# Distribution-Free Stochastic Simulation Methodology

# for Model Updating Under Hybrid Uncertainties

Von der Fakultät für Bauingenieurwesen und Geodäsie

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Ingenieurwissenschaften

**Dr.-Ing.**

genehmigte Dissertation

von

**Masaru Kitahara, M.Eng.**

**2022**

**Referent:** Prof. Dr.-Ing. Michael Beer

**Korreferent:** Prof. Dr.-Ing. Michael Haist

Prof. Dr.-Ing. Udo Nackenhorst

Prof. Dr.Eng. Chul-Woo Kim (Kyoto University)

Tag der Promotion: 03.06.2022

# Erklärung

Ich erkläre hiermit, dass die in dieser Dissertaion vorgestellten Ergebnisse auf meiner eigenen Arbeit beruhen und dass ich keine Arbeiten andere Personen vorgelegt habe und dass ich in allen Fällen, in denen ich auf Arbeiten anderer Personen Bezug genommen habe, diese in vollem Umfang und in angemessener Weise angegeben habe.

Masaru Kitahara

Hannover, März 2022

# Acknowledgments

This dissertaion is the summary of my work in the Institute for Risk and Reliability at Leibniz Universität Hannover. Over the last few years, a lot of people have helped me to grow, both academically and professionally. Without their support, this work would not have been possible.

My first greatest gratitude goes to my supervisor, Prof. Dr.-Ing. Michael Beer. He offered me the opportunity to study and work in his institute, and has continuously supported me to obtain the funds for my PhD study. His expertise on the academic area of uncertainty quantification allowed me to quickly learn the latest topics in this field such as the imprecise probabilities and gave me ideal guidelines for my research. The study in the Institute for Risk and Reliability has largely broadened my research view and helped me to put my ideas in shape. Michael always left me freedom to develop my own ideas but, at the same time, gave me advice at the right points in time during my research. This balance between freedom and guidance I found in his supervision was essential to my PhD study. He has also provided me a lot of opportunities to get involved in international academic networks, such as the Summer School on Smart Structures Technology at Sapienza University of Rome and 2nd Dynamics Workshop at University of Sheffield, as well as many international conferences especially ESREL 2019 that was organized by the institute. These opportunities have largely broadened my academic horizon. I am really grateful to him for all these achievements.

I would also like to show my deep appreciation to Dr. Matteo Broggi. He put my research focus on the Bayesian model updating and showed me his ideas about the application of staircase random variables. He always invented time when I had problems or question about my research or writing papers. Through our numberless and fruitful discussions, he always guided my research in the right direction. He also gave me the opportunity to participate in the NASA UQ challenge 2019. The findings through this challenge constituted the core of this thesis and were significant to make my work more visible. In this context, I would also like to thank Prof. Edoardo Patelli and Mr. Adolphus Lye who worked together on this challenge. The time they had in our intense and relevant discussions are highly appreciated. Over and above these scientific aspects, Matteo's personality made my doctoral experience rather unique and entertaining. I am very grateful to him.

Then, I would like to take this opportunity to thank Dr. Sifeng Bi, Prof. Pengfei Wei, and Dr. Jingwen Song for our collaboration and fruitful discussions. Sifeng gave me tremendous specific advice on the stochastic model updating.

I have learnt from him not only the latest findings in this field, but also all the aspects of research such as the way of thinking scientific problems, producing journal papers, and playing a role as reviewer. Pengfei provided me the Matlab codes of the AK-MCMC algorithm, which was the fundamental for my first development. Jingwen provided me with great assistance in using the generalized NISS method to address the NASA challenge. She also gave me many specific advices on my fifth development, together with Pengfei, as the co-authors.

Further, I also want to say thank you to all of the dear colleagues in the Institute for Risk and Reliability, including but not limited to Marco, Jasper, Julian, Marius, George, Peihua, Danko, and Thomas. Over the years, the Institute has been a pole of attraction for a lot of excellent researchers from many different parts of the world, and has provided the ideal environment for the growth of ideas and collaborations. I had the great pleasure to work with all of them. Also, I am very grateful to Natalia for her supports not only for the work in the Institute but also for live in Germany.

I am highly honored that Prof. Dr-Ing. Michael Haist, Prof. Dr-Ing. Udo Nackenhorst, and Prof. Chul-Woo Kim reviewed my thesis and acted as referees. In particular, Prof. Chul-Woo Kim introduced myself to Michael when I was looking for a place to study and gave him a personal recommendation which led to this fruitful doctoral experience. I am really grateful to him and pleasure to have the opportunity to show him my achievements. I also really acknowledge the funding from DAAD (Deutscher Akademischer Austauschdienst) and DFG (Deutsche Forschungsgemensschaft) supported me to study in Germany.

My special thanks go to my family. My parents and sister have been always on my side sustaining my decisions. It was a great memory that I can spend one year in Germany with my father during his research stay at Hannover. Our countless scientific discussions sparked many research ideas, which are highly appreciated. Last but not least, I would like to express my deepest gratitude to my dear wife Sachiko for putting up with my late hours working on the thesis.

Masaru Kitahara

Hannover, March 2022

# Abstract

In the real world, a significant challenge faced in the safe operation and maintenance of infrastructures is the lack of available information or data. This results in a large degree of uncertainty and the requirement for robust and efficient uncertainty quantification (UQ) tools in order to derive the most realistic estimates of the behavior of structures. While the probabilistic approach has long been utilized as an essential tool for the quantitative mathematical representation of uncertainty, a common criticism is that the approach often involves insubstantiated subjective assumptions because of the scarcity or imprecision of available information. To avoid the inclusion of subjectivity, the concepts of imprecise probabilities have been developed, and the distributional probability-box (p-box) has gained the most attention among various types of imprecise probability models since it can straightforwardly provide a clear separation between aleatory and epistemic uncertainty.

This thesis concerns the realistic consideration and numerically efficient calibraiton and propagation of aleatory and epistemic uncertainties (hybrid uncertainties) based on the distributional p-box. The recent developments including the Bhattacharyya distance-based approximate Bayesian computation (ABC) and non-intrusive imprecise stochastic simulation (NISS) methods have strengthened the subjective assumption-free approach for uncertainty calibration and propagation. However, these methods based on the distributional p-box stand on the availability of the prior knowledge determining a specific distribution family for the p-box. The target of this thesis is hence to develop a distribution-free approach for the calibraiton and propagation of hybrid uncertainties, strengthening the subjective assumption-free UQ approach.

To achieve the above target, this thesis presents five main developments to improve the Bhattacharyya distance-based ABC and NISS frameworks. The first development is on improving the scope of application and efficiency of the Bhattacharyya distance-based ABC. The dimension reduction procedure is proposed to evaluate the Bhattacharyya distance when the system under investigation is described by time-domain sequences. Moreover, the efficient Bayesian inference method within the Bayesian updating with structural reliability methods (BUS) framework is developed by combining BUS with the adaptive Kriging-based reliability method, namely AK-MCMC. The second development of the distribution-free stochastic model updating framework is based on the combined application of the staircase density functions and the Bhattacharyya distance. The staircase density functions can approximate a wide range of distributions

arbitrarily close; hence the development achieved to perform the Bhattacharyya distance-based ABC without limiting hypotheses on the distribution families of the parameters having to be updated. The aforementioned two developments are then integrated in the third development to provide a solution to the latest edition (2019) of the NASA UQ challenge problem. The model updating tasks under very challenging condition, where prior information of aleatory parameters are extremely limited other than a common boundary, are successfully addressed based on the above distribution-free stochastic model updating framework. Moreover, the NISS approach that simplifies the high-dimensional optimization to a set of one-dimensional searching by a first-order high-dimensional model representation (HDMR) decomposition with respect to each design parameter is developed to efficiently solve the reliability-based design optimization tasks. This challenge, at the same time, elucidates the limitations of the current developments, hence the fourth development aims at addressing the limitation that the staircase density functions are designed for univariate random variables and cannot acount for the parameter dependencies. In order to calibrate the joint distribution of correlated parameters, the distribution-free stochastic model updating framework is extended by characterizing the aleatory parameters using the Gaussian copula functions having marginal distributions as the staircase density functions. This further strengthens the assumption-free approach for uncertainty calibration in which no prior information of the parameter dependencies is required. Finally, the fifth development of the distribution-free uncertainty propagation framework is based on another application of the staircase density functions to the NISS class of methods, and it is applied for efficiently solving the reliability analysis subproblem of the NASA UQ challenge 2019.

