Optimal design of stochastic complex systems with applications to deep excavations

Liu, Wangsheng

2019

Liu, W. (2019). Optimal design of stochastic complex systems with applications to deep excavations. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/106034>

<https://doi.org/10.32657/10220/48126>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**OPTIMAL DESIGN OF STOCHASTIC COMPLEX
SYSTEMS WITH APPLICATIONS TO DEEP
EXCAVATIONS**

LIU WANGSHENG

SCHOOL OF CIVIL & ENVIRONMENTAL ENGINEERING

2018

**OPTIMAL DESIGN OF STOCHASTIC COMPLEX
SYSTEMS WITH APPLICATIONS TO DEEP
EXCAVATIONS**

LIU WANGSHENG

SCHOOL OF CIVIL & ENVIRONMENTAL ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

2018

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

03/April/2019

.....

Date

.....

LIU WANGSHENG

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

03/April/2019

.....
Date

.....
Dr. CHEUNG SAI HUNG

Authorship Attribution Statement

This thesis contains material from one paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 3 is published as Liu Wang-Sheng, and Sai Hung Cheung. "Reliability based design optimization with approximate failure probability function in partitioned design space." *Reliability Engineering & System Safety* 167 (2017): 602-611. DOI: 10.1016/j.ress.2017.07.007.

The contributions of the co-authors are as follows:

- A/Prof Cheung Sai Hung provided the initial project direction and edited the manuscript drafts.
- I prepared the manuscript drafts. The manuscript was revised by A/Prof Cheung Sai Hung.

03/April/2019

.....
Date

.....

LIU WANGSHENG

Acknowledgements

First, I would like to express my thanks to my supervisor and mentor Dr. Cheung Sai Hung. He incited my original interest in the design optimization of stochastic complex systems and has given strong support along this amazing journey of PhD research. I have enjoyed every meeting with him which has always been aspiring and productive. His insightful guidance has helped mature many ideas presented in this work.

I would like to acknowledge my colleagues in our research group for their support in my research. They include Sahil Bansal, Shao Zhe, Fan Ming, Queen Suraajini Rajendran and Hao Changyu. I am also grateful to other supporting staff at CEE, NTU. Without you, this thesis would never be possible.

Special thanks to the reviewers and the members of the thesis committee for their helpful comments and suggestions.

And finally, last but by no means least, I would like to thank my wife Cao Wenjun, my parents and sister for their unconditional support over the years. You make me realize how lucky I am. I dedicate this milestone to you. Now it's time to set off to new adventures. Onwards.

Summary

Modern geotechnical systems, like any other civil engineering systems, tend to have a remarkable increase in their complexity. Also, uncertainties arise due to incomplete knowledge about the system and its environment. Addressing explicitly these uncertainties for a complex system, which may involve high nonlinearity and/or a large number of uncertain parameters, can be very challenging. Particularly, in the optimal design of deep excavations, one should take into account the sophisticated soil mechanical behaviors, the complicated soil-structure interaction, and spatial variability of soil properties.

Reliability-based design optimization (RBDO) aims to model the safety-under-uncertainty aspect of design optimization by introducing reliability constraints. Among many RBDO approaches, the stochastic-sampling-based ones provide the flexibility in the choice of stochastic sampling techniques and the complexity level of numerical models. However, the required computational efforts become prohibitively large especially when expensive-to-evaluate numerical physical models are involved.

In this contribution, an efficient decoupling framework for RBDO called Design Optimization in the PARTitioned Design Space (DOPADS) has been proposed. DOPADS decouples reliability constraints from the optimization loop by approximating failure probability functions (FPF) in the partitioned design space. A sufficient number of failure samples are generated using Markov chain Monte Carlo methods in each subspace, ensuring the accuracy of the approximation over the entire design space. The major gain in efficiency comes from the decoupling that explores the design space in a single reliability analysis instead of performing the reliability analyses repeatedly as in nested RBDO approaches. In the vanilla version of DOPADS, an FPF is approximated for each reliability constraint separately, which makes its computational complexity linear to the number of reliability

constraints. Therefore, another framework DOPADS_m has been proposed which is tailored to the cases with multiple reliability constraints. By taking advantage of the dependency between different failure events, DOPADS_m largely enhances the efficiency of DOPADS.

Design optimization is an ongoing process in which the initial design may need to be modified based on model updating using observed data. In the traditional reliability updating methods, one needs to update the distribution of model parameters in a Bayesian approach, then calculate the failure probability based on the set of the updated models. This process takes a long time which may lead to lagged decision making regarding the current reliability state. Thus, a new framework INstant REliability Updating (INREU) for reliability updating has been presented. The updated failure probability function is constructed before collecting data, making the instant plug-in query of the updated reliability possible.

The application of the proposed frameworks to the deep excavation cases in Singapore has been investigated. The spatial variability of the properties of the Singapore marine clay is modeled by the random field theory. Results show that the variability of soil properties plays an important role in the design of deep excavations.

Table of contents

Acknowledgements.....	I
Summary	III
Table of contents.....	V
List of figures.....	XI
List of tables.....	XV
Notation	XVII
Definitions.....	XXI
1 Introduction.....	1
1.1 Background	1
1.2 Scope.....	2
1.3 Objectives	3
2 Literature review	5
2.1 Modeling of stochastic complex systems.....	5
2.1.1 Characteristics of up-to-date models.....	5
2.1.2 Uncertainty representation	6
2.1.3 Sources of uncertainties	7
2.2 Uncertainty-based analysis and Monte Carlo Simulation.....	9
2.2.1 Crude Monte Carlo Simulation	10
2.2.2 Importance sampling.....	12

2.2.3	Accept-reject method.....	13
2.2.4	Markov chain Monte Carlo methods	14
2.3	Reliability analysis	17
2.3.1	Overview	17
2.3.2	First-order and second-order reliability methods	18
2.3.3	Subset simulation.....	19
2.4	Reliability-based design optimization	20
2.4.1	Overview	20
2.4.2	Existing approaches.....	21
2.4.3	Applications to geotechnical problems.....	25
2.5	Conclusions	26
3	Approximate failure probability function in the partitioned design space	29
3.1	Introduction	29
3.2	The proposed framework.....	30
3.2.1	Formulation of failure probability function.....	30
3.2.2	Predefined joint probability distribution $\mathbf{p}\Phi\boldsymbol{\varphi}$	32
3.2.3	Reliability analysis of the augmented system and failure samples.....	32
3.3	An iterative scheme for density estimation	32
3.3.1	Partitioning the design space	33
3.3.2	Generation of additional failure samples.....	35
3.3.3	Density estimation in each iteration	37
3.4	Implementation.....	40
3.4.1	Flowchart.....	40
3.4.2	Sample size in each iteration	41

3.4.3	Partition density values	42
3.4.4	Stopping criterion.....	42
3.5	Illustrative examples	42
3.5.1	Example 1: A straight column with two design variables.....	42
3.5.2	Example 2: A ten-bar truss with three design variables and multiple reliability constraints.....	48
3.5.3	Example 3: A ten-story shear building with a large number of random variables	52
3.6	Conclusions.....	55
4	Enhancing the efficiency of design optimization for stochastic complex systems with multiple reliability constraints	57
4.1	Introduction.....	57
4.2	Propagating failure samples for different failure events simultaneously	59
4.2.1	Basic idea	59
4.2.2	Sample size in each iteration.....	61
4.2.3	Partition density values	62
4.2.4	Stopping criterion.....	63
4.3	Summary of implementation.....	63
4.4	Illustrative examples	65
4.4.1	Example 1: A cantilever beam	66
4.4.2	Example 2: a ten-story shear building subjected to ten reliability constraints	73
4.5	Conclusions.....	76
5	Instant reliability updating by constructing the updated failure probability function	79
5.1	Introduction.....	79

5.2	Common Bayesian reliability updating formulations.....	81
5.3	A framework for instant reliability updating.....	82
5.3.1	Updated failure probability function	83
5.3.2	Measurement samples conditional on the failure or non-failure event.....	84
5.3.3	Iterative density estimation of measurement samples	85
5.3.4	Summary of implementation	86
5.4	Illustrative examples.....	87
5.4.1	Example 1: Linear model - An analytical example	87
5.4.2	Example 2: Ten-bar truss with two measurements of strain.....	92
5.5	Conclusions	97
6	Applications to deep excavations considering the spatial variability of soil properties	99
6.1	Modeling of deep excavations.....	99
6.1.1	Traditional models.....	99
6.1.2	Numerical models in commercial codes.....	101
6.1.3	Uncertainty sources	102
6.2	Random field and its discretization	103
6.3	Design optimization of a two-dimensional deep excavation against basal heave	105
6.3.1	Case description.....	105
6.3.2	Results and discussions	107
6.3.3	Conclusions	113
6.4	Update the reliability of an excavation using monitoring data.....	114
7	Concluding remarks and future work	119
7.1	Conclusions	119

7.2	Future works	122
7.2.1	Surrogate for reliability-based design optimization	122
7.2.2	Reliability-based design optimization with discrete design variables	122
7.2.3	Data-adaptive density estimation with confidence bands	123
7.2.4	Multidisciplinary design optimization	123
Appendix A: Proof of stationarity.....		125
Appendix B: Slip circle method for basal heave analysis		127
Bibliography		129
Academic resume.....		137

List of figures

Figure 1-1: Collapse of Nicoll Highway, Singapore.....	1
Figure 1-2: Design of complex systems under uncertainty.....	3
Figure 2-3: Uncertainty propagation in uncertain-based analysis.....	10
Figure 2-4: Diagram of Metropolis-Hastings algorithm in the 2-d space. The dashed lines are the contour levels of the probability density function.....	16
Figure 2-5: Flowchart of double-loop approaches for reliability-based design optimization	22
Figure 2-6: Flowchart of the decoupling approach to reliability-based design optimization based on the approximate failure probability function	24
Figure 3-1: Estimated density function based on failure samples and the failure probability function	31
Figure 3-2: Iterative density estimation in the partitioned 2-dimensional design space....	34
Figure 3-3: Failure probability function by using the smoothed histogram in $\mathcal{Lk} - 1$	38
Figure 3-4: Flowchart of design optimization in the partitioned design space (DOPADS).....	41
Figure 3-5: Schematic diagram of a straight column considering its buckling failure	43
Figure 3-6: Formulation of the column design optimization problem.....	44
Figure 3-7: Partitioning of the design space and partition density values	44
Figure 3-8: Approximated surface of the density function	45
Figure 3-9: Contour plot of the failure probability function	46
Figure 3-10: Profiles of the failure probability function for $b = 100, 250, 400mm$	46
Figure 3-11: Optimal solutions and objective function values for different allowable failure probabilities.....	47
Figure 3-12: Schematic diagram of a ten-bar cantilever truss	48
Figure 3-13: Formulation of the ten-bar truss design optimization problem	49

Figure 3-14: Contour plots of the failure probability functions for the three failure modes ($A2 = 2000mm^2$).....	51
Figure 3-15: Optimal solution and the feasible region for the ten-bar truss example	52
Figure 3-16: Schematic diagram of the ten-story shear building and a sample time history of the ground acceleration	53
Figure 3-17: Formulation of the ten-story shear building optimization problem.....	54
Figure 3-18: Partitioning of the design space and its partition density values	55
Figure 3-19: Contour plot of the failure probability function.....	55
Figure 4-1: Comparison of the approximation of failure probability functions in DOPADS and DOPADS_m	61
Figure 4-2: Flowchart of design optimization in the partitioned design space for problems involving multiple reliability constraints (DOPADS_m). Filled rectangles represent the parts involving the model evaluation which require a large amount of computation time	65
Figure 4-3: Schematic diagram of a cantilever beam.....	66
Figure 4-4: Propagation of failure samples in partitioned design space (Case C). k is the index of iteration and I is the index of reliability constraint. Only 200 out of 3000 samples are plotted in this figure for visualization.....	69
Figure 4-5: Contour plots of the failure probability functions	70
Figure 4-6: Means of the minimal objective values for different Ns	73
Figure 4-7: Schematic diagram of the ten-story shear building subjected to multiple reliability constraints	74
Figure 4-8: Change of the minimal objective values with the increase of Ns	75
Figure 5-1: Comparison of the updating processes. Panel (a): Conventional reliability updating has a large time lag between the measurement acquisition and reliability updating; and Panel (b): The proposed framework has a negligible time lag. The length of arrows represents the length of time lags but not in scale and the shaded areas represent the most time-consuming part.....	83
Figure 5-2: Sample generation and propagation in the k -th iteration	86
Figure 5-3: Flowchart of instant reliability updating.....	87

Figure 5-4: Iterative density estimation of failure and non-failure samples for the linear model	90
Figure 5-5: Support points and mean smoothed density functions based on GP (C.I. is short for Credible Interval)	90
Figure 5-6: Comparison of the UFPF from INREU with the ground truth.....	91
Figure 5-7: Change of the global evaluation metric <i>MISE</i> with the increase of sample size	92
Figure 5-8: Change of the pointwise evaluation metric <i>MSE</i> with the increase of sample size	92
Figure 5-9: Schematic diagram of the ten-bar truss and parameters of the model (S.D. is short for Standard Deviation).....	93
Figure 5-10: Iterative density estimation of failure and non-failure samples for the ten-bar truss	94
Figure 5-11: Comparison of the UFPF surface of INREU with the ground true	94
Figure 5-12: Change of the global evaluation metric <i>MISE</i> with the increase of sample size	96
Figure 5-13: Change of the pointwise evaluation metric <i>MSE</i> with the increase of sample size	96
Figure 6-1: Two modes of external instability failure in an excavation (Ou 2006)	100
Figure 6-2: Realizations of random fields with different length of fluctuations.....	104
Figure 6-3: Profile of the braced excavation site (not to scale)	106
Figure 6-4: Formulation of the design optimization problem of a deep excavation	107
Figure 6-5: Approximate failure probability function based on iterative density estimation	108
Figure 6-6: Comparison of the approximate failure probability function in each iteration and that by the crude Monte-Carlo simulation	108
Figure 6-7: Change of the optimal design by (a) DOPADS and (b) double-loop approach	109
Figure 6-8: Number of model evaluations required in DOPADS and the two-loop approach	110

Figure 6-9: Optimal wall depth against the CoV of soil density for different mean values	111
Figure 6-10: Optimal wall depth against the CoV of normalized undrained shear strength for different mean values.....	112
Figure 6-11: Optimal wall depth against the length of fluctuation in x and z directions .	112
Figure 6-12: Optimal wall depth for different allowable failure probabilities	113
Figure 6-13: Profile of the excavation site (not to scale)	114
Figure 6-14: 2D model and its mesh of the excavation site	115
Figure 6-15: Parametric study and the importance of soil properties. (a) δhm and δvm calculated using different soil properties; (b1) the importance of soil properties to δhm ; (b2) the importance of soil properties to δvm . The importance is measured by the Pearson correlation coefficient.....	116
Figure 6-16: Stochastic mesh of a random field and one realization of its stochastic elements	117
Figure 6-17: Updated failure probability function in terms of the measured maximum wall deflection.....	118
Figure B-1: Schematic diagram of slip circle method.....	127

List of tables

Table 2-1: Algorithm for crude Monte Carlo Simulation (non-parallel)	11
Table 2-2: Algorithm for Importance Sampling (non-parallel)	12
Table 2-3: Algorithm for accept-reject method (non-parallel).....	13
Table 2-4: Algorithm for Metropolis-Hastings algorithm	15
Table 2-5: Algorithm for Gibbs sampling.....	16
Table 3-1: Algorithm for modified Metropolis-Hastings algorithm	35
Table 3-2: Ranges of design variables and distribution parameters of random variables..	43
Table 3-3: Optimal solutions and objective function values for different allowable failure probabilities.....	47
Table 3-4: Ranges of design variables and distribution parameters of random variables for a ten-bar cantilever truss	48
Table 3-5: Partition density values and the number of support points for three failure modes	50
Table 4-1: Distribution parameters of random variables for the cantilever beam.....	67
Table 4-2: Optimal solutions and their objective values for three different cases	70
Table 4-3: Number of model evaluations in DOPADS_m and DOPADS (Case A)	71
Table 4-4: Number of model evaluations in DOPADS_m and DOPADS (Case B).....	71
Table 4-5: Number of model evaluations in DOPADS_m and DOPADS (Case C).....	72
Table 4-6: Number of model evaluations in DOPADS_m using early stopping strategy to determine the sample size	76
Table 5-1: Number of model evaluations and samples during the approximation of UFPF	95
Table 6-1: Summary of statistical characteristics of random variables	106
Table 6-2: Summary of statistical characteristics of uncertain soil properties	116

Notation

In general, this thesis uses bold letters to indicate vectors, upper case letters to indicate random variables (their realizations are denoted by lower case letters). The table in this section does not include symbols in Chapter 6.

Latin letters

A	An event
d	Dimension of the parameter space
n_θ	Dimension of the parameter space
n_φ	Dimension of the design space
$\mathbf{y}$	A vector of the model outputs (or responses)
$P(A)$	$P(A) = P_A$ Probability of an event A
P_F	Failure probability or probability of failure
$P_{\bar{F}}$	Probability of the non-failure event $\bar{F}$. The identity $P_{\bar{F}} = 1 - P_F$ holds.
$[P_F]$	Allowable failure probability of a reliability constraint in the reliability-based design optimization problem
T	Transpose of a matrix
$\mathbf{d}^* = [d_1^*, \dots, d_{n_d}^*]^T$	A vector of n_d measurements

Greek letters

Ω	Sample or parameter space
Ω_F	$\Omega_F = \{\boldsymbol{\theta}: g(\boldsymbol{\theta}) \leq 0\} \subseteq \Omega$, the failure region
Ω_d	Domain of the updated failure probability function or measurement space

Θ	$\Theta = [\Theta_1, \dots, \Theta_{n_\theta}]^T$ uncertain parameters modeled by a random variable vector
θ	$\theta = [\theta_1, \dots, \theta_{n_\theta}]^T \in \mathbb{R}^{n_\theta}$, a vector of the parameter values (or realization) of Θ
$\theta^{(i)}$	A sample generated from Θ
Φ	$\Phi = [\varphi_1, \dots, \varphi_{n_\varphi}]^T \in \mathbb{R}^{n_\varphi}$, a vector of design variables
β	Reliability index
$\{(\Phi, \theta)\}_F$	A set of failure samples in the augmented parameter space
ϵ_p	$\epsilon_p \sim p(\epsilon_p \theta)$

Indices

i	The index of sample
j	The index of dimension
k	The index of iteration or level
I	The index of the failure event for problems with multiple failure events
t	The index of bin or rectangle in density estimation

Functions and symbols

$\mathcal{P}_\Theta(\theta)$	Cumulative distribution function (CDF) of a random variable Θ
$p_\Theta(\theta)$	Probability density function (PDF) of a random variable Θ
$p_\Theta(\theta)$	Joint probability density function (PDF) of random variables θ
$p_\Phi(\Phi)$	Joint probability density function (PDF) of design variables Φ that are artificially considered as random variables
$P_F(\Phi)$ or $P(F \Phi)$	Failure probability function with respect to design variables Φ
$p(\Phi F)$ or $p_F(\Phi)$	Probability density function of design variables conditional on an event F
$p(\theta \mathbf{d}^*)$ or $p_{\mathbf{d}^*}(\theta)$	Probability density function of a random variable vector θ conditional on the measurements $\mathbf{d}^*$

$p(\mathbf{d}^* \boldsymbol{\theta})$	likelihood function that describes the distribution of prediction errors between the predicted outputs of a model and the measured outputs $\mathbf{d}^*$
$\hat{p}(\boldsymbol{\varphi} F)$	Estimated probability density function
$\tilde{p}(\boldsymbol{\varphi} F)$	Smoothed $\hat{p}(\boldsymbol{\varphi} F)$
$g(\cdot)$	The performance function: $g(\cdot) \leq 0$ failure event; $g(\cdot) = 0$ limit-state function. $\mathbf{g}(\cdot)$ denotes a set of the performance functions.
$\mathbf{h}(\cdot)$	A set of the model performance assessment functions; a set of predictive functions in INREU
$T(\cdot)$	Isoprobabilistic transformation which maps the original random variables $\boldsymbol{\Theta}$ to independent standard normal random variables $\mathbf{U}$
$\mathbb{I}_F(\cdot)$	Indicator function of the failure event F which equals to one if a system fails and zero otherwise.
$\mathbb{E}[\cdot]$	Expectation operator
$\mathbb{V}[\cdot]$	Variance operator
$\rho(\cdot)$	Autocorrelation function

Acronym

RBDO	Reliability-based design optimization
PDF	Probability density function
CDF	Cumulative distribution function
CoV	Coefficient of Variation
MCS	Monte Carlo simulation
IS	Importance sampling
MCMC	Markov chain Monte Carlo methods
M-H	Metropolis-Hastings algorithm
i.i.d	Independently and identically distributed
LSF	Limit-state function
SS	Subset simulation

FORM/SORM	First-order/Second-order reliability method
MPP	Most probable point
RIA	Reliability index approach
PMA	Performance measure approach
SORA	Sequential Optimization and Reliability Assessment
FPF	Failure probability function
DOPADS	Design optimization in the partitioned design space (the proposed framework in this thesis)
LHS/RHS	Left-hand side or right-hand side of an equation
BSP	Bayesian sequential partitioning
SQP	Sequential quadratic programming
KKT	Karush-Kuhn-Tucker (KKT) conditions
INREU	Instant reliability updating
UFPF	Updated failure probability function
GP	Gaussian processes
LOF	Length of fluctuation

Definitions

Markov chain Monte Carlo:	A class of methods for sampling from a target probability distribution by using Markov chains.
Reliability	The likelihood that a component or a system will perform its intended function without failure for a specified period of time under stated operating conditions
Reliability-based design optimization	A methodology that allows for the search of the optimal design meanwhile satisfying the reliability constraint(s) and deterministic constraint(s).
Density estimation	Construction of an estimate of unobservable underlying probability density function based on samples

1 Introduction

1.1 Background

Design of earth-retaining structures for deep excavations needs a very high level of reliability. Excavation fatalities and injuries, which will lead to serious social and economic consequences, constantly plague the geotechnical community. According to the U.S. Bureau of Labor Statistics (BLS) (BLS 2010), 271 workers died in trenching or excavation cave-ins from 2000 to 2006. Another example is the collapse of the Nicoll Highway in Singapore (see Figure 1-1), which reportedly caused four deaths and millions of dollars in economic losses. According to Committee of Inquiry (2005), it was due to the under-design of the temporary lateral earth support system.

Figure 1-1: Collapse of Nicoll Highway, Singapore

Design optimization aims to find an optimal design configuration such that the cost can be minimized while satisfying multiple requirements such as specifications, safety, serviceability, reliability, etc. Uncertainties due to incomplete information of the system itself as well as its environment must be recognized explicitly in order to achieve a sufficient reliability margin. Reliability-based design optimization (RBDO) provides a rational and

rigorous methodology to optimize the cost when the system is subjected to multiple probabilistic (reliability) constraints.

For complex systems such as deep excavations, it is a formidable task to obtain the optimal design by performing RBDO. The two main challenges are:

- (1) Modern numerical physical models often have thousands of degrees of freedom and the complexity level is still increasing. Despite the high fidelity, it requires a large amount of time to run a single simulation. For RBDO problems, one needs to check whether the reliability constraints are satisfied for each design configuration during the optimization, resulting in tens of millions of simulations.
- (2) Uncertainties exist extensively in the system as well as its environment. For a geotechnical problem, site investigations are conducted to characterize the distribution of uncertain model parameters; however, the scale of site investigations or field tests is limited due to the project budget. Furthermore, to represent the spatial variability of soil properties, the random field theory is often adopted, which leads to a large number of random variables.

1.2 Scope

The whole process of uncertainty-based design can be illustrated in Figure 1-2. It starts with the modeling of the physical system and its environment, in which the associated uncertainties are considered. To obtain the optimal design configuration, it is followed by performing RBDO. For the optimal design, one can obtain the probabilistic performance predictions by propagating uncertainties from the inputs to the outputs; meanwhile, measured data can be collected if monitoring and/or inspections are performed on the built system. Based on the information from the collected data and performance predictions, the initial design of the system can be updated and modified. Furthermore, the probabilistic predictions along with associated consequences result in the risk, based on which further decisions on whether or not to adjust the design can be made. In the scope of this thesis, our focus is on the development of efficient and accurate RBDO frameworks as well as the potential design adjustment if measured data are available as indicated in Figure 1-2.

For mathematical modeling of the physical complex system, it is assumed that numerical models that can be simulated in the state-of-the-art commercial codes. Though analytical models are efficient, their predictive capability is limited.

For uncertainty modeling, the process of uncertainty quantification has been excluded from the scope of this thesis. It is assumed that the distribution of random variables as well as the statistical characteristics of random fields are known a priori.

For uncertainty propagation, it is performed based on the stochastic sampling approach instead of analytical counterparts.

Figure 1-2: Design of complex systems under uncertainty

1.3 Objectives

- (1) Propose an accurate and efficient framework for RBDO based on stochastic sampling. The framework should be able to incorporate different commercial codes in a non-intrusive way and be robust to the number of random variables.
- (2) Propose a framework for fast reliability updating if measured data are available.
- (3) Apply the proposed frameworks to deep excavation systems in Singapore taking into account the spatial variability of soil properties.

2 Literature review

Summary

This chapter starts by introducing the essential components of reliability-based design optimization (RBDO), including the uncertainty modeling, uncertainty-based analysis and reliability methods. In uncertainty-based analysis, the emphasis is put on the uncertainty propagation using stochastic sampling. Next, existing approaches to RBDO are introduced and compared: the double-loop approaches, the single-loop approaches and decoupling approaches. It is also noted that the RBDO design methodology is gaining popularity in the geotechnical community.

2.1 Modeling of stochastic complex systems

2.1.1 Characteristics of up-to-date models

To mimic the behavior of the real-world systems, *mathematical models* can be constructed based on physical laws and/or empirical observations. In general, mathematical models can be divided into two groups: *analytical models* and *numerical models*. In analytical models, the relationships composing the model are simple enough to allow one to derive the *exact* solution of the quantity of interest. It is desirable to study a model in its analytical form, but many models are too complex to have analytical solutions in finite time. Alternatively, numerical models (or *simulations*) can be used to give approximate solutions within allowable tolerance. Stable numerical models are prevalent in practice thanks to their efficiency and sufficient accuracy.

With the development of modern computers and simulation techniques, at least three features can be observed in up-to-date numerical models

- (1) *Availability of commercial codes.* In many disciplines, powerful commercial codes have been a mutual tool for modeling to scientists and engineers. For example, generic finite-element codes such as Abaqus and Ansys have been widely used to imitate sophisticated mechanical behaviors including nonlinearities, plasticity, instabilities, impact dynamics, etc. Therefore, it is advantageous to use the numerical models in commercial codes as a building block of other algorithms or frameworks instead of intruding into the models.
- (2) *Expensive to evaluate.* It is not uncommon that the up-to-date models have thousands of degrees of freedom, thus requiring a large amount of time to run a simulation. Though this issue can be alleviated to a certain extent by the increase of computational capability of modern computers, there is no guarantee that the improved computing power can compensate the ever-increasing complexity level of models. What's worse, many of us may have limited access to computational resources. This feature justifies the necessity to improve the efficiency from an algorithmic perspective when the problem to solve involves expensive-to-evaluate models.
- (3) *Uncertainties.* It has been well recognized that ignoring the uncertainties can lead to dangerous even catastrophic situations. Generally, up-to-date models can incorporate uncertainties in the input (or the environment of the system) as well as in models themselves, then uncertainties are propagated by model evaluations to quantify the uncertain characteristics of the model outputs.

Realizing these features, this thesis focuses on stochastic complex systems whose numerical models can be evaluated by commercial codes.

2.1.2 Uncertainty representation

Many theories have been developed to describe the characteristics and information of uncertainties. Most widespread uncertainty representation theories include: *probability theory*, *interval analysis* (Rao and Berke 1997, Ferson et al. 2007), *evidence theory* (Dempster 1967, Shafer 1976, Helton et al. 2007), and *possibility theory* (Zadeh 1999). Readers can refer to two comprehensive literature reviews on uncertainty representation

(Oberkamp et al. 2001, Helton et al. 2004). In the scope of this thesis, the proposed methods are established on the basis of probability theory due to its sound theoretical foundation.

Probability $P(A)$ is a single-valued measure used to represent the uncertainty about the occurrence of the event A . Two different interpretations of probability exist: frequentist (objective) probability and Bayesian (subjective) probability. From the perspective of a frequentist, the probability of the event A is the empirical frequency if the random experiment is repeated independently for an infinite number of times. It is not practical to repeat the experiment endlessly to obtain the true value of $P(A)$. On the contrary, the Bayesian believes that probability quantifies a subjective degree of belief on a given event.

Despite different interpretations, these two kinds of probability are compatible to each other and share the same notations. Let $(\Omega, \mathcal{F}, P)$ be a probability space where Ω denotes a sample space with σ -algebra $\mathcal{F}$ and P is a probability measure. Usually, a random variable $\Theta: \Omega \rightarrow \mathbb{R}$ maps the sample space Ω to the real-valued space $\mathbb{R}$. A random variable Θ can be completely defined by its associated *cumulative distribution function* (CDF) $\mathcal{P}_\Theta(\theta)$

$$\mathcal{P}_\Theta(\theta) = P(\Theta \leq \theta) \quad (2-1)$$

where θ denotes a realization of Θ . For the continuous case, a *probability density function* (PDF) exists

$$p_\Theta(\theta) = \lim_{h \rightarrow 0, h > 0} \frac{P(\theta \leq \Theta \leq \theta + h)}{h} \quad (2-2)$$

In general, uncertain parameters in a mathematical model are treated as a vector of n_θ random variables $\boldsymbol{\Theta} = [\theta_1, \dots, \theta_{n_\theta}]^T$ with a predefined joint PDF $p_{\boldsymbol{\Theta}}(\boldsymbol{\Theta})$. In the context of engineering, the joint PDF can be determined based on the domain knowledge, engineering judgement, or density estimation techniques (Parzen 1962, Silverman 1986, Scott 2015).

2.1.3 Sources of uncertainties

In this thesis, two types of uncertainties are distinguished: *aleatory* and *epistemic* (Der Kiureghian and Ditlevsen 2009). Aleatory uncertainty deals with the inherent variation of a physical system and its environment. Since this type of uncertainty cannot be reduced by

collecting additional information or increasing the knowledge, it is also known as *irreducible uncertainty*. On the contrary, epistemic uncertainty (subjective uncertainty) is attributed to a lack of knowledge about the current system and environment, which can be reduced in means of acquiring the new knowledge. In the modeling of stochastic complex systems, epistemic uncertainty may manifest itself in terms of modeling uncertainty, parameter uncertainty and environment uncertainty. Both aleatory and epistemic uncertainties can be represented by the probability distributions. Among all uncertainty taxonomies, the taxonomy based on the aleatory and epistemic uncertainties is arguably the most widely used one.

There have been many studies attempting to distinguish different sources of uncertainties but to date there is little agreement (Walker et al. 2003, Der Kiureghian and Ditlevsen 2009, Soize 2010). In this thesis, they are summarized in the following list. This list is by no means exhaustive but can be used as a guidance to identify uncertainties and has been used throughout the thesis.

