{LL}^T Cholesky

decomposition:

$$\mathbf{K}_o = \mathbf{L}_o \mathbf{L}_o^T \quad (7.3)$$

and, together with the symmetry of $\mathbf{K}(\xi)$, it ensures that the following generalised eigenproblem formed by \mathbf{K}_o and $\mathbf{K}(\xi)$ has real eigenvalues $\mathbf{A}_\xi = \text{diag}(\lambda_1^\xi, \lambda_2^\xi, \dots, \lambda_N^\xi)$

and corresponding real eigenvectors $\mathbf{V}_\xi = (\mathbf{v}_1^\xi, \mathbf{v}_2^\xi, \dots, \mathbf{v}_N^\xi)$:

$$\mathbf{K}(\xi) \mathbf{V}_\xi = \mathbf{K}_o \mathbf{V}_\xi \mathbf{A}_\xi \quad (7.4)$$

with

$$\mathbf{V}_\xi^T \mathbf{K}(\xi) \mathbf{V}_\xi = \mathbf{A}_\xi, \quad \mathbf{V}_\xi^T \mathbf{K}_o \mathbf{V}_\xi = \mathbf{I} \quad (\text{Identity matrix}) \quad (7.5)$$

and

$$\lambda_1^\xi \leq \lambda_2^\xi \leq \dots \leq \lambda_N^\xi. \quad (7.6)$$

The eigenpairs $\mathbf{A}_\xi(\xi)$ and $\mathbf{V}_\xi(\xi)$ also possess the following properties:

$$\mathbf{A}_\xi(\mathbf{0}) = \text{diag}(\mathbf{0}), \quad \mathbf{V}_\xi(\mathbf{0}) = (\mathbf{L}_o^T)^{-1} \quad \text{and} \quad \mathbf{A}_\xi(-\xi) = -\mathbf{A}_\xi(\xi). \quad (7.7)$$

With the aid of the above eigen-decomposition, the original equation (7.1) can be decoupled to

$$(\mathbf{I} + \mathbf{A}_\xi) \mathbf{V}_\xi^{-1} \mathbf{u}(\xi) = \mathbf{V}_\xi^T \mathbf{b} \quad (7.8)$$

which leads to an explicit solution of $\mathbf{u}(\xi)$:

$$\mathbf{u}(\xi) = \mathbf{V}_\xi (\mathbf{I} + \mathbf{A}_\xi)^{-1} \mathbf{V}_\xi^T \mathbf{b} = \sum_{i=1}^N \frac{\mathbf{b}^T \mathbf{v}_i^\xi}{1 + \lambda_i^\xi} \mathbf{v}_i^\xi. \quad (7.9)$$

Clearly, the eigenvalues λ_i^ξ ($i=1,2,\dots,N$) represent the variation scale caused by the random variables at the structural response level. The solution at each sampling point of ξ will depend on the actual spectrum of λ_i^ξ .

It is shown in Chapter 5 that the stochastic matrix sum $\mathbf{K}_o + \mathbf{K}(\xi)$ is, in the context of probability, real symmetric and positive definite. Thus, following the definition of real symmetric and positive definite matrices, it can be concluded that

$$\lambda_i^\xi > -1 \quad \text{for all } 1 \leq i \leq N \quad (7.10)$$

which, when taking into consideration of the property $\lambda_i^\xi(-\xi) = -\lambda_i^\xi(\xi)$, further leads to

$$\lambda_i^\xi < 1 \quad \text{for all } 1 \leq i \leq N. \quad (7.11)$$

Combining Eqs. (7.6), (7.10) and (7.11) yields

$$-1 < \lambda_1^\xi \leq \lambda_2^\xi \leq \dots \leq \lambda_N^\xi < 1 \quad (7.12)$$

Hence, in Eq. (7.9), those extreme eigenvalues at the lower end of the above eigenspectrum will have major contributions to the random variation of the solution $u(\xi)$.

Using $K_o = L_o L_o^T$ as a preconditioner to Eq. (7.1) results in

$$(I + \widehat{K}(\xi)) L_o^T u(\xi) = L_o^{-1} b \quad (7.13)$$

with

$$\widehat{K}(\xi) = L_o^{-1} K(\xi) (L_o^T)^{-1}. \quad (7.14)$$

It is trivial to observe that A_ξ is also the eigenvalue matrix of $\widehat{K}(\xi)$, i.e.

$$\widehat{V}_\xi^T \widehat{K}(\xi) \widehat{V}_\xi = A_\xi \quad \text{and} \quad \widehat{V}_\xi^T \widehat{V}_\xi = I \quad (7.15)$$

with the eigenvector matrix $\widehat{V}_\xi = L_o^T V_\xi$. Thus, following the eigenspectrum (7.12), the spectral condition number of $I + \widehat{K}(\xi)$ can be estimated as

$$\text{cond}_2(I + \widehat{K}(\xi)) = \frac{1 + \lambda_N^\xi}{1 + \lambda_1^\xi} < \frac{2}{1 + \lambda_1^\xi}. \quad (7.16)$$

When the minimum eigenvalue λ_1^ξ is not very close to -1 , which is normally the case for practical problems; $\text{cond}_2(I + \widehat{K}(\xi))$ will be in order $O(1)$. Thus, it is concluded that K_o is in general a good preconditioner for Eq. (7.1). It is also worth highlighting that at problem scales SFEM can currently handle effectively, the Cholesky decomposition is still very computationally competitive in terms of both memory and CPU time costs. Particularly, with the continuing advance of computer hardware, the scale of problems that a direct solver can effectively solve is also increasing. Furthermore, in the current problem concerned, a large number of repeated uses of the triangular decomposition L_o of K_o will significantly offset the overhead associated with the initial computation of L_o which further increases the effectiveness of the Cholesky decomposition. Thus the lower triangular matrix L_o and the deterministic solution u_d

$$u_d = u(0) = K_o^{-1} b \quad (7.17)$$

are assumed available when needed.

It is normally difficult to compute A_ξ and V_ξ as the general (random) functions of random vector ξ , but it is computationally feasible to obtain the joint eigen properties of each matrix K_i and K_o , particularly the two extreme eigenvalues, λ_1^i and λ_N^i . Investigating the eigenstructures of these individual matrices can reveal certain features of A_ξ and V_ξ . This issue however will not be pursued further in this chapter.

7.2 Modified Monte Carlo Simulation: One Random Variable

Let

$$p_\xi(\varsigma) = \frac{1}{(2\pi)^{\frac{m}{2}}} e^{-\frac{\varsigma \cdot \varsigma}{2}} \quad \varsigma \in \mathbb{R}^m \quad (7.18)$$

denote the joint probability density function of the random vector ξ in Eq. (7.1), then the expectation of an arbitrary function f of the solution $u(\xi)$ can be generally written as

$$E(f(u(\xi))) = \int_{\mathbb{R}^m} f(u(\varsigma)) p_\xi(\varsigma) d\varsigma. \quad (7.19)$$

The standard Monte Carlo simulation estimates the above probability integral by taking the arithmetic mean of the integral function $f(u(\xi))$ over M points $\{\varsigma_1, \varsigma_2, \dots, \varsigma_M\}$ sampled according to the probability density function $p_\xi(\varsigma)$:

$$E(f(u(\xi))) \approx \frac{1}{M} \sum_{i=1}^M f(u(\varsigma_i)). \quad (7.20)$$

The most expensive operation in the Monte Carlo simulation is that involved in the solution of (7.1) at each Monte Carlo point. The new DMC (directed Monte Carlo) approach, aimed at improving the computational efficiency of solving the equations on the basis of the utilisation of the spatial proximity of Monte Carlo sampling points, will be developed, first for one random variable cases in this section and then extended to general multiple random variable cases in the next section. This modified Monte Carlo approach consists of two essential ingredients: (a) an iterative algorithm is employed to solve the

linear system of equations at each Monte Carlo sample; and (b) the sampling points are first spatially sorted so that a high quality initial solution to the current sample can be obtained based on the solutions at its neighbours already solved.

7.2.1 Preconditioned Conjugate Gradient Method

The iterative solution of a linear system of equations is well established, and many solution algorithms have been proposed. Among all iterative solvers developed, the PCG (preconditioned conjugate gradient) algorithm is the most popular iterative solver for a real symmetric positive definite system. It is interesting to note that the PCG algorithm has been recently established to be equivalent to a second-order time integration scheme [7.9]. The standard PCG algorithm for solving a linear system

$$Kx = b \quad (7.21)$$

with given preconditioning matrix M_p (also assumed real symmetric and positive definite), initial guess x_0 , required solution accuracy τ and maximum iteration number I_{\max} , is provided below for reference.

Algorithm 7.1 Standard PCG algorithm: $x = \text{PCG}(K, b, x_0, M_p, \tau, I_{\max})$

- Compute $r_0 = b - Kx_0$. If $\|r_0\|_2 < \tau \|b\|_2$ return with $x = x_0$; else compute initial search direction $p_0 = M_p^{-1}r_0$

- For $i = 0, 1, 2, \dots, I_{\max}$:

1. Update solution:

$$x_{i+1} = x_i + \alpha_i p_i$$

where

$$\alpha_i = \frac{p_i^T r_i}{p_i^T K p_i} = \frac{r_i^T M_p^{-1} r_i}{p_i^T K p_i}$$

2. Update residual:

$$r_{i+1} = r_i - \alpha_i K p_i$$

3. Check convergence: if $\|r_{i+1}\|_2 < \tau \|b\|_2$, return with $x = x_{i+1}$

4. Compute new search direction:

$$\mathbf{p}_{i+1} = \mathbf{M}_p^{-1} \mathbf{r}_{i+1} + \beta_i \mathbf{p}_i$$

where

$$\beta_i = -\frac{\mathbf{p}_i^\top \mathbf{K} \mathbf{M}_p^{-1} \mathbf{r}_{i+1}}{\mathbf{p}_i^\top \mathbf{K} \mathbf{p}_i} = \frac{\mathbf{r}_{i+1}^\top \mathbf{M}_p^{-1} \mathbf{r}_{i+1}}{\mathbf{r}_i^\top \mathbf{M}_p^{-1} \mathbf{r}_i}$$

As all the computational aspects of the standard PCG have been thoroughly investigated, no new development will be offered in this work. However there are two integrated issues that are essential to the success of the DMC method, and therefore deserve further discussions: (a) preconditioning; and (b) initial guess. The latter issue will be elaborated upon in the next subsection.

Preconditioning is the most important technique for PCG to achieve a fast convergence for practical problems by providing a possible significant reduction in the spectral condition number of the original matrix. IC (incomplete Cholesky) decomposition and its variants are most commonly used preconditioning schemes [7.10]. Let the IC decomposition of a matrix \mathbf{K} be denoted as $\text{IC}(\mathbf{K})$. For the current problem concerned, since the matrix $\mathbf{K}(\xi)$ varies at every sampling point, there exist several options for the preconditioning matrix \mathbf{M}_p :

- **IC0:** $\mathbf{M}_p = \text{IC}(\mathbf{K}_o)$, the IC decomposition of \mathbf{K}_o ;
- **SIC:** $\mathbf{M}_p = \text{IC}(\mathbf{K}_o + \mathbf{K}(\xi))$, the IC decomposition of $\mathbf{K}_o + \mathbf{K}(\xi)$;
- **K0:** $\mathbf{M}_p = \mathbf{K}_o = \mathbf{L}_o \mathbf{L}_o^\top$, the complete Cholesky decomposition of \mathbf{K}_o .