The above five developments have successfully strengthened the assumption-free approach for both uncertainty calibration and propagation thanks to the nature of the staircase density functions approximating arbitrary distributions. The efficiency and effectiveness of those developments are sufficiently demonstrated upon the real-world applications including the NASA UQ challenge 2019.

**Keywords:** Uncertainty quantification; Imprecise probabilities; Stochastic model updating; Bhattacharyya distance; Staircase density function.

# Zusammenfassung

In der realen Welt stellt der Mangel an verfügbaren Informationen oder Daten eine wesentliche Herausforderung für den sicheren Betrieb und die Instandhaltung von Infrastruktur dar. Daraus resultiert ein hohes Ausmaß an Unsicherheit und der Bedarf an robusten und effizienten Instrumenten zur Unsicherheitsquantifizierung (UQ), um möglichst realistische Schätzungen des Strukturverhaltens abzuleiten. Der probabilistische Ansatz dient seit langem als wesentliches Instrument zur quantitativen mathematischen Darstellung der Unsicherheit. Hierbei wird jedoch häufig kritisiert, dass dieser Ansatz aufgrund der Knappheit oder Ungenauigkeit der verfügbaren Informationen oft ungerechtfertigte subjektive Annahmen beinhaltet. Zur Vermeidung der Einbeziehung von Subjektivität wurden Konzepte für unscharfe Wahrscheinlichkeiten entwickelt, wobei die verteilungsbezogene Wahrscheinlichkeitsbox (P-Box) unter den verschiedenen Arten von ungenaue Wahrscheinlichkeitsmodellen die meiste Aufmerksamkeit erlangt hat, da sie eine klare Trennung zwischen aleatorischer und epistemischer Unsicherheit bieten kann.

Diese Arbeit befasst sich mit der realistischen Betrachtung sowie numerisch effizienten Kalibrierung und Fortpflanzung von aleatorischen und epistemischen Unsicherheiten (hybride Unsicherheiten) auf Basis der verteilungsbezogenen P-Box. Die jüngsten Entwicklungen, einschließlich der auf der Bhattacharyya-Distanz basierenden „approximate Bayesian computation (ABC)" und „non-intrusive imprecise stochastic simulation (NISS)", haben den subjektiven, annahmefreien Ansatz zur Kalibrierung und Fortpflanzung von Unsicherheit verbessert. Die Methoden setzen jedoch die Verfügbarkeit des Wissens über die Verteilungsfamilie der P-Box voraus. Das Ziel dieser Arbeit ist es daher, einen verteilungsfreien Ansatz für die Kalibrierung und Fortpflanzung hybrider Unsicherheiten zu entwickeln, der den subjektiven, annahmefreien UQ-Ansatz kräftigt.

Um das oben genannte Ziel zu erreichen, werden in dieser Arbeit fünf Hauptentwicklungen zur Verbesserung der Bhattacharyya-Distanz-basierten ABC- und NISS-Framework vorgestellt. Die erste Entwicklung betrifft die Erweiterung des Anwendungsbereichs und Verbesserung der Effizienz des Bhattacharyya-Distanz-basierten ABC. Hierbei werden Verfahren der Dimensionsreduktion zur Bewertung der Bhattacharyya-Distanz eingeführt, die System untersuchen, welche durch Sequenzen im Zeitbereich beschrieben werden. Darüber hinaus wird ein effizienter Bayes'scher Inferenzalgorithmus entwickelt, der auf „Bayesian updating with structural reliability methods (BUS)" basiert. Die BUS wird dazu mit „adaptive Kriging-Markov chain Monte Carlo (AK-MCMC)" kombiniert. Die

zweite Entwicklung des verteilungsfreien stochastischen Modellaktualisierungsframeworks basiert auf der verknüpften Anwendung der Treppendichtefunktionen und des Bhattacharyya-Distanz. Mit Treppenfunktionen lässt sich ein breites Spektrum von Verteilungsfunktionen beliebig nah approximieren; Wodurch die ABC, basierend auf der Bhattacharyya-Distanz, durchgeführt werden kann, ohne einschränkende Hypothesen über die Verteilungsfamilie der zu aktualisierenden Parameter treffen zu müssen. Beide Entwicklungen werden dann in die dritte Entwicklung integriert, um eine Lösung für die aktuellste Auflage (2019) des NASA UQ-Problems zu liefern. Diese Aufgabe der Modellaktualisierung beinhaltet anspruchsvolle Bedingungen, in denen die Vorkenntnisse der aleatorischen Parameter, bis auf eine gemeinsame Randbedingung, extrem limitiert sind. Ferner konnte dies mit dem verteilungsfreien stochastischen Framework zur Modellaktualisierung erfolgreich angegangen werden. Zusätzlich wird durch den Ansatz der NISS die hochdimensionale Optimierung mithilfe der inkludierten „high-dimensional model representation (HDMR)" in jeweils eindimensionale Probleme zerlegt. Diese Herausforderung verdeutlicht gleichzeitig die Grenzen der diskutierten Entwicklungen. Aus diesem Grund zielt die vierte Weiterentwicklung darauf ab, die Einschränkung zu beseitigen, dass die Treppendichtefunktionen für univariate Zufallsvariablen konzipiert sind und die Parameterabhängigkeiten nicht berücksichtigen können. Um die Verbundverteilung der korrelierten Parameter zu justieren, wird das verteilungsfreie stochastische Framework zur Modellaktualisierung erweitert, indem die aleatorischen Parameter als Gaußsche Copulas mit Randverteilungen als Treppendichtefunktionen charakterisiert werden. Dies stärkt weiter den annahmefreien Ansatz zur Kalibrierung von Unsicherheiten, bei dem kein Vorwissen über das Vorhandensein der Parameterabhängigkeiten erforderlich sind. Schließlich basiert die fünfte Entwicklung des Frameworks zur Fortpflanzung verteilungsfreien Unsicherheiten auf eine weitere Anwendung der Treppendichtefunktionen auf die NISS-Methodenklasse. Ferner wird dies auf das Teilproblem der Zuverlässigkeitsanalyse der NASA UQ Challenge 2019 angewandt, um dies effizient zu lösen.

Die oben genannten fünf Entwicklungen haben den annahmefreien Ansatz sowohl für die Kalibrierung als auch für die Fortpflanzung der Unsicherheit erfolgreich verstärkt, dank der Eigenschaft das Treppenfunktionen beliebige Verteilungen approximieren kann. Die Effizienz und Effektivität dieser Entwicklungen werden durch Anwendungen in der Praxis, einschließlich der NASA UQ Challenge 2019, ausreichend demonstriert.

**Schlüsselwörter:** Unsicherheitsquantifizierung; unscharfe Wahrscheinlichkeiten; stochastische Modellaktualisierung; Bhattacharyya-Distanz; Treppendichtefunktion

# List of Publications

1.  Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2021). Bayesian Model Updating in Time Domain with Metamodel-Based Reliability Method. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7(3), 04021030.

2.  Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2022). Nonparametric Bayesian stochastic model updating with hybrid uncertainties. *Mechanical Systems and Signal Processing*, 163, 108195.

3.  Lye, A., Kitahara, M., Broggi, M., and Patelli, E. (2022). Robust optimization of a dynamic Black-box system under severe uncertainty: A distribution-free framework. *Mechanical Systems and Signal Processing*, 167, 108522.

4.  Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2022). Distribution-free stochastic model updating of dynamic systems with parameter dependencies. *Structural Safety*. 97, 102227.

5.  Kitahara, M., Song, J., Wei, P., Broggi, M., and Beer, M. (2022). A distributionally robust approach for mixed aleatory and epistemic uncertainties propagation. *AIAA Journal*. (In press).