- *Model input uncertainty*. It is also known as external or environment uncertainty. This group of uncertainties is caused by the random acting from the environment on the system under consideration. Take the geotechnical system as an example. The uncertainties in the soil mass can be considered as the model input uncertainty for an underground structure. Sometimes, the soil mass and structure (e.g., earth-retaining walls) are modeled as a whole system by taking soil-structure interaction into consideration. In this case, the uncertainties in the soil mass can be treated as the internal ones.
- *Modeling uncertainty*. When it comes to the modeling of a stochastic complex system, assumptions and simplifications are inevitable, which give rise to the modeling uncertainty. For a geotechnical system, typical assumptions and simplifications occur in boundary conditions, model dimensionality, constitutive model of soil mass. It should be noted that the computational errors introduced during the solving of a model (e.g., discretization error in a finite-element model or approximation error in the statistical estimation by using sampling-based methods) are also abusively interpreted as modeling uncertainty.

- *Model parameter uncertainty.* It is related to the limited knowledge about the parameters (e.g., material properties and geometric properties) in a given mathematical model. The level of model parameter uncertainty changes in different phases as new information can be acquired. As a contrast to external uncertainty, modeling uncertainty together with model parameter uncertainty sometimes may be referred to as internal uncertainty.
- *Manufacturing and construction uncertainty.* Uncertainties may also arise from the human errors during the manufacturing or construction phase.

2.2 Uncertainty-based analysis and Monte Carlo Simulation

Uncertainty-based analysis refers to the process in which the uncertainties in the environment and the primary model itself are propagated through the model itself, causing the variability in the model outputs; then the performance of the model will be evaluated based on the characteristics of the uncertain outputs.

The detailed explanation of uncertainty-based analysis is shown in Figure 2-3. Let $\mathcal{M}$ be a model class that consists of the environment (or input) model $\mathbf{E}(\boldsymbol{\theta}_{in})$, the primary model $\mathbf{g}(\boldsymbol{\varphi}, \mathbf{q}, \boldsymbol{\theta}_s)$ and the performance assessment model $\mathbf{h}(\mathbf{y}, \boldsymbol{\theta}_{eval})$. The design variable vector $\boldsymbol{\varphi}$ are the controllable parameters in the primary model which can be adjusted in the design phase. The output vector $\mathbf{q}$ of $\mathbf{E}(\cdot)$ is treated as the input of the primary model; similarly, the output vector $\mathbf{y}$ of $\mathbf{g}(\cdot)$ is used to calculate the model performance in $\mathbf{h}(\cdot)$. Here $\mathcal{M}$ represents a model class because each realization of uncertain parameters $\boldsymbol{\theta} = [\boldsymbol{\theta}_{in}; \boldsymbol{\theta}_s; \boldsymbol{\theta}_{eval}]$ in $\mathcal{M}$ can provide a mapping from the parameter space to the output space. Another thing to notice is that the models in the model class appears in bold to indicate that the model may yield multiple outputs by evaluating multiple functions.

$$\mathcal{M} = \{\mathbf{E}(\boldsymbol{\theta}_{in}), \mathbf{g}(\boldsymbol{\varphi}, \mathbf{q}, \boldsymbol{\theta}_s), \mathbf{h}(\mathbf{y}, \boldsymbol{\theta}_{eval})\}$$

Figure 2-3: Uncertainty propagation in uncertain-based analysis

It is not unusual that the environment is modeled as part of the primary model and the uncertainty of the parameters θ_{eval} in $h(\cdot)$ can be ignored. In this case, the primary model can be defined as:

$$g(\boldsymbol{\varphi}, \boldsymbol{\theta}): \mathbb{R}^{n_\varphi} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}^{n_y} \quad (2-3)$$

where n_φ , n_θ and n_y denote the dimension of the design space, the random parameter space and the output space, respectively. The performance assessment model $h(\cdot)$ can also be expressed in terms of $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$.

For the sake of simplicity, one can rewrite the performance assessment model by dropping the design variables $\boldsymbol{\varphi}$ and considering one single output

$$z = h(\boldsymbol{\theta}) \quad (2-4)$$

Let us assume the uncertainties in the model are described by probability theory, i.e., $\boldsymbol{\theta}$ follows a joint PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. The goal of uncertainty-based analysis is to obtain statistical information about $Z = h(\boldsymbol{\theta})$, such as the mean value, the probability of an event related to the value z or the probability distribution of Z .

Here it is assumed that the statistical information one wants to know is the expected value of $\mathbf{Z}$, i.e., $\mu = \mathbb{E}[h(\boldsymbol{\theta})]$ where $\mathbb{E}$ is the expectation operator. It is not easy to calculate since it is common that the distribution of $\boldsymbol{\theta}$ is sophisticated and/or $h(\cdot)$ is implicit and expensive to evaluate. Fortunately, *Monte Carlo Simulation* (MCS) (Robert 2004, Asmussen and Glynn 2007, Zio 2013, Rubinstein and Kroese 2016), the most popular stochastic sampling technique, allows one to estimate μ by simulating $h(\cdot)$ with random sampling. As MCS is straightforward, flexible and powerful, its application can be found in almost every field and an extensive amount of research has been focusing on this topic since the early 20th century. In this section, four different MCS methods will be covered: *the crude MCS*, *importance sampling* (IS), *accept-reject method* and *Markov chain Monte Carlo* (MCMC) methods.

2.2.1 Crude Monte Carlo Simulation

In fact, the expected value μ is a multidimensional integral for continuous random variables $\boldsymbol{\theta}$

$$\mu = \mathbb{E}[Z] = \int_{\boldsymbol{\theta} \in \Omega \subset \mathbb{R}^{n_{\boldsymbol{\theta}}}} h(\boldsymbol{\theta}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (2-5)$$

To estimate μ , one can generate N independent samples $\{\boldsymbol{\theta}^{(i)}; i: 1 \rightarrow N\}$ from the distribution $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and compute the following average:

$$\hat{\mu}_{MCS} = \frac{1}{N} \sum_{i=1}^N h(\boldsymbol{\theta}^{(i)}) \quad (2-6)$$

According to the law of large numbers, the estimator $\hat{\mu}_{MCS}$ converges almost surely to the true value μ if μ exists. Also, $\hat{\mu}_{MCS}$ is unbiased and its variance can be estimated by:

$$\text{Var}[\hat{\mu}_{MCS}] = \frac{\sum_{i=1}^N (h(\boldsymbol{\theta}^{(i)}) - \hat{\mu}_{MCS})^2}{N(N-1)} \quad (2-7)$$

The algorithm for the crude MCS to generate samples from Z is summarized in Table 2-1.

Table 2-1: Algorithm for crude Monte Carlo Simulation (non-parallel)

crudeMonteCarlo ($N, p_{\boldsymbol{\theta}}(\boldsymbol{\theta}), h$)	
<hr/>	
Inputs: N : sample size, $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$: joint probability density function, h : model under consideration	
Output: samples $\{z\}$ from Z where $Z = h(\boldsymbol{\theta})$	
1:	$i \leftarrow 0$ and $\{z\} \leftarrow \emptyset$
2:	while $i < N$:
3:	generate a sample $\boldsymbol{\theta}^{(i)}$ from $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$
4:	compute $z^{(i)} = h(\boldsymbol{\theta}^{(i)})$
5:	update $i \leftarrow i + 1$, $\{z\} \leftarrow \{z\} + z^{(i)}$
6:	endwhile
7:	return $\{z\}$

The main advantages of the crude MCS include: (i) It is flexible. This method is robust to the high dimensionality of random variables since the convergence rate of the estimator does not depend on the dimension. Also, the models with different levels of complexity can be incorporated in the crude MCS without intruding into the models themselves, which is an attractive feature to those in favor of commercial codes as the models. (ii) The implementation is very straightforward and easy to understand.

However, one of the biggest issues in the crude MCS is that the convergence rate is very slow, i.e., $\mathcal{O}(N^{-1/2})$. It implies that to achieve one more decimal digit of accuracy, one must increase the sample size by one hundred time. As a result, the computational burden

of the crude MCS tends to prohibitively large when a large number of samples are required to achieve a satisfactory level of accuracy (e.g., the quantity to be estimated is a small probability) and/or the model involved is expensive to evaluate.

2.2.2 Importance sampling

The crude MCS discussed in Section 2.2.1 requires a large sample size to achieve an estimator with a small variance. Importance Sampling (IS) (Rubinstein and Kroese 2016) is an alternative MCS method, which computes the estimating properties (e.g., expectation) of a distribution by sampling from an instrumental PDF $q(\boldsymbol{\theta})$. It is also known as a variance reduction technique.

Let $q(\boldsymbol{\theta})$ be an instrumental PDF supported on $\Omega \subset \mathbb{R}^{n_\theta}$ which is the same to the support of the original PDF $p_\theta(\boldsymbol{\theta})$. The mean value can be written as the expectation with respect to $q(\boldsymbol{\theta})$:

$$\mu = \int_{\boldsymbol{\theta} \in \Omega \subset \mathbb{R}^{n_\theta}} h(\boldsymbol{\theta}) \frac{p_\theta(\boldsymbol{\theta})}{q(\boldsymbol{\theta})} q(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbb{E}_q \left[\frac{h(\boldsymbol{\theta}) p_\theta(\boldsymbol{\theta})}{q(\boldsymbol{\theta})} \right] \quad (2-8)$$

Instead of sampling from $p_\theta(\boldsymbol{\theta})$, one can sample from $q(\boldsymbol{\theta})$ and adjust $h(\boldsymbol{\theta})$ with the likelihood ratio (or weight) $p_\theta(\boldsymbol{\theta})/q(\boldsymbol{\theta})$. Now the IS estimator of μ can be expressed as:

$$\hat{\mu}_{IS} = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} \frac{h(\boldsymbol{\theta}^{(i)}) p_\theta(\boldsymbol{\theta}^{(i)})}{q(\boldsymbol{\theta}^{(i)})} \quad (2-9)$$

where $\{\boldsymbol{\theta}^{(i)}; i: 1 \rightarrow N_{IS}\}$ are a set of i.i.d samples from $q(\boldsymbol{\theta})$. The IS estimator is unbiased, and its variance can be estimated by:

$$\text{Var}[\hat{\mu}_{IS}] = \frac{\sum_{i=1}^{N_{IS}} \left(\frac{h(\boldsymbol{\theta}^{(i)}) p_\theta(\boldsymbol{\theta}^{(i)})}{q(\boldsymbol{\theta}^{(i)})} - \hat{\mu}_{MCS} \right)^2}{N_{IS}(N_{IS} - 1)} \quad (2-10)$$

It can be observed in Equation (2-7) that the selection of $q(\boldsymbol{\theta})$ is of crucial importance in reducing the variance of the estimator. The optimal PDF $q(\boldsymbol{\theta})$ can be derived by minimizing the variance. An algorithm that uses IS to generate samples from Z and their associated weights is shown in Table 2-2. After that, the mean and variance values of the IS estimator can be calculated by using Equation (2-9) and Equation (2-10), respectively.

Table 2-2: Algorithm for Importance Sampling (non-parallel)

imptntSampling($N, p_{\theta}(\theta), q(\theta), h$)

Inputs: N : sample size, $p_{\theta}(\theta)$: original joint probability density function, $q(\theta)$: instrumental probability density function, h : model under consideration

Output: samples $\{z\}$ from Z and their associated weights $\{w\}$ where $Z = h(\theta)$

```

1:    $i \leftarrow 0, \{z\} \leftarrow \emptyset, \text{ and } \{w\} \leftarrow \emptyset,$ 
2:   while  $i < N$ :
3:       generate a sample  $\theta^{(i)}$  from  $q(\theta)$ 
4:       compute  $z^{(i)} = h(\theta^{(i)})$  and  $w^{(i)} = p_{\theta}(\theta^{(i)})/q(\theta^{(i)})$ 
5:       update  $i \leftarrow i + 1, \{z\} \leftarrow \{z\} + z^{(i)}, \{w\} \leftarrow \{w\} + w^{(i)}$ 
6:   endwhile
7:   return  $(\{z\}, \{w\})$ 

```

2.2.3 Accept-reject method

The accept-reject method generates independently and identically distributed (i.i.d) samples from the target PDF $\pi(\theta)$ whose normalizing constant is unknown by using a proposed distribution with the instrumental PDF $q(\theta)$. The basic idea behind this method is that one can generate a sample from $\pi(\theta)$ and accept it with probability $\pi(\theta)/(Mq(\theta))$ such that the accepted samples will follow the target distribution. Here the finite bound M is a constant. The procedure for the accept-reject method to generate N samples is shown in Table 2-3.

Table 2-3: Algorithm for accept-reject method (non-parallel)

acceptReject($N, \pi(\theta), q(\theta), M$)

Inputs: N : sample size, $\pi(\theta)$: target joint probability density function with unknown normalizing constant, $q(\theta)$: instrumental probability density function, M : a constant finite bound

Output: samples $\{\theta\}$

```

1:    $i \leftarrow 0$  and  $\{\theta\} \leftarrow \emptyset$ 
2:   while  $i < N$ :
3:       generate a sample  $u^{(i)}$  from the uniform  $U(0,1)$ 
4:       generate a sample  $\theta^{(i)}$  from  $q(\theta)$ 
5:       if  $u^{(i)} < \pi(\theta)/(Mq(\theta^{(i)}))$ :

```

```

6         update  $i \leftarrow i + 1$ ,  $\{\theta\} \leftarrow \{\theta\} + \theta^{(i)}$ 
7     endif
8: endwhile
9: return  $\{\theta\}$ 

```

The constant M is equal to the average number of iterations required to accept a sample, which indicates the efficiency of the stochastic sampling method. A small M is preferred while the inequality $\pi(\theta) \leq (Mq(\theta))$ must be satisfied for all θ .

2.2.4 Markov chain Monte Carlo methods

Markov chain Monte Carlo (MCMC) methods started in 1953 as Metropolis et.al presented the Metropolis algorithm in their pioneering article (Metropolis et al. 1953) to simulate a state that a liquid is in equilibrium with its gas phase. Later, Hastings (Hastings 1970) proposed a simulation scheme to generalize the Metropolis algorithm, which is often referred to as Metropolis-Hastings (M-H) algorithm. A special case of M-H algorithm, called Gibbs sampling, was first discussed in (Geman and Geman 1987) for finding the mode of the posterior. Later studies (Tanner and Wong 1987, Gelfand and Smith 1990) showed that the Gibbs sampling can be applied to various kinds of Bayesian inference.

In general, MCMC is defined as a class of methods for sampling from a target probability distribution $\pi(\theta)$ by using Markov chains. For a sequence of random variable vectors $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}$, it is a *Markov chain* if the conditional distribution $\theta^{(t+1)} | \theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}$ depends on $\theta^{(t)}$ only, i.e., $p(\theta^{(t+1)} | \theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}) = p(\theta^{(t+1)} | \theta^{(t)})$. Since time t is discrete, a Markov chain is in fact a discrete-time Markov process.

Three important concepts in MCMC is discussed as follows:

- (4) *Stationarity*. A Markov chain is stationary if and only if the marginal distribution of $\theta^{(t)}$ does not depend on t . Note that stationarity implies stationary transition probabilities but not vice versa.
- (5) *Reversibility*. A Markov chain is considered to be reversible if the distribution pair $(\theta^{(t)}, \theta^{(t+1)})$ is exchangeable. Stationarity is a necessary but not sufficient condition for reversibility. Reversibility is an important feature since (i) the conditions for Central Limit Theorem (CLT) are simpler if Markov chain is

reversible; (ii) the asymptotic variance can be estimated at the presence of reversibility.

- (6) *Ergodicity*. Ergodic Markov chains, also known as irreducible chains, are able to transit from any state to all possible states. Also, an ergodic Markov chain is a stationary chain where the same statistical characteristics can be observed in a realization of the chain and the realizations of a marginal distribution, which is useful to indicate the convergence of the Markov chain.

A wide range of tasks can be performed by using MCMC, including Bayesian inference, likelihood inference (implicit likelihood) and multidimensional integration. The crude MCS can be regarded as a special case of MCMC when the information about the past samples is not exploited. However, it must be noted that samples generated under MCMC schemes are correlated as they “*learn from the past samples*”. In the following, the algorithms for M-H algorithm and Gibbs sampling are introduced.

Metropolis-Hastings algorithm

Assume that our aim is to sample from an unnormalized target PDF $\pi(\boldsymbol{\theta})$. In M-H algorithm, given the current state $\boldsymbol{\theta}^{(i)}$, a candidate state $\boldsymbol{\theta}_c^{(i+1)}$ can be generated from a proposal conditional probability density $q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i)})$, which will be accepted with a certain probability. The algorithm for M-H algorithm is shown in Table 2-4 with a schematic diagram in Figure 2-4.

Table 2-4: Algorithm for Metropolis-Hastings algorithm

samplerMH ($N, \pi(\boldsymbol{\theta}), \boldsymbol{\theta}^{(1)}, q(\boldsymbol{\theta} \cdot)$)	
Inputs: N : sample size, $\pi(\boldsymbol{\theta})$: target joint probability density function with unknown normalizing constant, $\boldsymbol{\theta}^{(1)}$: initial state, $q(\boldsymbol{\theta} \cdot)$: proposed conditional probability density function given the current state	
Output: samples $\{\boldsymbol{\theta}\}$	
1:	$i \leftarrow 1, \boldsymbol{\theta}^{(1)}$ and $\{\boldsymbol{\theta}\} \leftarrow \boldsymbol{\theta}^{(1)}$
2:	while $i < N$:
3:	generate a sample $u^{(i)}$ from the uniform $U(0,1)$
4:	generate a candidate sample $\boldsymbol{\theta}_c^{(i+1)}$ from $q(\boldsymbol{\theta} \boldsymbol{\theta}^{(i)})$


```

5      if  $u^{(i)} < \pi(\theta_c^{(i+1)})q(\theta^{(i)}|\theta_c^{(i+1)}) / (\pi(\theta^{(i)})q(\theta_c^{(i+1)}|\theta^{(i)}))$ :
6          update  $\theta^{(i+1)} = \theta_c^{(i+1)}$ ,  $\{\theta\} \leftarrow \{\theta\} + \theta^{(i+1)}$  and  $i \leftarrow i + 1$ 
7      endif
8:  endwhile
9:  return  $\{\theta\}$ 

```

Figure 2-4: Diagram of Metropolis-Hastings algorithm in the 2-d space. The dashed lines are the contour levels of the probability density function.

It can be proven that the transition probability in M-H algorithm is reversible with respect to $\pi(\theta)$. A burn-in period is required to throw away some samples at the beginning of a Markov chain as they may not accurately represent the target distribution.

Gibbs sampling

Gibbs sampling can be seen as a special case of M-H algorithm, whose transition probability is constructed by the full-conditional distributions. Consider a target distribution $\pi(\theta)$ where $\theta = [\theta_1, \dots, \theta_{n_\theta}]^T$, in which samples of one component θ_j can be generated from its full-conditional distribution $\pi(\theta_j | \theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_{n_\theta})$. Gibbs samplers are commonly used to efficiently sample from the posterior distribution in Bayesian model updating (Ching et al. 2006, Bansal 2015, Cheung and Bansal 2017). The implementation of Gibbs sampling algorithm is summarized in Table 2-5.

Table 2-5: Algorithm for Gibbs sampling

samplerGibbs($N, \pi(\boldsymbol{\theta}), \boldsymbol{\theta}^{(1)}$)

Inputs: N : sample size, $\pi(\boldsymbol{\theta})$: target distribution which is full-conditional, $\boldsymbol{\theta}^{(1)}$: initial state

Output: samples $\{\boldsymbol{\theta}\}$

```

1:   $i \leftarrow 1, \boldsymbol{\theta}^{(1)}$  and  $\{\boldsymbol{\theta}\} \leftarrow \boldsymbol{\theta}^{(1)}$ 
2:  while  $i < N$ :
3:     $j \leftarrow 1$ 
4:    for  $j \leq n_\theta$ 
5:      generate a sample  $\theta_j^{(i+1)}$  from  $\pi(\theta_j | \theta_1^{(i+1)}, \dots, \theta_{j-1}^{(i+1)}, \theta_{j+1}^{(i)}, \dots, \theta_{n_\theta}^{(i)})$ 
6:      update  $j \leftarrow j + 1$ 
7:    endfor
8:    update  $\{\boldsymbol{\theta}\} \leftarrow \{\boldsymbol{\theta}\} + \boldsymbol{\theta}^{(i+1)}$  and  $i \leftarrow i + 1$ 
9:  endwhile
10: return  $\{\boldsymbol{\theta}\}$ 

```

2.3 Reliability analysis

2.3.1 Overview

In the context of engineering, *reliability* is defined as the likelihood that a component or a system will perform its intended function without ‘failure’ for a specified period of time under stated operating conditions (Yao et al. 2011). In the context of structural engineering, a failure event is often defined as the performance of the structure not satisfying the serviceability or/and ultimate limit state requirements. As the model performance can vary drastically when subjected to even a small variation in the model parameters or environment, reliability analysis is very essential for ensuring a system to perform as designed. The failure probability is a common measure of the reliability, i.e., the smaller probability of failure is, the higher the reliability becomes.

Given a model whose uncertain parameters are represented by a vector of n_θ random variables $\boldsymbol{\theta} = [\theta_1, \dots, \theta_{n_\theta}]^T$, the failure probability P_F for a specific event F can be written as:

$$P_F = \int_{\boldsymbol{\theta} \in \Omega \subseteq \mathbb{R}^{n_\theta}} \mathbb{I}_F(\boldsymbol{\theta}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbb{E}[\mathbb{I}_F(\boldsymbol{\theta})] \quad (2-11)$$

where $\mathbb{I}_F(\cdot)$ is the binary indicator function of the failure event, which equals to one if $\boldsymbol{\theta}$ belongs to the failure region Ω_F and zero otherwise. The failure region Ω_F is defined by

the performance function $g(\cdot)$, i.e., $\Omega_F = \{\boldsymbol{\theta}: g(\boldsymbol{\theta}) \leq 0\}$. Therefore, the corresponding limit-state function (LSF) $g(\boldsymbol{\theta}) = 0$ divides the parameter space Ω into the failure and safety regions. It should be noted that in Section 2.2 the performance function $g(\cdot)$ which predicts the output was named the primary model, and the indicator function $\mathbb{I}_F(\cdot)$ for reliability analysis is an example of the performance assessment model $h(\cdot)$, as illustrated in Figure 2-3.

A general solution for Equation (2-11) is difficult sometimes even impossible to obtain. Hence, a considerable amount of literature has been focusing on developing efficient and accurate ways to calculate the failure probability in means of approximation or sampling (Cornell 1969, Hasofer and Lind 1974, Der Kiureghian et al. 1987, Der Kiureghian and Dakessian 1998, Au and Beck 2001b, Katafygiotis et al. 2007, Katafygiotis and Cheung 2007, Cheung et al. 2011). In this section, two well-known approximation methods: first-order and second-order reliability methods (FORM and SORM) are presented. Then, a stochastic-sampling-based method called Subset Simulation (SS) (Au and Beck 2001b) is presented, discussing its efficiency in dealing with cases of small failure probabilities. The crude MCS and IS introduced in Section 2.2.1 and Section 2.2.2 can be easily implemented to estimate the failure probability by replacing $h(\cdot)$ with $\mathbb{I}_F(\cdot)$ in Equation (2-6) and Equation (2-9), thus will not be further discussed in this section.

2.3.2 First-order and second-order reliability methods

FORM/SORM are based on the linear or quadratic approximation of the original LSF around the *most probable point* (MPP) using Taylor series approximation. In FORM (Cornell 1969, Hasofer and Lind 1974), it searches for the MPP $\mathbf{u}^*$ in the independent standard normal space in which the original vector of random variables $\boldsymbol{\theta}$ is transformed to an independent standard normal vector $\mathbf{U}$ by isoprobabilistic transformation T (Rosenblatt 1952, Nataf 1962, Lebrun and Dutfoy 2009)

$$\mathbf{u}^* = \underset{\mathbf{u}}{\operatorname{argmin}}(\|\mathbf{u}\| | g(T^{-1}(\mathbf{u})) = 0) \quad (2-12)$$

where $g(\cdot)$ is the performance function, $T^{-1}(\cdot)$ inversely maps $\mathbf{u}$ to $\boldsymbol{\theta}$ and $\|\mathbf{u}\|$ is the ℓ^2 -norm of $\mathbf{u}$. The *reliability index* β defined as:

$$\beta = \|\mathbf{u}^*\| \quad (2-13)$$

is the minimal Euclidean distance from the origin to the MPP in the standard normal space.

Accordingly, the failure probability is given by:

$$P_F = \Phi(-\beta) \quad (2-14)$$

where $\Phi(\cdot)$ is the standard normal CDF.

In FORM, the error brought by the linear approximation tends to be large when the LSF is highly nonlinear and/or multiple MPPs exist. Various approaches have been suggested to alleviate this issue. For example, SORM (Der Kiureghian et al. 1987) approximates the LSF based on the second-order information of Taylor series expansion; multiple MPPs can be found to improve the accuracy of FORM (Der Kiureghian and Dakessian 1998).

2.3.3 Subset simulation

The basic idea behind Subset simulation (SS) (Au and Beck 2001b) is to convert the original rare failure event F into a sequence of m more frequent events, i.e., $F_1 \supset F_2 \cdots \supset F_m = F$. In a specific level $k = 1, 2, \dots, m$, define $F_k = \{\boldsymbol{\theta} | g(\boldsymbol{\theta}) \leq c_k\}$ where c_k is the threshold for this level and the thresholds of all levels satisfy $c_1 > c_2 > \dots > c_m = 0$. Hence, the small failure probability can be expressed as a product of a series of larger probabilities

$$P_F = P_{F_1} \prod_{k=2}^m P(F_k | F_{k-1}) \quad (2-15)$$

In each level except the final one, the threshold c_k ($k = 1, 2, \dots, m - 1$) is determined adaptively so that P_{F_1} or $P(F_{k+1} | F_k)$ is approximately equal to a fixed value P_0 ($0.1 \leq P_0 \leq 0.3$ has been suggested in (Zuev et al. 2012)). Specifically, in the first level ($k = 1$), a number of samples $\{\boldsymbol{\theta}\}_1$ are generated from the distribution $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and their outputs $\{g(\boldsymbol{\theta})\}_1$ are calculated by evaluating the performance function; c_1 is the $100P_0$ -th percentile of $\{g(\boldsymbol{\theta})\}_1$. Likewise, for $k = 2, \dots, m - 1$, samples $\{\boldsymbol{\theta}\}_k$ distributed as $p_{\boldsymbol{\theta}}(\boldsymbol{\theta} | F_{k-1})$ are generated using modified M-H algorithm (Au and Beck 2001b) and c_k is assigned to the $100P_0$ -th percentile of the samples' outputs $\{g(\boldsymbol{\theta})\}_k$ such that $P(F_k | F_{k-1})$ equals to P_0 approximately. In the final level $k = m$ (the procedure stops when $c_m \leq 0$ and one can reset c_m to zero), $P(F_m | F_{m-1})$ can be estimated by the

proportion of samples whose outputs (the values of the performance function) are smaller than zero.

2.4 Reliability-based design optimization

2.4.1 Overview

Reliability-based design optimization (RBDO), or *reliability-based optimization* (RBO), is a methodology that allows for optimizing the design configurations of a system while satisfying its reliability constraint(s). A typical formulation of a RBDO problem is given by:

$$\begin{aligned} \min_{\boldsymbol{\varphi}}: \quad & c(\boldsymbol{\varphi}) \\ \text{s. t. :} \quad & \begin{cases} P_{F_I}(\boldsymbol{\varphi}) \leq [P_{F_I}]; & I = 1, 2, \dots, n_{rc} \\ h_J(\boldsymbol{\varphi}) \leq 0; & J = 1, 2, \dots, n_{dc} \end{cases} \end{aligned} \quad (2-16)$$

where the deterministic design variables $\boldsymbol{\varphi} = [\varphi_1, \dots, \varphi_{n_\varphi}]^T \in \Omega_\varphi \subseteq \mathbb{R}^{n_\varphi}$ can be changed over the design space Ω_φ to minimize the objective function $c(\cdot)$. For $I = 1, 2, \dots, n_{rc}$, the inequality constraint $P_{F_I}(\boldsymbol{\varphi}) \leq [P_{F_I}]$ represents a *reliability constraint*, meaning that the occurrence probability of the failure event F_I given the design configuration $\boldsymbol{\varphi}$ must not exceed the allowable failure probability $[P_{F_I}]$. It implies that the system should always perform within prescribed safety margin by sustaining a certain reliability level. The associated failure event F_I is defined by the performance function $g_I(\boldsymbol{\varphi}, \boldsymbol{\theta})$, i.e., $F_I = \{\boldsymbol{\theta} \in \Omega_\theta | g_I(\boldsymbol{\varphi}, \boldsymbol{\theta}) \leq 0\}$. For $J = 1, \dots, n_{dc}$, the inequality constraint $h_J(\boldsymbol{\varphi}) \leq 0$ is a deterministic constraint. It can be easily observed from Equation (2-16) that RBDO distinguishes from the traditional deterministic design optimization with the presence of reliability constraints.

It is worth noting that RBDO problems can be formulated in other ways (Vanmarcke 1973, Enevoldsen and Sørensen 1994). For instance, one may desire to minimize the failure probability of a system under design under fixed costs. In Equation (2-16), the objective $c(\cdot)$ (typically, the cost function) is deterministic since no uncertain parameters $\boldsymbol{\theta}$ are involved; in cases that the cost function $c(\boldsymbol{\varphi}, \boldsymbol{\theta})$ has uncertain parameters, the objective function can be expressed in terms of $c(\boldsymbol{\varphi}, \boldsymbol{\theta})$, e.g., $\mathbb{E}[c(\boldsymbol{\varphi}, \boldsymbol{\theta})]$. Also, the performance function sometimes is written as $g_I(\boldsymbol{\theta}|\boldsymbol{\varphi})$ to indicate that the distribution of $\boldsymbol{\theta}$ is

conditional on the design variables $\boldsymbol{\varphi}$ (e.g., the mean value of an uncertain parameter is treated as a design variable).

RBDO provides a rational and rigorous framework for explicitly taking various sources of uncertainties into consideration. However, the computational efforts required to solve a RBDO problem can be prohibitively large. As can be seen in Equation (2-16), it is straightforward to solve this problem in a double-loop approach. In detail, in the outer loop, deterministic optimization is performed to search for the best solution based on the information gained from the reliability analysis in the inner loop, as shown in Figure 2-5. A single run of reliability analysis can be computationally expensive especially when implicitly numerical performance functions are involved; it is doomed to worsen for RBDO solved in a double-loop approach as repeated reliability analyses are inevitable. To this end, a lot of research has been devoted to alleviating the computational burden (e.g., Chen et al. (1997b), Royset et al. (2001), Du and Chen (2004), Du et al. (2008)). Works highly relevant to this thesis will be reviewed in the next section. For more complete surveys on RBDO, readers are referred to Schuëller and Jensen (2008), Aoues and Chateauneuf (2010), Valdebenito and Schuëller (2010), Yao et al. (2011).

2.4.2 Existing approaches

This section discusses three categories of approaches to solving a RBDO problem: double-loop approaches, single-loop approaches and decoupling approaches. For the sake of simple notation, the deterministic constraints are dropped and only one reliability constraint is considered in the following formulations of RBDO.

Double-loop approaches

Despite its simplicity, a double-loop approach often involving repeated reliability analyses tends to impose excessive computational burden, as illustrated in Figure 2-5. Approximate reliability methods (e.g., FORM/SORM discussed in Section 2.3.2) have been applied to efficiently perform the reliability analysis in a double-loop approach (Enevoldsen and Sørensen 1994, Lee and Kwak 1995, Tu et al. 2001, Lin et al. 2011)

$$\begin{aligned} \min_{\boldsymbol{\varphi}}: \quad & c(\boldsymbol{\varphi}) \\ \text{s. t.}: \quad & \beta(\boldsymbol{\varphi}) \leq [\beta] \end{aligned} \quad (2-17)$$

where $\beta(\boldsymbol{\varphi})$ is the reliability index for design configuration $\boldsymbol{\varphi}$ which is calculated by FORM/SORM and $[\beta]$ denotes the allowable reliability index. Compared to the reliability constraint Equation (2-16), the reliability is measured in terms of the reliability index rather than the failure probability. So a double-loop approach using β is also known as a *reliability index approach* (RIA) (Tu et al. 2001, Lin et al. 2011).

