Advanced surrogate modeling and machine learning methods in computational stochastic mechanics

By
Ioannis Kalogeris

School of Civil Engineering
Institute of Structural Analysis and Antiseismic Research

National Technical University of Athens

Supervisor: Associate Professor Vissarion Papadopoulos

A thesis submitted for the degree of
Doctor of Philosophy

January, 2020

APPROVAL

PhD THESIS EXAMINATION COMMITTEE:

Associate Professor Vissarion Papadopoulos
(Supervisor and Principal Advisor of the Committee)
National Technical University of Athens
School of Civil Engineering

Emeritus Professor Manolis Papadrakakis
(Member Advisor of the Committee)
National Technical University of Athens
School of Civil Engineering

Assistant Professor Dimitrios Vamvatsikos
(Member Advisor of the Committee)
National Technical University of Athens
School of Civil Engineering

Professor Andreas Boudouvis
(Member of the Examination Committee)
National Technical University of Athens
School of Chemical Engineering

Professor Demetris Koutsoyiannis
(Member of the Examination Committee)
National Technical University of Athens
School of Civil Engineering

Professor Konstantinos Spiliopoulos
(Member of the Examination Committee)
National Technical University of Athens
School of Civil Engineering

Associate Professor Nikolaos Lagaros
(Member of the Examination Committee)
National Technical University of Athens
School of Civil Engineering

To my parents.

Acknowledgments

Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people.

First and foremost, i want to thank my father, my mother, my brother and my sister for believing in me and encouraging me to follow my dreams. There are no words to express my gratitude for their constant support and unconditional love throughout these years and my life in general.

This thesis would not have been possible without the guidance, support and motivation of my supervisor Associate Professor Vissarion Papadopoulos. Thank you for your patience, your continuous belief in me and for being so generous with your time and intellect. But most importantly, thank you for the excellent example you have provided for me.

Besides my advisor, i would like to thank the rest of my thesis committee for their insightful comments and encouragement. I would like to particularly thank Emeritus Professor Manolis Papadrakakis and assistant professor Dimitris Vamvatsikos for participating in my supervising committee and for being ever available for consultation.

I want to express my gratitude to my dear friend and colleague Dr. Dimitris Giovanis for encouraging me to pursue a doctoral degree and for mentoring me throughout the process. Our conversations were always a source of inspiration to me.

I also thank my esteemed colleagues for the stimulating discussions, the sleepless nights we were working together before deadlines, and for all the fun we have had in the last five years.

Special thanks to all my friends for their understanding, encouragement, and above all, their friendship.

I gratefully acknowledge the funding received towards my Ph.D from the European Council Advanced Grant MASTER - *Mastering the computational challenges in numerical modeling and optimum design of CNT reinforced composites*, from the State Scholarships Foundation Grant *Research projects for excellency, IKY/SIEMENS* and from the European Regional Development Fund and Greek national funds under the Grant HEAT - *Optimal multiscale design of innovative materials for heat exchange applications.*

Advanced surrogate modeling and machine learning methods in computational stochastic mechanics

Abstract

Parameter uncertainty quantification and methods to predict uncertainty propagation in structural problems have become an essential part of the analysis and design of engineering applications. In the effort to include unpredictable factors in modeling, the stochastic finite element method (SFEM) has been developed over the past decades. SFEM can be seen as an extension of the classic finite element method to systems with inherent uncertainties in their initial conditions, materials or geometry. The governing equations of such systems are, in general, stochastic partial differential equations and SFEM can be seen as a powerful computational tool for solving this type of equations. From a mathematical point of view, analytic solutions to stochastic partial differential equations can be attained only in a few cases, while for complex realistic systems numerical methods are the only viable option. Undoubtedly, the Monte Carlo simulation is the most popular and widely applied technique in the field of SFEM. However, this method relies on a large number of random model simulations and its computational cost becomes unaffordable in cases where each run of the computer model takes minutes or hours to complete. This thesis presents a series of methodologies based on surrogate modeling and machine learning techniques, that have been implemented in the framework of SFEM in order to bypass the computational demands of the direct Monte Carlo simulation.

The first methodology proposes a reformulation of the Probability Density Evolution Method (PDEM), which is typically used to study the evolution of stochastic dynamic systems, to the case of general stochastic systems. In addition, the Streamline Upwind/ Petrov Galerkin scheme has been successfully employed to solve the corresponding generalized density evolution equations, which increased significantly the accuracy of the method and reduced the computational burden. In the second methodology, PDEM has been utilized for the estimation of stochastic limit loads and stochastic load-displacement curves in the frame of nonlinear structural analysis. The third methodology proposes an extension of the intrusive Spectral Stochastic Finite Element Method for geometrically nonlinear framed structures, where the stochastic incremental displacements are projected onto the Polynomial Chaos basis and their values are obtained by solving an augmented system of nonlinear equations. Lastly, a methodology has been put forth for the construction of an emulator of the computational model using a manifold learning tool, namely the Diffusion Maps algorithm, which is utilized in order to reduce to computational burden of the Monte Carlo simulation.

ΠΕΡΙΛΗΨΗ ΤΗΣ ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

με τίτλο

‘Προχωρημένες μέθοδοι υποκατάστατης μοντελοποίησης και μηχανικής μάθησης στην υπολογιστική στοχαστική μηχανική’

Η ποσοτικοποίηση της αβεβαιότητας στις παραμέτρους ενός μηχανικού συστήματος και οι μέθοδοι για τον προσδιορισμό της επιρροής της στην απόκριση αυτού αποτελούν ένα ουσιώδες κομμάτι της ανάλυσης και του σχεδιασμού των κατασκευών. Για το σκοπό αυτό αναπτύχθηκε η μέθοδος των στοχαστικών πεπερασμένων στοιχείων, η οποία έχει ως στόχο τη μελέτη συστημάτων, στα οποία ενυπάρχουν αβεβαιότητες στις παραμέτρους του συστήματος (πχ. ιδιότητες υλικών), στις συνοριακές συνθήκες, στη γεωμετρία και στη φόρτιση. Η πιθανοτική περιγραφή των αβέβαιων ποσοτήτων απαιτεί τη χρησιμοποίηση της θεωρίας των στοχαστικών ανελίξεων (ή πεδίων), δηλαδή τυχαίων μεταβλητών, οι οποίες παρουσιάζουν μια εξέλιξη στον χώρο ή/και στο χρόνο. Ως αποτέλεσμα, η απόκριση ενός τέτοιου συστήματος δεν θα είναι πλέον ντετερμινιστική, αλλά θα περιγράφεται και αυτή από κάποιο στοχαστικό πεδίο, το οποίο καλείται η μέθοδος των στοχαστικών πεπερασμένων στοιχείων να προσδιορίσει.

Η κυριότερη και πιο διαδεδομένη μέθοδος στην κατηγορία των στοχαστικών πεπερασμένων στοιχείων είναι η μέθοδος προσομοίωσης Monte Carlo. Η μέθοδος αυτή στηρίζεται σε τυχαίες δειγματοληψίες από τον χώρο των τυχαίων παραμέτρων του προβλήματος και για κάθε μία από αυτές εκτελείται μια ντετερμινιστική ανάλυση του συστήματος. Έπειτα, η στοχαστική απόκριση του προβλήματος προκύπτει από τη στατιστική επεξεργασία των αποτελεσμάτων των ντετερμινιστικών επιλύσεων. Στην πραγματικότητα, η μέθοδος αυτή είναι η μόνη ικανή να χειριστεί στοχαστικά προβλήματα στα οποία εμπλέκονται μη γραμμικότητες, δυναμικές φορτίσεις, προβλήματα αστάθειας, παραμετρικές διεγέρσεις κλπ. Γι' αυτό και χρησιμοποιείται ως μέθοδος αναφοράς με τα αποτελέσματα της οποίας συγκρίνονται τα αποτελέσματα άλλων μεθόδων. Όμως, παρά το γεγονός ότι η μέθοδος αυτή μπορεί θεωρητικά να εφαρμοστεί σε οποιουδήποτε τύπου στοχαστικό πρόβλημα, εύκολα γίνεται αντιληπτό, ότι σε λεπτομερή προσομοιώματα με πολλούς βαθμούς ελευθερίας ή/και σε μη-γραμμικά προβλήματα, όπου η διάρκεια της μιας ανάλυσης κυμαίνεται από μερικά λεπτά έως μερικές ώρες, το υπολογιστικό κόστος αυτής της προσέγγισης καθίσταται ασύμφορο.

Ένας αποτελεσματικός τρόπος για τη μείωση του υπολογιστικού κόστους της μεθόδου

Monte Carlo είναι με χρήση τεχνικών υποκατάστατης μοντελοποίησης (surrogate modeling techniques). Ένα υποκατάστατο μοντέλο (ή μετα-μοντέλο) είναι μια προσέγγιση του αρχικού, σύνθετου, προσομοιώματος με κάποιο απλούστερο που έχει 'σχεδόν' ισοδύναμη συμπεριφορά. Η κατασκευή του υποκατάστατου μοντέλου βασίζεται είτε σε κάποιες παραδοχές για το συναρτησιακό σχήμα του μοντέλου, είτε στην αξιοποίηση της πληροφορίας που παρέχεται από έναν περιορισμένο αριθμό επιλύσεων του πλήρους προσομοιώματος. Στα πλαίσια της στοχαστικής ανάλυσης, ορισμένες ιδιαίτερα διαδεδομένες τεχνικές είναι η μέθοδος των στοχαστικών διαταραχών (stochastic perturbation method), η μέθοδος της στοχαστικής συγκατάταξης (stochastic collocation technique) και η μέθοδος της παλινδρόμησης Γκαουσιανής ανέλιξης (Gaussian process regression). Μια εναλλακτική προσέγγιση για την δημιουργία υποκατάστατων μοντέλων είναι μέσω της χρήσης τεχνικών μηχανικής μάθησης, οι οποίες θα επιτρέψουν την κατασκευή απλούστερων μαθηματικών εκφράσεων ικανών να αντικαταστήσουν το προσομοίωμα πεπερασμένων στοιχείων. Το κύριοτερο παράδειγμα σε αυτή την κατηγορία είναι τα τεχνητά νευρωνικά δίκτυα με τις διάφορες παραλλαγές τους (deep feed-forward networks, convolutional networks, recurrent networks, deep belief networks κ.ά.).

Τα τελευταία χρόνια αναπτύχθηκε η Μέθοδος της Εξέλιξης της Πυκνότητας Πιθανότητας (Probability Density Evolution Method), η οποία προτάθηκε από τους Li and Chen και βασίζεται στο θεώρημα διατήρησης της πιθανότητας που διέπει ένα στοχαστικό δυναμικό σύστημα. Σύμφωνα με το θεώρημα αυτό, σε ένα τυχαίο χωρίο στον χώρο των καταστάσεων του συστήματος, η χρονική μεταβολή της πιθανότητας είναι ίση με την πιθανότητα που εισέρχεται/εξέρχεται από το σύνορο του χωρίου για αυτό το χρονικό διάστημα. Η άμεση εφαρμογή αυτού του θεωρήματος οδηγεί στις μερικές διαφορικές εξισώσεις Liouville και Fokker-Planck, οι οποίες περιγράφουν τη χρονική εξέλιξη της απο κοινού συνάρτησης πυκνότητας πιθανότητας όλων των καταστάσεων του συστήματος. Παρότι οι εξισώσεις αυτές περιέχουν όλη την πληροφορία για τη στοχαστική απόκριση του συστήματος, στην πράξη δεν χρησιμοποιούνται καθώς δεν επιδέχονται αναλυτική ή αριθμητική λύση λόγω του

μεγάλου αριθμού των όρων που περιέχουν.

Η μέθοδος εξέλιξης της συνάρτησης πυκνότητας πιθανότητας αναδιατυπώνει τις εξισώσεις αυτές από τη σκοπιά του τυχαίου γεγονότος (Λαγκρανζιανή περιγραφή) και με αυτόν τον τρόπο οι πολυδιάστατες διαφορικές εξισώσεις μετατρέπονται σε ένα σύνολο απλούστερων διαφορικών εξισώσεων, που καλούνται γενικευμένες εξισώσεις εξέλιξης πιθανότητας. Χωρίς βλάβη της γενικότητας, ας θεωρήσουμε ένα στοχαστικό δυναμικό πρόβλημα της μορφής:

$$\mathbf{M}(\boldsymbol{\theta})\ddot{\mathbf{u}} + \mathbf{C}(\boldsymbol{\theta})\dot{\mathbf{u}} + \mathbf{f}(\boldsymbol{\theta}, \mathbf{u}) = \mathbf{F}(\boldsymbol{\theta}, t) \tag{1}$$

με $\mathbf{M}$, $\mathbf{C}$, $\mathbf{f}$, $\mathbf{F}$ να είναι αντίστοιχα το μητρώο μάζας, το μητρώο απόσβεσης, το διάνυσμα της δύναμης επαναφοράς και το διάνυσμα της εξωτερικής διέγερσης. Επίσης, με $\boldsymbol{\theta}$ συμβολίζεται το τυχαίο διάνυσμα που περιέχει όλες τις τυχαίες μεταβλητές του συστήματος. Το θεώρημα διατήρησης για αυτό το πρόβλημα εκφράζεται μέσω της σχέσης:

$$\frac{D}{Dt}\int_{\Omega_t \times \Omega_\theta} p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)\,\mathrm{d}\mathbf{u}\mathrm{d}\boldsymbol{\theta} = 0 \tag{2}$$

όπου $\Omega_t \times \Omega_\theta$ ένα τυχαίο χωρίο στον διευρυμένο χώρο των καταστάσεων και των τυχαίων μεταβλητών, $\frac{D}{Dt}$ είναι η ολική παράγωγος, και $p_{\mathbf{u}\boldsymbol{\theta}}$ είναι η από κοινού συνάρτηση πυκνότητας πιθανότητας των $(\mathbf{u}, \boldsymbol{\theta})$. Για την εφαρμογή της μεθόδου της εξέλιξης πιθανότητας, απαιτείται η διαμέριση του δειγματόχωρου $\Omega_{\boldsymbol{\theta}}$ σε n_{pt} διακριτά σύνολα τέτοια ώστε $\cup_{q=1}^{n_{pt}}\Omega_q = \Omega_{\boldsymbol{\theta}}$ και για κάθε ένα από αυτά τα σύνολα προσδιορίζεται ένα χαρακτηριστικό στοιχείο $\boldsymbol{\theta}_q \in \Omega_q$. Έπειτα, θεωρούμε ένα υποκατάστατο σύστημα, του οποίου η απόκριση για κάθε $\boldsymbol{\theta} \in \Omega_q$ είναι σταθερή και ίση με $u(\boldsymbol{\theta}_q, t)$, δηλαδή την απόκριση του αρχικού συστήματος για το δεδομένο $\boldsymbol{\theta}_q$. Η διαδικασία αυτή θα μετατρέψει την προηγούμενη εξίσωση σε ένα σύνολο εξισώσεων της μορφής:

$$\frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}_q, t)}{\partial t} + \dot{u}(\boldsymbol{\theta}_q, t)\frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}_q, t)}{\partial u} = 0, \quad q = 1, 2, ..., n_{pt} \tag{3}$$

οι οποίες πλέον αναφέρονται σε έναν βαθμό ελευθερίας u και λύνονται ξεχωριστά σε κάθε χωρίο Ω_q. Για την επίλυση των παραπάνω εξισώσεων απαιτείται η γνώση της ποσότητας $\dot{u}(\boldsymbol{\theta}_q, t)$, η οποία θα προκύψει από μια ντετερμινιστική επίλυση του συστήματος για $\boldsymbol{\theta} = \boldsymbol{\theta}_q$. Η τελική λύση του προβλήματος θα προκύψει από την υπέρθεση όλων των λύσεων $p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}_q, t)$ για $q = 1, ..., n_{pt}$.

Στην παρούσα διατριβή διατυπώνεται μια νέα θεώρηση της μεθόδου εξέλιξης της συνάρτησης πυκνότητας πιθανότητας, η οποία εκφράζει το θεώρημα διατήρησης της πιθανότητας στο χώρο αντί στο χρόνο, καθιστώντας τη μέθοδο κατάλληλη και για την επίλυση στατικών στοχαστικών προβλημάτων. Για ένα σύστημα της μορφής:

$$\mathbf{K}(\boldsymbol{\theta})\mathbf{u} = \mathbf{F}(\boldsymbol{\theta}) \tag{4}$$

οι αντίστοιχες γενικευμένες εξισώσεις διατήρησης πιθανότητας που θα προκύψουν είναι οι:

$$\frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}_q, x)}{\partial x} + \frac{du(\boldsymbol{\theta}_q, x)}{dx} \frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}_q, x)}{\partial u} = 0, \quad q = 1, 2, ..., n_{pt} \tag{5}$$

Οι εξισώσεις αυτές πλέον περιγράφουν την κατανομή της πιθανότητας σε κάθε θέση x του φορέα. Η διαφορά με τις προηγούμενες εξισώσεις είναι ότι εδώ απαιτείται η γνώση της παραγώγου της απόκρισης ως προς τη θέση x, αντί του χρόνου t. Επιπλέον, οι συνοριακές συνθήκες της διαφορικής προκύπτουν από τις συνοριακές συνθήκες του φυσικού προβλήματος, καθώς στις στηρίξεις του φορέα γνωρίζουμε με πιθανότητα 1 ότι η απόκριση θα είναι μηδενική.

Το υπολογιστικό κέρδος της μεθόδου έγκειται στο γεγονός ότι οι επιλύσεις του συστήματος που απαιτούνται είναι σημαντικά λιγότερες από αυτές που θα απαιτούσε η μέθοδος Monte Carlo . Ωστόσο, το όφελος αυτό μειώνεται λόγου του κόστους επίλυσης των μερικών διαφορικών εξισώσεων (5). Επιπρόσθετα, οι εξισώσεις αυτές είναι εξισώσεις μεταφοράς και οι αρχικές τους συνθήκες είναι τύπου 'shock', καθιστώντας την αριθμητική επίλυσή τους ιδιαίτερα δυσχερή. Για το σκόπο αυτό, στη βιβλιογραφία προτείνεται μια πα-

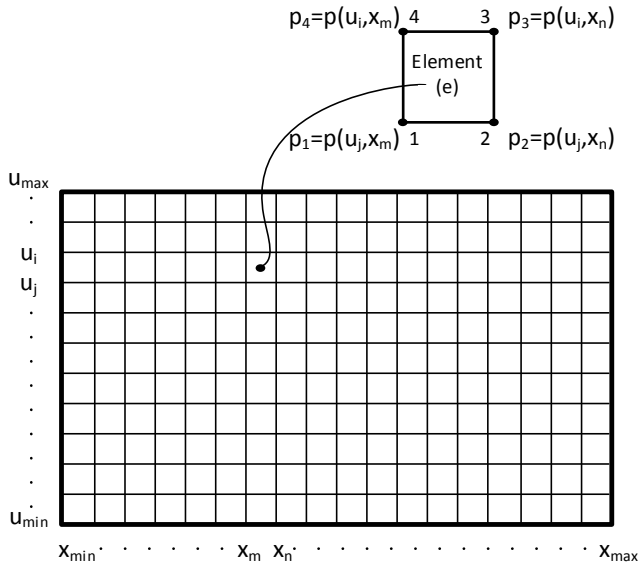
ραλλαγή του αριθμητικού σχήματος πεπερασμένων διαφορών Lax-Wendroff , εφοδιασμένο με έναν όρο flux limiter για να επιτευχθεί αριθμητική εξομάλυνση των ταλαντώσεων λόγω αρχικών συνθηκών. Όμως, το σχήμα αυτό τείνει να εισάγει σημαντική αριθμητική διάχυση και, επίσης, η σύγκλιση του υπόκειται στον περιορισμό Courant-Friedrichs-Lewy (CFL), ο οποίος επιβάλλει ιδιαίτερα μικρά βήματα Δx για την επίλυση της διαφορικής, αυξάντας έτσι το κόστος επίλυσης. Για το λόγο αυτό, στην παρούσα ερευνητική προσπάθεια διευρενήθηκαν δύο εναλλακτικά σχήματα πεπερασμένων στοιχείων, τα Discontinuous Galerkin time-marching scheme και Streamline Upwind/PetrovGalerkin finite element scheme , τα οποία προσαρμόστηκαν κατάλληλα στον τύπο του προβλήματος. Από τη διερεύνηση αυτή αναδείχθηκαν σημαντικά πλεονεκτήματα σε ακρίβεια και υπολογιστικό χρόνο του σχήματος Streamline Upwind/Petrov-Galerkin έναντι των άλλων, καθώς αυτό είναι ανεξάρτητο του περιορισμού CFL. Η εφαρμογή αυτού του σχήματος απαιτεί η διαφορική (5) να γραφεί στη μορφή

$$\boldsymbol{w}\nabla p = 0 \tag{6}$$

όπου $\boldsymbol{w} = (1, \alpha(x))$ το πεδίο ταχυτήτων, με $\alpha(x) = \frac{du(x)}{dx}$ και $\nabla p = (\frac{\partial p}{\partial x}, \frac{\partial p}{\partial u})$. Έπειτα, το χωρίο $\Omega = \Omega_u \times \Omega_x = [u_{min}, u_{max}] \times [x_{min}, x_{max}]$, στο οποίο υπολογίζεται η διαφορική, διακριτοποιείται σε πεπερασμένα στοιχεία που υπολογίζουν την πιθανότητα να εμφανιστεί στη θέση x_m η απόκριση u_j, όπως φαίνεται στο σχήμα 1.

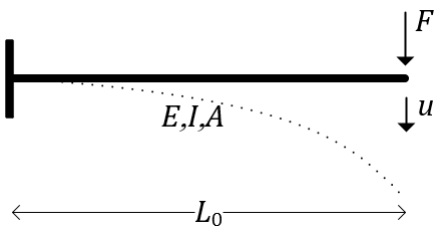
Η ακρίβεια της προτεινόμενης μεθοδολογίας και τα αντίστοιχα υπολογιστικά οφέλη επαληθεύονται μέσω αριθμητικών εφαρμογών, που περιλαμβάνουν στατικά συστήματα πεπερασμένων στοιχείων με τυχαίες παραμέτρους στις ιδιότητες των υλικών και τη φόρτιση.

Στη συνέχεια της διατριβής, προτείνεται μια νέα διατύπωση της μεθόδου της εξέλιξης της συνάρτησης πυκνότητας πιθανότητας, προσαρμοσμένη στη στοχαστική μη-γραμμική ανάλυση κατασκευών και ειδικότερα στην οριακή ανάλυση. Αν θεωρήσουμε για παράδειγμα μια κατασκευή, όπως ο πρόβολος του σχήματος 2, για τον οποίο οι παράμετροι του υλικού και της διατομής περιγράφονται από ένα τυχαίο διάνυσμα $\boldsymbol{\theta} = (E, I, A)$, τότε γίνεται

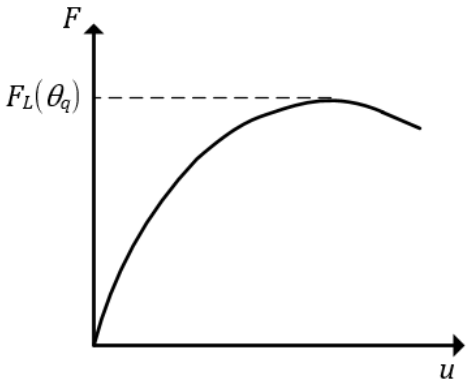


Σχήμα 1: Διακριτοποίηση του χωρίου της διαφορικής σε τετραπλευρικά πεπερασμένα στοιχεία με ένα βαθμό ελευθερίας ανά κόμβο, που αντιστοιχεί στην πιθανότητα να εμφανιστεί η συγκεκριμένη τιμή μετατόπισης σε αυτή τη θέση

εύκολα αντιληπτό ότι το οριακό φορτίο F_L της κατασκευής θα είναι διαφορετικό για κάθε διαφορετική τιμή του διανύσματος $\boldsymbol{\theta}$.



Σχήμα 2: Πρόβολος με κατακόρυφο φορτίο στο άκρο και $\boldsymbol{\theta} = (E, I, A)$ το διάνυσμα τυχαίων παραμέτρων



Σχήμα 3: Οριακό φορτίο κατάρρευσης για δεδομένες τιμές των τυχαίων παραμέτρων $\boldsymbol{\theta_q}$

Για την εφαρμογή της μεθόδου εξέλιξης της συνάρτησης πυκνότητας πιθανότητας, το οριακό φορτίο κατάρρευσης F_L κατά την μη-γραμμική ανάλυση αποκτά την έννοια ψευδο-χρόνου και η πιθανότητα κατάρρευσης του φορέα εκφράζεται μέσω αντίστοιχων γενικευμένων διαφορικών εξισώσεων εξέλιξης πιθανότητας, η οποίες επιλύονται με τη μέθοδο των χαρακτηριστικών. Αρχικά, ορίζουμε μια συνάρτηση

$$i(F_L, \theta_q) = \begin{cases} 1, & \text{αν } F_L >= F_L(\theta_q) \\ \\ 0, & \text{αν } F_L < F_L(\theta_q) \quad \text{φορ } q = 1, 2, ..., n_{pt} \end{cases} \tag{7}$$

τέτοια ώστε η i να παίρνει την τιμή 1 όταν η κατασκευή καταρρέει και 0 διαφορετικά.

Δηλαδή,

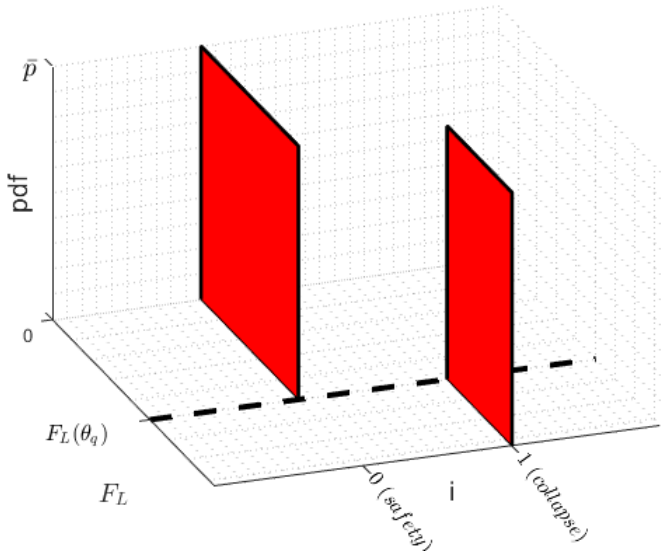
$$i(F_L, \theta_q) = H(F_L - F_L(\theta_q)) \quad q = 1, 2, ..., n_{pt} \tag{8}$$

Στην περίπτωση αυτή, η από κοινού συνάρτηση πυκνότητας πιθανότητας $p_{i\boldsymbol{\theta}}$ θα εξαρτάται από τη συνάρτηση $i(F_L, \boldsymbol{\theta})$ και από το οριακό φορτίο F_L. Έτσι, οι γενικευμένες εξισώσεις διατήρησης πιθανότητας μπορούν να γραφούν ως:

$$\frac{\partial p_{i\boldsymbol{\theta}}(i, \theta_q, F_L)}{\partial F_L} + \frac{di(\theta_q, F_L)}{dF_L} \frac{\partial p_{i\boldsymbol{\theta}}(i, \theta_q, F_L)}{\partial i} = 0, \ q = 1, 2, ..., n_{pt} \tag{9}$$

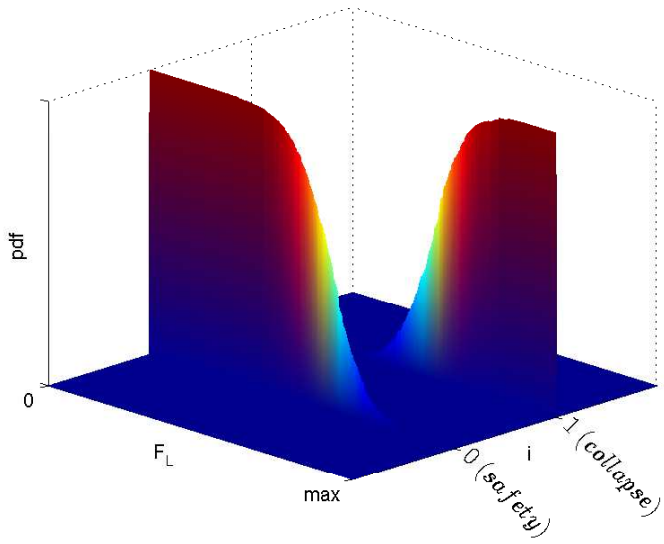
Η παραπάνω εξισώσεις επιλύονται αναλυτικά με τη μέθοδο των χαρακτηριστικών και η λύση τους δίνεται από το τύπο:

$$p_q(i, F_L) = \delta\left(i - H\left(F_L - F_L\left(\boldsymbol{\theta}_q\right)\right)\right) \bar{p}_q \tag{10}$$



Σχήμα 4: Λύση μια γενικευμένης εξίσωσης διατήρησης πιθανότητας για δεδομένο θ_q

Η υπέρθεση των λύσεων για όλα τα θ_q με $q = 1, 2, ..., n_{pt}$ θα οδηγήσει σε ένα σχήμα όπως αυτό του σχ. 5, και το προφίλ των λύσεων για $i = 1$ θα είναι ουσιαστικά η συνάρτηση κατανομής πιθανότητας του κρίσιμου φορτίου αστοχίας.



Σχήμα 5: Συνάρτηση πυκνότητας πιθανότητας της κατάστασης αστοχίας για διάφορες τιμές του οριακού φορτίου F_L

Επιπρόσθετα, για την εύρεση του στοχαστικού δρόμου ισορροπίας κατά τη μη-γραμμική ανάλυση κατασκευών προτείνεται μια αναπροσαρμογή των αλγορίθμων ελέγχου μετατόπισης και ελέγχου φορτίου. Στην πρώτη περίπτωση υπολογίζονται τα στοχαστικά χαρακτηριστικά του φορτίου στον κόμβο ελέγχου για τις διάφορες τιμές της επιβαλλόμενης μετατόπισης μέσω των εξισώσεων:

$$\frac{\partial p_{F\boldsymbol{\theta}}(F, \theta_q, \bar{u})}{\partial \bar{u}} + \frac{dF(\theta_q, \bar{u})}{d\bar{u}} \frac{\partial p_{F\boldsymbol{\theta}}(F, \theta_q, \bar{u})}{\partial F} = 0, \quad q = 1, 2, ..., n_{pt} \tag{11}$$

Κατ΄ αντιστοιχία, στη δεύτερη περίπτωση υπολογίζονται τα στοχαστικά χαρακτηριστικά

της μετατόπισης για τις διάφορες τιμές του επιβαλλόμενου φορτίου.

$$\frac{\partial p_{\bar{u}\boldsymbol{\theta}}(\bar{u},\theta_q,F)}{\partial F}+\frac{d\bar{u}(\theta_q,F)}{dF}\frac{\partial p_{\bar{u}\boldsymbol{\theta}}(\bar{u},\theta_q,F)}{\partial \bar{u}}=0,\ \ q=1,2,...,n_{pt} \tag{12}$$

Για την επαλήθευση της ακρίβειας της προτεινόμενης μεθοδολογίας παρατίθενται αριθμητικές εφαρμογές κατασκευών με τυχαιότητα στις παραμέτρους των υλικών και ατέλειες στις γεωμετρία τους. Η αριθμητική διερεύνηση κατέδειξε τα σημαντικά υπολογιστικά οφέλη της μεθόδου, καθώς αυτή είναι ικανή να επιτύχει την ίδια ακρίβεια με την προσομοίωση Monte Carlo με πολύ λιγότερες όμως επιλύσεις του ντετερμινιστικού προβλήματος.

Σε επόμενο στάδιο της διδακτορικής διατριβής μελετάται μια άλλη μέθοδος υποκατάστατης μοντελοποίησης, αυτή των Φασματικών Στοχαστικών Πεπερασμένων στοιχείων, η οποία αναπτύχθηκε από τους Ghanem & Spanos και είναι ιδιαίτερα διαδεδομένη για τη στοχαστική ανάλυση κατασκευών. Η μέθοδος αυτή βασίζεται σε δύο παραδοχές για το πρόβλημα. Αρχικά, θεωρούμε ότι τα στοχαστικά πεδία που περιγράφουν τις παραμέτρους του προβλήματος μπορούν να γραφούν σε ανάπτυγμα σε σειρά Karhunen-Loéve. Παραδείγματος χάριν, το ανάπτυγμα Karhunen-Loéve ενός Γκαουσιανού στοχαστικού πεδίου $a(\mathbf{x},\theta)$ με γνωστή συνάρτηση αυτοσυσχέτισης $C(\mathbf{x}_1,\mathbf{x}_2)$ γράφεται ως:

$$\hat{a}(\mathbf{x},\theta)=a_0(\mathbf{x})+\sum_{j=1}^{M}\sqrt{\lambda_j}\xi_j(\theta)\varphi_j(\mathbf{x}) \tag{13}$$

όπου, λ_j οι ιδιοτιμές της C, $\varphi_j(\mathbf{x})$ τα αντίστοιχα ιδιοδιανύσματα και $\xi_j(\theta)$ είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την κανονική κατανομή με μέση τιμή 0 και τυπική απόκλιση 1.

Στη συνέχεια γίνεται η θεώρηση ότι το άγνωστο στοχαστικό πεδίο της απόκρισης ενός συστήματος μπορεί να εκφραστεί μέσω ενός αναπτύγματος πολυωνυμικού χάους με πεπερασμένους όρους:

$$\mathbf{U}(\theta)=\sum_{k=0}^{P-1}\mathbf{U}_k\Psi_k(\boldsymbol{\xi}) \tag{14}$$

όπου $\Psi_k(\boldsymbol{\xi})$ είναι μια οικογένεια πολυδιάστατων πολυωνύμων Hermite. Στο ανάπτυγμα περιέχονται P όροι, όπου

$$P = \sum_{q=0}^{p} \binom{M+q-1}{q} \tag{15}$$

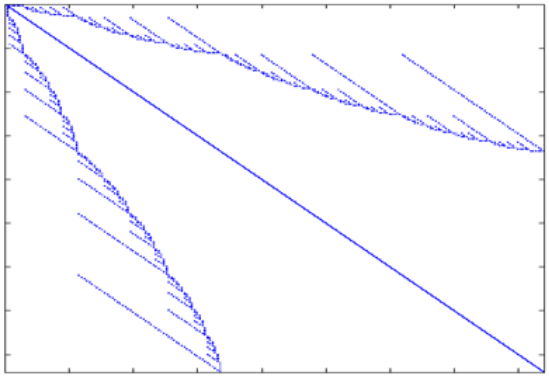
με p να είναι ο μέγιστος βαθμός των πολυωνύμων στο ανάπτυγμα. Οι όροι αυτοί κατασκευάζουν μια ορθοκανονική βάση ενός υπόχωρου του $\mathscr{L}^2(\mathbb{R})$ και αναζητούμε τους άγνωστους συντελεστές $\mathbf{U}_k$ της απόκρισης ως προς τη βάση. Υπό αυτή την έννοια, και αυτή η μέθοδος ανήκει στην κατηγορία των τεχνικών υποκατάστατης μοντελοποίησης, καθώς θεωρούμε ότι το (δυνητικά) απειροδιάστατο στοχαστικό πεδίο της απόκρισης προσεγγίζεται επαρκώς από το πεπερασμένο ανάπτυγμα πολυωνυμικού χάους. Το ανάπτυγμα της απόκρισης εισάγεται στις εξισώσεις του συστήματος και αναζητούμε τους συντελεστές του αναπτύγματος που θα ελαχιστοποιούν το σφάλμα:

$$\epsilon_{M,P} = \sum_{j=0}^{M} \sum_{k=0}^{P-1} \boldsymbol{K}_j \mathbf{U}_k \xi_j(\theta) \Psi_k(\theta) - \mathbf{F} \tag{16}$$

Αυτό επιτυγχάνεται εφαρμόζοντας τη μέθοδο των προβολών Galerkin, όπου αναζητούμε τις τιμές των συντελεστών για τις οποίες το σφάλμα είναι ορθογώνιο στα στοιχεία της βάσης:

$$\mathbb{E}[\epsilon_{M,P} \cdot \Psi_l] = 0 \quad l = 0, 1, ..., P-1 \tag{17}$$

Με αυτό τον τρόπο, θα προκύψει ένα διευρυμένο σύστημα $N \times P$ εξισώσεων όπου N ο αριθμός των βαθμών ελευθερίας του ντετερμινιστικού προβλήματος και P ο αριθμός τον όρων στο ανάπτυγμα πολυωνυμικού χάους. Στο σχήμα 6 φαίνεται η μορφή του αντίστοιχου διευρυμένου μητρώου του συστήματος στην περίπτωση που έχουμε 6 όρους στο ανάπτυγμα Karhunen-Loéve και μέχρι τέταρτης τάξης πολυώνυμα στο ανάπτυγμα πολυωνυμικού χάους. Στο σχήμα αυτό κάθε κουκκίδα αντιστοιχεί σε ένα μητρώο $N \times N$, ίσης διάστασης δηλαδή με το ντετερμινιστικό μητρώο δυσκαμψίας του συστήματος.



Σχήμα 6: Διευρυμένο μητρώο δυσκαμψίας για $M = 6, p = 4$

Η επίλυση του διευρυμένου συστήματος εξισώσεων θα προσδιορίσει τους συντελεστές του αναπτύγματος της άγνωστης στοχαστικής απόκρισης.

Η εφαρμογή αυτής της μεθόδου είναι άμεση για γραμμικά προβλήματα, ωστόσο σε προβλήματα με μη-γραμμικότητα υλικού ή γεωμετρίας απαιτείται ειδική διατύπωση των αντίστοιχων εξισώσεων. Στην παρούσα διατριβή διατυπώθηκε το μαθηματικό και υπολογιστικό πλαίσιο για την εφαρμογή της μεθόδου σε προβλήματα ραβδωτών φορέων με μη-γραμμικότητα γεωμετρίας. Σε τέτοιου τύπου προβλήματα προτείνεται η έκφραση των άγνωστων βηματικών αποκρίσεων, όπως αυτές προκύπτουν στα πλαίσια των μη-γραμμικών αλγορίθμων επίλυσης Newton-Raphson, σε ανάπτυγμα πολυωνυμικού χάους. Για παράδειγμα, ας θεωρήσουμε τη γραμμικοποιημένη μορφή του μη-γραμμικού συστήματος των εξισώσεων ισορροπίας:

$$\mathbf{K}_t^{i-1}\delta\mathbf{u}^i = \mathbf{R}^i - \mathbf{f}^{i-1} \tag{18}$$

όπου, $\mathbf{K}_t^{i-1}$, $\mathbf{R}^i$ και $\mathbf{f}^{i-1}$ είναι αντίστοιχα το εφαπτομενικό μητρώο στιβαρότητας, το διάνυσμα της εξωτερικής φόρτισης και το διάνυσμα εσωτερικών δράσεων. Ο άνω δείκτης i αναφέρεται στην i επανάληψη του αλγόριθμου επίλυσης και ο όρος $\delta\mathbf{u}^i$ είναι οι βηματικές μετατοπίσεις της επανάληψης i, οι οποίες αθροίζονται σε αυτές του προηγούμενου βήματος

$\mathbf{u}^{i-1}$ ώστε να προκύψει η συνολική μετατόπιση:

$$\mathbf{u}^i = \mathbf{u}^{i-1} + \delta\mathbf{u}^i \tag{19}$$

Έπειτα, οι βηματικές αποκρίσεις $\delta\mathbf{u}^i$ αναπτύσσονται σε σειρά πολυωνυμικού χάους και μέσω του επαναληπτικού αλγόριθμου υπολογίζονται οι αντίστοιχοι συντελεστές του αναπτύγματος. Δηλαδή:

$$\mathbf{K}_t^{i-1} \sum_{k=0}^{P-1} \delta\boldsymbol{\alpha}_k^i \Psi_k(\boldsymbol{\xi}) = \mathbf{R}^i - \mathbf{f}^{i-1} \tag{20}$$

με,

$$\delta\mathbf{u}^i = \sum_{k=0}^{P-1} \delta\boldsymbol{\alpha}_k^i \Psi_k(\boldsymbol{\xi}) \tag{21}$$

Το ανάπτυγμα της απόκρισης $\mathbf{u}$ οδηγεί στην σχέση:

$$\mathbf{u}^i = \sum_{k=0}^{P-1} \boldsymbol{\alpha}_k^i \Psi_k(\boldsymbol{\xi}) \tag{22}$$

όπου

$$\boldsymbol{\alpha}_k^i = \boldsymbol{\alpha}_k^{i-1} + \delta\boldsymbol{\alpha}_k^i \qquad k = 0, 1, ..., P-1 \tag{23}$$

Η προβολή του σφάλματος της προσέγγισης

$$\epsilon_P = \Big(\sum_{k=0}^{P-1} \mathbf{K}_t^{i-1} \delta\boldsymbol{\alpha}_k^i \Psi_k \Big) - \mathbf{R}^i + \mathbf{f}^{i-1} \tag{24}$$

στις συναρτήσεις της βάσης και η ελαχίστοποίηση αυτού θα καταλήξει στο ακόλουθο διευρυμένο επαναληπτικό σχήμα:

$$\boldsymbol{\mathcal{K}}_t^{i-1} \delta\boldsymbol{\alpha}^i = \boldsymbol{\mathcal{R}}^i - \boldsymbol{\mathcal{f}}^{i-1} \tag{25}$$

όπου $\boldsymbol{\mathcal{K}}_t^{i-1}$, $\boldsymbol{\mathcal{R}}^i$ και $\boldsymbol{\mathcal{f}}^{i-1}$ είναι το διευρυμένο εφαπτομενικό μητρώο στιβαρότητας και τα διευρυμένα διανύσματα εξωτερικής φόρτισης και εσωτερικών δράσεων κατά την i-επανάληψη, αντίστοιχα. Ο υπολογισμός τους γίνεται μέσω των εσωτερικών γινομένων των $\mathbf{K}_t^{i-1}$, $\mathbf{R}^i$, $\mathbf{f}^{i-1}$ με τις συναρτήσεις Ψ_k. Ενδεικτικά, παρουσιάζεται η μορφή του εφαπτομενικού μητρώου δυσκαμψίας

$$\boldsymbol{\mathcal{K}}_t^e = \begin{bmatrix} \boxed{\mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_0]} & \mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_1] & \cdots & \mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_{P-1}] \\ \mathbb{E}[\mathbf{K}_t^e \Psi_1 \Psi_0] & \mathbb{E}[\mathbf{K}_t^e \Psi_1 \Psi_1] & \cdots & \mathbb{E}[\mathbf{K}_t^e \Psi_1 \Psi_{P-1}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[\mathbf{K}_t^e \Psi_{P-1} \Psi_0] & \mathbb{E}[\mathbf{K}_t^e \Psi_{P-1} \Psi_1] & \cdots & \mathbb{E}[\mathbf{K}_t^e \Psi_{P-1} \Psi_{P-1}] \end{bmatrix} \tag{26}$$

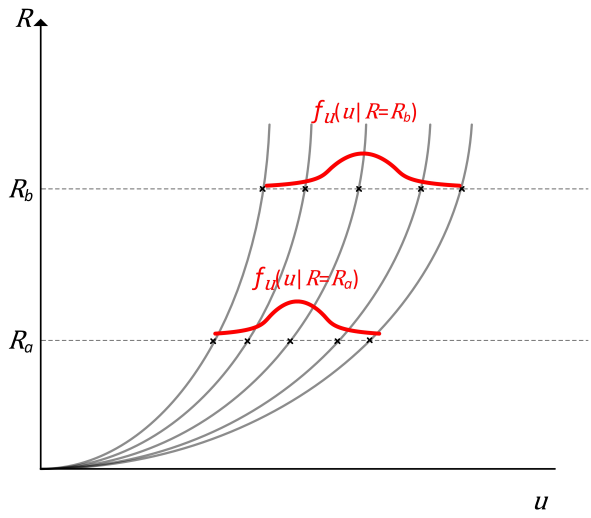
με

$$\begin{aligned} &\mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_0] = \\ &= \begin{bmatrix} \int_{-\infty}^{\infty} \mathbf{K}_t^e(1,1) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} & \cdots & \int_{-\infty}^{\infty} \mathbf{K}_t^e(1,n) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} \\ \vdots & \ddots & \vdots \\ \int_{-\infty}^{\infty} \mathbf{K}_t^e(n,1) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} & \cdots & \int_{-\infty}^{\infty} \mathbf{K}_t^e(n,n) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} \end{bmatrix} \end{aligned} \tag{27}$$

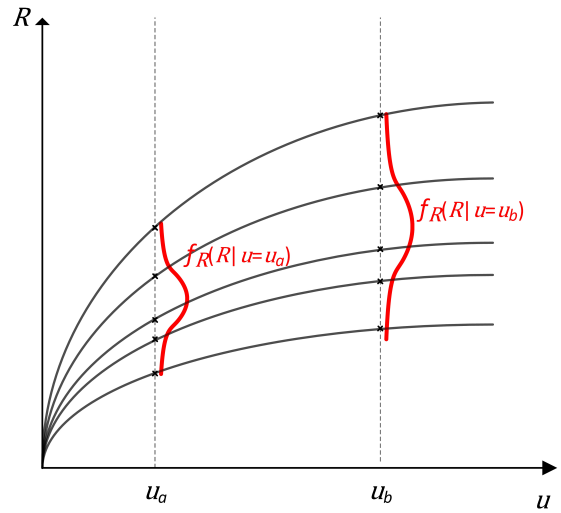
Για τον προσδιορισμό των στοιχείων του μη-επαυξημένου μητρώου δυσκαμψίας $\mathbf{K}_t$ χρησιμοποιείται η συστροφική διατύπωση του στοιχείου δοκού (co-rotational beam formulation). Όμως, η εξάρτηση των στοιχείων του μητρώου από τις στοχαστικές αποκρίσεις σε κάθε βήμα οδηγεί σε ορισμένες ιδιαίτερα πολύπλοκες εκφράσεις προς ολοκλήρωση κατά τη μόρφωση του $\boldsymbol{\mathcal{K}}_t$. Για το λόγο αυτό αναπτύσσεται μια μέθοδος επαναπροβολής των πολύπλοκων εκφράσεων σε απλούστερα αναπτύγματα πολυωνυμικού χάους ούτως ώστε να απλοποιηθεί η μορφή των ολοκληρωμάτων και να δίνονται πλέον από κλειστές μαθηματικές σχέσεις.

Για την ολοκλήρωση της προτεινόμενης μεθοδολογίας απαιτείται επίσης και η κατάλ-

ληλη επέκταση των αλγορίθμων ελέγχου φορτίου και ελέγχου μετατόπισης στα πλαίσια της μεθόδου των φασματικών στοχαστικών πεπερασμένων στοιχείων. Η πρώτη περίπτωση μας επιτρέπει να υπολογίσουμε τα στοχαστικά χαρακτηριστικά της απόκρισης για δεδομένη τιμή του φορτίου σε προβλήματα που παρουσιάζουν αύξηση της δυσκαμψίας «geometric stiffening», όπως φαίνεται στο σχήμα 7. Αντίστοιχα, ο στοχαστικός αλγόριθμος ελέγχου μετατόπισης μας επιτρέπει να υπολογίσουμε τα στοχαστικά χαρακτηριστικά του φορτίου για δεδομένη τιμή της απόκρισης, όπως φαίνεται στο σχήμα 8.



Σχήμα 7: Πιθανοτικά χαρακτηριστικά της μετατόπισης για δεδομένη τιμή του φορτίου



Σχήμα 8: Πιθανοτικά χαρακτηριστικά του φορτίου για δεδομένη τιμή της μετατόπισης

Στο τελευταίο στάδιο της διατριβής διερευνάται η χρήση τεχνικών μηχανικής μάθησης για τη μείωση του κόστους της μεθόδου Monte Carlo. Συγκεκριμένα, προτείνεται η αντικατάσταση των χρονοβόρων προσομοιωμάτων πεπερασμένων στοιχείων με υποκατάστατα μοντέλα που θα παρουσιάζουν την ίδια συμπεριφορά, αλλά με ελάχιστο κόστος επίλυσης. Προς αυτή την κατεύθυνση, μελετάται η εφαρμογή ενός αλγορίθμου μηχανικής μάθησης σε πολλαπλότητες, του αλγορίθμου των χαρτών διάχυσης (diffusion maps), με στόχο να μειωθεί η διαστατικότητα (dimensionality) του χώρου λύσεων ενός παραμετρικού προβλήματος πεπερασμένων στοιχείων. Ο αλγόριθμος αυτός βασίζεται στην υπόθεση ότι οι λύσεις ενός παραμετροποιημένου συστήματος, όπως είναι ένα στοχαστικό σύστημα, αποτελούν σημεία μιας πολλαπλότητας εμφυτευμένης σε έναν Ευκλείδειο χώρο αρκετά μεγαλύτερης διάστασης (όση οι βαθμοί ελευθερίας του προβλήματος πεπερασμένων στοιχείων). Τότε η εφαρμογή του αλγορίθμου των χαρτών διάχυσης σε ένα μικρό αριθμό λύσεων του μοντέλου θα μας επιτρέψει να αναγνωρίσουμε την εσωτερική διαστατικότητα της πολλαπλότητας, στην οποία κείτονται οι λύσεις του συστήματος. Η διάσταση αυτή εξαρτάται από τον αριθμό των τυχαίων μεταβλητών στο σύστημα και ενδέχεται να είναι τάξεις μεγέθους μικρότερη από τον

Ευκλείδειο χώρο στον οποίο εμφυτεύεται.

Η μεθοδολογία για την εφαρμογή του αλγορίθμου των χαρτών διάχυσης συνίσταται στα ακόλουθα βήματα:

1. Κατασκευή ενός αρχικού συνόλου δεδομένων $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_N] \in \mathbb{R}^{d \times N}$ αποτελούμενου από N λύσεις του d-διάστατου προσομοιώματος πεπερασμένων στοιχείων για N διαφορετικές τιμές των τυχαίων παραμέτρων.

2. Χρήση ενός μέτρου για την προσδιορισμό της συνάφειας μεταξύ των σημείων του συνόλου δεδομένων, όπως ο Γκαουσιανός πυρήνας: $k(\boldsymbol{u}_i, \boldsymbol{u}_j) = exp\left(-\|\boldsymbol{u}_i - \boldsymbol{u}_j\|^2/\epsilon^2\right)$

3. Κατασκευή του $N \times N$ μητρώου συνάφειας $L_{ij} = k(\boldsymbol{u}_i, \boldsymbol{u}_j)$ και του αντίστοιχου διαγώνιου μητρώου $D_{ii} = \sum_{q=1}^{N} L_{iq}$.

4. Κανονικοποίηση του $\boldsymbol{L}$ με μια παράμετρο α: $\boldsymbol{L}^{(\alpha)} = \boldsymbol{D}^{-\alpha}\boldsymbol{L}\boldsymbol{D}^{-\alpha}$. Στην παρούσα μεθοδολογία η παράμετρος α επιλέγεται ίση με 1 ώστε ο αλγόριθμος να αναγνωρίζει τη διαφορική γεωμετρία της πολλαπλότητας στην οποία κείτονται τα σημεία ανεξάρτητα από το μέτρο της πιθανότητας.

5. Στη συνέχεια κατασκευάζεται το $N \times N$ μητρώο μετάβασης πιθανότητας $\boldsymbol{P} = (\boldsymbol{D}^{\alpha})^{-1}\boldsymbol{L}^{(\alpha)}$. Τα στοιχεία του μητρώου αυτού εκφράζουν την πιθανότητα αν βρισκόμαστε στο σημείο του χώρου $\boldsymbol{u}_i$ στο βήμα τ να βρεθούμε στο σημείο $\boldsymbol{u}_j$ στο επόμενο βήμα $\tau + 1$. Ο αλγόριθμος των χαρτών διάχυσης ερμηνεύει αυτή την πιθανότητα ως μια έννοια άπόστασης΄ μεταξύ των σημείων στην πολλαπλότητα.

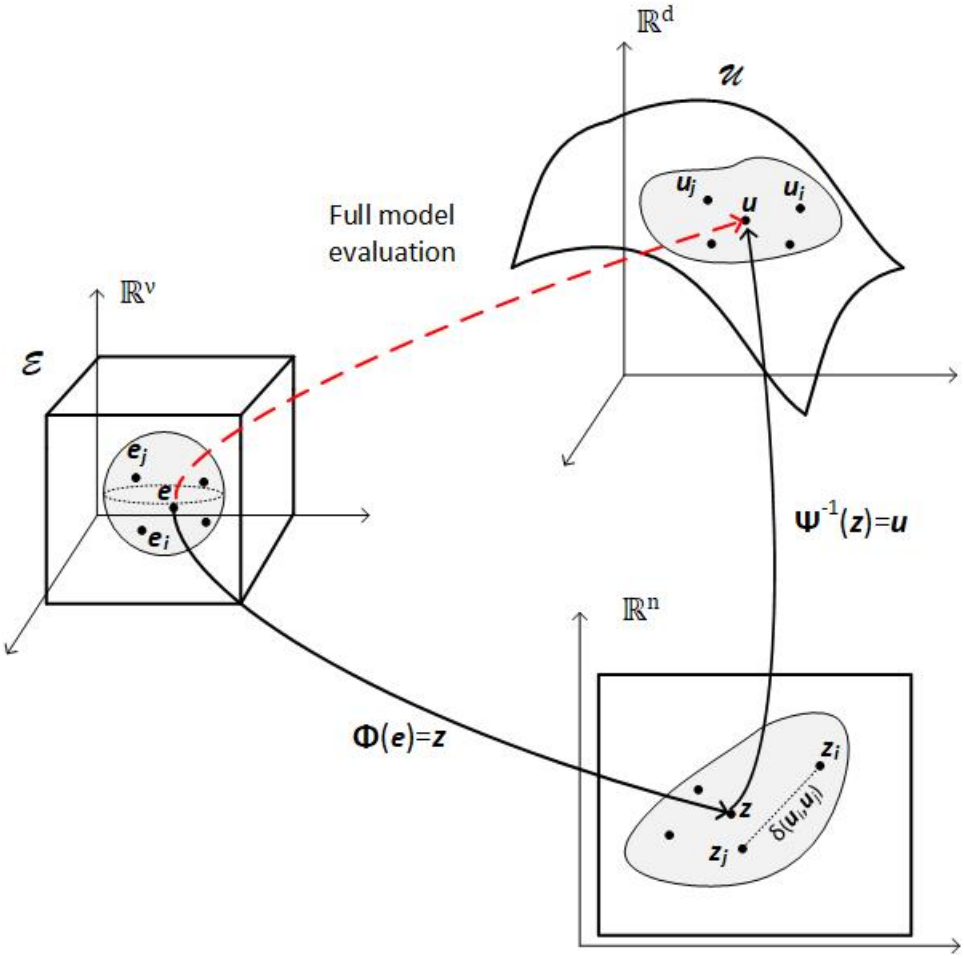
6. Στη συνέχεια, υπολογίζονται οι ιδιοτιμές λ_j και τα αντίστοιχα ιδιοδιανύσματα $\boldsymbol{\Psi}_j$ του $\boldsymbol{P}$ και από αυτά κρατάμε τα n σημαντικότερα, όπου στην πράξη $n << d$.

7. Από τις ιδιοτιμές και τα ιδιοδιανύσματα που υπολογίστηκαν προηγουμένως κατασκευάζεται μια απεικόνιση μειωμένης διάστασης των σημείων του αρχικού συνόλου δεδομένων στον χώρο των χαρτών διάχυσης. Ειδικότερα, σε κάθε $\boldsymbol{u}_i \in \mathbb{R}^d$ αντιστοιχίζεται το $\boldsymbol{z}_i \in \mathbb{R}^n$ μέσω της σχέσης:

$$\boldsymbol{z}_i = \boldsymbol{\Psi}(\boldsymbol{u}_i) \triangleq \left[\lambda_1\psi_i^1, \lambda_2\psi_i^2, ..., \lambda_n\psi_i^n\right] \tag{28}$$

Έχοντας καταφέρει να μειώσουμε τη διαστατικότητα του προβλήματος, στη συνέχεια, κατασκευάζεται κατάλληλος αλγόριθμος πολυωνυμικής παρεμβολής για την απεικόνιση νέων σημείων $\boldsymbol{e} \in \mathbb{R}^{\nu}$ από τον δειγματόχωρο των τυχαίων μεταβλητών του προβλήματος στον χώρο των χαρτών διάχυσης. Έπειτα, μέσω μιας παρεμβολής με ακτινικές συναρτήσεις

(radial basis function interpolation) τα σημεία από τον χώρο των χαρτών διάχυσης απεικονίζονται στον χώρο των λύσεων του προβλήματος. Η διαδικασία αυτή παρουσιάζεται στο σχήμα 9, όπου βλέπουμε ότι με την προτεινόμενη μεθοδολογία παρακάμπτεται πλήρως η ανάγκη για την διατύπωση των εξισώσεων ισορροπίας του συστήματος και τη μετέπειτα επίλυσή τους.



Σχήμα 9: Απεικονίσεις μεταξύ του χώρου των παραμέτρων $\mathcal{E} \in \mathbb{R}^\nu$ στον χώρο των χαρτών διάχυσης και από εκεί στην πολλαπλότητα $\mathcal{U}$ που περιέχεται στον περιβάλλοντα Ευκλείδειο χώρο των λύσεων $\mathbb{R}^d$

Για τη βελτίωση της προτεινόμενης μεθοδολογίας αξιοποιείται επιπρόσθετα η μετρική του χώρου των χαρτών διάχυσης για τη δημιουργία ενός σχήματος δειγματοληψίας, το οποίο θα αναγνωρίζει περιοχές του χώρου των λύσεων, στις οποίες υπάρχει μεγάλη μεταβλητότητα των λύσεων. Η απόσταση μεταξύ δύο σημείων $\boldsymbol{u}_i$ και $\boldsymbol{u}_j$ στον χώρο των χαρτών διάχυσης δίνεται από τη σχέση:

$$\delta\left(\boldsymbol{u}_i, \boldsymbol{u}_j\right) = \|\Psi(\boldsymbol{u}_i) - \Psi(\boldsymbol{u}_j)\| = \|\boldsymbol{z}_i - \boldsymbol{z}_j\| \tag{29}$$

Η μετρική αυτή δίνει μια πιο ακριβή εκτίμηση της πραγματικής (γεωδαισιακής) απόστασης των σημείων πάνω στην πολλαπλότητα. Ο προτεινόμενος αλγόριθμος δειγματοληψίας υπολογίζει για κάθε σημείο $\boldsymbol{u}_i$ του αρχικού συνόλου δεδομένων την απόσταση του με τους κ-γείτονες του στον χώρο των χαρτών διάχυσης. Αν αυτή η απόσταση είναι μεγαλύτερη από μια προκαθορισμένη τιμή για κάποιο $\boldsymbol{u}_j$, αυτό σημαίνει ότι η λύση υφίσταται μεγάλες μεταβολές στην περιοχή του $\boldsymbol{u}_i$ και απαιτούνται επιπλέον λύσεις στην περιοχή αυτή για την ακριβή αναγνώριση της γεωμετρίας της πολλαπλότητας. Έτσι, επιλέγεται ένα νέο διάνυσμα τιμών των παραμέτρων $\boldsymbol{e}_{new}$ ως ο κυρτός συνδυασμός των διανυσμάτων των τιμών των παραμέτρων που οδήγησαν στις προαναφερθείσες λύσεις, δηλαδή:

$$\boldsymbol{e}_{new} = c\boldsymbol{e}_i + (1-c)\boldsymbol{e}_j \tag{30}$$

Εν συνεχεία, εκτελείται μια επίλυση του πλήρους μοντέλου των πεπερασμένων στοιχείων και υπολογίζεται το νέο διάνυσμα λύσεων $\boldsymbol{u}_{new}$. Εφαρμόζοντας τον μεταχηματισμό Nyström βρίσκουμε την εικόνα του στον χώρο των χαρτών διάχυσης και ελέγχουμε αν πλέον ικανοποιείται το κριτήριο για τις αποστάσεις. Η διαδικασία αυτή επαναλαμβάνεται έως ότου όλες οι 'γειτονιές' των σημείων είναι επαρκώς σκιαγραφημένες.