The first option, denoted IC0, uses the IC of \mathbf{K}_o as the preconditioning matrix for all the sampling points. The apparent attractive feature is its computational efficiency as it is generated only once. It is argued however that its performance may not be sufficient when large scale random variations are present. The second choice, denoted SIC, seems desirable because a new IC decomposition is generated for $\mathbf{K}_o + \mathbf{K}(\xi)$ at each sampling point of ξ . The obvious disadvantage of this option is the relatively expensive computational costs involved in the generation of $\text{IC}(\mathbf{K}_o + \mathbf{K}(\xi))$ at the sampling points.

It is this concern that is addressed in [7.8] which proposes to express $\text{IC}(\mathbf{K}_o + \mathbf{K}(\xi))$ as a series of ξ based on the polynomial chaos expansion (see e.g. Ref. [7.1] or Section 6.2.4) to reduce the computational costs. Note that both IC and SIC decompositions may fail when the diagonal terms become negative during the decomposition. Although artificial amendments can be made to resume the procedure, the resulting decomposition often does not perform well.

The third option, denoted K0, uses the complete Cholesky decomposition of \mathbf{K}_o as the preconditioning matrix. This option is highly recommended in the present work supported by the analysis conducted in Section 7.1 for the following reasons: (a) unlike the first two schemes, this option always achieves a good convergence regardless of the properties of \mathbf{K}_o , unless the random variation is very high; and (b) the computational procedure involved in the decomposition is very stable while the other IC decompositions can fail. Note that the K0 scheme with further improvement has already been adopted in [7.5]. In Section 7.4, numerical examples will be provided to compare the performance of these three preconditioning schemes.

Generally a good initial approximation can also reduce the number of PCG iterations. However, it may not be easy to provide such an approximation apart from some special cases, as will be further discussed later. If no other information is available, the deterministic solution \mathbf{u}_d may always be considered as a fairly good candidate.

Note that the above discussion is applied to both one and multiple random variable cases.

7.2.2 Exploitation of Spatial Proximity of Monte Carlo Points

When only one random variable ξ is present, Eq. (7.1) is reduced to

$$(\mathbf{K}_o + \xi \mathbf{K}) \mathbf{u}(\xi) = \mathbf{b}. \quad (7.22)$$

Suppose that the total number of Monte Carlo samples required is M and all the sampling points, $\mathbb{M} = \{\xi_1, \xi_2, \dots, \xi_M\}$, for the random variable ξ are generated before the solution of Eq. (7.22) is performed. Normally these points bear no spatial sequence in \mathbb{R} while generated. However, when they are *sorted* in ascending order, their spatial

relationship becomes apparent. Therefore these sampling points in the set M can be assumed to be in order

$$\zeta_1 < \zeta_2 < \dots < \zeta_M. \quad (7.23)$$

This *reorganisation* of the sampling points makes no difference to the standard Monte Carlo simulation, but it is crucial for the modified Monte Carlo approach proposed here.

Following the reorder, the key to the effective exploitation of the localisation is how to provide a better guess for the current sample if the solutions at previous or neighbouring samples are known. This can be approached in two slightly different ways.

First, suppose that the solution $\mathbf{u}(\xi)$ to (7.22) at $\xi = \zeta$ is obtained and the solution $\mathbf{u}(\xi)$ at $\xi = \zeta + \Delta\zeta$ is sought:

$$(\mathbf{K}_o + (\zeta + \Delta\zeta)\mathbf{K})\mathbf{u}(\zeta + \Delta\zeta) = \mathbf{b}. \quad (7.24)$$

By employing a first order Taylor series approximation, $\mathbf{u}(\zeta + \Delta\zeta)$ has the following approximation

$$\mathbf{u}(\zeta + \Delta\zeta) \approx \mathbf{u}(\zeta) + \left. \frac{d\mathbf{u}(\xi)}{d\xi} \right|_{\xi=\zeta} \Delta\zeta \quad (7.25)$$

where the gradient $d\mathbf{u}(\xi)/d\xi$ can be readily obtained from (7.22) as

$$\frac{d\mathbf{u}(\xi)}{d\xi} = -(\mathbf{K}_o + \xi\mathbf{K})^{-1} \mathbf{K}\mathbf{u}(\xi). \quad (7.26)$$

However, to evaluate the above gradient is equivalent to the solution of a new linear system, and therefore using (7.25) to find an approximation to $\mathbf{u}(\zeta + \Delta\zeta)$ is not very attractive. Nevertheless, it is straightforward to compute the gradient at $\xi = 0$ since it is assumed that the Cholesky decomposition of \mathbf{K}_o is available:

$$\left. \frac{d\mathbf{u}(\xi)}{d\xi} \right|_{\xi=0} = -\mathbf{K}_o^{-1} \mathbf{K}\mathbf{u}_d. \quad (7.27)$$

A better approach is to utilise additional information to approximate the gradient $d\mathbf{u}(\xi)/d\xi$. Now suppose that the solutions $\mathbf{u}(\xi)$ at the sampling points from ζ_1 to ζ_i have been found and the aim is to use these results to obtain a high quality initial guess $\mathbf{u}_0(\xi)$ at $\xi = \zeta_{i+1}$ for the corresponding PCG iterations. This is basically an

interpolation/extrapolation problem, and many standard numerical schemes are available. Here the following two schemes are proposed for use:

- Linear interpolation/extrapolation scheme. In this scheme, the gradient $du(\xi)/d\xi$ at $\xi = \zeta_i$ is approximated by a forward Euler difference as

$$\left. \frac{du(\xi)}{d\xi} \right|_{\xi=\zeta_i} \approx \frac{1}{\zeta_i - \zeta_{i-1}} (u(\zeta_i) - u(\zeta_{i-1})), \quad (7.28)$$

then it follows that

$$u_0(\zeta_{i+1}) \approx u(\zeta_i) + \frac{\zeta_{i+1} - \zeta_i}{\zeta_i - \zeta_{i-1}} (u(\zeta_i) - u(\zeta_{i-1})). \quad (7.29)$$

- Cubic spline interpolation/extrapolation scheme. In this scheme, a cubic spline is generated first to interpolate all the existing solutions (componentwise) and then is used to compute an initial approximation solution at $\xi = \zeta_{i+1}$.

The first scheme is very simple to implement. Its performance is however dependent on the spacings $\zeta_i - \zeta_{i-1}$ and $\zeta_{i+1} - \zeta_i$. At the expense of more computational costs, the second scheme usually provides better approximation if the spacings between the sampling points are all fairly uniform. However, random samples cannot always satisfy this requirement and occasionally behave irrationally. Particularly when $\zeta_i - \zeta_{i-1}$ is significantly larger or smaller than the previous spacings, the predicated initial solution at ζ_{i+1} is normally poor, while on the other hand the linear scheme is not sensitive to the spacings. In general, as will be demonstrated numerically in Section 7.4, the spline scheme exhibits better performance when a relatively small number of sampling points are present, while the linear scheme behaves satisfactory when a large number of samples are present.

In both schemes, the first two solutions need to be computed differently. The first solution $u(\zeta_1)$ may be obtained probably with u_d as the initial guess, which is then used as the initial approximation for the next solution $u(\zeta_2)$. When the sample values ζ_i ($i=1, 2, \dots, M$) are roughly symmetric about the origin, which is true for most cases, there exists an alternative sample processing strategy. The sorted sample sequence (7.23) can be split into two parts so that the samples in the first part are all smaller than zero and larger than zero in the second part:

$$\zeta_1 < \cdots < \zeta_j < 0 < \zeta_{j+1} < \cdots < \zeta_M. \quad (7.30)$$

Now treat $\mathbf{u}_d = \mathbf{u}(0)$ as the first solution, and since the gradient $du(\xi)/d\xi$ at $\xi = 0$ is available from (7.27), (7.25) can be used to compute the initial guess for $\xi = \zeta_j$ to obtain the solution $\mathbf{u}(\zeta_j)$. Then the solution from ζ_{j-1} down to ζ_1 can be found following the above two interpolation/extrapolation schemes. The second part can be processed, starting from ζ_{j+1} until ζ_M , in a similar manner. This latter strategy is numerically proved to be slightly more effective.

The main steps involved in the DMC approach proposed above are summarised in Algorithm 7.2, where the sampling points are assumed to be processed from ζ_1 .

Algorithm 7.2 Directed Monte Carlo simulation – one random variable:

$$\text{DMC1}(K_o, K, b, f, M, M, p_\xi(\zeta))$$

- If M is empty, generate M random samples $M = \{\zeta_1, \zeta_2, \dots, \zeta_M\}$ according to the given probability density function $p_\xi(\zeta)$. Sort the samples in M in ascendant order.
- Compute $\mathbf{u}_d = K_o^{-1}b$ (and $\left. \frac{du(\xi)}{d\xi} \right|_{\xi=0} = -K_o^{-1}K\mathbf{u}_d$ if required).
- Set initial preconditioning matrix M_p (and update it for the second point if required).
- Compute the first two solutions:

$$\mathbf{u}_1 = \text{PCG}(K_o + \zeta_1 K, b, \mathbf{u}_d, M_p, \tau, I_{\max})$$

$$\mathbf{u}_2 = \text{PCG}(K_o + \zeta_2 K, b, \mathbf{u}_1, M_p, \tau, I_{\max})$$
- For $k = 3, 4, \dots, M$:
 1. Update preconditioning matrix M_p if required.
 2. Compute initial guess $\mathbf{u}_0(\zeta_i)$ based on the linear or spline scheme.

3. Obtain solution

$$\mathbf{u}_i = \text{PCG}(\mathbf{K}_o + \varsigma_i \mathbf{K}, \mathbf{b}, \mathbf{u}_0(\varsigma_i), M_p, \tau, I_{\max})$$

- Calculate the expectation:

$$E(f(\mathbf{u}(\xi))) = \frac{1}{M} \sum_{i=1}^M f(\mathbf{u}(\varsigma_i))$$

7.3 Modified Monte Carlo Simulation: Multiple Random Variable

The DMC solution strategy proposed in the previous section cannot be readily extended to multiple random variables. The main challenge is the fact that spatial points in an m -dimensional domain do not have an in-built “order” as in the one dimensional case. In this section, this difficulty is overcome by further modifying the standard Monte Carlo method with the introduction of a hyper-spherical transformation to the simulation.

7.3.1 Hyper-Spherical Transformation

Let \mathbb{S}_m denote the surface of a “unit m -sphere”, i.e. an m -dimensional hyper-sphere of unit radius. That is, the relation

$$\hat{\xi} \cdot \hat{\xi} = \hat{\xi}_1^2 + \hat{\xi}_2^2 + \cdots + \hat{\xi}_m^2 = 1 \quad (7.31)$$

holds for any point $\hat{\xi} = (\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_m)$ on \mathbb{S}_m . The total area of \mathbb{S}_m is

$$A_m = \frac{2\pi^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)} \quad (7.32)$$

where $\Gamma(\bullet)$ is the gamma function.

With the assistance of the above notation, the Gaussian random vector $\xi = (\xi_1, \xi_2, \dots, \xi_m)$ in Eq. (7.1) can be expressed as

$$\xi = \xi_r \hat{\xi} \quad (7.33)$$

where $\hat{\xi}$ is a random vector on \mathbb{S}_m and

$$\xi_r = \sqrt{\xi \cdot \xi} = (\xi_1^2 + \xi_2^2 + \dots + \xi_m^2)^{\frac{1}{2}} \geq 0 \quad (7.34)$$

is a random variable. This is basically a general hyper-spherical transformation where ξ_r and $\hat{\xi}$ represent respectively the “magnitude” (or “radius”) and the unit “direction” of ξ , as illustrated in Figure 7.1 in a polar coordinate sense.