Figure 2-5: Flowchart of double-loop approaches for reliability-based design optimization

Alternatively, the expression of the reliability constraint can be transformed into a constraint on the performance function by the *performance measure approach* (PMA) (Der Kiureghian et al. 1994, Tu et al. 2001)

$$\begin{aligned} \min_{\boldsymbol{\varphi}}: \quad & c(\boldsymbol{\varphi}) \\ \text{s. t.}: \quad & g(\boldsymbol{\varphi}, T^{-1}(\mathbf{u}^{iF})) \leq 0 \end{aligned} \quad (2-18)$$

where $\mathbf{u}^{iF}$ is determined by the inverse FORM analysis, i.e.,

$$\mathbf{u}^{iF} = \underset{\mathbf{u}}{\operatorname{argmin}}(g(\boldsymbol{\varphi}, T^{-1}(\mathbf{u})) ||\mathbf{u}|| = [\beta]) \quad (2-19)$$

The results in several studies suggested that in general PMA outperforms RIA in stability and efficiency (Lee et al. 2002). Later, the efficiency of PMA was further enhanced with the use of the features of the performance function (Youn et al. 2003).

The preceding double-loop approaches use approximation reliability methods FORM/SORM as a building block; as a result, the issues (e.g., inaccuracy in cases with highly nonlinear LSFs) that these reliability methods are faced with will still bother. Though

the double-loop approaches do not reject more advanced reliability methods (e.g., SS), they should be employed with caution to avoid excessive computational burden.

Single-loop approaches

In principle, single-loop approaches (Chen et al. 1997a, Kuschel and Rackwitz 1997, Agarwal et al. 2007, Liang et al. 2008) take advantage of an equivalent optimization formulation in which the reliability constraints in Equation (2-17) are replaced by the Karush-Kuhn-Tucker (KKT) optimality conditions. Thus, the inner loop of reliability analysis can be avoided.

The single-loop nature of these approaches offers an attractive advantage over the double-loop counterparts: their efficiency becomes almost equivalent to the deterministic optimization without sacrificing the accuracy (Liang et al. 2008). However, numerical instability of the single-loop approaches has been observed in (Aoues and Chateaneuf 2010); furthermore, these approaches are established based on the same approximation concept in FORM or inverse FORM, thus cannot be applied to cases with highly nonlinear LSFs.

Decoupling

Probably the most prevalent decoupling approach to RBDO is so called Sequential Optimization and Reliability Assessment (SORA) method (Du and Chen 2004). In this approach, a series of deterministic optimization and reliability assessments is performed. In each iteration, the reliability in each reliability constraint is calculated after the decoupled deterministic optimization using RIA or PMA. To those reliability constraints that are not satisfied, a shifting vector will be applied to the corresponding design variables to push them into the feasible region. The iteration continues until the shifting distance approaches to zero. SORA still requires repeated reliability analyses though the number of reliability assessments is less than the double-loop approaches thanks to the sequential decoupling.

Alternatively, one can construct an explicit and deterministic *failure probability function* (FPF) $\hat{P}_F(\boldsymbol{\varphi})$ in lieu of the original one $P_F(\boldsymbol{\varphi})$ which is commonly implicit and is expensive to evaluate. By replacing $P_F(\boldsymbol{\varphi})$ in Equation (2-16) with the approximated

FPF $\hat{P}_F(\boldsymbol{\varphi})$, a RBDO problem is converted into a deterministic optimization problem. Figure 2-6 provides an illustration of this decoupling approach.

Figure 2-6: Flowchart of the decoupling approach to reliability-based design optimization based on the approximate failure probability function

It is nontrivial to approximate the explicit FPFs since it requires the calculation of the failure probabilities for different design configurations. Early efforts (Kanda and Ellingwood 1991, Gasser and Schuëller 1997, Jensen 2005, Ching and Hsieh 2007a, Ching and Hsieh 2007b) focused on the usage of the parametric families of functions as the approximate FPFs, particularly the exponential family of functions. Gasser and Schuëller (1997) used a quadratic function to approximate the logarithm of the FPF. To fit the quadratic function, the failure probabilities of the selected interpolation points were calculated by using response surface method and IS. Similarly, Jensen (2005) adopted a linear function to approximate the logarithm of the FPF for a linear dynamic system subjected to stochastic excitations. The approximation was locally adjusted to a reliability database, in which the reliability analysis was performed by using IS. The accuracy of the approximation over the entire design space depends on the number as well as the location of support points.

Another strategy for constructing the approximate FPF was proposed by Au (2005), in which design variables are artificially treated as random variables so that the FPF is proportional to the PDF of failure samples according to Bayes' Theorem. Following this

idea, maximum entropy principle was adopted to estimate the PDF belonging to the exponential family (Ching and Hsieh 2007a, Ching and Hsieh 2007b). One remarkable merit of this strategy is that it only requires one single run of stochastic-sampling-based reliability analysis to obtain the approximate FPF. Another merit is that different advanced stochastic-sampling-based reliability methods (e.g., IS and SS) can be incorporated into this decoupling approach, making it more robust to the nonlinearity of LSFs and the dimensionality. The basic idea of this strategy has been employed in this thesis to propose a new RBDO framework, which will be discussed in detail in the next chapter.

In a recent work, Yuan (2013) proposed a so-called weighted approach for a special class of RBDO problems in which design variables are the distribution parameters of uncertain random variables. By sampling from an instrumental PDF, the estimate of the FPF can be given by the weighted sum of the values of the indicator function. This approach was implemented for different stochastic-sampling-based reliability methods, including the crude MCS, IS and SS.

2.4.3 Applications to geotechnical problems

Traditionally, working stress design (WSD) was a dominant design methodology in the geotechnical community. In WSD, a lumped factor of safety, defined as the ratio of the characteristic resistance to the characteristic load effect, is used to consider all uncertainties associated with a geotechnical system. For a specific design configuration in WSD, its factor of safety should not exceed the prescribed allowable threshold. However, the drawbacks of WSD have been well recognized: (i) the choice of the factor of safety relies on the experience and judgement of engineers to a large extent; (ii) the risk associated with the design cannot be explicitly quantified. It implies that the failure probabilities of two systems with the same factors of safety may be different.

The current design methodology for most countries in the world is the load and resistance factor design (LRFD). LRFD is a term commonly used in USA design codes, which is also known as the limit state design in Canada and the partial factor approach in Eurocode. In the LRFD methodology, all limit states are checked in a multiple-factor format. The factors, usually including load, resistance and importance factors, are derived

based on the variability of parameters, the modeling error and failure consequences. Unlike WSD, LRFD uses a set of factors instead of a lumped factor of safety.

The design methodology RBDO provides a unified treatment of the uncertainties in a geotechnical system, thus becoming increasingly popular. Recent years have seen an increasing trend in the application of RBDO to a wide range of geotechnical systems such as foundations (Wang 2009, Wang 2011, Wang 2013, Huffman et al. 2015), slopes (Foye et al. 2006, Wang et al. 2013), drilled shafts (Wang 2011, Wang et al. 2011) and earth-retaining systems (Low 2005, Wu et al. 2011, Wang 2013).

Phoon et al. (2003) used FORM to quantify the design risk of transmission line structure foundation, and also proposed a simplified RBDO approach for practical implementations based on the factors calibrated by FORM. This idea of reliability approximation continued in the methods developed by Low (2005) and Zhang et al. (2012). In essence, these methods are implemented in a double-loop way where the outer loop performs the optimization and the candidate optimal solution is evaluated by the simplified reliability analysis in the inner loop. To address these issues, the framework of expanded reliability-based design has been studied by some researchers (Ching and Hsieh 2007b, Wang 2009, Wang et al. 2011, Wang 2013) over the last decade following the decoupling idea first presented in Au (2005). The approach is robust with respect to the number of random variables and the degree of system nonlinearity.

2.5 Conclusions

This chapter first covered the essential components for performing the uncertainty-based analysis including the representation of different uncertainties and the algorithms for Monte Carlo Simulation (MCS). Then, the development of a specific type of the uncertainty-based analysis, namely reliability analysis, was discussed. Last, a design methodology called reliability-based design optimization (RBDO) was discussed, in which the optimal design of a system is searched without violating the reliability constraints. Existing approaches to RBDO as well as their applications to geotechnical problems were discussed and compared.

Considering the characteristics of up-to-date models for stochastic complex systems, it is preferable a RBDO approach which can (i) handle a large number of random variables;

(ii) handle several design variables (e.g., less than ten); (iii) incorporate commercial codes without intruding; (iv) achieve the acceptable accuracy within a prescribed number of performance function (or model) evaluations. Among the four performance metrics, the efficiency is measured in terms of the number of performance function evaluations because the runtime of other processes is negligible compared to the model evaluation.

The literature discussed in this chapter can be summarized as follows:

- (1) There are various types of uncertainties that need to be identified, modeled and quantified. Uncertainties can be represented by probability theory, interval analysis, evidence theory, possibility theory, etc. Among them, probability theory is the most widely used one due to its sound theoretical foundation.
- (2) In uncertainty-based analysis, the methods using stochastic sampling techniques can generate samples from the distribution of a quantity of interest so that statistical characteristics of the quantity can be summarized based on samples. They are flexible as they can literally incorporate any models with any levels of complexity. But the issue of the slow convergence rate needs to be addressed.
- (3) As a building block of RBDO, reliability analysis itself can be computationally expensive. Approximation methods (e.g., FORM/SORM) have been proposed to tackle this issue but their accuracy is poor for the cases with highly nonlinear performance functions or multiple most probable points. Alternatively, advanced stochastic-sampling-based reliability methods (e.g., IS and SS) have become increasingly popular.
- (4) Most of the current studies on RBDO have focused on performing the design optimization in a double- or single-loop way based on the approximation of the LSF. Like the approximation methods for reliability analysis, they suffer from the poor accuracy if the model is complex. On the other hand, the decoupling approaches that approximate the explicit failure probability function are promising to excel at the four performance metrics. In the geotechnical community, the design methodology is shifting from LRFD to RBDO with the advance of computing capability and model simulation techniques.

3 Approximate failure probability function in the partitioned design space

Summary

This chapter presents an efficient framework for RBDO, named Design Optimization in the PARTitioned Design Space (DOPADS), which is robust to complex systems involving computationally expensive numerical models and/or a large number of random variables. DOPADS is based on a decoupling approach in which the FPF is approximated in the partitioned design space. Following the idea proposed by Au (2005), the FPF is proportional to the PDF of the design variable components of failure samples. To enhance the accuracy of the FPF approximation, DOPADS adaptively partitions the design space into several subspaces and generates sufficient failure samples accordingly. Three illustrative examples involving structural systems subjected to static or dynamic loadings are discussed to demonstrate the efficiency and accuracy of the proposed framework.

3.1 Introduction

In the previous chapter, the typical formulation of a RBDO problem is introduced in Equation (2-16) and existing approaches to RBDO are reported in Section 2.4.2. The formulation in Equation (2-16) will continue to be used in this chapter. Among different approaches, the decoupling approaches (Kanda and Ellingwood 1991, Gasser and Schuëller 1997, Jensen 2005, Ching and Hsieh 2007a, Ching and Hsieh 2007b) based on the explicit

approximation of the FPF are attractive to researchers and practitioners due to its fully decoupling nature.

The key step in this class of approaches is the approximation of the FPF $P_F(\boldsymbol{\varphi})$ ahead of the deterministic optimization loop, which is by no means an easy task. Some previous research into the approximate FPF (Kanda and Ellingwood 1991, Gasser and Schuëller 1997, Jensen 2005) focused on the parametric families of functions, particularly the exponential family. But the accuracy of the approximate FPF is influenced largely by the number of support points and their locations. Each support point requires a single run of reliability analysis for the corresponding design configuration, which means the repeated reliability analyses are required in the construction procedure of the approximate FPF. Another issue of the parametric family is that the approximation error can be very large in the cases when the actual FPF does not follow well the predefined family.

In this chapter, a novel framework DOPADS for RBDO in which the approximate FPF is nonparametric is presented. It partitions the entire design space Ω_φ into several subspaces in an iterative manner. By generating sufficient failure samples in each subspace, one can obtain the distribution information about the failure samples even in the low-density regions.

The rest of this chapter is organized as follows. In Section 3.2, the FPF is related to the PDF of failure samples based on Bayes' theorem as well as the augmented parameter space. Section 3.3 proposes an iterative scheme for density estimation in the partitioned design space. Section 3.4 summarizes the implementation procedures of DOPADS. The accuracy and efficiency of the proposed framework are discussed in Section 3.5 using three illustrative examples.

3.2 The proposed framework

3.2.1 Formulation of failure probability function

Consider the RBDO problem in Equation (2-16) that involves n_φ deterministic design variables $\boldsymbol{\varphi} \in \Omega_\varphi \subseteq \mathbb{R}^{n_\varphi}$ and a vector of n_θ uncertain parameters $\boldsymbol{\theta} \in \Omega_\theta \subseteq \mathbb{R}^{n_\theta}$ where Ω_φ and Ω_θ denote the design space and the uncertain parameter space, respectively. To explore the design space in a single reliability analysis, an augmented

reliability formulation was proposed by Au (2005), where design variables $\boldsymbol{\varphi}$ are artificially considered as random variables with a prescribed joint probability distribution $p_{\boldsymbol{\Phi}}(\boldsymbol{\varphi})$ (e.g., uniform distribution) over the design space. In the augmented space $\Omega \subseteq \mathbb{R}^{n_{\varphi}+n_{\theta}}$, the FPF can be obtained based on Bayes' theorem

$$P_F(\boldsymbol{\varphi}) = P(F|\boldsymbol{\varphi}) = \frac{p(\boldsymbol{\varphi}|F)P(F)}{p_{\boldsymbol{\Phi}}(\boldsymbol{\varphi})} \quad (3-1)$$

where $P(F)$ is the failure probability of the augmented system

$$P(F) = \int_{\boldsymbol{\varphi} \in \Omega_{\varphi}} \int_{\boldsymbol{\theta} \in \Omega_{\theta}} \mathbb{I}_F(\boldsymbol{\varphi}, \boldsymbol{\theta}) p_{\boldsymbol{\Phi}}(\boldsymbol{\varphi}) p_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\varphi} d\boldsymbol{\theta} \quad (3-2)$$

and $p(\boldsymbol{\varphi}|F)$ is the PDF of $\boldsymbol{\varphi}$ conditional on the failure event F

$$p(\boldsymbol{\varphi}|F) = \frac{\int_{\boldsymbol{\theta} \in \Omega_{\theta}} \mathbb{I}_F(\boldsymbol{\varphi}, \boldsymbol{\theta}) p_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}}{P(F)} p_{\boldsymbol{\Phi}}(\boldsymbol{\varphi}) \quad (3-3)$$

Equation (3-1) implies that the FPF $P_F(\boldsymbol{\varphi})$ is proportional to $p(\boldsymbol{\varphi}|F)$, i.e., when $p_{\boldsymbol{\Phi}}(\boldsymbol{\varphi})$ is set to be a constant (uniform distribution), $P_F(\boldsymbol{\varphi}) \propto p(\boldsymbol{\varphi}|F)$. Given the design component of failure samples generated by the stochastic-sampling-based estimation of $P(F)$, the density function can be estimated via density estimation. Then, the density function is scaled to obtain the FPF with a scaling factor $P(F)/p_{\boldsymbol{\Phi}}(\boldsymbol{\varphi})$ based on Equation (3-1). Therefore, only one single reliability analysis is required. Figure 3-1 illustrates the basic idea.

Figure 3-1: Estimated density function based on failure samples and the failure probability function

3.2.2 Predefined joint probability distribution $p_{\Phi}(\boldsymbol{\varphi})$

In practice, the design space Ω_{φ} is generally defined by the lower bound $\boldsymbol{\varphi}^L$ and upper bound $\boldsymbol{\varphi}^U$. We artificially treat $\boldsymbol{\varphi}$ as uniformly distributed random variables, i.e., $\boldsymbol{\varphi} \sim \text{uniform}(\boldsymbol{\varphi}^L, \boldsymbol{\varphi}^U)$. Then, the value of the joint probability distribution function $p_{\Phi}(\boldsymbol{\varphi})$ equals to a constant:

$$p_{\Phi}(\boldsymbol{\varphi}) = \frac{1}{|\Omega_{\varphi}|} \quad (3-4)$$

where $|\Omega_{\varphi}|$ is the volume of the design space.

3.2.3 Reliability analysis of the augmented system and failure samples

It should be noted that the failure probability of the augmented system differs from that of the original system due to the additional uncertainty introduced by treating design variables as random ones.

In the context of DOPADS, different stochastic-sampling-based reliability methods discussed in Section 2.4 can be adopted to estimate $P(F)$ and to generate failure samples. It is suggested in DOPADS to use the crude MCS for the cases with large failure probabilities and SS (Au and Beck 2001b) for those with small failure probabilities (say $P(F) \leq 0.1$). Also notice that failure samples $\{(\boldsymbol{\varphi}^{(i)}, \boldsymbol{\theta}^{(i)}) | g(\boldsymbol{\varphi}^{(i)}, \boldsymbol{\theta}^{(i)}) \leq 0, i: 1 \rightarrow N\}$ are byproducts of the stochastic sampling in the reliability analysis.

3.3 An iterative scheme for density estimation

The core and critical part in DOPADS is the estimation of the density function $p(\boldsymbol{\varphi}|F)$. Let $\{(\boldsymbol{\varphi}, \boldsymbol{\theta})\}_F = \{(\boldsymbol{\varphi}^{(i)}, \boldsymbol{\theta}^{(i)}) | g(\boldsymbol{\varphi}^{(i)}, \boldsymbol{\theta}^{(i)}) \leq 0, i: 1 \rightarrow N\}$ denote a set of failure samples that have been generated after the reliability analysis of the augmented system. Then the design component $\{\boldsymbol{\varphi}\}_F$ follows the distribution $p(\boldsymbol{\varphi}|F)$ and can be used for density estimation (Parzen 1962, Silverman 1986, Scott 2015). For brevity, the term *failure samples* refers to the design component $\{\boldsymbol{\varphi}\}_F$, unless otherwise stated.

Since most failure samples are scattering over the high-density region, the accuracy of the estimated density function is poor over its low-density region. However, in the context of real-world RBDO problems, one is more concerned about the low-density region which corresponds to the low failure probabilities (say $10^{-2} \sim 10^{-4}$).

To resolve this issue, an iterative scheme for density estimation is proposed in DOPADS. Instead of estimating $p(\boldsymbol{\varphi}|F)$ over the entire design space, one can partition the design space iteratively into several subspaces such that failure samples can be propagated into the low-density region and generate the distribution information there.

3.3.1 Partitioning the design space

As shown in Figure 3-2, the initial design space Ω_φ is iteratively partitioned into $n_{it} + 2$ mutually exclusive and collectively exhaustive subspaces (also called regions) where n_{it} is the index of the final iteration

$$\Omega_\varphi = \mathcal{H}_0 \cup \mathcal{H}_1 \cup \cdots \cup \mathcal{H}_{n_{it}} \cup \mathcal{L}_{n_{it}} \quad (3-5)$$

In each subspace, sufficient failure samples are generated to ensure the accuracy of density estimation. Specifically, in the initial iteration $k = 0$, if the size of failure samples N_0 in Ω_φ is smaller than the prescribed value N_s , additional $(N_s - N_0)$ failure samples will be generated by modified Metropolis-Hastings algorithm (Au and Beck 2001b). The generation of additional samples will be further discussed in the next section. Then the density function estimator $\hat{p}_0(\boldsymbol{\varphi}|F)$ can be obtained by using multivariate density estimation techniques (Parzen 1962, Silverman 1986, Scott 2015). Since the low-density region in Ω_φ has poor accuracy, the design space Ω_φ is partitioned into the high-density part $\mathcal{H}_0$ and the low-density part $\mathcal{L}_0$, and $\mathcal{L}_0$ will be further populated if the stopping criterion is not satisfied. Likewise, in iteration k ($k = 1, \dots, n_{it}$), it starts with the generation of additional failure samples in the previous low-density region $\mathcal{L}_{k-1}$ until the sample size reaches N_s ; once sufficient failure samples are available, the density function $\hat{p}_k(\boldsymbol{\varphi}|F)$ over $\mathcal{L}_{k-1}$ is estimated and subsequently is used to partition $\mathcal{L}_{k-1}$ into the current high-density region $\mathcal{H}_k$ and low-density one $\mathcal{L}_k$, i.e., $\mathcal{L}_{k-1} = \mathcal{H}_k \cup \mathcal{L}_k$. For $k = 0, 1, \dots, n_{it}$,

$$\mathcal{L}_k = \{\boldsymbol{\varphi} | \hat{p}_k(\boldsymbol{\varphi}|F) < p_k^*\} \quad (3-6a)$$

$$\mathcal{H}_k = \{\boldsymbol{\varphi} | \hat{p}_k(\boldsymbol{\varphi}|F) \geq p_k^*\} \quad (3-6b)$$

where p_k^* is the partition density value in iteration k , which is determined adaptively according to the distribution of failure samples in the subspace.

Figure 3-2: Iterative density estimation in the partitioned 2-dimensional design space

According to the law of total probability, the final estimator $\hat{p}(\boldsymbol{\varphi}|F)$ can be written as:

$$\hat{p}(\boldsymbol{\varphi}|F) = \sum_{k=0}^{n_{it}} \hat{p}(\boldsymbol{\varphi}|F \cap \mathcal{H}_k)P(\mathcal{H}_k|F) + \hat{p}(\boldsymbol{\varphi}|F \cap \mathcal{L}_{n_{it}})P(\mathcal{L}_{n_{it}}|F) \quad (3-7)$$

where

$$\hat{p}(\boldsymbol{\varphi}|F \cap \mathcal{H}_0)P(\mathcal{H}_0|F) = \begin{cases} \hat{p}_0(\boldsymbol{\varphi}|F), & \boldsymbol{\varphi} \in \mathcal{H}_0 \\ 0, & \boldsymbol{\varphi} \notin \mathcal{H}_0 \end{cases} \quad (3-8a)$$

$$\hat{p}(\boldsymbol{\varphi}|F \cap \mathcal{H}_k)P(\mathcal{H}_k|F) = \begin{cases} \hat{p}_k(\boldsymbol{\varphi}|F)P(\mathcal{L}_{k-1}|F), & \boldsymbol{\varphi} \in \mathcal{H}_k \\ 0, & \boldsymbol{\varphi} \notin \mathcal{H}_k \end{cases} \quad (k = 1, \dots, n_{it}) \quad (3-8b)$$

$$\hat{p}(\boldsymbol{\varphi}|F \cap \mathcal{L}_{n_{it}})P(\mathcal{L}_{n_{it}}|F) = \begin{cases} \hat{p}_{n_{it}}(\boldsymbol{\varphi}|F)P(\mathcal{L}_{n_{it}-1}|F), & \boldsymbol{\varphi} \in \mathcal{L}_{n_{it}} \\ 0, & \boldsymbol{\varphi} \notin \mathcal{L}_{n_{it}} \end{cases} \quad (3-8c)$$

In Equation (3-8), $P(\mathcal{L}_{k-1}|F)$ ($k = 1, \dots, n_{it}$) is the probability of failure samples falling into the region $\mathcal{L}_{k-1}$. For $k = 1$, $P(\mathcal{L}_0|F)$ can be estimated by the ratio of failure samples scattering over $\mathcal{L}_0$ to those in the entire design space Ω_φ . For $k = 2, \dots, n_{it}$,

$$P(\mathcal{L}_{k-1}|F) = P(\mathcal{L}_0|F) \prod_{t=1}^{k-1} P(\mathcal{L}_t|F \cap \mathcal{L}_{t-1}) \quad (3-9)$$

where $P(\mathcal{L}_t|F \cap \mathcal{L}_{t-1})$ can be estimated by the ratio of failure samples scattering over the current low-density region $\mathcal{L}_t$ to those over the previous low-density region $\mathcal{L}_{t-1}$ in the t -th iteration. Notice that once the partition density p_k^* is fixed for each iteration, $P(\mathcal{L}_k|F \cap \mathcal{L}_{k-1})$ (or $P(\mathcal{L}_0|F)$) is fixed; and vice versa. Both can be used to determine the boundary between the high- and low-density parts, which will be detailed in Section 3.4.

3.3.2 Generation of additional failure samples

At the beginning of the k -th ($k = 1, \dots, n_{it}$) iteration, there are a small set of N_k failure samples scattering over the previous low-density region $\mathcal{L}_{k-1}$ which follow the distribution $p_k(\boldsymbol{\varphi}|F)$. To enhance the density estimator $\hat{p}_k(\boldsymbol{\varphi}|F)$, the sample size in $\mathcal{L}_{k-1}$ should be increased from N_k to N_s . In other words, $(N_s - N_k)$ additional failure samples following the distribution $p_k(\boldsymbol{\varphi}|F)$ are required to populate $\mathcal{L}_{k-1}$.

MCMC is a set of powerful methods for simulating samples from a target distribution $\pi(\boldsymbol{\varphi}, \boldsymbol{\theta})$ as discussed in Section 2.2.4. In DOPADS, the target distribution is:

$$\pi(\boldsymbol{\varphi}, \boldsymbol{\theta}) = p_k(\boldsymbol{\varphi}, \boldsymbol{\theta}|F) \propto \mathbb{I}_F(\boldsymbol{\varphi}, \boldsymbol{\theta}) \mathbb{I}_{\mathcal{L}_{k-1}}(\boldsymbol{\varphi}) p_{\Phi}(\boldsymbol{\varphi}) p_{\Theta}(\boldsymbol{\theta}) \quad (3-10)$$

where $\mathbb{I}_{\mathcal{L}_{k-1}}(\cdot)$ is an indicator function which equals to one if a sample belongs to $\mathcal{L}_{k-1}$ and zero otherwise. A MCMC algorithm, called modified Metropolis-Hastings (Au and Beck 2001b), is used due to its robustness to high dimensionality.

Consider a current sample $(\boldsymbol{\varphi}, \boldsymbol{\theta}) = (\varphi_1, \dots, \varphi_{n_{\varphi}}, \theta_1, \dots, \theta_{n_{\theta}})$ following $\pi(\boldsymbol{\varphi}, \boldsymbol{\theta})$. In the modified Metropolis-Hastings algorithm, for each component φ_j ($j = 1, \dots, n_{\varphi}$) (or θ_j ($j = 1, \dots, n_{\theta}$)), a one-dimensional PDF $q_j(\xi|\varphi_j)$ (or $q_j(\xi|\theta_j)$) centered at φ_j (or θ_j) with symmetric property is proposed, i.e., $q_j(\xi|\varphi_j) = q_j(\varphi_j|\xi)$ ($q_j(\xi|\theta_j) = q_j(\theta_j|\xi)$). Generation of the next sample $(\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1}$ is summarized in Table 3-1.

Table 3-1: Algorithm for modified Metropolis-Hastings algorithm

$\text{samplerMMH}((\boldsymbol{\varphi}, \boldsymbol{\theta}), q_j(\xi \varphi_j) (j = 1, \dots, n_{\varphi}), q_j(\xi \theta_j) (j = 1, \dots, n_{\theta}), p_{\Phi}(\boldsymbol{\varphi}), p_{\Theta}(\boldsymbol{\theta}), \mathbb{I}_F(\boldsymbol{\varphi}, \boldsymbol{\theta}), \mathbb{I}_{\mathcal{L}_{k-1}}(\boldsymbol{\varphi}))$	
Inputs: $(\boldsymbol{\varphi}, \boldsymbol{\theta})$: current sample, $q_j(\xi \varphi_j)$ and $q_j(\xi \theta_j)$: proposal probability density function for each component, $p_{\Phi}(\boldsymbol{\varphi})$ and $p_{\Theta}(\boldsymbol{\theta})$: initial PDF, $\mathbb{I}_F(\boldsymbol{\varphi}, \boldsymbol{\theta})$: indicator function of the failure event, $\mathbb{I}_{\mathcal{L}_{k-1}}(\boldsymbol{\varphi})$: indicator function of the sample in $\mathcal{L}_{k-1}$	
Output: next sample $(\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1}$	
1:	for $j = 1$ to n_{φ} :
2:	generate a sample u_j from the uniform distribution $U(0,1)$
3:	generate a candidate component ξ_j from $q_j(\xi \varphi_j)$
4:	if $u_j < \frac{p(\xi_j)q_j(\varphi_j \xi_j)}{p(\varphi_j)q_j(\xi_j \varphi_j)}$ where $p(\cdot)$ is a marginal PDF of $p_{\Phi}(\boldsymbol{\varphi})$ or $p_{\Theta}(\boldsymbol{\theta})$:

```

5:      accept component  $\tilde{\varphi}_j = \xi_j$ 
6:      else:
7:      reject component  $\tilde{\varphi}_j = \varphi_j$ 
8:      endif
9:      endfor
10:   for  $j = 1$  to  $n_\theta$ :
11:       generate a sample  $u_j$  from the uniform
12:       distribution  $U(0,1)$ 
13:       generate a candidate component  $\xi_j$  from  $q_j(\xi|\theta_j)$ 
14:       if  $u_j < \frac{p(\xi_j)q_j(\theta_j|\xi_j)}{p(\theta_j)q_j(\xi_j|\theta_j)}$ :
15:           accept component,  $\tilde{\theta}_j = \xi_j$ 
16:       else:
17:           reject component,  $\tilde{\theta}_j = \theta_j$ 
18:       endif
19:   candidate sample  $(\tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\theta}}) = (\tilde{\varphi}_1, \dots, \tilde{\varphi}_{n_\varphi}, \tilde{\theta}_1, \dots, \tilde{\theta}_{n_\theta})$ 
20:   if  $\mathbb{I}_F(\tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\theta}}) = 1$  and  $\mathbb{I}_{\mathcal{L}_{k-1}}(\tilde{\boldsymbol{\varphi}}) = 1$ :
21:       accept the candidate sample,  $(\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} = (\tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\theta}})$ 
22:   else:
23:       reject the candidate sample,  $(\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} = (\boldsymbol{\varphi}, \boldsymbol{\theta})$ 
24:   endif

```

To summarize, in the component level of ‘accept-reject’, one accepts the candidate component ξ_j with the acceptance ratio; and in the sample level of ‘accept-reject’, one accepts the candidate sample $(\tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\theta}})$ only if the system fails and $\tilde{\boldsymbol{\varphi}}$ lies in the subspace $\mathcal{L}_{k-1}$. The algorithm displayed in Table 3-1 is repeated until N_s samples are generated. Note that N_k initial failure samples following $p_k(\boldsymbol{\varphi}, \boldsymbol{\theta}|F)$ are available, thus N_k Markov chains can be used to generate N_s samples. The generated set of N_s failure samples is denoted as $\{(\boldsymbol{\varphi}, \boldsymbol{\theta})\}_{F, \mathcal{L}_{k-1}}$. It is obvious that its design component $\{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}}$ follows $p_k(\boldsymbol{\varphi}|F)$.

If the current sample $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ follows $p_k(\boldsymbol{\varphi}, \boldsymbol{\theta}|F)$, the next sample $(\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1}$ generated by modified Metropolis-Hastings algorithm will also follow the same distribution. The proof is given in Appendix A.