Συνοψίζοντας, τα βήματα της προτεινόμενης μεθοδολογίας για την ανάλυση στοχαστικών προβλημάτων είναι τα ακόλουθα:

1. Εκτέλεση N Monte Carlo επιλύσεων και κατασκευή του $d \times N$ μητρώου των λύσεων: $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_N]$.

2. Εφαρμογή του αλγορίθμου των χαρτών διάχυσης στο σύνολο των λύσεων $\boldsymbol{U}$ και εύρεση της απεικόνισης μειωμένης διάστασης $\boldsymbol{\Psi}$.

3. Αξιοποίηση της μετρικής στον χώρο των χαρτών διάχυσης για τον εντοπισμό περιοχών στην πολλαπλότητα όπου απαιτούνται περισσότερα σημεία.

4. Προσδιορισμός των πρόσθετων τιμών $\boldsymbol{e}$ που απαιτούνται και εκτέλεση των αντίστοιχων επιλύσεων του πλήρους μοντέλου. Η κάθε μια λύση απεικονίζεται στον χώρο των χαρτών διάχυσης μέσω του μετασχηματισμού Nyström.

5. Πλήρης ανάλυση Monte Carlo: Για κάθε νέο $\boldsymbol{e}$ χρησιμοποιείται η πολυωνυμική παρεμβολή για την εύρεση της εικόνας του στον χώρο των χαρτών διάχυσης και έπειτα η παρεμβολή με ακτινικές συναρτήσεις για τον προσδιορισμό της απόκρισης.

Οι αριθμητικές εφαρμογές που μελετήθηκαν έδειξαν ότι η προτεινόμενη προσέγγιση μπορεί να μειώσει δραματικά το υπολογιστικό κόστος της μεθόδου Monte Carlo χωρίς σημαντικές απώλειες στην ακρίβεια.

Contents

List of Figures

0
Introduction

0.1 Motivation

The first step towards studying the behavior of a mechanical system is to make accurate mathematical idealizations of its properties. This implies that the parameters of the system (system geometry, materials, etc.), the response parameters (displacement, stress, strain, etc.) and the relationships between these quantities, have to be mathematically expressed in such a way that the mathematical model is in good agreement with the physical system. In this direction, extensive research has been carried out in order to improve the structural models and the constitutive equations that describe the system. At the same time, the progress of computer science and the corresponding increase in computational power led to the development of powerful numerical tools capable of approximating the solutions of any system under consideration. The finite element method is undoubtedly the most popular and widely-used approach nowadays for the solution of such problems.

Despite the improvement in the constitutive models and the computational tools, in order to achieve a more reliable representation of the physical system, randomness in the system parameters and loading conditions, as well as the uncertainty in their estimation and their spatial variability must also be taken into account. Considering them to be of deterministic nature using average characteristics was a simplification

dictated mainly by the limitations of computational capacity and resulted in crude approximations of the system's behavior. Over the past decades the so called stochastic finite element method (SFEM) was developed, in the effort to incorporate randomness in the modeling of engineering systems. SFEM is an extension of the classical deterministic FE approach to the stochastic framework, and it involves a variety of methods for treating the response variability calculation. The most eminent method in this category is the well-known Monte Carlo simulation (MCS) [143]. This method is without a doubt the most versatile and easily implemented stochastic analysis method, which can be employed in any type of problem. In practice, due to its generality, MCS serves as a means of verifying the accuracy of other stochastic procedures. However, it possesses the drawback that in order to achieve high levels of accuracy, a large number of repeated simulations is required, which results in significant computational costs. One approach to alleviate this often intractable computational burden is through customized solution strategies for the repeated solution of near-by systems of algebraic equations required in MCS [149, 113, 111]. Moreover, variants of the direct MCS such as the importance sampling method [138], the subset simulation method [9] and the line sampling method [83]) can be applied in certain types of problems in order to reduce the number of model runs required.

From another perspective, an effective approach to bypass the direct MCS is through a process known as surrogate modeling. A surrogate model (response surface model, metamodel or emulator) is an approximation of the original computational model, which is typically built from a limited set of runs of the original model and/or some regularity assumptions for its functional shape. For the purposes of stochastic analysis, particularly popular surrogate modeling techniques are the stochastic collocation technique [10], that employs generalized polynomial chaos interpolating functions, and the Gaussian process regression (Kriging) [19], that fits a Gaussian stochastic process to the solution. Another surrogate modeling approach to reduce the cost of MCS is to utilize machine learning methods in order to replace the detailed model with simpler mathematical functions, which will be able to emulate the system's input-output relation with a controlled loss of accuracy. Among the most widely used machine learning methods for this task are the Neural networks [66] and deep neural networks [136] with a vast number of applications ranging from computer vision [67, 74, 80] to engineering problems [71, 72, 114, 112]. Also, reduced order modeling (ROM) methods have been successfully employed in order to reduce the computational cost of single model run in the context of MCS. These

methods aim at replacing the full-order system by one of significantly smaller dimension in order to speed-up the computations, while maintaining a desired level of accuracy. The basic idea is to project the governing equations of the physical system onto a subspace of reduced dimension. The most commonly preferred approach to identify this subspace is the Proper Orthogonal Decomposition (POD) [124, 3, 53] method, where the basic idea is that the subspace spanned by the eigenvectors of an ensemble of system solutions constitutes a basis of the total solution space and can accurately describe the system.

An important stochastic method that has received great attention in the last decade leading to elegant Galerkin-based formulations is the spectral stochastic finite element method (SSFEM) [57, 160, 58, 59, 60, 151, 150]. This method relies on the assumption that the unknown stochastic response of a system can be expressed by a Polynomial Chaos (PC) series expansion with finite terms. The PC expansion is a surrogate modeling technique, which simplifies the model by approximating its functional shape with a series of multivariate polynomials that are orthogonal with respect to the distributions of the input random variables. Under this assumption, the Galerkin projection technique can be employed for the estimation of the optimal coefficients of the PC basis functions in the expansion. This approach requires the solution of an augmented system of algebraic equations with respect to the corresponding deterministic one, which can become quite challenging due to the increased memory and computational resources required. This increase in the dimensions of the coefficient matrix to be inverted can be dramatic in certain cases, such as those involving large variations and/or non-Gaussian fields and may lead to a disproportionate increase of the computational cost, especially in problems involving a large number of degrees of freedom (dofs). This implication renders in such cases the SSFEM intrusive solution a more expensive method, compared to non-intrusive brute force Monte Carlo simulation. In a recent publication [149] the numerical performance of intrusive versus non-intrusive methods is critically assessed on the basis of up-to-date solvers and solution strategies, specifically tailored to the needs of these analyses.

As an alternative to the above methodologies, a new stochastic methodology has been recently proposed by Li and Chen [92], namely the probability density evolution method (PDEM), which has been successfully employed in a variety of stochastic problems [89, 91, 93, 118]. The foundation of this method lies on the principle of preservation of probability and, in a sense, introduces a new framework for solving the conservation equation that governs the flow of probability in the system under consideration. In

general, the principle of preservation of probability is expressed through the celebrated Liouville and Fokker-Planck partial differential equations, which estimate the joint pdf of the system's stochastic response at every time instant. It becomes evident that these equations carry the complete probabilistic information about the system, but despite their significance, their applicability to general stochastic finite element systems is limited to problems involving only a few degrees of freedom. This is due to the fact that the number of terms they involve depends on the number of degrees of freedom in the system. Similar type of equations have been studied by Cho et al., Venturi et al. and Wang et al. in computational fluid dynamics with uncertain parameters [158, 156, 155] as well as stochastic oscillators [30], but they also suffer from dimensionality issues.

The distinction between the previously mentioned approaches and PDEM is that in PDEM the probability conservation equation is re-derived from the random event perspective [27], giving rise to a new type of probability conservation equations, called the generalized density evolution equations (GDEE). Through these equations, the multidimensional joint pdf of the system's response is decoupled into a set of one-dimensional pdfs, corresponding to each dof. The random event perspective refers to the partitioning of the probability space formed by the system's random variables into disjoint sets, with each set being characterized by a representative set of values of the random parameters. Then, for each of the disjoint sets an advection type pde is defined and the solution of the GDEE is obtained from the superimposition of the solutions of these pde's, which are significantly easier to obtain. It should be mentioned that each of these pde's contains a flux term given by the time derivative of the response along the dof under consideration. This term is evaluated through a deterministic analysis of the system for the corresponding representative parameter values. Therefore, the number of deterministic analyses that PDEM requires is equal to the number of the disjoint sets in the partition and it can be remarkably smaller than those needed in the direct MCS. In essence, the surrogate modeling technique employed in this method that renders it more efficient than MCS is the random event perspective and the associated assumptions made to solve the resulting GDEEs. With this approach, the intractable Liouville and Fokker-Planck equations can be accurately approximated via a series of one-dimensional pde's, that are easier to compute.

As stated, the great advantage of PDEM is that it can achieve high levels of accuracy with the requirement of a relatively small number of repeated simulations. However, this method suffers from two drawbacks. The first drawback refers to the fact that

this method is mostly applied to dynamic problems, where the pdf of the quantity of interest evolves with respect to time. The second drawback comes from the numerical solution of the advection pde's. Despite their simple form, in most applications they involve discontinuous 'shock' solutions which are notoriously difficult for numerical schemes to handle. In the research presented in this dissertation both these issues are addressed. More specifically, in [105] a reformulation of the classic PDEM is proposed, which expresses the principle of preservation of probability in space instead of time, thus extending the applicability of the method to static problems. Also, the use of two numerical schemes, namely the time-marching Discontinuous Galerkin scheme endowed with shock capturing terms and the Streamline Upwind/ Petrov Galerkin (SUPG) finite element scheme were investigated for the solution of the advection type pde's. The analysis results indicated that the SUPG method was more accurate and more efficient than the commonly preferred Lax-Wendroff finite difference scheme.

A major challenge posed in computational stochastic mechanics is the extention of stochastic methodologies to nonlinear problems. Undoubtedly, linear analysis in structural mechanics is a well-established approach which offers the engineer a quick and cost-effective estimation of the response of a structural system. However, in many cases this approach fails to produce an accurate representation of the actual structural behavior and the reason for this lies in the effect nonlinearities have on the structure's behavior and, ultimately, its capacity. They can be classified into two broad categories, namely, material nonlinearity and geometric nonlinearity. The former refers to the nonlinear relationship between stress and strain for an actual material and is introduced in the computational model by updating the system's constitutive matrix at each analysis step. The latter refers to changes in geometry and/or boundary conditions and the effect these changes have on the system response, which in many cases can be quite significant and may lead to unexpected results (e.g. buckling, P-Δ effects). The purpose of nonlinear analysis of structures is to trace the equilibrium path and in classical nonlinear FEM, this is achieved by taking the equilibrium in the deformed configuration leading to a system of nonlinear equations that are solved in the context of a Newton-Raphson incremental-iterative scheme. An important issue that arises at this point is the selection of a suitable path-following technique able to predict the full range of the load-displacement curve. In problems where the load displacement curve exhibits snap-through's, a load control algorithm [166] fails to converge to the exact solution and a displacement control algorithm [15] is commonly applied, while in problems with snap-back's the arc length

method [43, 139] is preferred since it is the only capable of achieving convergence. In the general premise of nonlinear analysis, efficient methods to determine the collapse load of given structural model can also be sought in the field of limit analysis [97, 98, 154, 148].

Needless to say, the existense of imperfections in geometry, material and section properties plays a major role in the stucture's inelastic behavior [101, 133, 104]. In terms of modeling these imperfections, it is only reasonable to assume that they are of stochastic nature. As a consequence, the extension of nonlinear analysis to stochastic systems is essential for more realistic and accurate representations, especially for the purposes of system reliability analysis. Even though the issue of structural nonlinearities is thoroughly addressed for deterministic systems, this is not the case for stochastic problems. An attempt towards this direction can be found in [108, 133], where the nonlinear behavior of shell structures with random material and geometric imperfections is investigated by employing the Monte Carlo method. Similarly, the Monte Carlo method was also employed in [110], where material and geometrically nonlinear finite element analyses were performed to evaluate the stochastic buckling load of I-section portal frames, and in [120] to estimate the reliability of MDOF-systems. An overview of the available approaches to quantify the effect of material and geometric imperfections in nonlinear analysis can be found in [48]. Reviewing the available literature one quickly comes to the conclusion that the Monte Carlo simulation technique is the most widely accepted approach to include uncertainty in the nonlinear framework. Other methods that address problems with nonlinearities in the context of stochastic analysis and reliability include the response-surface technique [52, 12, 13], the Probability Density Evolution Method [25, 94, 28, 163], the β-unzipping method [153], bounding techniques [62] and analytical stochastic response analysis [64]. Stochastic buckling analysis has also received great attention, since the existence of random imperfections leads to a wide scattering of the buckling loads and the approaches to estimate their variability include either the assumption of imperfections with variable amplitude in the form of the critical eigenmode for the perfect structure [132, 76], or the more sophisticated representation of the imperfections as random fields [134, 7, 150, 42, 77].

In this dissertation, two novel formulations of PDEM are proposed [78], which are customized for the needs of stochastic nonlinear analysis. More specifically, the first formulation aims at evaluating the cumulative distribution function of the critical load of a structure, while the second formulation aims at capturing the probabilistic characteristics of the full load displacement curve in the context of Newton-Raphson

incremental-iterative schemes. The main advantage of the proposed approaches is that they allow for a quantification of the effect of uncertainties on the structural capacity, with only a small number of deterministic analysis compared to the MCS.

As mentioned, SSFEM has attracted the major interest in the scientific community with significant impact in the recent years. Even though this method has been widely applied to linear systems, its extension to nonlinear problems is marginal. A few works have appeared linking SSFEM to material nonlinearity [5, 6, 56, 129, 140], while in [2, 8] the application of SSFEM to nonlinear continuum elements is addressed. Also, in [116] and [46] efficient programming techniques are proposed in order to tackle the issue of excessive computational cost associated with the application of SSFEM in nonlinear stochastic problems. In this research, a novel mathematical formulation is presented which enables the coupling of SSFEM with the geometrically nonlinear analysis of framed structures [106]. The proposed approach relies on the co-rotational beam formulation and the idea is to project the stochastic part of the incremental displacements, formulated in the framework of a Newton-Raphson solution scheme, to a polynomial chaos expansion basis. Subsequently, the nonlinear solution is obtained by introducing the expansion coefficients in the iterative solution algorithm as unknowns to be evaluated. With this approach, an augmented system of nonlinear equations is produced and the values of the coefficients are obtained through the convergence of the iterative algorithm. Additionally, the nonlinear load and displacement control algorithm are extended to their stochastic counterparts and a technique to drastically reduce the computational cost of this approach is presented.

Recent advances in machine learning and more specifically in manifold learning gave rise to a number of novel methodologies that take into account the geometrical structure of the space that data live on, in order to enhance the quality of the surrogate models. The main assumption in manifold learning is that the data points of a given data set lie on a low-dimensional manifold embedded in an ambient higher-dimensional Euclidean space. The goal is to identify the parameters that describe the manifold and thus estimate its intrinsic dimensionality. In practice, however, the manifold is approximated by a limited number of discrete points, which, in engineering problems, usually correspond to a set of system (detailed model) evaluations. Therefore, inferring its complete structure from this limited information is a difficult endeavor. Manifold learning is assigned to this task and the available approaches can be mainly divided into linear and nonlinear methods. In the general case, linear methods such as the well-known Principal Component Analysis (PCA)

[75], which is also known as Karhunen-Loève Transform or Singular Value Decomposition, and the Multidimensional Scaling algorithm (MDS) [41] are shown to behave poorly. This is due to the fact that they can only see flat Euclidean structures and thus fail to discover the curved and nonlinear structures of the data. On the other hand, nonlinear methods, which include the Kernel PCA [135], the Isomap [152], locally linear embedding (LLE) [131], Hessian eigenmaps [47], Laplacian eigenmaps [16], local tangent space alignment (LTSA) [165] and the Diffusion Maps (DMAPs) [37], are more accurate at capturing the manifold geometry.

In [4] the geometric properties of the Grassmann manifold have been exploited in order to adapt precomputed ROMs to changes in physical and modeling parameters. Also, in [61] the idea was proposed to project the solution of a high-dimensional parametrized system onto the Grassmann manifold and obtain low-cost system solutions by using an interpolation scheme on the Grassmann manifold. From another perspective, the challenge of statistical sampling on manifolds defined by limited data has been addressed in [145], where the Diffusion Maps algorithm was utilized in order to identify the structure of the manifold, where the realizations of a given random vector lie and to effectively provide a low-dimensional data parametrization. Using this parametrization, a reduced-order Itô stochastic differential equation was developed and its solution provided additional realizations, which were statistically consistent with the initial set of realizations.

In the current research a surrogate modeling strategy based on the Diffusion Maps manifold learning algorithm is proposed [79] that can be used to accelerate computations in parametrized physical systems. After performing a small number of Monte Carlo simulations, the corresponding system solutions are collected in a matrix, which constitutes the problem's initial data set. Then, the DMAP algorithm is applied on the data set, in order to obtain a parametrization of the solution manifold. This is achieved by computing the eigenvalues and eigenfunctions of a diffusion operator of the data set. The basic idea is that the basis functions provided by the DMAP algorithm are able to capture more information than the ones provided by other projection methods such as the POD [146]. Another interesting feature of DMAP is that a metric, similar to the Euclidean metric, can be defined in the DMAP space, which gives a better approximation of the 'true' distance between two points on the manifold. This metric structure is further utilized for the development of a sampling refinement methodology that will guarantee an accurate resolution of the solution space. Also, two clustered interpolation

schemes are proposed; the first uses local polynomial interpolation in order to map points from the parameter space to the DMAP space, and the second relies on a radial basis function (RBF) interpolation scheme that will send a point from the DMAP space back to the ambient Euclidean space. The clustering upon which local interpolations are established is achieved by utilizing the DMAP distance. The aforementioned tools are the computational tools required for the construction of a surrogate model that takes into account the DMAP space geometry. The main advantage of this approach is that the detailed finite element model is substituted by simpler mathematical relations that are easy to calculate and thus the need to formulate and solve the governing equations of the system is bypassed. In addition, while most machine learning methods depend to some degree upon user's 'intuition' (e.g. selection of optimal network architectures in neural networks), the proposed methodology is implemented in a straightforward manner, in the sense that all surrogate model parameters are calibrated based on the geometric features of the initial data set.

0.2 Outline

This thesis is organized in 9 chapters, outlined as follows:

Chapter 1 introduces the reader to the basic principles of probability theory, which include the definitions of probability spaces, random variables, distributions and moments. Also, it is demonstrated that the probability space equipped with the expectation operator becomes an inner product vector space.

Chapter 2 introduces stochastic fields together with important properties such as correlation structure, stationarity, ergodicity and power spectral density. The discretization of stochastic processes via two popular series expansions, the Karhunen-Lòve series expansion and the Spectral Represenation, is described in detail.

Chapter 3 provides the basic concepts of the finite element method. Starting from the abstract formulation of the problem, the process of deriving the weak form of the equations of equilibrium is analytically explained and their solution with the finite element method is presented.

Chapter 4 introduces the concept of surrogate modeling and highlights its efficacy in modeling complex physical systems. The implementation aspects of some popular surrogate modeling techniques that were used in this thesis, namely the Polynomial Chaos expansion, the Radial Basis Function approximation, the Nyström extension and

the Diffusion Maps algorithm are illustrated.

Chapter 5 introduces the Stochastic Finite Element method, which is the computational framework required to study engineering systems influenced by randomness in their properties. The basic principles of some of the most emiment methods in this category, that is, the Monte Carlo simulation (MCS), the Spectral Stochastic Finite Element Method (SSFEM) and the Probability Density Evolution Method (PDEM), are described in detail.

Chapter 6 presents a reformulation of the classic PDEM to address the case of general stochastic finite element systems. The application of the Streamline Upwind/Petrov Galerkin (SUPG) method and the Discontinuous Galerkin time-marching scheme, as alternatives to the commonly used Lax-Wendroff finite difference scheme for the solution of the corresponding generalized density evolution equations, is investigated. Numerical demonstrations are presented to highlight the merits of the SUPG method in terms of accuracy and computational cost.

Chpater 7 presents an application of PDEM for the estimation of stochastic limit loads. First the necessary mathematical framework is developed, where a flow of probability with respect to the limit load is introduced. The partial differential equations describing this flow are solved with the method of characteristics and their solution gives the probabilistic characteristics of the structure's limit load. Moreover, the displacement control method is extended in the frame of PDEM in order to produce the stochastic nonlinear path of equilibrium for the structure under investigation. Numerical applications are presented, which showcase the method's advantages with respect to the Monte Carlo simulation.

Chapter 8 deals with the extension of SSFEM to geometrically nonlinear framed structures. In this chapter, the co-rotational beam formulation is extended to its stochastic counterpart and stochastic versions of the displacement and load-control algorithms are developed. This formulation involves the repeated evaluation of high-dimensional integrals of polynomials and is shown to be computationally inefficient. To this end, the idea of re-projecting complex non-polynomial functions onto the lower-dimensional polynomials of the Polynomial Chaos expansion is proposed and it is shown to reduce significantly the computational cost. Numerical examples are presented in the end of the chapter to demonstrate the advantages of the proposed methodology.

Chapter 9 presents a methodology to construct surrogate models of detailed physical systems for the purposes of non-intrusive Monte Carlo simulation. The diffusion maps algorithm, is utilized for the identification of the geometry of a low-dimensional

manifold where the system's solutions lie. The manifold's intrinsic dimensionality, being significantly smaller than the ambient Euclidean space of the solution vectors, allows for the construction of efficient interpolation schemes that are able to map points from the parameter space to the diffusion maps space and then to the solution space. With this approach the need to formulate and solve the governing equations of the system is bypassed, thus leading to significant computational gains. Furthermore, the metric in the DMAP space is further utilized for the development of a sampling scheme in order to accurately delineate the solution manifold. The accuracy and efficiency of the proposed methodology is demonstrated through a set of numerical examples of structural uncertainty quantification.

Lastly, Chapter 10 discusses the conclusions drawn from this research and presents a summary of the contributions.

1
Probability Theory

1.1 Introduction

The emergence of the concept of mathematical probability dates back in the sixteenth century, where Gerolamo Cardano gave an early definition of probability as the ratio of favourable outcomes to the total number of possible outcomes. During the seventeenth and eighteenth centuries, great minds of this era, including Pierre de Fermat, Blaise Pascal, Abraham de Moivre, Adrien-Marie Legendre and Pierre Simon de Laplace contributed to the development of what we nowadays call *probability* [65]. However, it wasn't until 1931, when Andrey Nikolaevich Kolmogorov introduced notions from measure theory to the study of probability and, in essence, laid the modern axiomatic foundations of the theory of probability [82].

In this chapter, the basic principles and fundamental concepts of probability theory are illustrated. This review aims to provide the definitions of concepts such as random events, probability spaces, random variables, distribution functions, moments and conditional probabilities, which are used throughout the text. A detailed study of the field of probability theory is beyond the scope of this work, however, a more demanding reader is referred to [73, 18, 130, 115]

1.2 Probability space

Let us consider an experiment and denote the set of all possible outcomes of the experiment as Θ, called the *sample space*. The sample space is classified into three categories according to its cardinality, that is, *finite*, *countably infinite* (or *denumerable*) and *uncountable*. More specifically, Θ is said to be finite if the number of elements $\theta \in \Theta$ is finite, countably infinite if a bijection (one-to-one correspondence and onto) between every $\theta \in \Theta$ and the infinite set of natural numbers can be established and uncountable if it is not countable. For instance, for the experiment of rolling a die, Θ is finite since $\Theta = \{1, 2, 3, 4, 5, 6\}$. On the other hand, for the experiment of performing stress tests on concrete specimens to estimate their yield stress, the sample space is uncountable.

A non-empty collection of subsets $\mathcal{F}$ of Θ is called a *σ-algebra* on Θ, if it satisfies the following properties:

1. the empty set belongs to $\mathcal{F}$: $\emptyset \in \mathcal{F}$
2. $\mathcal{F}$ is closed under complements: $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $\mathcal{F}$ is closed under countable unions: $A_i \in \mathcal{F},\ i \in I,\ I =$ a countable set$\Rightarrow \cup_{i \in I} A_i \in \mathcal{F}$

Every element of $\mathcal{F}$ is called *event*, or *$\mathcal{F}$-measurable* subset of Θ. The pair $(\Theta, \mathcal{F})$ constitutes a *measurable space*.

Let us also define a real-valued function μ on $\mathcal{F}$ with the following property:

$$\mu\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right) \text{ for } A_i \in \mathcal{F},\ A_i \cap A_j = \emptyset,\ i \neq j \tag{1.1}$$

The function μ is called a *measure* and the triple $(\Theta, \mathcal{F}, \mu)$ a *measure space*. Also, a measure with the property $\mu(\Theta) < \infty$ is called a *finite measure* and the scaled version of this measure, $\mu(A)/\mu(\Theta)$, takes values in the $[0, 1]$ domain. A set function $P : \mathcal{F} \to [0, 1]$ that satisfies the following properties:

1. $P(\Theta) = 1$
2. $P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$ for $A_i \in \mathcal{F},\ A_i \cap A_j = \emptyset,\ i \neq j$

is called a *probability measure* or simply *probability* and the triple $(\Theta, \mathcal{F}, P)$ constitutes the *probability space*.

1.3 Properties of Probability Measure

In this section some properties of the probability measure P, which are particularly useful for applications, are provided.

- $P(A) \leq P(B),\ A \subseteq B,\ A, B \in \mathcal{F}$
- $P(A) = 1 - P(A^c),\ A \in \mathcal{F}$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B), \quad A, B \in \mathcal{F}$
- The above equation generalizes to the union of multiple events as

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i=2}^{n} \sum_{j=1}^{i-1} P(A_i \cap A_j)$$
$$+ \sum_{i=3}^{n} \sum_{j=2}^{i-1} \sum_{k=1}^{j-1} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} P\left(\cap_{q=1}^{n} A_q\right),\ A_i \in \mathcal{F}$$

- Given a partition of Θ, A_i such that $\cup_{i=1}^{n} A_i = \Theta,\ A_i \cap A_j = \emptyset,$ for $i \neq j$ and an event $B \in \mathcal{F}$

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i)$$

- Let $(\Theta, \mathcal{F}, P)$ be a probability space and an event $B \in \mathcal{F}$. A new probability measure can be defined on $(\Theta, \mathcal{F})$ under the assumption that B has occurred. This new measure is referred to as the *conditional probability*. The probability of the event A occurring conditional on the event B is given by the following formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad A, B \in \mathcal{F} \text{ and } P(B) > 0$$

- Let $A_i \in \mathcal{F},\ i = 1, 2, \cdots, n$ be a partition of Θ.

  The *law of total probability* states:

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i) P(A_i)$$

  The *Bayes formula* states:

$$P(A_j|B) = \frac{P(A_j) P(B|A_j)}{P(B)} = \frac{P(A_j) P(B|A_j)}{\sum_{i=1}^{n} P(A_i) P(B|A_i)}$$

1.4 Random Variables

A (continuous) *random variable* defined over a probability space $\{\Theta, \mathcal{F}, P\}$ is a function X with domain Θ and codomain $\mathbb{R}$, that is:

$$\begin{aligned} X :&\Theta \to \mathbb{R} \\ &\theta \mapsto X(\theta) \end{aligned}$$

The *cumulative distribution function (cdf)* of a random variable X is the function:

$$F(x) = P(X \leq x),\ \infty < x < \infty$$

which satisfies the following properties:

1. $0 \leq F(x) \leq 1$
2. F is a non-decreasing function of x
3. $F(-\infty) = 0$ and $F(\infty) = 1$
4. F is right continuous: $lim_{x \to x_0^+} F(x) = F(x_0)$ for every x

The derivative of the cdf is called the *probability density function (pdf)*

$$\begin{aligned} f(x) &= \frac{dF(x)}{dx} \\ &= P(x \leq X \leq x + dx) \end{aligned}$$

A pdf satisfies the normalization condition

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Also, the following relations are equivalent

$$\begin{aligned} P(a \leq x \leq b) &= F(b) - F(a) \\ &= \int_a^b f(x)dx \end{aligned}$$

A d-dimensional vector whose coordinates are random variables is called a *random vector* $\boldsymbol{X}$:

$$\begin{aligned}\boldsymbol{X} :&\Theta \to \mathbb{R}^d \\ &\theta \mapsto \boldsymbol{X}(\theta)\end{aligned}$$

The definitions of the pdf and cdf are straightforwardly extended for random vectors. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_d) \in \mathbb{R}^d$ be a random vector with d coordinates. Its *joint cumulative distribution function* is defined as:

$$F(x_1, x_2, \ldots, x_d) = P(X_1 \leq x_1, \ldots, X_d \leq x_d), \quad -\infty < x_1, \ldots, x_d < \infty$$

and the corresponding *joint probability density function* is:

$$f(x_1, x_2, \ldots, x_d) = \frac{\partial^d}{\partial_1 \cdots \partial_d} F(x_1, \ldots, x_d)$$

Also, the joint cdf can be obtained by appropriate integration of the joint pdf, that is:

$$F(x_1, x_2, \ldots, x_d) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f(u_1, \ldots, u_d) du_1 \ldots du_d, \quad -\infty < x_1, \ldots, x_d < \infty$$

The *marginal cumulative distribution function* of the variable X_m, for $m = 1, ..., d$ is obtained by integrating out the rest of the variables in the random vector. More specifically,

$$\begin{aligned}F_{X_m}(x_m) &= P(X_m \leq x_m) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x_m} \cdots \int_{-\infty}^{\infty} f(u_1, \ldots, u_d) du_1 \ldots du_d, \quad -\infty < x_m < \infty\end{aligned}$$

and the corresponding *marginal probability density function* as

$$f_{X_m}(x_m) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_d) dx_1 \ldots dx_{m-1} dx_{m+1} \ldots dx_d$$

Now, let us consider a random vector $\boldsymbol{X} \in \mathbb{R}^d$ whose coordinates X_i satisfy the

following relation:

$$F(x_1, \ldots, x_d) = F_{X_1}(x_1) \ldots F_{X_d}(x_d)$$

or, equivalently,

$$f(x_1, \ldots, x_d) = f_{X_1}(x_1) \ldots f_{X_d}(x_d)$$

Then, the X_i's are said to be *independent* random variables and the above equations are the necessary and sufficient conditions for *independence*. Intuitively, the notion of independence suggests that the value of a random variable X_i, obtained through an experiment, does not give any further knowledge about the value of another random variable X_j in the same experiment.

Lastly, given two random variables X, Y and their joint pdf f, the *conditional probability density function* of Y given X is defined by the following relation:

$$f_{Y|X} = \begin{cases} \dfrac{f(x,y)}{f_X(x)}, & 0 < f_X(x) < \infty \\ 0, & \text{otherwise} \end{cases}$$

1.5 Moments

Let us first define a linear operator acting on random variables defined over a probability space $(\Theta, \mathcal{F}, P)$. This operator, acting on a (continuous) random variable X, is defined as:

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

and is called *expectation operator*, or simply *expectation* of a random variable X. For random variables with finite expectations, it is straightforward to prove that the following properties of $\mathbb{E}$ hold:

1. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
2. $\mathbb{E}[aX] = a\mathbb{E}[X]$
3. $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ iff X, Y independent

The expected value of a function of X, $g(X)$ given that X has a probability density function $f(x)$ is given by:

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

The above relations extend to multidimensional cases, where f is replaced by the corresponding joint density.

We define the n-th *moment about the origin* of a random variable X as the expectation of X^n and denote it as μ_n, that is:

$$\mu_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f(x)dx$$

For $n = 1$, μ_1, usually denoted with m, is called the *mean value* of X and for $n = 2$, μ_2 is the *mean square*.

Similarly, we define the n-th *moment about the mean* or n-th *central moment* of a random variable X as the expectation of $(X - m)^n$ and denote it as s_n, that is:

$$s_n = \mathbb{E}[(X - m)^n] = \int_{-\infty}^{\infty} (x - m)^n f(x)dx$$

The first central moment s_1 is zero and the second central moment s_2 is called the *variance*, usually denoted as σ^2, where $\sigma = \sqrt{s_2}$ is the *standard deviation*. The third and forth central moments are used in the definition of the *skewness* $\gamma = \frac{s_3}{\sigma^3}$ and *kurtosis* $\kappa = \frac{s_4}{\sigma^4}$. Skewness and kurtosis are both descriptors of the shape of a pdf, where the former measures the asymmetry of a pdf and the latter its 'tailedness'.

Some other important quantities that can be defined using the moments are:

- The *coefficient of variation* for a random variable X

$$c.o.v = \frac{\sigma}{m}$$

- The *covariance* between two random variables X, Y

$$Cov[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]E[Y]$$

  Note that in the case $X = Y$, $Cov[X, X] \doteq Var[X]$ where $Var[X]$ is the variance of X and

$$Var[X] = \mathbb{E}[X^2] - \mathbb{E}^2[X] = \mu_2 - m^2$$

- The *correlation coefficient*

$$\rho = \frac{\mathbb{E}[XY]}{\sqrt{Var[X]}\sqrt{Var[Y]}}$$

Lastly, we define the *conditional expectation* of a random random variable Y given the event $X = x$ has occurred, as

$$\begin{aligned}\mathbb{E}[Y|X = x] &= \int_{-\infty}^{\infty} y f(y|x) dy \\ &= \int_{-\infty}^{\infty} \frac{y f(x, y)}{f_X(x)} dy\end{aligned}$$

1.6 $\mathscr{L}_2$-Probability space

The random variables defined over a probability space $(\Theta, \mathscr{F}, P)$ satisfying the property:

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} |x|^2 f(x) dx < \infty$$

are elements of a vector space, henceforth denoted as $\mathscr{L}_2$, defined over the field of real numbers $\mathbb{R}$. In $\mathscr{L}_2$ the addition operation corresponds to the usual addition of two real-valued random variables, and the operation of scalar multiplication corresponds to the usual multiplication of a real-valued random variable by a real number.

The vector space is also endowed with an inner-product structure induced be the expectation operator $E[\cdot]$:

$$\langle X, Y \rangle \doteq \mathbb{E}[XY]$$

One can easily verify from the definition of the expectation that for $X, Y, Z \in \mathscr{L}_2$ the following inner product properties are satisfied:

1. $\langle X, Y \rangle = \langle Y, X \rangle$
2. $\langle X, X \rangle \geq 0$ and $\langle X, X \rangle = 0$ iff $P(X = 0) = 1$ $(X \equiv 0)$
3. $\langle aX, Y \rangle = a\langle X, Y \rangle$ for $a \in \mathbb{R}$
4. $\langle X + Y, Z \rangle = \langle X, Z \rangle + \langle Y, Z \rangle$

The $\mathscr{L}_2$ space is also *complete*, which means that every *Cauchy sequence* in $\mathscr{L}_2$ converges in $\mathscr{L}_2$. As a consequence, $\mathscr{L}_2$ is a *Hilbert space* (complete inner product space) and, thus, there exists an orthonormal basis that spans $\mathscr{L}_2$. This particular conclusion will prove very useful in the subsequent formulations.

2

Stochastic Processes

2.1 Introduction

In probability theory, a stochastic process is a mathematical object which is defined as the collection of an infinite number of random variables. Their study originated from the need to model dynamic systems that seemed to randomly change over time. The motion of a particle immersed in a fluid is such an example. Each time the particle is immersed in the fluid, it will follow a different trajectory as a result of the random microscopic collisions with the molecules of the fluid. This particular example of a stochastic process is known as the *Brownian motion*, named after the botanist Robert Brown, who first observed it in 1827. Since then stochastic processes have found applications in various fields in science and engineering, including image and signal processing, earthquake prediction, weather forecasting and finance.

The theory of stochastic processes is an important field in mathematics combining elements of different branches such as probability theory, measure theory, set theory, Fourier analysis and real analysis. The aim of this chapter is to revisit the fundamental concepts on stochastic processes, that will be used throughout this text.

2.2 Definitions

Let us recall, that a random variable was defined as a rule for assigning to every outcome θ of an an experiment a number $X(\theta)$. In certain applications, however, the experiment 'evolves' with respect to some deterministic parameter t belonging in an interval I. For instance, this would be the case of an engineering system subjected to random dynamic loads over a time interval $I \subseteq \mathbb{R}^+$. Then, the response at a material point of this system would be described by a collection of random variables $\{X(t)\}$ indexed by the parameter $t \in I$, rather than a single random variable. This collection of random variables over the interval I is called a *stochastic process* and is denoted by $\{X(t), t \in I\}$ or X. In this regard, a stochastic process can be seen as a generalization of the concept of a random variable, in the sense that it assigns to every outcome θ of the experiment, a function $X(t, \theta)$, referred to as a *realization* or a *sample function*.

A stochastic process is called a *continuous-time real-valued stochastic process* if I is the real axis and each random variable in the collection takes on values in $\mathbb{R}$. If I is the set of integers, then the process is called *discrete-time real-valued stochastic process*. Furthermore, if the process takes on values in $\mathbb{R}^d$, $d > 1$, the the process is called $\mathbb{R}^d$-*valued stochastic process* and the notation $\{\boldsymbol{X}(t), t \in I\}$ or, simply, $\boldsymbol{X}$ will be used. Lastly, if $\boldsymbol{X}$ is indexed by some space coordinate $\boldsymbol{s} \in D \subseteq \mathbb{R}^n$ rather than time t, then $\{\boldsymbol{X}(\boldsymbol{s}), \boldsymbol{s} \in D\}$ is called a *random field*. The summarize the above we provide the following two definitions:

Let $\boldsymbol{X} : I \times \Theta \to \mathbb{R}^d$ be a vector-valued function of two arguments, $t \in I$ and $\theta \in \Theta$, where $I \subseteq \mathbb{R}$ and $(\Theta, \mathscr{F}, P)$ denotes a probability space. Then, $\boldsymbol{X}$ is said to be an $\mathbb{R}^d$-valued stochastic process, if $\boldsymbol{X}(t)$ is an $\mathbb{R}^d$-valued random variable on the probability space $(\Theta, \mathscr{F}, P)$ for each $t \in I$.

Similarly, let $\boldsymbol{X} : D \times \Theta \to \mathbb{R}^d$ be a vector-valued function of two arguments, $\boldsymbol{s} \in D$ and $\theta \in \Theta$, where $D \subseteq \mathbb{R}^n$ and $(\Theta, \mathscr{F}, P)$ denotes a probability space. Then, $\boldsymbol{X}$ is said to be an $\mathbb{R}^d$-valued random field, if $\boldsymbol{X}(\boldsymbol{s})$ is an $\mathbb{R}^d$-valued random variable on the probability space $(\Theta, \mathscr{F}, P)$ for each $\boldsymbol{s} \in D$.

2.3 Statistics of Stochastic Processes

Let $\{\boldsymbol{X}(t), t \geq 0\}$ be an $\mathbb{R}^d$-valued stochastic process defined over a probability space $(\Theta, \mathscr{F}, P)$. Let $n \geq 1$ be an integer, $t_i \geq 0$, $i = 1, ..., n$, be arbitrary distinct times and

set $\mathfrak{X}_n = (\boldsymbol{X}(t_1), ..., \boldsymbol{X}(t_n))$. The *finite dimensional distributions of order* n of $\boldsymbol{X}$ are the distributions of the random vectors $\mathfrak{X}_n$:

$$F_n\left(\boldsymbol{x}^{(1)}, ..., \boldsymbol{x}^{(n)}; t_1, ..., t_n\right) = P\left(\cap_{i=1}^{n}\{\boldsymbol{X}(t_i) \in \prod_{k=1}^{d}(-\infty, x_{i,k}]\}\right) \tag{2.1}$$

where $\boldsymbol{x}^{(i)} = (x_{i,1}, ..., x_{i,d}) \in \mathbb{R}^d$. For example, if X is a real-valued ($d = 1$) stochastic process, its finite dimensional distributions are:

$$F_n\left(x_1, ..., x_n; t_1, ..., t_n\right) = P\left(X(t_1) \geq x_1, ..., X(t_n) \geq x_n\right) \tag{2.2}$$

Then, the corresponding *finite dimensional densities* of X can be derived as:

$$f_n\left(x_1, ..., x_n; t_1, ..., t_n\right) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_n\left(x_1, ..., x_n; t_1, ..., t_n\right) \tag{2.3}$$

A complete determination of the statistical properties of a stochastic process can only be attained through the knowledge of the function $F_n\left(x_1, ..., x_n; t_1, ..., t_n\right)$ for every x_i,t_i and n. However, in most applications, the available information on the process allow us to estimate at most its first and second order finite dimensional densities. From these, the following averages can be obtained as follows:

- The mean $m(t)$ of X is the expected value of each random variable $X(t)$:

  $$m(t) = \mathbb{E}\left[X(t)\right] = \int_{-\infty}^{\infty} x f(x; t) dx$$

  It should be mentioned that from every process X, a zero-mean process can be obtained, called *centered* process, simply by subtracting the mean, that is

  $$\tilde{X}(t) = X(t) - m(t)$$

- The autocorrelation $R_X(t_1, t_2)$ of X is the expected value of the product $X(t_1)X(t_2)$:

  $$R_X(t_1, t_2) = \mathbb{E}\left[X(t_1)X(t_2)\right] = \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; t_1, t_2) dx_1 dx_2$$

  In the case $t_1 = t_2 = t$ the autocorrelation function gives the mean-square moment of each random variable $X(t)$, that is, $R_X(t, t) = \mathbb{E}\left[X^2(t)\right]$.

- The autocovariance $C_X(t_1, t_2)$ of X is the expectation of the product $(X(t_1) - m(t_1))$ $(X(t_2) - m(t_2))$, which is the covariance of the random variables $X(t_1)$ and $X(t_2)$:

$$\begin{aligned} C_X(t_1, t_2) &= \mathbb{E}\left[(X(t_1) - m(t_1))\,(X(t_2) - m(t_2))\right] \\ &= \int_{-\infty}^{\infty} (x_1 - m(t_1))\,(x_2 - m(t_2))\, f(x_1, x_2; t_1, t_2) dx_1 dx_2 \end{aligned}$$

  After some algebra, it can be shown that:

$$C_X(t_1, t_2) = R_X(t_1, t_2) - m(t_1)m(t_2)$$

  and in the case $t_1 = t_2 = t$, $C(t, t)$ equals the variance of each random variable $X(t)$. Also, it is evident, that for zero mean processes $C_X(t_1, t_2) \equiv R_X(t_1, t_2)$.

- The correlation coefficient $r_X(t_1, t_2)$ is defined as the ratio:

$$r_X(t_1, t_2) = \frac{C_{X(t_1,t_2)}}{\sqrt{C_{X(t_1,t_1)} C_{X(t_2,t_2)}}}$$

- Correlation and covariance between two stochastic process X and Y, known as *cross-correlation* and *cross-covariance* are defined in a similar fashion:

$$R_{XY}(t_1, t_2) = \mathbb{E}\left[(X(t_1))\,(Y(t_2)))\right]$$

  and

$$C_{XY}(t_1, t_2) = \mathbb{E}\left[(X(t_1) - m_X(t_1))\,(Y(t_2) - m_Y(t_2))\right]$$

2.4 Classes of Stochastic Processes

2.4.1 Stationary processes

A stochastic process X is called *stationary in the strict sense* or *strictly stationary* if

$$f_n\left(x_1, ..., x_n; t_1, ..., t_n\right) = f_n\left(x_1, ..., x_n; t_1 + \tau, ..., t_n + \tau\right) \tag{2.4}$$

for any $n \geq 1$ and time shift τ. The above definition implies that a process is strictly stationary if its joint pdf does not change when shifted in time. This assumption is

quite restrictive and rarely met in real life phenomena. A milder assumption is that of *wide-sense stationarity* or *weak stationarity.*

A stochastic process X is called *stationary in the wide sense* or *weakly stationary* if its mean is constant

$$\mathbb{E}\left[X(t)\right] = m \ (= cnst) \tag{2.5}$$

and its autocorrelation depends only on the time lag $\tau = t_1 - t_2$

$$R_X(t_1, t_2) = R(t_1 - t_2) = R(\tau) = \mathbb{E}\left[X(\tau)^2\right] \tag{2.6}$$

2.4.2 Ergodic processes

A stochastic process X is said to be *ergodic* if the ensemble averages are equal to time averages. In other words, if every statistical property of the process can be obtained from one realization (of sufficient length) of the process. Weaker ergodicity conditions are that of *ergodicity of the mean* and *ergodicity of the autocorrelation* expressed through equations (2.7) and (2.8), respectively :

$$m = \mathbb{E}\left[X(t)\right] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)dt \tag{2.7}$$

$$R(\tau) = \mathbb{E}\left[X(t)X(t+\tau)\right] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(t)X(t+\tau)dt \tag{2.8}$$

Note that an ergodic process is always stationary, but the reverse is not always true.

2.4.3 Gaussian processes

A stochastic process X is said to be *Gaussian* if the joint pdf of the random variables $\{X(t_1), ..., X(t_n)\}$ is Gaussian for any n and t_i, $i = 1, ..., n$. A weakly-stationary Gaussian process possesses the favorable property of being completely defined by its mean value m, its standard deviation σ and its autocorrelation function. The pdf of the random variable $X(t)$ is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp^{-\frac{(x-m)^2}{2\sigma^2}} \tag{2.9}$$

This implies that a weakly-stationary Gaussian process is also stationary in the strict sense.

2.4.4 Translation processes

Generally, when a stochastic process X is not Gaussian it is practically impossible to estimate the joint pdf of the random variables. However, a class of non-Gaussian stochastic processes with given marginal distribution and second moment information can be defined via a nonlinear marginal transformation (translation), g, of an underlying Gaussian field, $G(t)$, that is

$$X(t) = g\left(G(t)\right) = F^{-1} \circ \Phi\left(G\right) \tag{2.10}$$

where F is an arbitrary distribution with density f, Φ is the cdf of the Gaussian distribution $N(0,1)$ and G is a Gaussian process with $\mathbb{E}[G(t)] = 0$ and $\mathbb{E}[G(t)^2] = 1$. The marginal distribution of X is F and the finite dimensional density of order n of X is given by

$$f_n(x_1, ..., x_n) = \frac{1}{\sqrt{(2\pi)^n det(\boldsymbol{\rho})}} \prod_{p=1}^{n} \frac{f(x_p)}{\phi(y_p)} exp\left(-\frac{1}{2}\boldsymbol{y}^T \boldsymbol{\rho}^{-1} \boldsymbol{y}\right) \tag{2.11}$$

where, $\boldsymbol{\rho} = \mathbb{E}\left[G(t_p)G(t_q)\right]$, $y_p = \Phi^{-1} \circ F(x_p),\ p, q = 1, ..., n$, ϕ is the density of $N(0,1)$, and $\boldsymbol{y} = (y_1, ..., y_n)$.

However, the existence of a translation field with the desired characteristics is not always guaranteed. This is due to the fact that the marginal distribution $X(t)$ imposes constraints to its correlation structure. Therefore, the following compatibility equation must be satisfied between F and $R_X(\tau)$

$$R_X^T(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F^{-1} \circ \Phi\left(G_1\right) F^{-1} \circ \Phi\left(G_2\right) f(G_1, G_2; R_G(\tau)) dG_1 dG_2 \tag{2.12}$$

with $G_1 = G(t)$, $G_2 = G(t+\tau)$ and $f(G_1, G_2; R_G(\tau))$ denotes the joint density of G_1 and G_2. If F and $R_X(\tau)$ are proven to be incompatible then one has to resort to translation fields that match approximately the target marginal distribution and/or autocorrelation function [63, 99, 141].

2.5 Power spectral Density

Let X be a weakly stationary stochastic process in $L^2(\Theta, \mathcal{F}, P)$. It can be easily verified that the autocorrelation function $R_X(t_1, t_2)$ is positive definite, that is

$$\sum_{i=1}^{n}\sum_{j=1}^{n} R_X(t_i, t_j)c_i c_j \geq 0 \tag{2.13}$$

for all finite sequences of points $t_1, ..., t_n$ and all choices of real numbers $c_1, ..., c_n$.

Bochner's Theorem states that a continuous function $r : \mathbb{R} \to \mathbb{C}$ is positive definite if and only if it admits a representation of the form

$$r_X(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau}\mathcal{S}(\omega)d(\omega) \tag{2.14}$$

where $\mathcal{S}$ is a real-valued, positive and bounded function. The autocorrelation function R_X of a weakly stationary process satisfies the conditions of Bocher's theorem and therefore it has a representation of the form of eq. (2.14). In this context, the function $\mathcal{S}$ is called the *spectral density function* of X and the variable ω is the *angular frequency*. Under this prism, the autocorrelation and the spectral density functions are Fourier pairs that satisfy the following equations, known as Wiener-Khinchin relations:

$$R_X(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau}\mathcal{S}(\omega)d\omega \tag{2.15}$$

$$\mathcal{S}_X(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}R(\tau)d\tau \tag{2.16}$$

Using Euler's formula, $e^{i\omega\tau} = cos(\omega\tau) + isin(\omega\tau)$ and using the fact that $sin(\omega\tau)$ is anti-symmetric, while $cos(\omega\tau), R_X(\tau)$ and $\mathcal{S}_X(\omega)$ are symmetric functions, then the integrals of the above equations can be simplified as:

$$R_X(\tau) = 2\int_{0}^{\infty} cos(\omega\tau)\mathcal{S}(\omega)d\omega \tag{2.17}$$

$$\mathcal{S}_X(\omega) = \frac{1}{\pi}\int_{0}^{\infty} cos(\omega\tau)R(\tau)d\tau \tag{2.18}$$

From eq. (2.17) one observes that the area under $\mathcal{S}_X(\omega)$ is $R_X(0)$. If $\mathbb{E}[X(t)] = 0$,

this area is the variance of X.

2.6 Discretization of a stochastic processes

2.6.1 The Karhunen-Loève series expansion

The Karhunen-Loève expansion of a zero-mean random process $X(t,\theta)$ is based on the spectral decomposition of its covariance function defined as

$$C_X(t_i,t_j) = \sigma_X(t_i)\cdot\sigma_X(t_j)\cdot\rho(t_i,t_j) \tag{2.19}$$

where ρ is the correlation coefficient. By definition, $C_X(t_i,t_j)$ is bounded, symmetric and has the following spectral or eigen-decomposition:

$$C_X(t_i,t_j) = \sum_{n=1}^{\infty}\lambda_n\varphi_n(t_i)\varphi_n(t_j) \tag{2.20}$$

where φ_n and λ_n are orthogonal deterministic eigenfunctions and eigenvalues of the covariance function, respectively, derived from the solution of the homogeneous Fredholm integral equation of the second kind for the covariance kernel:

$$\int_{\mathcal{D}} C_X(t_i,t_j)\cdot\varphi_n(t_j)\mathrm{d}t_j = \lambda_n\cdot\varphi_n(t_i) \tag{2.21}$$

where $\mathcal{D}$ is the domain in which the stochastic processes is defined. The key to KL expansion is to obtain the eigenvalues and eigenfunctions by solving eq.(2.21). Because an analytical solution of eq.(2.21) is tractable only is special cases, in general a numerical solution is the only resort.

The eigenfunctions form a complete orthogonal set satisfying the equation:

$$\int_{\mathcal{D}}\varphi_k(t)\varphi_l(t)\mathrm{d}t = \delta_{kl} \tag{2.22}$$

where δ_{kl} is the Kronecker-delta function. Any realization of $X(t,\omega)$ can thus be expanded over this basis as follows :

$$X(t,\theta) = \sum_{i=1}^{\infty}\sqrt{\lambda_n}\cdot\varphi_n(t)\cdot\xi_n(\theta), \quad t\in\mathcal{D} \tag{2.23}$$

where $\xi_n(\theta)$ is a set of uncorrelated random variables with mean $\mathbb{E}\big[\xi_n(\theta)\big] = 0$ and covariance function $\mathbb{E}\big[\xi_k(\theta) \cdot \xi_l(\theta)\big] = \delta_{kl}$ which can be expressed as

$$\xi_n(\theta) = \frac{1}{\sqrt{\lambda_n}} \int_{\mathcal{D}} X(t,\theta) \cdot \varphi_n(t) \mathrm{d}t \tag{2.24}$$

Equation (2.23) is known to converge in the mean square sense for any distribution of $X(t,\theta)$. The KL expansion of a Gaussian process has the property that $\xi_n(\theta)$ are independent standard normal variables. For practical implementation, the series is approximated by a finite number of terms M, giving

$$X(t,\theta) \approx \hat{X}(t,\theta) = \sum_{n=1}^{M} \sqrt{\lambda_n} \cdot \varphi_n(t) \cdot \xi_n(\theta) \tag{2.25}$$

The corresponding covariance function is then approximated by

$$\hat{C}_X(t_i,t_j) = \sum_{n=1}^{M} \lambda_n \varphi_n(t_i) \varphi_n(t_j) \tag{2.26}$$

This truncated series is optimal in the mean square since the eigenvalues λ_n of eq. (2.25) are converging fast to zero (Fig.2.1). Thus, the choice of the covariance eigenfunction basis $\{\varphi_n(t)\}$ is optimal in the sense that the mean square error resulting from a truncation after the M-th term is minimized.

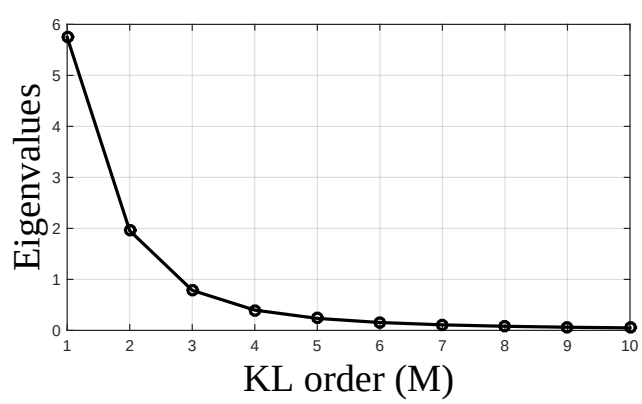


Figure 2.1: Decaying eigenvalues from the solution of the Fredholm integral of the second kind for $M = 10$.

The variance error e_{var} after truncating the expansion in M terms can be easily computed as

$$e_{var} = \mathrm{Var}\big[X(t,\theta) - \hat{X}(t,\theta)\big] = \sigma_X^2 - \sum_{n=1}^{M} \lambda_n \cdot \varphi_n^2(t) \tag{2.27}$$

The righthand side of the above equation means that the KL expansion always under-represents the true variance of the field.

For random processes where the analytical solution of the Fredholm integral equation is intractable, a numerical solution is necessary. One major category of such solution schemes are the expansion methods such as the Galerkin, the collocation and the Rayleigh-Ritz methods. Galerkin methods are essentialy error minimization schemes with respect to some residual calculated over the entire domain of the solution. Assuming that each eigenfunction $\varphi_n(t)$ of $C_X(t_i, t_j)$ may be represented by its expansion over a polynomial basis $\{h_i(\cdot)\}$, defined in the solution space, as:

$$\varphi_n(t) = \sum_{i=1}^{\infty} d_i^n \cdot h_i(t) \tag{2.28}$$

where d_i^n are unknown coefficients to be estimated, the Galerkin procedure targets to an optimal approximation of the eigenfunctions $\varphi_n(\cdot)$ after truncating the above series in N terms and computing the residual as:

$$\varepsilon_N(t) = \sum_{i=1}^{N} d_i^n \cdot \left[\int_{\mathcal{D}} C_X(t_i, t_j) \cdot h_i(t_j) \mathrm{d}t_j - \lambda_j \cdot h_i(t) \right] \tag{2.29}$$

Requiring the residual to be orthogonal to the space spanned by the same basis we get

$$< \varepsilon_N, h_j > := \int_{\mathcal{D}} \varepsilon_N(t) \cdot h_j(t) \mathrm{d}t = 0, \quad j = 1, \ldots, N \tag{2.30}$$

which leads to the following matrix eigenvalue equation:

$$\mathbf{C} \cdot \mathbf{D} = \mathbf{\Lambda} \cdot \mathbf{B} \cdot \mathbf{D} \tag{2.31}$$

where

$$\mathbf{B}_{ij} = \int_{\mathcal{D}} h_i(t) \cdot h_j(t) dt \tag{2.32}$$

$$\mathbf{C}_{ij} = \int_{\mathcal{D}} \int_{\mathcal{D}} C_X(t_i, t_j) \cdot h_i(t_j) \mathrm{d}t_j \tag{2.33}$$

$$\mathbf{D}_{ij} = d_i^j \tag{2.34}$$

$$\mathbf{\Lambda}_{ij} = \delta_{ij} \cdot \lambda_j \tag{2.35}$$

where $\mathbf{C}, \mathbf{D}, \mathbf{B}$ and $\mathbf{\Lambda}$ are $N \times N$-dimensional matrices. This generalized algebraic eigenvalue problem of eq. (2.31) can be solved for $\mathbf{D}$ and $\mathbf{\Lambda}$ and with backsubstitution we can estimate the eigenfunctions of the covariance kernel. This solution scheme can be implemented using piecewise polynomials for the basis $\{h_i(\cdot)\}$ of the expansion.

2.6.2 The Spectral Representation series expansion

The spectral representation method was proposed by Shinozuka and Deodatis [142] and generates sample functions that are ergodic in the mean value and autocorrelation. Its main property is that it expands the stochastic field to a series of trigonometric functions with random phase angles. For a zero-mean, one-dimensional stationary stochastic process $X(t, \theta)$ with autocorrelation function $R_X(t_i, t_j)$ and two-sided power spectral function $S_X(\omega)$ it was proven in [142], that

$$X(t, \theta) = \sqrt{2} \cdot \sum_{k=0}^{\infty} \left(2S_X(\omega_k) \cdot \Delta\omega\right)^{\frac{1}{2}} \cdot \cos(\omega_k \cdot t + \Phi_k) \tag{2.36}$$

where, $\omega_k = k\Delta\omega$ for $k = 1, ...$ are the frequencies of the power-spectrum $(2S_X)$, $(2S_X(\omega_k)$ are the values of the spectrum at each frequency ω_k and Φ_k are the random phase angles, which are random variables following the uniform distribution $\mathcal{U}\,[0, 2\pi]$.

For practical applications, in order to have a realization of this $X(t, \theta)$ we need to truncate the summation of eq. (2.36) after N terms.

$$\hat{X}(t, \theta) = \sqrt{2} \cdot \sum_{n=0}^{N-1} A_n \cdot \cos(\omega_n \cdot t + \Phi_n) \tag{2.37}$$

where,

$$
\begin{aligned}
A_n &= \left(2S_X(\omega_k)\cdot\Delta\omega\right)^{\frac{1}{2}} \quad \text{for } n = 0, 1, \ldots, N-1 \\
\omega_n &= n\cdot\Delta\omega \\
\Delta\omega &= \frac{\omega_u}{N} \\
A_0 &= 0 \quad \text{or } S_X(\omega_0 = 0) = 0
\end{aligned}
\tag{2.38}
$$

The coefficient A_0 is chosen zero such that the temporal mean value averaged over the whole simulation time $T_0 = \frac{2\pi}{\Delta\omega}$ of the generated stochastic process $\hat{X}(t,\theta)$ remains zero in each generated sample. This is because if some power spectral contribution is added at $\omega = 0$, a random variable term is always present, shifting the temporal (sample) average apart from being zero. In order to avoid having to impose this condition the frequency shifting theorem was proposed in [164]. However, this approach has the side effect of doubling the period of the simulated field.