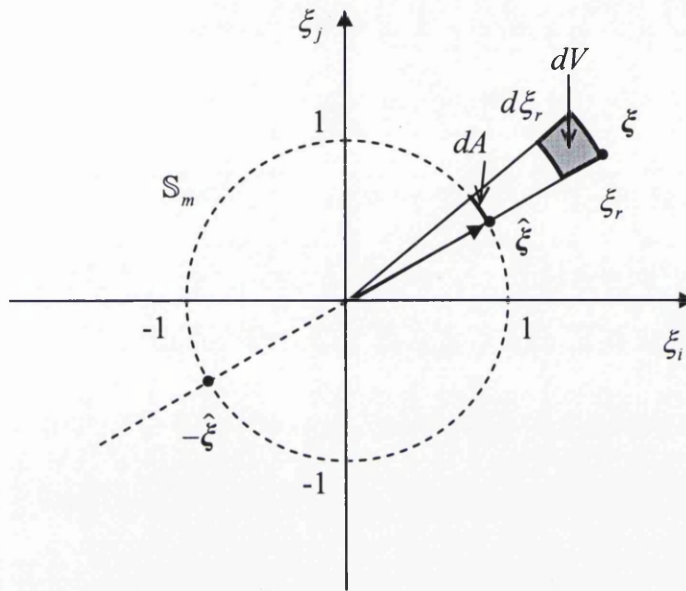


Figure 7.1 2D illustration of hyper-spherical transformation

In Eq. (7.33), $\hat{\xi}$ and ξ_r represent respectively the normalized orientation and the collective variation scale of the mutually independent standard Gaussian random variables ξ_i ($i=1,2,\dots,m$), and their probability distributions are uniquely determined by the joint probability distribution of ξ . Specifically, $\hat{\xi}$ and ξ_r are independent, and their probability density functions can be obtained as

$$p_{\xi_r}(\xi_r) = \begin{cases} \left(\xi_r^{m-1} e^{-\frac{\xi_r^2}{2}} \right) / \left(2^{\frac{m}{2}-1} \Gamma\left(\frac{m}{2}\right) \right) & \xi_r > 0 \\ 0 & \xi_r \leq 0 \end{cases} \quad (7.35)$$

and

$$p_{\hat{\xi}}(\hat{\xi}) = \begin{cases} 1/A_m & \hat{\xi} \cdot \hat{\xi} = 1 \\ 0 & \text{other} \end{cases} \quad (7.36)$$

respectively. As $\xi_r \in \mathbb{R}^+$, where \mathbb{R}^+ denotes the space formed by zero and all positive real numbers, the introduced hyper-spherical transformation (7.33) permits the following decomposition of the original ‘‘Cartesian’’ space \mathbb{R}^m

$$\mathbb{R}^m = \mathbb{R}^+ \times \mathbb{S}_m. \quad (7.37)$$

Without giving a detailed derivation, the volume element $d\zeta$ in the space \mathbb{R}^m has the following relationship with an infinitesimal ‘‘line’’ element $d\zeta_r$ on the ‘‘radial’’ direction and an infinitesimal ‘‘surface’’ element dA on \mathbb{S}_m :

$$d\zeta = \zeta_r^{m-1} d\zeta_r dA. \quad (7.38)$$

With the above preparations, the expectation of the function $f(\mathbf{u}(\xi))$, which is originally calculated in the Cartesian space \mathbb{R}^m in (7.19), can now be cast in the hyper-spherical space as

$$E(f(\mathbf{u}(\xi))) = \int_{\mathbb{S}_m} \int_{\mathbb{R}^+} f(\mathbf{u}(\zeta_r, \hat{\zeta})) \frac{1}{(2\pi)^{\frac{m}{2}}} e^{-\frac{\zeta_r^2}{2}} \zeta_r^{m-1} d\zeta_r dA. \quad (7.39)$$

Using the probability density functions $p_{\hat{\zeta}}(\hat{\zeta})$ and $p_{\zeta_r}(\zeta_r)$, the above expression can be rewritten as

$$E(f(\mathbf{u}(\xi))) = \int_{\mathbb{S}_m} f_{\xi_r}(\hat{\zeta}) p_{\hat{\zeta}}(\hat{\zeta}) dA, \quad (7.40)$$

in which

$$f_{\xi_r}(\hat{\zeta}) = \int_{\mathbb{R}^+} f(\mathbf{u}(\zeta_r, \hat{\zeta})) p_{\zeta_r}(\zeta_r) d\zeta_r. \quad (7.41)$$

It is noted that, when $\xi = \xi_r \hat{\xi}$ is fixed at $\hat{\xi} = \hat{\zeta}$ on \mathbb{S}_m , $\mathbf{u}(\xi)$ is the solution to the following linear equations with one random variable ξ_r :

$$(\mathbf{K}_o + \xi_r \mathbf{K}_{\hat{\zeta}}) \mathbf{u}(\xi_r, \hat{\zeta}) = \mathbf{b} \quad (7.42)$$

in which

$$\mathbf{K}_{\hat{\zeta}} = \hat{\zeta}_1 \mathbf{K}_1 + \hat{\zeta}_2 \mathbf{K}_2 + \cdots + \hat{\zeta}_m \mathbf{K}_m \quad (7.43)$$

is a deterministic matrix. Eqs. (7.40-7.43) readily suggest a new Monte Carlo simulation procedure in which the solution strategy developed in the previous section for the system with one random variable can now play an important part.

7.3.2 Modified Monte Carlo Simulation in Hyper-Spherical Space

Now the Monte Carlo simulation can be undertaken in two different spaces as follows:

I) Generate a set of the required number ($M_{\hat{\xi}}$) of sampling points on \mathbb{S}_m :

$$\mathbb{M}_{\hat{\xi}} = \{\hat{\varsigma}_1, \hat{\varsigma}_2, \dots, \hat{\varsigma}_{M_{\hat{\xi}}}\}. \quad (7.44)$$

II) At each sample $\hat{\varsigma}_i \in \mathbb{M}_{\hat{\xi}}$, compute

$$f_{\xi_r}(\hat{\varsigma}_i) = \int_{\mathbb{R}^+} f(\mathbf{u}(\varsigma_r, \hat{\varsigma}_i)) p_{\xi_r}(\varsigma_r) d\varsigma_r \quad (7.45)$$

which can be performed effectively by employing the one-dimensional DMC simulation (i.e. Algorithm 7.2) using M_{ξ_r} samples from the set

$$\mathbb{M}_{\xi_r} = \{\varsigma_r^1, \varsigma_r^2, \dots, \varsigma_r^{M_{\xi_r}}\} \quad (7.46)$$

for ξ_r . That is,

$$f_{\xi_r}(\hat{\varsigma}_i) \approx \frac{1}{M_{\xi_r}} \sum_{j=1}^{M_{\xi_r}} f(\mathbf{u}(\varsigma_r^j, \hat{\varsigma}_i)). \quad (7.47)$$

III) Then, the final result can be computed as

$$E(f(\mathbf{u}(\xi))) \approx \frac{1}{M_{\hat{\xi}}} \sum_{i=1}^{M_{\hat{\xi}}} f_{\xi_r}(\hat{\varsigma}_i). \quad (7.48)$$

The above procedure is referred to as the directed Monte Carlo simulation for multiple random variables.

Note that in the above discussion no coordinate system is explicitly specified for the computation associated with \mathbb{S}_m . Such a system may be needed when generating sampling points on \mathbb{S}_m . In fact, any valid coordinate system can be chosen in principle. For instance, the so called hyper-spherical coordinate system, a generalisation of the 3D spherical coordinate system, can be used, in which the independent variables are $m-1$ "angles" $\{\theta_1, \theta_2, \dots, \theta_{m-1}\}$, and any point $\hat{\xi}$ on \mathbb{S}_m can be expressed as

$$\hat{\xi} = \begin{pmatrix} \hat{\xi}_1 \\ \hat{\xi}_2 \\ \vdots \\ \hat{\xi}_k \\ \vdots \\ \hat{\xi}_{m-1} \\ \hat{\xi}_m \end{pmatrix} = \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \cos \theta_2 \\ \dots \dots \\ \left(\prod_{i=1}^{k-1} \sin \theta_i \right) \cos \theta_k \\ \dots \dots \\ \sin \theta_1 \dots \sin \theta_{m-2} \cos \theta_{m-1} \\ \sin \theta_1 \dots \sin \theta_{m-2} \sin \theta_{m-1} \end{pmatrix} \quad (7.49)$$

where $\theta_i \in [0, \pi]$ for $i=1, 2, \dots, m-2$ and $\theta_{m-1} \in [0, 2\pi)$. The random vector $\hat{\xi}$ will be uniformly distributed on S_m when all the angles θ_i ($i=1, 2, \dots, m-1$) are uniformly distributed in their value ranges.

However, the use of a coordinate system on S_m can be totally avoided. Suppose that the required number $M_{\hat{\xi}}$ of samples on S_m is larger than the required number M_{ξ_r} of the samples for ξ_r . First generate $M_{\hat{\xi}}$ sampling points of ξ from which two sets of samples for both ξ_r and $\hat{\xi}$, M_{ξ_r} and $M_{\hat{\xi}}$, can be obtained following Eqs. (7.33-7.34). Then when computing $f_{\xi_r}(\hat{\xi}_i)$ at each sampling point $\hat{\xi}_i \in M_{\hat{\xi}}$, randomly choose any M_{ξ_r} consecutive samples from M_{ξ_r} for ξ_r .

Although $\xi_r \in \mathbb{R}^+$ in Eqs. (7.33-7.34), due to the spherical symmetry of the distribution of ξ , the random variation ξ_r can be extended to the whole real number axis \mathbb{R} by restricting the directional vector $\hat{\xi}$ on half of S_m . The benefits of extending ξ_r to \mathbb{R} are twofold: (a) to reduce the number of MC sample points required by half, thereby reducing the total MC simulation times; and (b) to increase the number of sampling points of ξ_r used at each $\hat{\xi}_i$ to compute $f_{\xi_r}(\hat{\xi}_i)$, thereby maximising the computational gain of the one-dimensional DMC simulation. Another added benefit is that the important issue of fully understanding the eigenstructure of the general matrix $K(\xi)$ becomes more tractable, although this is outside the scope of this chapter.

Finally, the main steps involved in the directed Monte Carlo simulation for

multiple random variables are outlined in Algorithm 7.3.

Algorithm 7.3 Directed Monte Carlo simulation – multiple random variables:

$$\text{DMCm}\left(\mathbf{K}_o, \{\mathbf{K}_i\}, \mathbf{b}, f, m, M_{\hat{\xi}}, M_{\xi_r}\right)$$

- Generate $M_{\hat{\xi}}$ standard Gaussian random vector samples $\mathbb{M} = \{\varsigma_1, \varsigma_2, \dots, \varsigma_{M_{\hat{\xi}}}\}$ for ξ , and create two sets of sample points $\mathbb{M}_{\xi_r} = \{\varsigma_r^1, \varsigma_r^2, \dots, \varsigma_r^{M_{\hat{\xi}}}\}$ and $\mathbb{M}_{\hat{\xi}} = \{\hat{\varsigma}_1, \hat{\varsigma}_2, \dots, \hat{\varsigma}_{M_{\hat{\xi}}}\}$ for ξ_r and $\hat{\xi}$ respectively.
- Loop over all sampling points $\hat{\varsigma}_i = (\hat{\varsigma}_1^i, \hat{\varsigma}_2^i, \dots, \hat{\varsigma}_m^i) \in \mathbb{M}_{\hat{\xi}}$:
 1. Set $\mathbf{K}_{\hat{\xi}}^i = \hat{\varsigma}_1^i \mathbf{K}_1 + \hat{\varsigma}_2^i \mathbf{K}_2 + \dots + \hat{\varsigma}_m^i \mathbf{K}_m$
 2. Choose randomly M_{ξ_r} consecutive samples from \mathbb{M}_{ξ_r} to form a subset $\mathbb{M}_{\xi_r}^s \subset \mathbb{M}_{\xi_r}$
 3. Compute: $f_{\xi_r}(\hat{\varsigma}_i) = \text{DMC1}\left(\mathbf{K}_o, \mathbf{K}_{\hat{\xi}}^i, \mathbf{b}, f, \mathbb{M}_{\xi_r}^s, M_{\xi_r}, p_{\xi_r}(\varsigma_r)\right)$
- Calculate the expectation: $E(f(\mathbf{u}(\xi))) = \frac{1}{M_{\hat{\xi}}} \sum_{i=1}^{M_{\hat{\xi}}} f_{\xi_r}(\hat{\varsigma}_i)$

The DMC method for multiple random variables, however, has implications with regard to the solution accuracy in comparison with the standard Monte Carlo simulation, since both the Monte Carlo formulation and the sampling strategy are now essentially changed. This issue will be discussed further in conjunction with numerical simulations in the next section.