# Contents

# Chapter 1

# Introduction

## 1.1   Research background

Over the past few decades, civil infrastructures such as highway bridges have been aging in many countries, and more and more infrastructures are being classified as structurally deficient. For instance, in Japan as seismically-prone region, the use of seismic isolators has been promoted after the 1995 Kobe Earthquake for mitigating the risk of seismic damages on highway bridges, however aging deterioration is being confirmed in many rubber bearings after the service years exceeding ten years. Isolated bridges are typically designed to absorb earthquake energy by the rubber bearings and reduce the seismic force on superstructures; hence, deterioration of the rubber bearings can significantly affect the seismic capacity of the entire bridge systems. In fact, the rubber bearings ruptured during the 2011 off the Pacific coast of Tohoku Earthquake and 2016 Kumamoto Earthquake, causing the loss of serviceability of several highway bridges. More recently, the Morandi Bridge in Genova, Italy collapsed in 2018, which caused 43 deaths. Visual inspections of the bridge in 2015 confirmed deterioration of the stays in the pier, where the collapse occurred. Retrofitting for ensuring the necessary support was planned in 2017; however, the collapse transpired prior to the procedure. As such, structural deficiencies in key infrastructures can result in the massive consequences of possible failures, and the development of a framework for safe operation and maintenance of the infrastructures in order to ensure the infrastructure resilience is one of our centennial challenges.

*Structural health monitoring* (SHM) has been establishing itself as a key instrument for condition assessment and service life monitoring of existing structures mainly due to the recent advancements in sensing technology. The use of monitoring data is becoming increasingly common for assessing the safety and serviceability of structural designs and deriving decisions regarding the maintenance and repair. Moreover, the exponential growth in computational capacities facilitates the employment of computationally more and more complex models, enabling simulations of real structures in an ever-increasing level of detail. In structural engineering community, for instance, finite element (FE) models with hundreds to even millions of degrees of freedom are increasingly being utilized for analyzing, e.g., internal forces and displacements of the structure as a whole in a very detailed manner. Hence, model-based SHM approaches have drawn

particular attentions for coupling real structures and numerical models in order to derive the most realistic estimates of the behavior of structures, in which *model updating* techniques (e.g., the vibration-based FE model updating), that are aimed at calibrating model parameters based on measurements such that the best possible fit is obtained between model predictions and measurements, plays a central role.

However, no matter how sophisticated the models become, model predictions can show an incomplete agreement with the actual behavior of the structures. This discrepancy is typically caused by uncertainties in the model parameters and the models themselves. The former type of uncertainty is referred to as *parameter uncertainty*, and it indicates the inherently variable parameters or the lack of knowledge about the true values of the parameters. Typical examples of the inherent variability include the material properties, geometry dimensions, and boundary conditions of the structures varying due to the manufacturing processes, damaged conditions, or surroundings, whereas an example of the lack of knowledge is that the exact loading conditions of the structures during extreme events such as earthquakes are unknown a priori. On the other hand, the latter type of uncertainty is referred to as *modeling uncertainty*, and it is caused by the simplifications and approximations that have to be made for numerically representing real structures. Typical examples of such are the linearized representation of nonlinear structural behaviors, frictionless assumption of mechanical joints, and discretization over the time and space. In addition to the aforementioned uncertainties in simulations, uncertainties are also inevitable in measurements, e.g., the noise from environmental influences and the measurement system errors, and such type of uncertainties is referred to as *measurement uncertainty*.

Consequently, the main interest of current researches in the fields of SHM and model updating concerns methods for *uncertainty quantification* (UQ) of both simulations and measurements. UQ includes various tasks, e.g., *uncertainty characterization*, *uncertainty calibration*, and *uncertainty propagation*, aimed at providing a quantitative description and reduction of uncertainties based on non-deterministic modelling approaches. The conventional non-deterministic modelling approaches rely on the probabilistic descriptions; however, the construction of probabilistic models involves a significant amount of subjective judgements or assumptions because quantitative data from SHM is often very scarce and collecting more data may not be possible. To avoid the inclusion of subjectivity, the imprecision and vagueness in the modelling can be considered by using the concepts of *imprecise probabilities*, which combine probabilistic and set-theoretical frameworks to provide the bounds on probabilities for the events of interest. This thesis mainly focuses on the state-of-the-art developments in the field of imprecise probabilities, and the key challenge addressed is the realistic consideration and efficient quantitative evaluation of uncertainties under significantly limited prior information on the underlying probabilistic models.

### 1.1.1  Uncertainty characterization

Uncertainties can be generally classified into two categories: *aleatory uncertainty* and *epistemic uncertainty* (Der Kiureghian and Ditlevsen, 2009). Aleatory uncertainty (also called irreducible uncertainty, objective uncertainty, and variability) denotes the intrinsic variation or randomness of a phenomenon. An example of aleatory uncertainty is the variable mass of any part that can be produced by a manufacturing process. This kind of uncertainty is irreducible since it is inherent of the phenomenon. In fact, the uncertainty in the mass of the parts because of the manufacturing process can only be reduced by modifying the manufacturing and quality control processes. Epistemic uncertainty (also called reducible uncertainty, subjective uncertainty, and lack of knowledge), on the other hand, denotes the lack of knowledge or data about a phenomenon. An example of epistemic uncertainty is the constant but unknown mass of a specific part that can be produced by a manufacturing process. As the mass of the part can measure directly, if more information is acquired, this type of uncertainty can be reduced. Furthermore, if sufficient information is given, epistemic uncertainty can, in principle, be eliminated.

Among different sources of uncertainty, modeling and measurement uncertainties should both be categorized as epistemic uncertainty. In fact, modeling uncertainty can be reduced by improving numerical approximations and using higher fidelity models. Measurement uncertainty can also be reduced by using more accurate methods of measurement. Comparatively, as mentioned earlier, parameter uncertainty is categorized into either aleatory or epistemic uncertainty depending on the circumstances. However, a clear distinction between aleatory and epistemic uncertainty is not always trivial in characterizing the model parameters. For example, the variation in the mass of any part which can be produced by a manufacturing process is impossible to be precisely determined if only a small number of samples can be collected from the population. In this case, the uncertainty in the mass of the parts is considered to be the combination of aleatory and epistemic uncertainty, and such combination is in general referred to in literature as *hybrid uncertainties* or *mixed uncertainties*. By obtaining more samples of the manufactured parts, the variation in the mass of the parts can be more accurately determined as a result of the reduction of epistemic uncertainty. If one can obtain a sufficiently large number of samples, then epistemic uncertainty can be eliminated and only aleatory uncertainty remains in the mass of the parts due to its intrinsic variation.

Uncertainty characterization is in particular aimed at mathematically describing parameter uncertainty by means of *uncertainty models* (UMs). Several categories of the model parameters can be defined in this context based on their origin as aleatory or/and epistemic uncertainty:

Category I:     Parameters without any uncertainty, represented as explicit constants;

Category II:    Parameters with only aleatory uncertainty, represented as random variables with a fully prescribed

stochastic description;

Category III:   Parameters with only epistemic uncertainty, represented as unknown-but-fixed constants;

Category IV:   Parameters with both aleatory and epistemic uncertainties, represented as random variables with an

unknown-but-fixed stochastic description.

It should be emphasized that, in uncertainty characterization, it is of crucial importance to properly distinguish between aleatory and epistemic uncertainty because these two uncertainties have apparently different effects on the subsequent uncertainty calibration and propagation tasks. Take reliability analysis as an example of uncertainty propagation, where estimating the failure probability of the structure is one's main concern, and the presence of aleatory uncertainty results in a random failure because of the intrinsic property of the structure. Epistemic uncertainty, on the contrary, does not affect the true value of the failure probability, however it prevents us from estimating it precisely. These properties also hold for the two kinds of uncertainties in the category IV parameters.

The UMs in literature can be generally categorized into three groups based on the probabilistic or/and possibilistic information delivered, as probability models, non-probabilistic models, and imprecise probability models.