3.3.3 Density estimation in each iteration

After populating the subspace $\mathcal{L}_{k-1}$ with a sufficient number of N_s failure samples $\{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}} = \{\boldsymbol{\varphi}^{(i)}, i: 1 \rightarrow N_s\}$, our task is to estimate the PDF $\hat{p}_k(\boldsymbol{\varphi}|F): \mathcal{L}_{k-1} \rightarrow \mathbb{R}_{\geq 0}$ based on these samples. A wide variety of multivariate density estimation techniques (Parzen 1962, Silverman 1986, Scott 2015) can be used, such as the histogram, kernel estimators, and orthogonal series estimators.

For DOPADS, the density estimation selected should be able to handle (i) multivariate problems as it is common more than one design variables are involved; (ii) complex boundary; (iii) unknown class of the underlying densities. Therefore, DOPADS uses two nonparametric density estimation techniques, namely, smoothed histogram and Bayesian Sequential Partitioning (BSP) (Lu et al. 2013).

Smoothed histogram

In the smoothed histogram, the entire design space is binned using a regular grid and then the piecewise-constant estimator is smoothed using regression.

For each design variable φ_j ($j = 1, \dots, n_\varphi$), its range is divided into $n_{B,j}$ intervals. It generates a regular grid and a total number of $n_B = \prod_{j=1}^{n_\varphi} n_{B,j}$ rectangular bins B_t where $\cup_{t=1}^{n_B} B_t = \Omega_\varphi$ is shown in Figure 3-3. The piecewise-constant estimator $\hat{p}_k(\boldsymbol{\varphi}|F)$ can be expressed as:

$$\hat{p}_k(\boldsymbol{\varphi}|F) = \sum_{t=1}^{n_B} \hat{p}_{k,t}(\boldsymbol{\varphi}) \mathbb{I}_{B_t}(\boldsymbol{\varphi}), \quad \boldsymbol{\varphi} \in \{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}} \quad (3-11)$$

where $\mathbb{I}_{B_t}(\cdot)$ is an indicator function which equals to one if $\boldsymbol{\varphi} \in B_t$ and zero otherwise; $\hat{p}_{k,t}(\boldsymbol{\varphi}) = N_t/(N_s|B_t|)$ where $|B_t|$ represents the volume of the bin B_t and N_t is the size of samples in B_t .

Figure 3-3: Failure probability function by using the smoothed histogram in $\mathcal{L}_{k-1}$

The coordinates of all bin centers in Ω_φ and their density values can be considered as support points for smoothing the piecewise constant $\hat{p}(\boldsymbol{\varphi}|F)$. Many regression methods (Kleinbaum et al. 2013) including parametric and nonparametric ones can be chosen to fit the density function based on the complexity level of the underlying density function. Take second-order polynomial as an example. The smoothed estimator is given by:

$$\tilde{p}(\boldsymbol{\varphi}|F) = \mathbf{a}^T \mathbf{b} = a_0 + \sum_{j=1}^{n_\varphi} a_j \varphi_j + \sum_{j=1}^{n_\varphi} \sum_{k=j}^{n_\varphi} a_{jk} \varphi_j \varphi_k \quad (3-12a)$$

$$\mathbf{a} = [a_0, \dots, a_{n_\varphi}, a_{11}, a_{12}, \dots, a_{n_\varphi n_\varphi}]^T \quad (3-12b)$$

$$\mathbf{b} = [1, \varphi_1, \dots, \varphi_{n_\varphi}, \varphi_1^2, \varphi_1 \varphi_2, \dots, \varphi_{n_\varphi}^2]^T \quad (3-12c)$$

The coefficient vector is calculated by minimizing the mean squared error over all support points. To guarantee the solution existence, at least $\frac{n_\varphi(n_\varphi+3)+2}{2}$ support points are need.

Bayesian Sequential Partitioning (BSP)

Multivariate density estimation becomes very challenging as the number of design variables increases. That is, the required sample size N_s to achieve a certain level of accuracy increases exponentially with the dimension, also known as the curse of dimensionality. Unlike traditional histograms where the samples are binned into regular grids, BSP (Lu et

al. 2013) is an advanced nonparametric density estimation technique, which partitions the space adaptively based on the empirical distribution information of samples.

Consider a set of samples $\{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}}$. Define a partition $\mathcal{P}$ of $\mathcal{L}_{k-1}$ which consists of n_B disjoint rectangles $R_t (t = 1, \dots, n_B)$:

$$\mathcal{P} = \{R_1, \dots, R_{n_B}\} \quad (3-13)$$

where $\mathcal{L}_{k-1} = \bigcup_{t=1}^{n_B} R_t$ and $R_i \cap R_j = \emptyset$ for $i \neq j$.

To find the best partition, it is assumed in BSP that (i) the prior density of a partition is proportional to $\exp(-\lambda n_B)$ and (ii) the probability of a sample in any rectangle R_t follows the Dirichlet distribution with parameter $(\alpha, \dots, \alpha)$ where both λ and α are constants. Given samples $\{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}}$, the posterior probability of $\mathcal{P}$ can be written as:

$$P(\mathcal{P} | \{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}}) \propto \exp(-\lambda n_B) \frac{D(N_1 + \alpha, \dots, N_{n_B} + \alpha)}{D(\alpha, \dots, \alpha)} \prod_{t=1}^{n_B} \frac{1}{|R_t|^{N_t}} \quad (3-14)$$

where N_t is the sample count in R_t , $|R_t|$ is the volume of R_t , $D(\delta_1, \dots, \delta_{n_B}) = \prod_{t=1}^{n_B} \Gamma(\delta_t) / \Gamma(\sum_{t=1}^{n_B} \delta_t)$ in which $\Gamma(\cdot)$ denotes the gamma function. BSP further defines the score of a partition as $\log(P(\mathcal{P} | \{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}}))$.

BSP indirectly generates samples of partitions by drawing weighted samples of binary paths using sequential importance sampling. In detail, assume that there are m samples of binary paths $\{g_t^{(1)}, \dots, g_t^{(m)}\}$ in the level t . For a path $g_t = (d_1, \dots, d_t)$, d_i ($i = 1, \dots, t$) represents a binary cut that decides which rectangle to cut and along the midpoint of which dimension to cut in the selected rectangle. As a result, a level- t path results in t rectangles. Since a path g_t generates a partition $\mathcal{P}(g_t)$ by sequential binary cuts, m path samples lead to m partitions in each level. Binary cutting continues until the partition with the highest score is found.

Assume the best partition $\mathcal{P}^* = \{R_1^*, \dots, R_{n_B}^*\}$ has been found in the level n_B . A piecewise-constant estimator of $p_k(\boldsymbol{\varphi} | F)$ can be given by:

$$\hat{p}_k(\boldsymbol{\varphi} | F) = \sum_{t=1}^{n_B} \hat{p}_{k,t}(\boldsymbol{\varphi}) \mathbb{I}_{R_t^*}(\boldsymbol{\varphi}), \quad \boldsymbol{\varphi} \in \{\boldsymbol{\varphi}\}_{F, \mathcal{L}_{k-1}} \quad (3-15)$$

where $\mathbb{I}_{R_t^*}(\cdot)$ is an indicator function which is one when $\boldsymbol{\varphi} \in R_t^*$ and zero otherwise; $\hat{p}_{k,t}(\boldsymbol{\varphi}) = N_t / (N_s |R_t^*|)$ is the density value over the rectangle R_t^* . Like the smoothed

histogram, the BSP piecewise-constant estimator can be smoothed by using regression methods.

3.4 Implementation

3.4.1 Flowchart

The flowchart for the proposed framework DOPADS is illustrated in Figure 3-4.

Initialization: assign uniform distributions to design variables $\boldsymbol{\varphi}$ whose distribution intervals are the ranges of $\boldsymbol{\varphi}$.

Step 1: In this step, the design space is partitioned iteratively into a number of subspaces and the density function in each subspace is estimated accordingly. In the initial iteration $k = 0$, failure samples in Ω_{φ} are generated by stochastic-sampling-based reliability methods; an estimator of the density function $\hat{p}_0(\boldsymbol{\varphi}|F)$ is constructed using density estimation techniques as described in Section 3.3.3; then the subspace for the next iteration $\mathcal{L}_0$ is identified with the partition density value p_0^* . For iterations $k \geq 1$, first, additional failure samples in $\mathcal{L}_{k-1}$ are generated by modified Metropolis-Hastings algorithm as described in Section 3.3.2; then, density estimation and subspace identification are performed in the same way as the first iteration. In each iteration, the stopping criterion will be checked until satisfied. Refer to Section 3.4.4 for details about the stopping criteria.

Step 2: Once the iterative procedure in Step 1 stops, the approximated FPF $\hat{P}_F(\boldsymbol{\varphi})$ can be given in terms of $\tilde{p}(\boldsymbol{\varphi}|F)$, $\hat{P}(F)$ and $p_{\Phi}(\boldsymbol{\varphi})$ using Equation (3-1). Specifically, Equation (3-7) gives the density function over the entire space $\tilde{p}(\boldsymbol{\varphi}|F)$ based on the estimator in each iteration; $\hat{P}(F)$ can be estimated using MSC or SS in the initial iteration in Step 1; and $p_{\Phi}(\boldsymbol{\varphi})$ is a constant for uniform distributions specified in initialization.

Step 3: since $\hat{P}_F(\boldsymbol{\varphi})$ has been expressed as an explicit function of $\boldsymbol{\varphi}$, the feasible design region of the reliability constraint $P_F(\boldsymbol{\varphi}) \leq [P_F]$ can be determined explicitly. Thus, the original RBDO problem is decoupled into a conventional constrained nonlinear optimization problem, which can be solved efficiently by many standard algorithms (Snyman 2005), such as Sequential Quadratic Programming (SQP), Trust Region Reflective Algorithm (TRRA), Active Set Algorithm (ASA), etc. In the content of DOPADS, SQP is used for optimization. Refer to (Boggs and Tolle 2000, Boyd and Vandenberghe 2004) for

further details of SQP. Note that when multiple reliability constraints are involved, FPF for each reliability constraint is calculated separately.

Figure 3-4: Flowchart of design optimization in the partitioned design space (DOPADS)

3.4.2 Sample size in each iteration

The size of failure samples N_s is critical to the implementation of the proposed method in terms of efficiency and accuracy. A ‘sufficient’ number of failure samples should be generated in each iteration in order to guarantee the accuracy of density estimation. The ‘sufficiency’ depends on the density function structure and the accuracy requirement. In this section, the sample size is set to $N_s = 2000$ for all three illustrative examples, which gives good accuracy with acceptable computational burden.

3.4.3 Partition density values

As discussed in Section 3.3.1, users can predefine either the partition density value p_k^* or the probability ratio $P(\mathcal{L}_k|F \cap \mathcal{L}_{k-1})$ (or $P(\mathcal{L}_0|F)$ for the initial iteration), which is denoted as P_k^* for cleanness, since either p_k^* or P_k^* can separate the high-density region and low-density one. If p_k^* is predefined, P_k^* can be estimated by $N_{k,low}/N_s$ where $N_{k,low}$ is the number of failure samples in $\mathcal{L}_k$ for the k -th iteration; Otherwise, if P_k^* is predefined, p_k^* can be given by the $100(1 - P_k^*)$ -th percentile of the density values of N_s failure samples. In the illustrative example, the latter one is chosen and set P_k^* to 0.1 for each iteration. To derive p_k^* , the density value for each failure sample is estimated, then their values are sorted in the descending order, and p_k^* is given by the 90-th percentile of density values. In this way, 10% of failure samples in $\mathcal{L}_k$ will be in the next subspace.

3.4.4 Stopping criterion

One can stop partitioning in Step 1 when the feasible region with respect to the reliability constraints is included by the union of the high-density regions. That is, $p_k^* \leq \frac{[P_F]p_{\Phi}(\boldsymbol{\varphi})}{\hat{P}(F) \prod_{t=0}^{k-1} P_t^*}$, equivalent to $[P_F] \geq P_F(\boldsymbol{\varphi}^*) = \frac{p_k^* \hat{P}(F) \prod_{t=0}^{k-1} P_t^*}{p_{\Phi}(\boldsymbol{\varphi})}$ where $P_F(\boldsymbol{\varphi}^*)$ is the failure probability value corresponding to p_k^* . It implies that p_k^* is small enough for the problem under consideration and there is no further concern about the region of the design space where the failure probability is too small.

3.5 Illustrative examples

3.5.1 Example 1: A straight column with two design variables

Consider a long straight column with a rectangular cross section shown in Figure 3-5. The cross section of the column is b in width and h in height. The length is L and the Young's modulus is E . The column is fixed at one end and free at the other, subjected to a deterministic axial load F_a at the free end.

The width b and the height h are considered as design variables, i.e., $\boldsymbol{\varphi} = [b, h]^T$ where $b \in [100, 400](mm)$ and $h \in [100, 400](mm)$. The length L and the Young's

modulus E are taken as random variables, i.e., $\boldsymbol{\theta} = [L, E]^T$. The range of design variables and distribution parameters of random variables are given in Table 3-2.

Figure 3-5: Schematic diagram of a straight column considering its buckling failure

Table 3-2: Ranges of design variables and distribution parameters of random variables

Variable	Distribution	Interval	Mean	CoV
b (mm)	deterministic	[100, 400]	--	--
h (mm)	deterministic	[100, 400]	--	--
F_a (kN)	deterministic	1460	--	--
E (MPa)	lognormal	--	10000	0.15
L (mm)	lognormal	--	3000	0.05

The objective function is the total area of the cross section, which is proportional to the weight as well as the cost. In the reliability constraint, the failure is defined as the elastic buckling of the column, which means that the applied load exceeds the critical buckling load. The formulation of the column design optimization is given in Figure 3-6.

$$\begin{aligned}
 &\min_{\boldsymbol{\varphi}}: && c(\boldsymbol{\varphi}) = bh \\
 &\text{s. t. :} && \begin{cases} P_F(\boldsymbol{\varphi}) \leq [P_F] \\ b - h \geq 0 \\ 100\text{mm} \leq b \leq 400\text{mm} \\ 100\text{mm} \leq h \leq 400\text{mm} \end{cases} \\
 &\text{where:} && g(\boldsymbol{\varphi}, \boldsymbol{\theta}) = \frac{\pi^2 E b h^3}{12 L^2} - F_a
 \end{aligned}$$

Figure 3-6: Formulation of the column design optimization problem

After six iterations, the procedure of density function estimation stops as the stopping criterion $p_5^* = 2.4 \times 10^{-4} \leq 7.4 \times 10^{-4}$ is satisfied. As shown in Figure 3-7, the entire design space is partitioned into seven regions with corresponding partition density values. In each subspace, the density values of support points are calculated, and the density functions are approximated. Finally, the approximated failure probability surface for the entire design space is calculated as illustrated in Figure 3-8.

Figure 3-7: Partitioning of the design space and partition density values

Figure 3-8: Approximated surface of the density function

To verify the proposed framework, the crude MCS is used to approximate the failure probability. In the crude MCS, the design space is uniformly meshed, and the failure probability of each grid point is approximated by performing the direct MCS. For this example, we set 31 grid points for both b and h , i.e., $b = 100, 110, \dots, 400(mm)$ and $h = 100, 110, \dots, 400(mm)$. The contour plot of the failure probability computed by DOPADS and the crude MCS are depicted in Figure 3-9. Also, profiles of the failure probability surface ($b = 100, 250, 400mm$) are depicted in Figure 3-10. Results show that DAPADS has good accuracy compared to the crude MCS.

Since the computational burden of optimization is negligible after decoupling, the efficiency of DOPADS can be measured by the total number of model evaluations N_{eval} . In this example, $N_{eval} = 23100$ for DOPADS and $N_{eval} = 2.21 \times 10^9$ for the crude MCS, which indicates the high efficiency of the proposed framework.

Figure 3-9: Contour plot of the failure probability function

Figure 3-10: Profiles of the failure probability function for $b = 100, 250, 400\text{mm}$

The constrained nonlinear optimization algorithm SQP is used in the proposed method to find the optimal solution. Five cases with different allowable failure probabilities ($[P_F] = 0.1, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$) are considered. Finally, the optimal solutions $\boldsymbol{\varphi}^*$ and the corresponding objective function values $c(\boldsymbol{\varphi}^*)$ are compared to those obtained by the crude MCS. Figure 3-11 and Table 3-3 show that results of the proposed method are consistent to those of the crude MCS. Note that the optimal solutions of the crude MCS do not mean the true optimal as their accuracy is determined by the resolution of design space meshing.

Figure 3-11: Optimal solutions and objective function values for different allowable failure probabilities

Table 3-3: Optimal solutions and objective function values for different allowable failure probabilities

$[P_F]$	DOPADS		Crude MCS	
	$\boldsymbol{\varphi}^*$	$c(\boldsymbol{\varphi}^*)$	$\boldsymbol{\varphi}^*$	$c(\boldsymbol{\varphi}^*)$
0.1	(209.6, 209.6)	43932	(210.0, 210.0)	44100
10^{-2}	(219.1, 219.1)	48004	(220.0, 220.0)	48400
10^{-3}	(223.2, 223.2)	49818	(210.0, 240.0)	50400

10^{-4}	(230.9, 230.9)	53315	(230.0, 230.0)	52900
10^{-5}	(237.4, 237.4)	56358	(230.0, 240.0)	55200

3.5.2 Example 2: A ten-bar truss with three design variables and multiple reliability constraints

Consider the design of a ten-bar cantilever truss shown in Figure 3-12. The cross-sectional areas of horizontal, vertical, and diagonal members are A_1 , A_2 and A_3 , respectively, which are considered as design variables. All the members are made of aluminum with Young's modulus E . The length of horizontal and vertical members is L . The truss is subjected to two vertical loads P_1 and P_2 at Node 5 and Node 4, respectively. In this example, E , P_1 and P_2 are assumed to be random. In summary, design variables $\boldsymbol{\varphi} = [A_1, A_2, A_3]^T$ and random variables $\boldsymbol{\theta} = [E, P_1, P_2]^T$ are considered. The range of design variables and distribution parameters of random variables are given in Table 3-4.

Figure 3-12: Schematic diagram of a ten-bar cantilever truss

Table 3-4: Ranges of design variables and distribution parameters of random variables for a ten-bar cantilever truss

Variable	Distribution	Interval	Mean	CoV
A_1 (mm^2)	deterministic	[6000, 10000]	--	--
A_2 (mm^2)	deterministic	[1000, 3000]	--	--
A_3 (mm^2)	deterministic	[5000, 9000]	--	--
E (MPa)	lognormal	--	69000	0.10
P_1 (kN)	lognormal	--	300	0.20
P_2 (kN)	lognormal	--	300	0.20

In this example, the total volume of the truss is minimized with the displacement and stress constraints (see the formulation in Figure 3-13). Three failure modes are defined as the displacement of Node 4 exceeding the allowable tip displacement, or the stress of Bar 3 and Bar 7 exceeding the allowable stress of aluminum. Since different reliability levels are required for serviceability (i.e., displacement) and ultimate (i.e., stress) limit state, the allowable failure probability is 10^{-1} and 10^{-3} , respectively.

$$\begin{aligned}
& \min_{\boldsymbol{\varphi}}: & c(\boldsymbol{\varphi}) &= (4A_1 + 2A_2 + 4\sqrt{2}A_3)L \\
& \text{s. t. :} & \begin{cases} P_{Fd}(\boldsymbol{\varphi}) \leq [P_{Fd}] = 10^{-1} \\ P_{F\sigma 3}(\boldsymbol{\varphi}) \leq [P_{F\sigma 3}] = 10^{-3} \\ P_{F\sigma 7}(\boldsymbol{\varphi}) \leq [P_{F\sigma 7}] = 10^{-3} \\ 6000mm^2 \leq A_1 \leq 10000mm^2 \\ 1000mm^2 \leq A_2 \leq 3000mm^2 \\ 5000mm^2 \leq A_3 \leq 9000mm^2 \end{cases} \\
& \text{where:} & \begin{cases} g_d(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [d] - d_{n4}, [d] = 80mm \\ g_{\sigma 3}(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [\sigma] - \sigma_{b3}, [\sigma] = 140MPa \\ g_{\sigma 7}(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [\sigma] - \sigma_{b7} \end{cases}
\end{aligned}$$

Figure 3-13: Formulation of the ten-bar truss design optimization problem

The density functions were estimated separately for three failure modes. The partition density values and the number of support points are shown in Table 3-5. Since the failure probability of the three failure modes is not sensitive to the cross-sectional areas of vertical members A_2 , we visualize the failure probability contours in two-dimensional design space of A_1 and A_3 by setting $A_2 = 2000mm^2$, as shown in Figure 3-14. Again, the crude

MCS is used to verify DOPADS. The numbers of grid points for design variables A_1, A_2, A_3 are 41, 21, 41, respectively. It can be seen that the contours obtained by the proposed framework are consistent to those obtained by the crude MCS. In this example, the total number of model evaluations for the three failure modes is $N_{eval} = 37600$ in DOPADS and $N_{eval} = 6.81 \times 10^{11}$ in the crude MCS, indicating the efficiency of the proposed framework highly outperforms the crude MCS.

Table 3-5: Partition density values and the number of support points for three failure modes

Failure mode	Iteration	p_i^*	Number of support points
Mode 1	0	0.62	1097
	1	0.60	903
Mode 2	0	0.08	467
	1	0.07	782
	2	0.25	751
Mode 3	0	5.64×10^{-2}	619
	1	9.19×10^{-2}	667
	2	8.01×10^{-2}	488
	3	8.25×10^{-2}	226

Figure 3-14: Contour plots of the failure probability functions for the three failure modes ($A_2 = 2000mm^2$)

The feasible region of the design space as well as the optimal solution is identified after approximating the failure probability functions of the three failure modes, as shown in Figure 3-15. The optimal solution by the proposed method is $\boldsymbol{\varphi}^* = (7595.3, 1000, 8303.4)$ with the object function value $c(\boldsymbol{\varphi}^*) = 7.1 \times 10^5$, which is close to the result obtained by the crude MCS $\boldsymbol{\varphi}^* = (7600, 1000, 8300)$ with the objective function value $c(\boldsymbol{\varphi}^*) = 7.1 \times 10^5$.

Figure 3-15: Optimal solution and the feasible region for the ten-bar truss example

3.5.3 Example 3: A ten-story shear building with a large number of random variables

A ten-story planar linear shear building subjected to earthquake excitations is considered as shown in Figure 3-16. This example is modified from (Ching and Hsieh 2007a) as an illustrative example of dynamic systems with a large number of random variables. Assume that the values of the lumped mass are $m_{1-3} = 3.0\lambda_m \times 10^6 \text{ kg}$, $m_{4-6} = 2.5\lambda_m \times 10^6 \text{ kg}$, and $m_{7-10} = 2.0\lambda_m \times 10^6 \text{ kg}$ for the first three stories, story 4 to story 6, and the top four stories, respectively, where $\lambda_m \in [0.5, 2.0]$ is the mass scaling parameter; the values of the inter-story stiffness are $k_{1-3} = 5\lambda_k \times 10^6 \text{ N/m}$, $k_{4-6} = 3\lambda_k \times 10^6 \text{ N/m}$, $k_{7-10} = \lambda_k \times 10^6 \text{ N/m}$, where $\lambda_k \in [0.5, 2.0]$ is the stiffness scaling parameter; and the classical modal damping of each story is also an independent Gaussian random variable with mean value 0.04 and CoV 0.20. In this example, the design variables are the scalar parameters $\boldsymbol{\varphi} = [\lambda_m, \lambda_k]^T$.

The duration of earthquake ground motion is 30 seconds with the sampling interval $\Delta t = 0.01\text{s}$, and the total number of random variables is 3011. A stochastic model, referred

to as the modulated filtered white noise, is used to describe the ground motion. The ground acceleration $a(t)$ is:

$$a(t) = [\omega_1^2 \quad 2\omega_1\xi_1 \quad -\omega_2^2 \quad -\omega_2\xi_2] \begin{bmatrix} v_1 \\ \dot{v}_1 \\ v_2 \\ \dot{v}_2 \end{bmatrix} \quad (3-16)$$

where

$$\frac{d}{dt} \begin{bmatrix} v_1 \\ \dot{v}_1 \\ v_2 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\omega_1^2 & -2\omega_1\xi_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \omega_2^2 & 2\omega_2\xi_2 & -\omega_2^2 & -\omega_2\xi_2 \end{bmatrix} \begin{bmatrix} v_1 \\ \dot{v}_1 \\ v_2 \\ \dot{v}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ w(t) \\ 0 \\ 0 \end{bmatrix} \quad (3-17)$$

The parameters of the filter are deterministic, $\omega_1 = 15.0 \text{ rad/s}$, $\xi_1 = 0.8$, $\omega_2 = 0.3 \text{ rad/s}$, $\xi_2 = 0.995$. $w(t)$ denotes a Gaussian white noise with the autocorrelation function $[w(t)w(t+\tau)] = I\delta(\tau)h^2(t)$, where I is the intensity of the white noise, and

$$h(t) = \begin{cases} 0 & t \leq 0 \\ \frac{t}{2} & 0s < t \leq 2s \\ 1 & 2s < t \leq 10s \\ \exp[-0.1(t-10)] & t > 10s \end{cases} \quad (3-18)$$

The objective of this RBDO problem is to minimize the sum of the stiffness of ten stories. Failure of the building is defined as the exceedence of maximum inter-story drift Δd over 0.06m during the earthquake ground motion and the allowable probability of failure $[P_F]$ is 0.01 (see the formulation in Figure 3-17).

Figure 3-16: Schematic diagram of the ten-story shear building and a sample time history of the ground acceleration

$$\begin{array}{ll}
 \min_{\boldsymbol{\varphi}}: & c(\boldsymbol{\varphi}) = 3k_{1-3} + 3k_{4-6} + 4k_{7-10} \\
 & = 28\lambda_k \times 10^6 N/m \\
 \text{s. t. :} & \begin{cases} P_F(\boldsymbol{\varphi}) \leq [P_F] = 10^{-2} \\ 0.5 \leq \lambda_m \leq 2.0 \\ 0.5 \leq \lambda_k \leq 2.0 \end{cases} \\
 \text{where:} & g(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [\Delta d] - \Delta d, [\Delta d] = 60mm
 \end{array}$$

Figure 3-17: Formulation of the ten-story shear building optimization problem

After three iterations, the procedure of density function estimation stops as the stopping criteria $p_2^* = 1.20 \leq 1.48$ is satisfied. Figure 3-18 illustrates the design space is partitioned into four subspaces by corresponding partition density values. In each subspace, the failure probability function is fitted using failure probability values of support points. The contour plot of the failure probability function is shown in Figure 3-19.

To verify the proposed framework, the contour plot of DOPADS is compared with that of SS. SS is used in a brute-force way where the design space is uniformly meshed with 16 grid points in each direction, generating a total number of 256 combinations of design configurations. As shown in Figure 3-19, the contour lines of the proposed method and the brute-force SS are close to each other. Therefore, the optimal solution $\boldsymbol{\varphi}^*$, which in this case are the minimal λ_k satisfying $P_F(\boldsymbol{\varphi}) \leq 10^{-2}$ obtained by the proposed method and brute-force SS, are (0.50,0.55) and (0.50,0.56) respectively. The total number of model evaluations $N_{eval} = 15400$ in DOPADS, which is much smaller than $N_{eval} = 8.04 \times 10^5$ in the brute-force SS.

Figure 3-18: Partitioning of the design space and its partition density values

Figure 3-19: Contour plot of the failure probability function

3.6 Conclusions

In this chapter, an efficient decoupling framework DOPADS for RBDO problems based on the approximation of the FPF is presented. The novel idea behind the proposed framework is that instead of approximating FPFs over the entire design space by calculating the failure

probability for different design configurations repeatedly, one can iteratively partition the design space into several subspaces based on the distribution of failure samples in each subspace. In this way, failure samples are propagated into regions with small failure probability and the global accuracy of the approximate FPF can be improved.

Numerical results of three illustrative examples show that the proposed framework is advantageous in terms of efficiency and accuracy even if a large number of random variables are involved. DOPADS can be applied to different types of RBDO problems including those involving multiple reliability constraints, small allowable failure probability, sophisticated numerical models and high dimensional uncertain parameter space. Moreover, due to readily approximated FPFs, a deterministic optimization procedure can be performed seamlessly even if the allowable failure probabilities are changed as shown in the first illustrative example. But it should be noted that despite the advanced density estimation techniques adopted, the number of design variables in DOPADS is usually less than five due to the curse of dimensionality.

4 Enhancing the efficiency of design optimization for stochastic complex systems with multiple reliability constraints

Summary

It is common that the design of a system involves several different reliability constraints. This section presents a new efficient framework DOPADS_m, which is an extension to DOPADS, to deal with RBDO problems subjected to multiple reliability constraints. It can approximate different failure probability functions simultaneously by taking advantage of the dependency between failure events, as opposed to tediously repeating the approximation for each reliability constraints in DOPADS. The efficiency enhancement is remarkable as demonstrated in two illustrative examples.

4.1 Introduction

The reliability constraints represent a variety of design requirements considering the uncertainties in a system. For example, an earth-retaining system of a deep excavation has a limit on the maximal deflection of the earth-retaining wall, a limit on the maximal ground settlement, stability requirements that ensure the excavation will not collapse during the excavation activity, etc. Therefore, it is not uncommon that the system under study is subjected to multiple reliability constraints.

In DOPADS, the way to deal with multiple reliability constraints is straightforward: FPFs are approximated separately for different reliability constraints as shown in the second

illustrative example in Section 3.5. As a result, the computational complexity increases linearly with the number of reliability constraints, probably leading to prohibitive computational burden especially when expensive-to-evaluate models are involved.

One popular strategy to alleviate the computational burden may be the use of surrogates (also referred to as metamodels or response surfaces) in lieu of high fidelity numerical models (Youn and Choi 2004, Dubourg et al. 2011). Despite the efficiency in the evaluation of a surrogate, the computational investment in constructing a surrogate may outweigh the corresponding gain for cases with many model parameters. Additionally, the effect of the modeling uncertainty in the surrogate on the optimal solution is difficult to quantify. In the scope of this thesis, surrogates are not included.

This chapter proposed a new framework DOPADS_m for design optimization with emphasis on the enhancement of efficiency. In contrast to the previous work where FPFs are approximated separately for different reliability constraints, this paper can evaluate all the reliability constraints simultaneously based on the distribution of failure samples. To ensure the accuracy of the approximated FPF, failure samples are propagated iteratively from the high-density regions to low-density ones in order to obtain the distribution information therein. The novelty of this chapter lies in the way failure samples are propagated. That is, in each iteration, it is in the union of low-density regions of all failure events where additional failure samples are drawn for the next iteration. The similar idea of propagating samples into the union of individual target events was discussed in (Li et al. 2015, Bansal and Cheung 2017) in the context of reliability analysis for multiple performance functions. As we shall see, the required size of failure samples in each iteration is determined by the overlapping of individual low-density regions. Therefore, the efficiency of the FPF approximation can be improved as long as the dependence exists between different failure events in reliability constraints.

The rest of this section is organized as follows. Section 4.2 presents the proposed DOPADS_m with focus on both fundamentals and implementational details. Section 4.3 summarizes the implementation flowchart of DOPADS_m. The efficiency enhancement of the proposed method in dealing with multiple reliability constraints is demonstrated in Section 4.4 by two illustrative examples. Conclusions are summarized in Section 4.5.