7.4 Numerical Experiments

In this section, two examples will be used to provide a full assessment of several numerical aspects of the proposed DMC simulation methodology for the solution of Eq. (7.1). First the performance of PCG using different preconditioners and prediction

schemes for initial approximations are investigated in terms of both PCG iterations and CPU costs mainly for one random variable cases. Then the solution accuracy of the DMC approach against the standard MC simulation in terms of the number of MC sample points is examined.

7.4.1 Problem Descriptions

Both examples are elastic plane stress problems with a constant Poisson's ratio ($\nu(\mathbf{x}, \omega) \equiv 0.3$) and stochastic Young's modulus. Furthermore, it is assumed that both Young's modulus have unit mean value, i.e. $E(E(\mathbf{x}, \omega)) = 1.0$, while the following two covariance functions R_1 and R_2 are given respectively for the two problems:

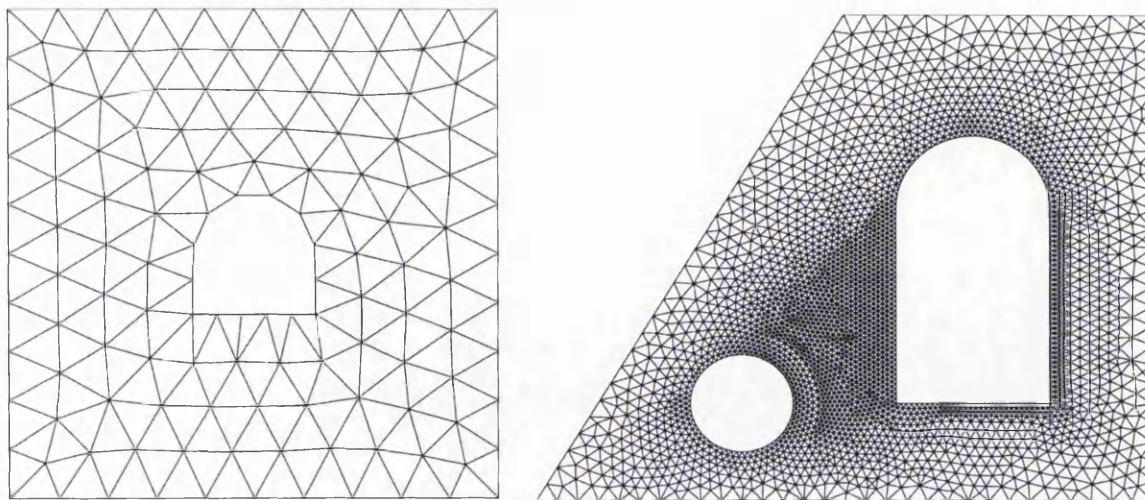
$$R_1(\mathbf{x}_1, \mathbf{x}_2) = \text{Cov}(E(\mathbf{x}_1, \omega), E(\mathbf{x}_2, \omega)) = 0.0729 \times \sigma_E^2 e^{-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2}{2.0^2}} \quad (7.50)$$

$$R_2(\mathbf{x}_1, \mathbf{x}_2) = \text{Cov}(E(\mathbf{x}_1, \omega), E(\mathbf{x}_2, \omega)) = 0.0729 \times \sigma_E^2 e^{-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2}{0.3^2}} \quad (7.51)$$

in which σ_E is a free parameter that can be used to vary the scale of the random variation. Following the SFEM formulation described in Chapter 5, a stochastic system of linear algebraic equations is obtained in the form of (7.1), in which the integer m is arbitrarily chosen in these examples without considering the accuracy of the F-K-L representation. In order to avoid ill conditioned systems occurring in Monte Carlo simulations, the sample value range of each normalised Gaussian random variable is restricted to $[-3, 3]$.

The finite element meshes of the two problems, with a total number DOFs of 274 and 5965 respectively, are shown in Figure 7.2. The maximum dimensions of the two problems in metres are 16×16 and 6×5 respectively.

The first example has a very small number of DOFs. It is used mainly for investigating the convergence features of the numerical techniques developed earlier. The small scale also makes it possible to examine the solution accuracy of the new method compared with the standard procedure within a reasonable time scale. For this example, the first nine terms in the F-K-L representation are considered, resulting in nine Gaussian



(a) Example 1 (274 DOFs)

(b) Example 2 (5986 DOFs)

Figure 7.2 Finite element discretizations of two examples

Table 7.1 Eigenvalues of the stochastic field and associated stiffness matrices: Example 1

Index	Eigenvalue of stochastic field (λ^E)	Min eigenvalue of stiffness matrices (λ_1)	Max eigenvalue of stiffness matrices (λ_N)
1	0.08889	0.0493	0.1101
2	0.08483	-0.1108	0.1113
3	0.08483	-0.1107	0.1110
4	0.07884	-0.1088	0.1088
5	0.07884	-0.0988	0.1008
6	0.07048	-0.1001	0.1008
7	0.07048	-0.0991	0.0995
8	0.06206	-0.0942	0.0942
9	0.06206	-0.0884	0.0855

random variables ξ_i and associated matrices \mathbf{K}_i . The eigenvalue λ_i^E corresponding to each random variable ξ_i and the two extreme eigenvalues of the matrix \mathbf{K}_i are listed in Table 7.1 at the variance scale of $\sigma_E = 1.0$. It is observed that for the given covariance function, the eigenvalues of all the nine random variables have a similar value and that, except for \mathbf{K}_2 which is symmetric positive definite, the remaining matrices are all indefinite, but their minimum and maximum eigenvalues are almost the same in magnitude.

Table 7.2 Eigenvalues of the stochastic field and associated stiffness matrices: Example 2

Index	Eigenvalue of stochastic field (λ^E)	Min eigenvalue of stiffness matrices (λ_1)	Max eigenvalue of stiffness matrices (λ_N)
1	0.02211	0.0022371	0.0105175
2	0.02091	-0.0100223	0.0099748
3	0.02091	-0.0085306	0.0098457
4	0.02062	-0.0099338	0.0107368
5	0.02062	-0.0098814	0.0107376

The second example has a more complex structural configuration with in total 5640 elements and 5986 DOFs, much larger than the first example. The main purpose of this example is to further confirm the results obtained in the first example and particularly to investigate the efficiency of the proposed methodology in terms of CPU time costs. For this example, the first five stochastic variables and the corresponding matrices are involved in the computation. Their eigenvalues at the variance scale of $\sigma_E = 1.0$ are listed in Table 7.2, and similar features to the first example can be observed.

In both examples, a randomly generated external load vector \mathbf{b} is considered.

7.4.2 Performance of the Directed Monte Carlo Method and the Associated Numerical Techniques

The aim of the first investigation is to use the two examples to establish the convergence properties of PCG, in the context of the DMC method for one random variable, when employing the different associated numerical techniques for solving the linear system of equations of (7.42). The equations are created as follows: A number of random variables and the corresponding matrices are arbitrarily selected from all the matrices available in each example, and the random vector $\hat{\xi}$ is given randomly to construct $\mathbf{K}_{\hat{\xi}}$. A total of $M_{\hat{\xi}}$ sample values for ξ_r are generated following the procedure outlined in the previous section.

7.4.2.1 Convergence of Preconditioning and Initial Approximation Schemes

The first batch of tests uses Example 1 to assess the performance of PCG with

different preconditioning and initial approximation prediction schemes in terms of PCG iterations.

The average numbers of iterations required to achieve the prescribed convergence of PCG per MC sample are recorded in Tables 7.3 and 7.4 for the following different combinations of the integrated numerical techniques and parameters in the DMC method: 1) the number of random matrices included, m ; 2) the variance scale σ_E ; 3) the three preconditioning schemes; 4) the two interpolation/extrapolation schemes for the prediction of the initial approximation; 5) the given convergence tolerance τ ; and 6) different numbers of Monte Carlo samples $M_{\xi_r} = 10, 24, 50, \text{ and } 100$ used for ξ_r . Note that the same number of negative and positive samples is used and these two sets of samples are separately processed starting from the smallest absolute value of the samples as outlined in Section 7.2.2.

Tables 7.3 and 7.4 list the results for $m=3$ and $m=9$ respectively with six cases included in total. For each case, the performance of the standard MC simulation with PCG as the solver using the three different preconditioning schemes is also included for comparison as a benchmark.

In Table 7.3 ($m=3$), three matrices K_3 , K_6 and K_9 are used. Considering Case 1 ($\sigma_E = 1.0$, $\tau = 1.0 \times 10^{-5}$) as a basic case, Cases 2 and 4 increase σ_E to 2.5 to represent large scale random variation conditions, and Cases 3 and 4 increase the solution accuracy to $\tau = 1.0 \times 10^{-6}$. For both Cases 1 and 3 ($\sigma_E = 1.0$), the two minimum and maximum eigenvalues of K_i and K_o are respectively -0.1104 and 0.1139, while for Cases 2 and 4 ($\sigma_E = 2.5$), the two eigenvalues proportionally increase in magnitude to -0.2760 and 0.2847.

In Table 7.4, all the nine matrices are used, but only $\sigma_E = 1.0$ is considered since a further increase of σ_E will be very likely to result in an indefinite K_{ξ} . Two levels of solution accuracy are considered in Cases 5 and 6 respectively.