### *Precise probability models*

Probability models (also called precise probability models by comparing imprecise probability models) are based on the well-established probability theory, and are typically suitable for characterizing the category II parameters. The probability models require the definition of a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $\Omega$ denotes the sample space containing all possible outcomes, $\mathcal{F}$ being the $\sigma$-algebra as the set of events containing zero or more outcomes, and $\mathcal{P}: \mathcal{F} \to [0, 1]$ the probability measure that assigns the probability to an event $\omega \in \mathcal{F}$. In this context, a random variable $X$ is defined by the mapping $X: (\Omega, \mathcal{F}) \to (\mathcal{D}_X, \mathcal{B}(\mathcal{D}_X))$, where $\mathcal{D}_X \subset \mathbb{R}$ represents the support domain of $X$ and $\mathcal{B}(\mathcal{D}_X)$ denotes the Borel $\sigma$-algebra on $\mathcal{D}_X$. The above definition holds that $X^{-1}(x) = \{\omega \in \Omega | X(\omega) \in x\} \in \mathcal{F}$ for the event $x \in \mathcal{B}(\mathcal{D}_X)$, and the probability of the event $x$ is expressed as:

$$\mathcal{P}_X(x) = \mathcal{P} \circ X^{-1}(x) = \mathcal{P}\{\omega: X(\omega) \in x\} \tag{1.1}$$

A typical representation form of the probability models that fully characterizes the random variable $X$ is the cumulative distribution function (CDF) $F_X$ assigning the probability to the event $\{X \leq x\}$, and is expressed as:

$$F_X(x) = \mathcal{P}_X(X \leq x) \tag{1.2}$$

Its derivative, denoted as $f_X(x)$, is known as the probability density function (PDF). The PDF means the likelihood of $X$ being in the neighborhood of $x$. In the definition, the CDF increases monotonically with $x$, and the PDF holds that $f_X(x) \geq 0$ for all $x \in \mathcal{D}_X$. Furthermore, the dependence structure among multiple variables can be represented using a joint CDF $F_{X_1, X_2, \cdots, X_{n_x}}(x_1, x_2, \cdots, x_{n_x}) = \mathcal{P}_X(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_{n_x} \leq x_{n_x})$. According to Sklar's theorem (Nelsen, 2006), the joint CDF with marginal distributions $F_{X_i}(x_i) = \mathcal{P}_X(X_i \leq x_i)$ can be uniquely written as:

$$F_{X_1, X_2, \cdots, X_{n_x}}(x_1, x_2, \cdots, x_{n_x}) = C\left(F_{X_1}(x_1), F_{X_2}(x_2), \cdots, F_{X_{n_x}}(x_{n_x})\right) \tag{1.3}$$

where $C$ denotes a copula, which contains all information on the dependence structure.

***Non-probabilistic models***

Non-probabilistic models (Faes and Moens, 2020) are set-theoretical models and typically used for characterizing the category III parameters. The typical non-probabilistic models include the interval/convex models, fuzzy set model, and the associated possibility theory (Helton et al., 2010). The non-probabilistic models assign bounds between which the parameters are considered to lie, without assigning the likelihood to each value within the bounds.

An interval model requires the definition of a crisp set $\mathcal{E}$ which, considering the sample space $\Omega$, the event $\omega$ will be included in the set if the binary membership function (also called the characteristic function) is $\xi(\omega) = 1$ and vice versa if $\xi(\omega) = 0$. In this context, an interval variable $Y$ is expressed by the bounds as:

$$Y = \left[\underline{y}, \overline{y}\right] = \left\{y \in \mathcal{D}_Y \,\middle|\, \underline{y} \leq y \leq \overline{y}\right\} \tag{1.4}$$

where $\mathcal{D}_Y \subset \mathbb{R}$ denotes the support domain of $Y$. The interval model generally does not consider dependency between multiple variables. The interval model with multiple variables is thus expressed by the Cartesian product $\boldsymbol{Y} = \times_{i=1}^{n_y} Y_i = \times_{i=1}^{n_y} \left[\underline{y_i}, \overline{y_i}\right]$, and the support domain is denoted as a rectangle or hyper-rectangle. A convex model, on the other hand, is defined by a set where, for any two points in the set, all points along the connecting line between the two points are also included in the set. The convex models are in many ways similar to the interval models, but also allow to consider dependency between multiple variables. In fact, an interval model of multiple variables can be interpreted as a specific case of the hyper-rectangular convex model. One of the most widely used convex models is the ellipsoid model (Jiang et al., 2013), and can be expressed by $\boldsymbol{Y} = \{\boldsymbol{y} \in \mathcal{D}_Y | (\boldsymbol{y} - \boldsymbol{c}_Y)'\boldsymbol{G}(\boldsymbol{y} - \boldsymbol{c}_Y) \leq 1\}$, where $\boldsymbol{G}$ denotes a symmetric positive-

definite characteristic matrix, which determines the size and orientation of the ellipsoid, and $\boldsymbol{c_Y}$ is the ellipsoidal centre. The characteristic matrix $\mathbf{G}$ demonstrates the degree of dependence between the variables. The marginal model of each variable $Y_i$ can be still expressed by an interval $Y_i = \left[\underline{y_i}, \overline{y}_i\right]$.

A fuzzy set model can be regarded as an extension of the intervals. Considering the sample space $\Omega$, a fuzzy set is defined by $\tilde{\mathcal{E}} = \{\omega, \xi(\omega) | \omega \in \Omega\}$, where $\xi: \Omega \to [0, 1]$ means a continuous membership function. Figure 1.1 shows the difference between the crisp and fuzzy sets. The triangular and trapezoidal fuzzy sets are herein depicted. In the above context, for a fuzzy variable $Y$, the support is defined by $\mathcal{Y} = \left[\underline{y}, \overline{y}\right] = \{y \in \mathcal{D}_Y | \xi_Y(y) > 0\}$. In addition, the $\alpha$-cut level interval can be similarly defined as:

$$Y_\alpha = \left[\underline{y}_\alpha, \overline{y}_\alpha\right] = \{y \in \mathcal{D}_Y | \xi_Y(y) \geq \alpha\} \tag{1.5}$$

where $\alpha \in [0, 1]$ indicates the $\alpha$-cut level. The $\alpha$-cut procedure can provide nested intervals as $Y_{\alpha_i} \subseteq Y_{\alpha_j}$ when $\alpha_i \geq \alpha_j$. Therefore, larger membership value indicates higher risk that the unknown true value $Y^*$ might be lie outside of the $\alpha$-cut interval. If the analyst does not want to take any risk, a large interval with the high confidence, i.e., $\mathcal{Y}$, should be used, whilst one can derive a narrower interval with less confidence by taking a certain degree of risk.
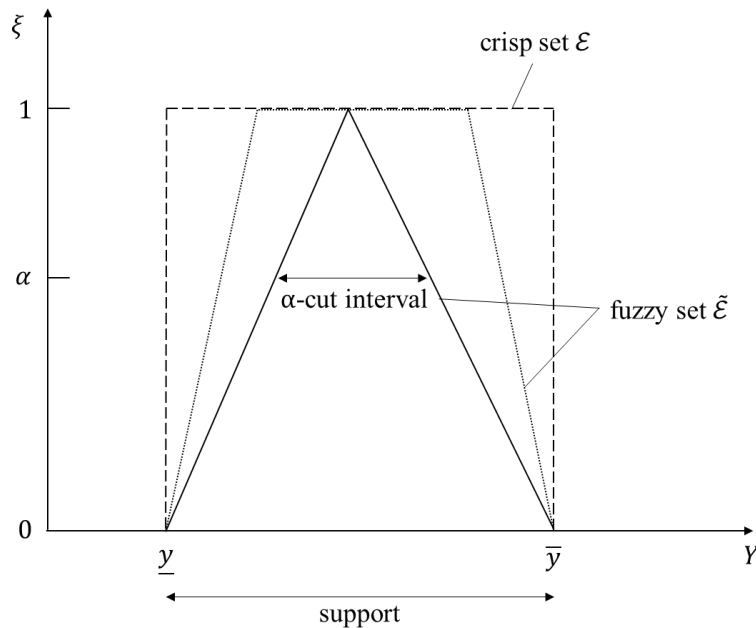


Figure 1.1 Illustration of crisp and fuzzy sets.

The possibility theory can be derived from the fuzzy set model. As the probability theory associates the probability distribution with the random variable, the possibility theory assigns the possibility distribution to the fuzzy variable $Y$.

Consider a possibility space $(\Omega, r)$, where $r(y) = \xi_Y(y)$ indicates the possibility distribution function and provides a confidence measure that is assigned to each event within the support $\mathcal{Y} = \left[\underline{y}, \overline{y}\right] \subset \mathbb{R}$. Intuitively, $r(y) = 1$ means that $y$ is entirely possible in the sense that sense that no contradictory information is available, and $r(y) = 0$ indicates that $y$ is not possible in light of the available information. The possibility theory provides two measures of likelihood, i.e., possibility and necessity, for each subset $Y \subseteq \mathcal{Y}$ as:

$$\text{Pos}(Y) = \sup_y \{\xi_Y(y)\} \tag{1.6}$$

$$\text{Nec}(Y) = 1 - \text{Pos}(Y^c) = \inf_y \{1 - \xi_Y(y)\} \tag{1.7}$$

It holds that $\text{Nec}(Y) \leq \text{Pos}(Y)$. In this context, the fuzzy variable $Y$ is characterized by both the cumulative possibility function (upper bound) and cumulative necessity function (lower bound) as:

$$\text{Pos}(Y \leq y) = \sup_y \min\left\{\xi_Y(y), \xi_y(y)\right\} \tag{1.8}$$

$$\text{Nec}(Y \leq y) = \inf_y \max\left\{1 - \xi_Y(y), \xi_y(y)\right\} \tag{1.9}$$

***Imprecise probability models***

Imprecise probability models (Beer et al., 2013) are the combination of probabilistic and set theoretical methods, and are capable aleatory and epistemic uncertainty to be handled separately within a unified framework, thanks to the hierarchical model structure. Hence, they are especially useful to characterize the category IV parameters. The typical imprecise probability models include the Dempster-Shafer structure (Dempster, 1967; Shafer, 1976), probability boxes (Faes et al., 2021), and fuzzy probability model (Stein et al., 2013).