4.2 Propagating failure samples for different failure events simultaneously

4.2.1 Basic idea

Consider a RBDO problem formulated in Equation (2-16) with n_{rc} reliability constraints. The aim of the proposed framework is to approximate $P_{F_I}(\boldsymbol{\varphi})$ ($I = 1, \dots, n_{rc}$) efficiently without loss of accuracy. It is achieved based on the idea that failure samples for different failure events F_I ($I = 1, \dots, n_{rc}$) can be propagated simultaneously over regions of interest in the design space due to the overlapping of failure regions.

As discussed in Section 3, since $P_{F_I}(\boldsymbol{\varphi})$ is proportional to $p(\boldsymbol{\varphi}|F_I)$, given the constants $P(F_I)$ and $p_{\Phi}(\boldsymbol{\varphi})$ in Equation (3-1), the problem is converted into how to approximate $p(\boldsymbol{\varphi}|F_I)$ simultaneously. Again, to ensure the accuracy of the final estimators $\tilde{p}(\boldsymbol{\varphi}|F_I)$ ($I = 1, \dots, n_{rc}$), two strategies are adopted: (1) iteratively propagate failure samples from high-density regions into low-density ones; (2) increase the sample size to N_s at the beginning of iteration k .

Assume that at the beginning of the initial iteration, $N_{b,0}$ failure samples generated by reliability analysis are scattering over the original design space Ω_{φ} . The samples, denoted as $\{\boldsymbol{\varphi}\}_{\cup F_I}$ hereafter, follow the distribution $p(\boldsymbol{\varphi}|\cup_{I=1}^{n_{rc}}(F_I \cap \Omega_{\varphi}))$. It is easy to observe that, if samples $\{\boldsymbol{\varphi}\}_{\cup F_I}$ are distributed as $p(\boldsymbol{\varphi}|\cup_{I=1}^{n_{rc}}(F_I \cap \Omega_{\varphi}))$, then $\{\boldsymbol{\varphi}\}_{F_I}$ will follow the distribution $p(\boldsymbol{\varphi}|F_I \cap \Omega_{\varphi})$ as $(F_I \cap \Omega_{\varphi}) \subseteq \cup_{I=1}^{n_{rc}}(F_I \cap \Omega_{\varphi})$ always holds. If $N_{b,0} < N_0$, additional $(N_0 - N_{b,0})$ samples should be simulated using stochastic sampling algorithms, e.g., modified Metropolis-Hastings algorithm (Au and Beck 2001a) as discussed in Section 3.3.2. For the failure event F_I ($I = 1, \dots, n_{rc}$), $N_{0,I}$ out of N_0 failure samples $\{\boldsymbol{\varphi}\}_{F_I}$ can be used to estimate $\hat{p}_0(\boldsymbol{\varphi}|F_I \cap \Omega_{\varphi})$ and the design space Ω_{φ} can be partitioned into the high-density part $\mathcal{H}_{0,I}$ and the low-density part $\mathcal{L}_{0,I}$, i.e., $\Omega_{\varphi} = \mathcal{H}_{0,I} \cup \mathcal{L}_{0,I}$:

$$\mathcal{H}_{0,I} = \{\boldsymbol{\varphi} | \hat{p}_0(\boldsymbol{\varphi}|F_I \cap \Omega_{\varphi}) > p_{0,I}^*\} \quad (4-1a)$$

$$\mathcal{L}_{0,I} = \{\boldsymbol{\varphi} | \hat{p}_0(\boldsymbol{\varphi}|F_I \cap \Omega_{\varphi}) \leq p_{0,I}^*\} \quad (4-1b)$$

where $p_{0,I}^*$ is the partition density value for the failure event F_I in the initial iteration. Little distribution information about the low-density regions is available for each failure

event. To propagate samples, the low-density region $\mathcal{L}_0$ of this iteration is defined as the union of low-density regions

$$\mathcal{L}_0 = \bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{0,I}) \quad (4-2)$$

In iteration k ($k = 1, \dots, n_{it}$), at the beginning $N_{b,k}$ failure samples $\{\boldsymbol{\varphi}\}_{\cup F_I}$ from the previous iteration scatter over the previous low-density region $\mathcal{L}_{k-1}$, which follow the distribution $p(\boldsymbol{\varphi} | \bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k-1,I}))$. Again, we start with the increase of the sample size to N_k , which is followed by the estimation of $\hat{p}_k(\boldsymbol{\varphi} | F_I \cap \mathcal{L}_{k-1,I})$ based on the subset $\{\boldsymbol{\varphi}\}_{F_I}$ of entire failure sample set $\{\boldsymbol{\varphi}\}_{\cup F_I}$ for each failure event (see Section 3.3.3 for details of multivariate density estimation). The previous low-density region $\mathcal{L}_{k-1,I}$ for the failure event F_I is partitioned into two parts $\mathcal{H}_{k,I}$ and $\mathcal{L}_{k,I}$

$$\mathcal{H}_{k,I} = \{\boldsymbol{\varphi} | \hat{p}_k(\boldsymbol{\varphi} | F_I \cap \mathcal{L}_{k-1,I}) > p_{k,I}^*\} \quad (4-3a)$$

$$\mathcal{L}_{k,I} = \{\boldsymbol{\varphi} | \hat{p}_k(\boldsymbol{\varphi} | F_I \cap \mathcal{L}_{k-1,I}) \leq p_{k,I}^*\} \quad (4-3b)$$

where $p_{k,I}^*$ is the partition density value for the failure event F_I in the iteration k . The low-density region of this iteration can be expressed as:

$$\mathcal{L}_k = \bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k,I}) \quad (4-4)$$

After the stopping criterion is satisfied, the final estimators $\hat{p}(\boldsymbol{\varphi} | F_I)$ can be estimated by assembling the density estimators in each subspace

$$\hat{p}(\boldsymbol{\varphi} | F_I) = \sum_{k=0}^{n_{it}} \hat{p}(\boldsymbol{\varphi} | F_I \cap \mathcal{H}_{k,I}) P(\mathcal{H}_{k,I} | F_I) + \hat{p}(\boldsymbol{\varphi} | F_I \cap \mathcal{L}_{n_{it},I}) P(\mathcal{L}_{n_{it},I} | F_I) \quad (4-5)$$

where

$$\hat{p}(\boldsymbol{\varphi} | F_I \cap \mathcal{H}_{0,I}) P(\mathcal{H}_{0,I} | F_I) = \begin{cases} \hat{p}_0(\boldsymbol{\varphi} | F_I), & \boldsymbol{\varphi} \in \mathcal{H}_{0,I} \\ 0, & \boldsymbol{\varphi} \notin \mathcal{H}_{0,I} \end{cases} \quad (4-6a)$$

$$\begin{aligned} & \hat{p}(\boldsymbol{\varphi} | F_I \cap \mathcal{H}_{k,I}) P(\mathcal{H}_{k,I} | F_I) \\ &= \begin{cases} \hat{p}_k(\boldsymbol{\varphi} | F_I) P(\mathcal{L}_{k-1,I} | F_I), & \boldsymbol{\varphi} \in \mathcal{H}_{k,I} \\ 0, & \boldsymbol{\varphi} \notin \mathcal{H}_{k,I} \end{cases} \quad (k = 1, \dots, n_{it}) \end{aligned} \quad (4-6b)$$

$$\hat{p}(\boldsymbol{\varphi} | F_I \cap \mathcal{L}_{n_{it},I}) P(\mathcal{L}_{n_{it},I} | F_I) = \begin{cases} \hat{p}_{n_{it}}(\boldsymbol{\varphi} | F_I) P(\mathcal{L}_{n_{it}-1,I} | F_I), & \boldsymbol{\varphi} \in \mathcal{L}_{n_{it},I} \\ 0, & \boldsymbol{\varphi} \notin \mathcal{L}_{n_{it},I} \end{cases} \quad (4-6c)$$

Figure 4-1 compares the approximation of FPF's in DOPADS and DOPADS_m for cases with two reliability constraints ($n_{rc} = 2$) in the iteration k . In Figure 4-1(a), failure samples are propagated independently for the two failure events. However, as shown in Figure 4-1(b), there exists failure samples in $\{\boldsymbol{\varphi}\}_{\cup F_I}$ that satisfy $(F_1 \cap \mathcal{L}_{k-1,1}) \cap (F_2 \cap$

$\mathcal{L}_{k-1,2}$) as long as $\mathcal{L}_{k-1,1}$ and $\mathcal{L}_{k-1,2}$ have overlapping regions and failure events F_1 and F_2 are dependent. Note that the first condition is met in most situations. This kind of failure samples contain the distribution information about both failure events and can be used for the estimation of $p_k(\boldsymbol{\varphi}|F_1 \cap \mathcal{L}_{k-1,1})$ and $p_k(\boldsymbol{\varphi}|F_2 \cap \mathcal{L}_{k-1,2})$.

Figure 4-1: Comparison of the approximation of failure probability functions in DOPADS and DOPADS_m

4.2.2 Sample size in each iteration

Like DOPADS, the efficiency of DOPADS_m is evaluated in terms of the total number of model evaluations N_{eval} . It reflects the efficiency well because compared with the simulation of samples, which is the most computationally expensive part in DOPADS_m

when sophisticated numerical models are involved, the computation time of other tasks such as the deterministic optimization is negligible.

To achieve a certain level of accuracy, in iteration k ($k = 0, \dots, n_{it}$), N_k failure samples distributed as $p(\boldsymbol{\varphi} | \bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k-1,I}))$ (or $p(\boldsymbol{\varphi} | \bigcup_{I=1}^{n_{rc}} (F_I \cap \Omega_\varphi))$ for $k = 0$) are used for density estimation after the generation of additional samples, among which $N_{k,I}$ failure samples are distributed as $p(\boldsymbol{\varphi} | F_I \cap \mathcal{L}_{k-1,I})$ (or $p(\boldsymbol{\varphi} | F_I \cap \Omega_\varphi)$ for $k = 0$). The sample size N_k is determined in an adaptive way such that there are at least N_s failure samples for each failure event, i.e., $N_{k,I} \geq N_s, \forall I = 1, \dots, n_{rc}$. Note that

$$P\left(F_I \cap \Omega_\varphi \middle| \bigcup_{I=1}^{n_{rc}} (F_I \cap \Omega_\varphi)\right) \approx \frac{N_{0,I}}{N_0} \quad (4-7a)$$

$$\begin{aligned} P\left(F_I \cap \mathcal{L}_{k-1,I} \middle| \bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k-1,I})\right) &= \frac{P(F_I \cap \mathcal{L}_{k-1,I})}{P(\bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k-1,I}))} \\ &\approx \frac{N_{k,I}}{N_k} \quad (k = 1, \dots, n_{it}) \end{aligned} \quad (4-7b)$$

Thus, the following inequality should be satisfied:

$$N_k \geq \rho_{k,I} N_s, \forall I = 1, \dots, n_{rc} \quad (4-8)$$

where $\rho_{k,I} = P(\bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k-1,I})) / P(F_I \cap \mathcal{L}_{k-1,I})$ (or $P(\bigcup_{I=1}^{n_{rc}} (F_I \cap \Omega_\varphi)) / P(F_I \cap \Omega_\varphi)$ for $k = 0$).

As $\rho_{k,i}$ may be different for different failure events, we set N_k to be $\max(\rho_{k,I}) N_s$ to make sure inequality (4-8) always holds true. It is easy to observe that $\rho_{k,I} \in [1, n_{rc}]$ whose value is determined by the dependence between $F_I \cap \mathcal{L}_{k-1,I}$ and $\bigcup_{I=1}^{n_{rc}} (F_I \cap \mathcal{L}_{k-1,I})$. Therefore, the total sample size N_k is bounded by:

$$N_s \leq N_k \leq n_{rc} N_s \quad (4-9)$$

It implies that the more dependence exists between $F_I \cap \mathcal{L}_{k-1,I}$, the more efficiency we will gain. Also, in the worst case where there is no dependence, N_k will equal to $n_{rc} N_s$, indicating that DOPADS_m degrades to DOPADS.

4.2.3 Partition density values

As shown in Equation (4-1) and Equation (4-3), in iteration k ($k = 1, \dots, n_{it}$), the previous low-density region $\mathcal{L}_{k-1,I}$ (or Ω_φ for $k = 0$) for the failure event F_I is partitioned into parts $\mathcal{H}_{k,I}$ and $\mathcal{L}_{k,I}$ by the partition density value $p_{k,I}^*$. The partition density $p_{k,I}^*$ is

determined to ensure that the probability for failure samples in $\mathcal{L}_{k-1,I}$ to scatter in $\mathcal{L}_{k,I}$ is a constant P^*

$$P(\mathcal{L}_{k,I}|F_I \cap \mathcal{L}_{k-1,I}) = P^*, \forall I = 1, \dots, n_{rc} \text{ and } \forall k = 1, \dots, n_{it} \quad (4-10)$$

For $k = 0$, $P(\mathcal{L}_{0,I}|F_I \cap \Omega_\varphi) = P^*$.

To implement it, assume that the density estimator is $\hat{p}_k(\boldsymbol{\varphi}|F_I \cap \mathcal{L}_{k-1,I})$ and failure samples $\{\boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(2)}, \dots, \boldsymbol{\varphi}^{(N_{k,I})}\}$ are scattering over $\mathcal{L}_{k-1,I}$. First, we calculate the density values for all $N_{k,I}$ failure samples, which is denoted as $\{\hat{p}_k(\boldsymbol{\varphi}^{(1)}|F_I \cap \mathcal{L}_{k-1,I}), \dots, \hat{p}_k(\boldsymbol{\varphi}^{(N_{k,I})}|F_I \cap \mathcal{L}_{k-1,I})\}$. Then sort them in descending order. In this way, the partition density value $p_{k,I}^*$ is given by the $100(1 - P^*)$ percentile of the sorted density values. As a result, the low-density region $\mathcal{L}_{k,I}$ in this iteration contains about $100P^*\%$ of the total $N_{k,I}$ failure samples. To calculate the final estimator, in Equation (4-5), the probability of failure samples in $\mathcal{L}_{k-1,I}$ can be written as:

$$P(\mathcal{L}_{k-1,I}|F_I) = (P^*)^k, \forall k = 1, \dots, n_{it} \quad (4-11)$$

4.2.4 Stopping criterion

For a failure event F_I ($I = 1, \dots, n_{rc}$), it exits from the iteration procedures when the failure probability $P_{F_I}^*$ corresponding to the partition density value $p_{n_{it},I}^*$ is smaller than the allowable failure probability of this failure event

$$P_{F_I}^* = \frac{p_{n_{it},I}^*(P^*)^{n_{it}}P(F_I)}{p_{\Phi}(\boldsymbol{\varphi})} \leq [P_{F_I}] \quad (4-12)$$

It means that the union of high-density regions $\bigcup_{k=0}^{n_{it}} \mathcal{H}_{k,I}$ has covered the feasible region of the reliability constraint. Thus, from the perspective of optimization, there is no need to further propagate samples into the low-density region $\mathcal{L}_{n_{it},I}$.

4.3 Summary of implementation

The proposed DOPADS_m can deliberately decouple the reliability analysis from the RBDO problems by approximating FPFs. The steps involved are summarized in the flowchart in Figure 4-2. In the initialization stage, a RBDO problem is formulated by identifying design variables, uncertain model parameters, the objective function, reliability constraints and deterministic constraints. Note that in DOPADS_m, design variables $\boldsymbol{\varphi}$ are

artificially treated as uniformly distributed random variables, i.e., $\boldsymbol{\varphi} \sim \text{uniform}(\boldsymbol{\varphi}^L, \boldsymbol{\varphi}^U)$. Meanwhile, the values of hyperparameters P^* and N_s are assigned.

The procedure of the FPF approximation can be roughly divided into three parts. First, we perform the pilot reliability analysis to calculate the failure probabilities of all failure events of interest, which also generates failure samples over the design space Ω_φ . In the second part, we estimate the density functions in an iterative manner to improve their accuracy. Specifically, in each iteration k ($k = 0, \dots, n_{it}$), using stochastic sampling algorithms such as modified Metropolis-Hastings algorithm, DOPADS_m increases the size of failure samples to N_k . The density distribution of these failure samples is estimated based on multivariate density estimation techniques. Then DOPADS_m further identifies the low-density region for the next iteration. When the iteration procedure stops, density function estimators which describe the density distributions in each subspace are aggregated into the final estimators $\hat{p}(\boldsymbol{\varphi}|F_I)$ ($I = 1, \dots, n_{rc}$). Finally, we scale the final density function estimators to FPFs.

With the help of approximated FPFs, we convert a RBDO problem into the conventional deterministic optimization problem. Again, we use SQP, which has been well implemented and documented in many numerical computing packages (e.g., Optimization Toolbox in Matlab), to perform the deterministic optimization.

Figure 4-2: Flowchart of design optimization in the partitioned design space for problems involving multiple reliability constraints (DOPADS_m). Filled rectangles represent the parts involving the model evaluation which require a large amount of computation time

4.4 Illustrative examples

In this section, we present two illustrative examples to demonstrate the implementational details as well as the efficiency improvement of DOPADS_m compared to the original version DOPADS. In the first example, design of a cantilever beam that has two design variables is studied. Both dynamic and static reliability constraints, whose failure events have different level of dependence, are considered. Despite the simplicity of the model, this two-dimensional example is suitable for the demonstration and visualization purposes. The second example is an engineering example where a shear building subjected to uncertain earthquake excitation is considered. This example consists of five design variables and ten reliability constraints. Moreover, a large number of uncertain model parameters (2000+)

demonstrates the effectiveness and the efficiency of the proposed method in high-dimensional cases.

4.4.1 Example 1: A cantilever beam

This example is adapted from (Sues et al. 2000, Bichon et al. 2009) and is tailor-made to suit the research purpose of this paper. The cantilever beam shown in Figure 4-3 has a deterministic length $L = 100.0 \text{ in}$ and constant cross-sectional area along the beam. It is subjected to uncertain horizontal load $X \text{ (lbf)}$ and vertical load $Y \text{ (lbf)}$. The objective of the design is to minimize the total weight by adjusting the design variables, i.e., the width $b \in [2, 4](\text{in})$ and the height $h \in [2, 4](\text{in})$ of the cross section. Equivalently, the objective can be expressed as the minimization of the cross-sectional area bh .

Figure 4-3: Schematic diagram of a cantilever beam

Three reliability constraints are considered, which are related to maximum normal stress, maximum displacement and natural frequency, respectively. The LSFs and their corresponding reliability constraints are listed as follows.

- (1) The first failure event is defined as yielding at a corner of the fixed end of the beam

where the normal stress is maximum

$$F_s = \left\{ (\boldsymbol{\varphi}, \boldsymbol{\theta}) \left| g_s(\boldsymbol{\varphi}, \boldsymbol{\theta}) = \sigma_y - \frac{600}{bh} \left(\frac{X}{b} + \frac{Y}{h} \right) \leq 0 \right. \right\} \quad (4-13)$$

where σ_y is the yield stress. The allowable failure probability $[P_{F,s}]$ for the yielding failure is 1.358×10^{-3} , which corresponds to the reliability index $\beta = 3.0$. The reliability constraint is given by:

$$P_{F,s}(\boldsymbol{\varphi}) \leq [P_{F,s}] = 1.358 \times 10^{-3} \quad (4-14)$$

- (2) The second failure event is defined as the maximum displacement at the free end of the beam exceeding the allowable tip displacement

$$F_d = \left\{ (\boldsymbol{\varphi}, \boldsymbol{\theta}) \left| g_d(\boldsymbol{\varphi}, \boldsymbol{\theta}) = [d] - \frac{4L^3}{Ebh} \sqrt{\left(\frac{X}{b^2}\right)^2 + \left(\frac{Y}{h^2}\right)^2} \leq 0 \right. \right\} \quad (4-15)$$

where $[d]$ is the allowable tip displacement, which is set to 2.25 in in this example; E is Young's modulus, which is considered as a random variable. The corresponding reliability constraint is given by:

$$P_{F,d}(\boldsymbol{\varphi}) \leq [P_{F,d}] = 1.358 \times 10^{-3} \quad (4-16)$$

(3) The third failure event is defined as the first natural frequency falling into the unfavorable frequency range $53 \leq \omega_1 \leq 67$ (rad/s)

$$F_{\omega_1} = \{(\boldsymbol{\varphi}, \boldsymbol{\theta}) | g_{\omega_1}(\boldsymbol{\varphi}, \boldsymbol{\theta}) = |60 - \omega_1| - 7 \leq 0\} \quad (4-17)$$

where

$$\omega_1 = 1.875^2 \sqrt{\frac{EI}{\rho AL^4}} = 1.015 \sqrt{\frac{Eh^2}{\rho L^4}} \quad (4-18)$$

Again, the density ρ is considered as a random variable. The required reliability level in this constraint is lower than that in the previous two constraints, i.e., the allowable failure probability is 6.210×10^{-3} , which corresponds to the reliability index $\beta = 2.5$. Its reliability constraint is expressed as:

$$P_{F,\omega_1}(\boldsymbol{\varphi}) \leq [P_{F,\omega_1}] = 6.210 \times 10^{-3} \quad (4-19)$$

It is easy to observe that failure events in reliability constraints (1) and (2) highly depend on each other while the failure event in the dynamic reliability constraint (3) has low dependence on the other two static reliability constraints. The statistical properties of uncertain parameters in this model are summarized in Table 4-1.

Table 4-1: Distribution parameters of random variables for the cantilever beam

Variable	Distribution	Mean	CoV
X (lbf)	normal	500	0.20
Y (lbf)	normal	1000	0.10
σ_y (psi)	normal	40000	0.05
E (psi)	normal	29×10^6	0.05
ρ (lb/in ³)	normal	0.29	0.05

Comparative study

Three cases are studied to demonstrate the efficiency of DOPADS_m compared to DOPADS: Case A considers reliability constraints (1) and (2); Case B considers reliability constraints (2) and (3); Case C considers all three reliability constraints. In each case, DOPADS, DOPADS_m and the crude MCS method are used respectively to calculate the optimal design. We set $P^* = 0.1$ and $N_s = 3000$.

Figure 4-4 shows how DOPADS_m propagates failure samples of all three failure events simultaneously from high-density regions to low-density regions in Case C. It takes five iterations for failure samples to cover the feasible regions in reliability constraints (1) and (2), while in reliability constraint (3) it only takes four iterations. The reason is that the allowable failure probability in reliability constraint (3) is larger than the other two constraints. As shown in Figure 4-4, reliability constraints (1) and (2) share a lot of common samples in each iteration due to the dependency of failure events, thus can reduce the required size of failure samples. The contour plots of the final FPFs that correspond to three reliability constraints are illustrated in Figure 4-5.

Figure 4-4: Propagation of failure samples in partitioned design space (Case C). k is the index of iteration and I is the index of reliability constraint. Only 200 out of 3000 samples are plotted in this figure for visualization.

Figure 4-5: Contour plots of the failure probability functions

After approximating FPFs, the optimal solutions can be obtained by the deterministic optimization, the results of which are summarized in Table 4-2. Here the crude MCS is used to validate DOPADS_m and DOPADS. In the crude MCS, failure probabilities are calculated using direct Monte Carlo simulation for 441 gridpoints in design space, which are generated by 21 evenly spaced points in both dimensions. For each gridpoint, a total number of 10^8 simulations are run to calculate the failure probability. It can be observed that the minimum objective values $b \times h$ obtained by DOPADS_m and DOPADS are close to those obtained by the crude MCS, indicating that the proposed method has a good accuracy in this example.

Table 4-2: Optimal solutions and their objective values for three different cases

Case	DOPADS_m			DOPADS			Crude MCS		
	b	h	$b \times h$	b	h	$b \times h$	b	h	$b \times h$
Case A	2.82	3.45	9.71	2.79	3.47	9.69	2.6	3.7	9.62
Case B	2.42	3.98	9.64	4.00	2.40	9.60	2.4	4.0	9.60
Case C	2.66	3.65	9.72	2.65	3.73	9.88	2.6	3.7	9.62

We measure the efficiency in terms of the total number of model evaluations. Table 4-3 to Table 4-5 compare the number of model evaluations required in each stage of DOPADS_m and DOPADS for Case A, Case B and Case C, respectively. In Case A, the total number is $N_{eval} = 24745$ for DOPADS_m and $N_{eval} = 30012$ for DOPADS. DOPADS_m improves the efficiency by reducing 17.6% of the total number. In Case B, the

total number is 24547 for DOPADS_m and 27173 for DOPADS. Compared to Case A, fewer evaluations are needed in Case B since the allowable failure probability in reliability constraint (2) is relatively large. However, in Case B, by using DOPADS_m, the number of model evaluations is only reduced by 9.7%. It is predictable considering the low dependence between the failure events involved. As listed in Table 4-5, in Case C, the total number of model evaluations is $N_{eval} = 28175$ for DOPADS_m and $N_{eval} = 42156$ for DOPADS. Because of adding one more reliability constraint in Case C, the total number is also increased compared to the other two cases. However, in this case DOPADS_m outperforms DOPADS greatly by reducing 33.2% of the total number.

In summary, DOPADS_m enhances the efficiency of DOPADS by propagating samples simultaneously. The more dependence different failure events have, the more efficiency we can gain. In addition, the efficiency enhancement is more remarkable when more reliability constraints are considered.

Table 4-3: Number of model evaluations in DOPADS_m and DOPADS (Case A)

Framework	Reliability constraint	Stage						
		Initial reliability analysis	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
DOPADS_m	--	1000	3078	3858	4560	4930	4431	2888
DOPADS	(1)	1000	2588	2819	2869	2871	2588	--
	(2)	1000	2535	2825	2835	2893	2535	--

Table 4-4: Number of model evaluations in DOPADS_m and DOPADS (Case B)

Framework	Reliability constraint	Stage						
		Initial reliability analysis	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
DOPADS_m	--	1000	4792	4205	4334	4388	2886	2942
DOPADS	(1)	1000	2588	2819	2869	2871	2883	--
	(3)	1000	2624	2807	2878	2834	--	--

Table 4-5: Number of model evaluations in DOPADS_m and DOPADS (Case C)

Framework	Reliability constraint	Stage						
		Initial reliability analysis	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
DOPADS_m	--	1000	5177	4424	4860	5202	4598	2914
	(1)	1000	2588	2819	2869	2871	2883	--
DOPADS	(2)	1000	2535	2825	2835	2893	2894	--
	(3)	1000	2669	2715	2865	2895	--	--

Effects of the sample size on the optimal solutions

In the previous comparative study, the hyperparameter is fixed to $N_s = 3000$ which ensures that in each iteration there are at least 3000 failure samples for each failure event F_l ($l = 1, 2, 3$) to perform the density estimation. On one hand, the increase of N_s can improve the accuracy of density estimation; on the other hand, since N_s determines the total number of failure samples N_k in each iteration ($k = 0, 1, 2, \dots$), a larger N_s will lead to more computational burden. In other words, the hyperparameter N_s is critical to balance the accuracy and efficiency in DOPADS_m.

Figure 4-6 shows how the means of minimal objective values change with the increase of N_s in all three cases. To obtain the means, both DOPADS_m and DOPADS are repeated independently for 30 times with different random seeds. In Figure 4-6, the optimization results of DOPADS_m and DOPADS are very close to each other when $N_s = 4000$. Furthermore, the means of optimal objective values of DOPADS_m appear to converge earlier than those of DOPADS: the fluctuation observed in DOPADS_m becomes small when $N_s \geq 2500$, but the results from DOPADS have not stabilized for Case A and Case B.

Figure 4-6: Means of the minimal objective values for different N_s

4.4.2 Example 2: a ten-story shear building subjected to ten reliability constraints

This example is adapted from Example 3 in Section 3.5.3 for the illustration of DOPADS_m. As illustrated in Figure 4-7, the ten-story planar shear building with $3m$ vertical distance between stories is subjected to earthquake excitation. The design variables are denoted as $\boldsymbol{\varphi} = [a_{1\sim 2}, a_{3\sim 4}, a_{5\sim 6}, a_{7\sim 8}, a_{9\sim 10}]^T$ where $a_{j\sim(j+1)} \in [250, 600]$ (mm) is the width or height of the columns in Story j and Story $(j+1)$ by assuming the two neighboring stories have square columns with the same dimensions. The lumped mass of each story m_j ($j = 1, \dots, 10$) follows the normal distribution whose CoV is 0.10 and mean values are 3×10^6 kg for the Story 1-4, 2.5×10^6 kg for Story 5-8 and 2.0×10^6 kg for Story 9-10. The Young's Modulus E_j ($j = 1, \dots, 10$) are independent and identically distributed normal random variables with mean value 30000 MPa and CoV 0.10. Classical modal damping of each story ζ_j ($j = 1, \dots, 10$) is also independent and identically distributed, which is assumed to follow the normal distribution with mean value 0.04 and CoV 0.20. The objective of this example is to minimize the total cross-sectional areas of all columns, i.e., $\sum_{j=1}^5 a_{(2j-1)\sim(2j)}^2$.

Figure 4-7: Schematic diagram of the ten-story shear building subjected to multiple reliability constraints

The stochastic ground motion is generated by the modulated filtered white noise as introduced in Section 3.5.3. The ground acceleration at any time t can be simulated in this way: first draw a white noise sample for $w(t)$, then calculate $[v_1, \dot{v}_1, v_2, \dot{v}_2]^T$ by solving the differential equation in Equation (3-17), and finally a sample of ground acceleration at time t can be obtained by using Equation (3-16). In this example, the duration of earthquake ground motion is 20s with the sampling interval 0.01s. As a result, we need 2001 standard normally distributed random variables to generate a sample time-history of ground acceleration as shown in Figure 4-7. The total number of random variables considered in this example is $2001 + 5 + 10 + 10 + 10 = 2036$.

The failure events corresponding to the reliability constraints are defined as the maximum inter-story drift during the earthquake period exceeding 60mm. For $j = 1, \dots, 10$,

$$F_j = \left\{ (\boldsymbol{\varphi}, \boldsymbol{\theta}) \mid g_j(\boldsymbol{\varphi}, \boldsymbol{\theta}) = 0.06 - \max_{0 \leq t \leq 20s} |\Delta d_{j,t}| \leq 0 \right\} \quad (4-20)$$

where $\Delta d_{j,t}$ is the inter-story drift between Story $(j - 1)$ and Story j . In this example, the allowable failure probability for all failure events is 0.01 and the reliability constraints can be expressed as:

$$P_{F,j}(\boldsymbol{\varphi}) \leq [P_{F,j}] = 0.01 \quad (4-21)$$

Unlike the previous example in which the LSFs are given in their closed forms, this dynamic example requires to solve the differential equation and thus can be computationally prohibitive when many evaluations are needed. Since $n_{gp} \times 10^8$ simulations where n_{gp} is the number of gridpoints should be run in the crude MCS, it is infeasible to use the crude MCS to validate DOPADS_m in this example. With a fixed N_s , one can always approximate the FPFs using DOPADS_m to obtain the optimal solutions. However, it may be doubtful whether the sample size N_s is sufficiently large to ensure an accurate result. To resolve this issue, in this example we keep increasing N_s by ΔN_s in each region of the initial partitioned design space until the early stopping criterion is met: the increase of N_s stops in round t ($t = 1, 2, 3, \dots$) if the condition $t - t^* \geq r_e$ is satisfied where t^* is the best round in which the optimal objective value is smallest and r_e is the number of the early stopping rounds. This strategy is suggested to be adopted when the models are computationally expensive to evaluate.

The change of the optimal objective values is shown in Figure 4-8. We set $r_e = 2$ so that the increase of N_s will stop if the minimal optimal objective value has been found in the last two rounds. Both DOPADS_m and DOPADS start with $N_s = 3000$ in the first round. Then approximation of FPFs and deterministic optimization are performed to obtain the optimal solution. In the following rounds, additional $\Delta N_s = 1000$ failure samples are generated in the design space which has been partitioned in the first round. For DOPADS_m, the minimal objective value is $9.86 \times 10^5 \text{ mm}^2$ when $t^* = 2$; for DOPADS, the minimal objective is $9.98 \times 10^5 \text{ mm}^2$ when $t^* = 4$.