In eq.(2.38) ω_u is usually applied as the upper-cut off frequency after which the power spectrum becomes practically zero. In order to estimate this frequency we use the following criterion:

$$
\int_0^{\omega_u} S_X(\omega)\mathrm{d}\omega = (1-\epsilon)\int_0^{\infty} S_X(\omega)\mathrm{d}\omega \tag{2.39}
$$

where $\epsilon \ll 1$ is the 'admissible relative error'. The target autocorrelation function $R_{\hat{X}}(\tau)$ is given by

$$
R_{\hat{X}}(\tau) = \int_{-\omega_u}^{\omega_u} S_X(\omega)e^{i\omega\tau}\mathrm{d}\omega = \int_0^{\omega_u} 2S_X(\omega)\cos\omega\tau\mathrm{d}\omega \tag{2.40}
$$

The difference between these two functions

$$
\epsilon^*(\tau) = R_X(\tau) - R_{\hat{X}}(\tau) = \int_{\omega_u}^{\infty} 2S_X(\omega)\cos(\omega\tau)\mathrm{d}\omega \tag{2.41}
$$

corresponds to the mean square simulation error due to the truncation of the spectral density function for $|\omega| \geqslant \omega_u$, which is termed 'truncation error'.

One sample function of the stochastic process can be generated by replacing the phase

angles $\Phi_0, \ldots, \Phi_{N-1}$ in eq. (2.37) with their corresponding sample values $\phi_0(\omega), \ldots, \phi_{N-1}(\omega)$, as these can be generated by some random number generator as follows:

$$\hat{X}(t,\theta) = \sqrt{2} \cdot \sum_{n=0}^{N-1} A_n \cdot \cos(\omega_n \cdot t + \phi_n(\theta)) \tag{2.42}$$

It must be mentioned that the step Δt of the generated sample functions must satisfy the following condition in order to avoid aliasing.

$$\Delta t \leq \frac{\pi}{\omega_u} \tag{2.43}$$

The sample functions generated by eq.(2.42) are obviously bounded by

$$\left|\hat{X}(t,\theta)\right| \leq \sqrt{2} \sum_{n=0}^{N-1} A_n \tag{2.44}$$

3

The finite element method

3.1 Introduction

Without a doubt the development of accurate mathematical models is essential for the study of physical systems. Under appropriate assumptions concerning the underlying physical mechanisms and the laws that govern them, these models are most often described by a differential equation satisfying specific boundary conditions, also referred to as a *boundary value problem* (BVP). However, in most practical problems, where complicated geometries, loading conditions, material properties and nonlinearities are involved, solving analytically the corresponding BVPs is often intractable. Therefore, the only alternative is to find approximate solutions using numerical methods. In this regard, the *finite element method* [166, 125, 14] is the most powerful computational technique used to obtain numerical solutions of BVPs in engineering.

The aim of this chapter is to revisit the basic principles of the finite element method. Starting from the abstract formulation of the BVP for the case of 3d elasticity, the variational BVP is obtained. Then, the finite element approximation technique is used, in order to convert the differential equations describing the problem into a system of linear equations.

3.2 The abstract problem

Most problems in the theory of elasticity can be written in the following abstract manner:

Let $\mathscr{V}$ be a normed linear space over $\mathbb{R}$. Let $\mathscr{J} : \mathscr{V} \to \mathbb{R}$ be a functional which can be written in the form:

$$\mathscr{J}(\boldsymbol{v}) = \frac{1}{2}a(\boldsymbol{v}, \boldsymbol{v}) - b(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in \mathscr{V} \tag{3.1}$$

where $a(\cdot, \cdot)$ is a continuous, symmetric bilinear form on $\mathscr{V}$ and b is an element of $\mathscr{V}^\star$, which is the dual space of $\mathscr{V}$. Typically, $\mathscr{J}$ represents the energy of some system under consideration. Then the problem consists in finding an element $\boldsymbol{u} \in \mathscr{V}$ such that

$$\boldsymbol{u} = \arg\min_{\boldsymbol{u} \in \mathscr{V}} \mathscr{J}(\boldsymbol{u}) \tag{3.2}$$

Often, instead of minimizing $\mathscr{J}$ over the entire space $\mathscr{V}$, a non-empty convex subset $\mathscr{K}$ of $\mathscr{V}$ is preferred, over which the minimum of $\mathscr{J}$ is sought, that is

$$\boldsymbol{u} = \arg\min_{\boldsymbol{u} \in \mathscr{K}} \mathscr{J}(\boldsymbol{u}) \tag{3.3}$$

If we denote the above problem as (P) then the next step is to examine the existence of a solution to (P) and, if such a solution does exist, its uniqueness.

Let $\mathscr{V}$ be a normed linear space. A bilinear form $a(\cdot, \cdot)$ on $\mathscr{V}$ is said to be $\mathscr{V}$-elliptic if there exists a constant $c > 0$ such that $\forall \boldsymbol{v} \in \mathscr{V}$:

$$a(\boldsymbol{v}, \boldsymbol{v}) \geq c\|\boldsymbol{v}\|^2 \tag{3.4}$$

The $\mathscr{V}$-ellipticity of $a(\cdot, \cdot)$ suggests that if $a(0, 0) = 0$, then $\boldsymbol{v} = 0$. Also, since $a(\cdot, \cdot)$ is bilinear and symmetric, then it defines an inner-product structure on $\mathscr{V}$. Thus, $\mathscr{V}$ acquires the structure of a Hilbert space. The existence of a unique solution to (P) is given by the following theorem:

Theorem 1. *Let $\mathscr{V}$ be a Banach space and $\mathscr{K}$ a closed convex subset of $\mathscr{V}$. Let $a(\cdot, \cdot)$ be a symmetric, $\mathscr{V}$-elliptic bilinear form. Then, there exists a unique solution for the problem (P).*

The following corollary stems from this theorem: If $\mathscr{K}$ is a subspace of $\mathscr{V}$ then the

solution $\boldsymbol{u}$ is characterized by

$$a(\boldsymbol{u}, \boldsymbol{v}) = b(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in \mathscr{K} \tag{3.5}$$

In order to prove the existence and uniqueness of the solution in theorem 1, the symmetry of the bilinear form was assumed, which provided the Hilbert space structure. If this condition is relaxed but instead $\mathscr{V}$ is assumed to be a Hilbert space and $\mathscr{V}$ is taken equal to $\mathscr{K}$, then the existence and uniqueness of the solution is given by the following theorem, known as the *Lax-Milgram Lemma*:

Theorem 2. *Let $\mathscr{V}$ be a Hilbert space, $a(\cdot,\cdot)$ be a continuous, $\mathscr{V}$-elliptic bilinear form and $b \in V^\star$. If (P) is the problem: find $\boldsymbol{u} \in \mathscr{V}$ such that $\forall \boldsymbol{v} \in \mathscr{V}$,*

$$a(\boldsymbol{u}, \boldsymbol{v}) = b(\boldsymbol{v}) \tag{3.6}$$

then (P) has a unique solution.

3.3 The boundary value problem

Let us denote with $\Omega \subset \mathbb{R}^n$ an open bounded subset and $\partial\Omega$ its boundary. Let us restrict our attention to the case $\Omega \subset \mathbb{R}^3$ and its boundary $\partial\Omega$ can be partitioned into two parts $\partial\Omega_D$ and $\partial\Omega_N$. According to the theory of elasticity, the equilibrium equation of a linear isotropic body occupying Ω which is subjected to body forces $\boldsymbol{f}$ in Ω and surface tractions $\bar{\boldsymbol{t}}$ in the boundary $\partial\Omega_N$ is:

$$\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{f} = 0 \ \text{ in } \Omega \tag{3.7}$$

where $\boldsymbol{\sigma}$ is the stress tensor and $\nabla \cdot \boldsymbol{\sigma}$ its divergence. Next, the strain tensor is defined as:

$$\boldsymbol{\epsilon} = \frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\right) \tag{3.8}$$

with $\boldsymbol{u} : \Omega \to \mathbb{R}^3$ being the vector field of displacements, that will be developed in the body under the prescribed loading conditions. The following relations hold between $\boldsymbol{\sigma}$, $\boldsymbol{\epsilon}$

and $\boldsymbol{u}$:

$$\begin{cases} \epsilon_{ij}(\boldsymbol{u}) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \\ \sigma_{ij}(\boldsymbol{u}) = \lambda\left(\sum_{k=1}^{3} \epsilon_{kk}(\boldsymbol{u})\right)\delta_{ij} + 2\mu\epsilon_{ij}(\boldsymbol{u}) \end{cases} \tag{3.9}$$

for $1 \leq i, j \leq 3$, or, equivalently,

$$\begin{aligned} \boldsymbol{\sigma} &= \lambda Tr(\boldsymbol{\epsilon})\mathbb{I}_{3\times 3} + 2\mu\boldsymbol{\epsilon} \\ &= \lambda(\nabla \cdot \boldsymbol{u})\mathbb{I}_{3\times 3} + 2\mu\boldsymbol{\epsilon} \end{aligned} \tag{3.10}$$

The latter equation is the well-known *Hooke's law* in 3D and λ, μ are positive constants, called the *Lame's coefficients*, given by:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{3.11}$$

$$\mu = \frac{E}{2(1+\nu)} \tag{3.12}$$

where E and ν are the modulus of elasticity and the Poisson coefficient, respectively. Eq. (3.10) can also be recast into the more familiar matrix form:

$$\{\boldsymbol{\sigma}\} = \boldsymbol{D}\{\boldsymbol{\epsilon}\} \tag{3.13}$$

with,

$$\{\boldsymbol{\sigma}\} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{bmatrix}, \ \{\boldsymbol{\epsilon}\} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \epsilon_{12} \\ \epsilon_{23} \\ \epsilon_{31} \end{bmatrix} \tag{3.14}$$

and

$$\boldsymbol{D} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} \end{bmatrix} \tag{3.15}$$

the 3D elasticity matrix.

Proceeding with the previous formulation, inserting eq. (3.10) to eq. (3.7) gives

$$-\lambda \nabla \cdot ((\nabla \cdot \boldsymbol{u})\, \mathbb{I}_{3\times 3}) - 2\mu \nabla \cdot \boldsymbol{\epsilon}(\boldsymbol{u}) = \boldsymbol{f} \tag{3.16}$$

and after some mathematical operations eq. (3.16) becomes

$$-(\lambda + \mu)\nabla(\nabla \cdot \boldsymbol{u}) - \mu \Delta \boldsymbol{u} = \boldsymbol{f} \tag{3.17}$$

where $\Delta \boldsymbol{u} = \nabla(\nabla \cdot \boldsymbol{u}) - \nabla \times (\nabla \times \boldsymbol{u})$ the Laplacian operator acting on $\boldsymbol{u}$. This equation is known as the *Navier-Cauchy* equation and the boundary value problem of eq. (3.7), that is

$$\begin{cases} -\sum_{j=1}^{3} \frac{\partial}{\partial x_j} \left(\sigma_{ij}(\boldsymbol{u})\right) = f_i, \text{ in } \Omega \text{ for } i = 1,2,3 \\ \boldsymbol{u} = \bar{\boldsymbol{u}} \text{ on } \partial\Omega_D \\ \boldsymbol{\sigma}\mathbf{v} = \bar{\boldsymbol{t}} \text{ on } \partial\Omega_N \end{cases} \tag{3.18}$$

becomes

$$\begin{cases} -(\lambda + \mu)\nabla(\nabla \cdot \boldsymbol{u}) - \mu \Delta \boldsymbol{u} = \boldsymbol{f} \text{ in } \Omega \\ \boldsymbol{u} = \bar{\boldsymbol{u}} \text{ on } \partial\Omega_D \\ \boldsymbol{\sigma}\mathbf{v} = \bar{\boldsymbol{t}} \text{ on } \partial\Omega_N \end{cases} \tag{3.19}$$

with $\mathbf{v}$ denoting the outward pointing unit normal vector. Equations (3.19) constitute the *system of linear elasticity.*

The system in eq. (3.19) is referred to as the *strong form* of the BVP and it requires the solution vector field to be at least two times continuously differentiable, that is, $\boldsymbol{u} \in C^2(\Omega)$. It turns out that this condition can be relaxed using the *weak*, or *variational formulation* of the BVP. This approach offers two advantages; the solution space is enlarged and the Neumann boundary conditions can be straightforwardly incorporated in the differential equation. In order to derive the variational BVP of eq. (3.19) some definitions have to be provided first.

The Sobolev space $\mathscr{W}^{m,p}(\Omega)$, with $m \in \mathbb{Z}^{\geq 0}$ and p and integer greater than 1, is given by:

$$\mathscr{W}^{m,p}(\Omega) = \{v \in \mathscr{L}^p(\Omega) | D^\alpha v \in \mathscr{L}^p(\Omega),\ \forall |\alpha| \leq m\} \tag{3.20}$$

where α is a multi-index denoting the n-tuple $(\alpha_1, \alpha_2, ..., \alpha_n)$, $\alpha_i \in \mathbb{Z}^{\geq 0}$ and $D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$ are the partial derivatives of order $|\alpha|$, which are understood in the weak sense. The Sobolev spaces with $p = 2$ are particularly important because they also form a Hilbert space $\mathscr{H}^m = \mathscr{W}^{m,2}$.

Consider the vector spaces

$$\begin{aligned} U &= \{\boldsymbol{u} \in \mathscr{W}^{1,2} : \boldsymbol{u} = \bar{\boldsymbol{u}} \text{ on } \partial\Omega_D\} \\ W &= \{\boldsymbol{w} \in \mathscr{W}^{1,2} : \boldsymbol{w} = 0 \text{ on } \partial\Omega_D\} \end{aligned} \tag{3.21}$$

with U being the space of admissible functions and W the space of test functions. Multiplying both sides of the equilibrium equation (3.7) with $\boldsymbol{w} \in W$ and integrating over Ω gives:

$$-\int_\Omega (\nabla \cdot \boldsymbol{\sigma}) \cdot \boldsymbol{w} d\boldsymbol{x} = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{w} d\boldsymbol{x} \tag{3.22}$$

The following integral formula holds for the left hand side of eq. (3.22)

$$\begin{aligned} \int_\Omega (\nabla \cdot \boldsymbol{\sigma}) \cdot \boldsymbol{w} d\boldsymbol{x} &= -\int_\Omega \boldsymbol{\sigma} : \nabla \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \cdot (\boldsymbol{\sigma}\mathbf{v}) d\boldsymbol{s} + \int_{\partial\Omega_D} \boldsymbol{w} \cdot (\boldsymbol{\sigma}\mathbf{v}) d\boldsymbol{s} \\ &= -\int_\Omega \boldsymbol{\sigma} : \nabla \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} d\boldsymbol{s} + \cancelto{0}{\int_{\partial\Omega_D} \boldsymbol{w} \cdot (\boldsymbol{\sigma}\mathbf{v}) d\boldsymbol{s}} \end{aligned} \tag{3.23}$$

where : denotes the tensor inner product and $d\boldsymbol{s}$ is the differential volume element on $\partial\Omega$. Also, the last term in eq. (3.23) is zero due to the compact support of $\boldsymbol{w}$. Therefore, eq. (3.22) can be rewritten as

$$\int_{\Omega} \boldsymbol{\sigma} : \nabla \boldsymbol{w} d\boldsymbol{x} = \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} d\boldsymbol{s} \tag{3.24}$$

or, due to the symmetry of $\boldsymbol{\epsilon}$

$$\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\boldsymbol{w}) d\boldsymbol{x} = \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} d\boldsymbol{s} \tag{3.25}$$

Applying eq. (3.10) to (3.25) we have:

$$\int_{\Omega} \Big(\lambda \left(\nabla \cdot \boldsymbol{u} \right) \mathbb{I}_{3\times 3} + 2\mu \boldsymbol{\epsilon} \left(\boldsymbol{u} \right) \Big) : \boldsymbol{\epsilon}(\boldsymbol{w}) d\boldsymbol{x} = \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} d\boldsymbol{s} \tag{3.26}$$

which is equivalent to

$$\int_{\Omega} \lambda \left(\nabla \cdot \boldsymbol{u} \right) \cdot \left(\nabla \cdot \boldsymbol{w} \right) + 2\mu \boldsymbol{\epsilon} \left(\boldsymbol{u} \right) : \boldsymbol{\epsilon}(\boldsymbol{w}) d\boldsymbol{x} = \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} d\boldsymbol{s} \tag{3.27}$$

Next, we define the bilinear form $a(\cdot, \cdot)$ by,

$$\begin{aligned} a(\boldsymbol{u}, \boldsymbol{w}) &= \int_{\Omega} \Big(\lambda \left(\nabla \cdot \boldsymbol{u} \right) \cdot \left(\nabla \cdot \boldsymbol{w} \right) + 2\mu \boldsymbol{\epsilon} \left(\boldsymbol{u} \right) : \boldsymbol{\epsilon}(\boldsymbol{w}) \Big) d\boldsymbol{x} \\ &= \int_{\Omega} \left(\lambda \left(\nabla \cdot \boldsymbol{u} \right) \cdot \left(\nabla \cdot \boldsymbol{w} \right) + 2\mu \sum_{i,j=1}^{3} \epsilon_{ij}(\boldsymbol{u}) \epsilon_{ij}(\boldsymbol{w}) \right) d\boldsymbol{x} \\ &= \int_{\Omega} \sum_{i,j=1}^{3} \sigma_{ij}(\boldsymbol{u}) \epsilon_{ij}(\boldsymbol{w}) d\boldsymbol{x} \end{aligned} \tag{3.28}$$

and the linear functional $b(\cdot)$ by:

$$b(\boldsymbol{w}) = \int_{\Omega} \boldsymbol{f} \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N} \boldsymbol{w} \bar{\boldsymbol{t}} d\boldsymbol{s} \tag{3.29}$$

Using the notation of operation theory, eq. (3.27) can be expressed as:

$$a(\boldsymbol{u}, \boldsymbol{w}) = b(\boldsymbol{w}) \tag{3.30}$$

Equations (3.27) and (3.30) are known as the *variational boundary value problem* or the weak formulation. It can be proven that a and b satisfy the conditions for Theorem 2 to hold, and, thus, the problem (P), i.e. $a(\boldsymbol{u}, \boldsymbol{w}) = b(\boldsymbol{w})$ admits a unique solution.

In this context, the request for the relation $a(\boldsymbol{u}, \boldsymbol{w}) = b(\boldsymbol{w})$ to be satisfied for all $\boldsymbol{w} \in \mathscr{W}$ can be interpreted via the *principle of virtual work* with $\frac{1}{2}a(\boldsymbol{u}, \boldsymbol{u})$ being the *strain energy* and $b(\boldsymbol{u})$ the *potential energy of exterior forces.* Generally, these equations cannot be solved analytically and the finite element method constitutes a generic numerical approach to approximate the solutions of (3.17) and its generalizations (nonlinear elasticity, dynamic, etc.).

3.5 The finite element method

In practice, obtaining an exact solution $\boldsymbol{u}$ of eq. (3.25) is not feasible in most applications of interest. Instead, an approximate solution is sought and the Galerkin approximation technique can provide it. The main idea behind this approach is to consider a finite-dimensional subspace $U_h \subseteq U$, spanned by a finite number of basis vectors $\{\boldsymbol{N}_i\}_{i=1}^{\bar{N}}$, and to find the best approximation $\boldsymbol{u}_h \in U_h$ of $\boldsymbol{u} \in U$. Since, $U_h \subseteq U$ and $span\{\boldsymbol{N}_1, ..., \boldsymbol{N}_{\bar{N}}\} = U_h$, then $\boldsymbol{u}_h$ will be a linear combination of the basis vectors, that is $\boldsymbol{u}_h = v_1\boldsymbol{N}_1(\boldsymbol{x}) + v_2\boldsymbol{N}_2(\boldsymbol{x}) + ... + v_{\bar{N}}\boldsymbol{N}_{\bar{N}}(\boldsymbol{x})$. Then, then problem consists in finding these coefficients $\boldsymbol{v} = (v_1, v_2, ..., v_{\bar{N}})^T$ of the linear combination that minimize the error $\|\boldsymbol{u} - \boldsymbol{u}_h\|$.

The Galerkin approximation considers $\boldsymbol{w} = \boldsymbol{N}_i$, $i = 1, ..., \bar{N}$ as test functions and by substituting $\boldsymbol{w}$ in eq. (3.30) one obtains:

$$a(\boldsymbol{N}_i, \sum_{j=1}^{\bar{N}} v_j\boldsymbol{N}_j) = b(\boldsymbol{N}_i), \quad \text{for } i = 1, ..., \bar{N} \tag{3.31}$$

and due to the linearity of a,

$$\sum_{j=1}^{\bar{N}} a(\boldsymbol{N}_i, \boldsymbol{N}_j)v_j = b(\boldsymbol{N}_i), \quad \text{for } i = 1, ..., \bar{N} \tag{3.32}$$

Equation (3.32) describes a linear system of equations of the form:

$$[\boldsymbol{K}][\boldsymbol{v}] = [\boldsymbol{P}] \tag{3.33}$$

where $[\boldsymbol{K}]$ is an $\bar{N} \times \bar{N}$ matrix with element $K_{ij} = a(\boldsymbol{N}_i, \boldsymbol{N}_j)$ and $[\boldsymbol{P}]$ is an $\bar{N} \times 1$ vector with elements $P_i = b(\boldsymbol{N}_i)$. The solution of this system will give the vector of coefficients $\boldsymbol{v}$ and thus the approximate solution $\boldsymbol{u}_h$ to the boundary value problem.

In a similar fashion, the finite element method, which is a variation of the Galerkin method, considers piecewise polynomials as the basis functions $\{N_i\}_{i=1}^n$. These polynomials are supported on small patches of Ω, called *finite elements* and are obtained by discretizing $\Omega \in \mathbb{R}^n$ into polygonal domains. The displacement vector field $\boldsymbol{u}_h \in U^h$ and test functions $\boldsymbol{w} \in W^h$ can be approximated within each element e via the basis functions as:

$$\boldsymbol{u}_h^e = \sum_i^n N_i^e \boldsymbol{v}_i^e = [\boldsymbol{N}^e][\boldsymbol{v}^e] \tag{3.34}$$

$$\boldsymbol{w}^e = \sum_i^n N_i^e \boldsymbol{w}_i^e = [\boldsymbol{N}^e][\boldsymbol{w}^e] \tag{3.35}$$

where $\boldsymbol{v}_i^e, \boldsymbol{w}_i^e$ are 3×1 vectors (case of 3D elasticity), n is the number of nodes in the element e and $[\boldsymbol{N}^e]$ is $3 \times 3n$ matrix with entries

$$[\boldsymbol{N}^e] = [N_1^e \mathbb{I}_{3\times3} \dots N_n^e \mathbb{I}_{3\times3}] \tag{3.36}$$

The strains and displacements within each element are related via the following relation:

$$\{\boldsymbol{\epsilon}(\boldsymbol{u}_h)\} = \sum_{i=1}^n [\boldsymbol{B}_i^e] \boldsymbol{v}_i^e = [\boldsymbol{B}^e][\boldsymbol{v}^e] \tag{3.37}$$

$$\{\boldsymbol{\epsilon}(\boldsymbol{w})\} = \sum_{i=1}^n [\boldsymbol{B}_i^e] \boldsymbol{w}_i^e = [\boldsymbol{B}^e][\boldsymbol{w}^e] \tag{3.38}$$

where the 6×3 matrix $[\boldsymbol{B}_i^e]$ is introduced, such that

$$[\boldsymbol{B}_i^e] = \begin{bmatrix} \frac{\partial N_i^e}{\partial x_1} & 0 & 0 \\ 0 & \frac{\partial N_i^e}{\partial x_2} & 0 \\ 0 & 0 & \frac{\partial N_i^e}{\partial x_3} \\ \frac{\partial N_i^e}{\partial x_2} & \frac{\partial N_i^e}{\partial x_1} & 0 \\ 0 & \frac{\partial N_i^e}{\partial x_3} & \frac{\partial N_i^e}{\partial x_2} \\ \frac{\partial N_i^e}{\partial x_3} & 0 & \frac{\partial N_i^e}{\partial x_1} \end{bmatrix} \tag{3.39}$$

and also

$$[\boldsymbol{B}^e] = [\boldsymbol{B}_1^e \dots \boldsymbol{B}_n^e] \tag{3.40}$$

Eq. (3.25) then must be satisfied for every element Ω^e, or

$$\int_{\Omega^e} \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\boldsymbol{w}) d\boldsymbol{x} = \int_{\Omega^e} \boldsymbol{f} \cdot \boldsymbol{w} d\boldsymbol{x} + \int_{\partial\Omega_N \cap \partial\Omega^e} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} d\boldsymbol{s} + \int_{\partial\Omega^e \setminus \partial\Omega_N} \boldsymbol{w} \cdot \boldsymbol{t}_b d\boldsymbol{s} \tag{3.41}$$

Using matrix notation the above equation can be written as

$$\int_{\Omega^e} \{\boldsymbol{\epsilon}(\boldsymbol{w})\}^T \boldsymbol{D} \{\boldsymbol{\epsilon}(\boldsymbol{u}_h)\} d\boldsymbol{x} = \int_{\Omega^e} \{\boldsymbol{w}\}^T \{\boldsymbol{f}\} d\boldsymbol{x} + \int_{\partial\Omega_N \cap \partial\Omega^e} \{\boldsymbol{w}\}^T \{\bar{\boldsymbol{t}}\} d\boldsymbol{s} + \int_{\partial\Omega^e \setminus \partial\Omega_N} \{\boldsymbol{w}\}^T \{\boldsymbol{t}_b\} d\boldsymbol{s} \tag{3.42}$$

The last term in eq. (3.42) accounts for the tractions developed between the element and its adjacent elements. Inserting eqs. (3.37) and (3.38) into eq. (3.42) gives:

$$\begin{aligned} \int_{\Omega^e} \left([\boldsymbol{B}^e][\boldsymbol{w}^e]\right)^T \boldsymbol{D} \left([\boldsymbol{B}^e][\boldsymbol{v}^e]\right) d\boldsymbol{x} = & \int_{\Omega^e} \left([\boldsymbol{N}^e][\boldsymbol{w}^e]\right)^T \{\boldsymbol{f}\} d\boldsymbol{x} + \int_{\partial\Omega_N \cap \partial\Omega^e} \left([\boldsymbol{N}^e][\boldsymbol{w}^e]\right)^T \{\bar{\boldsymbol{t}}\} d\boldsymbol{s} \\ & + \int_{\partial\Omega^e \setminus \partial\Omega_N} \left([\boldsymbol{N}^e][\boldsymbol{w}^e]\right)^T \{\boldsymbol{t}_b\} d\boldsymbol{s} \end{aligned} \tag{3.43}$$

Since $\boldsymbol{w}^e$ is arbitrary, it can be omitted from the above equation, that is

$$\cancel{[\boldsymbol{w}^e]^T} \int_{\Omega^e} [\boldsymbol{B}^e]^T \boldsymbol{D} [\boldsymbol{B}^e] d\boldsymbol{x} \; [\boldsymbol{v}^e] = \cancel{[\boldsymbol{w}^e]^T} \int_{\Omega^e} [\boldsymbol{N}^e]^T \{\boldsymbol{f}\} d\boldsymbol{x} + \cancel{[\boldsymbol{w}^e]^T} \int_{\partial\Omega_N \cap \partial\Omega^e} [\boldsymbol{N}^e]^T \{\bar{\boldsymbol{t}}\} d\boldsymbol{s}$$
$$+ \cancel{[\boldsymbol{w}^e]^T} \int_{\partial\Omega^e \setminus \partial\Omega_N} [\boldsymbol{N}^e]^T \{\boldsymbol{t}_b\} d\boldsymbol{s} \tag{3.44}$$

In essence, eq. (3.44) represents the $n \times n$ system of linear equations

$$[\boldsymbol{K}^e][\boldsymbol{v}^e] = [\boldsymbol{P}^e] + [\boldsymbol{P}_b^e] \tag{3.45}$$

where

$$[\boldsymbol{K}^e] = \int_{\Omega^e} [\boldsymbol{B}^e]^T \boldsymbol{D} [\boldsymbol{B}^e] d\boldsymbol{x} \tag{3.46}$$
$$[\boldsymbol{P}^e] = \int_{\Omega^e} [\boldsymbol{N}^e]^T \{\boldsymbol{f}\} d\boldsymbol{x} + \int_{\partial\Omega_N \cap \partial\Omega^e} [\boldsymbol{N}^e]^T \{\bar{\boldsymbol{t}}\} d\boldsymbol{s} \tag{3.47}$$
$$[\boldsymbol{P}_b^e] = \int_{\partial\Omega^e \setminus \partial\Omega_N} [\boldsymbol{N}^e]^T \{\boldsymbol{t}_b\} d\boldsymbol{s} \tag{3.48}$$

The vector $[\boldsymbol{P}_b^e]$ represents the force vector between the element e and its adjacent elements and does not need to be calculated since it cancels out when adding all element contributions to form the global system of equations

$$[\boldsymbol{K}][\boldsymbol{v}] = [\boldsymbol{P}] \tag{3.49}$$

The solution of the above linear system will eventually yield the coefficients for the approximate displacement field $\boldsymbol{u}_h$.

4

Surrogate modeling

4.1 Introduction

The great advances in computational technology over the last two decades enabled the use of high-fidelity physics-based mathematical models to describe complex physical systems, arising in all areas of engineering and science. These detailed models are better able to describe the underlying physics of the problem and thus provide accurate representations of the system's behavior. Nevertheless, for many applications such as nonlinear and/or dynamic systems with high-dimensionality, the computational cost can still be considerably high. This cost becomes prohibitive in the case of parametrized systems, which arise in the fields of stochastic analysis, sensitivity analysis or optimization, where a large number of repeated simulations is required. To circumvent this problem, surrogate modes may be employed in order to replace the original model by a function that emulates the complex system's behavior at significantly smaller cost per model evaluation.

To describe the problem in a more concrete setting, let as consider a computational model $\mathbb{M}$, which takes M-dimensional vectors $\mathbf{x}$ as inputs and maps them to $\mathbb{R}^N$:

$$\mathbb{M} : \mathbf{x} \in \mathcal{D} \subset \mathbb{R}^M \mapsto \mathbf{y} \in \mathbb{R}^n \tag{4.1}$$

A surrogate model $\tilde{\mathbb{M}}$ can then be defined as:

$$\tilde{\mathbb{M}} : \mathbf{x} \in \mathscr{D} \subset \mathbb{R}^M \mapsto \mathbf{y} \in \mathbb{R}^n \tag{4.2}$$

such that

$$\tilde{\mathbb{M}} \approx \mathbb{M} \tag{4.3}$$

The construction of the surrogate model can be based on an assumption about the functional shape of $\mathbb{M}$ and/or some partial information obtained from limited runs of the original model. Further, $\mathbb{M}$ is considered as a black box, in the sense that the inner mechanisms of the model are unknown and only the output $\mathbf{y} = \mathbb{M}(\mathbf{x})$ is accessible. As an example, a finite element model constitutes a surrogate model, since it approximates the solution of the system's governing equations that cannot be solved analytically.

Over the last decades a number of approaches has been proposed for constructing surrogate models. Among them the most popular are the polynomial chaos expansions [57, 161, 151], kriging [123], support vector machines [38], radial basis functions [21], space mapping[[11], principal component analysis[117, 75], kernel principal component analysis [137] and artificial neural networks [81, 128, 136]. In this chapter, the focus lies on the PC expansion and a class of machine learning algorithms, namely the radial basis function interpolation, the Nyström extension and the Diffusion Maps algorithm that were used in this thesis.

4.2 Polynomial Chaos expansion

Let us assume that the outcome of the model $\mathbb{M}$ is a random variable Y, which belongs to the Hilbert space $\mathscr{L}^2$ ($\mathbb{E}[Y^2] < \infty$). Thus, Y can be written as a series expansion with respect to set of orthonormal basis functions $\{\Psi_i\}_{i\in\mathbb{N}}$, which constitute the Hilbert basis of the space:

$$Y = \sum_{i\in\mathbb{N}} \hat{c}_i \Psi_i(\mathbf{x}) \tag{4.4}$$

where $\mathbf{x}$ is a random vector representing the input random variables, Ψ_i are the multivariate orthogonal polynomials and $\hat{c}_i$ the corresponding coefficients of the expansion. The type of polynomials is chosen according to the probability measures of the input random vector. For instance, for n-dimensional Gaussian random vectors the n-variate

Hermite polynomials are chosen, which are given by the following recurrent relation:

$$H_n(x_{i_1}, ..., x_{i_n}) = e^{\frac{1}{2}\mathbf{x}^T\mathbf{x}}(-1)^n \frac{\partial^n}{\partial_{i_1}, ..., \partial_{i_n}} e^{\frac{1}{2}\mathbf{x}^T\mathbf{x}} \tag{4.5}$$

Then, Y can be represented as:

$$Y = c_0 H_0 + \sum_{i_1=1}^{\infty} c_{i_1} H_1(x_{i_1}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty} c_{i_1 i_2} H_2(x_{i_1}, x_{i_2}) + ... \tag{4.6}$$

The above equation can be written more conveniently as:

$$Y = \sum_{i=0}^{\infty} \hat{c}_i \Psi_i(\mathbf{x}) \tag{4.7}$$

with $\Psi_i(\mathbf{x})$ being on an one-to-one correspondence with $H_n(x_{i_1}, ..., x_{i_n})$. The polynomial basis $\{\Psi\}$ of Hermite-Chaos forms a complete orthonormal basis, satisfying the following properties:

$$\Psi_0 = 1$$

$$\mathbb{E}[\Psi_i] = 0, \quad i > 0 \tag{4.8}$$

$$\mathbb{E}[\Psi_i \Psi_j] = \langle \Psi_i \Psi_j \rangle = \langle \Psi_i^2 \rangle \delta_{ij} \tag{4.9}$$

where, $\langle \cdot, \cdot \rangle$ denotes the inner product of two random vectors $\mathbf{x}$:

$$\langle f(\mathbf{x}), g(\mathbf{x}) \rangle = \int f(\mathbf{x}) g(\mathbf{x}) d\mu(\mathbf{x}) \tag{4.10}$$

with $d\mu$ being the Gaussian measure:

$$d\mu(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n}} e^{\frac{1}{2}\mathbf{x}^T\mathbf{x}} d\mathbf{x} \tag{4.11}$$

The one-dimensional and two-dimensional Hermite polynomial chaoses are given in tables 4.1 and 4.2, respectively.

Table 4.1: One-Dimensional Polynomial Chaoses

j	order of polynomial chaos	i^{th} Polynomial Chaos Ψ_i	$\langle \Psi_i^2 \rangle$
0	p=0	1	1
1	p=1	x_1	1
2	p=2	$x_1^2 - 1$	2
3	p=3	$x_1^3 - 3x_1$	6

Table 4.2: Two-Dimensional Polynomial Chaoses

j	order of polynomial chaos	i^{th} Polynomial Chaos Ψ_i	$\langle \Psi_i^2 \rangle$
0	p=0	1	1
1	p=1	x_1	1
2		x_2	1
3	p=2	$x_1^2 - 1$	2
4		$x_1^2 - 1$	1
5		$x_1 x_2$	2
6	p=3	$x_1^3 - 3x_1$	6
7		$x_1^2 x_2 - x_2$	2
8		$x_1 x_2^2 - x_1$	2
9		$x_2^3 - 3x_2$	6

4.3 Machine learning algorithms

Machine learning (ML) is an application of artificial intelligence that involves the study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on pattern recognition and inference instead. The process of learning begins with observations or data, such as examples, direct experience, or instructions, in order to look for patterns in data and make better decisions in the future based on the examples that the user provided. The primary aim is to allow the computers to learn automatically without human intervention or assistance and adjust actions accordingly. ML algorithms can be broadly classified into supervised or unsupervised. The algorithms in the first category are able a build a mathematical

model from a set of labeled data, that is, a set which contains both the inputs and the desired outputs. Then, the system can provide targets (predictions) for any new input after sufficient training. In contrast, unsupervised ML algorithms are used in order to construct a mathematical model from a set of data which contains only inputs and no desired output labels. Unsupervised learning is used for discovering patterns in the data and grouping the inputs into categories (feature learning). Other classes of ML algorithms include semi-supervised learning algorithms, active learning algorithms and reinforcement learning algorithms.

It becomes evident from the above that ML algorithms are perfect candidates for constructing surrogate models of parametrized systems. The data set that will be used to train the algorithm may come either from a limited number of runs of the original complex model (simulation-based approach) or some experiments (data-driven approach). In this section, the basic ideas behind three very popular ML algorithm, namely the RBF interpolation, the Nyström extension and the Diffusion Maps algorithm are illustrated.

4.3.1 Radial Basis function interpolation

Radial basis function methods can provide excellent interpolants for a large number of poorly distributed data points. Its ability to handle scarce and unevenly distributed data along with its algebraic simplicity renders this approach particularly popular in the field of machine learning. The implementation steps of this method are illustrated below.

Let $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, be a real function of d variables, and let the values $f(\mathbf{x}_i)$, $i = 1, ..., N$, at the specified locations $\mathbf{x}_i$, $i = 1, ..., N$, be given. The interpolation problem consists in constructing a function $p(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ such that:

$$p(\mathbf{x}_i) = f(\mathbf{x}_i), \quad \text{for } i = 1, ..., N \tag{4.12}$$

An N-dimensional vector space $\mathscr{A}$ of linear functionals on $\mathbb{R}^d$ is considered, which is spanned by a basis $a_j \in \mathscr{A}$, $j = 1, ..., N$. Then, the interpolant will take the following form:

$$p(\mathbf{x}) = \sum_{j=1}^{n} \lambda_j a_j(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d \tag{4.13}$$

Then, using eq. (4.12) the following linear system will arise

$$\sum_{j=1}^{n} \lambda_j a_j(\mathbf{x}_i) = f(\mathbf{x}_i), \quad i = 1, ..., N \tag{4.14}$$

that will determine the coefficients λ_j. These coefficients will be uniquely defined, provided that the matrix $\Phi_{ij} = a_j(\mathbf{x}_i)$, $1 \leq i, j \leq N$, of the system is nonsingular. Unfortunately, the nonsingularity of Φ can not be avoided in the general case, unless specific basis functions are chosen.

In the RBF interpolation the basis functions $\phi(\mathbf{x})$ are symmetric functions whose value depends only on the distance $r = \|\mathbf{x} - \mathbf{x}_i\|$ between the input and some fixed point $\mathbf{x}_i$. Examples of such functions include:

- Gaussian: $\phi(r) = e^{-(\epsilon r)^2}$
- Multiquadric: $\phi(r) = \sqrt{1 + (\epsilon r)^2}$
- Inverse multiquadric: $\phi(r) = \frac{1}{\sqrt{1+(\epsilon r)^2}}$
- Polyharmonic splines of odd degree: $\phi(r) = r^k, \quad k = 1, 3, 5, ...$
- Polyharmonic splines of even degree: $\phi(r) = r^{k-1} ln(r^r), \quad k = 2, 4, 6, ...$

The first three examples require the tuning of a shape parameter ϵ, while polyharmonic splines do not have such a dependence.

For a specific choice of ϕ, the basis functions are given as:

$$a_j(\mathbf{x}) = \phi(\|\mathbf{x} - \mathbf{x}_j\|), \quad j = 1, ..., N \tag{4.15}$$

and the linear system of eq. (4.14) becomes

$$\Phi \boldsymbol{\lambda} = \mathbf{f} \tag{4.16}$$

where $\boldsymbol{\lambda}$ and $\mathbf{f}$ are the vectors in $\mathbb{R}^N$ with components λ_i and $f(\mathbf{x}_i)$, $i = 1, ..., N$. Also, the matrix Φ has elements

$$\Phi_{ij} = \phi(\|\mathbf{x}_i - \mathbf{x}_j\|), \quad 1 \leq i, j \leq N \tag{4.17}$$

Therefore, Φ is a symmetric matrix and in the Gaussian and inverse multiquadric case it is guaranteed to be positive definite, as well.

Solving the system of eq. (4.16) will provide the coefficients $\boldsymbol{\lambda}$ and the interpolant p becomes

$$p(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i \phi(\|\mathbf{x} - \mathbf{x}_i\|), \quad \mathbf{x} \in \mathbb{R}^d \tag{4.18}$$

4.3.2 The Nyström extension

In most applications, it is essential to be able to evaluate the values of a function, which are known only on a given training set, to new samples. A popular technique widely used in the machine learning community when one wants to extend a function outside the range of samples is the Nyström extension [103, 54, 55, 159, 85].

To illustrate this method, let us consider a symmetric kernel $W(\mathbf{x}, \mathbf{y}) = W(\|\mathbf{x} - \mathbf{y}\|)$, i.e. $W(\mathbf{x}, \mathbf{y}) = e^{-\|\mathbf{x}-\mathbf{y}\|^2/\sigma^2}$, where $\mathbf{x}, \mathbf{y} \in \Omega \subset \mathbb{R}^d$. The eigenvalues and eigenfunctions of the kernel on the training data set Ω satisfy the equation:

$$\mu_l \phi_l(\mathbf{x}) = \sum_{\mathbf{y} \in \Omega} W(\mathbf{x}, \mathbf{y}) \phi_l(\mathbf{y}) \tag{4.19}$$

Since the kernel is well-defined in the entire space, then for any $\mathbf{x} \in \mathbb{R}^d$:

$$\bar{\phi}_l(\mathbf{x}) = \frac{1}{\mu_l} \sum_{\mathbf{y} \in \Omega} W(\mathbf{x}, \mathbf{y}) \phi_l(\mathbf{y}) \tag{4.20}$$

which is the Nyström extension of ϕ_l from Ω to $\mathbb{R}^d$. Using the fact that these eigenfunctions form a basis of the set of functions on the training data set, then any function f on Ω can be expressed as

$$f(\mathbf{x}) = \sum_{l} \langle \phi_l, f \rangle \phi_l(\mathbf{x}), \quad \mathbf{x} \in \Omega \tag{4.21}$$

and its extension to $\mathbb{R}^d$ is given by:

$$\bar{f}(\mathbf{x}) = \sum_l \langle \phi_l, f \rangle \bar{\phi}_l(\mathbf{x}), \quad \mathbf{x} \in \mathbf{R}^d \tag{4.22}$$

As a final remark, it should be mentioned that most kernels used in applications are associated with a scale parameter (i.e. σ for the Gaussian kernel) and for distances $\|\mathbf{x} - \mathbf{y}\|$ greater than a multiple of this parameter, the extension goes to zero. This means that the extension can be accurately extended to a distance proportional to the scale parameter, while for greater values the extension is unreliable.

4.3.3 The Diffusion Maps algorithm

The DMAP algorithm is a nonlinear dimensionality reduction method which focuses on discovering the underlying manifold that the data from a data set has been sampled from. This is achieved by computing a family of embeddings of the data set into Euclidean space whose coordinates can be obtained from the eigenvalues and eigenvectors of a diffusion operator on the data set. This algorithm belongs to a subarea of machine learning called manifold learning.

Following the approach described in [35] and more recently in [145], lets assume an ensemble of data (data set) $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N]$ is given, which consists of vectors $\boldsymbol{u}_i \in \mathbb{R}^d,\ i = 1, \cdots, N$. These vectors can be seen as N distinct realizations of an $\mathbb{R}^d$-valued random variable, whose probability distribution is unknown. Also, let $(\boldsymbol{U}, \mathscr{A}, \mu)$ be a measure space, where $\boldsymbol{U}$ is the data set, $\mathscr{A}$ its σ-algebra and μ represents the distribution of the points on $\boldsymbol{U}$. Next, a connectivity measure κ between two data points $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$, is constructed. This measure constitutes a kernel $\kappa : \boldsymbol{U} \times \boldsymbol{U} \to \mathbb{R}$ provided that the following two conditions are met:

- symmetry: $\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) = \kappa(\boldsymbol{u}_j, \boldsymbol{u}_i)$
- non negativity: $\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) \geq 0$

This kernel function provides a notion of similarity and describes the local interaction between the data points. Therefore, its selection should depend on the application at hand, since different kernels capture different features of the data set. Some common

examples of such similarity measures include the Gaussian and the Epanechnikov kernels:

$$\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) = e^{-(\|\boldsymbol{u}_i - \boldsymbol{u}_j\|/\varepsilon)^2} \tag{4.23}$$

and

$$\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) = \left(1 - \|\boldsymbol{u}_i - \boldsymbol{u}_j\|^2/\varepsilon^2\right) \mathbf{1}_{\{\|\boldsymbol{u}_i - \boldsymbol{u}_j\| < \varepsilon\}} \tag{4.24}$$

where, $\mathbf{1}_{\{\|\boldsymbol{u}_i - \boldsymbol{u}_j\| < \varepsilon\}}$ is an indicator function such that

$$\mathbf{1}_{\{\|\boldsymbol{u}_i - \boldsymbol{u}_j\| < \varepsilon\}} = \begin{cases} 1, & \text{if } \{\boldsymbol{u}_i - \boldsymbol{u}_j\| < \varepsilon\} \\ 0, & \text{otherwise} \end{cases} \tag{4.25}$$

It is evident that the proximity of the data points is affected by the choice of the parameter ε and its value is user-defined. For the case of the Gaussian kernel, a method to obtain reasonable values for ε has been proposed in [144], where the function $M(\varepsilon)$, defined as,

$$M(\varepsilon) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j; \varepsilon) \tag{4.26}$$

is computed and plotted on a log-log plot. A suitable value for ε is selected arbitrarily from the linear range of the plot.

Given $(\boldsymbol{U}, \kappa)$, a reversible Markov Chain can be constructed on $\boldsymbol{U}$ via a process known as graph Laplacian construction [31]. First the diffusion matrix $\boldsymbol{L}$ is constructed as:

$$L_{ij} = \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) \ , \quad i \text{ and } j \in \{1, ..., N\} \tag{4.27}$$

which is a symmetric matrix in $\mathbb{R}^{N \times N}$. Next, the anisotropic kernel is defined as

$$\hat{\kappa}_\alpha(\boldsymbol{u}_i, \boldsymbol{u}_j) = \frac{\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j)}{\left(\sum_{l=1}^{N} \kappa(\boldsymbol{u}_i, \boldsymbol{u}_l)\right)^\alpha \left(\sum_{l=1}^{N} \kappa(\boldsymbol{u}_j, \boldsymbol{u}_l)\right)^\alpha} \ , \quad i \text{ and } j \in \{1, ..., N\} \tag{4.28}$$

The corresponding anisotropic diffusion matrix $\boldsymbol{L}^{(\alpha)}$ is given by:

$$L_{ij}^{(\alpha)} = \hat{\kappa}_\alpha(\boldsymbol{u}_i, \boldsymbol{u}_j) = \frac{L_{ij}}{q_i^\alpha q_j^\alpha} \ , \quad \text{with} \ \ q_i = \sum_{l=1}^{N} L_{il} \tag{4.29}$$

or, in matrix form as:

$$\boldsymbol{L}^{(\alpha)} = \boldsymbol{D}^{-\alpha} \boldsymbol{L} \boldsymbol{D}^{-\alpha} \tag{4.30}$$

where $\boldsymbol{D}$ is the diagonal matrix such that:

$$D_{ii} = \sum_{q=1}^{N} L_{iq} \tag{4.31}$$

The choice of the parameter α provides a trade-off between the influence of the density of the points and that of the geometry of the manifold they are assumed to live on, over the spectrum of the diffusion. As stated in [35], when $\alpha = 0$, the weights of the graph given by the kernel κ are considered isotropic and the diffusion becomes the classical graph Laplacian normalization. In this case the influence of the density on the eigenvalues and eigenfunctions of the spectrum is maximal. On the other hand, when $\alpha = 1$, the diffusion approximates the Laplace-Beltrami operator and effectively provides a description of the Riemannian geometry of the data set, without taking into account the distribution of the points.

Proceeding with the formulation, the diagonal matrix $\boldsymbol{D}^{(\alpha)}$ is also introduced, given by

$$D_{ii}^{(\alpha)} = \sum_{q=1}^{N} L_{iq}^{(\alpha)} \tag{4.32}$$

Then, the matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$ defined as:

$$\mathbf{P} = \left(\mathbf{D}^{(\alpha)}\right)^{-1} \boldsymbol{L}^{(\alpha)} \tag{4.33}$$

will be a transition matrix in the sense that it has positive entries that satisfy the relation $\sum_{j=1}^{N} P_{ij} = 1$ for all $i = 1, ..., N$. Its elements P_{ij} express the probability of transition from point $\boldsymbol{u}_i$ to point $\boldsymbol{u}_j$ in one time step. Consequently, the probability of transition from $\boldsymbol{u}_i$ to $\boldsymbol{u}_j$ in t time steps is represented by the t-th power $\mathbf{P}^t$ of $\mathbf{P}$, with t termed

as the "diffusion time". Next, from the spectral analysis of $\boldsymbol{P}$ the n largest eigenvalues λ_k for $k = 0, ..., n$ and corresponding right eigenvectors $\{\boldsymbol{\psi}^k\}_{k=0}^n$ are computed and the family of "diffusion-maps" $\{\boldsymbol{\Psi}_t\}_{t\in\mathbb{N}}$ is defined as:

$$\boldsymbol{\Psi}_t(\boldsymbol{u}_i) \triangleq \left[\lambda_1^t \psi_i^1, \lambda_2^t \psi_i^2, ..., \lambda_n^t \psi_i^n\right] \tag{4.34}$$

where the eigenvector $\boldsymbol{\psi}^0$ is omitted in the above equation since it is constant, as a consequence of $\mathbf{P}$ being a transition matrix.

The mapping $\boldsymbol{\Psi}_t : \boldsymbol{U} \to \mathbb{R}^n$ embeds the data set into an n-dimensional Euclidean space, henceforth referred to as the Diffusion Maps space (DMAP space) and each component of $\boldsymbol{\Psi}_t(\boldsymbol{u_i})$ is referred to as the diffusion coordinate (DMAP coordinate). Collecting the DMAP coordinates of all the data points will lead to the construction of the diffusion maps basis $\hat{\boldsymbol{\Psi}}_t \in \mathbb{R}^{N\times n}$:

$$\hat{\boldsymbol{\Psi}}_t = \begin{bmatrix} \boldsymbol{\Psi}_t(\boldsymbol{u}_1) \\ \boldsymbol{\Psi}_t(\boldsymbol{u}_2) \\ \vdots \\ \boldsymbol{\Psi}_t(\boldsymbol{u}_N) \end{bmatrix} = \left[\lambda_1^t \boldsymbol{\psi}^1, \lambda_2^t \boldsymbol{\psi}^2, ..., \lambda_n^t \boldsymbol{\psi}^n\right] \tag{4.35}$$

It becomes evident that in eq. (4.35) the i-th row of $\hat{\boldsymbol{\Psi}}_t$ is the n-dimensional representation in the DMAP space of a particular d-dimensional vector $\boldsymbol{u}_i$ in the data set $\boldsymbol{U}$. Note that in the case $n = N$, $\hat{\boldsymbol{\Psi}}_t$ constitutes an algebraic basis of $\mathbb{R}^N$, while for $n < N$, $\hat{\boldsymbol{\Psi}}_t$ spans a subspace of dimension n of $\mathbb{R}^N$.

The DMAP space is equipped with a metric δ, which gives the diffusion distance between any two points $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ on the manifold. This distance is evaluated using the Euclidean distance between their diffusion coordinates $\boldsymbol{\Psi}_t(\boldsymbol{u}_i)$ and $\boldsymbol{\Psi}_t(\boldsymbol{u}_j)$, that is

$$\delta_t^2(\boldsymbol{u}_i, \boldsymbol{u}_j) = \sum_{l=1}^{n} (\lambda_l^t)^2 \left(\psi^l(\boldsymbol{u}_i) - \psi^l(\boldsymbol{u}_j)\right)^2 = \|\boldsymbol{\Psi}_t(\boldsymbol{u}_i) - \boldsymbol{\Psi}_t(\boldsymbol{u}_j)\|^2 \tag{4.36}$$

In order to simplify the notation the subscript t will be omitted in the subsequent formulation and the DMAP coordinates of a vector $\boldsymbol{u}_i$ will be denoted with $\boldsymbol{z}_i$, viz. $\boldsymbol{\Psi}(\boldsymbol{u}_i) = \boldsymbol{z}_i \in \mathbb{R}^n$. From a machine learning point of view, the diffusion distance possesses some very interesting features. More specifically, points that are closer with respect to the DMAP metric are those that are highly connected in the graph. This property provides a means for clustering the data, which can be used for the design of inference

algorithms in the context of machine learning.

Lastly, it should be mentioned that in certain applications the elements of a vector in the given data set might represent different physical quantities and differ by orders of magnitude. Since the DMAP algorithm relies on pairwise distances, in such cases it is advantageous to normalize first the given data set. Some popular approaches for this purpose are the $min - max$ normalization and the z-score normalization, which is presented here. Let $\boldsymbol{U}^0$ be the initial unnormalized data set. Then, the entries of the normalized version of the data set are given by the following relation:

$$u_{ij} = \frac{u_{ij}^0 - \mu_i}{\sigma_i} \quad \text{for } i = 1, ..., d \text{ and } j = 1, ..., N \tag{4.37}$$

where μ_i and σ_i are the statistical mean and standard deviation over all elements in the i-th line of $\boldsymbol{U}^0$.

5

The Stochastic Finite Element Method

5.1 Introduction

In an effort to achieve more reliable representations in real-life engineering problems, researchers quickly realized that this could not be accomplished without taking into account the role of uncertainties in the system under consideration. This notion led to the development of the stochastic finite element method (SFEM) and in the past few years significant advances have been accomplished in its application to structural engineering problems.

In this chapter, the derivation of the stochastic equations of equilibrium is illustrated. Next, the basic principles behind some of the most eminent stochastic methods for the solution of these equations are revisited. More specifically, the fundamentals of the Monte Carlo simulation, the Spectral Stochastic Finite Element Method and Probability Density Evolution Method are presented.

5.2 Stochastic equations of equilibrium

In the formulation of the finite element method presented in chapter 3, it was shown that, for a body Ω with Dirichlet boundary conditions on $\partial\Omega_D$ and Neumann boundary conditions on $\partial\Omega_N$, the displacement field within the body can be obtained via

appropriate interpolation from its nodal values, that is

$$\boldsymbol{u}_h^e = [\boldsymbol{N}^e][\boldsymbol{v}^e] \tag{5.1}$$

with $[\boldsymbol{N}^e]$ being the matrix of shape functions and $[\boldsymbol{v}^e]$ the nodal values of the displacement field. Then, from eqs. (3.39), (3.40) and (3.13) the strain and stress vectors are given by

$$\{\epsilon\} = [\boldsymbol{B}^e][\boldsymbol{v}^e] \tag{5.2}$$

and

$$\{\sigma\} = \boldsymbol{D}\{\epsilon\} = \boldsymbol{D}[\boldsymbol{B}^e][\boldsymbol{v}^e] \tag{5.3}$$

where $[\boldsymbol{B}^e]$ is the matrix that relates the components of strains to the nodal displacements and its elements are obtained by differentiation of the element shape functions. Also, $\boldsymbol{D}$ is the matrix describing the constitutive properties of the material. In the case these properties are random, i.e. they vary in time and/or in space and their variation can be described in the general case with a n-dimensional multivariate stochastic process, the constitutive matrix $\boldsymbol{D}$ becomes random and can be written in the form:

$$\boldsymbol{D}(\boldsymbol{x}, \theta) = \boldsymbol{D}_0 \cdot \big[1 + X(\boldsymbol{x}, \theta)\big] \tag{5.4}$$

where $\boldsymbol{x}$ is the position vector of the material points in Ω, $\boldsymbol{D}_0$ is the matrix containing the mean value of the material properties and $X(\boldsymbol{x}, \theta)$ is a zero-mean stationary stochastic process indicating the fluctuations around $\boldsymbol{D}_0$. Introducing the nodal force vector $\boldsymbol{P}^{(e)}$, from eq. (3.44) we get:

$$\int_{\Omega^e} \{\boldsymbol{\epsilon}^e\}^T \{\boldsymbol{\sigma}\} d\boldsymbol{x} = \int_{\Omega^e} [\boldsymbol{N}^e]^T \{\boldsymbol{f}\} d\boldsymbol{x} + \int_{\partial\Omega_N \cap \partial\Omega^e} [\boldsymbol{N}^e]^T \{\bar{\boldsymbol{t}}\} d\boldsymbol{s} + \int_{\partial\Omega^e \setminus \partial\Omega_N} [\boldsymbol{N}^e]^T \{\boldsymbol{t}_b\} d\boldsymbol{s} \Longleftrightarrow$$
$$\int_{\Omega^e} [\boldsymbol{B}^e]^T \boldsymbol{D} [\boldsymbol{B}^e] d\boldsymbol{x} \cdot [\boldsymbol{v}^e] = [\boldsymbol{P}^e] \tag{5.5}$$

Substituting eq.(5.4) to eq.(5.5) we can write:

$$\int_{\Omega^e} [\boldsymbol{B}^e]^T \boldsymbol{D}_0 \cdot \big[1 + X(\boldsymbol{x}, \theta)\big] [\boldsymbol{B}^e] d\boldsymbol{x} \cdot [\boldsymbol{v}^e] = [\boldsymbol{P}^e] \tag{5.6}$$

or

$$\left(\int_{\Omega^e}[\boldsymbol{B}^e]^T\boldsymbol{D}_0[\boldsymbol{B}^e]d\boldsymbol{x}+\int_{\Omega^e}[\boldsymbol{B}^e]^T\boldsymbol{D}_0X(\boldsymbol{x},\theta)[\boldsymbol{B}^e]d\boldsymbol{x}\right)\cdot[\boldsymbol{v}^e]=[\boldsymbol{P}^e] \tag{5.7}$$

The stiffness matrix associated with a given element of volume Ω_e reads :

$$\boldsymbol{K}^e=\int_{\Omega^e}[\boldsymbol{B}^e]^T\boldsymbol{D}_0[\boldsymbol{B}^e]d\boldsymbol{x}+\int_{\Omega^e}[\boldsymbol{B}^e]^T\boldsymbol{D}_0X(\boldsymbol{x},\theta)[\boldsymbol{B}^e]d\boldsymbol{x} \tag{5.8}$$

or

$$\boldsymbol{K}^e=\boldsymbol{K}_0^e+\Delta\boldsymbol{K}^e \tag{5.9}$$

where $\boldsymbol{K}_0^e=\int_{\Omega^e}[\boldsymbol{B}^e]^T\boldsymbol{D}_0[\boldsymbol{B}^e]d\boldsymbol{x}$ and $\Delta\boldsymbol{K}^e=\int_{\Omega^e}[\boldsymbol{B}^e]^T\boldsymbol{D}_0X(\boldsymbol{x},\theta)[\boldsymbol{B}^e]d\boldsymbol{x}$ are the deterministic and stochastic parts of the stochastic stiffness matrix, respectively. The global stochastic stiffness matrix can be expressed in the same manner as:

$$\boldsymbol{K}=\sum^{N_e}\boldsymbol{K}^e=\boldsymbol{K}_0+\Delta\boldsymbol{K} \tag{5.10}$$

where N_e is the total number of elements. Stochastic finite element analysis is based on the solution of the following equation:

$$\boldsymbol{P}=(\boldsymbol{K}_0+\Delta\boldsymbol{K})\cdot\mathbf{U} \tag{5.11}$$

where $\mathbf{P}$ is the external force vector (which may also be random) and $\mathbf{U}$ the nodal displacement vector.

5.3 The Monte Carlo simulation

The most commonly used method for the solution of eq.(5.11) is the non-intrusive Monte Carlo approximation. According to this approach, N_{sim} samples of the stochastic stiffness matrix $\Delta\mathbf{K}$ are generated and the deterministic problem of eq. (5.11) is solved N_{sim} times, generating a population for the response vector $\{\mathbf{U}_j\}$ with $j=1,\ldots,N_{sim}$. The variance of $\mathbf{U}$ is estimated after a statistical post-process of the population. Let u_i be the response of the degree of freedom (i). Its mean value and standard deviation are given by

$$\mathbb{E}[u_i] = \frac{1}{N_{sim}} \sum_{j=1}^{N_{sim}} u_i(j) \tag{5.12}$$

$$\sigma^2(u_i) = \frac{1}{N_{sim}-1} \sum_{j=1}^{N_{sim}} \left[u_i^2(j) - N_{sim} \cdot (\mathbb{E}[u_i])^2\right] \tag{5.13}$$

In a similar manner, higher order moments of response quantities can be estimated and finally, the entire pdf can be constructed from the statistical post-process of the N_{sim} samples.