The Dempster-Shafer (DS) structure is rooted in the evidence theory (also called the Dempster-Shafer's theory of evidence). The evidence theory loosens the strict assumption in the probability theory that the probability measure $\mathcal{P}$ is precisely known, by taking two measures, i.e., belief and plausibility, into account for each event $\omega$ in the sample space $\Omega$. Consider an evidence space $(\Omega, \mathcal{F}, m)$, where $\mathcal{F}$ denotes the countable collection of focal elements of $\Omega$ and $m$ is the basic probability assignment (BPA). Note that, in the probability theory, $\mathcal{F}$ is required to be a $\sigma$-algebra, while for the evidence theory, it is not necessary. The BPA of a focal element $\mathcal{X} = \left[\underline{x}, \overline{x}\right] \subset \mathbb{R}$ can be defined as:

$$m(\mathcal{X}) = \begin{cases} > 0 & \text{if } \mathcal{X} \in \mathcal{F} \\ 0 & \text{if } \mathcal{X} \subset \Omega \text{ and } \mathcal{X} \notin \mathcal{F} \end{cases} \tag{1.10}$$

$$\sum_{\mathcal{X} \in \mathcal{F}} m(\mathcal{X}) = 1 \tag{1.11}$$

The BPA associates the likelihood no longer with the specific event $x \in \mathcal{X}$, but rather with the focal element $\mathcal{X}$. Figure 1.2 illustrates the difference between the probability measure and BPA. The belief and plausibility of the event $x$ is then expressed as:

$$\text{Bel}(x) = \sum_{\mathcal{X} \subset x} m(\mathcal{X}) \tag{1.12}$$

$$\text{Pl}(x) = \sum_{\mathcal{X} \cap x \neq \emptyset} m(\mathcal{X}) \tag{1.13}$$

It holds that $\text{Bel}(x) \leq \mathcal{P}_X(x) \leq \text{Pl}(x)$, and the probability theory can be seen as a special case of the evidence theory when $\text{Bel}(x) = \mathcal{P}_X(x) = \text{Pl}(x)$. In the above context, a DS structure is expressed by $\left\{ \left( \mathcal{X}_1 = [\underline{x}_1, \overline{x}_1], m(\mathcal{X}_1) \right), \left( \mathcal{X}_2 = [\underline{x}_2, \overline{x}_2], m(\mathcal{X}_2) \right), \cdots, \left( \mathcal{X}_{n_x} = [\underline{x}_d, \overline{x}_d], m(\mathcal{X}_{n_x}) \right) \right\}$. With the DS structure representation, the random variable $X$ can be characterized by a set of the cumulative belief function, corresponding to the lower bound, and cumulative plausibility function, corresponding to the upper bound, as:

$$\text{Bel}(X \leq x) = \sum_{\mathcal{X}_i \subset x} m(\mathcal{X}_i) \tag{1.14}$$

$$\text{Pl}(X \leq x) = \sum_{\mathcal{X}_i \cap x \neq \emptyset} m(\mathcal{X}_i) \tag{1.15}$$
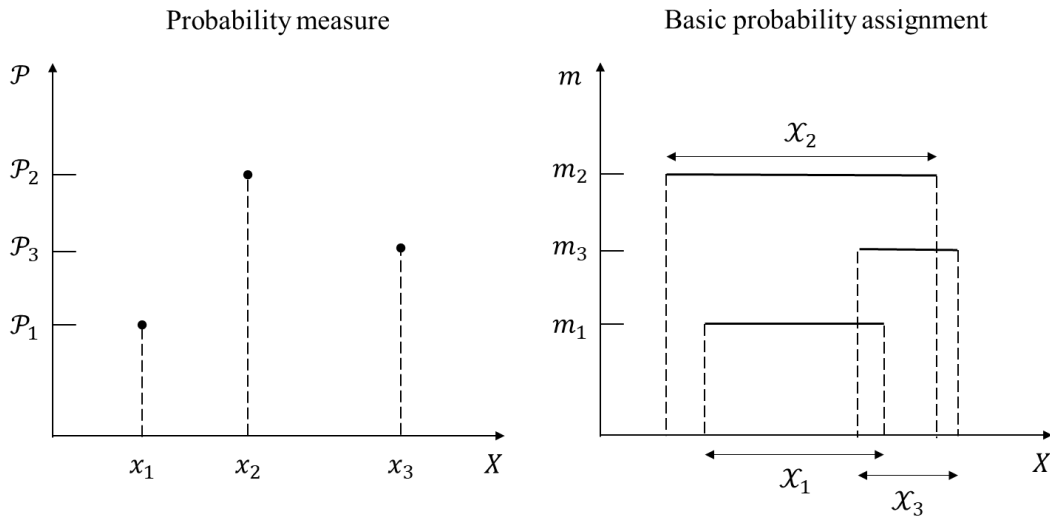


Figure 1.2 Illustration of probability measure and basic probability assignment.

Probability boxes (also called p-boxes) can be regarded as a combined use of the probability distributions and intervals. A p-box provides a pair of lower and upper CDFs $[\underline{F}_X, \overline{F}_X]$ for describing an unknown CDF $F_X$ of the random variable $X$. It holds that $\underline{F}_X(x) \le F_X(x) \le \overline{F}_X(x)$ for all $x \in \mathcal{D}_X$. These lower and upper CDFs are represented by the lower and upper probabilities of the event $\{X \le x\}$ as $\underline{F}_X(x) = \underline{\mathcal{P}}_X(X \le x)$ and $\overline{F}_X = \overline{\mathcal{P}}_X(X \le x)$, respectively. In case no further assumptions are made concerning the set of possible CDFs within the bounds, this type of p-boxes is usually denoted as *distribution-free p-boxes* (also called non-parametric p-boxes). This is the most general type of the p-boxes since any CDF complying with the lower and upper CDFs is admissible including, for example, non-smooth functions. In this context, a distribution-free p-box is expressed by the set of all the possible CDFs $\{F_X \in \mathbb{F} | \underline{F}_X(x) \le F_X(x) \le \overline{F}_X(x), x \in \mathcal{D}_X\}$, where $\mathbb{F}$ is the set of all CDFs on $\mathcal{D}_X$. The interval $[\underline{F}_X(x), \overline{F}_X(x)]$ represents the amount of epistemic uncertainty caused by incomplete information on $F_X$ for all $x \in \mathcal{D}_X$. The distribution-free p-boxes can be regarded as a continuous form of the DS structure. In fact, the belief function $\mathrm{Bel}(X \le x)$ is equivalent to $\underline{F}_X(x)$ and the plausibility function $\mathrm{Pl}(X \le x)$ is equivalent to $\overline{F}_X(x)$.