Figure 4-8: Change of the minimal objective values with the increase of N_s

As the sample simulation in the follow-up rounds is based on the partition from the first round, the computation cost in each round remains to be $\Delta N_s \times (n_{it} + 1)$ where n_{it} is the index of the final iteration. Table 4-3 summarizes the number of model evaluations for DOPADS_m. The total number of model evaluations in DOPADS_m is $N_s = 44177$, which takes about eight hours to run in a desktop with Intel Core i5-4570 CPU @ 3.2GHz. For DOPADS, the required total number of model evaluations is $N_s = 260975$. By taking advantage of the dependency of the ten failure events, DOPADS_m can reduce the computation time by 83.1% compared to DOPADS.

Table 4-6: Number of model evaluations in DOPADS_m using early stopping strategy to determine the sample size

Initial reliability analysis	Round 1				Round 2	Round 3	Round 4
	$k = 0$	$k = 1$	$k = 2$	$k = 3$			
4600	6905	6127	7913	6632	4000	4000	4000

4.5 Conclusions

It has been long recognized that decoupling the reliability analysis from the optimization loop can facilitate the procedure of RBDO. Though DOPADS is robust to the complexity of models and can handle a large number of random variables, its computation cost increases linearly with the number of reliability constraints involved. It will lead to prohibitive computational burden when many reliability constraints are taken into consideration.

In this section, a new framework DOPADS_m that enhances the efficiency of DOPADS by making use of the dependency between different failure events has been presented. The samples generated from different failure events can be propagated simultaneously from the high-density regions to low-density regions. It has been proven that the total number of sample simulations in DOPADS_m is always less than that required in DOPADS to achieve the same level of accuracy. Two illustrative examples with different levels of the model complexity have been investigated to show the effectiveness and efficiency of the proposed framework. In the first example, by studying the three cases (two reliability constraints involving highly dependent failure events, two reliability constraints involving lowly

dependent failure events, three reliability constraints), it has been shown that the efficiency enhancement becomes more notable in the case of higher dependence of failure events and more reliability constraints. In the second example that has more than two thousand random variables and ten reliability constraints, DOPADS_m can reduce the total required number of sample simulations by 83.1% compared to DOPADS.

Applications of RBDO to problems with multiple failure events are limited by the high computational burden posed by sophisticated numerical models. It has been shown in this chapter that the proposed DOPADS_m requires a reasonable amount of time and can alleviate this problem.

5 Instant reliability updating by constructing the updated failure probability function

Summary

Predictions by a design model often do not match well with the measured data from the experiments on the real system. It is crucial to update the design model to check whether the system sustains the prescribed reliability level in the construction phase; if not, immediate actions (e.g., design change) must be taken. The evaluation based on the updated model must be performed quickly to prevent potential catastrophic consequences or cost/schedule overrun. In this paper, we propose a framework for *instant reliability updating* (INREU) in which the *updated failure probability function* (UFPF) can be constructed before acquiring the measurements. With the help of UFPF, the process of reliability updating can be accomplished instantly by plugging measurements into the UFPF.

5.1 Introduction

In practice, mathematical models are used to mimic the behavior of a real-world system. While the level of complexity in models has continued to increase in order to predict the system behavior more precisely, uncertainties associated with these models must be addressed. Fortunately, thanks to the advances in sensing technologies, measured data from monitoring and/or inspection have become more widely available, which contain valuable information about the actual behavior of systems that can be used to reduce model uncertainties. This has led to the development of system identification and model updating

methods, aiming to improve models based on measurements so that one can subsequently update the performance assessment of a system (e.g., reliability in this paper).

Most often, reliability updating consists of two steps. First, one uses measured data to update uncertain model parameters. This is followed by the calculation of the probability of failure based on updated model parameters. The first step, often referred to as model updating, can be done in several approaches such as residual minimization (Mottershead et al. 2011), Bayesian inference (Katafygiotis and Beck 1998, Bernardo and Smith 2008, Beck 2010) and error-domain model falsification (Goulet and Smith 2013) (one can refer to (Simoen et al. 2015) for detailed comparison of different approaches). Among them, Bayesian inference is arguably one of the best tools for model updating since it can include all potentially usable information (i.e., measurements and prior knowledge) within a rigorous probabilistic framework. Alternatively, a new framework for Bayesian reliability updating has been proposed by Straub et al. (2016), which takes advantage of existing reliability methods to directly calculate the updated probability of failure without performing model updating.

Despite extensive theoretical developments, the application of reliability updating to real-world projects is restricted. In a responsive world, it is crucial that a massive amount of data from ubiquitous sensors can be processed timely with the corresponding assessments and decisions made quickly. However, neither aforementioned two-step nor one-step way can update the reliability immediately after the collection of measurement data when complex systems with models that are computationally expensive to evaluate are involved. For instance, in geotechnical engineering, it may take a few weeks to compute the updated probability of failure using regular personal computers if sophisticated soil constitutive relation and soil-structure interaction are considered.

In this section, we present a stochastic-sampling-based framework for instant reliability updating (INREU) by constructing the updated failure probability function (UFPF). The formulation of UFPF was first proposed by Cheung and Beck (2007), and later was adopted by Ching and Hsieh (2009) to update future reliability of systems with low dimensional measurements. Based on the formulation of UFPF, the accuracy of UFPF's mainly depends on the accuracies of density estimation of conditional samples. But neither of the papers

addressed this issue. In this paper, we present an iterative density estimation scheme in which conditional samples are propagated from the high-density regions to low-density ones, thus providing sufficient distribution information in low-density regions. Moreover, the accuracy of UFPF with respect to the sample size has also been investigated in illustrative examples. The UFPF allows one to enquire the updated probability of failure by plugging in measurements right after measured data are available. The proposed framework is robust to the number of random variables, system complexity and the degree of nonlinearity.

The structure of this chapter is as follows. In Section 5.2, we introduce two conventional Bayesian reliability updating formulations. Section 5.3 presents the proposed framework in detail. Section 5.4 demonstrates the usefulness of the proposed framework through two illustrative examples.

5.2 Common Bayesian reliability updating formulations

Bayesian reliability updating is a task of evaluating the probability of failure conditional on measurements. This part presents the conventional formulation for Bayesian reliability updating. The difficulties in the reliability updating as well as existing methods will be discussed.

Let $\mathbf{d}^* = [d_1^*, \dots, d_{n_d}^*]^T$ be a vector of n_d measurements and $\boldsymbol{\theta} = [\theta_1, \dots, \theta_{n_\theta}]^T$ be a vector of n_θ uncertain model parameters with the prior joint PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is often determined based on engineering experience and judgement or past investigation. By Bayes' theorem, the updated or posterior PDF $p_{\mathbf{d}^*}(\boldsymbol{\theta})$ can be expressed as:

$$p_{\mathbf{d}^*}(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|\mathbf{d}^*) = \frac{p(\mathbf{d}^*|\boldsymbol{\theta})p_{\boldsymbol{\theta}}(\boldsymbol{\theta})}{p(\mathbf{d}^*)} \propto p(\mathbf{d}^*|\boldsymbol{\theta})p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \quad (5-1)$$

where $p(\mathbf{d}^*|\boldsymbol{\theta})$ is the *likelihood function* which can be formulated by the probability distribution of the prediction errors describing the discrepancy between the predicted outputs of a model and the measured outputs. It should be noted that the prediction errors consist of both measurement errors and modeling errors (Beck and Katafygiotis 1998, Cheung et al. 2011). The denominator $p(\mathbf{d}^*) = \int p(\mathbf{d}^*|\boldsymbol{\theta})p_{\boldsymbol{\theta}}(\boldsymbol{\theta})d\boldsymbol{\theta}$ is the normalizing constant.

Given the updated PDF $p_{\mathbf{d}^*}(\boldsymbol{\theta})$, based on the law of total probability, the updated probability of failure $P_F(\mathbf{d}^*)$ is:

$$P_F(\mathbf{d}^*) = \int_{\boldsymbol{\theta} \in \Omega_{\theta} \subseteq \mathbb{R}^{n_{\theta}}} \mathbb{I}_F(\boldsymbol{\theta}) p_{\mathbf{d}^*}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbb{E}[\mathbb{I}_F(\boldsymbol{\theta})] \quad (5-2)$$

where $\mathbb{I}_F(\cdot)$ is the binary indicator function of the failure event, which equals to one if $\boldsymbol{\theta}$ belongs to the failure region Ω_F and zero otherwise. This two-step approach to reliability updating involves (i) the update of the probability distribution of model parameters in Equation (5-1) and (ii) the update of the failure probability in Equation (5-2).

Alternatively, substituting (5-1) into (5-2) gives rise to another expression of the updated failure probability $P_F(\mathbf{d}^*)$:

$$P_F(\mathbf{d}^*) = \frac{\int_{\boldsymbol{\theta} \in \Omega_{\theta} \subseteq \mathbb{R}^{n_{\theta}}} \mathbb{I}_F(\boldsymbol{\theta}) p(\mathbf{d}^* | \boldsymbol{\theta}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}}{p(\mathbf{d}^*)} \quad (5-3)$$

After choosing the likelihood function and the prior PDF, both Equations (5-2) and (5-3) can be used to perform the reliability updating. The multi-dimensional integrals in Equations (5-1) to (5-3) often cannot be evaluated analytically when the number of model parameters is not very small.

Two popular categories of methods for approximating the integrals are stochastic-sampling-based methods, especially MCMC-based ones (Ching and Chen 2007, Cheung and Beck 2009, Cheung and Bansal 2017), and Laplace's method and its generalizations (Beck and Katafygiotis 1998, Katafygiotis and Beck 1998, Beck 2010). In MCMC methods, first we generate samples from the posterior PDF $p_{\mathbf{d}^*}(\boldsymbol{\theta})$ which is known to up to a normalizing constant, then estimate the $P_F(\mathbf{d}^*)$ by averaging the values of the indicator function for the generated samples. Laplace's method needs to search for optimal points in the parameter space, which can be very challenging in high-dimensional cases. It is suggested by Beck (2010) that MCMC-based methods are preferred unless the parameters are globally identifiable and the amount of data is not too small.

5.3 A framework for instant reliability updating

In the conventional formulations, reliability updating is performed after measurements are acquired. As the implementation of model updating can be computationally demanding even

using MCMC when complex models are involved, these formulations are only feasible in cases where immediate responses to measured data are not required. However, in most real-world scenarios, the lag between data acquisition and model updating should be minimized to avoid catastrophic consequences.

This section presents a framework for instant reliability updating (INREU) by constructing the updated failure probability functions (UFPF). The term ‘instant’ is used here to distinguish updating process in the proposed framework from the conventional ones. That is, unlike the conventional updating process where one needs to acquire measurements first and then calculate the updated reliability, the proposed framework considers a measurement as a query point so that updating can be achieved instantly by plugging the query point into the UFPF. The comparison of the updating processes is illustrated in Figure 5-1.

Figure 5-1: Comparison of the updating processes. Panel (a): Conventional reliability updating has a large time lag between the measurement acquisition and reliability updating; and Panel (b): The proposed framework has a negligible time lag. The length of arrows represents the length of time lags but not in scale and the shaded areas represent the most time-consuming part.

5.3.1 Updated failure probability function

The core part of the proposed framework is the construction of the UFPF in terms of the measurements. Now let variables $\mathbf{d} = [d_1, \dots, d_{n_d}]^T$ represent a vector of measurements in

the UFPF. Then we denote the UFPF with domain Ω_d as $P_F(\mathbf{d}): \Omega_d \rightarrow [0, 1]$. Again, based on Bayes' theorem, the UFPF is given by:

$$P_F(\mathbf{d}) = P(F|\mathbf{d}) = \frac{p(\mathbf{d}|F)P(F)}{p(\mathbf{d})} \quad (5-4)$$

Based on the law of total probability, Equation (5-4) can be rewritten as:

$$P_F(\mathbf{d}) = \frac{p(\mathbf{d}|F)P(F)}{p(\mathbf{d}|F)P(F) + p(\mathbf{d}|\bar{F})P(\bar{F})} = \frac{1}{1 + \frac{p(\mathbf{d}|\bar{F})P(\bar{F})}{p(\mathbf{d}|F)P(F)}} \quad (5-5)$$

where $P(F)$ and $P(\bar{F})$ are the original probability of the failure and non-failure events, respectively, before updating; and $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$ are the PDFs of $\mathbf{d}$ conditional on F and $\bar{F}$, respectively, also referred to as the density functions of failure and non-failure samples. Like in DOPADS, the measurement variables are artificially treated as random variables following the prescribed PDF $p(\mathbf{d})$.

Regarding $P(F)$ and $P(\bar{F})$, the identity $P(\bar{F}) = 1 - P(F)$ holds as F and $\bar{F}$ are mutually exclusive and collectively exhaustive events. In the INREU framework, we only consider stochastic-sampling-based methods to conduct the reliability analysis for the original model. While the crude MCS can be straightforwardly implemented to estimate $P(F)$, it requires a prohibitively large amount of model evaluations when the failure probability is small. Advanced methods (e.g., SS (Au and Beck 2001b)) have proved to be far more efficient.

5.3.2 Measurement samples conditional on the failure or non-failure event

Let $\{\boldsymbol{\theta}\}_F = \{\boldsymbol{\theta} | \mathbb{I}_F(\boldsymbol{\theta}) = 1, \boldsymbol{\theta} \sim p_{\boldsymbol{\theta}}(\boldsymbol{\theta})\}$ and $\{\boldsymbol{\theta}\}_{\bar{F}} = \{\boldsymbol{\theta} | \mathbb{I}_F(\boldsymbol{\theta}) = 0, \boldsymbol{\theta} \sim p_{\boldsymbol{\theta}}(\boldsymbol{\theta})\}$ denote the sets of failure and non-failure parameter samples, respectively. As a byproduct of the preceding stochastic sampling for estimating $P(F)$, both $\{\boldsymbol{\theta}\}_F$ and $\{\boldsymbol{\theta}\}_{\bar{F}}$ distributed as $p(\boldsymbol{\theta}|F)$ and $p(\boldsymbol{\theta}|\bar{F})$ are readily available. Also, let $\mathbf{h}(\boldsymbol{\theta})$ denote the *predictive functions* for measurements. Without loss of generality, additive errors are considered here such that $\mathbf{d} = \mathbf{h}(\boldsymbol{\theta}) + \boldsymbol{\epsilon}_p$ where $\boldsymbol{\epsilon}_p \sim p(\boldsymbol{\epsilon}_p | \boldsymbol{\theta})$ is a vector of *prediction errors* which often depend on part of $\boldsymbol{\theta}$ (e.g., variances of Gaussian errors can be included in $\boldsymbol{\theta}$). Note that both the predictive functions $\mathbf{h}(\boldsymbol{\theta})$ and the performance functions $\mathbf{g}(\boldsymbol{\theta})$ are models of the studied system but have different aims: $\mathbf{h}(\boldsymbol{\theta})$ predicts the system responses related to the

measurement quantities whereas $\mathbf{g}(\boldsymbol{\theta})$ focuses on the responses related to the failure events.

For each failure parameter sample $\boldsymbol{\theta}_F^{(i)}$ in $\{\boldsymbol{\theta}\}_F$, if we draw an independent error sample $\boldsymbol{\epsilon}_p^{(i)}$ from $p(\boldsymbol{\epsilon}_p|\boldsymbol{\theta})$, the measurement sample $\mathbf{d}_F^{(i)} = \mathbf{h}(\boldsymbol{\theta}_F^{(i)}) + \boldsymbol{\epsilon}_p^{(i)}$ follows the distribution $p(\mathbf{d}|F)$. Similarly, a set of measurement sample $\mathbf{d}_{\bar{F}}^{(i)}$ for non-failure events that follows $p(\mathbf{d}|\bar{F})$ can be obtained. Then these samples are used to estimate the underlying density functions in Equation (5-5).

To summarize, the UFPF $P_F(\mathbf{d})$ can be constructed in a Bayesian approach as shown in Equation (5-5). Stochastic-sampling-based reliability methods can be employed to estimate the failure and non-failure probabilities, as well as to generate samples of measurements distributed as $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$, respectively (this step is referred to as the pilot stochastic simulation hereafter). With these samples at disposal, one can perform density estimation to estimate $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$.

5.3.3 Iterative density estimation of measurement samples

Given measurement samples $\{\mathbf{d}\}_F$ and $\{\mathbf{d}\}_{\bar{F}}$ for failure and non-failure events, respectively, we intend to estimate the density functions $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$ over the measurement space Ω_d . The iterative density estimation scheme presented in Section 3.3 is used. For brevity in this subsection, we do not distinguish between F and $\bar{F}$, and denote samples as $\{\mathbf{d}\}_A$ and its density function as $p_A(\mathbf{d})$ where A represents F or $\bar{F}$.

In essence, samples are propagated from the high-density regions to the low-density ones iteratively as shown in Figure 5-2. In each iteration, a piecewise-constant density estimator over the previous low-density region is constructed based on the increased samples. Readers may refer to Section 3.3 for details about the iterative scheme.

Figure 5-2: Sample generation and propagation in the k -th iteration

Note that the generation of additional samples in INREU is different from that described in Section 3.3. In INREU, we intend to perform the density estimation for measurement samples (output samples) instead of model parameter samples (input samples). Assume we have N_k parameter and error samples $\{(\boldsymbol{\theta}^{(i)}, \boldsymbol{\epsilon}_p^{(i)}), i: 1 \rightarrow N_s\}_{A, \mathcal{L}_{k-1}} \subset \{(\boldsymbol{\theta}, \boldsymbol{\epsilon}_p) | g(\boldsymbol{\theta}) \leq 0, (\mathbf{h}(\boldsymbol{\theta}) + \boldsymbol{\epsilon}_p) \in \mathcal{L}_{k-1}\}$ at the beginning of the k -th iteration whose corresponding measure samples $\{\mathbf{d}\}_{A, \mathcal{L}_{k-1}}$ follow $p_k(\mathbf{d}|A)$. To increase the sample size of $\{\mathbf{d}\}_{A, \mathcal{L}_{k-1}}$ from N_k to N_s , $(N_s - N_k)$ additional parameter and error samples are simulated using modified Metropolis-Hastings algorithm, thus the corresponding measurement samples $\{\mathbf{d}\}_{A, \mathcal{L}_{k-1}}$ follows $p_k(\mathbf{d}|A)$. The ‘accept-reject’ step of the sample level checks whether both A and $\mathcal{L}_{k-1}$ happen at the same time.

5.3.4 Summary of implementation

A flowchart that summarizes the steps of instant reliability updating (INREU) is shown in Figure 5-3. We start INREU by defining the inputs which include model parameters, prediction errors, the failure event and predictive functions for measurements. The next step

is to construct the UFPF which is expressed in terms of $P(F)$, $P(\bar{F})$, $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$ as shown in Equation (5-5). Among them, $P(F)$ and $P(\bar{F})$ are constants and can be estimated by stochastic-sampling-based methods such as MCS and SS; $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$ are density functions of $\mathbf{d}$. For estimating the density distribution of samples $\{\mathbf{d}\}_F$ and $\{\mathbf{d}\}_{\bar{F}}$, an iterative density estimation scheme ensures the estimation accuracy by propagating samples from high-density regions to low-density regions. Finally, actual measurements $\mathbf{d}=\mathbf{d}^*$ can be considered as query points to the UFPF and the probability of failure can be obtained instantly through plugging in.

Figure 5-3: Flowchart of instant reliability updating

5.4 Illustrative examples

5.4.1 Example 1: Linear model - An analytical example

First consider a standard linear model with Gaussian white noise:

$$h(\boldsymbol{\theta}) = \mathbf{x}^T \boldsymbol{\theta}, \quad y_m = h(\boldsymbol{\theta}) + \epsilon \quad (5-6)$$

Chapter 5. Instant reliability updating by constructing the updated failure probability function

where $\boldsymbol{\theta}$ is a vector of parameter of the linear model, $h(\boldsymbol{\theta})$ is the predictive function or its value, $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is the additive prediction error modeled as a Gaussian random variable with a mean value 0 and variance σ^2 , and y_m is the measured output.

For the prior PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \sim \mathcal{N}(\bar{\boldsymbol{\theta}}, \Sigma)$, the posterior PDF can be given analytically based on Bayes' theorem:

$$\begin{aligned} p_{\boldsymbol{\theta}}(\boldsymbol{\theta}|y_m) &\propto \exp\left(-\frac{(y_m - \mathbf{x}^T \boldsymbol{\theta})^2}{2\sigma^2}\right) \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^T \Sigma^{-1}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})\right) \\ &\propto \exp(\boldsymbol{\zeta}_1 + \boldsymbol{\eta}_1^T \boldsymbol{\theta} + \boldsymbol{\theta}^T \boldsymbol{\Lambda}_1 \boldsymbol{\theta}) \exp(\boldsymbol{\zeta}_2 + \boldsymbol{\eta}_2^T \boldsymbol{\theta} + \boldsymbol{\theta}^T \boldsymbol{\Lambda}_2 \boldsymbol{\theta}) \end{aligned} \quad (5-7)$$

where $\boldsymbol{\zeta}_1 = -\frac{y_m^2}{2\sigma^2}$, $\boldsymbol{\eta}_1 = \frac{y_m \mathbf{x}}{\sigma^2}$, $\boldsymbol{\Lambda}_1 = -\frac{\mathbf{x}\mathbf{x}^T}{2\sigma^2}$, $\boldsymbol{\zeta}_2 = -\frac{\bar{\boldsymbol{\theta}}^T \Sigma^{-1} \bar{\boldsymbol{\theta}}}{2}$, $\boldsymbol{\eta}_2 = \Sigma^{-1} \bar{\boldsymbol{\theta}}$, $\boldsymbol{\Lambda}_2 = -\frac{\Sigma^{-1}}{2}$.

Thus,

$$p_{\boldsymbol{\theta}}(\boldsymbol{\theta}|y_m) \propto \exp(\boldsymbol{\zeta}_3 + \boldsymbol{\eta}_3^T \boldsymbol{\theta} + \boldsymbol{\theta}^T \boldsymbol{\Lambda}_3 \boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}_p)^T \mathbf{A}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}_p)\right) \quad (5-8)$$

where $\mathbf{A} = \Sigma^{-1} + \frac{\mathbf{x}\mathbf{x}^T}{\sigma^2}$. Thus, the posterior PDF is $p_{\boldsymbol{\theta}}(\boldsymbol{\theta}|y_m) \sim \mathcal{N}(\bar{\boldsymbol{\theta}}_p, \mathbf{A}^{-1})$ where the mean $\bar{\boldsymbol{\theta}}_p = \mathbf{A}^{-1} \left(\frac{y_m \mathbf{x}}{\sigma^2} + \Sigma^{-1} \bar{\boldsymbol{\theta}} \right)$ and covariance matrix $\mathbf{A}^{-1}$.

Let the failure event be defined as:

$$F = \{\boldsymbol{\theta} | g(\boldsymbol{\theta}) = \theta_1 + \theta_2 - 1 \leq 0\} \quad (5-9)$$

As the LSF is linear, the probability of failure can be exactly determined using FORM:

$$P_F = \Phi(-\beta) \quad (5-10)$$

where β is the reliability index which equals to the minimum distance from the origin to the failure boundary in the standard Gaussian space and $\Phi(\cdot)$ is the CDF of the standard normal distribution.

Assume the prior PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \mathcal{N}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0.5^2 & 0 \\ 0 & 0.7^2 \end{bmatrix}\right)$, $\mathbf{x}^T = [1 \quad 1]$ and $\epsilon \sim \mathcal{N}(0, 0.8^2)$. For any possible measurement y_m in the domain of interest $[1.0, 4.5]$, the posterior PDF can be computed analytically according to Equation (5-8), and then the exact updated probability of failure can be obtained according to Equation (5-10). To validate INREU, the updated probabilities of failure are computed exhaustively for 36 equally spaced y_m 's over the range $[1.0, 4.5]$ as shown in Figure 5-4, which are regarded here as the ground truth.

Comparison of the updated failure probability function from INREU with the ground truth

To construct the UFPF, $p(y_m|F)$ and $p(y_m|\bar{F})$ are estimated separately by using the iterative scheme presented in Section 5.3.3. In each iteration, the sample size is $N_s = 2000$. From the pilot stochastic simulation, the original failure probability of the system is 0.1225 before updating. The iterative process stops when the union of the high-density regions identified in each iteration covers the domain of interest, which in this example is $y_m \in [1.0, 4.5]$. As shown in Figure 5-4, six iterations and two iterations are involved in the estimation of $p(y_m|F)$ and $p(y_m|\bar{F})$, respectively. The reason why fewer iterations are required for the estimation $p(y_m|\bar{F})$ is that non-failure samples are more likely to distribute over the domain of interest as illustrated in Figure 5-4(b). Support points from all iterations are used to smooth the estimators based on Gaussian processes (GP) (Rasmussen 2006).

Figure 5-5 shows the support points and the mean smoothed functions. Again, it illustrates the different distributions of failure and non-failure samples, i.e., the value of $p(y_m|F)$ decreases with the increase of the measurement y_m whereas $p(y_m|\bar{F})$ has a mode within the range $[1.0, 4.5]$.

Figure 5-4: Iterative density estimation of failure and non-failure samples for the linear model

Figure 5-5: Support points and mean smoothed density functions based on GP (C.I. is short for Credible Interval)

Figure 5-6 shows the UFPF based on the estimated density functions. It can be seen that the updated probability of failure decreases about four orders of magnitude with the increase of measurement. As shown in Figure 5-6, the UFPF has a good accuracy in comparison with the ground truth, and the accuracy in low-density regions can be improved during the iterative density estimation.

Figure 5-6: Comparison of the UFPF from INREU with the ground truth

Effect of the sample size on the UFPF

The sample size N_s per iteration is a crucial hyperparameter of INREU as it balances the efficiency and accuracy. To evaluate the effect of sample size on the accuracy of the estimated UFPF, we define two evaluation metrics:

$$MISE = \mathbb{E} \left\{ \int [P_F(\mathbf{d}) - \tilde{P}_F(\mathbf{d})]^2 d\mathbf{d} \right\} \quad (5-11a)$$

$$MSE(\mathbf{d}) = \mathbb{E} \{ [P_F(\mathbf{d}) - \tilde{P}_F(\mathbf{d})]^2 \} \quad (5-11b)$$

where $P_F(\cdot)$ and $\tilde{P}_F(\cdot)$ are UFPFs of the proposed INREU and the ground truth, respectively. $MISE$ is short for Mean Integrated Squared Error and can be treated as a global evaluation metric for density function estimator over the entire domain. MSE represents the Mean Squared Error, which is a pointwise evaluation metric.

30 independent replicates are run per sample size to compute the mean value of integrated squared errors and pointwise squared errors. Figure 5-7 depicts the convergence of $MISE$ with the increase of sample size from 200 to 6000. $MISE$ decreases quickly and converges when the sample size is larger than 2000. Similarly, as can be seen from Figure 5-8, all MSE s for three different measurements ($y_m = 1.5, 2.5, 3.5$) converge.

Figure 5-7: Change of the global evaluation metric $MISE$ with the increase of sample size

Figure 5-8: Change of the pointwise evaluation metric MSE with the increase of sample size

5.4.2 Example 2: Ten-bar truss with two measurements of strain

In this example, a ten-bar truss as shown in Figure 5-9 is studied. To update its reliability, we measure the strain of Bar 3 and Bar 7, i.e., $\mathbf{d} = [d_1, d_2]^T$ where d_1 and d_2 are strain measurements of Bar 3 and Bar 7, respectively. The failure event is defined as the vertical displacement of Node 4 Δd_{n4} exceeding the allowable displacement $[\Delta d_t] = 80mm$. Unlike the first illustrative example where measurement predictive function is directly related to the LSF in the definition of the failure event, this example measures the strain of bars while the failure event is concerned about the node displacement.

Statistical characteristics of uncertain parameters and values of deterministic parameters are given in Figure 5-9. Young's modulus of the ten bars and two loadings are considered as uncertain parameters, i.e., $\boldsymbol{\theta} = [E_1, \dots, E_{10}, \dots, P_1, P_2]^T$; the prediction errors of measurements $\epsilon_{p,1}$ and $\epsilon_{p,2}$ are modeled as independent Gaussian random variables. Before updating, the original failure probability P_F of the event $F = \{\boldsymbol{\theta} | [\Delta d_t] - \Delta d_{n4}(\boldsymbol{\theta}) \leq 0\}$ is 0.0113. P_F is calculated in the pilot stochastic simulation in INREU using SS.

Figure 5-9: Schematic diagram of the ten-bar truss and parameters of the model (S.D. is short for Standard Deviation)

Comparison of the updated failure probability function from INREU with the ground truth

The procedure of iterative estimation of $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$ is illustrated in Figure 5-10 with the sample size $N_s = 4000$ per iteration. It is after two iterations and four iterations respectively when the union of high-density regions covers the rectangular domain of interest $d_1 \in [-2.5 \times 10^{-3}, -1.5 \times 10^{-3}]$ (compressive strain) and $d_2 \in [0.8 \times 10^{-3}, 1.8 \times 10^{-3}]$ (tensile strain). After smoothing the piecewise-constant estimators of $p(\mathbf{d}|F)$ and $p(\mathbf{d}|\bar{F})$ using GP, the UFPF of INREU is calculated according to Equation (5-5).

In contrast to the exact ground truth in the first illustrative example, so called ground truth in this example is numerically given by calculating the probabilities of failure for 121 grid points based on Equation (5-3). As shown in Figure 5-11, the grid points are generated based on 11 equally spaced measurements for both d_1 and d_2 . For each grid point, we

Chapter 5. Instant reliability updating by constructing the updated failure probability function

calculate its updated failure probability based on Equation (5-3) using 10^6 samples. Thus, a total number of 2.42×10^8 model function evaluations (including both the predictive function and the limit-state function evaluations) have been performed to obtain the ground truth. Figure 5-11 also shows that the obtained UFPF surface of INREU is close to the ground truth, indicating a good accuracy of the proposed framework.

Figure 5-10: Iterative density estimation of failure and non-failure samples for the ten-bar truss

Figure 5-11: Comparison of the UFPF surface of INREU with the ground true

In this example, INREU is very efficient compared with the use of Equation (5-3) to calculate the ground truth. The total number of model evaluations is $N_{eval} = 88802$ as

listed in Table 5-1. In each level of the pilot SS (SS_1 and SS_2), we draw 1000 parameter samples $\{\theta\}$ to estimate the original failure probability which involves 1000 LSF evaluations. At the end of the pilot SS, 113 failure parameter samples $\{\theta\}_F$ and 1000 non-failure parameter samples $\{\theta\}_{\bar{F}}$ can be obtained, and the original failure probability is 0.0113. It is followed by the simulation of corresponding measurement samples $\{d\}$ (SS_d), which requires a total of 1130 predictive function evaluations. The iterative density estimation is conducted separately for $\{d\}_F$ and $\{d\}_{\bar{F}}$. For $\{d\}_F$, the sample size is increased from 113 to 4068 in the initial iteration ($k = 0$), and from 407 to 4070 in the first iteration ($k = 1$). Likewise, for $\{d\}_{\bar{F}}$, the sample size is increased from 1000 to 4000 in the initial iteration ($k = 0$); in the following iterations ($k = 1,2,3$), 10% of the samples in the previous iteration are used and its sample size is increased from 400 to 4000. Note that we keep every second sample in the Markov chain to reduce autocorrelation. The efficiency of INREU is more remarkable for cases where values of both LSF and predictive function can be obtained simultaneously within one simulation.