Table 7.3 Average PCG iterations per MC sample - $m = 3$: $\{K_3, K_6, K_9\}$

Case 1		$\sigma_E = 1.0, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.1104, 0.1139]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 10$	16.0	13.7	15.4	12.7	3.00	2.40
$M_{\xi_r} = 24$	13.6	8.54	13.5	8.75	1.66	1.33
$M_{\xi_r} = 50$	8.44	9.50	8.00	9.58	1.34	1.20
$M_{\xi_r} = 100$	7.22	11.5	7.02	10.9	1.06	1.32
Standard MC	26.5		26.3		4.75	
Case 2		$\sigma_E = 2.5, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.2760, 0.2847]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 10$	24.9	22.5	22.6	19.3	6.80	5.70
$M_{\xi_r} = 24$	23.0	18.8	20.9	16.4	5.21	7.64
$M_{\xi_r} = 50$	18.5	14.4	16.0	12.7	3.80	2.44
$M_{\xi_r} = 100$	16.0	14.3	13.8	12.2	3.13	2.50
Standard MC	30.1		27.2		8.63	
Case 3		$\sigma_E = 1.0, \tau = 1.0 \times 10^{-6}, \lambda \in [-0.1104, 0.1139]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 24$	19.7	12.4	19.1	12.4	3.20	1.67
$M_{\xi_r} = 50$	16.3	10.1	15.3	9.76	2.24	1.30
$M_{\xi_r} = 100$	14.1	11.1	13.7	11.3	1.90	1.43
Standard MC	28.7		28.6		5.68	
Case 4		$\sigma_E = 2.5, \tau = 1.0 \times 10^{-6}, \lambda \in [-0.2760, 0.2847]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 24$	27.3	21.5	24.3	19.0	7.13	4.91
$M_{\xi_r} = 50$	23.7	16.7	20.6	14.8	5.34	4.18
$M_{\xi_r} = 100$	22.4	15.0	20.0	13.1	4.62	2.79
Standard MC	32.5		29.3		10.2	

Table 7.4 Average PCG iterations per MC sample - $m = 9: \{K_1, K_2, \dots, K_9\}$

Case 5		$\sigma_E = 1.0, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.1104, 0.1139]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\epsilon_r} = 10$	20.1	17.9	18.6	16.4	4.60	3.90
$M_{\epsilon_r} = 24$	16.2	12.0	14.3	11.3	3.25	2.46
$M_{\epsilon_r} = 50$	14.7	11.1	12.9	9.98	2.81	2.14
$M_{\epsilon_r} = 100$	11.5	12.0	10.2	11.2	2.31	1.90
Standard MC	28.6		26.8		7.66	
Case 2		$\sigma_E = 2.5, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.1104, 0.1139]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\epsilon_r} = 10$	25.0	21.9	23.2	20.4	6.00	5.00
$M_{\epsilon_r} = 24$	22.5	15.3	21.2	15.2	4.50	2.96
$M_{\epsilon_r} = 50$	21.4	13.8	19.8	12.7	4.28	2.50
$M_{\epsilon_r} = 100$	19.1	12.9	17.4	12.5	3.69	2.14
Standard MC	30.9		28.9		9.17	

By examining the results presented in Tables 7.3 and 7.4, the following conclusions can be made:

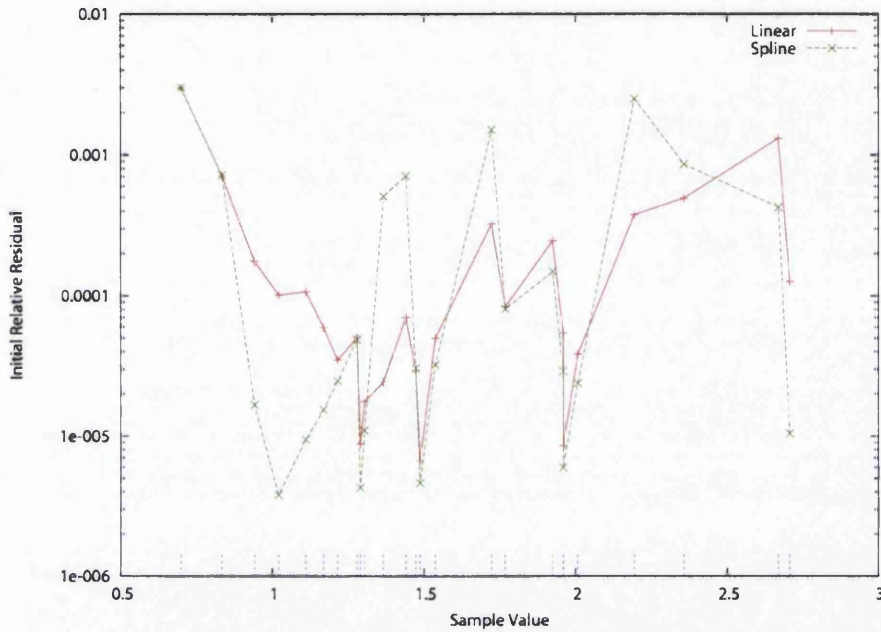
- I) Preconditioning schemes: As expected, the incomplete Cholesky decomposition of K_o , IC0, is the least effective scheme. Although the “stochastic” preconditioning scheme, SIC, can normally enhance the performance but the improvement is marginal and in general no more than 15% of reduction in iterations is achieved in all the test cases. On the contrary, the K0 scheme using the complete Cholesky decomposition of K_o exhibits a very good performance, with a reduction of iterations ranging from around 3 to 10 times compared to both IC0 and SIC schemes, and a typical reduction factor is around 5 or 6. A better performance of the K0 scheme is achieved at relatively smaller variation scales (Cases 1 and 3) and degrades slightly at larger variations (Cases 2 and 4), while the increase of the level of solution accuracy requirement τ makes the comparison more favourable to the K0 scheme. These results very positively support the earlier analytical analysis on

the eigenvalue properties of the stochastic matrix $K(\xi)$ made in Section 7.1. It is highlighted that the above conclusion applies not only to one random variable but also holds for general cases. It is also noticed that the K0 preconditioning scheme can also be further enhanced by, for instance, the use of the Neumann expansion of Eq. (7.22) as proposed in [7.5].

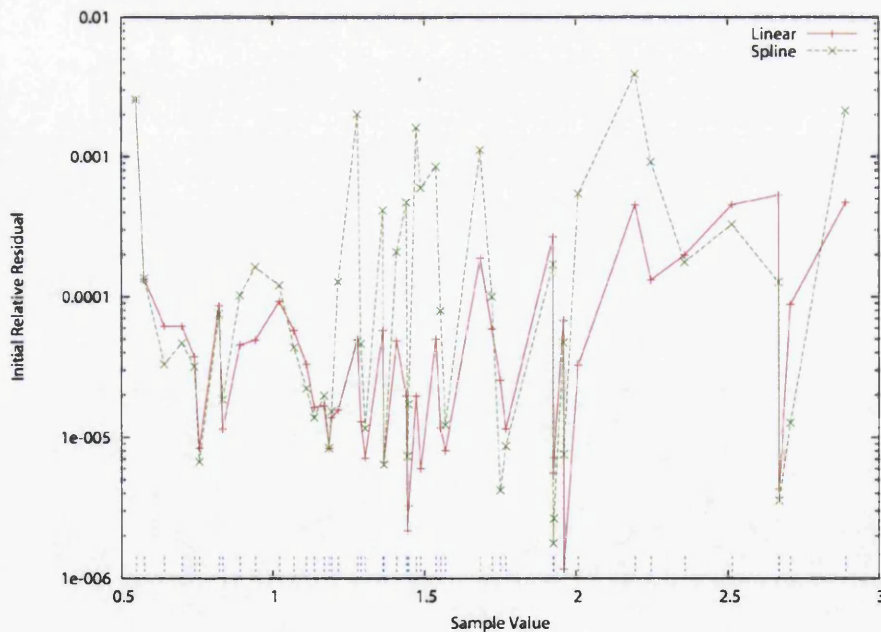
II) Initial approximation schemes. The ability to provide a good initial approximation for PCG solvers is the key to the success of the proposed DMC method and this is confirmed by all the test cases. In fact, compared with the standard MC method, the current PCG with an initial approximation achieves a reduction of iterations by 2 to 4 times, depending on the number of Monte Carlo samples used. Generally speaking, the required PCG iterations decrease with increase of the number of MC samples. It is possible that for a sufficiently large M_{ξ} the average iterations may be reduced to less than 1, thereby achieving an even greater improvement over the standard MC method. For the two interpolation/extrapolation approaches proposed, the spline scheme outperforms, in most cases, the linear scheme by an amount ranging from 10% up to 40% in terms of iterations, but the scheme becomes less effective in several cases. This phenomenon is purely dependent on the distribution of MC sampling points. When the samples are fairly evenly distributed, which is often the case when the number of samples is small, the spline extrapolation will be superior to the linear extrapolation. However, when some clusters occur in the samples, which is an inherent feature with a large number of random numbers, the accuracy of the spline extrapolation at points immediately after the clusters will suffer, leading to an increase of PCG iterations. Figure 7.3 illustrates this phenomenon by showing the accuracy of the initial approximations provided by the two extrapolation schemes for two different numbers of samples $M_{\xi} = 24$ and $M_{\xi} = 50$ in Case 1. Note that only the positive samples are used and the positions of these samples are marked along the axis- x , from which the irregular spacing pattern of the samples is clearly demonstrated. The defect of the spline scheme may be eliminated, however, by enhancing the standard spline interpolation algorithm for the current situation.

In summary, the proposed DMC method can indeed significantly enhance the

computational effectiveness of the Monte Carlo simulation for general one random variable problems. When equipped with the K0 preconditioning and linear/spline extrapolation schemes, it can typically achieve performance around 3 times faster than the standard Monte Carlo simulation when the number of samples involved is not too small.



(a) $M_{\xi_r} = 24$



(b) $M_{\xi_r} = 50$

Figure 7.3 Accuracy of initial approximations provided by linear and spline schemes in Case 1

7.4.2.2 CPU Time Cost Comparisons

The above observations are further examined by the second batch of tests using the second example. All the five matrices associated with the five variables are considered, i.e. $m = 5$. Table 7.5 lists the average number of PCG iterations per MC sample with the same test conditions as in the previous batch of tests except for the scales of variation σ_E which are 10 times larger.

Table 7.5 Average PCG iterations per MC sample – Example 2 ($m = 5$)

Case 7		$\sigma_E = 10.0, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.0830, 0.1300]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 10$	41.9	36.4	39.6	34.9	2.50	2.20
$M_{\xi_r} = 24$	33.5	29.8	31.5	28.7	1.95	1.18
$M_{\xi_r} = 50$	16.7	26.2	15.2	23.8	1.20	1.12
$M_{\xi_r} = 100$	9.89	34.6	9.84	32.8	0.77	1.30
Standard MC	77.74		74.90		4.14	
Case 8		$\sigma_E = 25.0, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.2175, 0.3250]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 10$	68.8	63.0	58.2	53.7	6.20	5.30
$M_{\xi_r} = 24$	60.1	46.4	50.3	38.3	4.67	3.25
$M_{\xi_r} = 50$	41.6	37.9	19.8	33.1	3.30	2.10
$M_{\xi_r} = 100$	27.6	38.4	22.0	36.5	2.46	2.39
Standard MC	91.0		80.0		7.57	

In essence, the new tests confirm all the observations made in the previous tests for the three preconditioning schemes and the two extrapolation schemes. As a matter of fact, the K0 preconditioning performs even better as it converges over 10 times faster than the other two counterparts at both normal and larger variance scales (Cases 7 and 8). Also the average number of iterations of the K0 preconditioning is reduced to less than 1 at $M_{\xi_r} = 100$ in Case 7. These reveal that the features of the proposed methodology are universal regardless of the finite element scale of problems to be modelled.

Although it is clear that the K0 preconditioning is superior to the other two preconditioning schemes in terms of PCG iterations, it is more costly at each iteration as the size of its Cholesky decomposition matrix L_o is larger than the other two IC decompositions, ignoring the extra cost associated with the re-generation of a new incomplete decomposition for SIC at each sample. Therefore it is necessary to examine the CPU time costs of all the numerical techniques at the same time. The relevant results are collected in Table 7.6.