On the other hand, in case the distribution family associated with $X$ can also be determined, this type of p-boxes is in general denoted as *distributional p-boxes* (also called parametric p-boxes). A distributional p-box yields a family of CDFs whose hyper-parameters $\theta_i$ (e.g., the mean and variance) are unknown but contained within intervals $[\underline{\theta}_i, \overline{\theta}_i]$, for $i = 1, \cdots, n_\theta$. These hyper-parameters are collected as $\boldsymbol{\theta} \in \mathcal{D}_\Theta$, where $\mathcal{D}_\Theta \subset \mathbb{R}$ indicates the support domain of $\boldsymbol{\theta}$, and is represented by the Cartesian product $\mathcal{D}_\Theta = \times_{i=1}^{n_\theta} \theta_i = \times_{i=1}^{n_\theta} [\underline{\theta}_i, \overline{\theta}_i]$ when the intervals are independent each other. In this context, a distributional p-box is represented by the set of the constrained possible CDFs $\{F_X(\cdot | \boldsymbol{\theta}) \in \mathbb{F} | \boldsymbol{\theta} \in \mathcal{D}_\Theta\}$. The distributional p-boxes can clearly distinguish between aleatory uncertainty, described by the distribution families, and epistemic uncertainty, represented by the intervals for the hyper-parameters $\boldsymbol{\theta}$. The lower and upper CDFs of the distributional p-box can be computed as:

$$\underline{F}_X(x) = \min\{F_X(x|\boldsymbol{\theta}) \in \mathbb{F} | \boldsymbol{\theta} \in \mathcal{D}_\Theta\} \tag{1.16}$$

$$\overline{F}_X(x) = \max\{F_X(x|\boldsymbol{\theta}) \in \mathbb{F} | \boldsymbol{\theta} \in \mathcal{D}_\Theta\} \tag{1.17}$$

Figure 1.3 illustrates the above two types of p-boxes. A distributional p-box on the left-hand side consists of a Gaussian distribution family with the mean bounded as $\mu = [-1.0, 1.0]$ and the variance bounded as $\sigma = [0.6, 1.4]$. The lower CDF $\underline{F}_X$ is obtained as a combination of the distribution with maximum mean and minimum variance (Realization #2)

and the distribution with maximum mean and maximum variance (Realization #4). The upper CDF $\overline{F}_X$ is computed in a similar manner by a combination of the distributions with minimum mean. Besides, a distribution-free p-box on the right-hand side is bounded by the exactly same lower and upper CDFs as the above distributional p-box. Note that this distribution-free p-box is not equivalent to the distributional p-box on the left-hand side because the former p-box also yields the CDFs within the boundary CDFs not belonging to the distribution family of the latter p-box (i.e., the Gaussian distribution family).
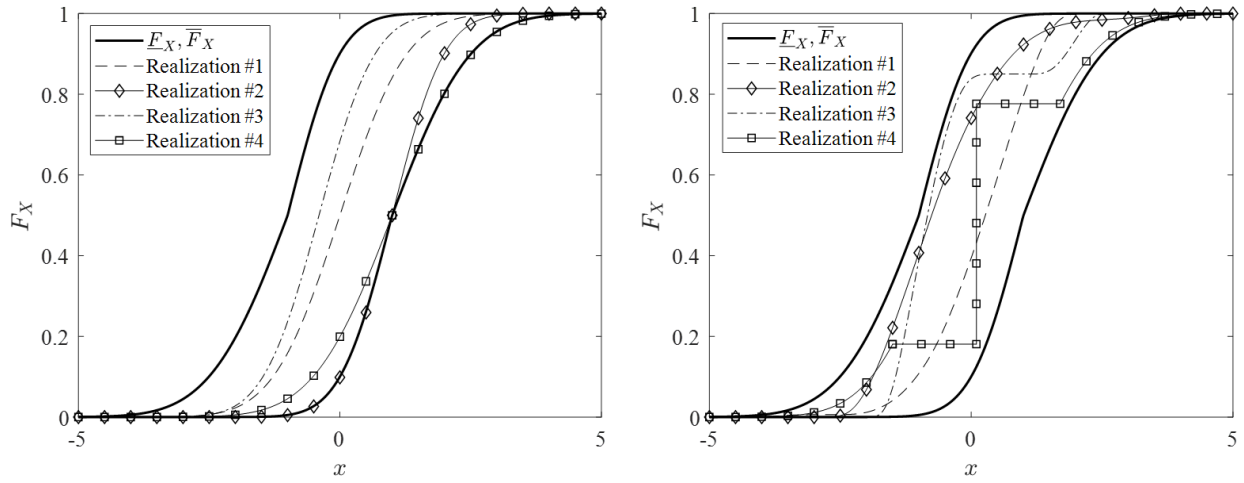


Figure 1.3 Illustration of distributional p-box (left figure) and distribution-free p-box (right figure).

A fuzzy distribution model can be regarded as an extension of the p-boxes as same way a fuzzy set model can be regarded as an extension of the intervals. Considering a fuzzy set $\tilde{\mathcal{E}}_{F_X} = \{F_X, \xi(F_X)|F_X \in \mathbb{F}\}$, the support of the fuzzy distribution can be defined by the interval $[\underline{F}_X, \overline{F}_X] = \{F_X \in \mathbb{F}|\xi(F_X) > 0\}$, then the corresponding $\alpha$-cut interval can be similarly defined as:

$$F_{X_\alpha} = [\underline{F}_{X_\alpha}, \overline{F}_{X_\alpha}] = \{F_X \in \mathbb{F}|\xi(F_X) \geq \alpha\} \tag{1.18}$$

In this context, the $\alpha$-cut procedure can provide nested intervals as $F_{X_{\alpha_i}}(x) \subseteq F_{X_{\alpha_j}}(x)$ for all $x \in \mathcal{D}_X$ when $\alpha_i \geq \alpha_j$.

Moreover, each $\alpha$-cut leads to a distribution-free p-box with the lower CDF $\underline{F}_{X_\alpha}$ and the upper CDF $\overline{F}_{X_\alpha}$. Besides to the above definition, it is also possible to define the fuzzy distribution by assigning a fuzzy set to the hyper-parameters $\boldsymbol{\theta}$ of the probability distribution. Considering a fuzzy set $\tilde{\mathcal{E}}_{\boldsymbol{\theta}} = \{\boldsymbol{\theta}, \xi(\boldsymbol{\theta})|\boldsymbol{\theta} \in \mathcal{D}_\Theta\}$, the support of the hyper-parameters is given by $\Theta = [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}] = \{\boldsymbol{\theta} \in \mathcal{D}_\Theta|\xi(\boldsymbol{\theta}) > 0\}$ and leads to a set of the possible CDFs $\{F_X(\cdot\,|\boldsymbol{\theta}) \in \mathbb{F}|\boldsymbol{\theta} \in \Theta\}$. Similarly, the $\alpha$-cut interval is defined as:

$$\boldsymbol{\theta}_\alpha = \left[\underline{\boldsymbol{\theta}}_\alpha, \overline{\boldsymbol{\theta}}_\alpha\right] = \{\boldsymbol{\theta} \in \mathcal{D}_\Theta | \xi(\boldsymbol{\theta}) \geq \alpha\} \tag{1.19}$$

and leads to a set of the possible CDFs $\{F_X(\cdot | \boldsymbol{\theta}) \in \mathbb{F} | \boldsymbol{\theta} \in \boldsymbol{\theta}_\alpha\}$. In this context, the $\alpha$-cut procedure can provide nested intervals as $\boldsymbol{\theta}_{\alpha_i} \subseteq \boldsymbol{\theta}_{\alpha_j}$ when $\alpha_i \geq \alpha_j$. Furthermore, each $\alpha$-cut leads to a distributional p-box with the interval valued hyper-parameters $\left[\underline{\boldsymbol{\theta}}_\alpha, \overline{\boldsymbol{\theta}}_\alpha\right]$. Therefore, larger $\alpha$ value implies higher risk that the unknown true probability distribution $F_X^*$ might be lie outside of the $\alpha$-cut interval.

### *Aleatory and epistemic spaces*

In this thesis, it is assumed that the category II parameters are described by the CDFs, the category III parameters are presented by the intervals, and the category IV parameters are expressed by the distributional p-boxes. In addition, the fuzzy sets and fuzzy distributions are supplementary utilized to determine the amount of epistemic uncertainty in the category III and IV parameters remaining after the uncertainty calibration process (which will be further discussed in the next subsection). The above assumption is consistent with many real-world UQ applications, such as the initial (2014) and latest (2019) editions of the NASA UQ challenge problem (Crespo et al., 2014; Crespo and Kenny, 2019). Indeed, most of the novel developments for uncertainty calibration and propagation in this thesis was motivated by the NASA UQ challenge 2019.