Table 5-1: Number of model evaluations and samples during the approximation of UFPE

Number of evaluations or samples	SS_1	SS_2	SS_D		$k = 0$	$k = 1$	$k = 2$	$k = 3$
LSF evaluations	1000	1000	0	F	$113 \times (2 \times 35)$	$407 \times (2 \times 9)$	--	--
				$\bar{F}$	$1000 \times (2 \times 3)$	$400 \times (2 \times 9)$	$400 \times (2 \times 9)$	$400 \times (2 \times 9)$
Predictive function evaluations	0	0	1130	F	$113 \times (2 \times 35)$	$407 \times (2 \times 9)$	--	--
				$\bar{F}$	$1000 \times (2 \times 3)$	$400 \times (2 \times 9)$	$400 \times (2 \times 9)$	$400 \times (2 \times 9)$
Failure samples $\{d\}_F$	0	0	113		4068	4070	--	--
Non-failure samples $\{d\}_{\bar{F}}$	0	0	1000		4000	4000	4000	4000

Effect of the sample size on the UFPF

MISE is clearly related to sample size as indicated in Figure 5-12. *MISE* decreases with larger sample size and converges when sample size exceeds 4000. *MSEs* with respect to different sample size for two measurements $[d_1, d_2] = [-2 \times 10^{-3}, 1.3 \times 10^{-3}]$ and $[d_1, d_2] = [-1.5 \times 10^{-3}, 8 \times 10^{-4}]$ are shown in Figure 5-13. The convergence of errors can be observed clearly for both measurements. Compared with the previous example where only one measurement is considered, this example requires a larger sample size for convergence.

Figure 5-12: Change of the global evaluation metric *MISE* with the increase of sample size

Figure 5-13: Change of the pointwise evaluation metric *MSE* with the increase of sample size

5.5 Conclusions

Now that system safety management requires prompt responses to monitoring or inspection data, this chapter offers a negligible latency-time solution for Bayesian reliability updating by proposing a novel framework INREU. The main novelty of this framework is that the updated failure probability function (UFPF) is constructed based on iterative density estimation, enabling the instant plug-in query of updated reliability. In the iterative density estimation scheme, failure samples are propagated from the high-density regions to the low-density ones, thus approximating the UFPF globally over the entire measurement range of interest.

Two simplified but representative illustrative examples were studied to show the accuracy and efficiency of the proposed framework in comparison with the ground truth. The fact that errors converge quickly with the increase of sample size indicates its usefulness in dealing with real-world systems in practice.

It is worth noting that the required sample size increases when more measurements are investigated. It is expected as a natural result of density estimation. Fortunately, as a sample-adaptive density estimation technique, the computational complexity of BSP is linear rather than exponential to the sample dimension. The application of INREU to cases with more measurements will be further investigated.

In contrast to conventional reliability updating, INREU can obtain the updated reliability instantly by plugging the measurements into the UFPF. Moreover, besides the probability of system failure, there is a great potential in updating the probability of other rare events such as the occurrence of natural hazards.

6 Applications to deep excavations considering the spatial variability of soil properties

Summary

In the civil engineering community, the soil in its natural state is arguably the most variable engineering material. It has long been recognized that the spatial variability of soil properties plays an important role in the probabilistic analysis of geotechnical problems. In this section, DOPADS (or DOPADS_m) is applied to the design optimization of a type of geotechnical systems – deep excavations considering different sources of uncertainties. Also, INREU is used to perform reliability updating using the monitoring data collected during the construction phase, thus ensuring the system under study sustains a certain level of reliability and providing information for the potential design change. Results show that the two proposed frameworks are computationally efficient, are flexible to implement, and are well suited for deep excavation problems that may involve sophisticated numerical models or/and a large number of random variables.

6.1 Modeling of deep excavations

6.1.1 Traditional models

For an excavation system, appropriate models are required for the analysis of the stability and deformation.

Stability analysis

History of underground space development sees many tragedies due to the collapse of excavations (instability). Stability analysis can be classified into two groups, i.e., internal and external stability. The former is related to the failure of the earth-retaining system (walls, struts, etc.), while the main concerns of the latter include the overall instability and basal heave. Here we introduce the overall instability and basal heave (see Figure 6-1).

Figure 6-1: Two modes of external instability failure in an excavation (Ou 2006)

The overall stability analysis is of great importance for projects in soft soil or/and sloping ground. In traditional models, the potential failure surface is assumed to be circular or of other geometries. Simple limit equilibrium analysis or computational limit analysis is conducted based on soil strength parameters.

Basal heave, which is a failure mode induced by excessive bottom heave, may occur in soft soil. Typical traditional methods for the basal heave analysis are Terzaghi's method (Terzaghi 1944) and Bjerrum and Eide's method (Bjerrum and Eide 1956).

Deformation analysis

Empirical and semi-empirical models (Peck 1969, Clough and O'Rourke 1990, Hsieh and Ou 1998, Long 2001, Yoo 2001, Moormann 2004, Leung and Ng 2007) were proposed to predict the soil deformation (or the ground movement) induced by excavation, e.g., the maximum horizontal displacement and the surface settlement profile. Based on field observations, Peck (1969) firstly proposed a design chart of excavation-induced ground

settlement for three different types of soil. Clough and O'Rourke (1990) further investigated the field observations of case studies, and proposed different settlement envelopes for different types of soil. They found that the settlement envelopes for sand or stiff clay tend to be triangular; while for soft to medium clay, trapezoidal envelopes were observed. Hsieh and Ou (1998) proposed a method to predict the settlement profiles for both spandrel and concave types using a regression analysis.

6.1.2 Numerical models in commercial codes

In recent years, a wide spectrum of commercial codes using numerical modelling techniques have been developed, thanks to the emerging computational techniques and the availability of higher computational power. These codes have become a standard tool for practitioners, as they can accommodate advanced constitutive models, model complex geometries and boundary conditions, etc.

Mathematical models of geotechnical systems are described by partial differential equations (PDEs). As solving PDEs analytically is tedious, numerical methods have been developed to obtain approximate solutions. The two classical methods dominate the geoengineering community, i.e., the finite difference method (FDM), and the finite element method (FEM). The FDM is an old and simple method based on the local Taylor expansions. Traditional FDM, where elements in the discretization mesh are topologically square, cannot handle discontinuities and complex geometries. However, given a uniform structured mesh, the FDM is generally advantageous over FEM in terms of computational efficiency (Morton and Mayers 2005). Rather than the differential form of the governing equations in the FDM, the FEM uses the integral form counterpart which does not rely on special mesh structures.

Last decades have seen the development of several commercial codes in geoengineering community. FLAC (Fast Lagrangian Analysis of Continua), released in 1986 by Itasca Consulting Group, was a two-dimensional explicit finite difference program for simulating the behavior of soil, rock, structure, etc. Though FLAC is a FDM software, its algebraic equations are identical to those solved by the FEM, and solved by dynamic relaxation. Several commercial codes based on the FEM have been widely used in the geoengineering

practice, such as Plaxis, Zsoil, and Slope/W. In addition to codes specializing in geoenvironmental analysis, general finite element codes Abaqus and Ansys provide more powerful tools for pre-processing and post-processing, thus facilitating the analysis of large-scale complex geotechnical systems.

While the deformation analysis is quite straightforward, the stability analysis requires a thorough investigation of the stress-strain state in these codes. There are two ways to conduct the stability analysis: the limit equilibrium method (LEM) and the strength reduction method (SRM). The LEM is a traditional method in which the factor of safety (FoS) is defined as the ratio of the resistance force and the driving force. Unlike the LEM whose FoS is in terms of force, the FoS of SRM is computed by gradually reducing the shear strength until the excavation collapses. The SRM is commonly integrated into the commercial codes to solve sophisticated problems as essentially failure arises from the shear stress exceeding the shear strength. Comparison of the LEM and the SRM can be found in publications by Duncan (1996), Cheng et al. (2007), Wei et al. (2009), Sloan (2013).

6.1.3 Uncertainty sources

There are two major sources of uncertainties in the modeling of deep excavations. The first one is attributed to soil properties. Theoretically, for a specific body of soil, its properties are deterministic in a given space and time. But in practice, the values of soil properties are unknown to geotechnical engineers until limited site investigations as well as laboratory experiments on soil samples are performed. Uncertainties, due to incomplete knowledge in soil properties, are categorized into parametric uncertainty (i.e., inherent soil spatial variation, which is caused by factors of soil formation), experimental uncertainty (variability of experimental measurements), structural uncertainty (caused by lack of knowledge of the underlying structure of data and model), statistical uncertainty (caused by lack of available data).

Another source is attributed to the modelling of the geotechnical system. Mathematical models are commonly used to represent the underlying physical model of a geotechnical system. Considering the fact that models are almost always only approximations to reality, particularly involving soil constitutive relation and soil-structure interaction, discrepancies

are expected between the model prediction and the real system response. In addition, some mathematical models are too complicated to be solved accurately, and so their solutions can only be approximated by numerical methods (e.g., FEM and FDM are commonly adopted in solving PDEs in commercial geotechnical codes). The calculated response is calibrated by the model factor, which is defined as the ratio of the calculated response to the measured response.

6.2 Random field and its discretization

Random field theory (Vanmarcke 1977) is considered as a rigorous treatment of spatial variability. Consider a continuous random field $\Theta(\mathbf{t})$ that represents a collection of random variables indexed by continuous coordinates $\mathbf{t}$. In geotechnical engineering, $\mathbf{t} \in \mathbb{R}^d$ where d is the dimension of the random field. For instance, $\mathbf{t} = [x, z]$ if the spatial variability in the x and z directions is studied in a two-dimensional random field. It is of practical interest to discrete a continuous random field into a finite number of stochastic elements, each of which can be represented by a random vector. That is, $\Theta(\mathbf{t}) = \hat{\Theta}(\mathbf{t}) = \Theta = [\theta_1, \dots, \theta_{n_\theta}]^T$ where n_θ is the number of random variables.

Different discretization schemes have been established and comprehensive literature reviews can be found in (Sudret and Der Kiureghian 2000) and (Stefanou 2009). In the following, the midpoint method (Der Kiureghian and Ke 1988) that uses the statistical characteristics at the centroid to represent the property over the entire stochastic element is adopted. For the j -th stochastic element, its associated random variable can be given by $\theta_j = \Theta(\mathbf{t}_j)$ where $\mathbf{t}_j$ is a vector of element centroid coordinates.

It should be noted that the stochastic elements for the discretization of random fields are not necessarily the elements used in the numerical methods (e.g., finite element or finite difference method). While the size of the former is determined by the characteristics of random fields (Ching and Phoon 2013), the size of the latter is determined by the distribution of stress and strain in the numerical model (e.g., small size elements are needed in the stress concentration region to guarantee the accuracy). Figure 6-2 illustrates the stochastic elements in the realizations of soil profiles. The dimension of stochastic elements is $2.5m \times 2.5m$.

A random field $\Theta(\mathbf{t})$ can be fully characterized by the inherent mean $\mathbb{E}(\Theta(\mathbf{t}))$, the inherent variance $\mathbb{V}(\Theta(\mathbf{t}))$ and the autocorrelation function $\rho(\mathbf{t}_i, \mathbf{t}_j)$. The autocorrelation function describes the spatial correlation between two points $\mathbf{t}_i$ and $\mathbf{t}_j$ in a random field and is defined as:

$$\rho(\mathbf{t}_i, \mathbf{t}_j) = \frac{CV(\Theta(\mathbf{t}_i), \Theta(\mathbf{t}_j))}{\sqrt{\mathbb{V}(\Theta(\mathbf{t}_i))} \sqrt{\mathbb{V}(\Theta(\mathbf{t}_j))}} \quad (6-1)$$

where $CV(\Theta(\mathbf{t}_i), \Theta(\mathbf{t}_j))$ denotes the covariance of $\Theta(\mathbf{t}_i)$ and $\Theta(\mathbf{t}_j)$. For stationary random fields, $\rho(\mathbf{t}_i, \mathbf{t}_j)$ is determined by the relative location between two points and thus can be simplified to $\rho(\boldsymbol{\tau}_{ij})$ where $\boldsymbol{\tau}_{ij} = \mathbf{t}_i - \mathbf{t}_j$. For example, the single exponential autocorrelation function for the two-dimensional random field is expressed as:

$$\rho(\boldsymbol{\tau}_{ij}) = \exp\left(-\frac{2|x_i - x_j|}{l_x} - \frac{2|z_i - z_j|}{l_z}\right) \quad (6-2)$$

where l_x and l_z are the length of fluctuation (LoF) in the x and z directions, respectively. The LoF describes the distance within which soil properties at two points will tend to be highly correlated. Therefore, as shown in Figure 6-2, a smaller LoF produces a rougher random field while a larger one produces a smoother random field.

Figure 6-2: Realizations of random fields with different length of fluctuations

6.3 Design optimization of a two-dimensional deep excavation against basal heave

6.3.1 Case description

The studied excavation site is $20m$ wide and $10m$ deep as shown in Figure 6-3. The diaphragm walls supported by three levels of struts are used as the earth-retaining system, among which the depth of the final strut h_s is $8.5m$.

The soil profile in Figure 6-3 is a typical example of excavation sites in Singapore. As the properties of the upper marine clay dominate the performance of the earth-retaining system, only the soil properties of this layer are treated as uncertain. Specifically, two soil properties, i.e., the density ρ and normalized shear strength s_u/σ'_v (i.e., undrained shear strength over the effective overburden stress), are modeled by random fields simultaneously in the calculation of the reliability against basal heave. The slip circle method is adopted for basal heave analysis as detailed in Appendix B. In this case, the LoFs $l_x = 12m$ in the x direction and $l_z = 5m$ in the z direction in conjunction with the single exponential autocorrelation function (Equation (6-2)) are assumed to characterize the spatial dependence for both ρ and s_u/σ'_v . Both random fields are discretized based on the midpoint method using an identical stochastic mesh scheme where the size of stochastic elements is $5m$ and $2.5m$ in the x and z directions respectively, which is about half of the LoFs. For the dimension of the soil profile being $60m$ wide and $25m$ deep, it results in $12 \times 10 \times 2 = 240$ random variables. Additionally, the surcharge q is modeled as a spatially-constant random variable. Table 6-1 lists the statistical characteristics of these random variables.

Figure 6-3: Profile of the braced excavation site (not to scale)

Table 6-1: Summary of statistical characteristics of random variables

Random variable	Distribution	Mean	CoV
Density ρ (kg/m^3)	Lognormal	1900	0.10
Normalized shear strength s_u/σ'_v	Lognormal	0.25	0.30
Surcharge q (kPa/m)	Normal	10.0	0.20

The objective of this design optimization problem (see Figure 6-4) is to minimize the depth of the diaphragm wall h_w , which implies the minimization of the construction cost. It is subject to a reliability constraint $P_F(h_w) \leq [P_F]$ where the failure event F is the basal heave of the excavation defined as the driving moment M_d exceeding the resistance moment M_r (see Appendix B for the calculation of M_d and M_r). The allowable failure probability $[P_F]$ depends on the perceived economic losses and social consequences. First, the case $[P_F] = 10^{-3}$ is investigated and the effects of $[P_F]$ on the optimal design will be discussed later.

$$\begin{array}{ll}
 \min: & h_w \\
 \text{s. t. :} & \begin{cases} P_F(h_w) \leq [P_F] \\ 12m \leq h_w \leq 24m \end{cases} \\
 \text{where:} & g(h_w, \boldsymbol{\theta}) = M_r(h_w, \boldsymbol{\theta}) - M_d(h_w, \boldsymbol{\theta})
 \end{array}$$

Figure 6-4: Formulation of the design optimization problem of a deep excavation

6.3.2 Results and discussions

Optimal design

The DOPADS optimization procedure starts with the approximation of the FPF. First, in the pilot stochastic simulation, the failure probability of the augmented system is calculated $P_F = 0.0842$ and failure samples with the size of 1000 are generated. As most of failure samples at this stage are scattering in the high-density region, we adopt the proposed iterative density estimation scheme to improve the accuracy of density estimation. The iterative procedure stops after six iterations as the failure probability that corresponds to the partition density value becomes smaller than $[P_F]$. The density values in each iteration using BSP are treated as support points and smoothed to obtain the final density estimator. Figure 6-5 presents the final FPF along with failure samples generated in each iteration. It should be noted that failure samples in Figure 6-5 are subsampled to the size of 50 for visualization purpose.

Figure 6-5: Approximate failure probability function based on iterative density estimation

To validate the approximate FPF, it is compared with the FPF calculated in a brute-force way. That is, the failure probabilities for 48 evenly spaced design configurations h_w between $12m$ and $24m$ are calculated using the crude MCS. Figure 6-6 shows that the final FPF ($k = 5$) agrees well with the result obtained by the crude MCS.

Additionally, the approximate FPFs based on failure samples from the first five iterations are also illustrated in Figure 6-6. It is clear that the approximate FPF converges to the brute-force FPF with the increase of iterations. In fact, the result of the initial iteration, which shows an unsatisfactory accuracy over the low-density region, is equivalent to the local approximation presented in (Ching and Hsieh 2007b, Wang 2009, Wang et al. 2011, Wang and Cao 2013) using expanded reliability-based design. Failure samples in the iterative density estimation scheme provide the distribution information over the entire design space, thus making the global approximation possible.

Given the final approximate FPF, the optimal design can be efficiently obtained by deterministic optimization routine SQP. The minimal diaphragm wall depth h_w^* satisfying the reliability constraint is found to be $20.87m$.

Figure 6-6: Comparison of the approximate failure probability function in each iteration and that by the crude Monte-Carlo simulation

Comparison to the double-loop approach

DOPADS in this paper differs from the double-loop approach in the way that reliability constraints are evaluated. DOPADS no longer requires any reliability analyses after approximating the FPF. The differences of these two methods are compared in terms of convergence and efficiency in this subsection. For comparison purpose, SQP is used for both approaches.

The change of the optimal design by the decoupled and double-loop approaches is illustrated in Figure 6-7 for three different initial points. Both approaches can converge to the optimal design less than ten iterations, but it takes more iterations for the double-loop approach to converge due to the variability of failure probability estimated in each iteration. Also, it is clear from Figure 6-7 that the initial points for optimization can affect the number of iterations before convergence. It should be noted that required number of iterations is crucial for the double-loop approach as it is positively related to the number of reliability analyses.

Figure 6-7: Change of the optimal design by (a) DOPADS and (b) double-loop approach

The efficiency of these methods is measured in terms of the total number of model evaluations. In this example, the model evaluation refers to the calculation of M_d and M_r for basal heave failure. Figure 6-8 shows the change of the number of model evaluations during the optimization iterations. For DOPADS, the model evaluations are only performed for approximating the FPF, thus the number of model evaluations remains constant

regardless of the initial points. On the other hand, the number of model evaluations will increase when the double-loop approach is used. Specifically, for a small initial point (16m), the total number of model evaluations increases drastically as the candidate optimal design moves into the low failure probability region; for the initial point (22m) close to the final optimal design (20.87m), the change is steady. After convergence, the number of model evaluations in the double-loop approach exceeds greatly that in DOPADS, indicating better efficiency of DOPADS.

Figure 6-8: Number of model evaluations required in DOPADS and the two-loop approach

Effect of statistical properties of random fields

The soil density ρ and normalized undrained shear strength s_u/σ'_v are the two most dominant parameters in the design of deep excavations against basal heave. The effects of their inherent variability on design optimization is investigated for a wide range of CoVs and mean values using DOPADS.

It can be noted from Figure 6-9 and Figure 6-10 that the optimal wall depth becomes larger as CoV_ρ and CoV_{s_u/σ'_v} increase, implying more cost is needed for an excavation site whose soil properties have greater inherent variability. For mean values $m_\rho = 1700, 1800, 1900 \text{ kg/m}^3$, the wall depth increases with the soil density as shown in Figure

6.3. Design optimization of a two-dimensional deep excavation against basal heave

6-9 in order to restrict the probability of the basal heave; oppositely, Figure 6-10 indicates that one can design shorter diaphragm walls for a site with larger m_{s_u/σ'_v} . It can be explained by the fact that ρ contributes to the driving moment part while s_u/σ'_v contributes to the resistance moment part in the failure event formulation of basal heave. It can be observed that in Figure 6-10 the optimal wall depth increases drastically from 19.8m to 60.0m for $m_{s_u/\sigma'_v} = 0.20$ with the growth of CoV_{s_u/σ'_v} , meaning the optimal design becomes sensitive to the change of the variability of normalized undrained shear strength when the mean value of s_u/σ'_v is small.

Figure 6-9: Optimal wall depth against the CoV of soil density for different mean values

Figure 6-10: Optimal wall depth against the CoV of normalized undrained shear strength for different mean values

Besides inherent variability, spatial variability which can be characterized by the LoF should be taken into account for RBDO. In this case, the effects of vertical (z) and horizontal (x) LoF are studied separately. That is, the spatial variability in one direction is ignored by assigning a large LoF $500m$, while different values of LoF ($5m, 10m, 20m, 40m, 60m, 80m$) in the other direction were investigated. Note that the LoF in the z direction is much smaller than that in the x direction due to soil deposition, and values of LoF vary considerably from site to site. For illustrative purpose, the same set of LoF was used for both x and z directions. As shown in Figure 6-11, it is clear that smaller LoF in either x or z direction will result in smaller optimal wall depth. The result is consistent with the fact that smaller LoF of a random field with fixed inherent variability means smaller integrated variability, in which case the shorter earth-retaining wall is needed to stabilize the system against the basal heave.

Figure 6-11: Optimal wall depth against the length of fluctuation in x and z directions

Effect of different allowable failure probabilities

The results presented so far are calculated for a preassigned allowable failure probability $[P_F] = 10^{-3}$. It is not uncommon that allowable failure probability changes when the

decision makers reassess the acceptable risk level. Unlike coupled approaches in which one needs to rerun the whole RBGD if the allowable failure probability in reliability constraints is changed, in DOPADS we still can proceed to the deterministic optimization given the calculated FPF without additional model evaluations. 30 replicates are run for this case study and in each replicate, the corresponding values of optimal wall depth are obtained for $[P_F] = 0.1, 10^{-2}, 10^{-3}$. The boxplots in Figure 6-12 display the distribution of optimal points. The range of variation for $[P_F] = 10^{-3}$ remains small and no outlier exists (the maximum whisker length is specified as the interquartile range).

Figure 6-12: Optimal wall depth for different allowable failure probabilities

6.3.3 Conclusions

The variability of the Singapore marine clay properties (soil mass density and normalized undrained shear strength) greatly affects the optimal design result of the deep excavation against basal heave. For the inherent variability, the larger mean soil mass density results in the longer wall depth of the optimal design while the mean value of normalized undrained shear strength acts conversely; the variance of both soil properties can increase the length of the earth-retaining wall needed. For spatial variability, smaller length of fluctuation in either x or z direction will result in smaller optimal wall depth.

DOPADS is useful to the geotechnical design practitioners due to its accuracy and efficiency as demonstrated in the design example of a deep excavation against basal heave.

It provides great flexibility to numerical modeling, spatial variability modeling, reliability constraints, and deterministic optimization algorithms. It can be potentially applied to other geotechnical problems.

6.4 Update the reliability of an excavation using monitoring data

In the construction of an excavation, the maximum horizontal deflection of the earth-retaining wall is monitored closely to ensure the reliability of the system. Since the monitoring data is indirectly related to system responses (e.g., the ground settlement in this section) defining the system failure, model updating methods should be used to extract and utilize the information in the monitoring data. In this section, we adopt INREU to update the reliability of an excavation.

The excavation is 24 m wide and 7 m deep. The soil profile of the excavation is typical among many construction sites in Singapore as shown in Figure 6-13. Different from the case in Section 6.3, a sheetpile wall with three levels of temporary steel struts rather than a diaphragm wall is used as the earth-retaining system in this section.

Figure 6-13: Profile of the excavation site (not to scale)

In this example, the maximum horizontal deflection δ_{hm} of the sheetpile wall is treated as the measurement. The failure is defined as the maximum ground settlement δ_{vm}

exceeding its allowable value $[\delta_{vm}] = 26 \text{ mm}$. A commercial code FLAC (Fast Lagrangian Analysis of Continua, ITASCA), which employs the finite difference method for the solution of PDEs, is used to calculate δ_{hm} and δ_{vm} in each simulation. Figure 6-14 shows the 2D model and its finite difference mesh in FLAC. It can be observed that the mesh has been densified where stress concentration and rapid strain changing may occur to ensure accurate simulation.

Figure 6-14: 2D model and its mesh of the excavation site

Three soil properties, i.e., Young's modulus E , Density ρ , and Poisson's ratio ν , are identified as the key factors to the ground settlement and the sheetpile wall deflection as shown in Figure 6-15. Their spatial variability is modeled by homogeneous random fields whose statistical properties are summarized in Table 6-2. An exponential correlation function in Equation (6-2) is adopted to characterize their spatial dependence where $l_x = 12m$ and $l_z = 5m$.

The midpoint method is adopted to discretize a random field into random variables where statistical characteristics at the centroid of a stochastic element represents the variability over the entire stochastic element. As shown in the stochastic mesh in Figure 6-16, the size of stochastic elements is $\bar{l}_x = 6m$ and $\bar{l}_z = 2.5m$, resulting in 312 random variables for each random field. Thus, the total number of uncertain model parameters involved in this model is $312 \times 3 = 936$. The prediction error of the maximum wall deflection follows the Gaussian distribution with mean 0 and standard deviation 1 mm . Comparing Figure 6-14 and Figure 6-16, it can be noted that the stochastic mesh is not necessarily the same to the finite difference mesh as the former is determined by the spatial

variability whereas the latter is determined by the distribution of stress and strain in the numerical model.

Figure 6-15: Parametric study and the importance of soil properties. (a) δ_{hm} and δ_{vm} calculated using different soil properties; (b1) the importance of soil properties to δ_{hm} ; (b2) the importance of soil properties to δ_{vm} . The importance is measured by the Pearson correlation coefficient.

Table 6-2: Summary of statistical characteristics of uncertain soil properties

Soil property	Distribution	Mean	CoV
Young's modulus E (MPa)	Lognormal	12	0.20
Density ρ (kg/m ³)	Lognormal	1700	0.10
Poisson's ratio ν	Lognormal	0.45	0.02

Figure 6-16: Stochastic mesh of a random field and one realization of its stochastic elements

As each simulation takes about two minutes in a desktop with four-core i5 CPU, it is not affordable to evaluate results using MISE or MSE metric in which millions of simulations are required. Instead, five independent replicates were run to approximate the UFPF to illustrate its variance. The average original failure probability is 0.2680. Two iterations are run in the iterative density estimation for both failure and non-failure samples. Note that by using finite difference method, one can obtain the results of predictive function and LSF in one model evaluation. The average number of model evaluations in INREU is 17100.

As shown in Figure 6-17, in the domain of interest $\delta_{hm} \in [30, 90]$ (mm), the failure probability increases with the increase of the measured maximum deflection of the sheetpile wall. Also, the discrepancy between different replicates turns out to be small, indicating small variance of the approximated UFPF.

Figure 6-17: Updated failure probability function in terms of the measured maximum wall deflection

7 Concluding remarks and future work

7.1 Conclusions

In this thesis, a novel framework DOPADS is proposed for stochastic complex systems to deal with reliability-based design optimization (RBDO) problems. The efficiency of DOPADS is further enhanced for system design problem with multiple reliability constraints by taking advantage of the dependency between failure events. Additionally, a framework called INREU is proposed to update the designed system instantly once the measured data are available.

Chapter 2 reviews the existing literature on the methods for uncertainty-based analysis with special emphasis on the stochastic-sampling based ones such as Monte Carlo simulation, importance sampling, Markov chain Monte Carlo methods, etc. As a basis of RBDO, a large body of research has been devoted to developing reliability methods, attempting to propagate uncertainties to the system responses of interest and quantify the failure probability fast and accurately. It can be concluded from the literature that the computational efforts required to perform the reliability analysis are prohibitively large when expensive-to-evaluate numerical models are involved, making RBDO tasks even more challenging due to the potentially nested reliability analyses in the optimization process. To alleviate this issue, different approaches including double-loop, single-loop and decoupling approaches have been proposed.

A novel framework DOPADS is detailed in Chapter 3, which addresses computational issues (i.e., computational burden and high dimensionality of random variables) in the design optimization of stochastic complex systems. In DOPADS, the failure probability function (FPF) is approximated globally in the partitioned design space, making sure there are sufficient information about failure events near the optimal solution. Illustrative

examples are used to demonstrate the implementation details as well as to validate DOPADS. Results reveal that

- (1) The optimal design by DOPADS can achieve a high accuracy which has been compared to the solution by the brute-force crude MCS. It can be attributed to the propagation of failure samples from the high-density region to the low-density one where the optimal design will be found in the context of engineering.
- (2) DOPADS is efficient. By decoupling the reliability analysis from the optimization loop, repeated reliability analyses can be avoided. Once decoupled, only deterministic optimization follows which needs a neglectable amount of computing time. In essence, in DOPADS non-overlapping failure samples will be generated in the partitioned design space, thus improving the efficiency in stochastic-sampling-based RBDO.
- (3) DOPADS enjoys flexibility in dealing with different kinds of RBDO problems. First, it can deal with a large number of random variables (e.g., 2036 in Section 4.4.2). Second, sophisticated numerical models (e.g., those simulated in commercial codes) can be incorporated in DOPADS in a nonintrusive way. Third, DOPADS is robust to the change of allowable failure probabilities due to the global approximate FPF. Fourth, it is possible for the user to choose alternative stochastic sampling techniques and density estimation techniques.

Chapter 4 extends DOPADS to deal with cases of multiple reliability constraints more efficiently. Instead of approximating FPFs separately, in the new framework DOPADS_m, FPFs are approximated simultaneously by taking advantage of the dependence between failure events. The efficiency enhancement, which is measured by the reduction of the number of model evaluations compared to that required in DOPADS, is affected by the dependency between failure events and the number of reliability constraints. It is important to highlight that the total number of samples in each iteration can be bounded by N_s and $n_{rc}N_s$ where n_{rc} is the number of reliability constraints and N_s is the required sample size per failure event. This confirms that DOPADS_m will never perform worse than DOPADS in terms of efficiency.

Design optimization is an ongoing process in which the initial design can be modified if the model of interest and its associated performance is updated using observed data. In Chapter 5, a new framework INREU for reliability updating was presented. In traditional reliability updating methods, first one needs to update the distribution of model parameters in a Bayesian approach, then calculates the failure probability of the updated model. The calculation of the failure probability shall take a very long time if the simulation involves expensive-to-evaluate numerical models, leading to lagged decisions regarding the current reliability state. In INREU, the updated failure probability function (UFPF) is constructed before collecting data, making the instant plug-in query of updated reliability possible. The main advantages of INREU are:

- (1) INREU provides a negligible latency-time solution to Bayesian reliability updating. It is of great significance to the systems that are highly uncertain and whose failure will lead to catastrophic results.
- (2) The accuracy of the approximate UFPF is ensured globally over the entire measurement range due to the iterative density estimation scheme.
- (3) Two representative illustrative examples have shown that INREU can be applied to real-world systems accurately and efficiently.

Finally, the proposed frameworks have been applied to the deep excavation cases in Singapore in Chapter 6. These cases are featured by (i) the spatial variability of the properties of the Singapore marine clay and (ii) complex systems involving soil-structure interaction and sophisticated soil constitutive relation. The spatial variability of soil properties is modeled by the random field theory. The results have been found that the variability of the density and normalized undrained shear strength of the Singapore marine clay greatly affects the optimal design of a deep excavation against basal heave. Smaller length of fluctuation in either x or z direction will result in smaller optimal depth of the earth-retaining wall.