5.4 The Spectral Stochastic Finite Element Method

The spectral stochastic finite element method is an extension of the deterministic finite element method for problems involving random material properties and\or loading conditions. For instance, if a material property, such as the modulus of elasticity is considered to be of stochastic nature, that is, $E = E(\theta)$, with θ denoting the random event, then the system's response will also be stochastic, $\mathbf{U} = \mathbf{U}(\theta)$. This implies that both E and $\mathbf{U}$ are elements of the $\mathscr{L}^2(\Theta, \mathscr{F}, \mathscr{P})$ probability space, and thus they can be projected onto appropriate bases of this space.

5.4.1 SSFEM for linear systems

In the SSFEM approach, it is preferred to use the Karhunen-Loéve expansion when the stochastic input parameters are Gaussian random fields and the Polynomial Chaos expansion for the system's response. More particularly, the truncated KL expansion for a random field $a(\mathbf{x}, \theta)$, describing a stochastic input parameter, with given covariance function $C(\mathbf{x}_1, \mathbf{x}_2)$ is defined as

$$\hat{a}(\mathbf{x}, \theta) = a_0(\mathbf{x}) + \sum_{j=1}^{M} \sqrt{\lambda_j} \xi_j(\theta) \varphi_j(\mathbf{x}) \tag{5.14}$$

where, λ_j are the eigenvalues of C, $\varphi_j(\mathbf{x})$ are the associated eigenfunctions and $\xi_j(\theta)$ are independent standard normal variables.

Using conventional notation, the equation of equilibrium of a linear static system of size $N \times N$ reads

$$\boldsymbol{K}\mathbf{U} = \mathbf{F} \tag{5.15}$$

In the above equation, $\mathbf{U}$ is the $N \times 1$ vector of the unknown displacements, resulting from the finite element discretization, $\mathbf{F}$ the corresponding loading vector and $\boldsymbol{K}$ the $N \times N$ global stiffness matrix, which is obtained after assembling the element stiffness matrices $\boldsymbol{K}^e$:

$$\boldsymbol{K}^e = \int_{\Omega_e} \boldsymbol{B}^T \boldsymbol{D} \boldsymbol{B} d\Omega_e \tag{5.16}$$

where, $\boldsymbol{D}$ and $\boldsymbol{B}$ stand for the elasticity and the constitutive matrix, respectively.

Let us consider now the case the modulus of elasticity E is a Gaussian random field. Then, the elasticity matrix in point $\mathbf{x}$ can be written as:

$$\boldsymbol{D}(\mathbf{x}, \theta) \equiv E(\mathbf{x}, \theta) \boldsymbol{D}_0 \tag{5.17}$$

with $\boldsymbol{D}_0$ being a constant matrix. The Karhunen-Loéve expansion of E reads:

$$E(\mathbf{x}, \theta) = E_0(\mathbf{x}) + \sum_{j=1}^{M} \sqrt{\lambda_j} \xi_j(\theta) \varphi_j(\mathbf{x}) \tag{5.18}$$

where, $E_0(\mathbf{x})$ is the mean value of the field in point $\mathbf{x}$. Using eqs. (5.17) and (5.18), eq. (5.16) takes the form:

$$\boldsymbol{K}^e(\theta) = \boldsymbol{K}_0^e + \sum_{j=1}^{\infty} \boldsymbol{K}_j^e \xi_j(\theta) \tag{5.19}$$

In the above equation, $\boldsymbol{K}_0^e$ is the element's mean stiffness matrix and $\boldsymbol{K}_j^e$ are deterministic matrices obtained by:

$$\boldsymbol{K}_j^e = \sqrt{\lambda_j} \int_{\Omega_e} \varphi_j(\mathbf{x}) \boldsymbol{B}^T \boldsymbol{D}_0 \boldsymbol{B} d\Omega_e \tag{5.20}$$

Assembling the above element contributions in a manner similar to the deterministic

case, the equation of equilibrium for the stochastic case becomes

$$\Big(\boldsymbol{K}_0 + \sum_{j=1}^{M} \boldsymbol{K}_j \xi_j(\theta)\Big)\mathbf{U}(\theta) = \mathbf{F} \tag{5.21}$$

$\boldsymbol{K}_0$ is the mean global stiffness matrix and $\boldsymbol{K}_j$ are deterministic matrices obtained by the assembly of $\boldsymbol{K}_j^e$.

Next, the system's response is projected in a PC basis as follows

$$\mathbf{U}(\theta) = \sum_{k=0}^{P-1} \mathbf{U}_k \Psi_k(\boldsymbol{\xi}) \tag{5.22}$$

where, $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{P-1} = \{\Psi_k((\xi_1(\theta), ..., \xi_M(\theta)\}_{k=0}^{P-1}$ is the PC basis, consisting of the M-dimensional zero mean and orthogonal Hermite polynomials of order p. The order of P in Eq. (5.22) is determined by the following formula

$$P = \sum_{q=0}^{p} \binom{M+q-1}{q} \tag{5.23}$$

It should be noted that other polynomials from the Askey tree can also be used, depending on the random variables' type used to represent the field, leading to the so-called Generalized Polynomial Chaos expansion [161]. By denoting $\xi_0(\theta) \equiv 1$, eq. (5.21) can be more compactly written as

$$\left(\sum_{j=0}^{M} \boldsymbol{K}_j \xi_j(\theta)\right)\left(\sum_{k=0}^{P-1} \mathbf{U}_k \Psi_k(\theta)\right) = \mathbf{F} \tag{5.24}$$

with the following residual due to the truncation error

$$\epsilon_{M,P} = \sum_{j=0}^{M}\sum_{k=0}^{P-1} \boldsymbol{K}_j \mathbf{U}_k \xi_j(\theta)\Psi_k(\theta) - \mathbf{F} \tag{5.25}$$

Galerkin minimization of the residual in the mean square sense leads to

$$\mathbb{E}[\epsilon_{M,P} \cdot \Psi_l] = 0 \quad l = 0, 1, ..., P-1 \tag{5.26}$$

Introducing the notation

$$c_{jkl} = \mathbb{E}[\xi_j \Psi_k \Psi_l] \tag{5.27}$$

$$\mathbf{F}_l = \mathbb{E}[\Psi_l \mathbf{F}] \tag{5.28}$$

eq. (5.24) can be rewritten as

$$\sum_{j=0}^{M} \sum_{k=0}^{P-1} c_{jkl} \boldsymbol{K}_j \mathbf{U}_k = \mathbf{F}_l \tag{5.29}$$

and denoting

$$\boldsymbol{K}_{kl} = \sum_{j=0}^{M} c_{jkl} \boldsymbol{K}_j \tag{5.30}$$

eq. (5.29) rewrites as

$$\sum_{k=0}^{P-1} \boldsymbol{K}_{kl} \mathbf{U}_k = \mathbf{F}_l \tag{5.31}$$

In the above equation, each $\mathbf{U}_k$ is a N-dimensional vector and each $\boldsymbol{K}_{kl}$ is a matrix of size $N \times N$. The P different equations can be cast in a matrix form of size $NP \times NP$ as follows:

$$\bar{\boldsymbol{K}} \bar{\mathbf{U}} = \bar{\mathbf{F}} \tag{5.32}$$

where,

$$\bar{\boldsymbol{K}} = \begin{bmatrix} \sum_{j=0}^{M} c_{j,0,0} \boldsymbol{K}_j & \sum_{j=0}^{M} c_{j,1,0} \boldsymbol{K}_j & \cdots & \sum_{j=0}^{M} c_{j,P-1,0} \boldsymbol{K}_j \\ \sum_{j=0}^{M} c_{j,0,1} \boldsymbol{K}_j & \sum_{j=0}^{M} c_{j,1,1} \boldsymbol{K}_j & \cdots & \sum_{j=0}^{M} c_{j,P-1,1} \boldsymbol{K}_j \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=0}^{M} c_{j,0,P-1} \boldsymbol{K}_j & \sum_{j=0}^{M} c_{j,1,P-1} \boldsymbol{K}_j & \cdots & \sum_{j=0}^{M} c_{j,P-1,P-1} \boldsymbol{K}_j \end{bmatrix} \tag{5.33}$$

and

$$\bar{\mathbf{U}} = [\mathbf{U}_0, \mathbf{U}_1, ..., \mathbf{U}_{P-1}]^T \tag{5.34}$$

$$\bar{\mathbf{F}} = [\mathbf{F}_0, \mathbf{F}_1, ..., \mathbf{F}_{P-1}]^T \tag{5.35}$$

It should be noted here that, in the case that $\mathbf{F}$ is deterministic

$$\mathbf{F}_k = \{\mathbf{0}\} \quad k > 0 \tag{5.36}$$

Solving the augmented system of equations in (5.32), will give the PC expansion coefficients $\mathbf{U}_k$, which, in turn, will produce the final solution

$$\mathbf{U}(\theta) = \sum_{k=0}^{P-1} \mathbf{U}_k \Psi_k(\theta) \tag{5.37}$$

5.4.2 SSFEM for linear systems with non-Gaussian parameters

In the formulation presented in the previous section, Gaussian random fields were considered for the modeling of the system's random parameters. This assumption facilitates computations, however, it can be applied only to problems with small coefficient of variation (∼15-20%), since it can be physically inconsistent in cases where the random parameters are not allowed to take negative values. Such is the case of the modulus of elasticity, where negative values of E would result in a displacement vector of infinite variance, and hence the solution would not belong to the $\mathscr{L}^2(\Theta, \mathscr{F}, \mathscr{P})$ probability space. As a consequence the PCE approximation is not convergent in the $\mathscr{L}^2$ sense. An approach to overcome this drawback is presented in [56], where the PCE is used to represent the material property instead of the KL expansion, and will be briefly illustrated here.

The main idea in this approach is to identify an underlying Gaussian field for the corresponding non-Gaussian field. Then, all the equations in the previous section remain valid under some minor modifications. More specifically, the coefficients c_{jkl} in eq. (5.27) will now be given by the relation:

$$c_{jkl} = \mathbb{E}[\Psi_j \Psi_k \Psi_l] \tag{5.38}$$

and the upper bound in the summations in the previous equations will become P instead of M.

To further illustrate this process, let us consider the case where E is represented by a lognormal random field, typically used in most applications. In this case, the field can be represented by the exponential of an underlying Gaussian field, denoted as $E^G(\mathbf{x}, \theta)$, which has the following KL expansion:

$$E^G(\mathbf{x}, \theta) = E_0^G(\mathbf{x}) + \sum_{j=1}^{M} \sqrt{\lambda_j} \xi_j(\theta) \varphi_j(\mathbf{x}) \tag{5.39}$$

Then, $E\left(\mathbf{x}, \theta\right) = exp(E^G(\mathbf{x}, \theta))$ and can be represented using the PCE as:

$$E\left(\mathbf{x}, \theta\right) = exp(E^G(\mathbf{x}, \theta)) = \sum_{j=0}^{P-1} e_j(\mathbf{x}) \Psi_j(\theta) \tag{5.40}$$

The values of the coefficients $e_j(\mathbf{x})$ can be obtained using the orthogonality property of the Ψ_j variables, that is

$$e_j(\mathbf{x}) = \frac{\mathbb{E}\left[\Psi_j E\left(\mathbf{x}, \theta\right)\right]}{\mathbb{E}\left[\Psi_j^2\right]} \tag{5.41}$$

For the calculation of the numerator in the above equation, closed form relations can be found in [56].

5.5 The Probability Density Evolution Method

5.5.1 The principle of preservation of probability

Let us consider a multi-dimensional stochastic dynamical system

$$\dot{\mathbf{Y}}(t) = \mathbf{A}\left[\mathbf{Y}(t), t\right] + \mathbf{B}\left[\mathbf{Y}(t), t\right] \boldsymbol{\xi}(t), \quad \mathbf{Y}(t_0) = \mathbf{Y}_0 \tag{5.42}$$

where $\mathbf{Y} = (Y_1, Y_2, ..., Y_m)^T$ is the m-th order state vector, $\mathbf{A} = (A_1, A_2, ...A_m)^T$ is the m-order linear or nonlinear operator, $\mathbf{B} = [B_{ij}]$ the m by r load influence matrix, $\boldsymbol{\xi}(t) = (\xi_1(t), \xi_2(t), ..., \xi_r(t))^T$ is the r-dimensional excitation vector and $\mathbf{Y}_0$ is the deterministic or random initial value vector.

If the stochastic excitations are white noise processes, then eq. (5.42) can be written

as an equivalent Itô stochastic differential equation

$$d\mathbf{Y}(t) = \mathbf{A}\left[\mathbf{Y}(t), t\right] dt + \mathbf{B}\left[\mathbf{Y}(t), t\right] d\mathbf{W}(t), \quad \mathbf{Y}(t_0) = \mathbf{Y}_0 \tag{5.43}$$

with $\mathbf{W}(t) = (W_1(t), W_2(t), ..., W_r(t))^T$ is an r-dimensional Wiener process vector, whose increments satisfy the Itô rules, that is,

$$\mathbb{E}\left[d\mathbf{W}(t)\right] = 0 \tag{5.44}$$

$$\mathbb{E}\left[d\mathbf{W}(t) d\mathbf{W}^T(t)\right] = \mathbf{D}dt \tag{5.45}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator and $\mathbf{D}$ is an $r \times r$ constant, symmetric positive definite matrix.

According to the principle of preservation of probability, *in any arbitrary fixed domain in the state space, the increment of probability during a time interval is equal to the probability transited through the boundary during the same time interval* [27]. This definition is similar in nature to the conservation of mass theorem in continuum mechanics.

Using some mathematical formalism, the aforementioned statement can be written as:

$$\Delta P_{\mathcal{D}} = \Delta P_{\partial\mathcal{D}} \tag{5.46}$$

where $\mathcal{D}$ is an arbitrary domain and $\partial\mathcal{D}$ its boundary. $\Delta P_{\mathcal{D}}$ denotes the change in probability over $\mathcal{D}$ for the time interval $[t, t + \Delta t]$ and is equal to:

$$\begin{aligned} \Delta P_{\mathcal{D}} &= \int_{\mathcal{D}} p(\boldsymbol{y}, t + \Delta t \mid \boldsymbol{y}_0, t_0) d\boldsymbol{y} - \int_{\mathcal{D}} p(\boldsymbol{y}, t \mid \boldsymbol{y}_0, t_0) d\boldsymbol{y} \\ &= \int_{\mathcal{D}} \left(\frac{\partial p}{\partial t} \Delta t \right) d\boldsymbol{y} + o(\Delta t) \end{aligned} \tag{5.47}$$

where $p(\boldsymbol{y}, t \mid \boldsymbol{y}_0, t_0)$ is the transition probability function and $o(\Delta t)$ the infinitesimal quantity.

In eq. (5.46) $\Delta P_{\partial\mathcal{D}}$ denotes the probability transiting through the boundary of the domain, which can be broken into two parts:$\Delta P_{\partial\mathcal{D}}^{drift}$ due to drift and $\Delta P_{\partial\mathcal{D}}^{diffusion}$ due to the diffusion. These quantities are given by the following formulas:

$$\Delta P_{\partial\mathcal{D}}^{drift} = -\int_{\mathcal{D}} \sum_{l=1}^{m} \frac{\partial \left[A_l\left(\boldsymbol{y},t\right)p\right]}{\partial y_l} d\boldsymbol{y}\Delta t + o(\Delta t) \tag{5.48}$$

$$\Delta P_{\partial\mathcal{D}}^{diffusion} = \int_{\mathcal{D}} \frac{1}{2} \sum_{l=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 \left(\sigma_{kl} p\right)}{\partial y_l \partial y_k} d\boldsymbol{y}\Delta t + o(\Delta t) \tag{5.49}$$

with $\boldsymbol{\sigma} = [\sigma_{ij}] = \mathbf{B}(\boldsymbol{y},t)\mathbf{D}\mathbf{B}(\boldsymbol{y},t)^T$ is the $m \times m$ diffusion matrix.

Inserting eqs. (5.47),(5.48) and (5.49) into eq. (5.46), dividing both sides by Δt and letting $\Delta t \to 0$ yields

$$\begin{aligned}\int_{\mathcal{D}} \frac{\partial p}{\partial t} d\boldsymbol{y} = &-\int_{\mathcal{D}} \sum_{l=1}^{m} \frac{\partial \left[A_l\left(\boldsymbol{y},t\right)p\right]}{\partial y_l} d\boldsymbol{y} \\ &+ \int_{\mathcal{D}} \frac{1}{2} \sum_{l=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 \left(\sigma_{kl} p\right)}{\partial y_l \partial y_k} d\boldsymbol{y}\end{aligned} \tag{5.50}$$

Since the domain $\mathcal{D}$ is arbitrarily chosen, then the integrand quantities in the above equation must always be equal to zero, namely,

$$\frac{\partial p}{\partial t} + \sum_{l=1}^{m} \frac{\partial \left(A_l p\right)}{\partial y_l} - \frac{1}{2} \sum_{l=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 \left(\sigma_{kl} p\right)}{\partial y_l \partial y_k} = 0 \tag{5.51}$$

In the special case that $\mathbf{D} = \mathbf{0}$ the above equation becomes:

$$\frac{\partial p}{\partial t} + \sum_{l=1}^{m} \frac{\partial \left(A_l p\right)}{\partial y_l} = 0 \tag{5.52}$$

Eqs. (5.51) and (5.52) are the celebrated Fokker Planck and Liouville equations, respectively. These partial differential equations give the probability density functions of the system at any time instant t, provided that the initial conditions are known. From a computational point of view, these equations, despite their elegance, can only be solved for small scale systems. Indeed, upon inspection of these equations it becomes evident that for large scale finite element models with many degrees of freedom ($m >> 1$) solving them is impractical (if not intractable) even with the most advanced numerical tools due to the large number of terms in the equations.

5.5.2 Fundamentals of the probability density evolution method

The principle of preservation of probability states that, the inflows, outflows and change in storage of probability in a control volume of the system must be in balance. As shown in the previous section, the above statement is mathematically expressed in the probability density evolution equations, namely the Liouville and the Fokker-Planck equations, which can only be solved for systems with few degrees of freedom. It wasn't until the past few years, that a new perspective came up [92, 27], that of the random event description and the above equation was re-derived, leading to a new family of generalized density evolution equations. Consider, without loss of generality, an n-dimensional stochastic dynamic system, whose equation of motion reads:

$$\mathbf{M}(\boldsymbol{\theta})\ddot{\mathbf{u}} + \mathbf{C}(\boldsymbol{\theta})\dot{\mathbf{u}} + \mathbf{f}(\boldsymbol{\theta}, \mathbf{u}) = \mathbf{F}(\boldsymbol{\theta}, t) \tag{5.53}$$

where $\mathbf{M}, \mathbf{C}, \mathbf{f}, \mathbf{F}$ are the mass and damping matrices, the restoring force vector and the excitation vector respectively; $\ddot{\mathbf{u}}$,$\dot{\mathbf{u}}$,$\mathbf{u}$ are the acceleration, velocity and displacement vectors of the structure relative to the ground; $\boldsymbol{\theta} = (\boldsymbol{\eta}, \boldsymbol{\zeta})$ the vector of all random parameters, with $\boldsymbol{\eta}$ being the vector of random parameters involved in the physical properties of the system and $\boldsymbol{\zeta}$ the vector of random parameters involved in the excitation. In equation (5.53), all the randomness in the system is expressed with the parameter $\boldsymbol{\theta}$, therefore the augmented system $(\mathbf{u}(t), \boldsymbol{\theta})$ is a probability preserved system. In this case the principle of preservation of probability holds, that is:

$$\frac{D}{Dt}\int_{\Omega_t \times \Omega_\theta} p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)\, \mathrm{d}\mathbf{u}\mathrm{d}\boldsymbol{\theta} = 0 \tag{5.54}$$

where, $\dfrac{D}{Dt}$ denotes the total or material derivative, $\Omega_t \times \Omega_\theta$ is any arbitrary domain in the augmented state space at time t, $\Omega_{\boldsymbol{\theta}}$ is the distribution range of the random vector $\boldsymbol{\theta}$ and $p_{\mathbf{u}\boldsymbol{\theta}}$ denotes the joint pdf of $(\mathbf{u}(t), \boldsymbol{\theta})$

After a series of mathematical manipulations and using Reynold's transfer theorem [92, 90] eq. (5.54) written as

$$\int_{\Omega_0 \times \Omega_\theta} \left(\frac{\partial p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)}{\partial t} + \sum_{j=1}^{m} \dot{u}_j(\boldsymbol{\theta}, t) \frac{\partial p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)}{\partial u_j} \right) \mathrm{d}\mathbf{u}\mathrm{d}\boldsymbol{\theta} = 0 \tag{5.55}$$

where m is the total number of d.o.f's. Since eq. (5.55) holds for any arbitrary domain

$\Omega_0 \times \Omega_\theta \in \Omega \times \Omega_{\boldsymbol{\theta}}$,then for any arbitrary $\Omega_\theta \in \Omega_{\boldsymbol{\theta}}$,

$$\int_{\Omega_\theta} \left(\frac{\partial p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)}{\partial t} + \sum_{j=1}^{m} \dot{u}_j(\boldsymbol{\theta}, t) \frac{\partial p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)}{\partial u_j} \right) \mathrm{d}\boldsymbol{\theta} = 0 \tag{5.56}$$

which results in the following partial differential equation:

$$\frac{\partial p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)}{\partial t} + \sum_{j=1}^{m} \dot{u}_j(\boldsymbol{\theta}, t) \frac{\partial p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u}, \boldsymbol{\theta}, t)}{\partial u_j} = 0 \tag{5.57}$$

or in the case where m = 1, equation (5.57) reduces to

$$\frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t)}{\partial t} + \dot{u}(\boldsymbol{\theta}, t) \frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t)}{\partial u} = 0 \tag{5.58}$$

Equations (5.57), (5.58) are the so-called generalized density evolution equations (GDEEs) and possess some significant advantages over the classical probability density evolution equations (Liouville, FP, etc.). More specifically, they reduce the high-dimensional partial differential equations, which can be unfeasible to solve numerically for problems with many dof's, into a series of one-dimensional deterministic partial differential equations. To solve the pde's we need the initial conditions for the problem, which for most of the cases are of the form:

$$p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t) \mid_{t=t_0} = \delta(u - u_0) p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \tag{5.59}$$

Then, the solution to eq. (5.58) leads to the marginal pdf of $u(t)$ as:

$$p_u(u, t) = \int_{\Omega_{\boldsymbol{\theta}}} p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t) \mathrm{d}\boldsymbol{\theta} \tag{5.60}$$

Taking into account that eq. (5.56) holds for any $\Omega_q \in \Omega_{\boldsymbol{\theta}}$,then if we partition Ω_θ into sub-domains, Ω_q' s, $q = 1, 2, ..., n_{pt}$, such that $\Omega_i \cap \Omega_j = \oslash$, $\forall i \neq j$ and $\cup_{q=1}^{n_{pt}} \Omega_q = \Omega_{\boldsymbol{\theta}}$, eq. (5.58) becomes (for $m = 1$):

$$\int_{\Omega_q} \left(\frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t)}{\partial t} + \dot{u}(\boldsymbol{\theta}, t) \frac{\partial p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t)}{\partial u} \right) \mathrm{d}\boldsymbol{\theta} = 0, \qquad q = 1, 2, ..., n_{pt} \tag{5.61}$$

The probability corresponding to the sub-domain Ω_q is then expressed as:

$$P_q = \int_{\Omega_q} p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}, \;\; q = 1, 2, ..., n_{pt} \tag{5.62}$$

If we denote

$$p_q(u,t) = \int_{\Omega_q} p_{u\boldsymbol{\theta}}(u, \boldsymbol{\theta}, t) \mathrm{d}\boldsymbol{\theta}, \;\; q = 1, 2, ..., n_{pt} \tag{5.63}$$

then,

$$p(u,t) = \sum_{q=1}^{n_{pt}} p_q(u,t) \tag{5.64}$$

The procedure described above is referred to as the probability density evolution method and it is evident that the accuracy of this method depends solely on two parameters: a) the partitioning of the probability domain and b) the accurate solution of the corresponding pde's. For the latter, it has to be mentioned that, solving these pde's is not a trivial task, due to the fact that these equations are of the pure advection form with, in most cases, discontinuous initial conditions (shock-type). Therefore, the introduction of artificial dissipation (viscosity) is necessary to ensure the numerical stability of the solution, while adding too much dissipation will result in a severe smearing of the solution.

5.5.3 Surrogate modeling approach for the solution of the GDEE

The basic steps needed to solve the GDEE in eq. (5.58), as proposed by Li and Chen [92], are outlined here. These steps are slightly modified to fit the case of general stochastic systems and are the following:

Step 1: Discretize the probability-assigned space and select representative point sets (random events) $\boldsymbol{\theta}_q = (\theta_{q,1}, \; \theta_{q,2}, \; ..., \theta_{q,s})$ with $q = 1, 2, ..., n_{pt}$, where n_{pt} is the cardinal number of the point set. To each point set determine the assigned probabilities P_q' s via eq.(5.63).

Step 2: Discretize the time-domain into m partitions. For the prescribed $\boldsymbol{\theta} = \boldsymbol{\theta}_q$ solve the discretized deterministic equilibrium equation (eq.(5.53)) with a standard FE solver, to evaluate $\dfrac{du\,(\boldsymbol{\theta}_q, t_m)}{dt}$, where $t_m = m\Delta t \;\; (m = 1, 2...)$, Δt is the time step, $0 \leq t_m \leq T$, T being the total duration of the time history.

Step 3: Introducing $\frac{du\left(\boldsymbol{\theta}_q, t_m\right)}{dt}$ into the generalized density evolution equations and taking into account the discretized version of eq. (5.59), that is, $p_q(u,t)\mid_{t=t_0}=\delta(u-u_0)P_q$, solve eq. (5.58) with finite difference method, or a finite element method, as proposed in the next chapters, to obtain the numerical solution of $p_q(u,x)$

Step 4: Repeat steps 2, 3 for $q=1,2,...,n_{pt}$ and take the numerical integration in eq. (5.63) to compute the numerical solution of $p_u\left(u,x\right)$.

For the solution of the GDEE a surrogate model was employed. More specifically, in step 2 of the algorithm it was assumed that the system's output for any $\boldsymbol{\theta}$ in each of the disjoint sets Ω_q is constant and equal to the response of the system for $\boldsymbol{\theta}=\boldsymbol{\theta}_q$. This indicates the importance of the scheme chosen for partitioning the probability-assigned space to the method's overall performance.

5.5.4 Partitioning of the probability space using the Rotational Quasi-symmetric point method

The strategy of selecting representative point sets at step 1 of the solution algorithm is of paramount importance to the efficiency of the method. Apart from the grid-type partitioning of the probability domain, many other strategies have been successfully employed, depending on the nature of the problem that significantly reduce the computational cost of the method, such as the Number Theoretical Method [95, 24], Tangent Sphere Method [26], Q-SPM [162], RQ-SPM [23] etc.
In [157] a cubature formula for the numerical integration of Gaussian-weighted multi-dimensional integrals,was constructed as:

$$\frac{1}{(2\pi)^{s/2}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{-\boldsymbol{x}^T\boldsymbol{x}/2}f(\boldsymbol{x})dx_1\cdots dx_s=\sum_{k=1}^{N}a_k f(\boldsymbol{x}_k) \tag{5.65}$$

where, $\boldsymbol{x}_k$ are the quasi-symmetric points, a_k are the corresponding weights and s the dimension. If $f(\boldsymbol{x})=1$ for all $\boldsymbol{x}\in R^s$ then this formula provides a partition of the normal random-variate space. In this context, the integral points in the formula are equivalent to the representative points in PDEM, while the weights represent the probability assigned to each of them. Although in [162] it was proven that this formula behaves well when applied in PDEM, nevertheless, there were still some deficiencies that limited its applicability. To further enhance the accuracy of the method the idea was

put forward to rotate the quasi-symmetric points by some specified angles ϑ_{ij}'s in a manner that reduces the generalized F-discrepancy, (GF-discrepancy) [23]. Thus, eq. (5.65) becomes

$$\frac{1}{(2\pi)^{s/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\boldsymbol{x}^T \boldsymbol{x}/2} f(\boldsymbol{x}) dx_1 \cdots dx_s = \sum_{k=1}^{N} a_k f(\boldsymbol{R}\boldsymbol{x}_k) = \sum_{k=1}^{N} a_k f(\boldsymbol{y}_k) \tag{5.66}$$

with $\boldsymbol{y}_k = \boldsymbol{R}\boldsymbol{x}_k$ the points after the rotational transform.
These angles ϑ_{ij}'s can be obtained through an optimization process (e.g Genetic Algorithm) that minimizes the GF-discrepancy of the set.

5.5.5 Partitioning of the probability space using the Number Theoretical Method

NTM is another approach to find an efficient cubature formula for high-dimensional integration. For this method, an appropriate integer generator vector $(n, z_1, z_2, ..., z_s)$ is selected first, from which the point set $\{x_{j,k},\ k = 1, 2, \cdots, n; j = 1, 2, \cdots, s\}$ is produced as

$$x_{j,k} = \left\{ \frac{k z_j}{n} \right\}, \quad k = 1, 2, \cdots, n; j = 1, 2, \cdots, s \tag{5.67}$$

where, $\{\cdot\}$ denotes the fractional part of the bracketed quantity and n is the cardinal number of the point-set. The integer generator vector is computed based on number theory and in [69] and [51] a list of suitable generator vectors for different s and n is provided. The point set generated with this approach has the property of being uniformly scattered over the s-dimensional unit hypercube $\mathscr{C}^s = [0, 1]^s$. Then, using the affine transformation

$$\tilde{\theta}_{j,k} = 2\lambda(x_{j,k} - 0.5), \quad k = 1, 2, \cdots, n; j = 1, 2, \cdots, s \tag{5.68}$$

the point set is mapped to the corresponding point set over the hyper-cube $[-\lambda, \lambda]^s$ in the standardized probability space. The assigned probability of the point $(\tilde{\theta}_{1,k}, \tilde{\theta}_{2,k}, \cdots, \tilde{\theta}_{s,k})$ reads

$$\tilde{P}_k = p_{\boldsymbol{\theta}}(\tilde{\theta}_{1,k}, \tilde{\theta}_{2,k}, \cdots, \tilde{\theta}_{s,k}) \cdot V_k \quad k = 1, 2, \cdots, n$$
$$= \frac{1}{(2\pi)^{s/2}} e^{-\frac{\tilde{\theta}_{1,k}^2 + \tilde{\theta}_{2,k}^2 + \cdots + \tilde{\theta}_{s,k}^2}{2}} \cdot V_k \tag{5.69}$$

where V_k is the representative volume of the element given by the expression $V_k = (2\lambda)^s/n$.

While this method works well for low dimensions i.e $s < 10$, for higher dimensions n becomes too large, leading to an excessive number of structural analyses. To overcome this drawback in [95] the spherical symmetry of the probability space $\Omega_{\boldsymbol{\theta}}$ is exploited. That is, the selected points of the final set, will be those satisfying the condition

$$\sum_{j=1}^{s} \tilde{\theta}_{j,q} \leq (r_0 \lambda)^2, \quad q = 1, 2, \cdots, n; j = 1, 2, \cdots, N_{sel} \tag{5.70}$$

with r_0 being the bounded radius coefficient, chosen such that $0 < r_0 \leq \sqrt{s}$. It is proven, that with this approach, we obtain $N_{sel} \ll n$ without any noticeable loss in the accuracy.

5.5.6 Lax-Wendroff finite difference method with TVD scheme

A modified version of the Lax-Wendroff scheme proposed in [92] is commonly employed for solving eq. (5.58). The implementation steps of this scheme are the following. Denoting $p_u(u_j, \boldsymbol{\theta}, x_m) = p_{j,m}$, then eq. (5.58) is discretized into the form

$$p_{j,m+1} = p_{j,m} - r_L \Big[\frac{1}{2}(g_m + |g_m|)(p_{j,m} - p_{j-1,m}) + \frac{1}{2}(g_m -$$
$$- |g_m|)(p_{j+1,m} - p_{j,m}) \Big] - \frac{1}{2}(1 - |r_L g_m|)|r_L g_m|$$
$$\Big[\chi(r^+_{j+\frac{1}{2}}, r^-_{j+\frac{1}{2}})(p_{j+1,m} - p_{j,m}) - \chi(r^+_{j-\frac{1}{2}}, r^-_{j-\frac{1}{2}})$$
$$(p_{j,m} - p_{j-1,m}) \Big] \tag{5.71}$$

where $t_m = m \cdot \Delta t$ and Δt is the time step in the difference method; $r_L = \frac{\Delta t}{\Delta u}$ is the lattice ratio; Also,

$$r^{+}_{j+\frac{1}{2}} = \frac{p_{j+2,m} - p_{j+1,m}}{p_{j+1,m} - p_{j,m}}, \; r^{-}_{j+\frac{1}{2}} = \frac{p_{j,m} - p_{j-1,m}}{p_{j+1,m} - p_{j,m}}$$
$$r^{+}_{j-\frac{1}{2}} = \frac{p_{j+1,m} - p_{j,m}}{p_{j,m} - p_{j-1,m}}, \quad r^{-}_{j-\frac{1}{2}} = \frac{p_{j-1,m} - p_{j-2,m}}{p_{j,m} - p_{j-1,m}} \tag{5.72}$$

The coefficients g_m take the value

$$g_m = \frac{1}{2}\left(\frac{du\left(\boldsymbol{\theta}, t_{m-1}\right)}{dt} + \frac{du\left(\boldsymbol{\theta}, t_{m}\right)}{dt}\right) \tag{5.73}$$

$\chi(r^+, r^-)$ is the flux limiter, which is constructed based on the Roe-Sweby flux limiter

$$\chi_{sb}(r^-) = max(0, min(2r^-, 1), min(r^-, 2)) \tag{5.74}$$

and

$$\chi(r^+, r^-) = H(-g_m)\chi_{sb}(r^+) + H(g_m)\chi_{sb}(r^-) \tag{5.75}$$

with H being the Heaviside function.

A necessary (but not sufficient) condition for the stability of the scheme is the well-known Courant-Friedrichs-Lewy (CFL) condition, which for eq. (5.71) reads: $|r_L g_m| \leq 1$.

6

A Galerkin-based formulation of the probability density evolution method for general stochastic finite element systems

6.1 Introduction

A key aspect of PDEM lies in solving accurately the resulting deterministic pde' s. These pde' s express mathematically the conservation of probability and have the form of pure advection, that is, there is no diffusion term in the equations. In such cases numerical finite difference solutions tend to be unstable, unless some diffusion is added artificially in the scheme. Upwind methods, introduce such artificial diffusion, but, even though they lead to a numerically stable formulation, they may result in a considerable "smearing" of the solution [86, 45, 122, 127]. Moreover, in many practical applications the initial conditions involve discontinuities (shock, sharp jumps) and resolving these discontinuities may produce 'unwanted' wiggles (spurious oscillations). In [92] a modified version of the Lax-Wendroff (LW) finite difference scheme, enhanced with the Total Variation Diminishing (TVD) property, is proposed for solving the pde's. This scheme is proven very efficient and numerically stable, but can be prone to overly diffuse solutions. Besides that, in order to ensure its stability, the well known Courant-Friedrichs-Lewy(CFL) [39]

condition is necessary to be satisfied. This condition defines the incremental step of the scheme, with the subsequent increase in the computational cost.

In an attempt to improve the accuracy of the Lax-Wendroff (LW) finite difference schemes, an alternative numerical solution of the PDEM is presented in this chapter, which is based on a Galerkin finite element method approximation. More specifically, the GDEE is reformulated in the framework of time-marching Discontinuous Galerkin (DG) scheme [68, 126, 32, 33, 34], as well as Streamline Upwind/Petrov Galerkin (SUPG) FE scheme [49, 20] for the solution of the corresponding pde's. Numerical results are presented, which demonstrated that the use of the SUPG, endowed with a shock capturing term [70], provides more accurate results with respect to LW and DG, especially in the areas near the boundary conditions at practically lower computational effort. The decrease in the computational effort is mainly attributed to the fact that the SUPG method is not restricted by the CFL condition. In addition, a reformulation of the classical GDEE is proposed, which implements the principle of probability preservation in space instead of time, making this approach suitable for the stochastic analysis of finite element systems. The advantages of the FE Galerkin methods and in particular the SUPG over finite difference schemes, like the modified Lax-Wendroff (LW), which is the most frequently used method for the solution of the GDEE, are illustrated with numerical examples and explored further.

6.2 PDEM for general stochastic systems

A reformulation of the GDEE for the analysis of general stochastic systems can be derived in a straightforward manner by following exactly the same steps deployed in section 5.5.2. Consider the equation of equilibrium of a multi-degree-of freedom system

$$\mathbf{K}(\boldsymbol{\theta})\mathbf{u} = \mathbf{F}(\boldsymbol{\theta}) \tag{6.1}$$

where, $\mathbf{u}$ stands for the displacement vector along the dimension of one global degree of freedom. In this case our interest lies in finding the evolution of probability density function, along the material points or in other worlds, as a function of the position vector of the system, henceforth denoted as x in the case of one dimensional structures. In this

case the principle of preservation of probability is expressed as:

$$\frac{D}{Dx}\int_{\Omega_x\times\Omega_\theta} p_{\mathbf{u}\boldsymbol{\theta}}(\mathbf{u},\boldsymbol{\theta},x)\,\mathrm{d}\mathbf{u}\mathrm{d}\boldsymbol{\theta}=0 \tag{6.2}$$

Thus, following the same procedure described in section 5.5.2, eq. (5.58) will in this case take the form:

$$\frac{\partial p_{u\boldsymbol{\theta}}(u,\boldsymbol{\theta},x)}{\partial x}+\frac{du(\boldsymbol{\theta},x)}{dx}\frac{\partial p_{u\boldsymbol{\theta}}(u,\boldsymbol{\theta},x)}{\partial u}=0 \tag{6.3}$$

The difference with the previous case is that, now, the rate or 'velocity' of u is with respect to the length of the structure and x replaces t in all the equations of section 5.5.2.

Also, the initial conditions for the problem will be:

$$p_{u\boldsymbol{\theta}}(u,\boldsymbol{\theta},x)\mid_{x=x_0}=\delta(u-u_0)p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \tag{6.4}$$

Then the marginal pdf of $u(x)$ can be written as

$$p_u(u,x)=\int_{\Omega_{\boldsymbol{\theta}}} p_{u\boldsymbol{\theta}}(u,\boldsymbol{\theta},x)\mathrm{d}\boldsymbol{\theta} \tag{6.5}$$

For example, if we consider the generic element of figure 6.1 as a typical part of stochastic static system, it is obvious that the randomness involved in the loading q and the bending stiffness EI of the element, will lead to a random field $u(x)$, along the length of the element, which can be evaluated via eq. (6.3) with the initial conditions expressed in eq. (6.4). The initial conditions can be found at the edge nodes of the element. For instance, if we have fixed support at node i, then the displacement at this position will be zero with certain probability, and the initial condition reads: $p_{u\boldsymbol{\theta}}(u,\boldsymbol{\theta},x)\mid_{x=0}=\delta(u-0)p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$.

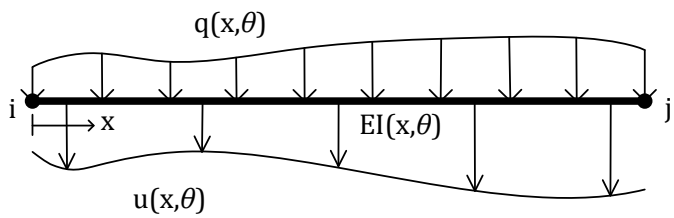


Figure 6.1: Generic element with the randomness involved in the loading and the system parameters

Solving the deterministic pde (eq. (6.3)) accurately and efficiently is a vital component of the whole procedure. When faced with this task, one quickly realizes that there is a wide range of methods to do so, including finite difference schemes, finite volume and finite element methods. The selection of the optimum scheme though isn't always an easy task. On the one hand, it depends on the problem under consideration (e.g. existence of steep gradients) as well as on the scheme chosen, as some schemes tend to be more dispersive and others more dissipative. From the three methods mentioned, finite difference schemes are the easiest to implement but they have certain drawbacks. For instance, central finite difference schemes lead to unstable solutions, while upwind methods suffer from overdissipation. In addition, finite difference schemes depend heavily upon the so-called Courant-Friedrichs-Levy (CFL) condition, which poses a restriction on the incremental step of the method. In this work we investigated two alternative to the widely used up to now LW schemes, based on Galerkin finite element approximations for solving eq. (6.3). More specifically, we employed a time-marching discontinuous Galerkin scheme (DG-FEM) and the StreamlineUpwind/PetrovGalerkin (SUPG) method. The motivation behind that, was the fact that eq. (6.3) describes a conservation law and resembles the 1D hyperbolic advection equation. Both these methods have been successfully used for treating this type of pde's, each with its own merits [68, 49].

6.3 Space-time finite elements: The StreamlineUpwind/Petrov-Galerkin formulation

The streamline-upwind/Petrov-Galerkin concept (SUPG) was introduced by Hughes and Brooks [20, 70], and the basic idea behind this method is to add numerical diffusion

along the streamlines in a consistent manner. In the above mentioned references, its application in the linear scalar advection equation is demonstrated. Since eq. (6.3) resembles the advection equation, we can rewrite it in the following form:

$$\mathbf{w}\nabla p = 0 \tag{6.6}$$

where, $\mathbf{w} = (1, \alpha(x))$ the velocity field, $\alpha(x)$ is the flux $\dfrac{du(x)}{dx}$ and $\nabla p = (\frac{\partial p}{\partial x}, \frac{\partial p}{\partial u})$. A finite element partition of the computational domain $\Omega = \Omega_u \times \Omega_x \subset R^2$ is denoted by $\{\Omega^e\}$ for $e = 1, 2, ..., N_{el}$. Since the domain is two dimensional, we can use quadrilateral elements to discretize Ω. Figure 6.2 illustrates this discretization of the physical domain with quadrilateral elements, where each d.o.f corresponds to the probability at this position.

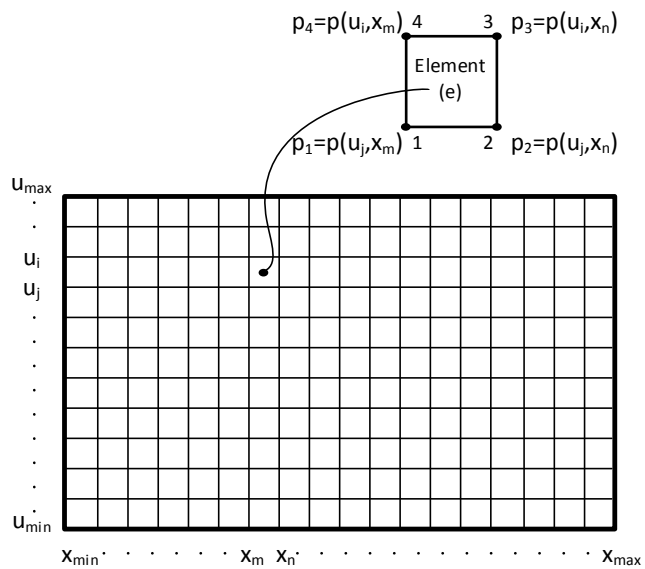


Figure 6.2: Discretization of the physical domain into 4-noded quadrilateral finite elements with 1 degree of freedom per node, corresponding to the probability assigned to the displacement value and this position

The boundary conditions are Dirichlet on all of the boundary $\partial\Omega$, that is, $p = g$. We can define the space of trial solutions as $\mathscr{V} = \{p | p \in \mathscr{H}^1(\Omega), p = g \text{ on } \partial\Omega\}$, where

$\mathscr{H}^1 = \{p : \Omega \to R | p, \frac{\partial p}{\partial x}, \frac{\partial p}{\partial u} \in L_2(\Omega)\}$ is the Sobolev space and we also define the space $\mathscr{V}_0 = \{\psi | \psi \in \mathscr{H}^1(\Omega), \psi = 0 \text{ on } \partial\Omega\}$. The basic idea of Petrov-Galerkin approximation is to specify a weak formulation in which the space of test (weighting) functions is taken to be different than the space of the trial solutions. More specifically, the test space is the space spanned by functions of the form:

$$\tilde{\psi} = \psi + \tau \mathbf{w} \nabla \psi \tag{6.7}$$

where $\psi \in \mathscr{V}_0$ is the Galerkin-type weighting function and τ is a coefficient, which for each element e is given by the following expression [20]:

$$\tau_e = \frac{\mathsf{a}\lambda_e}{2|\mathbf{w}_e|} \tag{6.8}$$

In the above expression λ_e is the characteristic length of the e element given by $\lambda_e = min\left(\frac{\lambda_x}{cos\vartheta}, \frac{\lambda_u}{sin\vartheta}\right)$ with λ_x, λ_u being the rectangle's lengths in x and u direction, respectively, and $\vartheta = arctan\left(|\frac{w_u}{w_x}|\right)$. In the following equations we will use the subscript h to refer to the discrete finite element problem. The weak form of eq. (6.6) can be written as

$$\int_\Omega \psi^h \mathbf{w} \nabla p^h d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau \mathbf{w} \nabla \psi^h \mathbf{w} \nabla p^h d\Omega_e = 0 \tag{6.9}$$

In eq. (6.8) a is the upwind function, which depends on the element Peclet number Pe ($Pe = \infty$ for pure advection). The expression $\mathsf{a} = coth(Pe) - \frac{1}{Pe}$, with a being equal to 1 in our case, can be used.

In order to further enhance the accuracy of the method, we add a discontinuity capturing term in eq. (6.9) based on the methodology presented in [70]. Then, eq. (6.9) is rewritten as follows:

$$\int_\Omega \psi^h \mathbf{w} \nabla p^h d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_1 \mathbf{w} \nabla \psi^h \mathbf{w} \nabla p^h d\Omega_e + \\ + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_2 \mathbf{w} \nabla \psi^h \mathbf{w}_\parallel \nabla p^h d\Omega_e = 0 \tag{6.10}$$

where, $\mathbf{w}_{\|}$ denotes the projection of $\mathbf{w}$ onto ∇p^h, that is,

$$\mathbf{w}_{\|} = \begin{cases} \dfrac{\mathbf{w}\nabla p^h}{|\nabla p^h|_2^2}, & \text{if } \nabla p^h \neq 0 \\ \\ 0, & \text{if } \nabla p^h = 0 \end{cases} \tag{6.11}$$

According to [70], $\tau_1 = \tau$ and $\tau_2 = max\left(0, \tau_{\|} - \tau\right)$, where $\tau_{\|}$ is computed as indicated in eq. (6.8) , but using $\mathbf{w}_{\|}$ instead of $\mathbf{w}$. In a matrix notation, eq. (6.10) can be rewritten as:

$$\boldsymbol{\mathcal{K}}\,\mathbf{p} = 0 \tag{6.12}$$

where,

$$\boldsymbol{\mathcal{K}} = \boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{S}} \tag{6.13}$$

$$\mathcal{K}_{ij} = \mathcal{A}_{ij} + \mathcal{S}_{ij} \tag{6.14}$$

$$\mathcal{A}_{ij} = \int_{\Omega} \left(\mathbf{w}\nabla\varphi_j\right)\varphi_i \tag{6.15}$$

$$\mathcal{S}_{ij} = \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_1 \mathbf{w}\nabla\varphi_j \mathbf{w}\nabla\varphi_i + \sum_{e=1}^{n_{el}} \int_{\Omega_e} \tau_2 \mathbf{w}\nabla\varphi_j \mathbf{w}_{\|}\nabla\varphi_i \tag{6.16}$$

and $\{\varphi\}$ is the basis that spans the functional spaces.

6.4 Discontinuous Galerkin Finite Element Method (DG-FEM)

The discontinuous Galerkin method is a robust and compact finite element projection method, well suited for dealing with partial differential equations describing conservations laws (e.g. advection equation). It can be seen as a combination of the finite element method and the finite volume method, taking advantages from both. An important distinction between the DG method and the usual finite-element method is that in the DG-FEM the resulting equations are local to the generating element and the coupling of the elements is achieved using the appropriate numerical flux across the elements. Thus, each element may be thought of as a separate entity that merely needs to obtain

some boundary data from its neighbours and no global linear or nonlinear systems need to be solved. Moreover, a numerical flux can be chosen to be more or less dissipative, depending on the problem under consideration.

Consider the following equation representing a GDEE.

$$\frac{\partial p_q(u,x)}{\partial x} + \alpha_q(x)\frac{\partial p_q(u,x)}{\partial u} = 0, \text{ for } q = 1, 2, ...n_{pt} \tag{6.17}$$

under the initial condition:

$$p_q(u,0) = p_{q,0}(u) \tag{6.18}$$

where, $\alpha_q(x)$ is the flux $\dfrac{du_q(x)}{dx}$, $u \in [u_{min}, u_{max}] = \Omega_u$ and $0 \leq x \leq L$, with L being the length under consideration in the stochastic system. Without introducing any ambiguities, we will omit the subscript q in the above expression and introduce the subscript h to refer to the discrete solution. Partitioning the domain Ω_u by K non-overlapping elements $u \in \left[u_l^k, u_r^k\right] = D^k$ such that

$$\Omega_u \cong \Omega_u^h = \bigcup_{k=1}^{K} D^k \tag{6.19}$$

where u_l^k and u_r^k are the left and right displacements of element k, then on each of these elements the local solution can be expressed as a polynomial of order $N = N_p - 1$.

$$u \in D^k : p_h^k\left(u,x\right) = \sum_{i=1}^{N_p} p_h^k\left(u_i^k, x\right) l_i^k(u) \tag{6.20}$$

where $l_i^k\left(u\right)$ are the interpolating Lagrange polynomials and $p_h^k\left(u_i^k, x\right)$ refers to the nodal values. The global solution $p\left(u,x\right)$ is then assumed to be approximated by the N-th order polynomial approximation $p_h\left(u,x\right)$, that is:

$$p(u,x) \cong p_h(u,x) = \bigoplus_{k=1}^{K} p_h^k(u,x) \tag{6.21}$$

defined as the direct sum of the K local polynomial solutions $p_h^k\left(u,x\right)$.

Following the procedure described in [68] a strong DG-form of (6.17) for each element

k can be written in matrix notation as follows

$$\boldsymbol{M}^k \frac{d}{dx}\boldsymbol{p}_h^k + \boldsymbol{S}^k \alpha \boldsymbol{p}_h^k = \left[\boldsymbol{l}^k(u)(\alpha p_h^k - f^\star)\right]_{u_l^k}^{u_r^k} \tag{6.22}$$

where $\boldsymbol{M}^k$ and $\boldsymbol{S}^k$ are the element's mass and stiffness matrices respectively. More specifically:

$$M_{ij}^k = \int_{D_k} l_i^k(u) l_j^k(u) du \tag{6.23}$$

$$S_{ij}^k = \int_{D_k} l_i^k(u) \frac{dl_j^k(u)}{du} du \tag{6.24}$$

and $f^\star$ is a suitably chosen numerical flux, which controls the information entering and exiting the element. In our calculations we used the monotone Lax-Friedrichs flux [121, 87], while for the integration with respect to x we used a third-order three-stage Runge-Kutta scheme [68]. In addition, in order to prevent the well known oscillations appearing in the DG-FEM solution, the DG scheme was equipped with the MUSCL (Monotone Upstream-centered Scheme for Conservations Laws) slope limiter [87].
It must be mentioned here, that the stability of the DG method is controlled by the CFL condition, which generally depends on the selection of numerical flux, the order of elements used and the finite difference scheme employed. In our case, the inequality $\Delta x \leq \frac{1}{|\alpha|}\Delta u$ poses a reasonable bound for the choice of Δu and Δx.

6.5 Numerical Examples

Example 1. Portal frame

The simple example of the portal frame of figure 6.3 is chosen first, in order to test the accuracy of the proposed methodologies. For this frame of length and height equal to $L = 5m$, a HEB 220 cross-section is assigned to all its members. The random parameters considered are the modulus of elasticity E and the horizontal force P, which are assumed to follow the Gaussian distribution with mean $\bar{E} = 21000\frac{kN}{cm^2}$, coefficient of variation $COV_E = 0.10$, mean $\bar{P} = 30kN$ and $COV_P = 0.15$. In this case, our interest lies in finding the evolution of the probability density function of the horizontal displacement u along the length of its member of the frame. Geometric and material nonlinearities

are included in the model. For the material nonlinearity an elastic-perfectly plastic constitutive law is assummed with a yield stress of $F_y = 355MPa$. For the application of PDEM, a symmetric 10 by 10 grid was used to partition the two-dimensional probability space, leading to 100 model evaluations. In the absence of an analytic solution to our problem, the results of a large number of Monte Carlo simulations were considered as the 'exact solution' for comparison purposes.

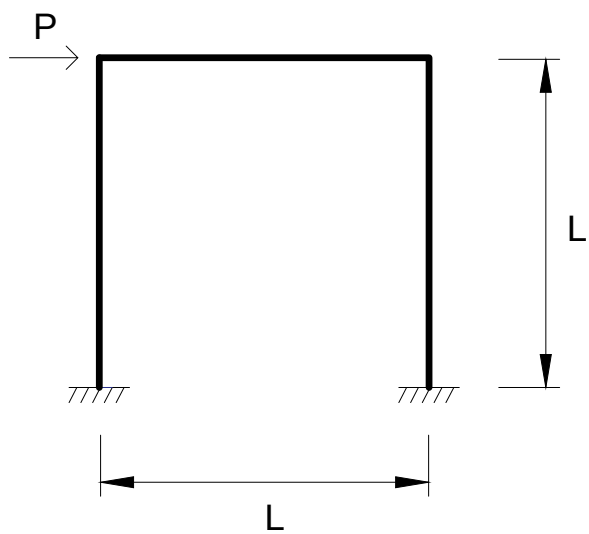


Figure 6.3: 1-story stochastic structure

Depicted in the following figures are the mean value and variance along the length of the path, the evolution of the pdf $p(u, x)$, and typical pdf profiles at certain positions (snapshots), computed with the various proposed methodologies and the brute force Monte Carlo Simulation (MCS). Figure 6.4 presents a comparison in the mean value between the MCS and the three different approaches, SUPG, LW and DG-FEM, respectively. We notice a perfect match in all cases. The same comparison in terms of variance is depicted in figure 6.5 for two different grids used to discretize the domain $\Omega_u \times \Omega_x$, namely, grid A with 150×600 elements and grid B with 450×1500. Figures 6.5a and 6.5b present the results for grid A and grid B, respectively. It can be seen in these figures that only SUPG has converged to the MC solution with the coarse grid A, while the DG and LW require a more dense mesh to achieve the same accuracy. The reason behind this

is excessive dissipation observed (see figure 6.7) in LW and DG-FEM, which leads to overly diffuse and spread out solutions. Nevertheless, in figure 6.5b it becomes apparent, that a finer grid tends to significantly improve the results. Besides that, from figures 6.4, 6.5 it is also worth noticing that the mean value and the variability of the horizontal displacement u, remains invariant along the length of the beam, since the term $\frac{du}{dx}$ in eq. (6.3) becomes equal to zero and therefore, $p(u, x)$ becomes constant.

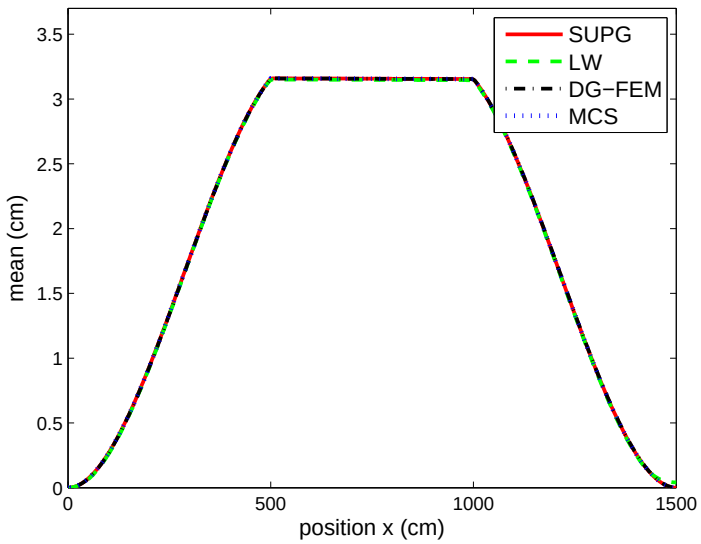
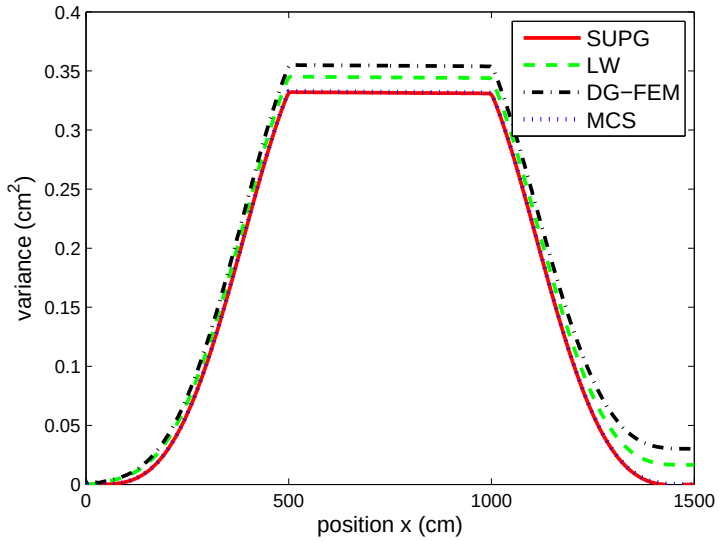
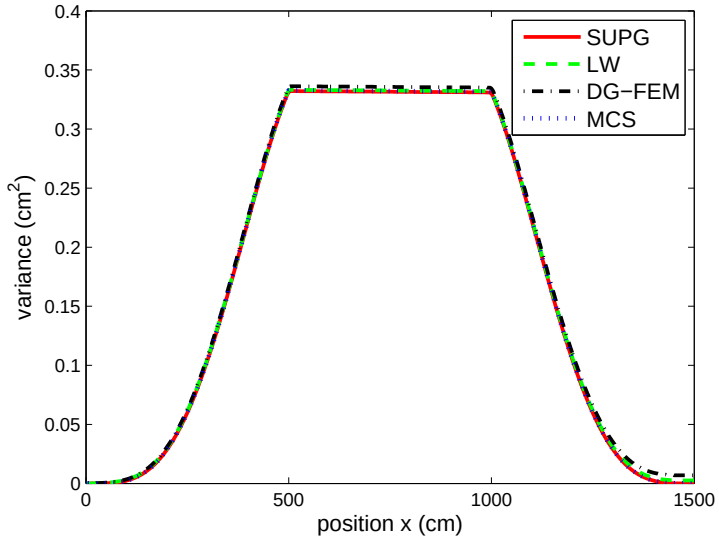


Figure 6.4: Comparison in mean value between the SUPG, the modified LW and the DG-FEM for a 150x600 grid with brute force Monte Carlo simulation



(a)



(b)

Figure 6.5: Comparison in variance between the SUPG, the modified LW and the DG-FEM for a a) 150x600 grid and b) 450x1500 grid with brute force Monte Carlo simulation

The faster convergence of SUPG is further illustrated in the convergence study of figure 6.6 from which, it is evident that SUPG has the highest convergence rate among

the three methods and achieves a considerably smaller error for the same number of elements in the Ω_u domain. The corresponding mesh sizes were selected as the minimum ones that satisfy the CFL condition for the LW and DG-FEM, in order for the solution to be stable. This CFL imposed restriction can be avoided, however, if we choose to employ the SUPG method, in which the discretization on Ω_u can be refined indepedently of the Ω_x discretization. This advantage of the SUPG will be illustrated in more detail in the next example.