Finally, a double-loop sampling scheme is introduced to generate samples of the above three category parameters by differentiating between aleatory and epistemic uncertainty, as illustrated in Figure 1.4. An outer loop samples from the interval valued parameters, i.e., the category III parameters and the hyper-parameters of the category IV parameters. These interval valued parameters only contain epistemic uncertainty; therefore without loss of generality, the Cartesian product of all intervals $\mathcal{D}_e = \times_{i=1}^{n_y} \left[\underline{y}_i, \overline{y}_i\right] \times_{i=1}^{n_\theta} \left[\underline{\theta}_i, \overline{\theta}_i\right]$ can be denoted as *epistemic space*. A reduced epistemic space can be achieved through the uncertainty calibration process when additional information is available. In addition, the true values of the epistemic parameters $\boldsymbol{e}^* \in \mathcal{D}_e$ will result if epistemic uncertainty is eliminated. The above sampling is simply performed uniformly within the intervals. Note that this does not mean that these parameters are uniformly distributed in nature. Each epistemic space realization results in a precise CDF for the category IV parameters. Then, an inner loop samples from the CDFs of both the category II and IV parameters. These CDFs only contain aleatory uncertainty and the CDF values are bounded by the unit hyper-rectangle $\mathcal{D}_a = (0, 1]^{n_a}$, where $n_a$ is the total number of the category II and IV parameters. Hence, the hyper-rectangle $\mathcal{D}_a$ can be denoted as *aleatory space*, and it is clear that $\mathcal{D}_a$ cannot be reduced by further collecting additional information.
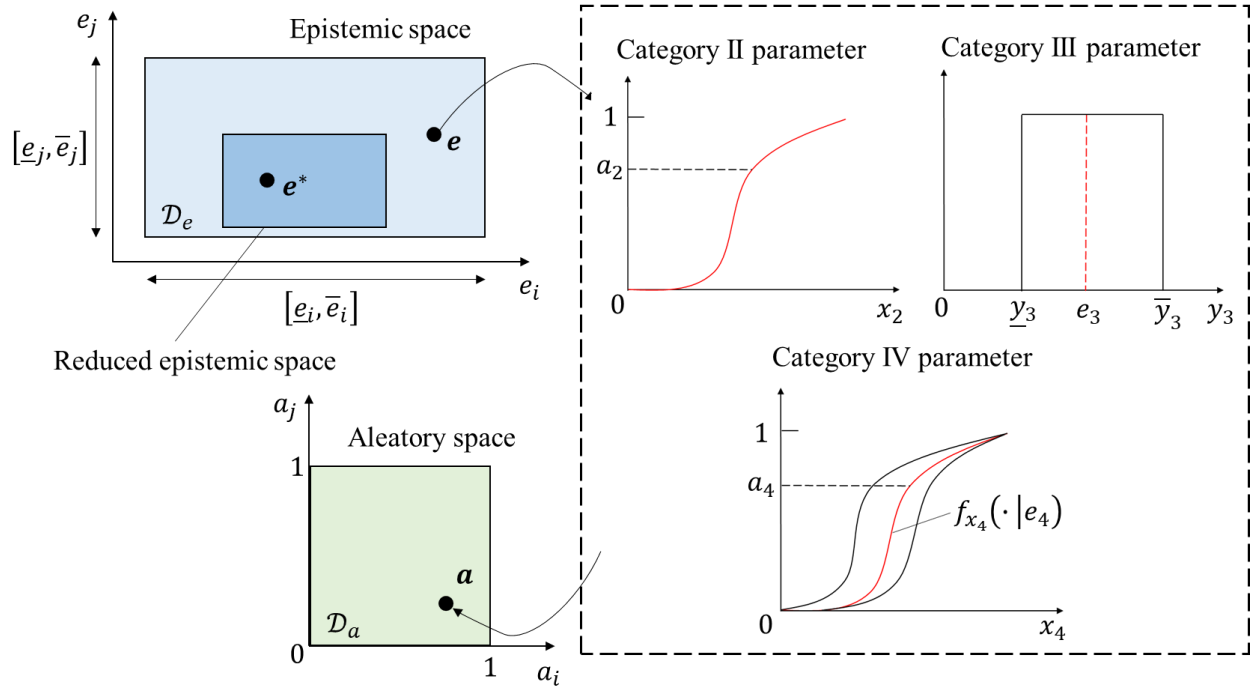
Figure 1.4 Double-loop sampling scheme in aleatory and epistemic spaces.

## 1.1.2  Uncertainty calibration

After the UM of all model parameters is constructed based on the available data for the parameters, the epistemic

space reflects the state of knowledge about the consistency of the numerical model to reality. The aim of uncertainty

calibration is to refine the UM by reducing the epistemic space, so that it improves the fidelity of the model predictions

to the actual behaviors of the system. This reduction is attained based on the availability of the quantitative data which

represents operating conditions of the system.

***Deterministic vs. stochastic model updating***

Uncertainty calibration can be generally classified into two paradigms: the deterministic and stochastic calibration

(Simoen et al., 2015). The deterministic calibration (also called the *deterministic model updating*) investigates only the

category III parameters, i.e., the parameters with only epistemic uncertainty, and is usually aimed at finding the optimal

parameter values which result in the best possible fit between the model predictions and observed data. This can be

attained by solving a (constrained) optimization problem where the objective is minimizing the discrepancy between

the model predictions and observations with respect to suitable response metrics which are sensitive to the variation of

the investigating parameters. The sensitivity method (Mottershead and Friswell, 1993) is one of the most successful

local approaches to the deterministic model updating, and in recent years has been utilized in a broad range of practical applications, including the correction of large-scale FE models (Mottershead et al., 2011). However, the deterministic calibration faces three main criticisms: (i) It assumes that the numerical model employed can fully describe the actual structure, i.e., it is not affected by modeling uncertainty; (ii) it also assumes that the observed data of the true UM $e^* \in \mathcal{D}_e$ can be ideally used, i.e., it is not affected by measurement uncertainty; (iii) it does not consider that the observed behaviors of the structure might inherently vary due to the presence of aleatory uncertainty.

The stochastic calibration (also called the *stochastic model updating*), on the other hand, investigates not only the category III but also the category IV parameters, i.e., aleatory uncertainty is present in the problem (Mares et al., 2006; Mottershead et al., 2006). In the similar sense with the deterministic model updating, the stochastic model updating is ultimately aimed at finding the optimal parameter distributions which result in the best possible fit between the clouds of the model predictions and observations. However, the inherent variability in the observed behaviors of the structure cannot be fully distinguished from the above two unavoidable sources of uncertainties, and thus it should be aimed at not a single set of the optimal epistemic parameters but a reduced epistemic space which achieves a trade-off between the fidelity-to-data and robustness-to-uncertainty. It is noted that, these two sources of uncertainties can exist regardless of the presence of aleatory uncertainty and model updating approaches under these uncertainties can, broadly speaking, be referred to as the stochastic model updating. The stochastic model updating methodologies in the literature can be generally categorized into two groups: probabilistic approaches, e.g., Monte Carlo based model updating (Sairajan and Aglietti, 2012; Bi et al., 2013) and perturbation methods (Khodaparast et al., 2008), and non-probabilistic approaches, e.g., interval model updating (Khodaparast et al., 2011; Govers et al., 2015). This thesis focuses on the probabilistic stochastic model updating.

***Bayesian model updating***

Among the probabilistic approaches, one of the most well-established is the Bayesian model updating framework (Beck and Katafygiotis, 1998; Katafygiotis and Beck, 1998). A key advantage of the Bayesian model updating lies in its capability to combine prior information on the parameters to be inferred with the observed data to yield a posterior stochastic representation of the parameters that indicates the state of knowledge after integrating additional information yielded by the observed data. This is achieved through the well-known Bayes' theorem:

$$P(\boldsymbol{\vartheta}|\boldsymbol{D}) = \frac{P_L(\boldsymbol{D}|\boldsymbol{\vartheta})P(\boldsymbol{\vartheta})}{P(\boldsymbol{D})} \tag{1.20}$$

whereby

- $P(\boldsymbol{\vartheta})$ represents the prior distribution of the parameters to be inferred $\boldsymbol{\vartheta}$;

- $P(\boldsymbol{\vartheta}|\boldsymbol{D})$ represents the posterior distribution of the parameters $\boldsymbol{\vartheta}$;

- $P(\boldsymbol{D})$ represents the evidence of the observed data $\boldsymbol{D}$;

- $P_L(\boldsymbol{D}|\boldsymbol{\vartheta})$ represents the likelihood function of the observed data $\boldsymbol{D}$ for an instance of $\boldsymbol{\vartheta}$.