Some limitations in this research much be pointed out, which include (i) the number of design variables in design optimization or measurements in reliability updating are less than ten even using the state-of-the-art density estimation techniques; (ii) it is difficult to collect sufficient data to support the determination of the distribution of random variables in real-

world projects; (iii) the computational burden (several tens of thousands of model evaluations) may remain prohibitive for some cases. However, the limitations above should not discourage the use of the proposed frameworks considering their benefits. Furthermore, we will partially address these limitations in the future works.

7.2 Future works

7.2.1 Surrogate for reliability-based design optimization

Surrogates (also referred to as metamodels or response surfaces) can be built in lieu of the original expensive-to-evaluate numerical models. Different types of surrogates can be constructed such as polynomials, Gaussian Processes, polynomial chaos expansion, artificial neural network and support vector machine.

The most critical part of efficiently building a surrogate is probably the adaptive infill strategy, also referred to as adaptive design of experiment or refinement scheme in some literature. Instead of densely filling the parameter space with support points, an adaptive infill strategy often starts with a coarse surrogate built upon a small number of support points, and then refine it progressively using judiciously selected support points. The selection of sequential support points is guided by active learning functions which are task-specific, i.e., the best next support points can be found by maximizing the information about the optimal solution as well as the parameter space. The computational burden can be alleviated to a large extent by adaptively-built surrogates.

Therefore, it is promising to further enhance the efficiency of the proposed frameworks with customized surrogates.

7.2.2 Reliability-based design optimization with discrete design variables

Besides continuous design variables, engineering design problems often have discrete ones such as material selection and size of elements. As for the design of deep excavations, the bracings are made in a factory based on specifications so that their dimensions can only be selected from a limited pool of candidates.

By far, only double-loop approach can be applied to RBDO problems with discrete design variables directly by adopting discrete programming techniques in the outer loop.

For single-loop and decoupling approaches, modifications are needed for the existing methods to deal with the mixture of discrete and continuous design variables. For DOPADS and DOPADS_m, the key to solving this problem is the approximation of the failure probability function at the presence of both discrete and continuous variables, which needs to be further studied in the future research.

7.2.3 Data-adaptive density estimation with confidence bands

Density estimation is a building block of the frameworks proposed in this thesis as it is found that the failure probability function and updated failure probability function are proportional to the density function. It will help to evaluate the accuracy of the approximate failure probability function theoretically if the confidence of density estimation can be assessed.

Traditional methods can build confidence bands of densities based on extreme value theory but the underlying true density must satisfy strict differentiability assumptions (Bickel and Rosenblatt 1973, Komlós et al. 1975). More recently, Giné and Nickl (2010) proposed a method for constructing data-adaptive confidence bands which is applicable to a wider class of densities. Nevertheless, these methods can only give confidence bands for densities involving no more than two random variables.

7.2.4 Multidisciplinary design optimization

Complex systems often involve multiple disciplines. In the design of an excavation, other requirements such as those on the minimal influence on the adjacent buildings, the design of dewatering schemes and the noise level of construction must be satisfied besides the primary safety requirements.

Multidisciplinary design optimization has emerged to search for the optimal solution while considering the coupling of all relevant disciplines since 1980th. It is advantageous over the traditional optimization method in which the optimal design is found by performing optimization separately for each discipline. However, there are at least two challenges that need to be addressed. First, it is difficult and even impossible to model some disciplines

such as human behavior. Second, there is little knowledge about the mechanism of the interactions between disciplines.

Appendix A: Proof of stationarity

The ‘detailed balance condition’, or ‘reversibility condition’, is a sufficient condition for $p_k(\boldsymbol{\varphi}, \boldsymbol{\theta}|F) \propto \mathbb{I}_F(\boldsymbol{\varphi}, \boldsymbol{\theta}) \mathbb{I}_{\mathcal{L}_{k-1}}(\boldsymbol{\varphi}) p_{\Phi}(\boldsymbol{\varphi}) p_{\Theta}(\boldsymbol{\theta})$ to be stationary distribution. That is, if the current sample $(\boldsymbol{\varphi}, \boldsymbol{\theta}) \sim p_k(\boldsymbol{\varphi}, \boldsymbol{\theta}|F)$, the next sample $(\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1}$ will also be distributed as $p_k(\boldsymbol{\varphi}, \boldsymbol{\theta}|F)$. For $(\boldsymbol{\varphi}, \boldsymbol{\theta}), (\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} \in \Omega_{\varphi} \times \Omega_{\theta} \subseteq \mathbb{R}^{n_{\varphi}+n_{\theta}}$, the detailed balance condition is given by:

$$p((\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} | (\boldsymbol{\varphi}, \boldsymbol{\theta})) p_k((\boldsymbol{\varphi}, \boldsymbol{\theta}) | F) = p((\boldsymbol{\varphi}, \boldsymbol{\theta}) | (\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1}) p_k((\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} | F) \quad (\text{A-1})$$

Considering the sample level of ‘accept-reject’ that enforces all failure samples to lie in $\mathcal{L}_{k-1}$ in the modified Metropolis-Hastings algorithm discussed in Section 3.3.2, it is sufficient to check the detailed balance condition for $(\boldsymbol{\varphi}, \boldsymbol{\theta}), (\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} \in \Omega_{\varphi, F} \cap \mathcal{L}_{k-1} \subseteq \mathbb{R}^{n_{\varphi}+n_{\theta}}$

$$p((\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1} | (\boldsymbol{\varphi}, \boldsymbol{\theta})) p_{\Phi}(\boldsymbol{\varphi}) p_{\Theta}(\boldsymbol{\theta}) = p((\boldsymbol{\varphi}, \boldsymbol{\theta}) | (\boldsymbol{\varphi}, \boldsymbol{\theta})_{+1}) p_{\Phi}(\boldsymbol{\varphi}_{+1}) p_{\Theta}(\boldsymbol{\theta}_{+1}) \quad (\text{A-2})$$

As a vector of dependent non-Gaussian random variables can always be transformed into a vector of independent standard normal ones by isoprobabilistic transformation T (Rosenblatt 1952, Nataf 1962, Lebrun and Dutfoy 2009), it is assumed that random variables in $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ are independent to each other. Thus, Equation (A-2) can be written as:

$$\begin{aligned} \prod_{j=1}^{n_{\varphi}} p_j(\varphi_{j,+1} | \varphi_j) p(\varphi_j) \prod_{j=1}^{n_{\theta}} p_j(\theta_{j,+1} | \theta_j) p(\theta_j) \\ = \prod_{j=1}^{n_{\varphi}} p_j(\varphi_j | \varphi_{j,+1}) p(\varphi_{j,+1}) \prod_{j=1}^{n_{\theta}} p_j(\theta_j | \theta_{j,+1}) p(\theta_{j,+1}) \end{aligned} \quad (\text{A-3})$$

It is sufficient for Equation (A-2) to hold true if:

$$\text{For } j = 1, \dots, n_{\varphi}, p_j(\varphi_{j,+1} | \varphi_j) p(\varphi_j) = p_j(\varphi_j | \varphi_{j,+1}) p(\varphi_{j,+1}) \quad (\text{A-4a})$$

$$\text{For } j = 1, \dots, n_{\theta}, p_j(\theta_{j,+1} | \theta_j) p(\theta_j) = p_j(\theta_j | \theta_{j,+1}) p(\theta_{j,+1}) \quad (\text{A-4b})$$

For succinctness, only Equation (A-4a) is proved as Equation (A-4b) can be derived easily by replacing φ with θ . For $\varphi_{j,+1} \neq \varphi_j$ (when the component is accepted), the LHS of Equation (A-4a) equals to the RHS

$$\begin{aligned} p_j(\varphi_{j,+1}|\varphi_j)p(\varphi_j) &= p(\varphi_j)q_j(\varphi_{j,+1}|\varphi_j)\min\left(1, \frac{p(\varphi_{j,+1})q_j(\varphi_j|\varphi_{j,+1})}{p(\varphi_j)q_j(\varphi_{j,+1}|\varphi_j)}\right) \\ &= p(\varphi_{j,+1})q_j(\varphi_j|\varphi_{j,+1})\min\left(1, \frac{p(\varphi_j)q_j(\varphi_{j,+1}|\varphi_j)}{p(\varphi_{j,+1})q_j(\varphi_j|\varphi_{j,+1})}\right) \\ &= p_j(\varphi_j|\varphi_{j,+1})p(\varphi_{j,+1}) \end{aligned} \tag{A-5}$$

Appendix B: Slip circle method for basal heave analysis

In Section 6.3, the failure event F -basal heave is considered to occur when the driving moment M_d exceeds the resistance moment M_r . As shown in Figure B-1, the moments can be expressed as:

$$M_d = \frac{Wr}{2} + \frac{qr^2}{2} \quad (\text{B-1a})$$

$$M_r = r \int_0^{\pi/2+\alpha} s_u \cdot r \cdot d\omega \quad (\text{B-1b})$$

where $r = h_w - h_s$ is the radius of the failure circle, W is weight of the soil mass between the wall and vertical failure plane as well as above the excavation bottom surface, q is the surcharge, and s_u is the undrained shear strength.

Figure B-1: Schematic diagram of slip circle method

The spatial variability of the density ρ and normalized shear strength s_u/σ'_v are modeled by random fields. For discretized random fields, samples of ρ are simulated first in each stochastic element, which are used to calculate the weight of soil mass W in Equation (B-1a) and the effective overburden stress σ'_v ; then samples of s_u/σ'_v are simulated, and s_u in each stochastic element can be obtained as product of s_u/σ'_v and the calculated σ'_v .

Bibliography

- Agarwal, H., C. K. Mozumder, J. E. Renaud and L. T. Watson (2007). "An inverse-measure-based unilevel architecture for reliability-based design optimization." Structural and Multidisciplinary Optimization **33**(3): 217-227.
- Aoues, Y. and A. Chateauneuf (2010). "Benchmark study of numerical methods for reliability-based design optimization." Structural and Multidisciplinary Optimization **41**(2): 277-294.
- Asmussen, S. and P. W. Glynn (2007). Stochastic simulation: Algorithms and analysis, Springer Science & Business Media.
- Au, S.-K. and J. L. Beck (2001a). "Estimation of small failure probabilities in high dimensions by subset simulation." Probabilistic Engineering Mechanics **16**(4): 263-277.
- Au, S. K. (2005). "Reliability-based design sensitivity by efficient simulation." Computers and Structures **83**(14): 1048-1061.
- Au, S. K. and J. L. Beck (2001b). "Estimation of small failure probabilities in high dimensions by subset simulation." Probabilistic Engineering Mechanics **16**(4): 263-277.
- Bansal, S. (2015). "A new Gibbs sampling based Bayesian model updating approach using modal data from multiple setups." International Journal for Uncertainty Quantification **5**(4).
- Bansal, S. and S. H. Cheung (2017). "On the evaluation of multiple failure probability curves in reliability analysis with multiple performance functions." Reliability Engineering & System Safety **167**: 583-594.
- Beck, J. L. (2010). "Bayesian system identification based on probability logic." Structural Control and Health Monitoring **17**(7): 825-847.
- Beck, J. L. and L. S. Katafygiotis (1998). "Updating models and their uncertainties. I: Bayesian statistical framework." Journal of Engineering Mechanics **124**(4): 455-461.
- Bernardo, J. M. and A. F. M. Smith (2008). Bayesian Theory.
- Bichon, B. J., S. Mahadevan and M. Eldred (2009). Reliability-based design optimization using efficient global reliability analysis. Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and material conference.
- Bickel, P. J. and M. Rosenblatt (1973). "On some global measures of the deviations of density function estimates." The Annals of Statistics: 1071-1095.
- Bjerrum, L. and O. Eide (1956). "Stability of strutted excavations in clay." Geotechnique **6**(1): 32-47.
- BLS (2010). Census of fatal occupational injuries (2000-2009). Washington, D.C., Bureau of Labor Statistics.

- Boggs, P. T. and J. W. Tolle (2000). "Sequential quadratic programming for large-scale nonlinear optimization." Journal of computational and applied mathematics **124**(1-2): 123-137.
- Boyd, S. and L. Vandenberghe (2004). Convex optimization, Cambridge university press.
- Chen, X., T. Hasselman, D. Neill, X. Chen, T. Hasselman and D. Neill (1997a). Reliability based structural design optimization for practical applications. 38th Structures, structural dynamics, and materials conference.
- Chen, X., T. K. Hasselman and D. J. Neill (1997b). Reliability based structural design optimization for practical applications. Proceedings of the 38th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference.
- Cheng, Y., T. Lansivaara and W. Wei (2007). "Two-dimensional slope stability analysis by limit equilibrium and strength reduction methods." Computers and Geotechnics **34**(3): 137-150.
- Cheung, S. H. and S. Bansal (2017). "A new Gibbs sampling based algorithm for Bayesian model updating with incomplete complex modal data." Mechanical Systems and Signal Processing **92**: 156-172.
- Cheung, S. H. and J. L. Beck (2007). New stochastic simulation method for updating robust reliability of dynamic systems. Proceedings of the 8th Engineering Mechanics Division Conference.
- Cheung, S. H. and J. L. Beck (2009). "Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters." Journal of Engineering Mechanics **135**(4): 243-255.
- Cheung, S. H., T. A. Oliver, E. E. Prudencio, S. Prudhomme and R. D. Moser (2011). "Bayesian uncertainty analysis with applications to turbulence modeling." Reliability Engineering & System Safety **96**(9): 1137-1149.
- Ching, J. and Y. C. Chen (2007). "Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging." Journal of Engineering Mechanics **133**(7): 816-832.
- Ching, J. and Y.-H. Hsieh (2007a). "Approximate reliability-based optimization using a three-step approach based on subset simulation." Journal of Engineering Mechanics **133**(4): 481-493.
- Ching, J. and Y.-H. Hsieh (2007b). "Local estimation of failure probability function and its confidence interval with maximum entropy principle." Probabilistic Engineering Mechanics **22**(1): 39-49.
- Ching, J. and Y.-H. Hsieh (2009). "Updating future reliability of nonlinear systems with low dimensional monitoring data using short-cut simulation." Computers & Structures **87**(13): 871-879.
- Ching, J., M. Muto and J. L. Beck (2006). "Structural model updating and health monitoring with incomplete modal data using Gibbs sampler." Computer-Aided Civil and Infrastructure Engineering **21**(4): 242-257.
- Ching, J. and K. K. Phoon (2013). "Effect of element sizes in random field finite element simulations of soil shear strength." Computers and Structures **126**(1): 120-134.
- Clough, G. W. and T. D. O'Rourke (1990). Construction induced movements of insitu walls. Design and performance of earth retaining structures, ASCE.

- Committee of Inquiry (2005). Report of the Committee of Inquiry into the incident at the MRT Circle Line worksite that led to the collapse of the Nicoll Highway on 20 April 2004. Singapore, Ministry of Manpower.
- Cornell, C. A. (1969). "A Probability-based Structural Code." Journal of the American Concrete Institute **66**(12): 974-985.
- Dempster, A. P. (1967). "Upper and lower probabilities induced by a multivalued mapping." The annals of mathematical statistics: 325-339.
- Der Kiureghian, A. and T. Dakessian (1998). "Multiple design points in first and second-order reliability." Structural Safety **20**(1): 37-49.
- Der Kiureghian, A. and O. Ditlevsen (2009). "Aleatory or epistemic? Does it matter?" Structural Safety **31**(2): 105-112.
- Der Kiureghian, A. and J.-B. Ke (1988). "The stochastic finite element method in structural reliability." Probabilistic engineering mechanics **3**(2): 83-91.
- Der Kiureghian, A., H.-Z. Lin and S.-J. Hwang (1987). "Second-order reliability approximations." Journal of Engineering Mechanics **113**(8): 1208-1225.
- Der Kiureghian, A., Y. Zhang and C.-C. Li (1994). "Inverse reliability problem." Journal of engineering mechanics **120**(5): 1154-1159.
- Du, X. and W. Chen (2004). "Sequential optimization and reliability assessment method for efficient probabilistic design." Journal of Mechanical Design **126**(2): 225-233.
- Du, X., J. Guo and H. Beeram (2008). "Sequential optimization and reliability assessment for multidisciplinary systems design." Structural and Multidisciplinary Optimization **35**(2): 117-130.
- Dubourg, V., B. Sudret and J.-M. Bourinet (2011). "Reliability-based design optimization using kriging surrogates and subset simulation." Structural and Multidisciplinary Optimization **44**(5): 673-690.
- Duncan, J. M. (1996). "State of the art: limit equilibrium and finite-element analysis of slopes." Journal of Geotechnical engineering **122**(7): 577-596.
- Enevoldsen, I. and J. D. Sørensen (1994). "Reliability-based optimization in structural engineering." Structural Safety **15**(3): 169-196.
- Ferson, S., V. Kreinovich, J. Hajagos, W. Oberkampf and L. Ginzburg (2007). "Experimental uncertainty estimation and statistics for data having interval uncertainty." Sandia National Laboratories, Report SAND2007-0939.
- Foye, K., R. Salgado and B. Scott (2006). "Assessment of variable uncertainties for reliability-based design of foundations." Journal of geotechnical and geoenvironmental engineering **132**(9): 1197-1207.
- Gasser, M. and G. I. Schuëller (1997). "Reliability-based optimization of structural systems." Mathematical Methods of Operations Research **46**(3): 287-307.
- Gelfand, A. E. and A. F. Smith (1990). "Sampling-based approaches to calculating marginal densities." Journal of the American statistical association **85**(410): 398-409.
- Geman, S. and D. Geman (1987). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. Readings in Computer Vision, Elsevier: 564-584.
- Giné, E. and R. Nickl (2010). "Confidence bands in density estimation." The Annals of Statistics **38**(2): 1122-1170.

- Goulet, J.-A. and I. F. Smith (2013). "Structural identification with systematic errors and unknown uncertainty dependencies." Computers & structures **128**: 251-258.
- Hasofer, A. M. and N. C. Lind (1974). "Exact and Invariant Second-moment Code Format." Journal of the Engineering Mechanics Division, ASCE **100**: 111-121.
- Hastings, W. K. (1970). "Monte Carlo sampling methods using Markov chains and their applications." Biometrika **57**(1): 97-109.
- Helton, J. C., J. D. Johnson and W. L. Oberkampf (2004). "An exploration of alternative approaches to the representation of uncertainty in model predictions." Reliability Engineering and System Safety **85**(1-3): 39-71.
- Helton, J. C., J. D. Johnson, W. L. Oberkampf and C. B. Storlie (2007). "A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory." Computer Methods in Applied Mechanics and Engineering **196**(37-40 SPEC. ISS.): 3980-3998.
- Hsieh, P.-G. and C.-Y. Ou (1998). "Shape of ground surface settlement profiles caused by excavation." Canadian geotechnical journal **35**(6): 1004-1017.
- Huffman, J. C., A. W. Strahler and A. W. Stuedlein (2015). "Reliability-based serviceability limit state design for immediate settlement of spread footings on clay." Soils and Foundations **55**(4): 798-812.
- Jensen, H. A. (2005). "Structural optimization of linear dynamical systems under stochastic excitation: a moving reliability database approach." Computer methods in applied mechanics and engineering **194**(12): 1757-1778.
- Kanda, J. and B. Ellingwood (1991). "Formulation of load factors based on optimum reliability." Structural Safety **9**(3): 197-210.
- Katafygiotis, L., T. Moan and S. H. Cheung (2007). "Auxiliary domain method for solving multi-objective dynamic reliability problems for nonlinear structures." Structural Engineering and Mechanics **25**(3): 347-363.
- Katafygiotis, L. S. and J. L. Beck (1998). "Updating models and their uncertainties. II: Model identifiability." Journal of Engineering Mechanics **124**(4): 463-467.
- Katafygiotis, L. S. and S. Cheung (2007). "Application of spherical subset simulation method and auxiliary domain method on a benchmark reliability study." Structural Safety **29**(3): 194-207.
- Kleinbaum, D., L. Kupper, A. Nizam and E. Rosenberg (2013). Applied regression analysis and other multivariable methods, Cengage Learning.
- Komlós, J., P. Major and G. Tusnády (1975). "An approximation of partial sums of independent RV's, and the sample DF. I." Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **32**(1-2): 111-131.
- Kuschel, N. and R. Rackwitz (1997). "Two basic problems in reliability-based structural optimization." Mathematical Methods of Operations Research **46**(3): 309-333.
- Lebrun, R. and A. Dutfoy (2009). "A generalization of the Nataf transformation to distributions with elliptical copula." Probabilistic Engineering Mechanics **24**(2): 172-178.
- Lee, J.-O., Y.-S. Yang and W.-S. Ruy (2002). "A comparative study on reliability-index and target-performance-based probabilistic structural design optimization." Computers & structures **80**(3-4): 257-269.

- Lee, J. H. and B. M. Kwak (1995). "Reliability-based structural optimal design using the Neumann expansion technique." Computers & structures **55**(2): 287-296.
- Leung, E. H. Y. and C. W. W. Ng (2007). "Wall and ground movements associated with deep excavations supported by cast in situ wall in mixed ground conditions." Journal of Geotechnical and Geoenvironmental Engineering **133**(2): 129-143.
- Li, H.-S., Y.-Z. Ma and Z. Cao (2015). "A generalized Subset Simulation approach for estimating small failure probabilities of multiple stochastic responses." Computers & Structures **153**: 239-251.
- Liang, J., Z. P. Mourelatos and J. Tu (2008). "A single-loop method for reliability-based design optimisation." International Journal of Product Development **5**(1-2): 76-92.
- Lin, P. T., H. C. Gea and Y. Jaluria (2011). "A modified reliability index approach for reliability-based design optimization." Journal of Mechanical Design **133**(4): 044501.
- Long, M. (2001). "Database for retaining wall and ground movements due to deep excavations." Journal of Geotechnical and Geoenvironmental Engineering **127**(3): 203-224.
- Low, B. (2005). "Reliability-based design applied to retaining walls." Geotechnique **55**(1): 63-75.
- Lu, L., H. Jiang and W. H. Wong (2013). "Multivariate density estimation by bayesian sequential partitioning." Journal of the American Statistical Association **108**(504): 1402-1410.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller (1953). "Equation of state calculations by fast computing machines." The journal of chemical physics **21**(6): 1087-1092.
- Moormann, C. (2004). "Analysis of wall and ground movements due to deep excavations in soft soil based on a new worldwide database." Soils and Foundations **44**(1): 87-98.
- Morton, K. W. and D. F. Mayers (2005). Numerical Solution of Partial Differential Equations: An Introduction, Cambridge University Press.
- Mottershead, J. E., M. Link and M. I. Friswell (2011). "The sensitivity method in finite element model updating: a tutorial." Mechanical systems and signal processing **25**(7): 2275-2296.
- Nataf, A. (1962). "Statistique mathematique-determination des distributions de probabilites dont les marges sont donnees." COMPTE RENDUS HEBDOMADAIRES DES SEANCES DE L ACADEMIE DES SCIENCES **255**(1): 42-&.
- Oberkampf, W. L., J. C. Helton and K. Sentz (2001). "Mathematical representation of uncertainty." Non-deterministic Approaches Forum.
- Ou, C.-Y. (2006). Deep excavation: theory and practice, CRC Press.
- Parzen, E. (1962). "On estimation of a probability density function and mode." The annals of mathematical statistics **33**(3): 1065-1076.
- Peck, R. B. (1969). Deep excavations and tunnelling in soft ground. Proc. 7th int. conf. on SMFE.
- Phoon, K.-K., F. H. Kulhawy and M. D. Grigoriu (2003). "Development of a reliability-based design framework for transmission line structure foundations." Journal of Geotechnical and Geoenvironmental Engineering **129**(9): 798-806.
- Rao, S. S. and L. Berke (1997). "Analysis of uncertain structural systems using interval analysis." AIAA journal **35**(4): 727-735.

- Rasmussen, C. E. (2006). "Gaussian processes for machine learning."
- Robert, C. P. (2004). Monte carlo methods, Wiley Online Library.
- Rosenblatt, M. (1952). "Remarks on a multivariate transformation." The annals of mathematical statistics **23**(3): 470-472.
- Royset, J. O., A. Der Kiureghian and E. Polak (2001). "Reliability-based optimal structural design by the decoupling approach." Reliability Engineering & System Safety **73**(3): 213-221.
- Rubinstein, R. Y. and D. P. Kroese (2016). Simulation and the Monte Carlo method, John Wiley & Sons.
- Schuëller, G. I. and H. A. Jensen (2008). "Computational methods in optimization considering uncertainties - An overview." Computer Methods in Applied Mechanics and Engineering **198**(1): 2-13.
- Scott, D. W. (2015). Multivariate density estimation: theory, practice, and visualization, John Wiley & Sons.
- Shafer, G. (1976). A mathematical theory of evidence, Princeton university press.
- Silverman, B. W. (1986). Density estimation for statistics and data analysis, CRC press.
- Simoen, E., G. De Roeck and G. Lombaert (2015). "Dealing with uncertainty in model updating for damage assessment: A review." Mechanical Systems and Signal Processing **56**: 123-149.
- Sloan, S. (2013). "Geotechnical stability analysis." Géotechnique **63**(7): 531-571.
- Snyman, J. (2005). Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms, Springer Science & Business Media.
- Soize, C. (2010). "Generalized probabilistic approach of uncertainties in computational dynamics using random matrices and polynomial chaos decompositions." International Journal for Numerical Methods in Engineering **81**(8): 939-970.
- Stefanou, G. (2009). "The stochastic finite element method: past, present and future." Computer Methods in Applied Mechanics and Engineering **198**(9): 1031-1051.
- Straub, D., I. Papaioannou and W. Betz (2016). "Bayesian analysis of rare events." Journal of Computational Physics **314**: 538-556.
- Sudret, B. and A. Der Kiureghian (2000). Stochastic finite element methods and reliability: a state-of-the-art report, Department of Civil and Environmental Engineering, University of California Berkeley, CA.
- Sues, R., M. Cesare, S. Pageau and J. Wu (2000). Reliability based MDO for aerospace systems. 8th Symposium on Multidisciplinary Analysis and Optimization, AIAA.
- Tanner, M. A. and W. H. Wong (1987). "The calculation of posterior distributions by data augmentation." Journal of the American statistical Association **82**(398): 528-540.
- Terzaghi, K. (1944). Theoretical Soil Mechanics, Chapman and Hall.
- Tu, J., K. K. Choi and Y. H. Park (2001). "Design potential method for robust system parameter design." AIAA journal **39**(4): 667-677.
- Valdebenito, M. A. and G. I. Schuëller (2010). "A survey on approaches for reliability-based optimization." Structural and Multidisciplinary Optimization **42**(5): 645-663.
- Vanmarcke, E. H. (1973). "Matrix formulation of reliability analysis and reliability-based design." Computers & Structures **3**(4): 757-770.

- Vanmarcke, E. H. (1977). "Probabilistic modeling of soil profiles." Journal of the geotechnical engineering division **103**(11): 1227-1246.
- Walker, W. E., P. Harremoës, J. Rotmans, J. P. van der Sluijs, M. B. van Asselt, P. Janssen and M. P. Kreyer von Krauss (2003). "Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support." Integrated assessment **4**(1): 5-17.
- Wang, L., J. H. Hwang, C. H. Juang and S. Atamturktur (2013). "Reliability-based design of rock slopes—a new perspective on design robustness." Engineering Geology **154**: 56-63.
- Wang, Y. (2009). "Reliability-Based Economic Design Optimization of Spread Foundations." Journal of Geotechnical and Geoenvironmental Engineering **135**(7): 954-959.
- Wang, Y. (2011). "Reliability-based design of spread foundations by monte carlo simulations." Geotechnique **61**(8): 677-685.
- Wang, Y. (2013). "MCS-based probabilistic design of embedded sheet pile walls." Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards **7**(3): 151-162.
- Wang, Y., S.-K. Au and F. H. Kulhawy (2011). "Expanded reliability-based design approach for drilled shafts." Journal of geotechnical and geoenvironmental engineering **137**(2): 140-149.
- Wang, Y. and Z. Cao (2013). "Expanded reliability-based design of piles in spatially variable soil using efficient Monte Carlo simulations." Soils and Foundations **53**(6): 820-834.
- Wei, W., Y. Cheng and L. Li (2009). "Three-dimensional slope failure analysis by the strength reduction and limit equilibrium methods." Computers and geotechnics **36**(1): 70-80.
- Wu, S.-H., C.-Y. Ou, J. Ching and C. Hsein Juang (2011). "Reliability-based design for basal heave stability of deep excavations in spatially varying soils." Journal of Geotechnical and Geoenvironmental Engineering **138**(5): 594-603.
- Yao, W., X. Chen, W. Luo, M. van Tooren and J. Guo (2011). "Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles." Progress in Aerospace Sciences **47**(6): 450-479.
- Yoo, C. (2001). "Behavior of braced and anchored walls in soils overlying rock." Journal of Geotechnical and Geoenvironmental Engineering **127**(3): 225-233.
- Youn, B. D. and K. K. Choi (2004). "A new response surface methodology for reliability-based design optimization." Computers & structures **82**(2-3): 241-256.
- Youn, B. D., K. K. Choi and Y. H. Park (2003). "Hybrid analysis method for reliability-based design optimization." Journal of Mechanical Design **125**(2): 221-232.
- Yuan, X. (2013). "Local estimation of failure probability function by weighted approach." Probabilistic Engineering Mechanics **34**: 1-11.
- Zadeh, L. A. (1999). "Fuzzy sets as a basis for a theory of possibility." Fuzzy sets and systems **100**: 9-34.
- Zhang, J., L. M. Zhang and W. H. Tang (2012). "Reliability-based optimization of geotechnical systems." Journal of Geotechnical and Geoenvironmental Engineering **137**(12): 1211-1221.

Bibliography

Zio, E. (2013). The Monte Carlo simulation method for system reliability and risk analysis, Springer.

Zuev, K. M., J. L. Beck, S.-K. Au and L. S. Katafygiotis (2012). "Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions." Computers & structures **92**: 283-296.

Academic resume

EDUCATION

Doctor of Philosophy (PhD) AUG/2014 – AUG/2018

Nanyang Technological University

Singapore

Master of engineering

AUG/2011 - AUG/2014

Tongji University

Shanghai, China

Bachelor of engineering

AUG/2007 - AUG/2011

Tongji University

Shanghai, China

PUBLICATIONS

[1] Liu Wang-Sheng, Cheung Sai Hung. Instant reliability updating by constructing the updated failure probability function. *Computer methods in applied mechanics and engineering* (under review)

[2] Liu Wang-Sheng, Cheung Sai Hung. Reliability based design optimization with approximate failure probability function in partitioned design space. *Reliability Engineering and System Safety* 167 (2017): 602-611

[3] Liu Wang-Sheng, Cheung Sai Hung. An efficient reliability-based optimization framework for geotechnical design considering spatial variability of soil properties. *Applied Mathematical Modelling* (under review)

-
- [4] Liu Wang-Sheng, Cheung Sai Hung. Design optimization of stochastic complex systems via iterative density estimation. *12th International Conference on Structural Safety and Reliability, ICOSSAR*, Aug 2017, Vienna, Austria
- [5] **(Patent)** Liu Wang-Sheng, Zhao Ming. Tongji University. A wireless deformation monitoring and early-warning system for shield tunnels. China; CN103968800 A, 2016.
- [6] Liu Wang-Sheng, Zhao Ming. A Reliability-based Early Warning System for Real-time Monitoring on RC Buildings. *Advanced Materials Research*, 446-449, 2012: P3422-3427