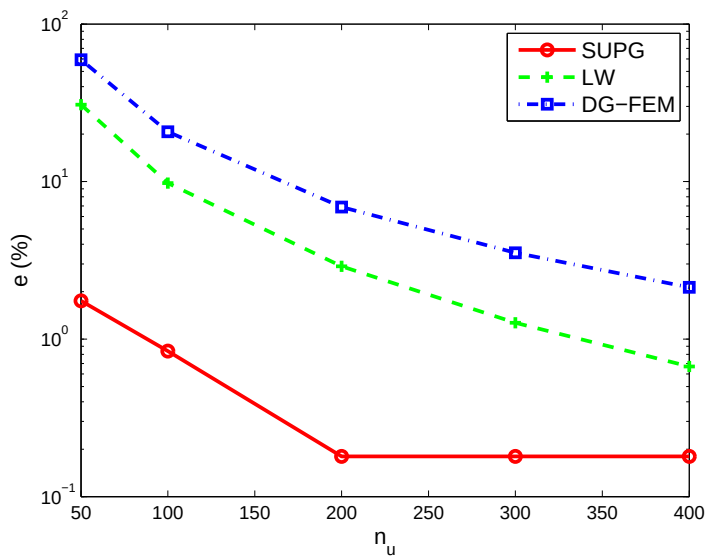
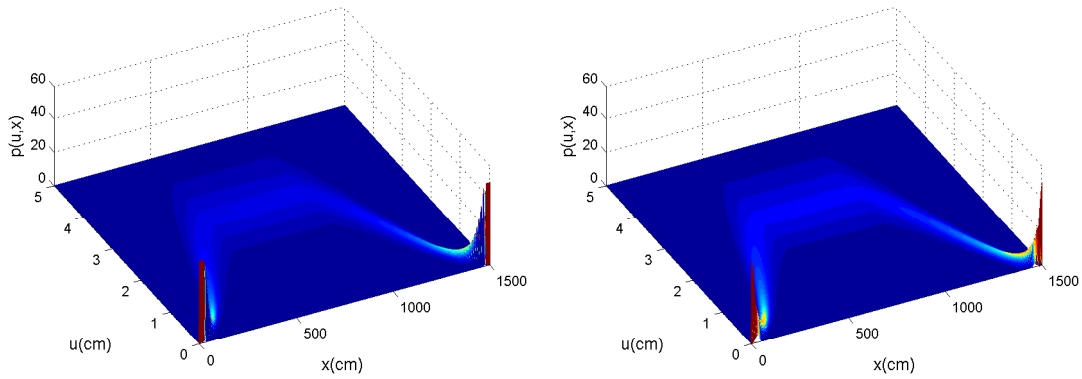


Figure 6.6: Convergence error (%) in variance as a function of the number of elements in the Ω_u domain for the SUPG, the modified LW and the DG-FEM

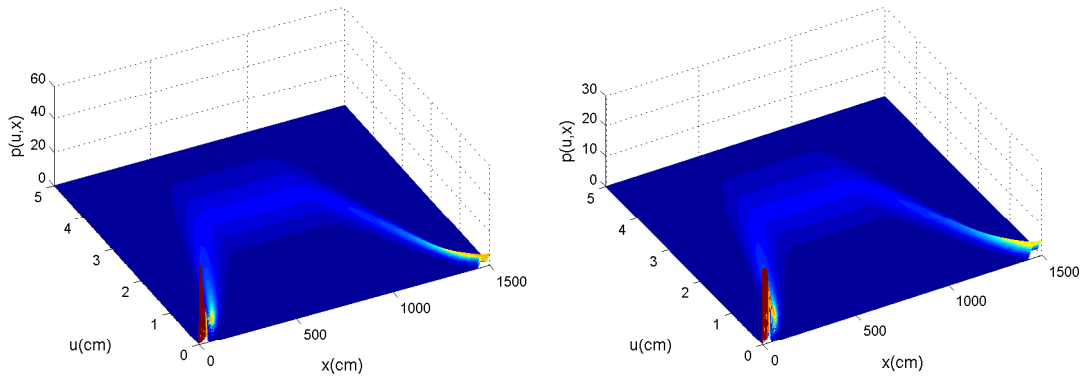
In figure 6.7 the evolution of $p(u, x)$ along each position of the structure is presented, where the 450×1500 grid was applied in all cases. In this figure it is demonstrated how the initial probability "flows" through the physical domain for a) the MCS method, b) the SUPG formulation, c) the modified Lax Wendroff scheme and d) the DG-FEM. The x-axis denotes the position of the structure, with x ranging from 0 to $3L$, the y-axis denotes the possible values of the horizontal displacement u and the z-axis the pdf values $p(u, x)$ in each position. An intersection with a vertical plane parallel to the y-axis will give the pdf profile at this position. Due to the geometry and boundary conditions of the structure, a symmetrical profile of the evolution of the pdf is expected, and also, in the locations of the supports ($x = 0$ and $x = 3L$) all the probability should be lumped

at the certain event of a zero displacement. From these figures it is obvious, that only the SUPG method equipped with the shock capturing terms in eq. (6.10) can capture this symmetry. The DG-FEM formulation and the modified LW scheme fail to do so, due to the artificial dissipation added in these methods. Even with an increase in the number of elements in the grid, the probability at this boundary could not be accurately captured.



(a) Monte Carlo Simulation

(b) StreamlineUpwind/Petrov Galerkin



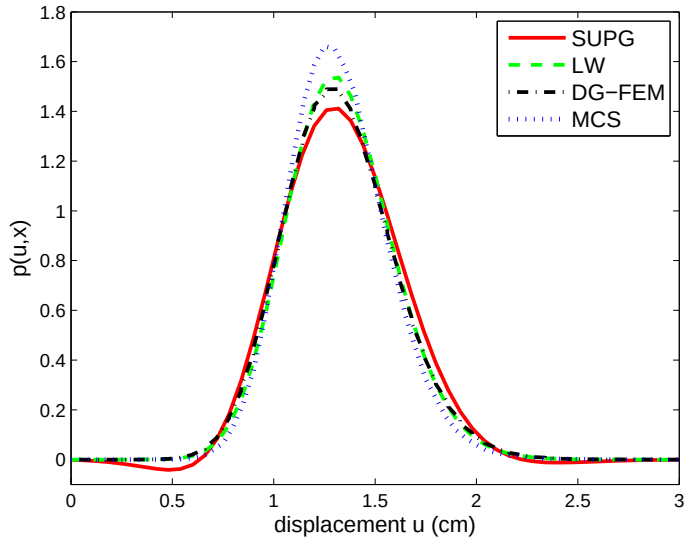
(c) Modified Lax-Wendroff

(d) Discontinuous Galerkin FEM

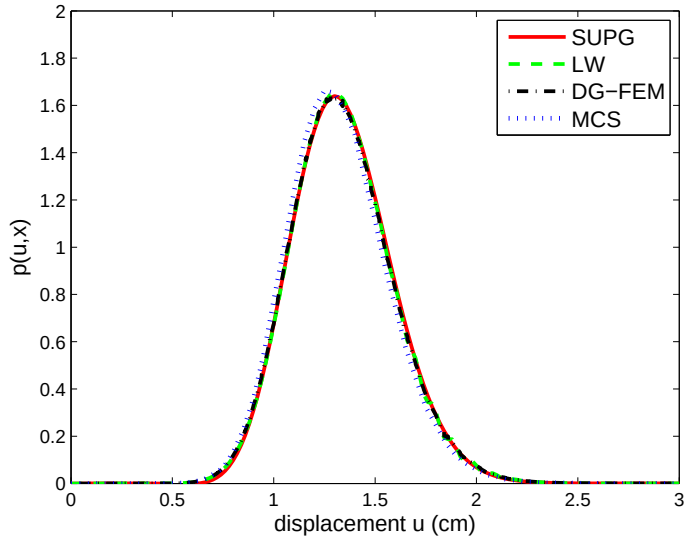
Figure 6.7: Evolution of the probability density function $p(u, x)$ for a) the Monte Carlo Simulation, b) the SUPG, c) the modified LW and d) the DG-FEM

Finally, figures 6.8, 6.9 plot the pdf profiles at the positions $x = 2.50m$ and $x = 7.50m$, respectively, for the two mesh sizes mentioned earlier. Inspection of these figures reveals some very useful conclusions for each method. Specifically, SUPG has the disadvantage

of not being a monotone method and thus allowing negative values in the neighbourhood of discontinuities, like in the case of the boundary conditions. This conclusion is more obvious in figure 6.8a, which is closer to the boundaries of the structure, and less in figure 6.9a. This problem, though, can be significantly ameliorated by considering a locally refined mesh near the boundaries (figures 6.8b, 6.9b) to resolve the discontinuities. On the other hand, the LW and DG-FEM possess the TVD property and thus the non-negativity of the solution is ensured. From figures 6.8a, 6.8b however, we observe that both these methods suffer from a noticeable "smearing" in the solution profile due to the excess dissipation added in the schemes. Again, the situation can be improved by using a finer grid (figures 6.8b, 6.9b).

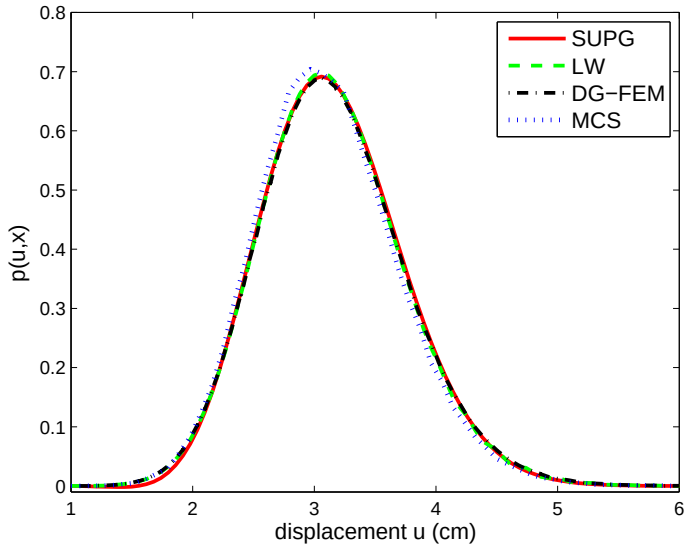


(a)

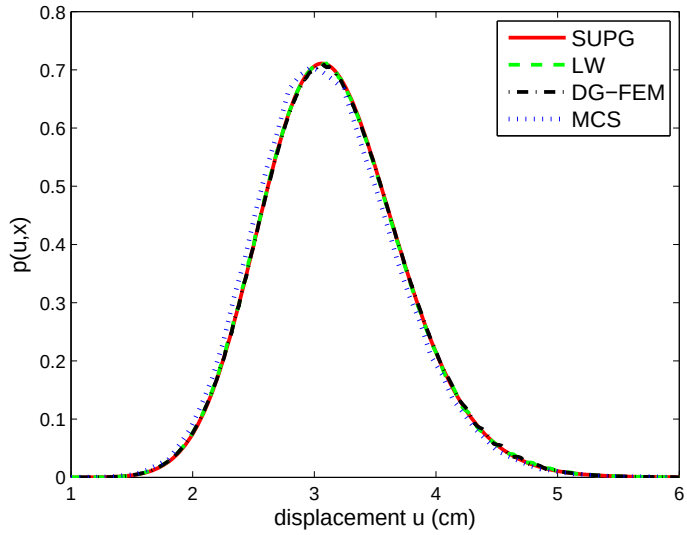


(b)

Figure 6.8: PDF profiles ta x=2.50m for the SUPG, the modified LW and the DG-FEM for a a) 150x600 grid and b) 450x1500 grid versus brute force Monte Carlo



(a)



(b)

Figure 6.9: PDF profiles ta x=7.50m for the SUPG, the modified LW and the DG-FEM for a a) 150x600 grid and b) 450x1500 grid versus brute force Monte Carlo

Example 2. Static response of a two-bay, three-story frame

A more realistic case which is the 3-storey, 2-bay structure of figure 6.10 is considered as the second test case. In this more complex structure, the path, along which we estimate the evolution of probability is not unique. In figure 6.10 the red (dashed) line denotes the path we chose to investigate, but alternative paths can be chosen as well, provided that the initial conditions are well defined. Initial conditions are not necessarily defined at some support of the structure. For instance if we have already evaluated the pdf at the upper left node of the first story, then we can proceed along the path denoted with the blue (solid) line in figure 6.10 defining as initial conditions the calculated pdf at this point. The loading, the moduli of elasticity of each column, the beams of each floor and the yielding stresses of the materials in each storey are considered as random parameters, with properties the ones presented in table 6.1. The partitioning of the corresponding probability domain plays an important role in the efficiency of the method and for this reason we chose to employ the Rotational Quasi-Symmetric Point Method (RQ-SPM) method in order to determine the representative point sets of the domain. Implementation aspects of the RQ-SPM can be found in [23]. Only 286 representative point sets are required for the method to achieve a partioning of the probability domain with an error of 3.58%, which can be considered as small enough. These sets can also be found in [23]. As in the previous example, we used HEB 220 for all section types of the frame members, all having a length of $L = 5m$. Geometrical and material nonlinearities were also included in the model.

Table 6.1: Random parameters of the structure

	$E_1 - E_4$	$E_5 - E_8$	$E_9 - E_{11}$	f_{y_1}	f_{y_2}	f_{y_3}	q
mean	210 GPa	200 GPa	190 GPa	425 MPa	355 MPa	275 MPa	10 $\frac{kN}{m^2}$
COV	0.2	0.2	0.2	0.2	0.2	0.2	0.2

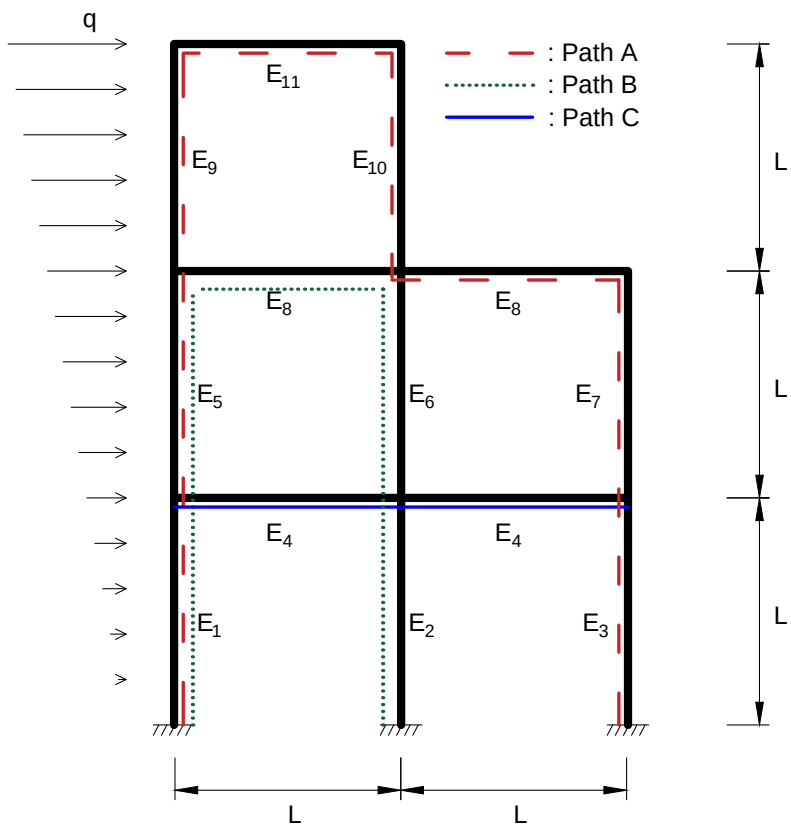


Figure 6.10: 3-story, 2-bay stochastic frame structure

Figures 6.11 and 6.12 present the mean and variance as a function of the position along the selected path, calculated with all aforementioned methodologies, for a 100×400 grid. From these figures it is apparent that, as in the previous example, in terms of mean all methods provide excellent results, while in terms of variance SUPG provides the most accurate results. The convergence rates in variance for each method are also investigated and the results are depicted in figure 6.13. Again, the superiority of SUPG is evident.

As mentioned previously, another advantage of the SUPG method, that has not been exploited yet, is the fact that this method is not restricted under the CFL condition, like the other two methods. Therefore, the Ω_x domain in SUPG can be discretized in as many elements as necessary to obtain an accurate solution indepedently of the Ω_u disretization. In the LW and DG-FEM cases the number of elements in Ω_x must obey the CFL condition and thus they should be significantly increased leading to a more expensive mesh. This advantage of the SUPG is demonstrated in figure 6.14, which plots the convergence error in variance against the total number of elements in the domain for two cases, namely 'SUPG1' and 'SUPG2'. In both cases the same number of elements for the discretization of Ω_u is used, but in the former case ('SUPG1') the Ω_x is discretized in as many elements as required by the CFL condition, while for the latter ('SUPG2') a smaller number is used, namely 160 elements for all analyses. From this figure, it can be observed that SUPG can achieve almost the same level of accuracy for a significantly coarser mesh with the consequent reduction in the computational burden.

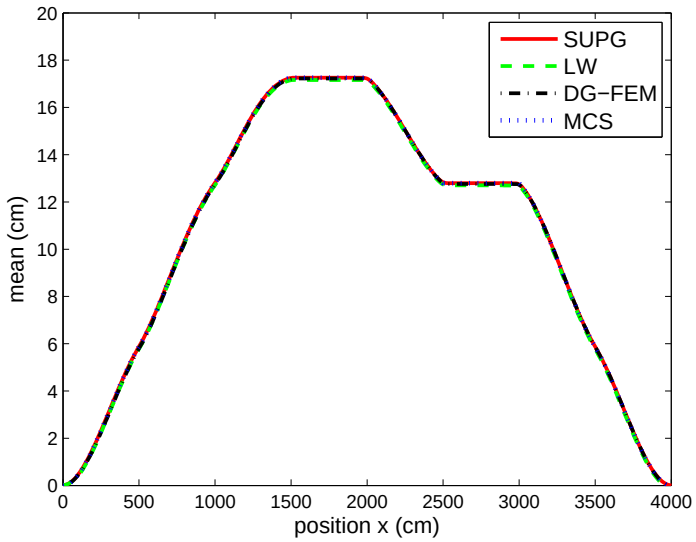


Figure 6.11: Comparison in mean value between the SUPG, the modified LW and the DG-FEM with brute force Monte Carlo simulation

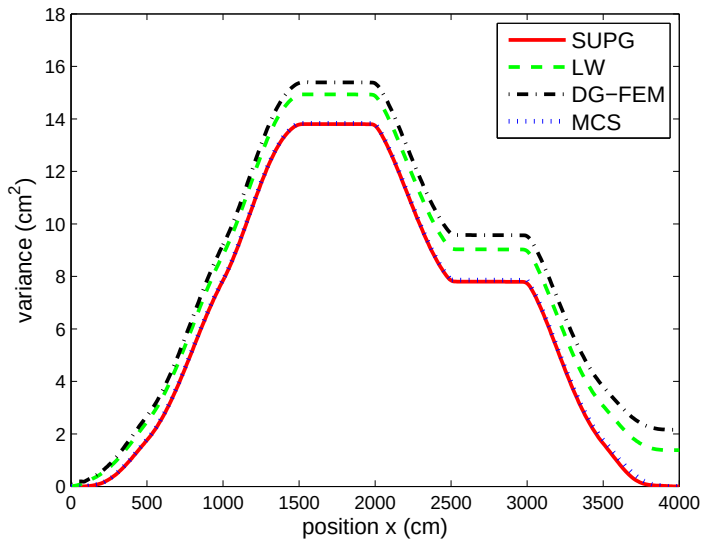


Figure 6.12: Comparison in variance between the SUPG, the modified LW and the DG-FEM with brute force Monte Carlo simulation

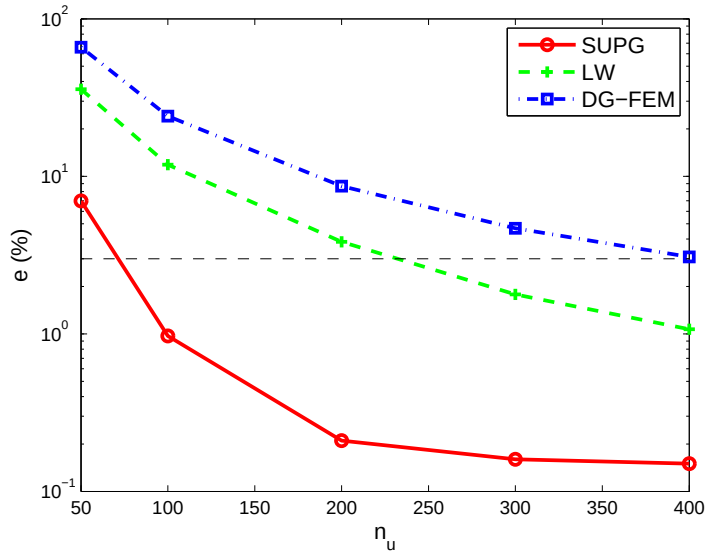


Figure 6.13: Convergence error (%) in variance as a function of the number of elements in the Ω_u domain for the SUPG, the modified LW and the DG-FEM cases

Figure 6.15 presents the evolution of the pdf along the length of the chosen path for

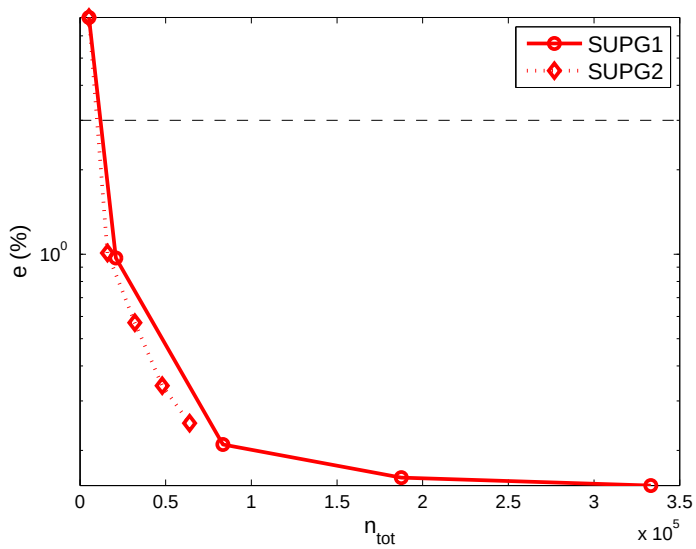
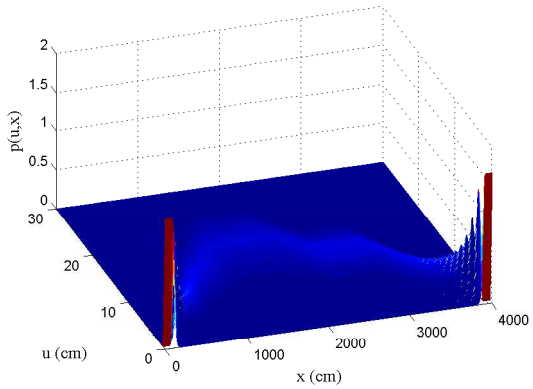
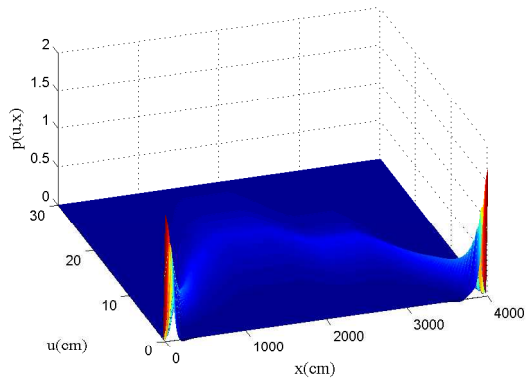


Figure 6.14: Convergenve error (%) in variance as a function of the total number of degrees of freedom n_{tot} in the $\Omega_u \times \Omega_x$ domain for the SUPG1 and SUPG2 cases

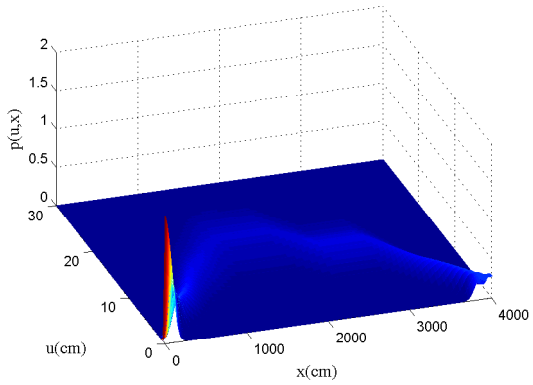
the a) MCS, b) SUPG, c) LW and d) DG-FEM formulation. Again, it is apparent that only the SUPG can capture accurately the probability at the end of the domain. As in the previous example, the variability of the horizontal response remains constant along the beams, a result that agrees with the physical interpretation. In figures 6.16 and 6.17 the pdf profiles at the positions $x = 10.0m$ and $x = 30.0m$ respectively, are presented. In all cases we used the same 100×400 grid, expect for the SUPG case where we also used a 200×400 grid for comparison purposes. The results are of fair accuracy but some weaknesses for each method can also be reported. More specifically, SUPG allows the pdf to take negative values, while LW and DG-FEM produce more "smeared" profiles. For the SUPG case, in particular, this problem could be ameliorated by considering a finer grid and this becomes evident from the results of the SUPG case with the denser mesh, where a noticeable improvement can be reported.



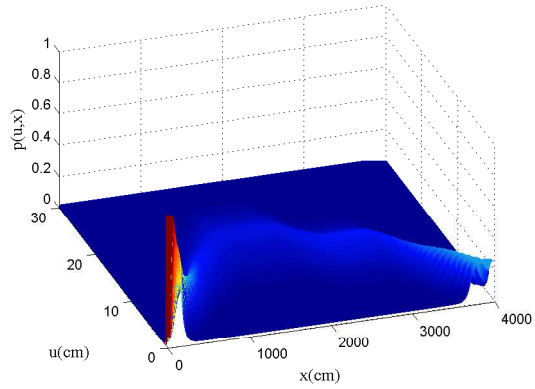
(a) Monte Carlo Simulation



(b) StreamlineUpwind/Petrov Galerkin



(c) Modified Lax-Wendroff



(d) Discontinuous Galerkin FEM

Figure 6.15: Evolution of the probability density function $p(u, x)$ for a) the Monte Carlo Simulation, b) the SUPG, c) the modified LW and d) the DG-FEM

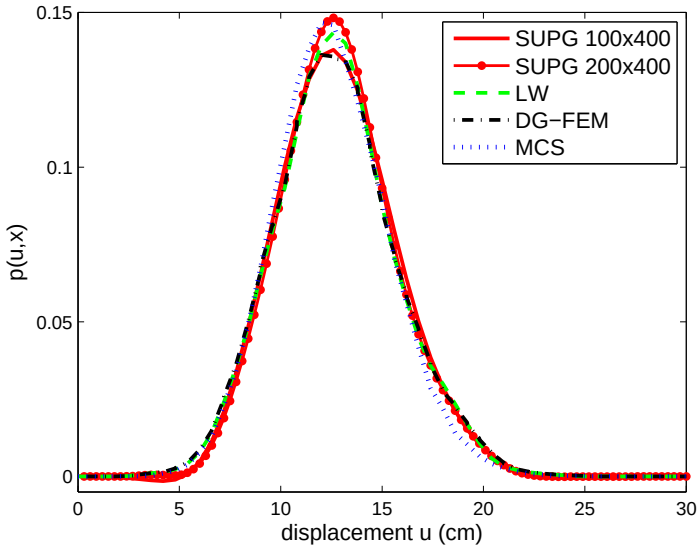


Figure 6.16: PDF profiles ta x=10.00m for a) SUPG 100x400, b) SUPG 200x400, c) modified LW and d) DG-FEM versus brute force Monte Carlo

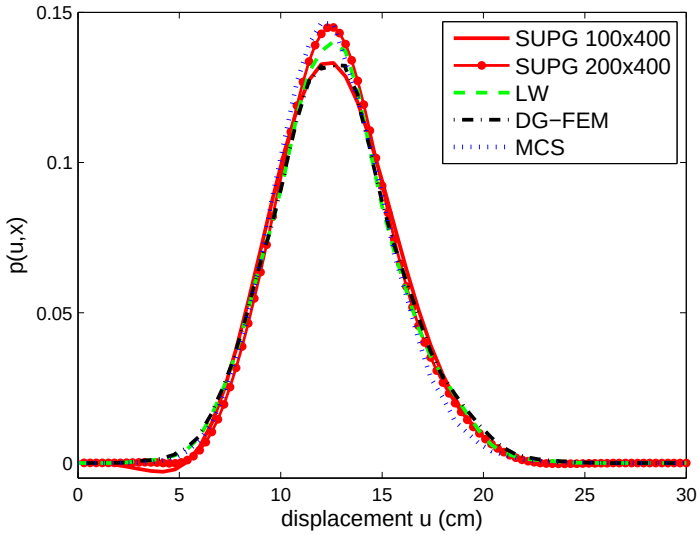


Figure 6.17: PDF profiles ta x=30.00m for a) SUPG 100x400, b) SUPG 200x400, c) modified LW and d) DG-FEM versus brute force Monte Carlo

Finally, an assessment of the performance of the proposed methodologies is attempted

here on the basis of computing time required to reach the solution. In order for the comparison to be fair we used the exact grid that gives an error of 3% in variance for each method. More specifically, we considered a 84×160, a 241×513 and a 400×851 grid for SUPG, LW and DG-FEM respectively. The computational cost for each method in terms of total CPU time is illustrated in table 6.2. From this table it is apparent that the SUPG method is the most efficient choice, while the LW and the DG-FEM require six to seven times the cost of SUPG.

Table 6.2: Computational cost for each method

	solving one pde	number of analyses	total CPU time
SUPG	0.289 *sec*	286	82.709 *sec*
Lax-Wedroff	1.829 *sec*	286	523.032 *sec*
DG-FEM	1.904 *sec*	286	544.663 *sec*

Example 3. Plane stress problem

Finally, the L-shaped plane stress problem of figure 6.18 is considered, in order to demonstrate the applicability of the proposed methodology to general stochastic finite element systems. In this example the modulus of elasticity E, the forces F and the Poisson ratio ν are assumed random and follow a Gaussian distribution with properties the ones presented in table 6.3. For the solution of the physical problem, the plate is discretized into 340 quadrilateral finite elements resulting in 682 dofs. In this case, our interest lies in estimating the evolution of the pdf of the vertical displacement u_v. It should be noted here that, despite the fact that the problem is two-dimensional, the one-dimensional version of the GDEE in eq. (5.58) can be implemented in the same way, that it was implemented in the previous example 2. Therefore, a path has to be selected at first, in order to apply the proposed methodology. Such a path is depicted with the red (dashed) line in figure 6.18 and its total length is $3.1m$.

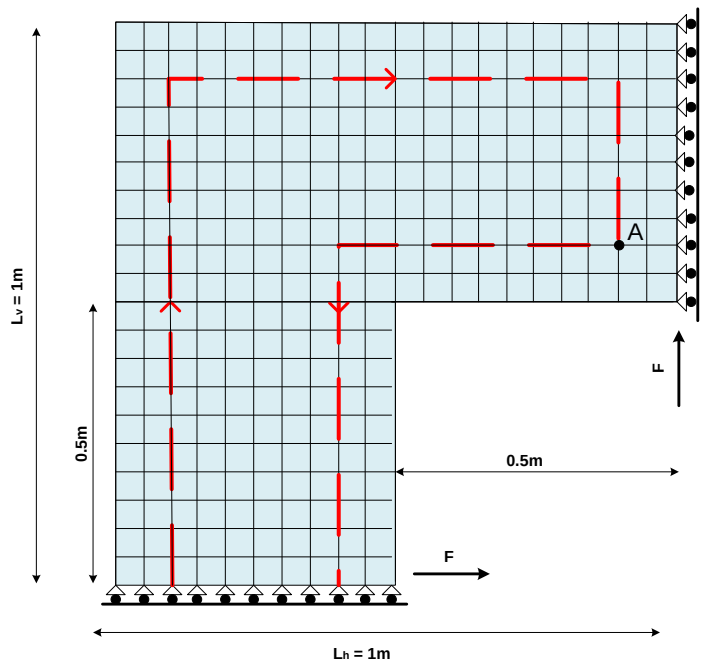


Figure 6.18: L-shaped plane stress problem

Table 6.3: Random parameters of the structure

	E	F	ν
mean	21 GPa	5000 kN	0.3
COV	0.2	0.2	0.2

Figure 6.19 plots the variance of the vertical displacement along the selected path computed with the PDEM with SUPG and MCS. As shown in this figure, the results are in close agreement. Figure 6.20 presents the evolution of the pdf of the vertical displacement $p(u_v, x)$ along the length of the selected path for the SUPG. It becomes obvious that SUPG indeed manages to capture accurately the probabilistic information at the boundaries. An intersection with a vertical plane parallel to the u_v-axis at position $x = 2.0$m, which corresponds to node A with coordinates (0.9 , 0.6), will give the pdf profile of the vertical displacement at this position, as illustrated in fig. 6.21. Again the

results between MCS and PDEM are in almost perfect match. Lastly, instead of the chosen path we could consider a path that runs through all the nodes of the problem and estimate the complete probabilistic characteristics of the vertical displacement. In this case though, boundary conditions would have to be imposed not only at the edges of the path, but also at the interior positions where the path reaches a support that constraints the vertical displacement. SUPG is the only method that enables insertion of such constraints in the resulting system of algebraic equations (eq. (6.12)). This property renders SUPG the only method that can be applied directly to such general finite element systems since it is the only capable of gathering the probability at the certain event of zero displacement at the supports. The results of this calculation are presented in the contour plots of the variance of the vertical displacement depicted in figures 6.22a and 6.22b for MCS and SUPG, respectively. These results demonstrate a fair accuracy of the SUPG with respect to the 'exact' MCS solution.

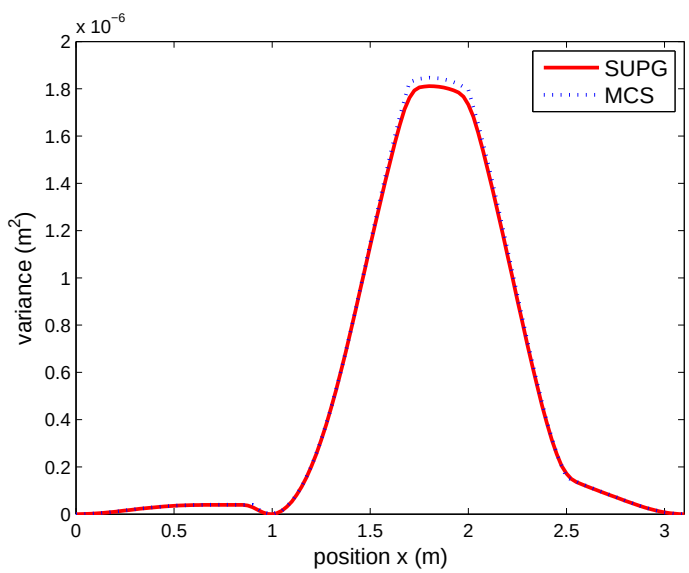


Figure 6.19: Comparison in variance between SUPG and brute force Monte Carlo simulation

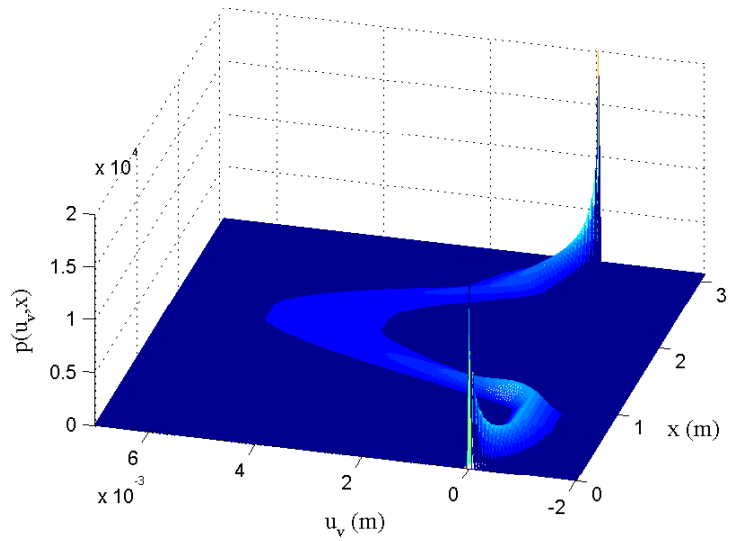


Figure 6.20: Evolution of the probability density function of the vertical displacement $p(u_v, x)$ obtained by the PDEM with the SUPG formulation

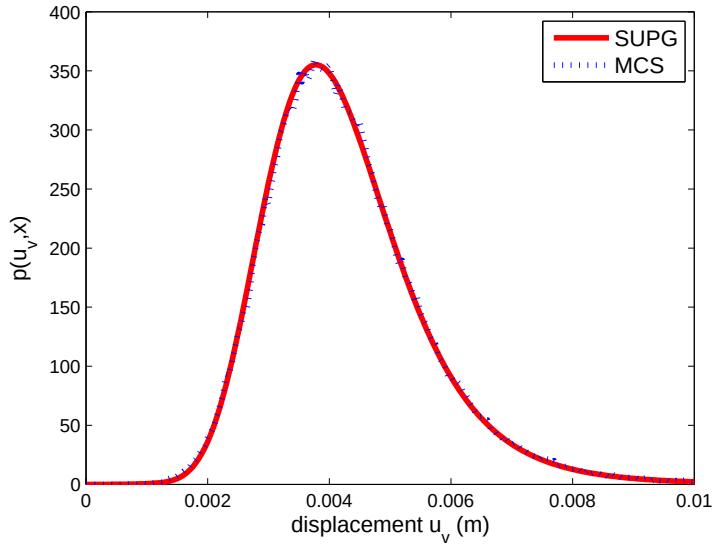
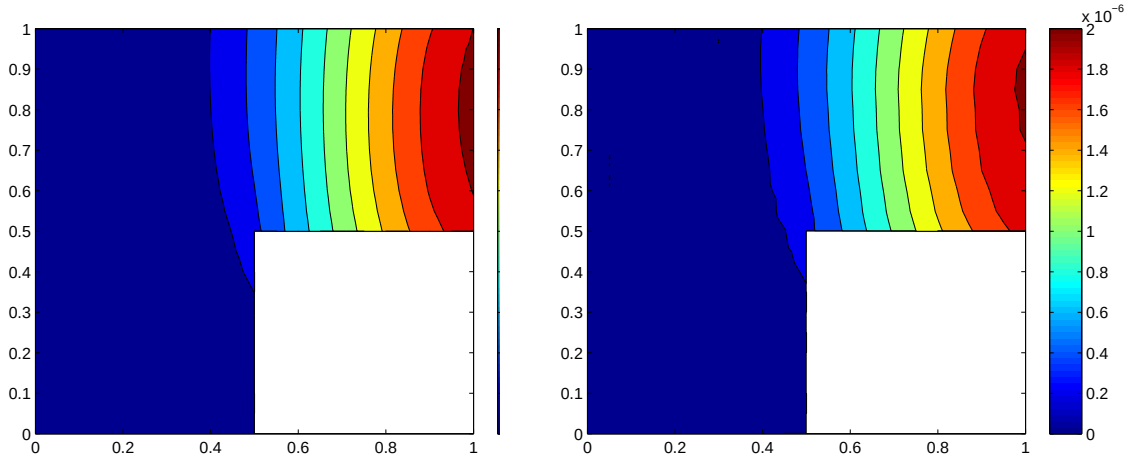


Figure 6.21: PDF profiles at x=2.0m for the SUPG and brute force Monte Carlo simulation



(a) Monte Carlo Simulation **(b)** StreamlineUpwind/Petrov Galerkin

Figure 6.22: Contour plots of variance for a) the Monte Carlo Simulation and b) the SUPG

6.6 Conclusions

In this chapter, an alternative methodology of applying the probability density evolution method in general stochastic finite element systems is outlined. With the proposed methodology the probability density function at a given position of the structure can be estimated via the solution of a series of pure advection pde's and thus, the complete probabilistic information of the system under consideration can be accurately and efficiently captured. In an attempt to increase the accuracy of the most frequently used finite difference LW schemes, which are used in order to solve the pde's involved in the method, two alternative Galerkin-based techniques were investigated, namely, the discontinuous Galerkin finite element method and the streamline upwind/Petrov Galerkin method. The results were compared to the LW equipped with the TVD property, as well as with brute force Monte Carlo Simulation. Our investigations showed SUPG to be more efficient, having higher convergence rates than the other two methods, while being the only method that could capture accurately the evolution of pdf near the boundaries of the structure.

7

Limit analysis of stochastic structures in the framework of the Probability Density Evolution Method

7.1 Introduction

In this chapter a Probability Density Evolution formulation is proposed for the limit analysis of stochastic systems, which can accurately and efficiently evaluate the effect the system's random parameters have on its nonlinear and limit response. The proposed formulation of the classic Probability Density Evolution Method reduces the corresponding Generalized Density Evolution equations (GDEEs), which are partial differential equations, to a system of ordinary differential equations, that can be efficiently solved with the method of characteristics [40, 86]. With this reformulation, the cumulative distribution function of the critical load of the structure can be accurately and efficiently evaluated. The estimation of stochastic limit buckling loads of imperfection sensitive structures is a natural extension of this method. In addition, a methodology is put forward for the estimation of the probabilistic characteristics of the full load-displacement curve for a stochastic nonlinear system in the context of Newton-Raphson incremental-iterative schemes. The main advantage of the proposed approaches is that they allow

for a quantification of the effect of uncertainties on the structural capacity, with only a small number of deterministic analyses compared to a Monte Carlo simulation. The applicability and validity of the proposed methodology for limit and buckling structural analysis is verified through extensive numerical investigations.

The remaining of the chapter is organized as follows: In section 2, the formulation of the PDEM in static cases is outlined. In section 3, the partitioning of the probability-assigned space, as well as the Karhunen-Loève expansion for random fields are described. Section 4 presents the implementation aspects of applying the PDEM in the estimation of stochastic limit loads. In section 5, an approach is put forward that enables the coupling between PDEM and Newton-Raphson solution algorithms. Lastly, in section 6, the proposed methodology is applied in test cases and the results obtained are critically assessed and discussed further.

7.2 PDEM for the estimation of stochastic limit loads

It is well understood that in a system where randomness is involved in the system parameters and/or loading, the critical engineering demand parameter at any limit state (i.e. collapse load) will not have a unique value, rather it will be a random variable describing the probabilistic limit response of the structure. The PDF of this limit parameter (i.e. limit load) variable can be efficiently estimated with a customized formulation of the PDEM methodology as follows:

Without loss of generality, consider the cantilever beam of figure 7.1, subjected to a vertical load at its end, F. The system parameters $\boldsymbol{\theta}$ (e.g. the modulus of elasticity, the moment of inertia, the cross-sectional area and the yield stress) are all considered as random.

Denoting $F_L(\boldsymbol{\theta})$ the limit collapse load as depicted in figure 7.2, this will be different for any random event θ_q (realization) in $\boldsymbol{\theta}$.

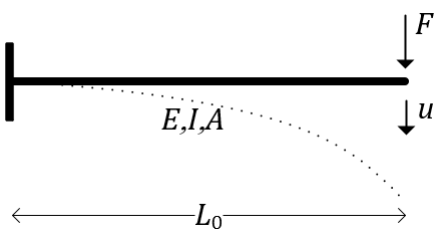


Figure 7.1: Cantilever under vertical load

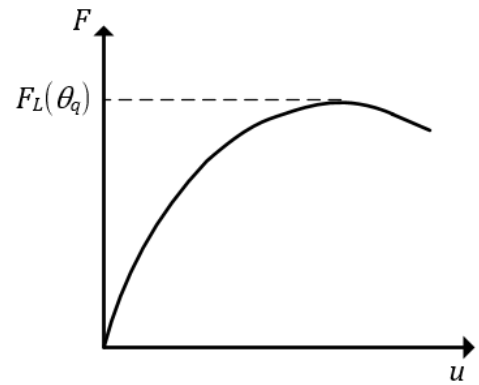


Figure 7.2: Limit collapse load for a given realization θ_q

It is obvious, that F_L will have a range of possible values with a different probability of occurrence assigned at each value. In other words, it will be a random variable, whose PDF we want to estimate. Defining an indicator function $i(F_L, \theta_q)$ as:

$$i(F_L, \theta_q) = \begin{cases} 1, & \text{if } F_L >= F_L(\theta_q) \\ \\ 0, & \text{if } F_L < F_L(\theta_q) \quad \text{for } q = 1, 2, ..., n_{pt} \end{cases} \tag{7.1}$$

then i takes the value of 1 when the structure collapses and 0 otherwise. In other words, i is the Heaviside Step function, that is

$$i(F_L, \theta_q) = H(F_L - F_L(\theta_q)) \;\; \text{for } q = 1, 2, ..., n_{pt} \tag{7.2}$$

Obviously, the joint pdf of $p_{i\boldsymbol{\theta}}$ will be a function of $i(F_L, \boldsymbol{\theta})$ and F_L. At this point it is interesting to note the correspondence between $i(F_L, \boldsymbol{\theta})$ and $u(x, \boldsymbol{\theta})$ as defined in eq. (6.3) of chapter 5, where now F_L plays the role of x. Then, the GDEE's that govern the flow of probability can be written as

$$\frac{\partial p_{i\boldsymbol{\theta}}(i,\theta_q,F_L)}{\partial F_L} + \frac{di(\theta_q,F_L)}{dF_L}\frac{\partial p_{i\boldsymbol{\theta}}(i,\theta_q,F_L)}{\partial i} = 0, \qquad q = 1,2,...,n_{pt} \quad (7.3)$$

The flux term $\frac{di(\theta_q,F_L)}{dF_L}$ in equation (7.3), is actually the Dirac function $\delta(F_L - F_L(\theta_q))$, defined as

$$\delta(F_L - F_L(\theta_q)) = \begin{cases} \infty, & \text{if } F_L = F_L(\theta_q) \\ \\ 0, & \text{if } F_L \neq F_L(\theta_q) \end{cases} \quad (7.4)$$

The initial conditions needed to solve these equations derive naturally from the fact that for zero compressive force F_L, we are certain that there will be no failure. This statement can be mathematically written as

$$p_{i\boldsymbol{\theta}}(i,\boldsymbol{\theta},F_L)\mid_{F_L=0} = \delta(i-0)\bar{p}_q \quad (7.5)$$

7.2.1 Solution of the GDEE with the method of characteristics

Having an analytical expression for the flux term, allows us to seek an analytical solution for the GDEE in (7.3), via the method of characteristics [40, 86]. This method is valid for any hyperbolic partial differential equation and it actually reduces the partial differential equation into a system of ordinary differential equations.
To simplify notation and without introducing any ambiguities, we rewrite eq. 7.3 as follows

$$\frac{\partial p}{\partial F_L} + \delta(F_L - \bar{F}_L)\frac{\partial p}{\partial i} = 0 \quad (7.6)$$

where, $\bar{F}_L$ denotes the value of the critical load for this particular GDEE, that is, $F_L(\theta_q)$. Transforming the first order PDE of eq. 7.6 into an ODE along some path s, we obtain

$$\frac{d}{ds}p\big(i(s),F_L(s)\big) = \mathcal{F}\big(p,i(s),F_L(s)\big) \quad (7.7)$$

where $\big(i(s), F_L(s)\big)$ is some trajectory (characteristic line).Applying now the chain rule for the differentiation with respect to s we obtain

$$\frac{d}{ds}p = \frac{\partial p}{\partial F_L}\frac{dF_L}{ds} + \frac{\partial p}{\partial i}\frac{di}{ds} \tag{7.8}$$

Setting $\frac{dF_L}{ds} = 1$ and $\frac{di}{ds} = \delta(F_L - \bar{F}_L)$ we get the left hand side of the PDE. In other words, along the characteristic lines $\big(i(s), F_L(s)\big)$ the original PDE becomes the ODE: $\frac{d}{ds}p = 0$, which means that the solution is constant along the characteristic line, or, $p\big(i(s), F_L(s)\big) = p(i, 0)$. Thus, the PDE solution degenerates to a solution of the following system of ODEs with unknowns the $F_L(s)$, $i(s)$ and $p(s)$. More specifically,

$$\frac{dF_L}{ds} = 1 \tag{7.9}$$

and by letting $F_L(0) = 0$, then $F_L = s$.

$$\frac{di}{ds} = \delta(F_L - \bar{F}_L) \tag{7.10}$$

letting $i(0) = k$ and using the fact that $F_L = s$ then eq.(7.10) becomes

$$\frac{di}{dF_L} = \delta(F_L - \bar{F}_L) \Rightarrow i = H(F_L - \bar{F}_L) + k \tag{7.11}$$

and thus

$$k = i - H(F_L - \bar{F}_L) \tag{7.12}$$

Next, we solve the equation

$$\frac{dp}{ds} = 0 \tag{7.13}$$

letting $p(0) = f(i(0)) = f(k) \overset{eq.(7.5)}{=} \delta(k - 0)\bar{p}_q$, then

$$p\big(i, F_L\big) = f(k) \overset{eq.(7.12)}{=} \delta\big(i - H(F_L - \bar{F}_L)\big)\bar{p}_q \tag{7.14}$$

In figure 7.3, a typical solution of a GDEE of this type is presented. From this figure it becomes apparent that the probability $\bar{p}$ assigned to each realization θ_q is lumped at the non-collapse event (safety) for as long as $F_L < F_L(\theta_q)$. When F_L takes the value of

$F_L(\theta_q)$, the PDF displays a jump, that gathers all probability at the event of collapse and remains constant for $F_L > F_L(\theta_q)$.

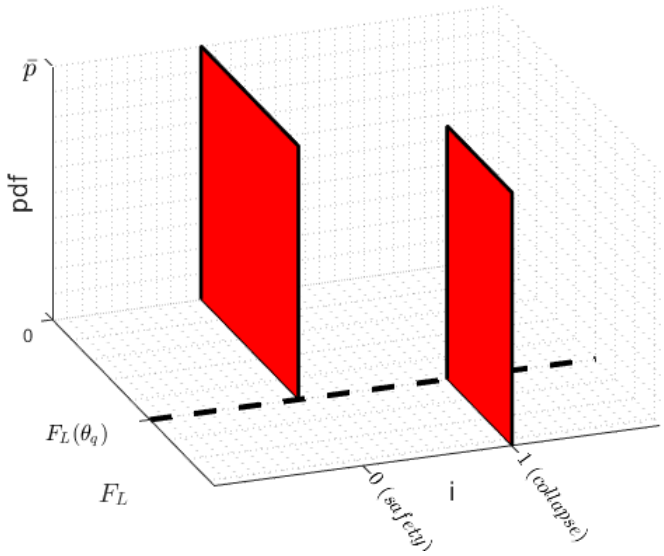


Figure 7.3: Solution of a GDEE for a given θ_q

Overlaying the solutions for all θ_q's for $q = 1, 2, ..., n_{pt}$ we obtain figure 7.4, and the solution profile for $i = 1$ is in fact the CDF of the critical load for structural failure. As will be shown, the main advantage of this approach is that it achieves almost the same accuracy with the MCS but with a significantly smaller number of deterministic analyses, which is a direct consequence of the partitioning of the probability-assigned space.

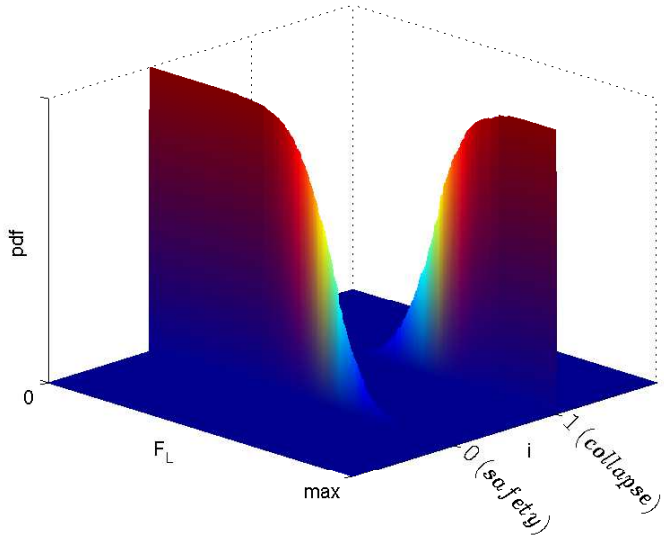


Figure 7.4: Probability density function of the F_L values

7.3 PDEM for the estimation of the complete stochastic nonlinear path

The purpose of deterministic nonlinear analysis of structures is to trace its equilibrium path. In classical nonlinear FEM, this is achieved by taking the equilibrium in the deformed configuration leading to a system of nonlinear equations that are solved in the context of a Newton-Raphson incremental iterative scheme, using a load or displacement control formulation. However, in the case of stochastic systems, the load-displacement curve will be a function of the system's random parameters, or in other words, it will be a random field. In this section, a methodology is put forward, that captures the complete probabilistic characteristics of the load-displacement curve, in the PDEM framework.

7.3.1 Displacement control in the framework of PDEM

In the deterministic displacement control method, an incremental displacement $\bar{u}$ is imposed along a monitored degree of freedom and the load $F(\bar{u})$, corresponding to this displacement is evaluated through an iterative Newton-Raphson procedure. In the stochastic case, though, a deterministic displacement along one dof can be the result of

various load values each with a different probability, due to the randomness in the system parameters, that is, $F(\bar{u}) = F(\bar{u}, \boldsymbol{\theta})$. Then, the corresponding GDEE's that govern the evolution of the pdf of F along the values of $\bar{u}$ are (similarly to equation (7.3))

$$\frac{\partial p_{F\boldsymbol{\theta}}(F, \theta_q, \bar{u})}{\partial \bar{u}} + \frac{dF(\theta_q, \bar{u})}{d\bar{u}} \frac{\partial p_{F\boldsymbol{\theta}}(F, \theta_q, \bar{u})}{\partial F} = 0, \quad q = 1, 2, ..., n_{pt} \tag{7.15}$$

Denoting

$$p_q(F, \bar{u}) = \int_{\Omega_q} p_{F\boldsymbol{\theta}}(F, \boldsymbol{\theta}, \bar{u}) \mathrm{d}\boldsymbol{\theta}, \ q = 1, 2, ..., n_{pt} \tag{7.16}$$

then, the joint pdf of F and $\bar{u}$ is

$$p(F, \bar{u}) = \sum_{q=1}^{n_{pt}} p_q(F, \bar{u}) \tag{7.17}$$

7.3.2 Load control in the framework of PDEM

In contrast to the previous case, in the formulation of the load control method, a load increment is applied at each step of the analysis. Here, the displacement of the monitored dof will be a random field with respect to F and its pdf can be evaluated by the following equations

$$\frac{\partial p_{\bar{u}\boldsymbol{\theta}}(\bar{u}, \theta_q, F)}{\partial F} + \frac{d\bar{u}(\theta_q, F)}{dF} \frac{\partial p_{\bar{u}\boldsymbol{\theta}}(\bar{u}, \theta_q, F)}{\partial \bar{u}} = 0, \quad q = 1, 2, ..., n_{pt} \tag{7.18}$$

Denoting

$$p_q(\bar{u}, F) = \int_{\Omega_q} p_{\bar{u}\boldsymbol{\theta}}(\bar{u}, \boldsymbol{\theta}, F) \mathrm{d}\boldsymbol{\theta}, \ q = 1, 2, ..., n_{pt} \tag{7.19}$$

then, the joint pdf of F and $\bar{u}$ is

$$p(\bar{u}, F) = \sum_{q=1}^{n_{pt}} p_q(\bar{u}, F) \tag{7.20}$$

7.4 Numerical examples

Two test cases were selected in order to demonstrate the applicability of the proposed methodolgies to general static systems, as well as their efficiency in comparison to the Monte Carlo Simulation method. A portal frame under pushover loading conditions and an imperfection sensitive shallow shell with random imperfections.

7.4.1 Example 1: Portal Frame

For the first test case, we considered the 2-storey, 1-bay frame of figure 7.5. The height of each storey is $4m$ and the width is $6m$. The columns and beams of the frame are HEB 220 steel beams with yield stress $f_y = 355Mpa$. All internal element sections are considered as elastic, while the plasticity is assumed concentrated in the element edge sections, modeled with nonlinear rotational springs following the Modified Ibarra Krawinkler Deterioration Model [96] of figure 7.6.

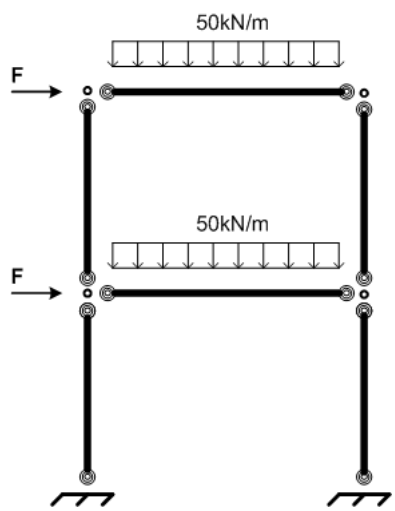


Figure 7.5: 2-storey, 1-bay frame

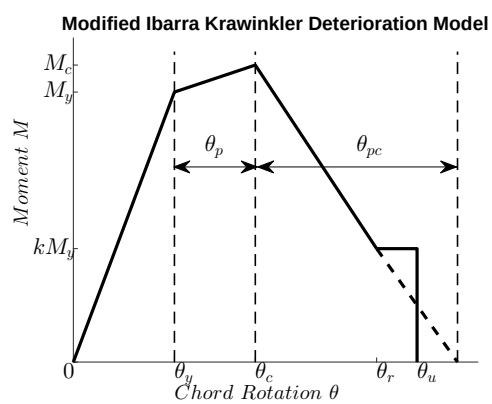


Figure 7.6: Moment-rotation law for rotational springs

The yield moment M_y and the rotational spring parameters θ_p, θ_{pc}, θ_u for each element are assumed to be random variables following the gaussian distribution, leading to a total of 24 random variables in the system under consideration. Their properties are given in table 7.1. For the partitioning of the resulting probability assigned space of the random variables the RQSPM [23] method was employed with 1072 points, leading to

the same number of required deterministic analyses of the system. As shown in figure 7.5 the frame's beams are subjected to a uniform loading of $50kN/m$, while two pushover equal lateral forces, F, act on each storey.

The limit (ultimate) value of F,F_L, at collapse is obviously a random variable, the properties of which depend both on the system's characteristics, as well as the random input parameters in it. Figure 7.7 plots the pushover curves of F versus the top displacement for a selection of values of the input random variables. The cumulative distribution function of the limit values of the lateral loads F_L is calculated using the procedure described in Section 4 and a plot of this CDF is demonstrated in figure 7.8. These results are compared to those obtained for 50000 MC simulations and an almost perfect match is reported. By numerical differentiation of this CDF we get the probability density function of F_L of figure 7.9. Figure 7.10 demonstrates the convergence in variance as a function of the number of MC simulations. The value of variance obtained from 180000 MC analyses, found to be $5.3092 \cdot 10^8 N^2$, is considered as an 'exact' reference value for comparison purposes. This reference value is denoted with the red solid line in figure 7.10, while the solid black line in the same figure denotes the variance obtained from PDEM after 1072 deterministic analyses. The value of variance from PDEM was found to be $5.3611 \cdot 10^8 N^2$, which results in an error less than 1%. From fig. 7.10 it is also evident, that in order for the MC to converge to the 'exact' value of variance with error lesser than 1% at least 44000 MC simulations are required. This illustrates the significant gain in computational time with the proposed approach. In figures 7.11 and 7.12 the mean load-displacement curve and its standard deviation are plotted, respectively, by coupling PDEM and the displacement control algorithm. The results are compared to MCS and fair accuracy can be reported. Lastly, figure 7.13 presents the "flow" of probability of F along the increments of u, as obtained from PDEM.