Table 7.6 Total CPU times (seconds) – Example 2 ($m = 5$)

Case 7		$\sigma_E = 10.0, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.0830, 0.1300]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 10$	9.64	8.56	16.9	15.1	1.29	1.26
$M_{\xi_r} = 24$	18.6	17.0	32.7	35.1	2.37	1.93
$M_{\xi_r} = 50$	19.9	32.3	40.1	64.7	3.75	5.60
$M_{\xi_r} = 100$	23.9	91.1	66.2	169.0	5.54	14.1
Standard MC	178.1		261.3		23.2	
Case 8		$\sigma_E = 25.0, \tau = 1.0 \times 10^{-5}, \lambda \in [-0.2175, 0.3250]$				
Preconditioning	IC0		SIC		K0	
Interpolation	Linear	Spline	Linear	Spline	Linear	Spline
$M_{\xi_r} = 10$	15.73	14.64	24.13	23.26	3.17	3.10
$M_{\xi_r} = 24$	33.43	26.51	52.18	47.17	6.03	5.72
$M_{\xi_r} = 50$	48.95	45.56	86.64	87.81	10.26	10.65
$M_{\xi_r} = 100$	66.78	101.2	147.7	189.2	17.47	26.21
Standard MC	210.0		283.2		34.73	

It is evident that although the overall speed-up of the K0 preconditioning in terms of CPU costs over the other two schemes is smaller than the speed-up in iterations, it is still about 4 to 5 times faster, confirming that K0 is indeed a very effective and robust preconditioner, at least for not very large scale matrices. As expected, a small reduction in iterations of the SIC scheme over the IC0 scheme cannot compensate for the extra costs incurred for the generation of the IC decomposition at each MC sample, making it an unattractive option unless further developments are undertaken, such as the work in [7.8].

It is noticeable that the spline extrapolation scheme becomes slower than the linear scheme in all cases. This is due to the fact that in the standard B-spline interpolation algorithm currently implemented, all the previous solutions are used to predict the initial solution for the current MC sample. It is possible however to modify the B-spline formulation so that only a few solutions are required for the interpolation and extrapolation, thereby making its computational cost comparable to that of the linear interpolation/extrapolation.

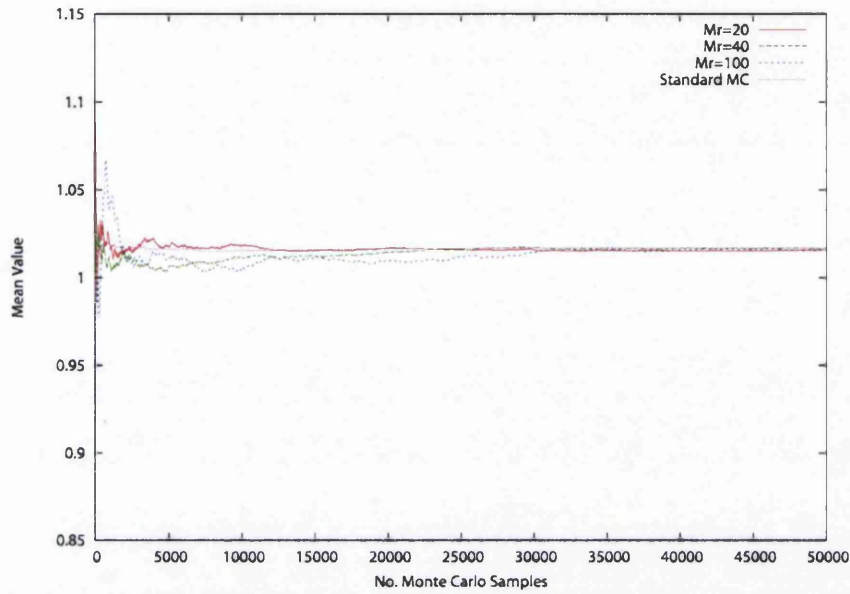
7.4.3 Solution Accuracy of the Directed Monte Carlo Method

The proposed DMC simulation method for multiple Gaussian random variables utilizes the hyper-spherical transformation to convert the original Monte Carlo simulation into two parts: a uniform distribution simulation on \mathbb{S}_m and a one-dimensional χ^2 -type distribution along the ξ_r direction, representing the global variation scale. The previous tests have established that the DMC method provides a very effective procedure to deal with the solution along the ξ_r direction, i.e. the probability distribution of the solution at each point on \mathbb{S}_m can be obtained efficiently. However, as the original goal is to solve Eq. (7.1) for any number of random variables, it is necessary to examine the solution accuracy of the DMC method for general cases in comparison with the standard MC simulation where the sampling is performed in real space \mathbb{R}^m .

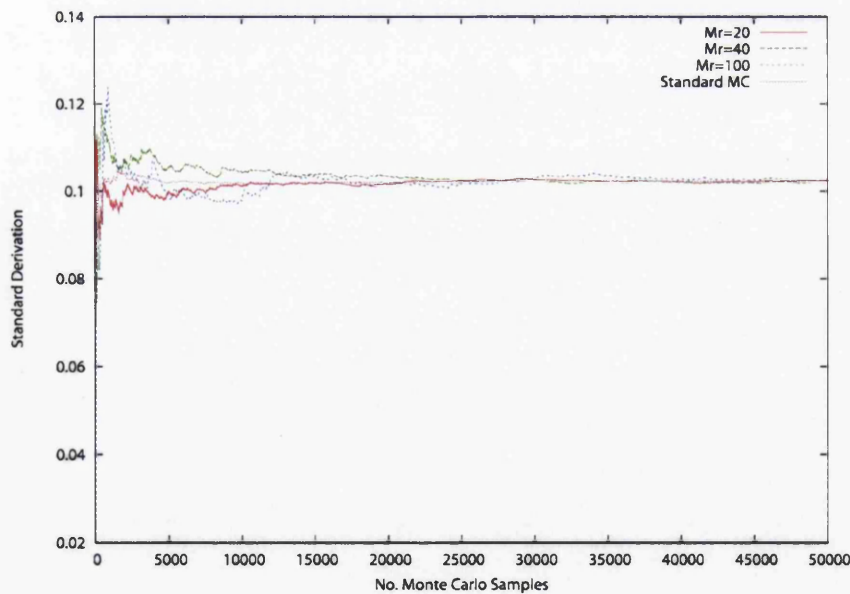
The solution accuracy of the DMC method in terms of the total number of Monte Carlo points is assessed using Example 1. Following the procedure outlined in Section 7.3, for a given m , different numbers of Monte Carlo samples, $M_{\hat{\xi}}$, are generated on \mathbb{S}_m . At each point $\hat{\xi}_i$, different numbers of sample points M_{ξ_r} are used for ξ_r to determine $f_{\xi_r}(\hat{\xi}_i)$ (refer to (7.47)). In the simulation, the expectation and standard deviation of the total strain energy of the structure $\mathbf{b}^T \mathbf{u}(\xi)$, normalized by the deterministic energy $\mathbf{b}^T \mathbf{u}_d$, are considered.

The computed expectation (mean value) and standard deviation of the normalized strain energy via the total number of Monte Carlo points $M = M_{\hat{\xi}} \times M_{\xi_r}$ up to 50,000 are

depicted in Figure 7.4 for the cases of $m = 9$ with $M_{\xi_r} = 20, 40$ and 100 respectively. Note that the corresponding numbers for M_{ξ} are 2500, 1250, and 500. The results computed by the standard Monte Carlo simulation, which corresponds to the special case of $M_{\xi_r} = 1$, are also shown in the figure for comparison.



(a) The mean value



(b) The standard deviation

Figure 7.4 Convergence histories of the normalized total strain energy versus the number of Monte Carlo samples: Example 1

The figure illustrates that the best convergence is achieved by the standard MC simulation, while the convergence of the DMC solution appears to be slightly less smooth and the oscillation at the small sample numbers increases with the increase of M_{ξ_r} . This is not surprising as for a fixed M , an increase of M_{ξ_r} is equivalent to the concentration of Monte Carlo samples along fewer directions, therefore reducing the randomness of these samples. This suggests that the current hyper-spherical transformation may lose, to a limited degree, the unique advantage of the standard MC simulation when the number of MC samples used is not sufficiently large. It can be argued that the improved computational efficiency of DMC at each sample on \mathbb{S}_m will permit the use of more sample points thereby compensating for the slight loss of solution accuracy.

However, it is more important to highlight that the hyper-spherical transformation, together with the ability of effectively (and accurately) determining $f_{\xi_r}(\hat{\xi}_i)$ at each point on \mathbb{S}_m , may provide a prospect of developing new simulation approaches within the general Monte Carlo solution framework that have potential to greatly improve the solution accuracy of the current DMC method. This is the line of research that is currently being pursued in order to further enhance the overall computational capability of the DMC method for general problems with multiple random variables.

7.5 Summary

This chapter proposes a modified Monte Carlo simulation procedure, DMC, for solving a stochastic system of linear algebraic equations. The basic idea of DMC for one random variable cases is to order the Monte Carlo samples so that when the samples are processed in sequence the previous obtained solutions can be utilised to provide a high quality initial approximation for the current point thereby significantly accelerating the convergence of iterative solvers. In the DMC method, PCG plays a central role and the two essential numerical techniques crucial to the success of the method include preconditioning and initial approximation predictions. It is proposed on the basis of analytical analysis, and later confirmed numerically, that the deterministic matrix K_0 can serve as a very effective preconditioning matrix. The numerical experiments conducted

demonstrate that the proposed DMC can indeed significantly enhance the computational effectiveness of the Monte Carlo simulation for general one random variable problems. When employing the K0 preconditioning and linear/spline extrapolation schemes, DMC can typically perform around 3 times faster than the standard Monte Carlo simulation when the number of samples used is not too small.

The extension of the DMC method to multiple random variable cases is realised by the adoption of a hyper-spherical transformation whereby any m -dimensional random vector in \mathbb{R}^m can be expressed by a random variable ξ_r , representing the global random variation scale, and a unit directional random vector on the unit “ m -sphere” S_m . Such a transformation permits the Monte Carlo calculation of the solution to be undertaken as a one random variable case along the ξ_r direction at each sample on S_m . Although the overall computational costs of the Monte Carlo simulation can be reduced in this way, it is at the expense of slightly losing solution accuracy when the total number of samples used is not sufficiently large. This observation indicates the aspect to be further pursued in order to improve the overall performance of the proposed DMC method for general multiple random variable problems.

The major results in this chapter have been reported in [7.11].

References

- [7.1] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements – A Spectral Approach*, Revised edition, Dover Publications, New York, 2003.
- [7.2] G.P. Lepage, A new algorithm for adaptive multidimensional integration, *Journal of Computational Physics*, 27 (1978) 192-203.
- [7.3] W.H. Press, S.A. Teukolsky and W.T. Vetterling et al., *Numerical Recipes in C: the Art of Scientific Computing*, 2nd Edition, Cambridge University Press, Cambridge, 1995.
- [7.4] R.G. Ghanem and R.M. Kruger, Numerical solution of spectral stochastic finite element systems, *Computer Methods in Applied Mechanics and Engineering*, 129 (3) (1996) 289-303.
- [7.5] M. Papadrakakis and V. Papadopoulos, Robust and efficient methods for

stochastic finite element analysis using Monte Carlo simulation, *Computer Methods in Applied Mechanics and Engineering*, 134 (3-4) (1996) 325-340.

- [7.6] M. Papadrakakis and A. Kotsopoulos, Parallel solution methods for stochastic finite element analysis using Monte Carlo simulation, *Computer Methods in Applied Mechanics and Engineering*, 168 (3-4) (1998) 305-320.
- [7.7] R.G. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, *Computer Methods in Applied Mechanics and Engineering*, 168 (1-4) (1999) 19-34.
- [7.8] C. Desceliers, R.G. Ghanem and C. Soize, Polynomial chaos representation of a stochastic preconditioner, *International Journal for Numerical Methods in Engineering*, 64 (5) (2005) 618-634.
- [7.9] Y.T. Feng, On the discrete dynamic nature of the conjugate gradient method, *Journal of Computational Physics*, 211 (1) (2006) 91-98.
- [7.10] M. Benzi, Preconditioning techniques for large linear systems: a survey, *Journal of Computational Physics*, 182 (2) (2002) 418-477.
- [7.11] Y.T. Feng, C.F. Li and D.R.J. Owen, A Directed Monte Carlo Solution of Linear Stochastic Algebraic System of Equations, *International Journal for Numerical Methods in Engineering*, Submitted.