Table 7.1: Random parameters of the structure

	M_y	θ_p	θ_{pc}	θ_u
mean	293.59 kNm	0.02	0.25	0.18
COV	0.2	0.1	0.1	0.1

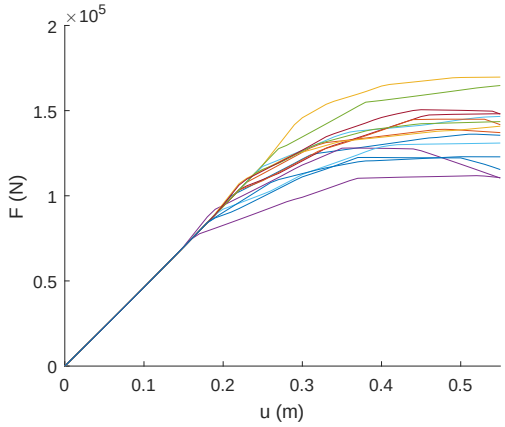


Figure 7.7: F versus top diplacement

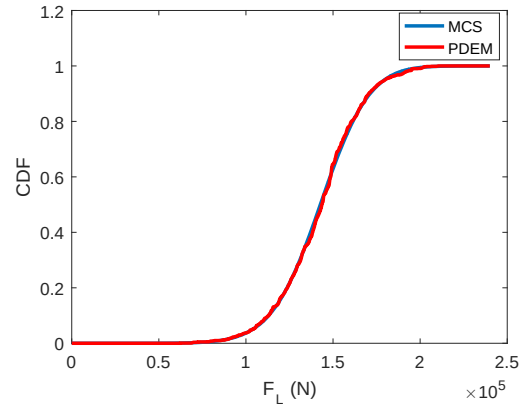


Figure 7.8: cdf of F_L for PDEM vs. MCS

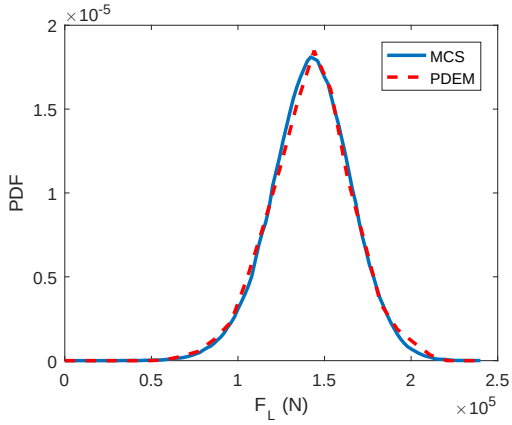


Figure 7.9: pdf of F_L for PDEM vs. MCS

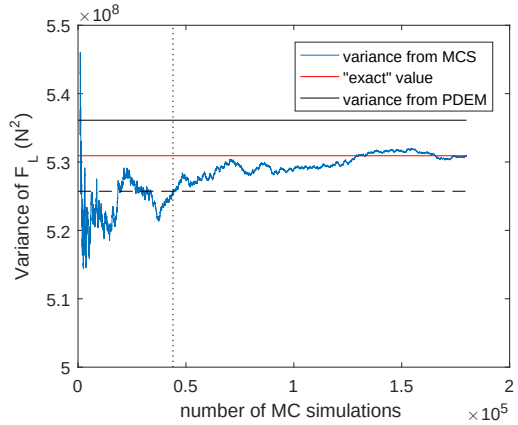


Figure 7.10: Convergence in variance of F_L

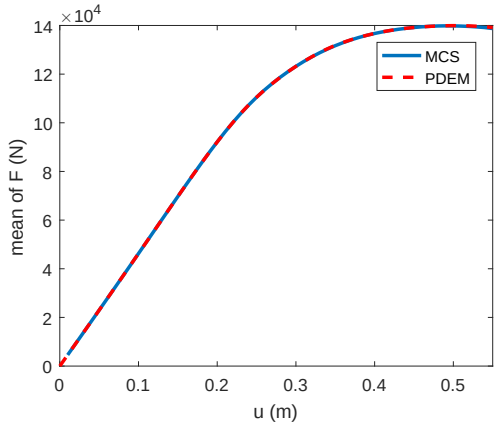


Figure 7.11: Mean value of F as a function of the displacement u

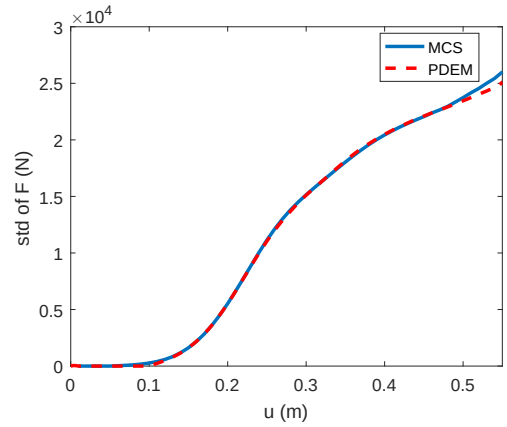


Figure 7.12: Standard deviation of F as a function of the displacement u

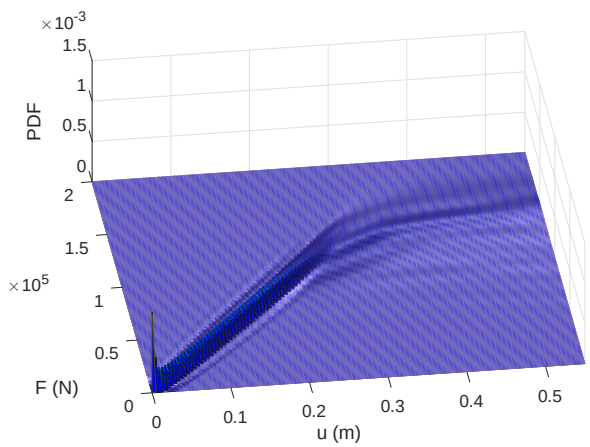


Figure 7.13: PDF of F as a function of the displacement

7.4.2 Example 2: Shallow Shell

The second test case is the hinged isotropic panel shown in figure 7.14, which has been extensively studied in [109, 107]. This example is of particular interest because of the dominant role of physical non-linearities on the structure's response. The geometrical and materials properties of the hinged isotropic panel can also be seen in figure 7.14. An elastoplastic model is adopted and geometrical non-linearities were also included in the

analysis. The problem is formulated in the nonlinear framework using the displacement control path following method.

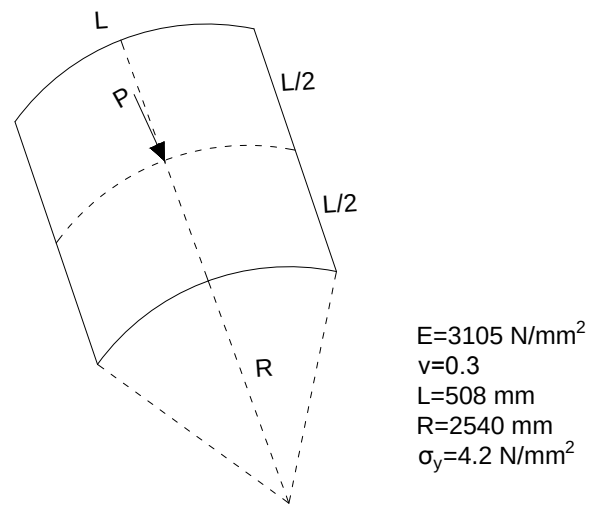


Figure 7.14: Hinged isotropic panel

The modulus of elasticity was modelled as a 2-D homogeneous lognormal stochastic field with mean value $3105MPa$ and standard deviation 10% of the mean. Since lognormal fields can be treated as the result of the exponentiation of an underlying Gaussian field, then $E(x, y)$ can be written in the form:

$$E(x, y) = exp(\mu_E + \sigma_E f(x, y)) \tag{7.21}$$

where μ_E is the mean value of the underlying Gaussian field, equal to 8.0358, σ_E the corresponding standard deviation equal to 0.0998 and $f(x, y)$ is a zero mean unit variance stochastic field corresponding to the variability of E. For the autocorrelation function of $f(x, y)$, the modified exponential kernel proposed by Spanos et al. [147] was employed. This kernel is differentiable at the origin which renders it computationally more efficient than the classic exponential kernel, and also, a more suitable choice to model natural phenomena. It can be written for 2-D fields as follows:

$$R_f(x_1, x_2; y_1, y_2) = exp\left[-\left(\hat{a}_x|x_1 - x_2| + \hat{a}_y|y_1 - y_2|\right)\right](1 + \hat{a}_x|x_1 - x_2|)(1 + \hat{a}_y|y_1 - y_2|) \tag{7.22}$$

Eq. (7.22) can be written as

$$R_f(x_1, x_2; y_1, y_2) = R_f^{1D}(x_1, x_2) R_f^{1D}(y_1, y_2) \tag{7.23}$$

with

$$R_f^{1D}(x_1, x_2) = exp\left(-\hat{a}_x |x_1 - x_2|\right)\left(1 + \hat{a}_x |x_1 - x_2|\right) \tag{7.24}$$

$$R_f^{1D}(y_1, y_2) = exp\left(-\hat{a}_y |y_1 - y_2|\right)\left(1 + \hat{a}_y |y_1 - y_2|\right) \tag{7.25}$$

where the superscript ^{1D} refers to the solution of the one-dimensional eigenvalue problem. Then, the spectral representation of the 2-D field can be obtained by products of one-dimensional solutions as follows:

$$\lambda_i = \lambda_{i_1}^{1D} \lambda_{i_2}^{1D} \tag{7.26}$$

$$\phi_i(x, y) = \phi_{i_1}(x) \phi_{i_2}(y) \tag{7.27}$$

The n greatest products in eq.(7.26) are selected with their corresponding eigenfunctions (eq.(7.27)). Then, the approximate field becomes:

$$f(x, y; \omega) = \sum_{i=1}^{n} \sqrt{\lambda_i} \xi_i(\omega) \phi_i(x, y) \tag{7.28}$$

In a similar fashion, for the representation of the imperfect geometry, the radius of the structure was considered as a 2D homogeneous Gaussian field

$$r(x, y) = R + g(x, y) \tag{7.29}$$

with $r(x, y)$ being the varying radius at each point of the structure, and $g(x, y)$ is again a zero mean unit variance homogeneous field, decomposed similarly to $f(x, y)$ via eq. (7.28). The following assumption for the corresponding correlation parameters was made: $\hat{a}_x = \hat{a}_y = 0.008$ for both cases.

For the decomposition of the field 5 KL terms were used in each direction and the 10 terms with the greatest contribution were retained from their products. Figure 7.15 illustrates the first 10 eigenvalues of the expansion for the 1-D kernel. The kernel approx-

imations for the 1-D and 2-D cases are depicted in figures 7.16 and 7.17, respectively. The average errors, given by eqs. (7.30) and (7.31) are 0.0078 for the 1-D case and 0.069 for the 2-D, which is deemed adequate for the purposes of this work. Also, in figures 7.18, 7.19 two random representations for the material and initial imperfection fields, are displayed, respectively.

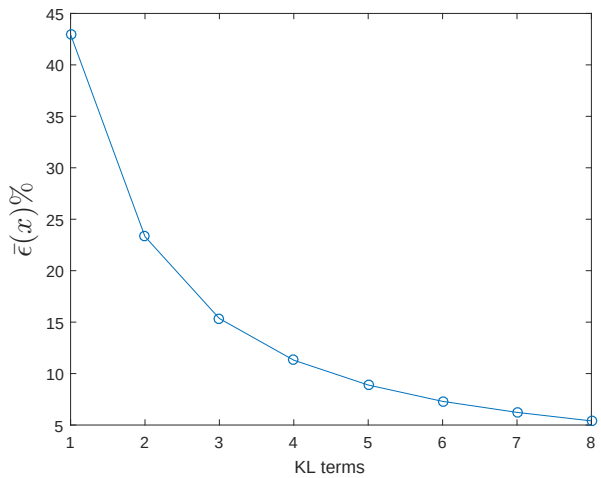


Figure 7.15: Eigenvalues of the 1D modified exponential kernel

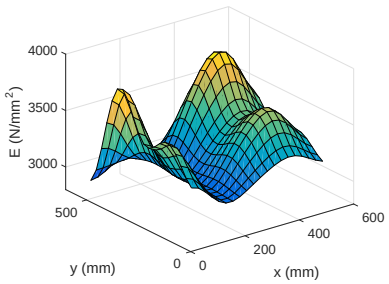


Figure 7.16: Five-term approximation of the 1D autocorrelation kernel

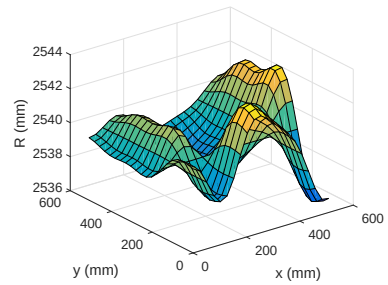


Figure 7.17: Ten-term approximation of the 2D autocorrelation kernel

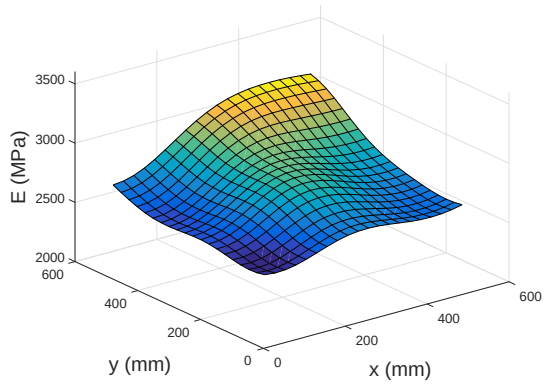


Figure 7.18: 2D material imperfection shape for $\sigma^2 = 1$ and $\hat{a}_x = \hat{a}_y = 0.008$

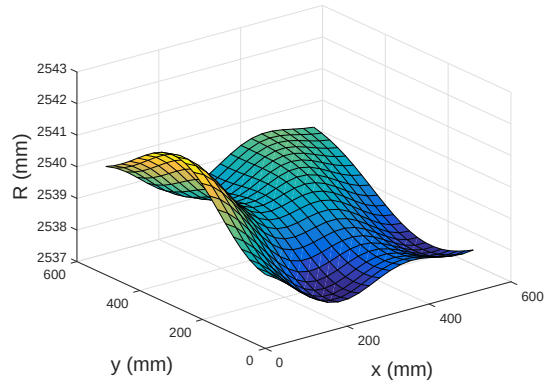


Figure 7.19: 2D initial imperfection shape for $\sigma^2 = 1$ and $\hat{a}_x = \hat{a}_y = 0.008$

$$err_{avg}^{1D} = \frac{\int_0^L |\sigma_f^2 - \sum_{i=1}^5 \lambda_i^{1D} \phi_i^2(x)| dx}{L} \tag{7.30}$$

$$err_{avg}^{2D} = \frac{\int_0^L \int_0^L |\sigma_f^2 - \sum_{i=1}^{10} \lambda_i \phi_i^2(x,y)| dx dy}{L \times L} \tag{7.31}$$

The partitioning of the probability-assigned space was achieved through the RQ-SPM, which required 276 representative point sets in the cases of the 10 random variables (stand alone case of material or geometry imperfections only) and 552 in the case of 20 random variables. The panel is discretized with a 20×20 mesh of quadrilateral shell elements, which is considered to be dense enough both for the convergence of the static nonlinear analyses, as well as for the accurate representation of the fluctuations of the fields [88]. The ultimate load for the perfect configuration with no imperfections at all is found to be $P_u = 1240N$.

The effect of material and geometry imperfections on the ultimate load is studied at first separately and then combined. In figures 7.20 and 7.21 the mean value of P and its standard deviation are plotted against the displacement, for the cases of material imperfections, geometric imperfections and their combination. Also, figures 7.22, 7.23 and 7.24 depict the CDF of the critical load P_u obtained from the proposed methodology, along with those obtained from a large number of Monte Carlo simulations.

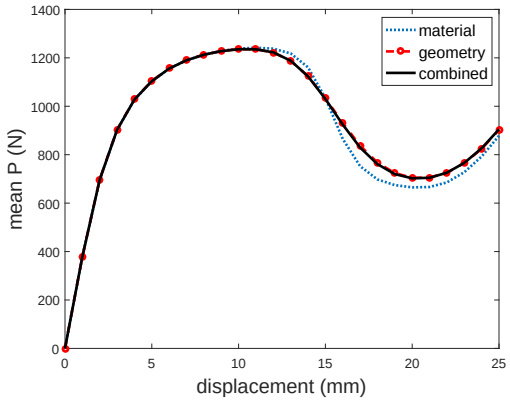


Figure 7.20: Mean value of P, as a function of the displacement for material, geometric and combined imperfections

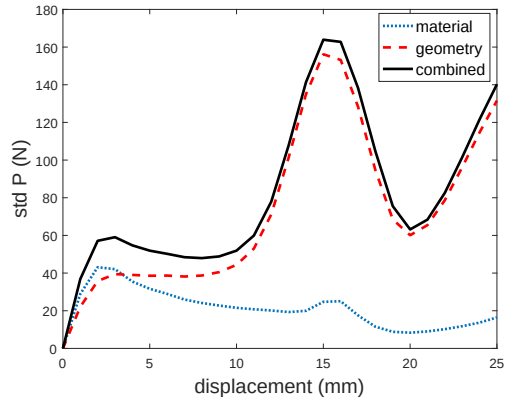


Figure 7.21: Standard deviation of P, as a function of the displacement for material, geometric and combined imperfections

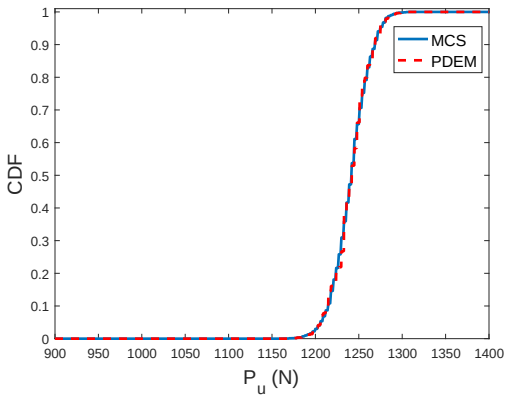


Figure 7.22: Cumulative distribution function of the critical load P_u for material imperfections

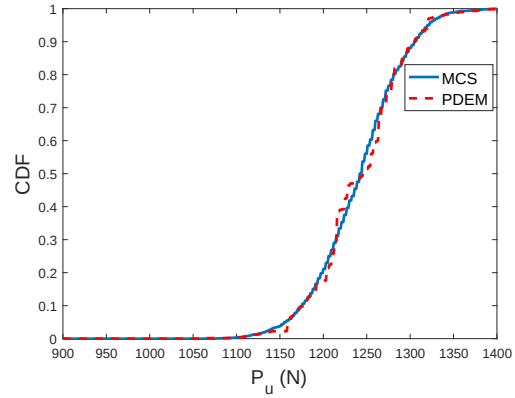


Figure 7.23: Cumulative distribution function of the critical load P_u for geometric imperfections

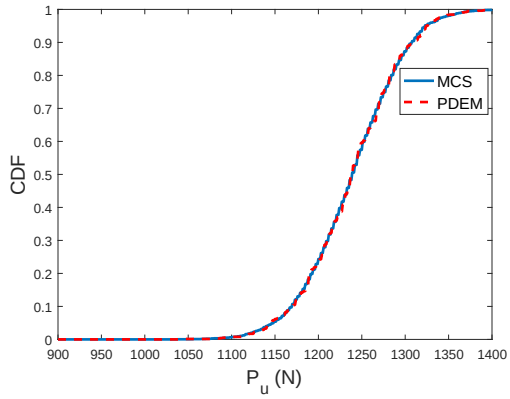


Figure 7.24: Cumulative distribution function of the critical load P_u for the combined geometric and material imperfections

In all cases the results are of fair accuracy. Hence, from the above inspection it becomes evident that the proposed method is capable of producing accurate results with only a small number of deterministic analyses, compared to the tens of thousands required for the convergence of an MC simulation.

7.5 Conclusions

In this chapter, a customized formulation of the PDEM for the limit analysis of stochastic structures was introduced, which relies on the law of conservation of probability. In the context of PDEM, the multidimensional probability space of the input's random variables was partitioned via the RQ-SPM method into subdomains and representative point sets, associated to each subdomain. The proposed approach enables the accurate and efficient estimation of the CDF of the critical load for the problem at hand by solving the deterministic equilibrium equation for the prescribed point sets and consequently, by solving the corresponding first-order partial differential equations that express the flow of probability, solved analytically using the method of characteristics. The method was naturally extended to the nonlinear analysis of stochastic structures, where a methodology for the estimation of the probabilistic characteristics of the full load displacement path via PDEM was put forward. The numerical tests performed demonstrated the applicability of the method as well as its accuracy and efficiency.

8

A spectral stochastic formulation for nonlinear framed structures

8.1 Introduction

The present chapter presents a stochastic formulation which enables the effective coupling of spectral stochastic finite elements with geometrically nonlinear analysis of framed structures [100, 44, 84]. This is achieved by projecting the stochastic part of the incremental displacements, formulated in the framework of a Newton-Raphson solution scheme, to a polynomial chaos expansion basis. The nonlinear solution is consequently achieved by introducing these expansion coefficients in the iterative solution algorithm as unknowns to be evaluated. With this approach, an augmented system of nonlinear equations is produced and the values of the coefficients are obtained through convergence of the iterative algorithm. In a similar to the deterministic problems manner, the nonlinear load or displacement control algorithms are extended to their stochastic counterparts and a discussion on the appropriate choice of such algorithms depending on the problem at hand is presented. This work focuses on stochastic beam finite element systems in which uncertainty is considered in the system properties. The efficiency of the proposed methodology is demonstrated in benchmark problems with strong geometrically nonlinear behavior. In order to verify the validity of the proposed approach the results

obtained are compared to those of Monte Carlo simulations and fair accuracy can be reported.

The body of this chapter has been divided to into four sections, in order to give the theoretical background for the implemented methodologies, with regards to the stochastic path following techniques. More specifically, the remaining of this paper is organized as follows: In section 2, the fundamentals of SSFEM are outlined and an extension of the method for co-rotational elements in the the context of nonlinear analysis is put forward. In section 3, a reformulation of the standard iterative solution algorithms, that is the load control and the displacement control, is presented. In section 4, the computational aspects of the proposed methodology are illustrated and discussed further. Lastly, section 5 includes numerical examples to test the validity and accuracy of the proposed methods.

8.2 SSFEM for geometrically nonlinear systems

Having established the preliminary background for linear systems, we are able to proceed with the formulation of the SSFEM for geometrically nonlinear systems. As mentioned previously, in a Langrangian description of a geometrically nonlinear problem the equation of equilibrium is expressed through Newton-Raphson incremental-iterative scenario

$$\mathbf{K}_t^{i-1}\delta\mathbf{u}^i = \mathbf{R}^i - \mathbf{f}^{i-1} \tag{8.1}$$

where, $\mathbf{K}_t^{i-1}$, $\mathbf{R}^i$ and $\mathbf{f}^{i-1}$ are the tangent stiffness matrix, the external load vector and the internal force vector, respectively. The superscript i refers to i-th iteration of the solution algorithm. In eq. (8.1) $\delta\mathbf{u}^i$ are the incremental displacements of iteration i, which may be added to the displacements of the previous iteration $\mathbf{u}^{i-1}$ in order to compute the total displacements as:

$$\mathbf{u}^i = \mathbf{u}^{i-1} + \delta\mathbf{u}^i \tag{8.2}$$

Randomness in the system properties will result in random displacement increments $\delta\mathbf{u}^i$ at each step of the iterative process. Therefore, the rational approach is to expand $\delta\mathbf{u}^i$ in a PC basis and evaluate the PC coefficients through the convergence of an

incremental-iterative process. More specifically, eq. (8.1) reads:

$$\mathbf{K}_t^{i-1} \sum_{k=0}^{P-1} \delta \boldsymbol{\alpha}_k^i \Psi_k(\boldsymbol{\xi}) = \mathbf{R}^i - \mathbf{f}^{i-1} \tag{8.3}$$

where,

$$\delta \mathbf{u}^i = \sum_{k=0}^{P-1} \delta \boldsymbol{\alpha}_k^i \Psi_k(\boldsymbol{\xi}) \tag{8.4}$$

PC expansion of displacement $\mathbf{u}$ leads to

$$\mathbf{u}^i = \sum_{k=0}^{P-1} \boldsymbol{\alpha}_k^i \Psi_k(\boldsymbol{\xi}) \tag{8.5}$$

where

$$\boldsymbol{\alpha}_k^i = \boldsymbol{\alpha}_k^{i-1} + \delta \boldsymbol{\alpha}_k^i \qquad k = 0, 1, ..., P-1 \tag{8.6}$$

Projecting the residual

$$\epsilon_P = \Big(\sum_{k=0}^{P-1} \mathbf{K}_t^{i-1} \delta \boldsymbol{\alpha}_k^i \Psi_k \Big) - \mathbf{R}^i + \mathbf{f}^{i-1} \tag{8.7}$$

to the PC basis and minimizing it in the mean square sense we obtain the following augmented incremental Newton-Raphson scheme

$$\boldsymbol{\mathcal{K}}_t^{i-1} \delta \boldsymbol{\alpha}^i = \boldsymbol{\mathcal{R}}^i - \boldsymbol{\mathcal{f}}^{i-1} \tag{8.8}$$

where $\boldsymbol{\mathcal{K}}_t^{i-1}$, $\boldsymbol{\mathcal{R}}^i$ and $\boldsymbol{\mathcal{f}}^{i-1}$ are the augmented tangent stiffness matrix, external load vector and internal force vector at the i-iteration, respectively, computed as the inner products of $\mathbf{K}_t^{i-1}$, $\mathbf{R}^i$, $\mathbf{f}^{i-1}$ with Ψ_k. Also, $\delta \boldsymbol{\alpha}_k^i$ is the vector of the expansion coefficients to be evaluated. A major issue that arises here is that both $\mathbf{K}_t$ and $\mathbf{f}$ are functions of the displacement vector $\mathbf{u}$ and the material properties (i.e Young's modulus E). Hence, Eq. (5.33) and (5.36) are no longer valid and have to be redefined.

The stochastic tangent stiffness matrix for each element e with n-dofs may be computed

as follows

$$\boldsymbol{\mathcal{K}}_t^e = \begin{bmatrix} \boxed{\mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_0]} & \mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_1] & \cdots & \mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_{P-1}] \\ \mathbb{E}[\mathbf{K}_t^e \Psi_1 \Psi_0] & \mathbb{E}[\mathbf{K}_t^e \Psi_1 \Psi_1] & \cdots & \mathbb{E}[\mathbf{K}_t^e \Psi_1 \Psi_{P-1}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[\mathbf{K}_t^e \Psi_{P-1} \Psi_0] & \mathbb{E}[\mathbf{K}_t^e \Psi_{P-1} \Psi_1] & \cdots & \mathbb{E}[\mathbf{K}_t^e \Psi_{P-1} \Psi_{P-1}] \end{bmatrix} \tag{8.9}$$

Also,

$$\boldsymbol{\mathcal{f}}^e = \mathbb{E}[\Psi_k \mathbf{f}^e] \tag{8.10}$$

$$\boldsymbol{\mathcal{R}}^e = \mathbb{E}[\Psi_k \mathbf{R}^e] \quad k = 1, .., P-1 \tag{8.11}$$

where, $\boldsymbol{\mathcal{K}}_t^e$, $\boldsymbol{\mathcal{f}}^e$ and $\boldsymbol{\mathcal{R}}^e$ are $nP \times nP$, $nP \times 1$ and $nP \times 1$, respectively. Let us also recall that $\mathbb{E}[\cdot]$ is the inner product of the $\mathscr{L}^2(\Theta, \mathscr{F}, \mathscr{P})$ space and is given from the following relation

$$\begin{aligned} \mathbb{E}[\cdot] &= \int_{-\infty}^{\infty} [\cdot] \frac{1}{(\sqrt{2\pi})^M} e^{-\boldsymbol{\xi}^T \boldsymbol{\xi}/2} d\boldsymbol{\xi} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} [\cdot] \frac{1}{(\sqrt{2\pi})^M} e^{-\xi_1^2/2} \cdots e^{-\xi_M^2/2} d\xi_1 \cdots d\xi_M \end{aligned} \tag{8.12}$$

For instance, the boxed entry $\mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_0]$ in eq. (8.9) reads

$$\begin{aligned} &\mathbb{E}[\mathbf{K}_t^e \Psi_0 \Psi_0] = \\ &= \begin{bmatrix} \int_{-\infty}^{\infty} \mathbf{K}_t^e(1,1) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} & \cdots & \int_{-\infty}^{\infty} \mathbf{K}_t^e(1,n) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} \\ \vdots & \ddots & \vdots \\ \int_{-\infty}^{\infty} \mathbf{K}_t^e(n,1) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} & \cdots & \int_{-\infty}^{\infty} \mathbf{K}_t^e(n,n) \Psi_0 \Psi_0 \frac{1}{(\sqrt{2\pi})^M} e^{-\frac{\boldsymbol{\xi}^T \boldsymbol{\xi}}{2}} d\boldsymbol{\xi} \end{bmatrix} \end{aligned} \tag{8.13}$$

Once these quantities are calculated for each element, then they are assembled to formulate the global matrices $\boldsymbol{\mathcal{K}}_t$, $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{f}}$.

The equations derived in this section are built upon the ideas proposed in [56],

where it is stated that each entry of $\mathcal{K}_t$, $\mathcal{R}$ and $\boldsymbol{f}$ can be represented using the PC decomposition. In the next section, the explicit relations involving these entries and the random parameter vector $\boldsymbol{\xi}$ will be derived in the frame of the stochastic co-rotational formulation for beam elements.

8.3 Spectral Stochastic co-rotational formulation for beams

In the previous equations it becomes evident, that an expression for the tangent stiffness and the vector of internal forces, which will incorporate their random nature, is required for the implementation of the augmented NR scheme of eq. (8.8). These expressions will stem from the kinematic equations that describe the system and differ for each finite element type. To illustrate the Langrangian formulation of the stochastic tangent stiffness matrix and vector of internal forces for a beam element, let us first consider the generic beam element of figure 8.1 from node 1 to node 2, under some nodal displacements $\mathbf{u} = \begin{pmatrix} u_1 & v_1 & z_1 & u_2 & v_2 & z_2 \end{pmatrix}^T$.

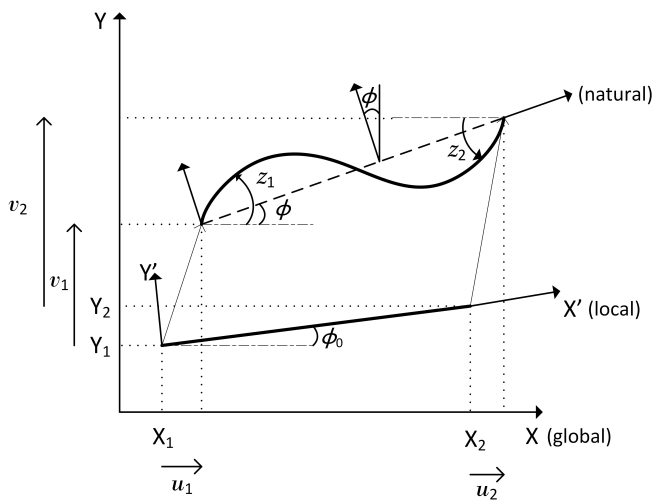


Figure 8.1: Generic Beam Element

In this figure, $\mathbf{X_0} = \begin{pmatrix} X_1 & Y_1 & X_2 & Y_2 \end{pmatrix}^T$ is the position vector in the undeformed configuration. By virtue of eq. (8.5), where the superscript i is omitted for the sake of brevity, $\mathbf{u}$ is written as

$$\mathbf{u}(\boldsymbol{\xi}) = \Big(\sum_{k=0}^{P-1} \alpha_{k1}\Psi_k(\boldsymbol{\xi}), \sum_{k=0}^{P-1} \alpha_{k2}\Psi_k(\boldsymbol{\xi}), \sum_{k=0}^{P-1} \alpha_{k3}\Psi_k(\boldsymbol{\xi}), \sum_{k=0}^{P-1} \alpha_{k4}\Psi_k(\boldsymbol{\xi}), \sum_{k=0}^{P-1} \alpha_{k5}\Psi_k(\boldsymbol{\xi}), \sum_{k=0}^{P-1} \alpha_{k6}\Psi_k(\boldsymbol{\xi}) \Big)^T \tag{8.14}$$

From eq. (8.14) it becomes evident, that the position vector in the deformed configuration, $\mathbf{X_n} = \mathbf{X_0} + \mathbf{u}$, will also be stochastic.

Let us denote L_0 and A the element's length and cross-sectional area in the undeformed configuration, respectively. Let us also consider, for the purposes of demonstration, the Young's modulus E as a Gaussian random field, which can be expressed in a KL expansion series as follows

$$E(x,\theta) = E_0 + \sum_{j=1}^{M} \lambda_j \varphi_j(x) \xi_j(\theta) \tag{8.15}$$

It should be mentioned that for non-Gaussian random fields all subsequent formulations will remain the same, but instead of the KL expansion, their PC expansion will be used, as described in section 5.4.2.

It follows that the element's length in the deformed configuration will be random, that is

$$\begin{aligned} L(\boldsymbol{\xi}) =& \left((X_1 + u_1 - X_2 - u_2)^2 + (Y_1 + v_1 - Y_2 - v_2)^2 \right)^{1/2} \\ =& \Bigg(\Big(X_1 + \sum_{k=0}^{P-1} \alpha_{k1}\Psi_k(\boldsymbol{\xi}) - X_2 - \sum_{k=0}^{P-1} \alpha_{k4}\Psi_k(\boldsymbol{\xi}) \Big)^2 \\ &+ \Big(Y_1 + \sum_{k=0}^{P-1} \alpha_{k2}\Psi_k(\boldsymbol{\xi}) - Y_2 - \sum_{k=0}^{P-1} \alpha_{k5}\Psi_k(\boldsymbol{\xi}) \Big)^2 \Bigg)^{1/2} \end{aligned} \tag{8.16}$$

Defining now the random rotation $\phi(\boldsymbol{\xi})$ as

$$\phi(\boldsymbol{\xi}) = 2arctan\left(\frac{1 - c(\boldsymbol{\xi})}{s(\boldsymbol{\xi})} \right) \tag{8.17}$$

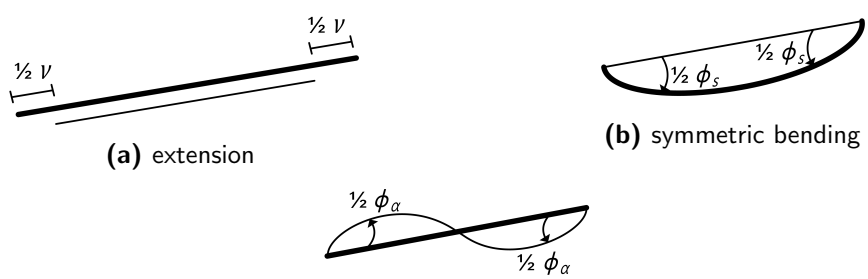
where $c(\boldsymbol{\xi})$ and $s(\boldsymbol{\xi})$ are the random cosine and sine, respectively, given as

$$\begin{aligned} c(\boldsymbol{\xi}) =& \frac{X_2 + u_2 - X_1 - u_1}{L(\boldsymbol{\xi})} \\ =& \frac{X_2 + \sum_{k=0}^{P-1} \alpha_{k4} \Psi_k(\boldsymbol{\xi}) - X_1 - \sum_{k=0}^{P-1} \alpha_{k1} \Psi_k(\boldsymbol{\xi})}{L(\boldsymbol{\xi})} \end{aligned} \tag{8.18}$$

$$\begin{aligned} s(\boldsymbol{\xi}) =& \frac{Y_2 + v_2 - Y_1 - v_1}{L(\boldsymbol{\xi})} \\ =& \frac{Y_2 + \sum_{k=0}^{P-1} \alpha_{k5} \Psi_k(\boldsymbol{\xi}) - Y_1 - \sum_{k=0}^{P-1} \alpha_{k2} \Psi_k(\boldsymbol{\xi})}{L(\boldsymbol{\xi})} \end{aligned} \tag{8.19}$$

In the co-rotational formulation, the vector of the stochastic nodal displacements $\mathbf{u}(\boldsymbol{\xi})$ is transformed to $\boldsymbol{\upsilon}(\boldsymbol{\xi})$, which is a reduced vector of nodal displacements, corresponding to the basic independent modes of the system's deformation modes expressed in the natural (co-rotational) system. These modes include the extension mode $\nu(\boldsymbol{\xi})$, the symmetric bending mode $\phi_s(\boldsymbol{\xi})$ and the anti-symmetric bending mode $\phi_a(\boldsymbol{\xi})$, which are shown in figure 8.2. Using these three modes any deformed shape of an element in the global system can be expressed as a linear combination of these modes.

$$\boldsymbol{\upsilon}(\boldsymbol{\xi}) = \begin{pmatrix} \nu(\boldsymbol{\xi}) \\ \phi_s(\boldsymbol{\xi}) \\ \phi_a(\boldsymbol{\xi}) \end{pmatrix} \tag{8.20}$$



(a) extension

(b) symmetric bending

(c) anti-symmetric bending

Figure 8.2: Natural modes

This transformation reads as

$$\nu(\boldsymbol{\xi}) = L(\boldsymbol{\xi}) - L_0 \tag{8.21}$$

$$\begin{aligned}\phi_s(\boldsymbol{\xi}) =& z_2 - z_1 \\ =& \sum_{k=0}^{P-1} \alpha_{k6}\Psi_k(\boldsymbol{\xi}) - \sum_{k=0}^{P-1} \alpha_{k3}\Psi_k(\boldsymbol{\xi})\end{aligned} \tag{8.22}$$

$$\begin{aligned}\phi_a(\boldsymbol{\xi}) =& z_2 + z_1 - 2(\phi(\boldsymbol{\xi}) - \phi_0) \\ =& \sum_{k=0}^{P-1} \alpha_{k6}\Psi_k(\boldsymbol{\xi}) + \sum_{k=0}^{P-1} \alpha_{k3}\Psi_k(\boldsymbol{\xi}) - 2(\phi(\boldsymbol{\xi}) - \phi_0)\end{aligned} \tag{8.23}$$

where ϕ_0 is the initial angle of the element. The rotation $\phi_a(\boldsymbol{\xi})$ is modified in order to be robust for arbitrary angles, as follows

$$\phi_a(\boldsymbol{\xi}) := mod(\phi_a(\boldsymbol{\xi}) + \pi, 2\pi) - \pi \tag{8.24}$$

Also, we define $\Phi(\boldsymbol{\xi}) = \dfrac{12E(\boldsymbol{\xi})I}{L(\boldsymbol{\xi})^2 G(\boldsymbol{\xi})A}$, with $G(\boldsymbol{\xi})$ being the shear modulus, and $\psi_a(\boldsymbol{\xi}) = \dfrac{1}{1+\Phi(\boldsymbol{\xi})}$. Then, the stress resultants corresponding to the deformation modes in the natural system (fig. 8.3) of the beam element are given as

$$N(\boldsymbol{\xi}) = \frac{E(\boldsymbol{\xi})A}{L_0}\nu(\boldsymbol{\xi}) \tag{8.25}$$

$$M_s(\boldsymbol{\xi}) = \frac{E(\boldsymbol{\xi})I}{L_0}\phi_s(\boldsymbol{\xi}) \tag{8.26}$$

$$M_a(\boldsymbol{\xi}) = 3\psi_a \frac{E(\boldsymbol{\xi})I}{L_0}\phi_a(\boldsymbol{\xi}) \tag{8.27}$$

$$Q(\boldsymbol{\xi}) = -\frac{2M_a(\boldsymbol{\xi})}{L_0} \tag{8.28}$$

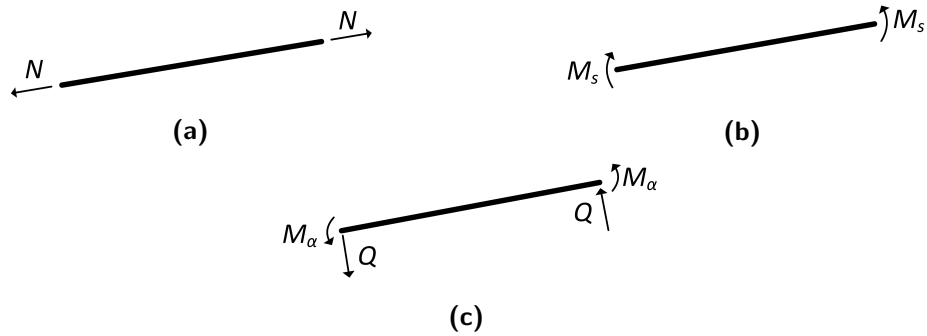


Figure 8.3: Stress resultants in the natural system

Following the deterministic co-rotational formulation [84] we define the following beam rotation matrix, which describes the transformation between the global and local frame of reference as

$$\boldsymbol{R}_e(\boldsymbol{\xi}) = \begin{bmatrix} [\hat{\boldsymbol{R}}(\boldsymbol{\xi})] & [0] \\ [0] & [\hat{\boldsymbol{R}}(\boldsymbol{\xi})] \end{bmatrix} \tag{8.29}$$

where

$$\hat{\boldsymbol{R}}(\boldsymbol{\xi}) = \begin{bmatrix} c(\boldsymbol{\xi}) & -s(\boldsymbol{\xi}) & 0 \\ s(\boldsymbol{\xi}) & c(\boldsymbol{\xi}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8.30}$$

The transformation from the natural system to the local system is given by the 6×3 matrix $\boldsymbol{S}$, where

$$\boldsymbol{S}(\boldsymbol{\xi}) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 2/L(\boldsymbol{\xi}) \\ 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & -2/L(\boldsymbol{\xi}) \\ 0 & 1 & 1 \end{bmatrix} \tag{8.31}$$

Furthermore, the stochastic tangent stiffness matrix $\boldsymbol{K}_d(\boldsymbol{\xi})$ of the element's natural

deformation modes reads

$$\boldsymbol{K}_d(\boldsymbol{\xi}) = \begin{bmatrix} \frac{E(\boldsymbol{\xi})A}{L(\boldsymbol{\xi})} & 0 & 0 \\ 0 & \frac{E(\boldsymbol{\xi})I}{L(\boldsymbol{\xi})} + \frac{L(\boldsymbol{\xi})N(\boldsymbol{\xi})}{12} & 0 \\ 0 & 0 & \frac{3\psi_a(\boldsymbol{\xi})E(\boldsymbol{\xi})I}{L(\boldsymbol{\xi})} + \frac{L(\boldsymbol{\xi})N(\boldsymbol{\xi})}{20} \end{bmatrix} \tag{8.32}$$

Also, in order to account for the effect of changes in $L(\boldsymbol{\xi})$ to the shear force, the stiffness matrix $\boldsymbol{K}_r(\boldsymbol{\xi})$ is introduced, where

$$\bar{\boldsymbol{K}}_r(\boldsymbol{\xi}) = \begin{bmatrix} 0 & -Q(\boldsymbol{\xi})/L(\boldsymbol{\xi}) & 0 \\ -Q(\boldsymbol{\xi})/L(\boldsymbol{\xi}) & N(\boldsymbol{\xi})/L(\boldsymbol{\xi}) & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{8.33}$$

and

$$\boldsymbol{K}_r(\boldsymbol{\xi}) = \begin{bmatrix} [\bar{\boldsymbol{K}}_r(\boldsymbol{\xi})] & [-\bar{\boldsymbol{K}}_r(\boldsymbol{\xi})] \\ [-\bar{\boldsymbol{K}}_r(\boldsymbol{\xi})] & [\bar{\boldsymbol{K}}_r(\boldsymbol{\xi})] \end{bmatrix} \tag{8.34}$$

Then, the element's stochastic tangent stiffness matrix in the local system is given by the following equation

$$\boldsymbol{K}_e(\boldsymbol{\xi}) = \boldsymbol{S}(\boldsymbol{\xi})\boldsymbol{K}_d(\boldsymbol{\xi})\boldsymbol{S}(\boldsymbol{\xi})^T + \boldsymbol{K}_r(\boldsymbol{\xi}) \tag{8.35}$$

Finally, the tangent stiffness matrix and the vector of internal forces in the global system are given from the following transformations

$$\mathbf{K}_t(\boldsymbol{\xi}) = \boldsymbol{R}_e(\boldsymbol{\xi})\boldsymbol{K}_e(\boldsymbol{\xi})\boldsymbol{R}_e(\boldsymbol{\xi})^T \tag{8.36}$$

$$\mathbf{f}(\boldsymbol{\xi}) = \boldsymbol{R}_e(\boldsymbol{\xi})\boldsymbol{S}(\boldsymbol{\xi})\boldsymbol{K}_d(\boldsymbol{\xi})\boldsymbol{v}(\boldsymbol{\xi}) \tag{8.37}$$

8.4 Stochastic path following techniques

8.4.1 Stochastic Load Control method (SLC)

In the formulation of the SLC method we assume that the applied load increment at each step is deterministic. Due to the random material properties, however, the response $\mathbf{u}$ is stochastic and is expressed through the PC coefficients $\boldsymbol{\alpha}$ (eq. (8.5)). The procedure is similar to the standard deterministic load control algorithm, with the difference that instead of the displacement components, we solve eq. (8.3) to find the PC expansions

coefficients $\delta\boldsymbol{\alpha}^i$. After each iteration their values are updated as described in eq. (8.6), namely,

$$\boldsymbol{\alpha}_k^i = \boldsymbol{\alpha}_k^{i-1} + \delta\boldsymbol{\alpha}_k^i \qquad k = 0, 1, ..., P-1$$

With $\boldsymbol{\alpha}_k^i$ known, the displacement vector can be evaluated according to eq. (8.5) as $\mathbf{u}^i = \sum_{k=0}^{P-1} \boldsymbol{\alpha}_k^i \Psi_k$.
The process is repeated until convergence is achieved ($\|\delta\boldsymbol{\alpha}^i\| \leq$ tolerance ε) and then, proceed with the next load increment. With this approach, the information we retrieve is the evolution of the pdf of the displacement vector $\mathbf{u}$ as a function of the deterministic load R.

8.4.2 Stochastic Displacement Control method (SDC)

For the SDC method the incremental displacements are also expanded in a PC series according to eq. (8.4), but in this case loading is considered as an imposed incremental displacement at the initial iteration of each increment, $\delta u^0(q)$, along the q-th dof, which has to be deterministic, that is

$$\delta u^0(q) = \sum_{k=0}^{P-1} \delta\alpha_k^0(q)\Psi_k(\boldsymbol{\xi}) \tag{8.38}$$

with

$$\delta\alpha_k^0(q) = \begin{cases} \delta u^0(q), & \text{for } k = 0 \\ 0, & \text{for } k > 0 \end{cases} \tag{8.39}$$

Furthermore, the incremental displacement along the q-th dof is considered to be zero for all subsequent iterations within the same increment, which means

$$\delta u^i(q) = \sum_{k=0}^{P-1} \delta\alpha_k^i(q)\Psi_k(\boldsymbol{\xi}) = 0 \quad \text{for } i \neq 0 \tag{8.40}$$

and

$$\delta\alpha_k^i(q) = 0 \qquad \text{for } k = 0, 1, ..P-1 \tag{8.41}$$

In contrast to the SLC, SDC considers the external load in the monitored dof $R^i(q)$ to be random and expands it in the same PC basis with coefficients $\beta_k^i(q)$, such that

$$R^i(q) = \sum_{k=0}^{P-1} \beta_k^i(q)\Psi_k(\boldsymbol{\xi}) \tag{8.42}$$

This assumption is justified because a deterministic displacement at one dof can be the result of various load values, each with a different probability, due to the randomness in the system properties. With this approach the number of unknowns in the system remains the same, hence, the system has a unique solution.

In order to calculate the PC coefficients, the residual ϵ_p is projected onto the PC basis and required to be zero. Thus, we obtain again eq. (8.8) but now unknowns can be found both in $\delta\boldsymbol{\alpha}^i$ and $\boldsymbol{\mathcal{R}}^i$. Therefore, we rearrange these vectors into two sub-vectors, one for the unknown coefficients (free) and one for the known (supported).

$$\delta\boldsymbol{\alpha}^i = \begin{bmatrix} \delta\boldsymbol{\alpha}_f^i \\ \delta\boldsymbol{\alpha}_s^i \end{bmatrix} \ , \ \boldsymbol{\mathcal{R}}^i = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\beta}^i \end{bmatrix} \ \text{and} \ \boldsymbol{\mathcal{f}}^{i-1} = \begin{bmatrix} \boldsymbol{\mathcal{f}}_f^{i-1} \\ \boldsymbol{\mathcal{f}}_s^{i-1} \end{bmatrix} \tag{8.43}$$

Subdividing $\boldsymbol{\mathcal{K}}_t^{i-1}$ into the following submatrices,

$$\boldsymbol{\mathcal{K}}_t^{i-1} = \begin{bmatrix} \boldsymbol{\mathcal{K}}_{t_{ff}}^{i-1} & \boldsymbol{\mathcal{K}}_{t_{fs}}^{i-1} \\ \boldsymbol{\mathcal{K}}_{t_{sf}}^{i-1} & \boldsymbol{\mathcal{K}}_{t_{ss}}^{i-1} \end{bmatrix} \tag{8.44}$$

the values of the coefficients are obtained as follows

$$\delta\boldsymbol{\alpha}_f^i = -(\boldsymbol{\mathcal{K}}_{t_{ff}}^{i-1})^{-1}(\boldsymbol{\mathcal{K}}_{t_{fs}}^{i-1}\delta\boldsymbol{\alpha}_s^i + \boldsymbol{\mathcal{f}}_f^{i-1}) \tag{8.45}$$

and

$$\boldsymbol{\beta}^i = \boldsymbol{\mathcal{K}}_{t_{sf}}^{i-1}\delta\boldsymbol{\alpha}_f^i + \boldsymbol{\mathcal{K}}_{t_{ss}}^{i-1}\delta\boldsymbol{\alpha}_s^i + \boldsymbol{\mathcal{f}}_s^{i-1} \tag{8.46}$$

After each iteration, the displacement coefficients are updated as

$$\boldsymbol{\alpha}_k^i = \boldsymbol{\alpha}_k^{i-1} + \delta\boldsymbol{\alpha}_k^i \qquad k = 0, 1, ..., P-1 \tag{8.47}$$

until convergence is reached. In contrast to the stochastic load control algorithm, the information obtained with this approach is the probabilistic characteristics of R for each value of the incremental deterministic displacement at the monitored dof.

8.5 Computational aspects of the proposed methodology

8.5.1 Formulation of the augmented tangent stiffness matrix

The estimation of the integrals of the matrices in equations (8.9), (8.10) and (8.11) is far from trivial, because the entries in these matrices are nonlinear functions with respect to $\boldsymbol{\xi}$. For this reason no analytic relation can be derived for their estimation and, thus the only feasible approach is to resort to numerical quadrature. Considering the existence of the exponential terms $e^{-\boldsymbol{\xi}^T\boldsymbol{\xi}/2}$, Gauss-Hermite quadrature [1] is a well-suited choice for this purpose, but other numerical integration schemes can be used as well. Nevertheless, these integrations must be repeated at each NR iteration, and might lead to an extensive computational cost, especially for large systems with many stochastic dimensions.

However, exploiting the scalability of the proposed methodology would reduce significantly the computational cost. For instance, the evaluation of the augmented tangent stiffness matrix for the element e is disjoint from the other elements and therefore, a different processing unit can be assigned for each evaluation of $\boldsymbol{\mathcal{K}}_t^e$. Also, the symmetrical properties of the element's stiffness matrix, as well as those of the augmented stiffness matrix can be further exploited. All analyses in this work were performed in the **MSolve** finite element analysis software [MSolve], where the above-mentioned modifications were incorporated.

An alternative approach, which does not require the evaluation of complex integrals, is described in [46]. The main idea is to project polynomial and non-polynomial functions represented by PCE onto a new variable of the same order as the initial ones. To give an example of multiplication of random variables, let us consider two random variables $\mathtt{a}$ and $\mathtt{b}$, approximated by a PC expansion, as follows

$$\hat{\mathtt{a}} = \sum_{j=0}^{P-1} \mathtt{a}_j \Psi_j \tag{8.48}$$

$$\hat{\mathtt{b}} = \sum_{k=0}^{P-1} \mathtt{b}_k \Psi_k \tag{8.49}$$

Under the assumption that their product, $\mathtt{c} = \mathtt{ab}$ can be represented as

$$
\begin{aligned}
\hat{\mathtt{c}} &= \sum_{l=0}^{P-1} \mathtt{c}_l \Psi_l \\
&= \sum_{j=0}^{P-1} \mathtt{a}_j \Psi_j \sum_{k=0}^{P-1} \mathtt{b}_k \Psi_k
\end{aligned} \tag{8.50}
$$

then, the coefficients $\mathtt{c}_l$ can be obtained by a Galerkin projection as

$$
\mathtt{c}_l = \sum_{j=0}^{P-1} \sum_{k=0}^{P-1} \mathtt{C}_{jkl} \mathtt{a}_j \mathtt{b}_k \quad \forall l \in \{0, 1, ..., P-1\} \tag{8.51}
$$

with

$$
\mathtt{C}_{jkl} = \frac{\mathbb{E}[\Psi_j \Psi_k \Psi_l]}{\mathbb{E}[\Psi_l^2]} \tag{8.52}
$$

Operations such as division of rv's, raise to a power or inverse of an rv can be performed in a similar fashion. For other nonlinear functions such as the *arctan* in (8.17) a direct integration procedure can be applied. For instance, if $\mathtt{d}$ is a function $\mathscr{F}(\mathtt{a})$ of the rv $\mathtt{a}$, then,

$$
\hat{\mathtt{d}} = \sum_{j=0}^{P-1} \mathtt{d}_j \Psi_j
$$

where,

$$
\mathtt{d}_j = \frac{\mathbb{E}[\mathscr{F}(\mathtt{a}) \Psi_j]}{\mathbb{E}[\Psi_j^2]} \tag{8.53}
$$

This idea was successfully applied in [2], where it was extended to operations on random matrices.

With this procedure we arrive at the following relations for the stochastic tangent stiffness matrix $\mathbf{K}_t$, the internal force vector $\mathbf{f}$ and the external load vector $\mathbf{R}$

$$\mathbf{K}_t = \sum_{j=0}^{P-1} \mathbf{K}_{t_j} \Psi_j \tag{8.54}$$

$$\mathbf{f} = \sum_{l=0}^{P-1} \mathbf{f}_l \Psi_l \tag{8.55}$$

$$\mathbf{R} = \sum_{l=0}^{P-1} \mathbf{R}_l \Psi_l \tag{8.56}$$

and a Galerkin projection of the equation of equilibrium will ultimately yield the following system of linear equations for the unknowns $\delta\boldsymbol{\alpha}_k$.

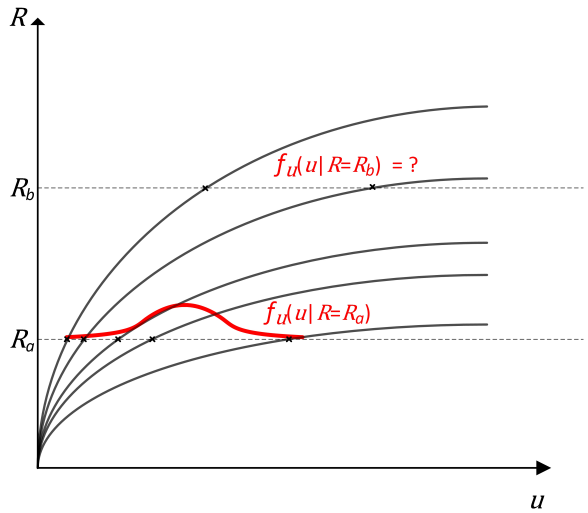
$$\mathbf{K}_{t_j} \delta\boldsymbol{\alpha}_k \cdot \mathbb{E}[\Psi_j \Psi_k \Psi_l] = \mathbf{R}_l \cdot \mathbb{E}[\Psi_l^2] - \mathbf{f}_l \cdot \mathbb{E}[\Psi_l^2] \tag{8.57}$$

8.5.2 Areas of applicability

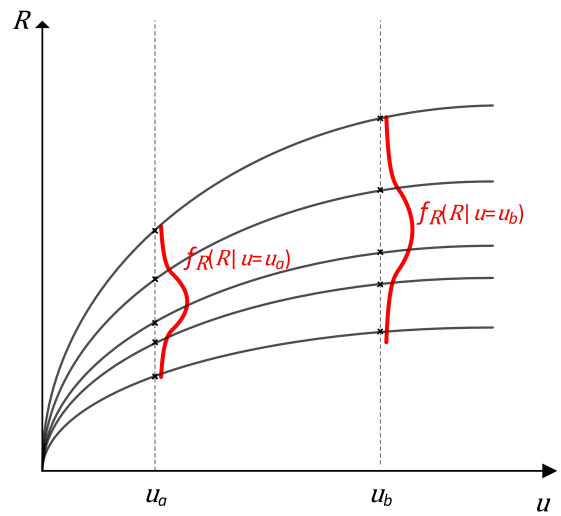
Another issue that should be addressed here is the areas of applicability of each of the proposed stochastic path following methodologies implemented. In deterministic cases, it is well known that the load control cannot capture the equilibrium path near and after a limit point, while the displacement control cannot trace paths that exhibit snap-back's. This notion is extended to their stochastic counterparts, as well. From section 8.4, it becomes evident that the choice of the stochastic iterative algorithm will produce a different interpretation of the results, since the load control evaluates the pdf of the displacement for each value of the deterministic load increment, while the displacement control evaluates the pdf of the load for each value of the deterministic incremental displacement of the monitored dof. This difference in the interpretation may pose some limitations in the applicability of SLC and SDC.

To illustrate this, let us consider two stochastic systems, one that exhibits geometrical softening and one that exhibits stiffening, as depicted in figures 8.4 and 8.5, respectively. Also, let's assume that the response of these systems is described by five equiprobable load-displacement curves. In figure 8.4a it is straightforward to answer which is the distribution of u, when the external load R becomes equal to R_a, as it intersects with all load-displacement curves. However, when $R = R_b$ the conditional pdf $f_u(u|R = R_b)$ cannot be estimated, due to the fact that only two of the five paths attain load values $\geq R_b$,

which has no physical interpretation. Instead, as shown in figure 8.4b the distribution of R given u can be obtained in a straightforward manner, which is the purpose of the SDC. Thus, in cases exhibiting geometric softening, a stochastic displacement control algorithm must be used. On the other hand, from figures 8.5a and 8.5b the opposite conclusion is drawn for problems exhibiting stiffening only, in which the distribution of the load R for each value of u can be straightforwardly estimated, only with the SLC algorithm.