Chapter 8

Concluding Remarks

The aim of this thesis as stated in Section 1.3.1 has been achieved. Namely, a numerical framework for elastostatics of random media has been formulated. It was pursued through four distinct and consecutive phases as follows

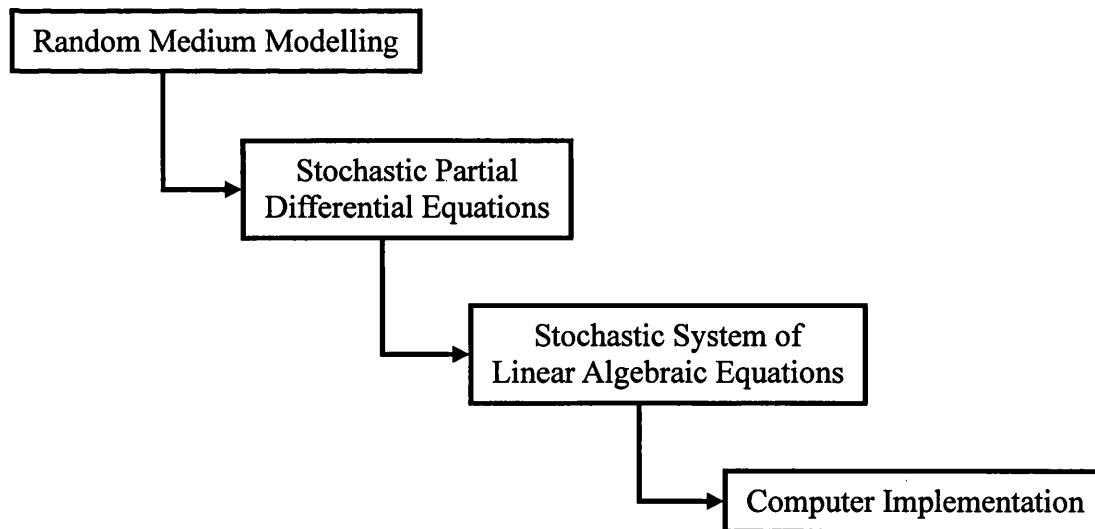


Figure 8.1 Research components of this work

As reviewed in Chapter 1, a number of important results on the topic of SFEM (especially the pioneering works of Shinozuka, Liu and Ghanem) have been reported in the past two decades; however, compared with the well established FEM, the SFEM is still in its infancy and many fundamental problems are still outstanding. Hence, in order to maintain a maximum flexibility during this work, no existing SFEM technique has been

taken for granted without independent investigation. Although the thesis is presented in a way emphasising its engineering background, equal attention has been paid both to practical engineering requirements and to mathematical rigour throughout this work.

In the following, a more detailed list of the achievements and conclusions of this work is provided. The thesis closes with suggestions for future research.

8.1 Achievements and Conclusions

8.1.1 Random Medium Modelling

In order to utilize existing mathematical results in SODE and SPDE, the possibility of developing a generalized white noise model for practical random media is first checked against common sense engineering requirements. The rejection of the white noise approach indicates that the non-singular spatial stochastic dependence of material properties is a fundamental characteristic of practical random media. Based on the recognition of the probabilistic essence of random media and driven by practical engineering requirements, the ERM model is consequently defined and its macro-scale properties including stationarity, continuity/differentiability and principles for material measurements are systematically explored.

It should be noted that wide-sense stationary stochastic fields and their first two statistical moments have long been used in various SFEM formulations to describe random material properties without addressing their suitability. Hence, the basis of the ERM model is not new. However, the contribution here is threefold:

- I) The investigation of the white noise approach makes the ERM model a natural choice in random medium modelling, which supports the past intuitive use of wide-sense stationary stochastic fields in SFEMs.
- II) The specific and detailed mathematical definition of ERM provides a solid foundation to interpret the governing SPDE system for elastostatics of random media.
- III) The exploration of macro-scale properties of ERM not only improves the general understanding of random media but also reveals some defects in existing SFEM

formulations, specifically the misuse of some non-differentiable second-order stochastic fields in random medium modelling.

In the ERM model, random material properties are implicitly defined by their first- and second- order statistical moments, while an explicit representation of the associated stochastic fields are required in the governing SPDE system for elastostatics of random media. Hence, based on the spectral representation theory of wide-sense stationary stochastic fields and the standard dimensionality reduction technology of principal component analysis, the F-K-L representation scheme is developed for the general elastic tensor of ERM.

Compared with the widely used K-L expansion based on the standard finite element method, the F-K-L representation scheme has the following advantages:

- I) The K-L expansion method is essentially limited to the representation of a single random material parameter, while the F-K-L representation scheme is designed for the general random elastic tensor and therefore can deal with any number of correlated random material properties.
- II) The K-L expansion method is based on FE meshes, and therefore its approximation accuracy can not be explicitly controlled and depends on the mesh density used. However, the F-K-L representation scheme is completely mesh free and independent of the specific shape of the random structure under consideration. In particular, it is achieved with an explicit and a priori error control.
- III) The eigenvalue decay rate in the F-K-L representation can be approximately predicted without solving any equation, which provides a useful function for evaluating the randomness scale of practical engineering systems with variable uncertainties.

In addition, during the development of the F-K-L representation scheme, an accurate and efficient quadrature algorithm for multidimensional oscillatory functions is obtained, which reduces the associated computational cost by up to several orders of magnitude. It can be expected that this novel numerical integration algorithm might also provide a key to resolving many other physical problems where integrations of similar multidimensional oscillatory functions are present.

8.1.2 Stochastic Partial Differential Equations

Mechanical behaviours of random media conform to the same principles of mechanics as homogeneous materials. Hence, elastostatics of random media is formally described by the same equations as in the deterministic case. However, the original PDE system becomes a SPDE system due to the presence of stochastic fields, and all the associated differential/integral operators need to be redefined in the context of probability. In this thesis, following the mathematical definition of ERM, mean square convergence is taken as the basic mode of convergence to interpret the governing SPDE system for elastostatics of random media (or more specifically ERM). From a mathematical viewpoint, it is of significant importance to investigate the existence, uniqueness and analytical properties of the solution for the resulting SPDE system. However, this important aspect is out side the scope of this thesis whose focus of solution is on its engineering aspect, i.e. developing effective numerical methods to approximately solve the corresponding SPDE system.

It should be noted that the mathematical investigation of the SPDE system for elastostatics of random media could be very challenging and its study might require fundamentally new developments of mathematical theories and methods. Indeed, even for the conventional PDE systems of mechanics, many basic questions regarding the existence, uniqueness and analytical properties of their solutions are still open. A well known example for this is the existence and smoothness of the solution for Navier-Stokes equations, which is listed in the CMI Millennium Problems [8.1].

8.1.3 The Stochastic System of Linear Algebraic Equations

Based on the F-K-L representation for the general elastic tensor of ERM and following a similar procedure as in the standard finite element method, the SPDE system for elastostatics of random media is discretized to obtain a stochastic system of linear algebraic equations. Two numerical techniques including the joint diagonalization method and the DMC (directed Monte Carlo) method are then proposed to solve the resulting stochastic linear algebraic system.

The joint diagonalization method provides a novel solution strategy for the stochastic system of linear algebraic equations. It simultaneously diagonalizes all the

matrices in the equation system via a sequence of similarity transformations, following which an explicit solution is obtained by inverting the sum of the diagonalized stochastic matrices. Unless all the matrices share exactly the same eigenstructure, the joint diagonalization can only be approximately achieved. Hence, the solution is approximate and corresponds to a particular average eigenstructure of the matrix family. Specifically, the classical Jacobi algorithm for the computation of eigenvalues of a single matrix is modified to accommodate multiple matrices and the resulting Jacobi-like joint diagonalization algorithm preserves the fundamental properties of the original version including its convergence and an explicit solution for the optimal Givens rotation angle.

Although the use of the proposed Jacobi-like joint diagonalization solution is limited in stochastic linear algebraic systems with small-scale matrices, the joint diagonalization method, as a general solution strategy for the stochastic system of linear algebraic equations, indicates a promising direction to develop efficient solvers for large-scale practical problems consisting of random media. In addition, as a decoupling technique, the major principle in the joint diagonalization solution strategy not only holds for static/stationary random medium problems in mechanics, but is also applicable to dynamic/transient problems involving random media.

In order to efficiently solve the stochastic system of linear algebraic equations containing large-scale matrices, a modified Monte Carlo method, namely DMC, is developed. The DMC method solves the stochastic linear algebraic system at each Monte Carlo sample with the PCG (preconditioned conjugate gradient) method. By utilizing the spatial proximity of Monte Carlo samples, high quality initial approximations are provided for the PCG solver to significantly reduce the total cost of Monte Carlo simulations. Furthermore, to achieve the best solution efficiency, different preconditioning matrices are compared in the DMC framework.

8.1.4 Computer Implementation

In the course of this work, an independent multipurpose simulation system, termed OMEGA, has been implemented on Windows XP System with Microsoft Visual C++. Starting from scratch provided the maximum flexibility in algorithm development and also allowed a high degree of modularity to be achieved. The OMEGA system is designed to

solve both stochastic and deterministic problems in the same framework. Although not stated in this thesis, flexible 3D visualization and manipulation facility, which has been partially demonstrated in [8.2-8.3], is also provided with the OMEGA platform.

8.2 Suggestions for Future Research

It may be said that the material presented in this thesis constitutes a meaningful tool for accelerating the evolution of the field of stochastic finite elements. However, compared with the well developed finite element method, the methodology presented in this thesis is incomplete, and further development is needed to advance the proposed numerical framework to a level where both reasonable mathematical rigour and sufficient computational efficiency are achieved so that it can be applied to solve large scale practical problems involving random media. Hence, the following suggestions are made for future research.

8.2.1 Material Modelling of Random Media

- The first key issue in random medium modelling is to extend the ERM model and the F-K-L representation from elasticity to plasticity such that creep, fatigue and their resulting fracture mechanisms in random media can be appropriately modelled. This is of crucial importance for a more accurate understanding and analysis of many random medium systems such as composites with metal matrix and random inclusions, large-scale metal structures with randomly distributed microcracks, and Micro-Electro-Mechanical Systems (MEMS) where micro-scale heterogeneous material properties need to be taken into account.
- In the F-K-L representation scheme of the ERM elastic tensor, Fourier spectral analysis is employed to examine the global “energy-frequency” distribution of the stochastic field corresponding to each individual random material parameter. Because of its prowess and simplicity, Fourier analysis has dominated data analysis efforts since soon after its introduction. Although the Fourier transform is valid under extremely general conditions, there are some crucial restrictions of the Fourier spectral analysis: the system must be linear; and the data must be strictly periodic or stationary; otherwise, the resulting spectrum will make little physical

sense because spurious harmonic components are inevitably induced by the nonlinearity and nonstationarity. Hence, when significant nonlinearity and nonstationarity are present in the constitutive relation of plastic random media, the performance of the F-K-L representation will diminish. To overcome this potential problem, methods for processing non-stationary data, including the spectrogram, wavelet analysis, the Wigner-Ville distribution, the evolutionary spectrum, the empirical orthogonal function expansion and the empirical mode decomposition, are suggested for further investigation.

8.2.2 Stochastic Partial Differential Equations

The principle discussed in Section 8.1.2 also holds for general static/steady-state and dynamic/transient problems regarding random media. That, by redefining the associated differential/integral operators in the context of probability, the conventional governing PDE systems in continuum mechanics can be transformed into SPDE systems to describe the behaviour of practical random medium systems. The resulting SPDEs should not only be able to describe the irregular spatial variation of material properties of random media, but also be able to accommodate discontinuities commonly existing in practical random media. Hence, it is suggested to take the mean square convergence, instead of the widely used almost sure convergence, as the basic mode of convergence to define continuity, differentiability and integrability in the corresponding SPDE system. However, for the system of SPDEs to be studied here the main point is the ability to solve the system rather than the existence or properties of a solution.