(a)



(b)

Figure 8.4: Stochastic load control for problems with geometric softening

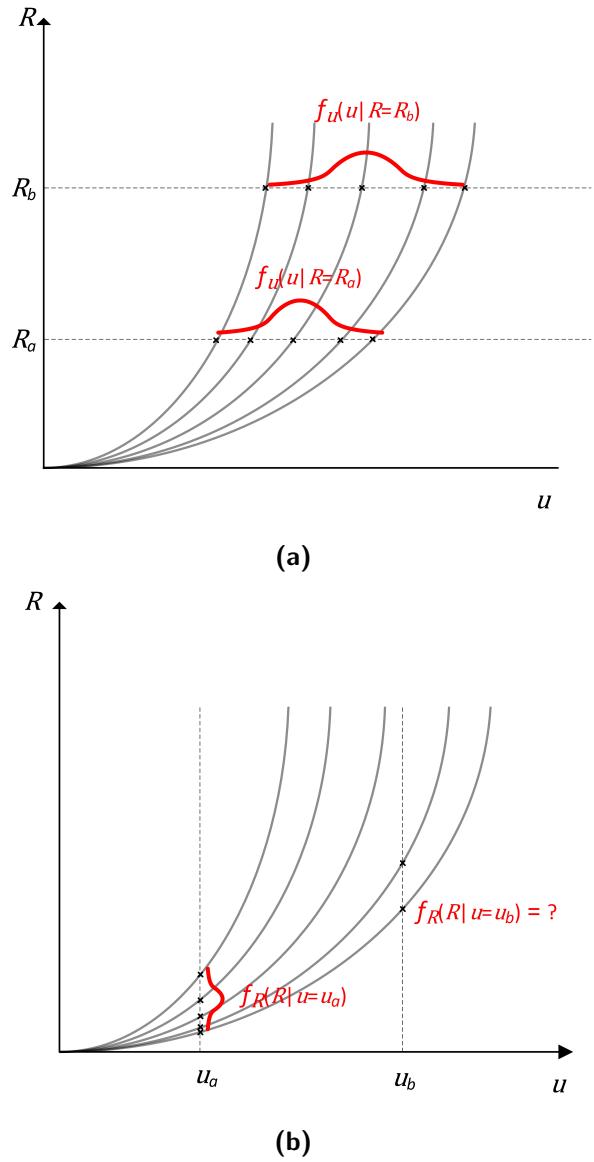


Figure 8.5: Stochastic displacement control for problems with geometric hardening

For the sake of completeness, the more general case, where the structure could exhibit both stiffening and softening, as well as snap-back's, should be addressed. In such cases, neither the stochastic load control, nor the displacement control would be able to capture the stochastic equilibrium path. For instance, problems involving random initial deflections that may lead to buckling cannot be tackled by the proposed methodology,

since the type of the post-buckling path cannot be predicted a priori. It is expected that a stochastic formulation of the arc-length control method might be able to handle this kind of problems, but such an implementation remains the scope of future work.

8.6 Numerical Examples

1. Two-bar shallow truss

For our first test case we consider the two-bar shallow truss depicted in figure 8.6. This structure exhibits highly nonlinear behavior and serves as benchmark example in many textbooks [44]. For this example, the following assumptions were made: $L_0 = 1.0m$ and $A = 25 \cdot 10^{-4} m^2$. Also, the Young's modulus E is considered to be a random variable following the normal distribution with mean $\bar{E} = 2 \cdot 10^8 \ kN/m^2$ and coefficient of variation $cov_E = 20\%$. Since E is random, the load-displacement curve $R - u$ will also be random and this becomes evident in figure 8.7, where the curve is plotted for various values of E. Also, the existence of snap-through's in the curves suggests that only the SDC method can trace the equilibrium path and the only information we can extract in such cases is the distribution of the load R for each value of u.

By applying the proposed methodology we obtain figures 8.8 and 8.9 which depict the mean value and standard deviation of R as a function of u, respectively, computed with the nonlinear SSFEM (NLSSFEM) for $p = 2$ and 3, as well as with brute force MCS. As can be seen from these figures the results are in perfect agreement even for a low $p = 2$ order.

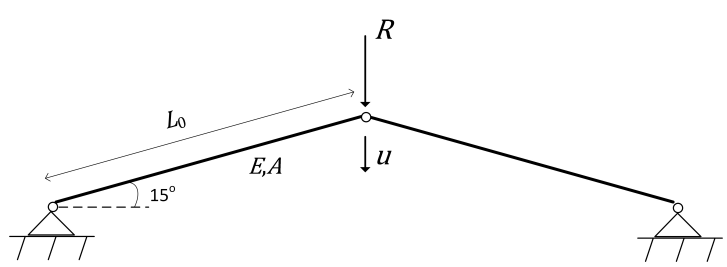


Figure 8.6: Swallow Truss

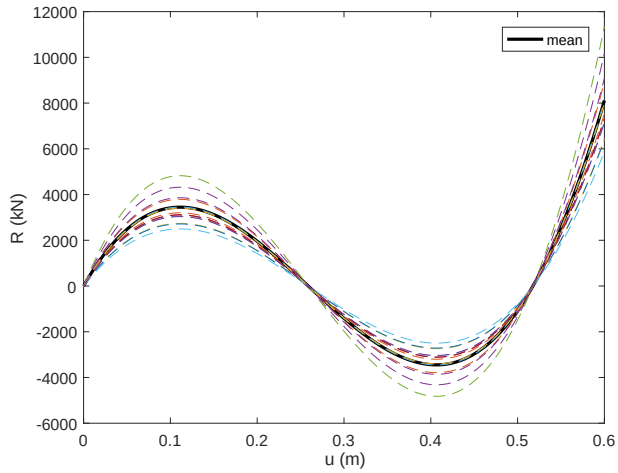


Figure 8.7: Load-Displacements curves for various E values

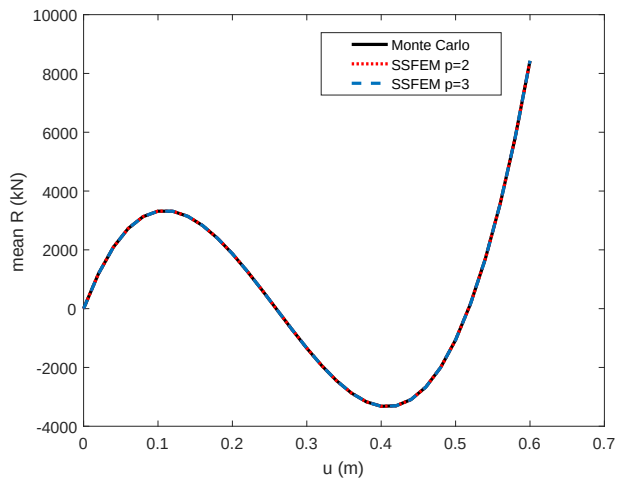


Figure 8.8: Mean value of R for each value of u

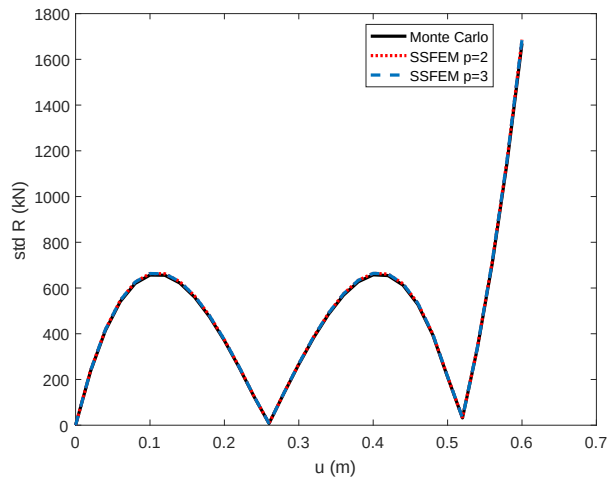


Figure 8.9: Standard deviation of R for each value of u

2. Cantilever with bending moment at the tip

For our second test case we consider the plain cantilever of figure 8.10 subjected to a bending moment at its free end. This example has the particularity that in the corresponding deterministic problem [84], as the bending moment M increases, the cantilever rolls up until both ends meet for $M_c = 2\pi EI/L_0$ (see fig.8.11). In this example, the parameters used are $L_0 = 3m$, $I = 8.333 \cdot 10^{-6}m^4$ and $A = 0.01m^2$. Also, the Young's modulus is considered as a lognormal random field along the length of the beam, $E(x)$, with mean value $2 \cdot 10^8 kN/m^2$ and standard deviation $2 \cdot 10^7 kN/m^2$. Since lognormal fields can be treated as the result of the exponentiation of an underlying Gaussian field, then $E(x)$ can be written in the form $E(x) = exp\,(\mu_E + \sigma_E f_E(x))$, where μ_E is the mean value of the underlying Gaussian field, equal to 19.1089, σ_E the corresponding standard deviation equal to 0.0998 and $f(x)$ is a zero mean unit variance stochastic field corresponding to the variability of E. For the autocorrelation function, the exponential kernel $R_f(x_1, x_2) = e^{-|x_1 - x_2|/b}$ was employed, with correlation length parameter $b = 3m$, while for the decomposition of the field 4 KL terms were used ($M_{KL} = 4$). Fig. 8.12 illustrates the first 10 eigenvalues of the kernel and fig. 8.13 the kernel approximation using 2, 3 and 4 terms in the expansion. The four-term approximation of the kernel is deemed accurate enough for the purposes of this example.

Also, the structure is discretized into 10 beam elements of equal length, which is

considered adequate for the representation of the field as well as for the convergence of the corresponding deterministic analysis. In fig. 8.14 the vertical displacement v, as well as the horizontal displacement u, are displayed for the deterministic case, when M reaches values up to $3M_c$. The behavior of this structure exhibits geometric stiffening, which suggests that only the SLC is applicable to this case.

Figures 8.15 and 8.16 depict the mean values of the horizontal displacement u and vertical displacement v, respectively, as functions of the external moment M. In both cases, the results obtained from the NLSSFEM with $p = 3$ are in perfect match to those obtained from the Monte Carlo simulation. Similarly, figures 8.17 and 8.18 display the standard deviations of u and v for each value of M. It becomes evident that for increasing values of the PC order greater convergence to the exact values, as obtained from the MC simulation, is attained. Again, for $p = 3$ fair accuracy can be reported.

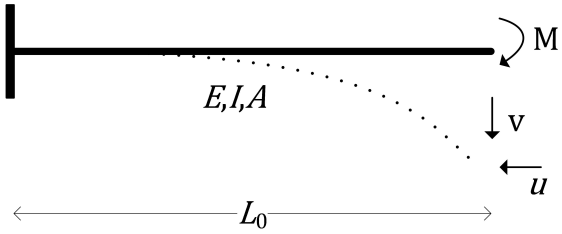


Figure 8.10: Plain cantilever with bending moment at the tip

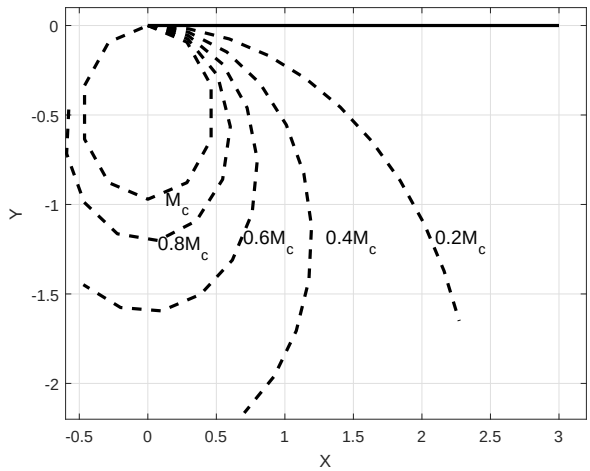


Figure 8.11: A full roll-up when $M = M_c$

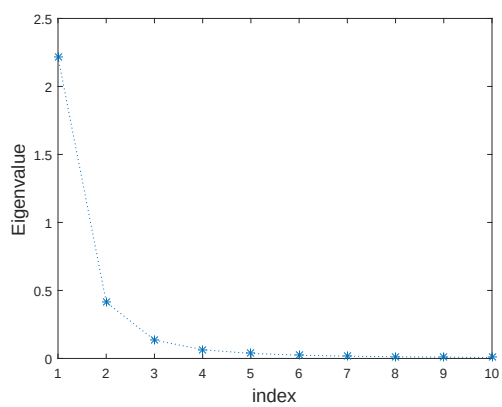


Figure 8.12: Eigenvalues of the autocorrelation kernel

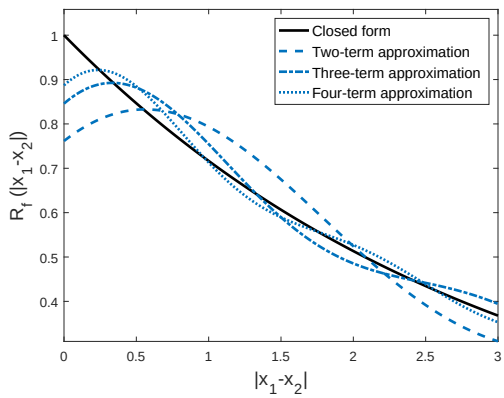


Figure 8.13: Four-term approximation of the autocorrelation kernel

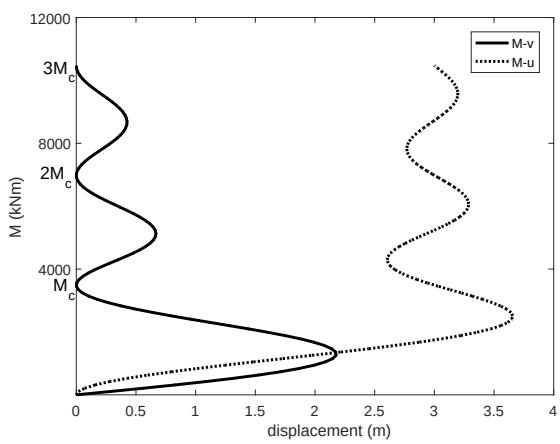


Figure 8.14: Deterministic Moment-Displacement curve for values up to $3M_c$

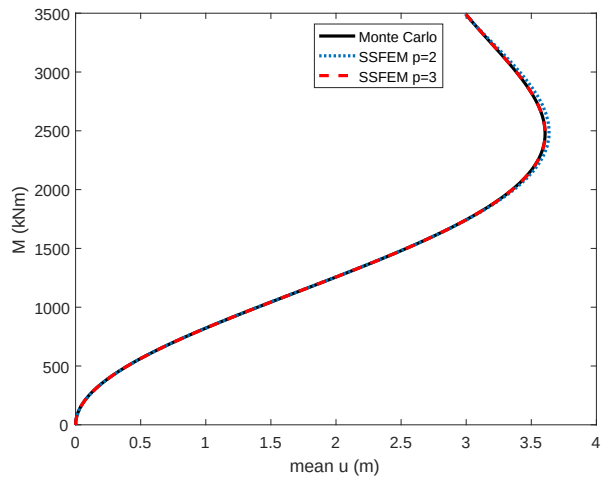


Figure 8.15: Mean value of u for each value of M

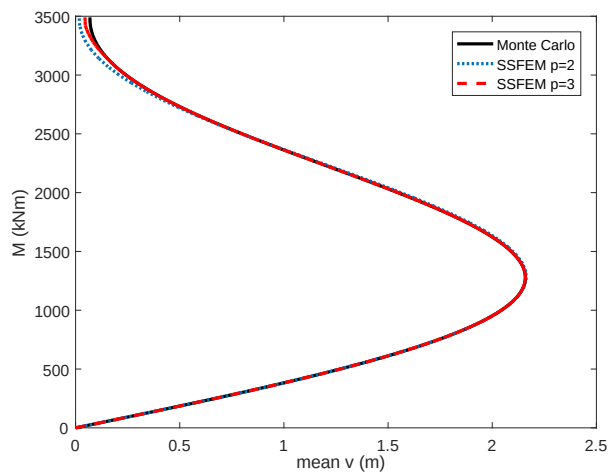


Figure 8.16: Mean value of v for each value of M

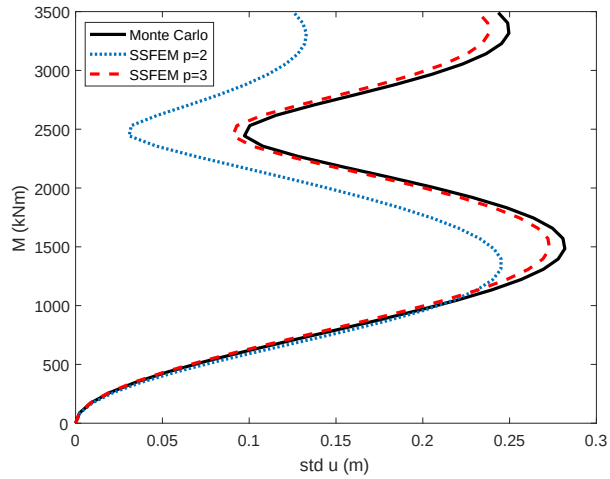


Figure 8.17: Standard deviation of u for each value of M

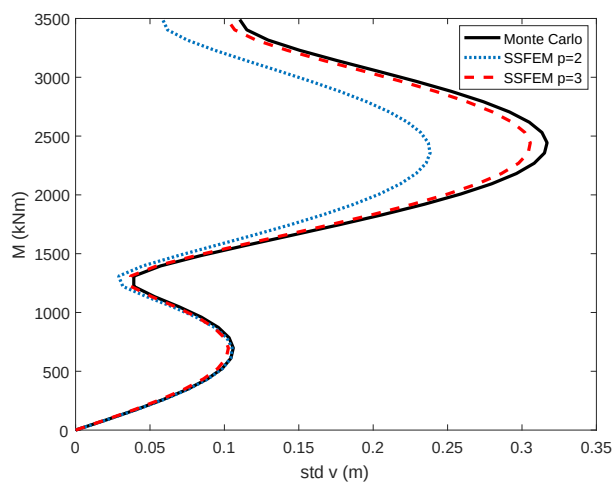


Figure 8.18: Standard deviation of v for each value of M

3. Arched beam

For the last test case we consider the arched beam depicted in figure 8.19. The structure consists of 92 beam elements and has a span of $10L_0$ and height $3L_0$, with $L_0 = 1m$. The structure is subjected to a concentrated vertical force R at its top point (node 1).

The cross-sectional area of each element is $A = 0.01m^2$ and the moment of inertia is $I = 8.333 \cdot 10^{-6}m^4$. For the construction of this structure, the following assumption was made. All steel elements comprising it, were manufactured simultaneously and then assembled in the following fashion. First, the elements of the top arched beam were assembled, then the diagonal elements and lastly the elements of the bottom arched beam, in a consecutive order. It is noted that the total length of the beam elements adds up to $36m$.

For the stochastic problem, the Young's modulus is considered as a lognormal random field, varying along the 'unfolded' total length of the structure. The mean value of the field is taken $2 \cdot 10^8 kN/m^2$ and the standard deviation $3 \cdot 10^7 kN$. The parameters of the underlying Gaussian field in this case are $\mu_E = 19.1027$, $\sigma_E = 0.1492$. Also, the exponential kernel $R_f(x_1, x_2) = e^{-|x_1 - x_2|/b}$ was used for the autocorrelation function, with correlation length parameter $b = 36m$. For the KL expansion, terms up to $M_{KL} = 4$ were found to be adequate for the purposes of the analysis, while polynomials of order 2 were considered, which led to 15 terms for the PCE. The load-displacement curve for this example exhibits geometric softening and thus we employed the SDC algorithm for this example.

In figure 8.20 the mean value of the load R vs. the top vertical displacement u is plotted and a comparison is performed between the results obtained from the proposed methodology for $p = 2$ and 10000 MC simulations. The standard deviation of R as a function of the displacement u is displayed in figure 8.21. From these figures, it is observed that the resulting standard deviation is around 13.1% of the mean. Figure 8.22 illustrates the pdf of the vertical load at $u = 0.20m$. In all cases we notice an almost perfect match between our approach and the Monte Carlo simulation. Lastly, it should be mentioned that, due to the large number of elements in this example and the use of multivariate polynomials, the re-projecting technique mentioned in section 8.5.1 was employed in order to avoid the excessive computational cost of the formulation of the stochastic stiffness matrix. With this approach, the total cost with SSFEM was $2800sec$, while for the MC simulations the cost was $5100sec$. It should be mentioned that in both cases no parallel programming techniques were used, in order for the comparison to be fair. Therefore, in terms of efficiency, the proposed methodology achieves the same accuracy as the MC, with a significant reduction of the computational time.

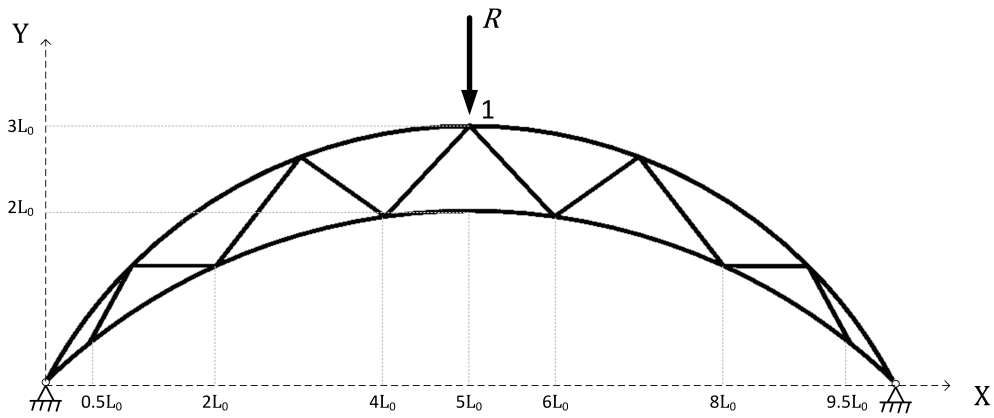


Figure 8.19: Arched beam

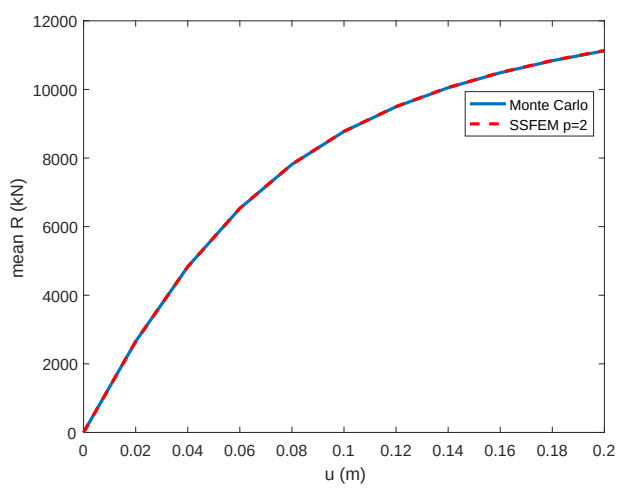


Figure 8.20: Mean value of R for each value of u

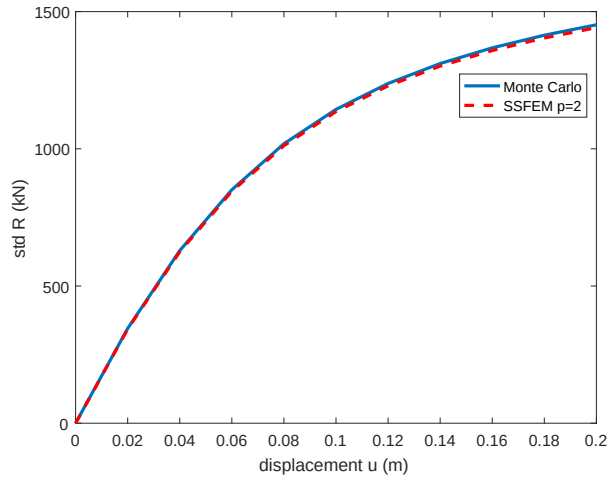


Figure 8.21: Standard deviation of R for each value of u

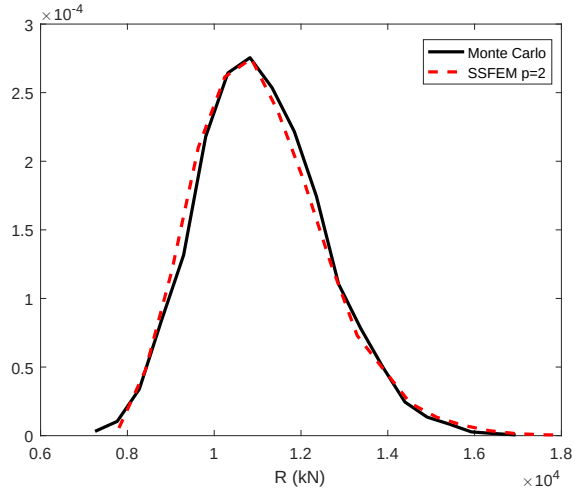


Figure 8.22: Comparison of the pdf of R for $u = 0.20m$ using MC and SSFEM with $p = 2$

8.7 Conclusions

In this chapter, a novel application of the Spectral Stochastic Finite Element Method is presented, which addresses the problem of stochasticity in systems with geometric

nonlinearities. For this purpose, two reformulations of the load control and displacement control method were developed, namely, the stochastic load control and the stochastic displacement control, both of which share the same applicability and limitations as their deterministic counterparts. Moreover, the stochastic kinematic equations for the co-rotational beam elements were presented.

Numerical examples are provided, in which the results obtained from the proposed methodology are compared to those obtained from brute force Monte Carlo simulations. In all cases fair accuracy is reported and the computational merits of the proposed approach are showcased. Future research involves the extension of these ideas to other solution algorithms, such as the arc-length, and other element types (shell elements, 3d beams etc.).

9

Diffusion maps-based surrogate modeling - An alternative machine learning approach

9.1 Introduction

In this chapter an alternative machine learning methodology is proposed, which utilizes nonlinear manifold learning techniques in the frame of surrogate modeling. Under the assumption that the solutions of a parametrized physical system lie on a low-dimensional manifold embedded in a high-dimensional Euclidean space, the goal is to unveil the manifold's intrinsic dimensionality and use it for the construction of a surrogate model, which will be used as a cost-efficient emulator of the high-dimensional physical system. To this purpose, a computational framework based on the Diffusion Maps algorithm is put forth herein, where a set of system solutions is used in order to identify the geometry of a low-dimensional space called the Diffusion Maps space. This space is completely described by a low-dimensional basis, which is constructed from the eigenvectors and eigenvalues of a diffusion operator on the data. The proposed approach exploits the Diffusion maps space's reduced dimensionality for the construction of locally clustered interpolation schemes between the parameter space, the diffusion maps space and the

solution space, which are cheap to evaluate and highly accurate. This way the need to formulate and solve the governing equations of the system is eliminated. In addition, a sampling methodology is proposed based on the metric of the Diffusion Maps space in order to efficiently sample the parameter space, thus ensuring the quality of the surrogate model. Even though it is exploited herein in the premises of uncertainty quantification, this methodology is applicable to any other problem type that depends on some parametric space (i.e. optimization, sensitivity analysis, etc). In the numerical examples it is shown that the proposed surrogate model is capable of high levels of accuracy, as well as significant computational gains.

The rest of this chapter is organized as follows. In section 2, a concise description of DMAPs for manifold learning in high-dimensional data spaces is presented, along with techniques to map points from the ambient space to DMAP space and the inverse. In section 3 the problem of low-dimensional representations for complex systems is addressed and a DMAP-based approach for constructing surrogate models is provided. Section 4 provides numerical examples to test the validity of the proposed methodology and Section 5 concludes with a brief summary and possible extensions.

9.2 Diffusion Maps

9.2.1 General framework

In this chapter, the methodology described in sec. 4.3.3 is adjusted to the problem of the stochastic analysis of finite element systems. Let us consider a data set $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N]$, which consists of vectors $\boldsymbol{u}_i \in \mathbb{R}^d,\ i = 1, \cdots, N$ corresponding to different solution vectors of a parametrized system. If the elements of the vectors differ by orders of magnitude it is advantageous to scale them first using the z-score normalization of eq. (4.37) and work with normalized version of the data set. Next, a connectivity measure κ between two data points $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$, is constructed using the Gaussian kernel:

$$\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) = e^{-(\|\boldsymbol{u}_i - \boldsymbol{u}_j\|/\varepsilon)^2} \tag{9.1}$$

The value of the scale parameter ε can be chosen based on the plot of eq. (4.26).

Subsequently, the symmetric diffusion matrix $\boldsymbol{L} \in \mathbb{R}^{N \times N}$ is constructed as:

$$L_{ij} = \kappa(\boldsymbol{u}_i, \boldsymbol{u}_j) \ , \quad i \text{ and } j \in \{1, ..., N\} \tag{9.2}$$

The anisotropic kernel is defined as

$$\hat{\kappa}(\boldsymbol{u}_i, \boldsymbol{u}_j) = \frac{\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j)}{\left(\sum_{l=1}^{N} \kappa(\boldsymbol{u}_i, \boldsymbol{u}_l)\right)\left(\sum_{l=1}^{N} \kappa(\boldsymbol{u}_j, \boldsymbol{u}_l)\right)} \ , \quad i \text{ and } j \in \{1, ..., N\} \tag{9.3}$$

and the corresponding anisotropic diffusion matrix $\hat{\boldsymbol{L}}$ is given by:

$$\hat{\boldsymbol{L}} = \boldsymbol{D}^{-\alpha} \boldsymbol{L} \boldsymbol{D}^{-\alpha} \tag{9.4}$$

where $\boldsymbol{D}$ is the diagonal matrix such that:

$$D_{ii} = \sum_{q=1}^{N} L_{iq} \tag{9.5}$$

For the definition of the anisotropic diffusion matrix, the parameter α (see eqs. (4.28) and (4.29)) was taken equal to 1 in order for the diffusion to approximate the Laplace-Beltrami operator and provide a description of the Riemannian geometry of the data set, ignoring the distribution of the points. This approach is justified because the aim in this work is to obtain an accurate parametrization of the manifold of the data set and, subsequently exploit the knowledge of its geometry to produce an accurate surrogate for the system under investigation.
Next, the transition matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$ defined as:

$$\mathbf{P} = \hat{\mathbf{D}}^{-1} \hat{\boldsymbol{L}} \tag{9.6}$$

where,

$$\hat{D}_{ii} = \sum_{j=1}^{N} \hat{L}_{ij} \tag{9.7}$$

From the eigenvalue decomposition of $\boldsymbol{P}$ the n largest eigenvalues λ_k for $k = 0, ..., n$ and corresponding right eigenvectors $\{\boldsymbol{\psi}^k\}_{k=0}^{n}$ are computed. Then, the mapping $\boldsymbol{\Psi} : \boldsymbol{U} \to \mathbb{R}^n$

that embeds the data set into an n-dimensional Euclidean space is given by:

$$\boldsymbol{\Psi}(\boldsymbol{u}_i) \triangleq \left[\lambda_1\psi_i^1, \lambda_2\psi_i^2, ..., \lambda_n\psi_i^n\right] \tag{9.8}$$

Collecting the DMAP coordinates of all the data points will lead to the construction of the diffusion maps basis $\hat{\boldsymbol{\Psi}} \in \mathbb{R}^{N\times n}$:

$$\hat{\boldsymbol{\Psi}} = \begin{bmatrix} \boldsymbol{\Psi}(\boldsymbol{u}_1) \\ \boldsymbol{\Psi}(\boldsymbol{u}_2) \\ \vdots \\ \boldsymbol{\Psi}(\boldsymbol{u}_N) \end{bmatrix} = \left[\lambda_1\boldsymbol{\psi}^1, \lambda_2\boldsymbol{\psi}^2, ..., \lambda_n\boldsymbol{\psi}^n\right] \tag{9.9}$$

Let us also recall the DMAP metric δ, which gives the diffusion distance between any two points $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ on the manifold:

$$\delta_t^2(\boldsymbol{u}_i, \boldsymbol{u}_j) = \sum_{l=1}^{n}(\lambda_l)^2\left(\psi^l(\boldsymbol{u}_i) - \psi^l(\boldsymbol{u}_j)\right)^2 = \|\boldsymbol{\Psi}(\boldsymbol{u}_i) - \boldsymbol{\Psi}(\boldsymbol{u}_j)\|^2 \tag{9.10}$$

9.2.2 Out-of-sample extension using the Nyström method

Using the Nyström extension of sec. 4.3.2 the coordinates in the DMAP space of new vector $\boldsymbol{u}$, which does not lie in the original data set, can be estimated. This scheme involves the following steps.

First, the N distances $\|\boldsymbol{u} - \boldsymbol{u}_i\|$ between the new vector and the N vectors in the data set are computed. Then, then values of the anisotropic kernel defined in the previous section are evaluated for the new point as:

$$\hat{\kappa}(\boldsymbol{u}_i, \boldsymbol{u}) = \frac{\kappa(\boldsymbol{u}_i, \boldsymbol{u})}{\sum_{j=1}^{N}\kappa(\boldsymbol{u}_i, \boldsymbol{u}_j)\sum_{j=1}^{N}\kappa(\boldsymbol{u}, \boldsymbol{u}_j)} \tag{9.11}$$

The l-th DMAP coordinate ψ^l of the new point is given as:

$$\psi^l(\boldsymbol{u}) = \lambda_l^{-1}\sum_{i=1}^{N}\frac{\tilde{\kappa}(\boldsymbol{u}_i, \boldsymbol{u})}{\sum_{j=1}^{N}\tilde{\kappa}(\boldsymbol{u}_j, \boldsymbol{u})}\psi_i^l \tag{9.12}$$

and the DMAP representation of $\boldsymbol{u}$ is straightforwardly evaluated as:

$$\boldsymbol{z} = \left[\Psi^1(\boldsymbol{u}), ..., \Psi^n(\boldsymbol{u})\right] = \left[\lambda_1\psi^1(\boldsymbol{u}, ..., \lambda_n\psi^n(\boldsymbol{u})\right] \tag{9.13}$$

It should be mentioned that this procedure will return a result even if $\boldsymbol{u}$ is not chosen to be exactly on the low-dimensional manifold.

9.2.3 'Inverse' Nyström mapping

So far, a mapping between the ambient Euclidean space and the low-dimensional DMAP space has been established. The inverse process, however, is not a trivial task. A number of approaches can be found in the literature for the mapping of a point in a low-dimensional space back in the high-dimensional space. To mention a few, such approaches include simulated annealing [50], polynomial interpolation [146], manifold regularization [17], Laplacian Pyramids [22], geometric harmonics [36] and interpolation using radial basis functions [119]. In a recent study [29], a detailed comparison among different interpolation schemes was performed and the results indicated the interpolation using RBF is capable of achieving high levels of accuracy at small computational cost.

Following the formulation presented in 4.3.1, the current problem setting consists in finding a vector $\boldsymbol{u} \in \mathbb{R}^d$ such that for the given vector $\boldsymbol{z} \in R^n$, $\boldsymbol{z} = \boldsymbol{\Psi}(\boldsymbol{u})$. Using the RBF method, the l-th coordinate of this vector can be obtained as:

$$u_l = \sum_{i=1}^{nn} c_{i,l} \bar{f}\left(\|\boldsymbol{z} - \boldsymbol{z}_i\|\right), \quad l = 1, ..., d \tag{9.14}$$

where $\bar{f}(\cdot)$ is a radial function that depends on the distance between two points in the DMAP space and nn are the nearest neighbors of $\boldsymbol{z}$ in the DMAP space. Among various choices of RBF functions, we restrict our attention to polyharmonic splines of degree p, with p being an odd integer, given by

$$\bar{f}\left(\|\boldsymbol{z} - \boldsymbol{z}_i\|\right) = \|\boldsymbol{z} - \boldsymbol{z}_i\|^p \tag{9.15}$$

Then, the coefficients $c_{i,l}$ of the summation in eq. (9.14) are obtained as

$$[c_{1,l}, ..., c_{nn,l}]^T = \boldsymbol{F}^{-1}[u_{1,l}, ..., u_{nn,l}]^T \tag{9.16}$$

where $\boldsymbol{F}$ is the $nn \times nn$ matrix, whose entries read, $F_{ij} = \|\boldsymbol{z}_i - \boldsymbol{z}_j\|^p$.

9.3 A DMAP-based approach for surrogate modeling

Let us consider the equation of equilibrium of a d-dof stochastic linear static system

$$\mathbf{K}(\boldsymbol{e})\boldsymbol{u}(\boldsymbol{e}) = \mathbf{R}(\boldsymbol{e}) \tag{9.17}$$

where $\boldsymbol{e} \in \mathbb{R}^{\nu}$ is a ν-dimensional vector denoting the random parameters of any source of uncertainty in the system. The main goals in this section are to first identify the geometry of the manifold $\mathcal{U} \in \mathbb{R}^{d}$, where the system's solutions $\boldsymbol{u}$ lie and subsequently, to utilize this knowledge in order to construct a surrogate model of the physical system. This model should be able to estimate accurately and efficiently the displacement vector $\boldsymbol{u}$ given a vector of parameter values $\boldsymbol{e}$ without solving eq. (9.17).

In order to find the image $\boldsymbol{u}$ of a given vector of parameters $\boldsymbol{e}$, it is proposed that, instead of a full model evaluation, $\boldsymbol{e}$ is mapped to its image $\boldsymbol{z}$ in the diffusion maps space via a local transformation $\boldsymbol{\Phi}(\boldsymbol{e}) = \boldsymbol{z}$, that will be established in sec. 9.3.1 and then the inverse Nyström mapping described in section 9.2.3 is used in order to obtain its image on the manifold $\mathcal{U}$. A schematic representation of this approach is demonstrated in fig. 9.1, which depicts the process of associating a new vector of parameters $\boldsymbol{e} \in \mathbb{R}^{\nu}$ with its corresponding solution vector $\boldsymbol{u} \in \mathbb{R}^{d}$ via the composition of maps $(\boldsymbol{\Psi}^{-1} \circ \boldsymbol{\Phi})(\boldsymbol{e})$, bypassing the need for a full model evaluation. In terms of implementation, the only requirement is the knowledge of the local geometry of the manifolds, which is inferred from the distances of a point in the manifold and its nearest neighbors.

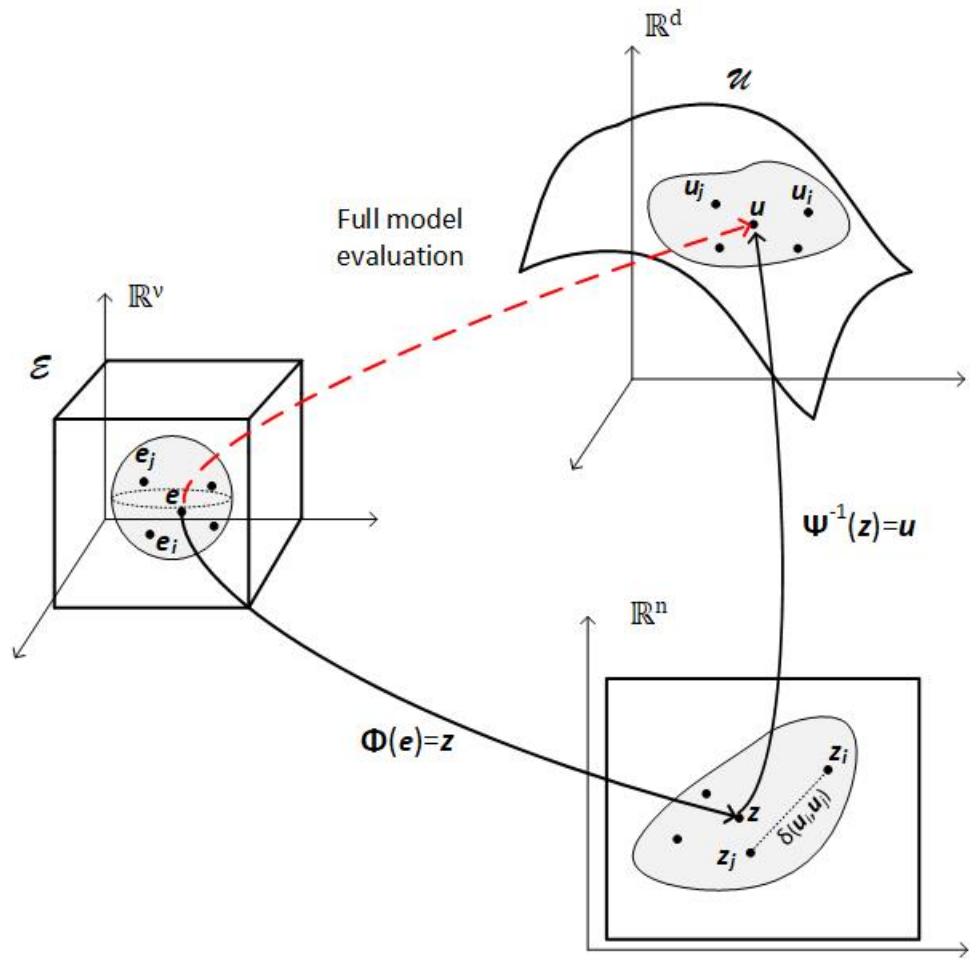


Figure 9.1: Mappings between $\mathscr{E} \in \mathbb{R}^\nu$ and the DMAP space, and the DMAP space and $\mathcal{U} \in \mathbb{R}^d$

9.3.1 Mapping between the parameter space the DMAP space using polynomial interpolation

The aim of this section is to establish a transformation $\boldsymbol{\Phi} : \mathbb{R}^\nu \to \mathbb{R}^n$, that will map a new point $\boldsymbol{e}$ from the parameter space to the DMAP space, that is, $\boldsymbol{\Phi}(\boldsymbol{e}) = \boldsymbol{z}$. First, the ρ closest points to $\boldsymbol{e}$ in the parameter set are determined and note both the DMAP coordinates $\{\boldsymbol{z}_l\}_{l=1}^{\rho}$ and the corresponding parameter coordinates $\{\boldsymbol{e}_l\}_{l=1}^{\rho} \in \mathbb{R}^\nu$ of these ρ points. For each coordinate $z^{(j)}$, $j = 1, ..., n$, these points are used to compute the polynomial coefficients for the polynomial interpolation using $\{\boldsymbol{e}_1, z_1^{(j)}\}$,

$\{\boldsymbol{e}_2, z_2^{(j)}\}$,...,$\{\boldsymbol{e}_\rho, z_\rho^{(j)}\}$. Denoting the result as $f^{(j)}$, then $z^{(j)} = f^{(j)}(\boldsymbol{e})$. This procedure is relatively fast owing to the fact that n is a small number. Typically, ρ can be taken equal to $\binom{m+\nu}{m}$, where m is the degree of polynomials considered in the interpolation. If ρ is taken greater than $\binom{m+\nu}{m}$, the system of equations will be overdetermined and the polynomial coefficients are fitted through least squares.

Lastly, since this approach depends on pairwise distances between points in the parameter space, if the elements of the parameter vector in the given data set represent different physical quantities we normalize them first according to the procedure mentioned in 9.2.1. Let us denote with $\left[\boldsymbol{e}^0\right]$ the $\nu \times N$ matrix of the unnormalized data set. Then, the entries of the normalized version of the data set are given by the following relation for $i = 1, ..., \nu$ and $j = 1, ..., N$:

$$e_{ij} = \frac{e_{ij}^0 - \mu_i}{\sigma_i} \quad \text{for } i = 1, ..., \nu \text{ and } j = 1, ..., N \tag{9.18}$$

with μ_i and σ_i in this case being the statistical mean and standard deviation over all elements in the i-th line of $[\boldsymbol{e}]^0$.

9.3.2 DMAP adaptive sampling algorithm

Besides the purposes of dimensionality reduction, diffusion maps can be further exploited in order to provide an efficient sampling method from the probability space. This stems from the fact that the DMAP space is associated with a metric, called the diffusion distance, which gives an approximation of the 'true' distance between two points on the manifold. As mentioned in sec. 9.2.1 this metric is evaluated as ([35])

$$\delta(\boldsymbol{u}_i, \boldsymbol{u}_j) = \|\boldsymbol{\Psi}(\boldsymbol{u}_i) - \boldsymbol{\Psi}(\boldsymbol{u}_j)\| \tag{9.19}$$

where $\|\cdot\|$ is the Euclidean norm. For every point in the parameter space $\boldsymbol{e}_i$ we can identify a neighborhood around it, consisting of its 'k'-nearest points from the initial set of parameter vectors. Next, the diffusion distances between the image $\boldsymbol{u}_i$ in the solution space of $\boldsymbol{e}_i$ and the corresponding images of the nearest points are computed. By examining these distances, it is straightforward to infer, how much the variation of the parameters in the region of $\boldsymbol{e}_i$ affected the solution. If a small distance in the parameter space leads to a large distance in the DMAP space, then this implies that the solution manifold is not adequately delineated at this region and more parameter vectors

with their respective solution vectors are needed. The aim is to attain an accurate resolution for the manifold, that will ensure the quality of the interpolation scheme, since the Diffusion Maps will now encapsulate the complete information about its geometry.

In order to specify the exact samples needed from the parameter space, that will lead to the accurate manifold resolution, the following algorithm is proposed. By appropriately selecting a prescribed value $\bar{\delta}$, the distances in the DMAP space between the corresponding solution vectors of a point $\boldsymbol{e}_i$ and its k-neighbors are checked if they satisfy the condition $\delta(\boldsymbol{u}_i, \boldsymbol{u}_j) < \bar{\delta}$ for $j = 1, .., k$. The number k can be taken equal to ρ, which is the number of neighbors used in the polynomial interpolation of section 9.3.1, since these are the points that will be utilized in order to construct the mapping that will send a new parameter vector to the DMAP space. If the above condition is not met between the two points $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$, then a new point in the parameter space $\boldsymbol{e}_{new}$ is considered as the convex combination of $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$, plus a small perturbation given by the random vector $\boldsymbol{w}$. Thus, $\boldsymbol{e}_{new} = c\boldsymbol{e}_i + (1-c)\boldsymbol{e}_j + \boldsymbol{w}$, with c initially taken equal to 0.5. The role of $\boldsymbol{w}$ is to prevent $\boldsymbol{e}_{new}$ from being a linear combination of $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$, since this might affect the rank of the matrices to be inverted in the polynomial interpolation described in the previous section. Its values can be taken equal to the difference $\boldsymbol{e}_i - \boldsymbol{e}_j$, multiplied element-wise by a vector of random variables uniformly distributed in the $[-0.05, 0.05]$ domain. Then, the full system is solved for the new parameter values and the response vector $\boldsymbol{u}_{new}$ is obtained. The image of $\boldsymbol{u}_{new}$ in the DMAP space is directly obtained using the Nyström extension of eq. (9.12) and the distance $\delta(\boldsymbol{u}_i, \boldsymbol{u}_{new})$ is re-evaluated and checked to satisfy the condition $\delta(\boldsymbol{u}_i, \boldsymbol{u}_{new}) < \bar{\delta}$. If the condition is still not met, the parameter c is increased until a point $\boldsymbol{u}_{new}$ is found, that satisfies the criterion. This process is repeated for all points in the parameter space, and once all additional points have been obtained, the DMAP algorithm is applied to the augmented data set of solution vectors.

An algorithm describing the above procedure is given below.

9.3.3 Construction of the DMAP surrogate

In the previous sections the basic features of the diffusion maps algorithm have been outlined and ways to utilize them in the context of surrogate modeling have been proposed. As a further step, all these methodologies are combined in a unified framework

Data: matrices of parameter vectors $[\boldsymbol{e}]$, solution vectors $\boldsymbol{U} = [\boldsymbol{u}_1, ..\boldsymbol{u}_N]$ and DMAP coordinates $\hat{\Psi} = [\boldsymbol{z}_1, \boldsymbol{z}_2, ..\boldsymbol{z}_N]^T$

Result: additional points $\boldsymbol{e}$ and $\boldsymbol{u}$

```
initialization
Select values for k and δ̄
for i = 1 : N do
    find the k-nearest neighbors of e_i in the parameter space
    for j = 1 : k do
        if δ(u_i, u_j) < δ̄ then
            exit if-loop and proceed with next point
        end
        else
            e_new = c e_i + (1 − c) e_j + w
            perform a full model evaluation for e_new and obtain u_new
            obtain z_new using the Nyström mapping
            store e_new, u_new, z_new
            if δ(u_i, u_new) < δ̄ then
                exit if-loop and proceed with next point
            end
            else
                increase c and repeat steps 9-12
            end
        end
    end
end
Perform the DMAP algorithm for the augmented data set of solutions
```

Algorithm 1: DMAP adaptive sampling algorithm

in order to construct a surrogate model, that can be used to accelerate computations, in the context of non-intrusive Monte Carlo simulations. To necessary steps for the algorithmic implementation of the proposed procedure can be summarized as follows:

1. Generate N samples of the vector of random parameters $\boldsymbol{e}$ and perform N Monte Carlo simulations to obtain the $d \times N$ matrix of solutions $\boldsymbol{U} = [\boldsymbol{u_1}, \boldsymbol{u_2}, ..., \boldsymbol{u_N}]$.

2. Given the matrix $\boldsymbol{U}$ of the N distinct solutions, utilize the machinery provided by the diffusion maps to unveil the underlying manifold structure (Sec. 9.2.1) and identify the low dimensional representation $\hat{\boldsymbol{\Psi}}$ in the DMAP space of the data set $\boldsymbol{U}$.

3. Utilize the metric in the DMAP space to locate areas of $\mathcal{U}$, where higher resolution is required and identify the additional points $\boldsymbol{e} \in \mathbb{R}^{\nu}$ with the methodology proposed in 9.3.2

4. For each additional point $\boldsymbol{e}$, perform a full model evaluation to obtain its image in $\mathcal{U}$ and then, via the out-of-sample extension of sec. 9.2.2, obtain its image in the diffusion maps space.

Start Monte Carlo simulations

5. For a given vector of parameter values $\boldsymbol{e}$ determine the mapping between the parameter space $\mathcal{E} \in \mathbb{R}^{\nu}$ and the diffusion maps space, and obtain its image $\boldsymbol{\Phi}(\boldsymbol{e}) = \boldsymbol{z} \in \mathbb{R}^n$ using the methodology of sec. 9.3.1.

6. Use the inverse Nyström transform of sec. 9.2.3 to map the point $\boldsymbol{z} \in \mathbb{R}^n$ back to the manifold $\mathcal{U}$.

9.4 Numerical study

In this section numerical examples are presented to test the accuracy and efficiency of the proposed methodology.

9.4.1 Test case 1: Cantilever

As a first illustrative example we consider the cantilever depicted in fig. 9.2, where the geometry and boundary conditions of the problem are also shown. The loading consists of a horizontal force $P_h = 1kN$, a vertical force $P = 10N$ and a concentrated moment $M = 100Nm$ applied at its right side. The cantilever's length L, the modulus of elasticity E and the moment of inertia I are all considered to be random variables with parameters those given in table 9.1.

Table 9.1: Random parameters

parameter	distribution	mean	CoV
L	Lognormal	$5m$	0.2
E	Lognormal	$5 \cdot 10^7 kN/m^2$	0.2
I	Lognormal	$52 \cdot 10^{-8} m^4$	0.2

The cantilever system is modeled as a 2D beam finite element system consisting of 10 elements with three active degrees of freedom at each node. Therefore, a solution vector will contain the 30 free degrees of freedom, or, $\boldsymbol{u} \in \mathbb{R}^{30}$ and the solution space will be a subset of $\mathbb{R}^{30}$.

Needless to say, that for this small-scale system there is no practical need to use a surrogate model, however it serves as a good basis for comparison since the relation between the solution vector $\boldsymbol{u}$ and the system parameters $\{E, I, L\}$ is highly nonlinear.

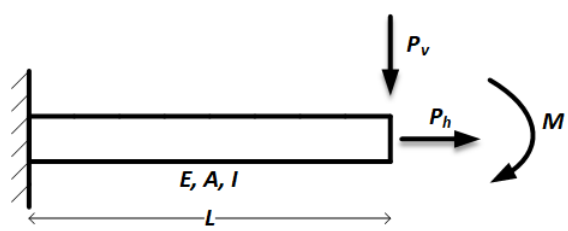
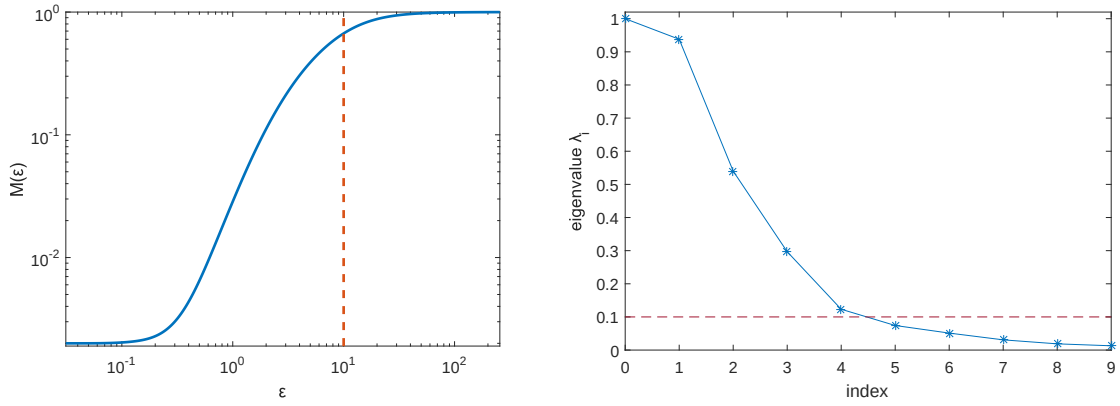


Figure 9.2: Cantilever with random parameters

The vertical displacement and the rotation at its free end are denoted with u_v and θ and are considered quantities of interest. In order to test the accuracy of the proposed

methodology, 50000 MC simulations were performed on the full model and the results were compared to those obtained from the DMAP-based surrogate model. For the construction of the surrogate several sets consisting of 500 to 2000 training points (full-model solutions) with an increment of 250 points were used and the comparison is made with respect to the probability density functions produced by both approaches. The value of the diffusion kernel parameter ε for the case $N = 500$ was chosen according to the plot of $M(\varepsilon)$ versus ε depicted in fig. 9.3a. Our investigation showed that optimal results in the sense of achieving maximum data compression without significant losses in accuracy were obtained using the highest values of ε at the end of the linear part of the spectrum in fig. 9.3a. In this case, ε was taken equal to 10 and the eigenvalues of the corresponding diffusion kernel are shown in fig. 9.3b. For the representation of the data set in the DMAP space, the first eigenvalues satisfying the criterion $\lambda_i > 0.1$ were retained, which led to four non-trivial eigenfunctions. The same approach was employed for all other cases where the optimal values of ε were obtained in the range $[10, 12]$.

After identifying the geometry of the DMAP space for all cases, the next step is to construct the interpolation schemes between a) the parameter space $\mathscr{E}$ and the DMAP space, and b) the DMAP space and the solution manifold $\mathscr{U}$. More specifically, for the polynomial interpolation second-order polynomials were used while for the RBF interpolation cubic splines ($p = 3$) over 50 nearest neighbors yielded the best results. In all cases the percent error given by $e\% = \|p_{MC}(\cdot) - p_{DMAP}(\cdot)\|_2 / \|p_{MC}(\cdot)\|_2 \times 100\%$ was considered, where $\|\cdot\|_2$ denotes the L_2 norm and $p_{MC}(\cdot)$, $p_{DMAP}(\cdot)$ are the pdf's of the respective quantities obtained via the MC simulation and the proposed surrogate, respectively. Apart from these quantities, the residual $res = \|\boldsymbol{u} - \tilde{\boldsymbol{u}}\|_2$, given by the L_2-norm of the difference between the true solution vector $\boldsymbol{u}$ and the approximate solution $\tilde{\boldsymbol{u}}$, is measured in order to quantify the quality of the surrogate model. Figures 9.4, 9.5 and 9.6 depict the pdf's of u_v, θ and the histogram of res for the cases $N = 500, 1500, 2000$. A convergence analysis of the error in the pdf's with respect to the number of samples is illustrated in fig. 9.7a (solid lines), while 9.7b plots the mean value of res as a function of the number of samples. Upon inspection of these figures one quickly comes to the conclusion that for small number of samples the quality of the results is poor, but can be significantly improved by increasing the sample size.