8.2.3 Numerical Solutions

For large-scale practical engineering systems composed of random media, computational methods are the only realistic choice to solve their governing SPDEs. The F-K-L representation (including its improvements suggested in Section 8.2.1) decomposes material properties of random media into a deterministic part and a random part, which in turn isolates variable uncertainties from the associated SPDE system. This decoupling technique provides a simple and powerful approach to pursue the solution of SPDEs arising from random medium mechanics, in that, standard finite element techniques can be applied to discretize the decoupled SPDE system. Specifically, for static/steady-state

problems, the final stochastic algebraic equation system has the form

$$\left(\sum_{i=1}^k \alpha_i A_i \right) \mathbf{x} = \mathbf{b}, \quad (8.1)$$

where α_i are random variables, A_i deterministic matrices, \mathbf{b} a deterministic/random vector, and \mathbf{x} the unknown vector to be sought; for dynamic/transient problems, the final stochastic algebraic equation system has the form

$$\left(\sum_{i=1}^{k_1} \alpha_i A_i \right) \ddot{\mathbf{x}} + \left(\sum_{i=1}^{k_2} \beta_i B_i \right) \dot{\mathbf{x}} + \left(\sum_{i=1}^{k_3} \chi_i C_i \right) \mathbf{x} = \mathbf{b}, \quad (8.2)$$

where α_i , β_i and χ_i are random variables, and A_i , B_i and C_i are deterministic matrices. The major task of this phase is to develop efficient numerical methods for the solution of Eqs. (8.1) and (8.2).

It is highlighted that, in contrast to conventional deterministic algebraic equation systems, the solution of stochastic Eqs. (8.1) and (8.2) is a much underdeveloped area in numerical analysis and no effective solution strategies are readily available, particularly for large scale problems. On the other hand, this situation can offer exciting opportunities to explore the problem in many different ways, and contributions made in this respect will be of benefit not only to the increasingly important stochastic modelling field but also to a wide range of scientific and engineering communities as well.

It is well known that for a deterministic system of linear algebraic equations $A\mathbf{x} = \mathbf{b}$, there exist various numerical algorithms developed for different types of matrices and different solution requirements, which are all explicitly or implicitly based on inverting the matrix under consideration. For a more general stochastic system of linear algebraic equations (8.1), it can be similarly expected that there will be different numerical algorithms based on the joint diagonalization of multiple matrices, which essentially give an approximate inverse of the matrix family. For the solution of the stochastic system of ordinary differential equations (8.2), it is clear that the joint diagonalization strategy can also be employed to approximately decouple the equation so that the complexity in obtaining the solution can be significantly reduced.

8.2.4 Verification and Application

Unlike the pure research in SPDEs which is based on highly abstract mathematical

definitions and driven by the elegance of mathematics, this work was initiated by practical applications. Its objective is not to develop a comprehensive SPDE theory, and instead the main motivation is to effectively analyze practical engineering systems composed of random media. Hence, the theory and algorithms developed in this work must be verified and further applied to practical engineering problems. Specifically, the following three classes of engineering systems are suggested for detailed investigation.

- Large-scale rock, concrete and metal infrastructures in which joints, flaws and microcracks commonly exist. A key issue in the construction and management of these engineering structures is risk assessment. The discontinuity and irregular spatial variation of material properties are the major challenges in the analysis of these engineering systems, and they cannot be met by the conventional deterministic analysis tools without introducing a conservative safety factor. The probabilistic material model provides a rational tool to describe random material properties and discontinuities through the medium. Hence, it is expected that a more accurate risk assessment can be achieved by the theory and computational methods to be developed in the future research, which in turn significantly reduce the associated costs in the design, manufacture and maintenance of these engineering infrastructures.
- Micro-electro-mechanical systems that generally range in size from a micrometer to a millimetre. Common applications of MEMS include inkjet printers, accelerometers in modern cars for airbag deployment in collisions, and disposable blood pressure sensors etc. Due to the difficulties in performing accurate experiments at the micro scale, the standard finite element analysis has been extensively used in the design of MEMS. However, as heterogeneous material properties are inevitably present at the scale these devices normally operate, the simulation results from FEM are often controversial. Hence, the analysis tools developed for general random media in the future work could be highly beneficial to the development of MEMS.
- Human tissues such as bones and teeth that have multi-scale irregular material structures. The material properties of human tissues bear significant difference between different individuals, and they also vary for the same individual, especially when the person is ill. In biomedical engineering, there has been an increasing need for providing computational tools to assist medical diagnosis and treatment. The

intrinsic randomness existing in human tissues makes stochastic modelling an attractive approach. It is therefore expected that the probabilistic material model and the associated analysis tools developed in the future work will find its use in biomedical engineering, especially in the analysis of bones and teeth.

An important phenomenon that requires careful treatment in all applications above is size-effects, which are commonly observed in most practical random media. To date, the only probabilistic model to account for size effects of random media is Weibull's statistical model [8.4-8.5], in which random media are simplified into a chain connected with a sequence of independent rings. Weibull's model is often used to qualitatively explain size effects, but it has been generally discarded in engineering analysis because it relies on adjustable artificial parameters [8.6]. Also, Weibull's model runs into trouble when the scale of the random medium tends to infinity. The failure of Weibull's model actually comes from its over simplification and not taking into consideration the interdependence of separate points within the medium. The ERM model (including its improvements suggested in Section 8.2.1) is based on the probabilistic essence of practical random media, and it fully takes into account the spatial stochastic dependence between different material properties. Hence, it is expected that the related future research could provide a meaningful tool for the analysis and prediction of size-effects in practical random media. This important aspect should be thoroughly investigated and verified in the context of all the above applications.

References

- [8.1] CMI Millennium Problems, <http://www.claymath.org/millennium/>.
- [8.2] S.M. Hu, C.F. Li and H. Zhang, Actual morphing: a physics-based approach for blending two topologically identical shapes, *ACM Symposium on Solid Modelling and Applications*, Genoa, Italy, June 9-11, 2004.
- [8.3] C.F. Li, Y.T. Feng and D.R.J. Owen, SMB: collision detection based on temporal coherence, *Computer Methods in Applied Mechanics and Engineering*, 195 (19-22) (2006) 2252–2269.
- [8.4] W. Weibull, A statistical theory of the strength of materials, *Royal Swedish Academy of Engineering Sciences*, 151 (1939) 1-45.

- [8.5] W. Weibull, A statistical distribution function of wide applicability, *Journal of Applied Mechanics – ASME*, 18 (1951) 293-297.
- [8.6] Z.P. Bažant and J. Planas, *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*, CRC Press, New York, 1998.

List of Publications

Items marked by superscript † are finished during the course of this work and items marked by superscript ‡ are directly related to the subject addressed in this thesis.

Journal Papers

- [1] C.F. Li, N. Hu, Y.J. Yin, H. Sekine and H. Fukunaga, Low-velocity impact-induced damage of continuous fiber-reinforced composite laminates: part I an FEM numerical model, *Composites: Part A Applied Science and Manufacturing*, 33 (8) (2002) 1055-1062, The most downloaded paper in Elsevier JCOMA 2003:
http://www.extranet.elsevier.com/listman/Sbelcher/CompositesA_July04.htm
- [2] C.F. Li, N. Hu, J.G. Cheng, H. Fukunaga and H. Sekine, Low-velocity impact-induced damage of continuous fiber-reinforced composite laminates: part II verification and numerical investigation, *Composites: Part A Applied Science and Manufacturing*, 33 (8) (2002) 1063-1072.
- [3] C.F. Li, J.G. Cheng, N. Hu and Z.H. Yao, A numerical simulation for impact damages of long-fibre reinforced resin laminates, *Acta Mechanica Sinica*, 34 (6) (2002) 989-997. (In Chinese)
- [4] C.F. Li, J.G. Cheng, N. Hu and Z.H. Yao, Delaminated composite laminates' responses under low-velocity impacts, *Acta Meteriae Compositae Sinica*, 20 (1) (2003) 38-45. (In Chinese)
- † [5] C.F. Li, Y.T. Feng and D.R.J. Owen, SMB: collision detection based on temporal coherence, *Computer Methods in Applied Mechanics and Engineering*, 195 (19-22) (2006) 2252-2269.
- † ‡ [6] C.F. Li, Y.T. Feng and D.R.J. Owen, Explicit solution to stochastic system of linear algebraic equations $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)x = b$, *Computer Methods in Applied Mechanics and Engineering*, 195 (44-47) (2006) 6560-6576.
- † ‡ [7] C.F. Li, Y.T. Feng, D.R.J. Owen and I.M. Davies, Fourier representation of random media fields in stochastic finite element modelling, *Engineering Computations*, 23 (7) (2006) 794-817.

- † ‡ [8] C.F. Li, Y.T. Feng, D.R.J. Owen, D.F. Li and I.M. Davies, A Fourier-Karhunen-Loève representation scheme for wide-sense stationary stochastic fields, *International Journal for Numerical Methods in Engineering*, Submitted.
- † [9] D.F. Li, J. Lu, C.F. Li and H. Qing, Solving multiple-crack interaction problems with infinite-dimensional linear systems, *Journal of Applied Mechanics – ASME*, Submitted.
- † ‡ [10] Y.T. Feng, C.F. Li and D.R.J. Owen, A directed Monte Carlo solution of linear stochastic algebraic system of equations, *International Journal for Numerical Methods in Engineering*, Submitted.
- † [11] C.F. Li, Y.T. Feng and D.R.J. Owen, On the quadrature of certain multidimensional oscillatory functions, *Computer Physics Communications*, In preparation.
- † ‡ [12] C.F. Li, Y.T. Feng and D.R.J. Owen, Elastostatics of elementary random media: part I – theory and algorithms, *Journal of the Mechanics and Physics of Solids*, In preparation.
- † ‡ [13] C.F. Li, Y.T. Feng and D.R.J. Owen, Elastostatics of elementary random media: part II – verification and application, *Journal of the Mechanics and Physics of Solids*, In preparation.

Conference Papers and Presentations

- † [14] S.M. Hu, C.F. Li and H. Zhang, Actual morphing: a physics-based approach for blending two topologically identical shapes, *ACM Symposium on Solid Modelling and Applications*, Genoa, Italy, June 9-11, 2004.
- † ‡ [15] C.F. Li, Y.T. Feng and D.R.J. Owen, Continuum mechanics and finite elements in stationary stochastic fields: the formulation, *The Third M.I.T. Conference of Computational Fluid and Solid Mechanics*, Massachusetts Institute of Technology, Cambridge, USA, June 14-17, 2005.
- † ‡ [16] Y.T. Feng, C.F. Li and D.R.J. Owen, Continuum mechanics and finite elements in stationary stochastic fields: computational issues, *The Third M.I.T. Conference of Computational Fluid and Solid Mechanics*, Massachusetts Institute of Technology, Cambridge, USA, June 14-17, 2005.
- † ‡ [17] C.F. Li, Y.T. Feng, D.R.J. Owen and I.M. Davies, A simple semi-analytic solution for Karhunen-Loève expansions of stationary stochastic fields, *The Seventh World Congress on Computational Mechanics*, Los Angeles, California, USA, July 16-22, 2006.
- † ‡ [18] Y.T. Feng, C.F. Li and D.R.J. Owen, Joint diagonalization solution for the stochastic system of linear algebraic equations $(\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_m A_m)x = b$, *The Seventh World Congress on Computational Mechanics*, Los Angeles, California, USA, July 16-22, 2006.