(a) $M(\varepsilon)$ vs ε **(b)** Eigenvalue spectrum for $\varepsilon = 10$

Figure 9.3: Selection of ε and number of eigenfunctions for the DMAP representation for $N = 500$

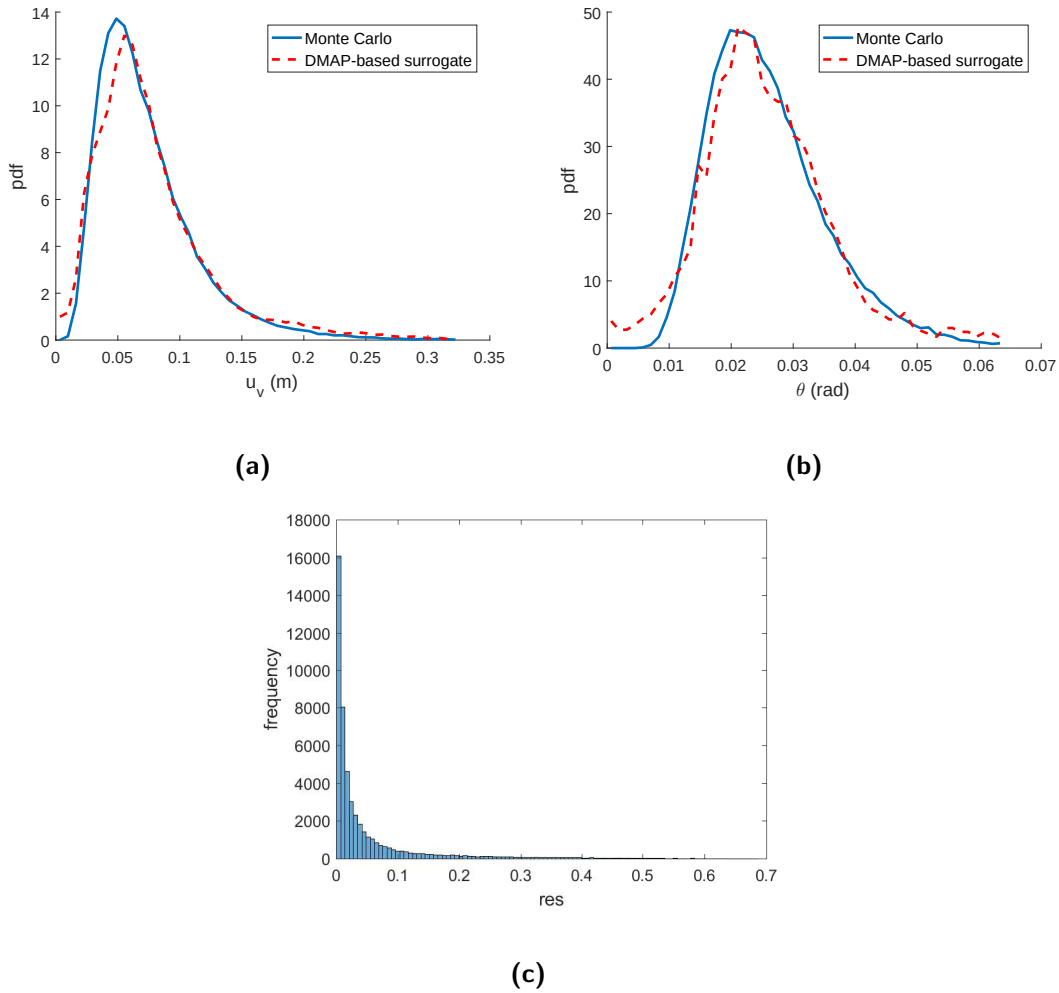


Figure 9.4: Comparison of the probability density functions for (a) u_v, (b) θ and (c) histogram of the residual for the case $N = 500$

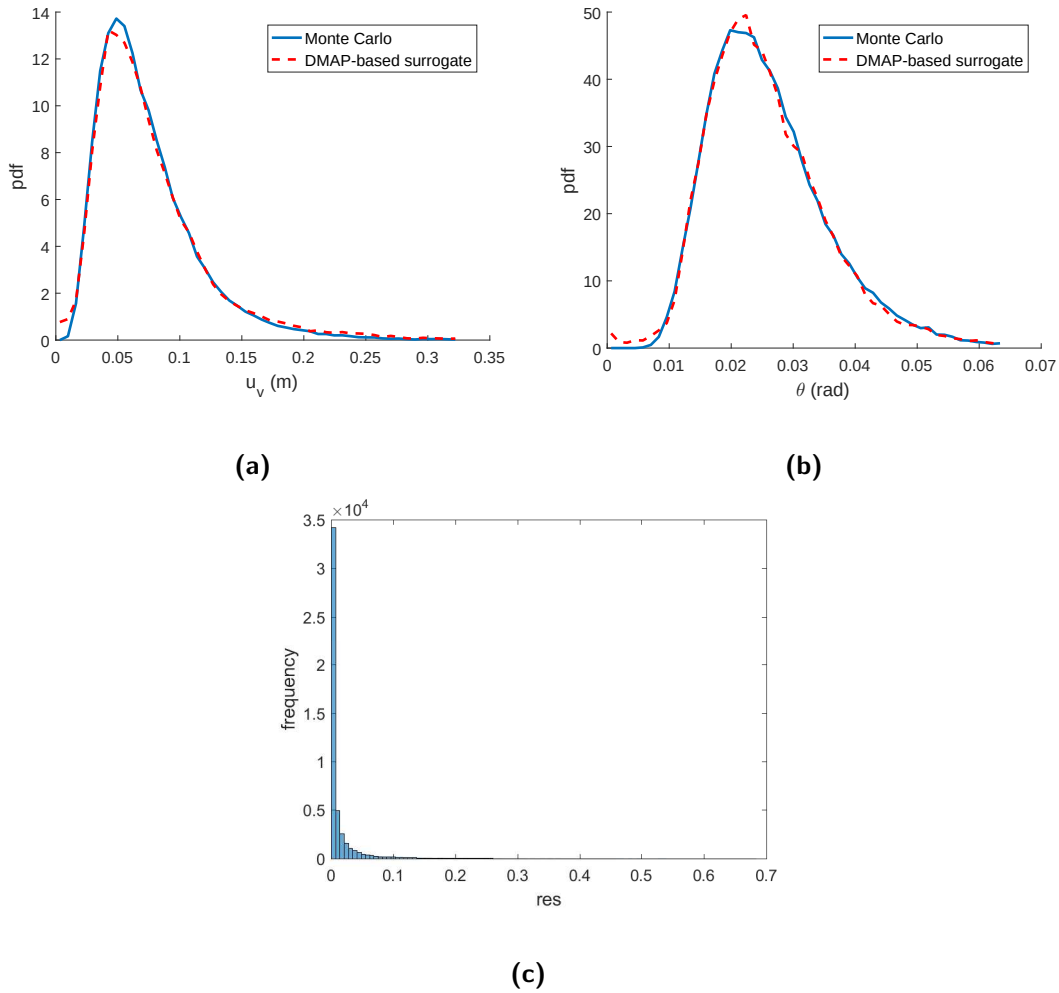


Figure 9.5: Comparison of the probability density functions for (a) u_v, (b) θ and (c) histogram of the residual for the case $N = 1500$

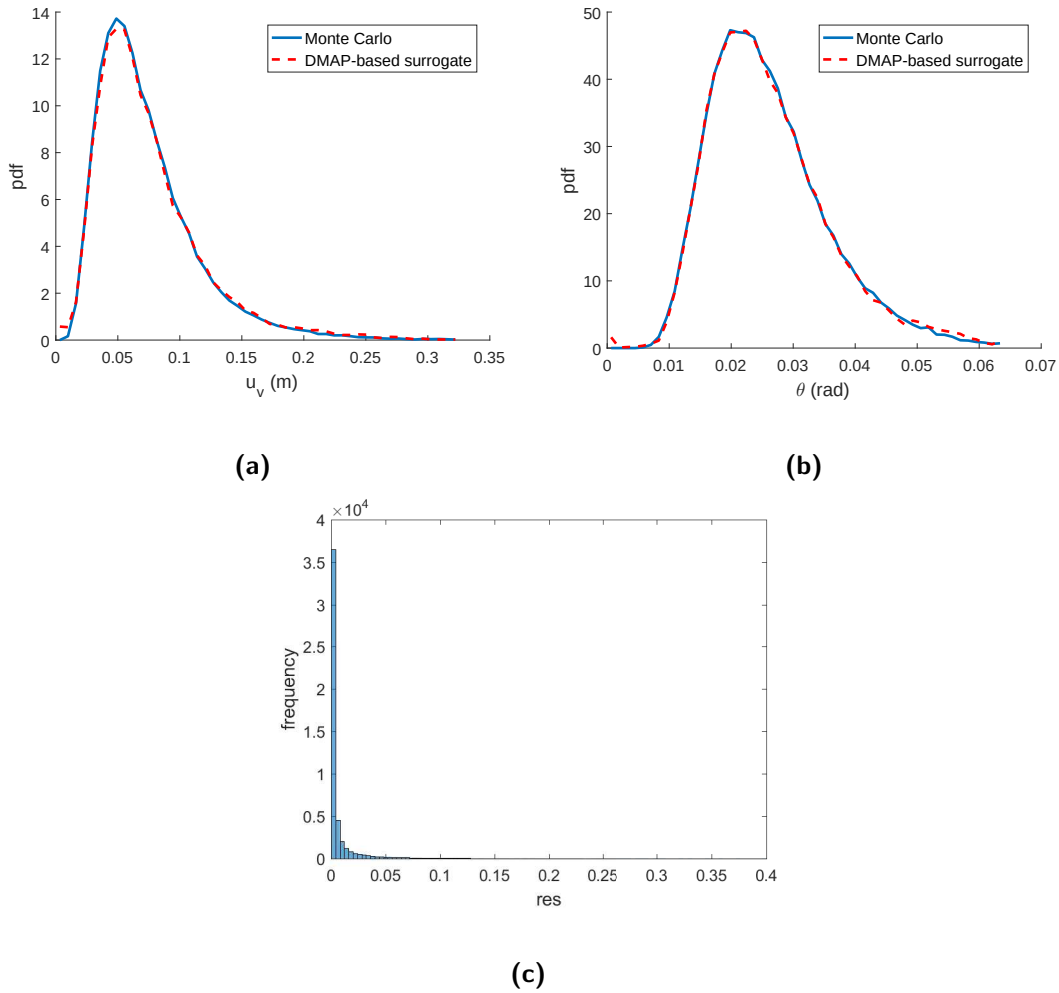
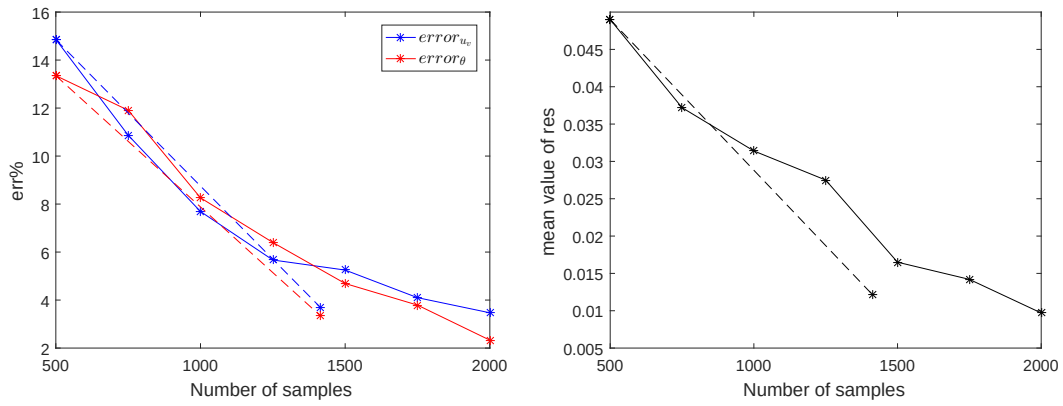


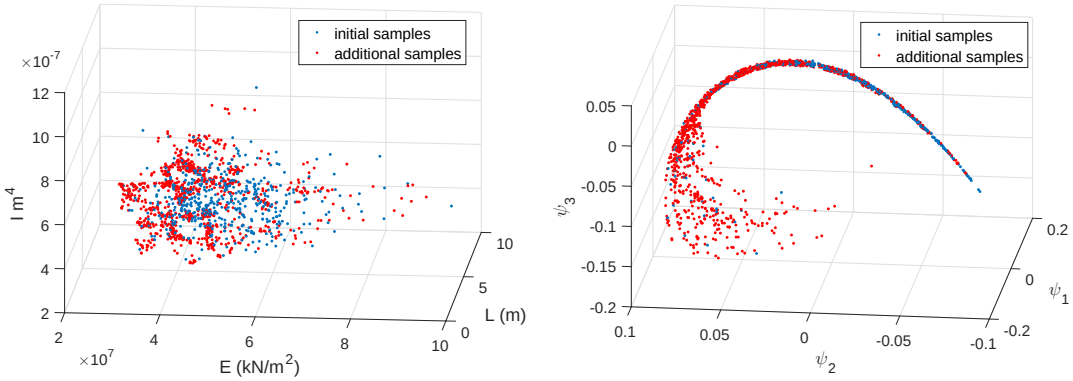
Figure 9.6: Comparison of the probability density functions for (a) u_v, (b) θ and (c) histogram of the residual for the case $N = 2000$



(a) **(b)**

Figure 9.7: Convergence analysis of (a) the error in the probability density functions of u_v, θ and (b) mean value of res

Next, the influence of the proposed sampling methodology in the accuracy of the surrogate model is investigated. For its application, the DMAP distances of the 10 nearest neighbors of every point in the parameter space were checked to satisfy the criterion $\bar{\delta} < 0.01$. With this approach, it was found that 914 additional samples were required, in order to have an accurate resolution of the solution space, resulting in a total of 1414 full model evaluations. The distributions of the additional samples in the parameter space and the DMAP space using the first three coordinates are illustrated in figs. 9.8a and 9.8b, respectively. Next, the DMAP algorithm is performed using $\varepsilon = 11$ and in figure 9.9 the pdfs of u_v and θ, as well as the histogram of res are depicted. In this case, the errors in the pdf's of the quantities of interest were 3.70% and 3.35%, while the mean value of the residual was 0.012, as shown in fig. 9.7(dashed lines). These results are significantly better than those obtained using 1500 random samples and comparable to the case $N = 2000$. However, the clustering of the additional points in the parameter space indicate that it is possible to further optimize the sampling algorithm.



(a) Distribution of additional samples in the parameter space

(b) Distribution of additional samples in the DMAP space

Figure 9.8

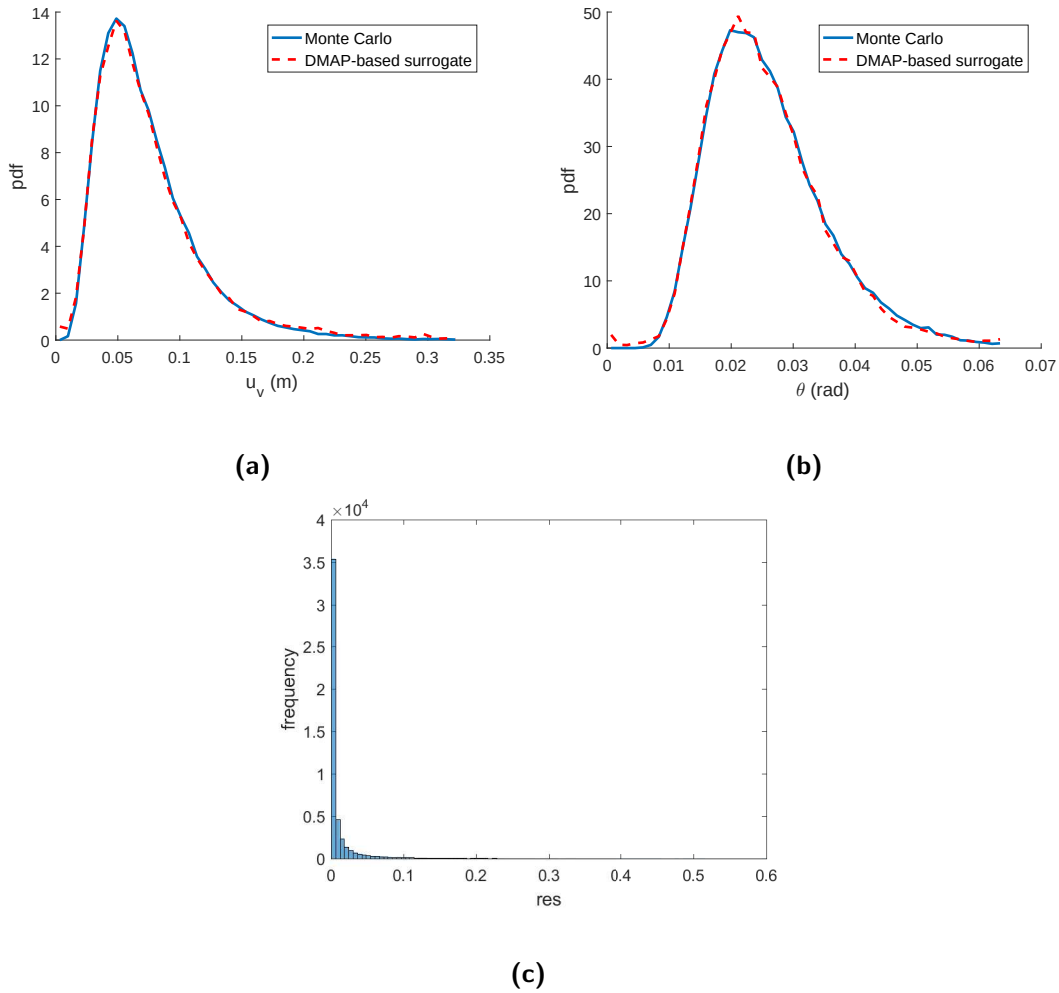


Figure 9.9: Comparison of the probability density functions for (a) u_v, (b) θ and (c) histogram of the residual using the proposed sampling methodology

9.4.2 Test case 2: 3-D elasticity problem

For the second test case, the 3D wrench-like model depicted in figure 9.10 is considered. The structure is subjected to two uniformly distributed vertical loads $q_1 = 4000kN/m^2$, $q_2 = 10000kN/m^2$ and a concentrated force $Q = 20kN$ as shown in this figure. Regarding the boundary conditions, all nodes at the rightmost surface of the structure are assumed to be pinned. The structure is discretized into 620 solid elements leading to 2916 degrees

of freedom in the system. The modulus of elasticity is modeled as an one-dimensional lognormal random field, which is the result of the exponentiation of an underlying Gaussian field with mean 16.817 and autocorrelation function $R(\tau_x) = \sigma^2 exp\left(-\frac{\tau_x}{b}\right)$ with correlation length parameters $b = 0.10$ and standard deviation $\sigma = 0.2936$. The Karhunen-Loève expansion was employed for the generation of 50000 random realizations for the field, where only the 3 largest terms were retained in the expansion. The decay in the eigenvalue spectrum is depicted in fig. 9.11. The lengths L_1, L_2, L_3 and L_4, were considered as random variables, with uniform distributions, as shown in table 9.2.

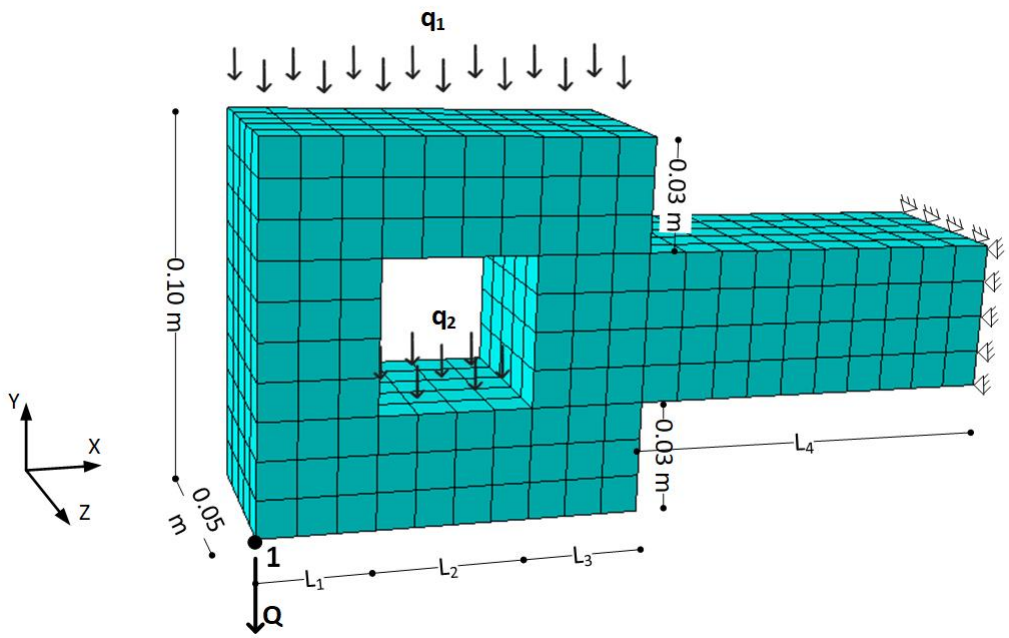


Figure 9.10: Wrench

Table 9.2: Random lengths of the system

random variable	distribution
L_1	$\mathcal{U}(0.02, 0.04)$ (m)
L_2	$\mathcal{U}(0.02, 0.06)$ (m)
L_3	$\mathcal{U}(0.02, 0.04)$ (m)
L_4	$\mathcal{U}(0.07, 0.13)$ (m)

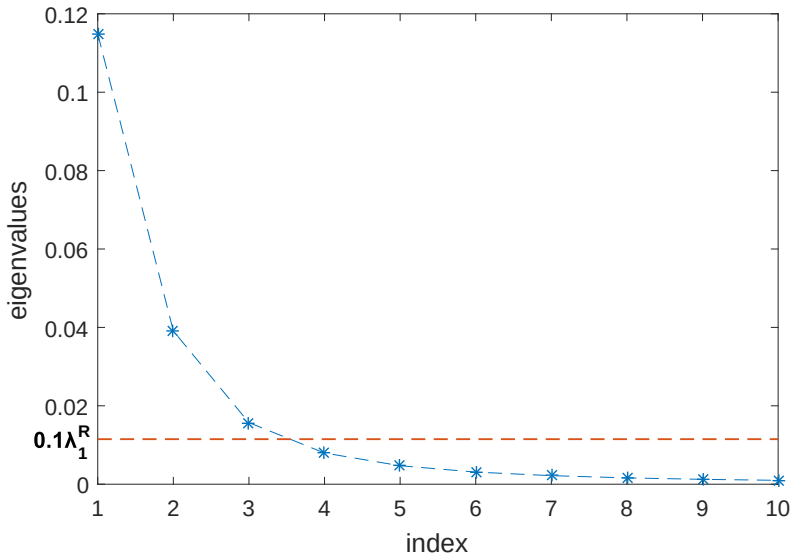
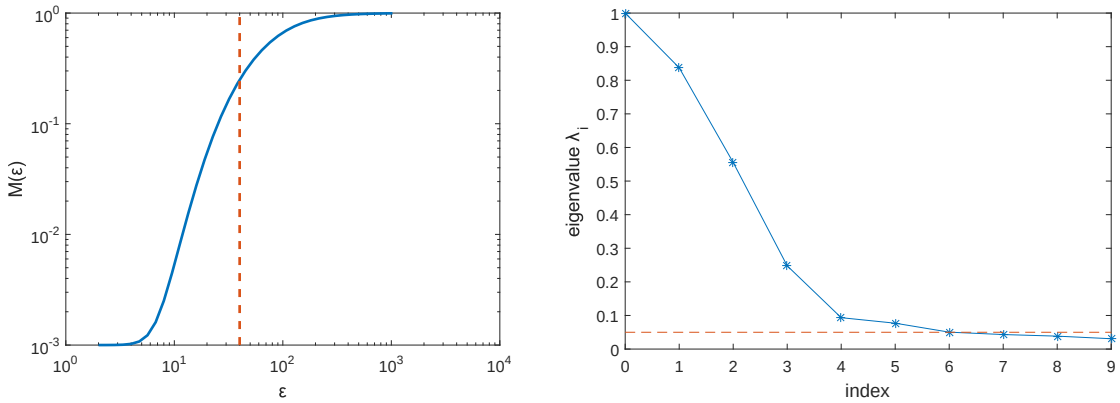


Figure 9.11: Eigenvalues of autocorrelation kernel



(a) $M(\varepsilon)$ vs ε **(b)** Eigenvalue spectrum for $N = 1000$

Figure 9.12: Selection of ε and number of eigenfunctions for the DMAP representation

In this example, the initial data set used for the learning of the surrogate model consisted of 1000 model evaluations and the DMAP algorithm was applied to the set

using $\varepsilon = 80$ (fig.9.12a). For this example, the criterion to select the eigenvalues was $\lambda_i > 0.05$, where it was found that 6 eigenvectors were adequate for the representation of the data in the DMAP space, as shown in 9.12b. Moreover, in this case the best results were obtained using first order polynomials for the polynomial interpolation and cubic splines over 20 nearest neighbors for the RBF scheme.

In order to test the quality of the surrogate model, the vertical displacement at the bottom corner node 1, u_1 was considered a quantity of interest, and the residual, $res = \|\boldsymbol{u} - \tilde{\boldsymbol{u}}\|_2$, between the exact and the approximate solution, was monitored as well. The exact solutions were obtained from 50000 MC simulations and the results were compared to those produced by the DMAP-based surrogate for the same number of analyses. Fig. 9.13a depicts the pdf of u_1 and 9.13b displays the histogram of res. The error with respect to the Monte Carlo simulation in this case is 20.58% and the mean value of res is 0.27. It becomes evident that the initial sample size is inadequate to produce accurate results.

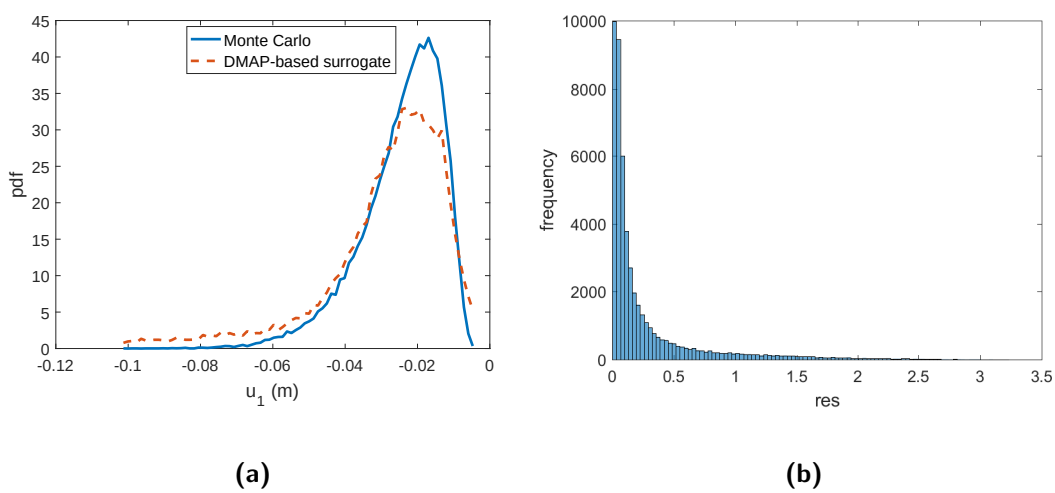
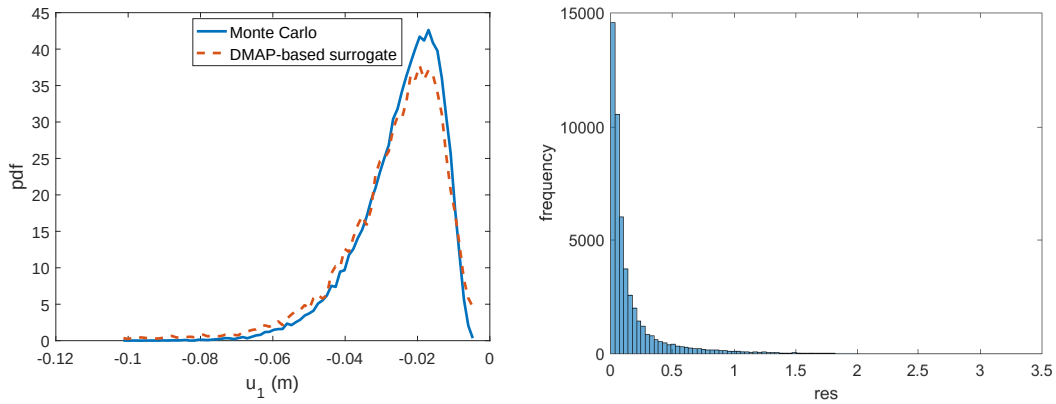


Figure 9.13: (a) Comparison of the probability density functions for u_1 and (b) histogram of the residual for the case $N = 1000$

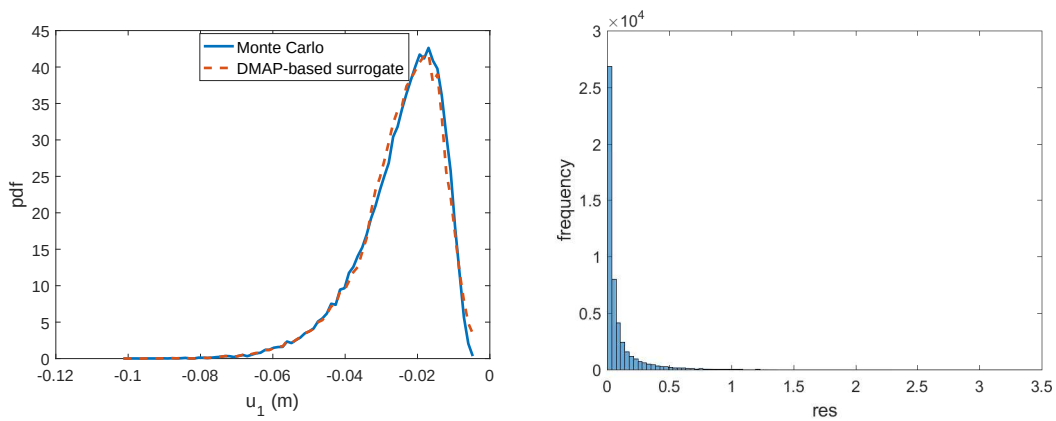
For this purpose, the convergence of the model with respect to the number of samples was investigated, where sets consisting of $N = 2000$ and 3000 model solutions were tested. The kernel parameters were taken equal to 90 and 95 in these cases, while the parameters in the polynomial interpolation and the RBF scheme remained the same.

The corresponding pdf's and histograms for these cases are given in figs. 9.14 and 9.15, where a noticeable improvement can be reported. In fig. 9.16 the results of the convergence analysis are plotted.



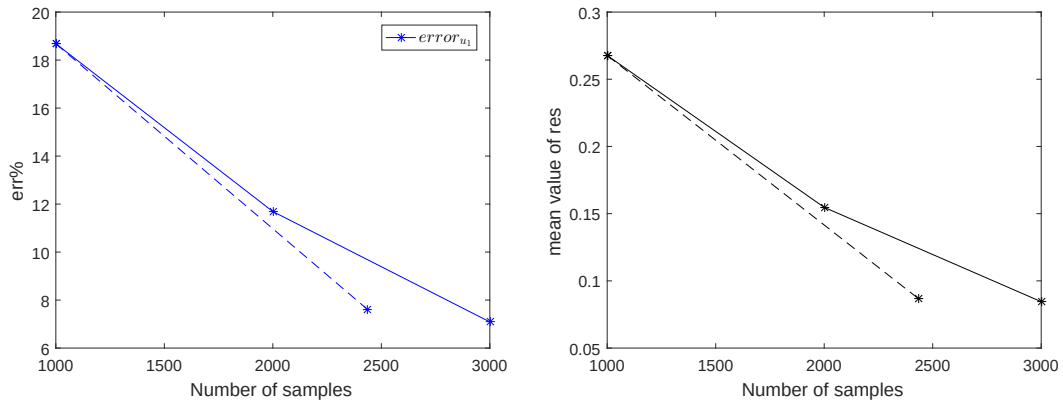
(a) **(b)**

Figure 9.14: (a) Comparison of the probability density functions for u_1 and (b) histogram of the residual for the case $N = 2000$



(a) **(b)**

Figure 9.15: (a) Comparison of the probability density functions for u_1 and (b) histogram of the residual for the case $N = 3000$



(a) **(b)**

Figure 9.16: Convergence analysis of (a) the error in the probability density function of u_1 and (b) mean value of res

Lastly, a comparison in terms of the computational cost for the 50000 full-model evaluations vs. the 50000 surrogate model evaluations is given in table 9.3, where a significant computational gain can be reported.

Table 9.3: Computational cost

	No. of analyses	CPU time (sec)
detailed model	50000	46532
surrogate model	50000	876

9.5 Conclusions

In this chapter a novel approach for constructing surrogate models based on the machinery provided by the Diffusion Maps has been demonstrated. Using the DMAP algorithm a significant dimensionality reduction was achieved for large scale stochastic systems. Also, the DMAP metric was utilized for the development of a sampling refinement strategy. Two mappings were proposed that would map a point from the parameter space to the DMAP space, and subsequently, from the DMAP space to the solution space. With these tools at hand, a methodology for the construction of a surrogate

model was developed, which was capable of emulating the system's input-output relation. The proposed methodology was applied to problems of large scale structural uncertainty quantification, where it was shown to reduce the computational cost with high levels of accuracy.

The applications presented herein involved linear stochastic systems and served as validation to the proposed methodology. A more meaningful approach would be to test the performance of DMAP-based surrogate models to more demanding problems such as nonlinear and/or dynamic systems in the frame of stochastic analysis and/or optimization. This direction remains in the scope of future work.

10

Summary - Innovation of thesis

This thesis presented a theoretical and numerical framework developed specifically to address the issue of uncertainty in the analysis of structural systems. In this context, a review of the most eminent intrusive and non-intrusive stochastic finite element methods was provided and their applications in structural analysis were further investigated. The focus of this research was to reduce the computational demands of the Monte Carlo simulation, by utilizing surrogate models and machine learning techniques. A summary of the main outcomes and innovations of this work is given below.

First, a reformulation of the classic PDEM was presented to address the case of general stochastic finite element systems. Also, the accuracy of various numerical schemes for the solution of the corresponding generalized density evolution equations was investigated. It was shown that the Streamline Upwind/Petrov Galerkin (SUPG) method, endowed with a shock capturing term is more accurate than the commonly used Lax-Wendroff finite difference scheme and the Discontinuous Galerkin time-marching scheme. Moreover, SUPG resulted in a decrease in the computational effort due to the fact that this method is not restricted by the Courant-Friedrichs-Lewy condition.

Furthermore, an application of PDEM for the estimation of stochastic limit loads was presented. In this regard, a flow of probability with respect to the limit load was proposed and the corresponding partial differential equations to describe this flow were developed. Moreover, the displacement control method was extended in the frame of

PDEM in order to produce the stochastic nonlinear path of equilibrium for the structure under investigation. The accuracy and efficiency of the proposed methodology was demonstrated through numerical examples.

In addition, an extension of SSFEM was proposed to address the issue of geometrical nonlinearities in framed structures. In this context, the co-rotational beam formulation was extended to its stochastic counterpart and stochastic versions of the displacement and load-control algorithms were developed in order to solve the augmented system of nonlinear equations. This formulation involved the repeated evaluation of high-dimensional integrals of polynomials and non-polynomial functions which proved to be computationally inefficient. For this purpose, the idea of re-projecting the complex non-polynomial functions onto the lower-dimensional polynomials of the Polynomial Chaos expansion was proposed and it was shown to reduce drastically the computational cost. The efficiency of the proposed methodology was verified through benchmark problems with strong geometrically nonlinear behavior.

Lastly, a methodology to construct surrogate models of detailed physical systems was proposed, which was combined with non-intrusive Monte Carlo simulation. A manifold learning algorithm, namely the diffusion maps algorithm was utilized for the identification of the geometry of a low-dimensional manifold where the system's solutions lie. The manifold's intrinsic dimensionality, being significantly smaller than the ambient Euclidean space of the solution vectors, allowed for the construction of efficient interpolation schemes that mapped points from the parameter space to the diffusion maps space and then to the solution space. This approach overrode the need to formulate and solve the governing equations of the system, and thus, it led to significant computational gains. In addition, the metric in the DMAP space was further utilized for the development of a sampling scheme in order to accurately delineate the solution manifold. The accuracy and efficiency of the proposed methodology was demonstrated through a set of numerical examples of structural uncertainty quantification.

References

[1] Abramowitz, M. & Stegun, I. (1972). *Handbook of Mathematical Functions*. Dover.

[2] Acharjee, S. & Zabaras, N. (2006). Uncertainty propagation in finite deformations - a spectral stochastic lagrangian approach. *Computer Methods in Applied Mechanics and Engineering*, 195(19-22), 2289–2312.

[3] Amsallem, D. & Farhat, C. (2008). Interpolation method for adapting reduced-order models and application to aeroelasticity. *American Institute of Aeronautics and Astronautics*, 46(7), 1803–1813.

[4] Amsallem, D., Zahr, M. J., & Farhat, C. (2012). Nonlinear model order reduction based on local reduced-order bases. *International Journal for Numerical Methods in Engineering*, 92(10), 891–916.

[5] Anders, M. & Hori, M. (1999). Stochastic finite element method for elasto-plastic body. *International Journal for Numerical Methods in Engineering*, 46(11), 1897–1916.

[6] Anders, M. & Hori, M. (2001). Three-dimensional stochastic finite element method for elasto-plastic bodies. *International Journal for Numerical Methods in Engineering*, 51(4), 449–478.

[7] Argyris, J., Papadrakakis, M., & Stefanou, G. (2002). Stochastic finite element analysis of shells. *Computer Methods in Applied Mechanics and Engineering*, 191(41–42), 4781 – 4804.

[8] Arnst, M. & Ghanem, R. (2012). A variational-inequality approach to stochastic boundary value problems with inequality constraints and its application to contact and elastoplasticity. *International Journal for Numerical Methods in Engineering*, 89(13), 1665–1690.

[9] Au, S.-K. & Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4), 263 – 277.

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

[11] Bandler, J. W., Cheng, Q. S., Dakroury, S. A., Mohamed, A. S., Bakr, M. H., Madsen, K., & Sondergaard, J. (2004). Space mapping: the state of the art. *IEEE Transactions on Microwave Theory and Techniques*, 52(1), 337–361.

[12] Baroth, J., Bodé, L., Bressolette, P., & Fogli, M. (2006). SFE method using hermite polynomials: An approach for solving nonlinear mechanical problems with uncertain parameters. *Computer Methods in Applied Mechanics and Engineering*, 195(44–47), 6479 – 6501.

[13] Baroth, J., Bressolette, P., Chauvière, C., & Fogli, M. (2007). An efficient SFE method using lagrange polynomials: Application to nonlinear mechanical problems with uncertain parameters. *Computer Methods in Applied Mechanics and Engineering*, 196(45–48), 4419 – 4429.

[14] Bathe, K. J. (2007). *Finite Element Procedures.*

[15] Batoz, J. L. & Dhatt, G. (1979). Incremental displacement algorithms for nonlinear problems. *International Journal for Numerical Methods in Engineering*, 14(8), 1262–1267.

[16] Belkin, M. & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.*, 15(6), 1373–1396.

[17] Belkin, M., Niyogi, P., & Sindhwani, V. (2006). Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *The Journal of Machine Learning Research*, (7), 2434.

[18] Bertsekas, D. & Tsitsiklis, J. (2008). *Introduction to Probability, 2nd Edition.* Athena Scientific.

[19] Bilionis, I., Zabaras, N., Konomi, B. A., & Lin, G. (2013). Multi-output separable gaussian process: Towards an efficient, fully bayesian paradigm for uncertainty quantification. *J. Comput. Physics*, 241, 212–239.

[20] Brooks, A. N. & Hughes, T. J. (1982). Streamline upwind/petrov-galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32(1–3), 199 – 259.

[21] Broomhead, D. & Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2, 321–355.

[22] Burt, P. & Adelson, E. (1983). The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4), 532–540.

[23] Chen, J. & Zhang, S. (2013). Improving point selection in cubature by a new discrepancy. *SIAM Journal on Scientific Computing*, 35(5), A2121–A2149.

[24] Chen, J.-B., Ghanem, R., & Li, J. (2009). Partition of the probability-assigned space in probability density evolution analysis of nonlinear stochastic structures. *Probabilistic Engineering Mechanics*, 24(1), 27 – 42.

[25] Chen, J. B. & Li, J. (2007). Development-process-of-nonlinearity-based reliability evaluation of structures. *Probabilistic Engineering Mechanics*, 22(3), 267 – 275.

[26] Chen, J.-B. & Li, J. (2008). Strategy for selecting representative points via tangent spheres in the probability density evolution method. *International Journal for Numerical Methods in Engineering*, 74(13), 1988–2014.

[27] Chen, J.-B. & Li, J. (2009). A note on the principle of preservation of probability and probability density evolution equation. *Probabilistic Engineering Mechanics*, 24(1), 51 – 59.

[28] Chen, J. B. & Li, J. (2010). Stochastic seismic response analysis of structures exhibiting high nonlinearity. *Computers & Structures*, 88(7–8), 395 – 412.

[29] Chiavazzo, E., Gear, C., Dsilva, C., Rabin, N., & Kevrekidis, I. (2014). Reduced models in chemical kinetics via nonlinear data-mining. *Processes*, 2(1), 112–140.

[30] Cho, H., Venturi, D., & Karniadakis, G. (2013). Adaptive discontinuous galerkin method for response-excitation pdf equations. *SIAM Journal on Scientific Computing*, 35(4), B890–B911.

[31] Chung, F. R. K. (1997). *Spectral Graph Theory.* American Mathematical Society.

[32] Cockburn, B., Lin, S.-Y., & Shu, C.-W. (1989). TVB runge-kutta local projection discontinuous galerkin finite element method for conservation laws III: One-dimensional systems. *Journal of Computational Physics*, 84(1), 90 – 113.

[33] Cockburn, B. & Shu, C.-W. (1998). The local discontinuous galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35(6), 2440–2463.

[34] Cockburn, B. & Shu, C.-W. (2001). Runge-kutta discontinuous galerkin methods for convection-dominated problems. *Journal of Scientific Computing*, 16(3), 173–261.

[35] Coifman, R. R. & Lafon, S. (2006a). Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1), 5 – 30. Special Issue: Diffusion Maps and Wavelets.

[36] Coifman, R. R. & Lafon, S. (2006b). Geometric harmonics: A novel tool for multiscale out-of-sample extension of empirical functions. *Applied and Computational Harmonic Analysis*, 21(1), 31 – 52. Special Issue: Diffusion Maps and Wavelets.

[37] Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., & Zucker, S. W. (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences*, 102(21), 7426–7431.

[38] Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.

[39] Courant, R., Friedrichs, K., & Lewy, H. (1928). Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische Annalen*, 100(1), 32–74.

[40] Courant, R. & Hilbert, D. (2008). *Methods of Mathematical physics: Partial differential equations, vol.2.* John Wiley & Sons.

[41] Cox, M. A. A. & Cox, T. F. (2008). *Multidimensional Scaling.* Berlin, Heidelberg: Springer Berlin Heidelberg.

[42] Craig, K. J. & Roux, W. J. (2008). On the investigation of shell buckling due to random geometrical imperfections implemented using Karhunen Loéve expansions. *International Journal for Numerical Methods in Engineering*, 73(12), 1715–1726.

[43] Crisfield, M. (1981). A fast incremental/iterative solution procedure that handles "snap-through". *Computers & Structures*, 13(1), 55 – 62.

[44] de Borst, R., Crisfield, M. A., Remmers, J. J., & Verhoosel, C. (2012). *Non-linear Finite Elemet Analysis of Solids and Structures.* Wiley.

[45] de Vahl Davis, G. & Mallinson, G. (1976). An evaluation of upwind and central difference approximations by a study of recirculating flow. *Computers and Fluids*, 4(1), 29 – 43.

[46] Debusschere, B. J., Najm, H. N., Pébay, P. P., Knio, O. M., Ghanem, R. G., & Maitre, O. P. L. (2004). Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM Journal on Scientific Computing*, 26(2), 698–719.

[47] Donoho, D. L. & Grimes, C. (2003). Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, 100(10), 5591–5596.

[48] Elishakoff, I. (2000). Uncertain buckling: Its past, present and future. *International Journal of Solids and Structures*, 37(46), 6869–6889.

[49] Elman, H. C. & Silvester, D. J. (2005). *Finite Elements and Fast Iterative Solvers : With Applications in Incompressible Fluid Dynamics.* Oxford University Press.

[50] Erban, R., Frewen, T. A., Wang, X., Elston, T. C., Coifman, R., Nadler, B., & Kevrekidis, I. G. (2007). Variable-free exploration of stochastic models: A gene regulatory network example. *The Journal of Chemical Physics*, 126(15), 155103.

[51] Fang, K. & Wang, Y. (1994). *Number theoretical methods in statistics.* Chapman.

[52] Faravelli, L. (1989). Response-surface approach for reliability analysis. *Journal of Engineering Mechanics*, 115(12), 2763–2781.

[53] Farhat, C., Chapman, T., & Avery, P. (2015). Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering*, 102(5), 1077–1110.

[54] Fowlkes, C., Belongie, S., Chung, F., & Malik, J. (2004). Spectral grouping using the nystrom method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2), 214–225.

[55] Fowlkes, C., Belongie, S., Chung, F., & Malik, J. (2004). Spectral grouping using the nyström method. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2), 214–225.

[56] Ghanem, R. (1999). Ingredients for a general purpose stochastic finite elements implementation. *Computer Methods in Applied Mechanics and Engineering*, 168(1), 19 – 34.

[57] Ghanem, R. & Spanos, P. (1990a). Polynomial chaos in stochastic finite elements. *Journal of Applied Mechanics, Transactions ASME*, 57(1), 197–202.

[58] Ghanem, R. & Spanos, P. (1990b). Polynomial chaos in stochastic finite elements. *Journal of Applied Mechanics, Transactions ASME*, 57(1), 197–202.

[59] Ghanem, R. & Spanos, P. (1991a). *Stochastic Finite Elements: A Spectral Approach.* Springer-Verlag.

[60] Ghanem, R. G. & Spanos, P. D. (1991b). Spectral stochastic finite element formulation for reliability analysis. *Journal of Engineering Mechanics*, 117(10), 2351–2372.

[61] Giovanis, D. & Shields, M. (2018). Uncertainty quantification for complex systems with very high dimensional response using grassmann manifold variations. *Journal of Computational Physics*, 364, 393 – 415.

[62] Greig, G. L. (1992). An assessment of high-order bounds for structural reliability. *Structural Safety*, 11(3–4), 213 – 225.

[63] Grigoriu, M. (2009). Existence and construction of translation models for stationary non-gaussian processes. *Probabilistic Engineering Mechanics*, 24, 545–551.

[64] Guo, A., Xu, Y., & Wu, B. (2002). Seismic reliability analysis of hysteretic structure with viscoelastic dampers. *Engineering Structures*, 24(3), 373 – 383.

[65] Hacking, I. (2006). *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference.* Cambridge University Press.

[66] Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation.* Prentice Hall, 2 edition.

[67] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *IEEE conference on Computer Vision and Pattern Recognition*, (pp. 770–778).

[68] Hesthaven, J. S. & Warburton, T. (2007). *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications. (1. ed.).* Springer Publishing Company, Incorporated.

[69] Hua, L. & Wang, Y. (1981). *Applications of number theory to numerical analysis.* Springer-Verlag.

[70] Hughes, T. J., Mallet, M., & Akira, M. (1986). A new finite element formulation for computational fluid dynamics: II. Beyond SUPG. *Computer Methods in Applied Mechanics and Engineering*, 54(3), 341 – 355.

[71] Hurtado, J. (2001). Neural networks in stochastic mechanics. *Archives of Computational methods in Engineering*, 8, 303–342.

[72] Hurtado, J. & Alvarez, D. (2001). Neural network-based reliability analysis: A comparative study. *Computer methods in Applied Mechanics and Engineering*, 191, 113–132.

[73] Jaynes, E. (2003). *Probability Theory: The Logic of Science.* Cambridge University Press.

[74] Ji, S., Yang, M., & Yu, K. (2012). 3d convolutional neural networks for human action recognition. *IEEE transactions on Pattern Analysis and Machine Intelligence.*

[75] Jolliffe, I. T. (1986). *Principal Component Analysis and Factor Analysis*, (pp. 115–128). Springer New York: New York, NY.

[76] Kala, Z. (2005). Sensitivity analysis of the stability problems of thin-walled structures. *Journal of Constructional Steel Research*, 61(3), 415 – 422.

[77] Kala, Z., Valeš, J., & Jönsson, J. (2017). Random fields of initial out of straightness leading to column buckling. *Journal of Civil Engineering and Management*, 23(7), 902–913.

[78] Kalogeris, I. & Papadopoulos, V. (2018). Limit analysis of stochastic structures in the framework of the probability density evolution method. *Engineering Structures*, 160, 304 – 313.

[79] Kalogeris, I. & Papadopoulos, V. (2019). Diffusion maps-based surrogate modeling: An alternative machine learning approach. *International Journal for Numerical Methods in Engineering*, n/a(n/a).

[80] Karpathy, A., Toderici, G., Shetty, S., Leung, T., & Sukthankar, L. (2014). Deep residual learning for image recognition. *IEEE conference on Computer Vision and Pattern Recognition*, (pp. 1725–1732).

[81] Kleene, S. C. (1956). Representation of events in nerve nets and finite automata. In C. Shannon & J. McCarthy (Eds.), *Automata Studies* (pp. 3–41). Princeton, NJ: Princeton University Press.

[82] Kolmogoroff, A. (1931). Über die analytischen methoden in der wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104(1), 415–458.

[83] Koutsourelakis, P., Pradlwarter, H., & Schuëller, G. (2004). Reliability of structures in high dimensions, part I: algorithms and applications. *Probabilistic Engineering Mechanics*, 19(4), 409 – 417.

[84] Krenk, S. (2009). *Non-linear Modeling and Analysis of Solids and Structures*. Cambridge University Press.

[85] Lafon, S., Keller, Y., & Coifman, R. R. (2006). Data fusion and multicue data matching by diffusion maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11), 1784–1797.

[86] LeVeque, R. J. (1992). *Numerical methods for conservation laws (2. ed.)*. Birkhäuser.

[87] LeVeque, R. J. (2002). *Finite volume methods for hyperbolic problems*. Cambridge texts in applied mathematics. Cambridge, New York: Cambridge University Press.

[88] Li, C. & Der Kiureghian, A. (1993). Optimal discretization of random fields. *Journal of Engineering Mechanics*, 119(6), 1136–1154.

[89] Li, J. (2016). Probability density evolution method: Background, significance and recent developments. *Probabilistic Engineering Mechanics*, 44, 111 – 117. Special Issue Based on Papers Presented at the 7th International Conference on Computational Stochastic Mechanics (CSM7).

[90] Li, J. & Chen, J. (2008). The principle of preservation of probability and the generalized density evolution equation. *Structural Safety*, 30(1), 65 – 77.

[91] Li, J., Chen, J., Sun, W., & Peng, Y. (2012). Advances of the probability density evolution method for nonlinear stochastic systems. *Probabilistic Engineering Mechanics*, 28, 132 – 142. Computational Stochastic Mechanics — CSM6.

[92] Li, J. & Chen, J. B. (2004). Probability density evolution method for dynamic response analysis of structures with uncertain parameters. *Computational Mechanics*, 34(5), 400–409.

[93] Li, J. & Chen, J.-B. (2006). The probability density evolution method for dynamic response analysis of non-linear stochastic structures. *International Journal for Numerical Methods in Engineering*, 65(6), 882–903.

[94] Li, J. & Chen, J. B. (2007a). The number theoretical method in response analysis of nonlinear stochastic structures. *Computational Mechanics*, 39(6), 693–708.

[95] Li, J. & Chen, J.-b. (2007b). The number theoretical method in response analysis of nonlinear stochastic structures. *Computational Mechanics*, 39(6), 693–708.

[96] Lignos, D. & Krawinkler, H. (2011). Deterioration modeling of steel components in support of collapse prediction of steel moment frames under earthquake loading. *Journal of Structural Engineering*, 137(11), 1291–1302.

[97] Maier, G. (1966). Quadratic programming and theory of elastic-perfectly plastic structures. *Meccanica*, 3(4), 265–273.

[98] Maier, G., Giannessi, F., & Nappi, A. (1982). Indirect identification of yield limits by mathematical programming. *Engineering Structures*, 4(2), 86 – 98.

[99] Masters, F. & Gurley, K. (2003). Non-gaussian simulation: cumulative distribution function map-based spectral correlation. *Engineering Mechanics*, 129, 1418–1428.

[100] Meek, J. & Tan, H. S. (1984). Geometrically nonlinear analysis of space frames by an incremental iterative technique. *Computer Methods in Applied Mechanics and Engineering*, 47(3), 261 – 282.

[101] Melcher, J., Kala, Z., Holický, M., Fajkus, M., & Rozlívka, L. (2004). Design characteristics of structural steels based on statistical analysis of metallurgical products. *Journal of Constructional Steel Research*, 60(3), 795 – 808.

[MSolve] MSolve. Open source numerical solver for computational mechanics problems. *https://github.com/mgroupntua*.

[103] Nyström, E. J. (1930). Über die praktische auflösung von integralgleichungen mit anwendungen auf randwertaufgaben. *Acta Math.*, 54, 185–204.

[104] Papadopoulos, V. & Iglesis, P. (2007). The effect of non-uniformity of axial loading on the buckling behaviour of shells with random imperfections. *International Journal of Solids and Structures*, 44(18-19), 6299–6317.

[105] Papadopoulos, V. & Kalogeris, I. (2016). A galerkin-based formulation of the probability density evolution method for general stochastic finite element systems. *Comput. Mech.*, 57(5), 701–716.

[106] Papadopoulos, V., Kalogeris, I., & Giovanis, D. G. (2019). A spectral stochastic formulation for nonlinear framed structures. *Probabilistic Engineering Mechanics*, 55, 90 – 101.

[107] Papadopoulos, V. & Papadrakakis, M. (2004). Finite element analysis of cylindrical panels with random initial imperfections. *Journal of Engineering Mechanics*, 130(8), 867–876.

[108] Papadopoulos, V. & Papadrakakis, M. (2005a). 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), 1405 – 1426.

[109] Papadopoulos, V. & Papadrakakis, M. (2005b). 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), 1405 – 1426.

[110] Papadopoulos, V., Soimiris, G., & Papadrakakis, M. (2013). Buckling analysis of I-section portal frames with stochastic imperfections. *Engineering Structures*, 47, 54 – 66.

[111] Papadrakakis, M. & Kotsopoulos, A. (1999). Parallel solutions methods for stochastic fea using monte carlo simulation. *Comp. Meth. Appl. Mech. Eng.*, 168, 305 – 320.

[112] Papadrakakis, M. & Lagaros, N. D. (2002). Reliability-based structural optimization using neural networks and monte carlo simulation. *Computer Methods in Applied Mechanics and Engineering*, 191(32), 3491 – 3507.

[113] Papadrakakis, M. & Papadopoulos, V. (1996). Robust and efficient solution techniques for the stochastic finite element analysis of space frames. *Comp. Meth. Appl. Mech. Eng.*, 134, 627 – 658.

[114] Papadrakakis, M., Papadopoulos, V., & Lagaros, N. D. (1996). Structural reliability analyis of elastic-plastic structures using neural networks and monte carlo simulation. *Computer Methods in Applied Mechanics and Engineering*, 136(1), 145 – 163.

[115] Papoulis, A. & P., U. (2002). *Probability, Random Variables and Stochastic Processes 4th Edition*. McGraw-Hill.

[116] Pawlowski, R. P., Phipps, E. T., Salinger, A. G., Owen, S. J., Siefert, C. M., & Staten, M. L. (2012). Automating embedded analysis capabilities and managing software complexity in multiphysics simulation, part ii: Application to partial differential equations. *Sci. Program.*, 20(3), 327–345.

[117] Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2, 559–572.

[118] Peng, Y., Chen, J., & Li, J. (2014). Nonlinear response of structures subjected to stochastic excitations via probability density evolution method. *Advances in Structural Engineering*, 17(6), 801–816.

[119] Powell, M. J. D. (1987). Algorithms for approximation. (pp. 143–167).

[120] Pradlwarter, H., Schuëller, G., & Dorka, U. (1998). Reliability of MDOF-systems with hysteretic devices. *Engineering Structures*, 20(8), 685 – 691.

[121] Qiu, J., Khoo, B. C., & Shu, C.-W. (2006). A numerical study for the performance of the runge–kutta discontinuous galerkin method based on different numerical fluxes. *Journal of Computational Physics*, 212(2), 540 – 565.

[122] Raithby, G. (1976). A critical evaluation of upstream differencing applied to problems involving fluid flow. *Computer Methods in Applied Mechanics and Engineering*, 9(1), 75 – 103.

[123] Rasmussen, C. E. (2004). *Gaussian Processes in Machine Learning*, (pp. 63–71). Springer Berlin Heidelberg.

[124] Rathinam, M. & Petzold, L. R. (2003). A new look at proper orthogonal decomposition. *SIAM Journal on Numerical Analysis*, 41(5), 1893–1925.

[125] Reddy, J. N. (2018). *Introduction to the finite element method.* McGraw-Hill Education, 4 edition.

[126] Reed, W. & Hill, T. (1973). Triangular mesh methods for the neutron transport equation. *Tech. Report, LA-UR-73-479, Los Alamos Scientific Laboratory.*

[127] Roache, P. J. (1972). On artificial viscosity. *Journal of Computational Physics*, 10(2), 169 – 184.

[128] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.

[129] Rosic, B. V. & Matthies, H. G. (2011). Stochastic galerkin method for the elastoplasticity problem with uncertain parameters. *Recent Developments and Innovative Applications in Computational Mechanics*, (pp. 303–310).

[130] Ross, S. (2012). *A First Course in Probability (9th Edition)*. Pearson.

[131] Roweis, S. T. & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2323–2326.

[132] Schafer, B., Grigoriu, M., & Peköz, T. (1998). A probabilistic examination of the ultimate strength of cold-formed steel elements. *Thin-Walled Structures*, 31(4), 271–288.

[133] Schenk, C. & Schuëller, G. (2003). Buckling analysis of cylindrical shells with random geometric imperfections. *International Journal of Non-Linear Mechanics*, 38(7), 1119 – 1132.

[134] Schillinger, D., Papadopoulos, V., Bischoff, M., & Papadrakakis, M. (2010). Buckling analysis of imperfect I-section beam-columns with stochastic shell finite elements. *Computational Mechanics*, 46(3), 495–510.

[135] Schölkopf, B., Smola, A., & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5), 1299–1319.

[136] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85 – 117.

[137] Scholkopf, B., Smola, A., & Muller, K. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5), 1299–1319.

[138] Schuëller, G. (2006). Developments in stochastic structural mechanics. *Archive of Applied Mechanics*, 75(10-12), 755–773.

[139] Schweizerhof, K. & Wriggers, P. (1986). Consistent linearization for path following methods in nonlinear fe analysis. *Computer Methods in Applied Mechanics and Engineering*, 59(3), 261 – 279.

[140] Sett, K., Jeremić, B., & Kavvas, M. L. (2011). Stochastic elastic–plastic finite elements. *Computer Methods in Applied Mechanics and Engineering*, 200(9), 997 – 1007.

[141] Shields, M., Deodatis, R., & Bocchini, P. (2011). A simple and efficient methodology to approximate a general non-gaussian stationary stochastic process by a translation process. *Probabilitstic Engineering Mechanics*, 26, 511–519.

[142] Shinozuka, M. & Deodatis, G. (1991). Simulation of stochastic processes by spectral representation. *Applied Mechanics Reviews*, 44(4), 191–204.

[143] Shinozuka, M. & Jan, C. M. (1972). Digital simulation of random processes and its applications. *Journal of Sound and Vibration*, 25(1), 111 – 128.

[144] Singer, A., Erban, R., Kevrekidis, I. G., & Coifman, R. R. (2009). Detecting intrinsic slow variables in stochastic dynamical systems by anisotropic diffusion maps. *Proceedings of the National Academy of Sciences*, 106(38), 16090–16095.

[145] Soize, C. & Ghanem, R. (2016). Data-driven probability concentration and sampling on manifold. *Journal of Computational Physics*, 321, 242 – 258.

[146] Sonday, B., Singer, A., Gear, C., & G. Kevrekidis, I. (2010). Manifold learning techniques and model reduction applied to dissipative pdes.

[147] Spanos, P. D., Beer, M., & Red-Horse, J. (2007). Karhunen-Loéve expansion of stochastic processes with a modified exponential covariance kernel. *Journal of Engineering Mechanics*, 133(7), 773–779.

[148] Spiliopoulos, K. V. & Patsios, T. N. (2010). An efficient mathematical programming method for the elastoplastic analysis of frames. *Engineering Structures*, 32(5), 1199 – 1214.

[149] Stavroulakis, G., Giovanis, D. G., Papadrakakis, M., & Papadopoulos, V. (2014). A new perspective on the solution of uncertainty quantification and reliability analysis of large-scale problems. *Computer Methods in Applied Mechanics and Engineering*, 276(0), 627 – 658.

[150] Stefanou, G. (2009). The stochastic finite element method: Past, present and future. *Computer Methods in Applied Mechanics and Engineering*, 198(9-12), 1031–1051.

[151] Sudret, B. & Kiureghian, A. D. (2000). Stochastic finite element methods and reliability: A state-of-the-art report.

[152] Tenenbaum, J., de Silva, V., & Langford, J. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500), 2319–2323.

[153] Thoft-Christensen, P. (1990). Consequence modified β-unzipping of plastic structures. *Structural Safety*, 7(2–4), 191 – 198.

[154] Tin-Loi, F. & Wong, M. (1989). Nonholonomic computer analysis of elastoplastic frames. *Computer Methods in Applied Mechanics and Engineering*, 72(3), 351 – 364.

[155] Venturi, D. & Karniadakis, G. (2012). New evolution equations for the joint response-excitation probability density function of stochastic solutions to first-order nonlinear pdes. *Journal of Computational Physics*, 231(21), 7450–7474.

[156] Venturi, D., Tartakovsky, D., Tartakovsky, A., & Karniadakis, G. (2013). Exact pdf equations and closure approximations for advective-reactive transport. *Journal of Computational Physics*, 243, 323–343.

[157] Victoir, V. (2004). Asymmetric cubature formulae with few points in high dimension for symmetric measures. *SIAM J. Numer. Anal*, 42, 209 – 227.

[158] Wang, P. & Tartakovsky, D. M. (2012). Uncertainty quantification in kinematic-wave models. *Journal of Computational Physics*, 231(23), 7868 – 7880.

[159] Williams, C. K. I. & Seeger, M. (2001). Using the nyström method to speed up kernel machines. In T. K. Leen, T. G. Dietterich, & V. Tresp (Eds.), *Advances in Neural Information Processing Systems 13* (pp. 682–688). MIT Press.

[160] Xiu, D. & Karniadakis, G. E. (2002a). The wiener–askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2), 619–644.

[161] Xiu, D. & Karniadakis, G. E. (2002b). The wiener–askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2), 619–644.

[162] Xu, J., Chen, J., & Li, J. (2012). Probability density evolution analysis of engineering structures via cubature points. *Computational Mechanics*, 50(1), 135–156.

[163] Xu, J. & Feng, D.-C. (2018). Seismic response analysis of nonlinear structures with uncertain parameters under stochastic ground motions. *Soil Dynamics and Earthquake Engineering*, 111, 149 – 159.

[164] Zerva, A. (1992). Seismic ground motion simulations from a class of spatial variability models. *Earthquake Engineering & Structural Dynamics*, 21(4), 351–361.

[165] Zhang, Z. & Zha, H. (2004). Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM Journal on Scientific Computing*, 26(1), 313–338.

[166] Zienkiewicz, O. C., Taylor, R. L., & Fox, D. (2014). *The Finite Element Method for Solid and Structural Mechanics*. Butterworth-Heinemann, 7 edition.