The prior distribution $P(\boldsymbol{\vartheta})$ reflects one's a priori knowledge on the calibrating parameters $\boldsymbol{\vartheta}$ before any observed data is available. The category III parameters and the hyper-parameters of the category IV parameters are the candidates of the calibrating parameters. Without loss of generality, independence between the calibrating parameters is assumed and the prior distribution is thus expressed by $P(\boldsymbol{\vartheta}) = P(\vartheta_1, \vartheta_2, \cdots, \vartheta_{n_\vartheta}) = \prod_{i=1}^{n_\vartheta} P(\vartheta_i)$. A priori knowledge may come in forms of expert opinions, lab-scale experiments, and previous uncertainty calibration results. In theory, any type of distribution can be chosen depending on the amount of available information, and the principle of maximum entropy is most widely employed for choosing the prior PDF. The principle of maximum entropy expresses that the probability distribution that best represents the current state of knowledge on the parameters $\boldsymbol{\vartheta}$ is the distribution resulting in the largest information entropy. The principle results in the prior distribution e.g., as a Gaussian distribution when a priori available information of a parameter consists of the mean and relative error, and a uniform distribution when only the lower and upper bounds are known for a parameter. It should be noted that, for multivariate cases, the principle always leads to independent prior PDF, which supports the above independent assumption between the calibrating parameters. It is also important to note that, by adopting the uniform distribution as the prior distribution, the posterior distribution will be simply proportional to the likelihood function. In this thesis, it is assumed that the category III parameters and the hyper-parameters of the category IV parameters are represented by the intervals; hence, a uniform distribution on the epistemic space $\mathcal{D}_e$ is used as the prior.

The posterior distribution $P(\boldsymbol{\vartheta}|\boldsymbol{D})$ reflects the updated knowledge on the parameters $\boldsymbol{\vartheta}$ based on new information obtained from the observed data $\boldsymbol{D}$. Except for some special cases, such as when conjugate prior distributions are used, the posterior distribution cannot be derived analytically. Alternatively, it is usually estimated through sampling methods, such as Markov chain Monte Carlo (MCMC) methods (Tierney, 1994). In general, the support domain of the posterior distribution occupies a much smaller volume than that of the prior, because combinations of the parameters that fail to describe the observed data are excluded from the posterior. In this context, the support domain of the posterior can be regarded as the reduced epistemic space. At the same time, the posterior distribution the maximum value of which is

normalized to one can serve as a membership function to treat the parameters as fuzzy variables. It is noted that, when the parameters $\boldsymbol{\vartheta}$ are the hyper-parameters of the category IV parameters, the resultant category IV parameters will be represented as the fuzzy distributions. With this procedure, a larger $\alpha$-cut level results in a smaller reduced epistemic space. When the $\alpha$-cut level is $\alpha = 1$, the reduced epistemic space is identical with the support domain of the posterior whilst, when the $\alpha$-cut level is $\alpha = 0$, it comprises a set of the crisp values that correspond to the most probable values (MPVs) of the posterior. If the posterior distribution is sharply peaked, the analyst might employ $\alpha = 0$ with the high confidence that the MPVs are equivalent to the true values of the epistemic parameters $\boldsymbol{e}^*$. Conversely, if the posterior distribution is relatively flat, the analyst might use an adequately small $\alpha$-cut level to ensure that the reduced epistemic space still include the true values $\boldsymbol{e}^*$.

The evidence (also called as marginal likelihood) $P(\boldsymbol{D})$ is expressed as:

$$P(\boldsymbol{D}) = \int_{\mathcal{D}_{\boldsymbol{\vartheta}}} P_L(\boldsymbol{D}|\boldsymbol{\vartheta})P(\boldsymbol{\vartheta})d\boldsymbol{\vartheta} \tag{1.21}$$

The evidence is a constant and serves as a normalization factor to ensure that the posterior distribution integrates to one. Furthermore, it measures the plausibility of the assumed model class $M$ given by the observed data $\boldsymbol{D}$. Therefore, one can use multiple candidate model classes to describe the observed data and associate the evidence with each model class to determine the most probable model class. This procedure is referred to as Bayesian model class selection (Beck and Yuen, 2004). However, it is typically non-trivial to compute the evidence due to the multi-dimensional integral in Equation (1.21), and is thus usually estimated numerically using MCMC methods (see, e.g., Ching and Cheng, 2007; Cheung and Beck, 2009). On the contrary, the evidence does not affect the shape of the posterior distribution and can be neglected in sampling from the posterior distribution. In this manner, the posterior distribution would be known up to a proportionality, i.e.,

$$P(\boldsymbol{\vartheta}|\boldsymbol{D}) \propto P_L(\boldsymbol{D}|\boldsymbol{\vartheta})P(\boldsymbol{\vartheta}) \tag{1.22}$$

and point-wise for any generated samples of $\boldsymbol{\vartheta}$. Generally, methods for sampling from the posterior distribution are not necessarily suitable to estimate the evidence, and vice versa. However, the Bayesian inference algorithms used in this thesis can address both tasks simultaneously as summarized later.

The likelihood function $P_L(\boldsymbol{D}|\boldsymbol{\vartheta})$ is the key component in the Bayesian model updating and reflects the degree of agreement between the model predictions $M(\boldsymbol{\vartheta})$ and observed data $\boldsymbol{D}$. The mathematical relation between the model predictions and observed data can be generally expressed by $\boldsymbol{D} = M(\boldsymbol{\vartheta}) + \epsilon$, where $\epsilon$ indicates the error caused by the

three sources of uncertainties (i.e., parameter uncertainty, modeling uncertainty, and measurement uncertainty). The

likelihood function will return a high value for an instance of $\boldsymbol{\vartheta}$ when the corresponding error $\epsilon$ is small and low value

when the error is large. Assuming that the observations $\boldsymbol{D}^{(k)} \in \mathbb{R}^{1 \times n}$ ($k = 1, \cdots, N_{\mathrm{obs}}$), comprising $n$ response metrics,

are independently and identically distributed, the likelihood function can be then expressed as:

$$P_L(\boldsymbol{D}|\boldsymbol{\vartheta}) = \prod_{k=1}^{N_{\mathrm{obs}}} P\left(\boldsymbol{D}^{(k)}\middle|\boldsymbol{\vartheta}\right) \tag{1.23}$$

It is noteworthy that the likelihood function is a function of $\boldsymbol{\vartheta}$ and not of $\boldsymbol{D}$. When $\boldsymbol{\vartheta}$ comprises only the category III

parameters, the model outputs $\boldsymbol{z}(\boldsymbol{\vartheta}) \in \mathbb{R}^{1 \times n}$: $\boldsymbol{z} = M(\boldsymbol{\vartheta})$, for an instance of $\boldsymbol{\vartheta}$, would be deterministic. In this context,

the uncertainty in the observations is purely due to measurement uncertainty. Since the model outputs are deterministic,

the assumption that the error between the model outputs and observations follows a Gaussian distribution with zero

means and fixed variances is often made to assign a probability distribution to each $\boldsymbol{\vartheta}$ instance. It leads to the common

choice of the likelihood function by the Gaussian distribution:

$$P_L(\boldsymbol{D}|\boldsymbol{\vartheta}) = \prod_{k=1}^{N_{\mathrm{obs}}} \left(\prod_{i=1}^{n} \frac{1}{\sigma_i \sqrt{2\pi}}\right) \exp\left[-\sum_{i=1}^{n} \frac{\left(D_i^{(k)} - z_i(\boldsymbol{\vartheta})\right)^2}{2\sigma_i^2}\right] \tag{1.24}$$

However, the criticism exists that the magnitude of the errors is often not precisely known. On the other hand, when $\boldsymbol{\vartheta}$

comprises the hyper-parameters of the category IV parameters (and the category III parameters), then the model outputs

$\boldsymbol{z}(\boldsymbol{\vartheta}) \in \mathbb{R}^{N_{\mathrm{sim}} \times n}$, for an instance of $\boldsymbol{\vartheta}$, would be stochastic. In this context, the variability in the observations is due to

the inherent property of the system while they are also contaminated by measurement uncertainty. Because the model

outputs are stochastic, they compose the probability distribution and thus the probability $P\left(\boldsymbol{D}^{(k)}\middle|\boldsymbol{\vartheta}\right)$ in Equation (1.23)

can be computed by $P\left(\boldsymbol{D}^{(k)}\middle|\boldsymbol{\vartheta}\right) = f_{\boldsymbol{z}(\boldsymbol{\vartheta})}\left(\boldsymbol{D}^{(k)}\right)$. Nevertheless, the PDF of the model outputs is required to be estimated

respectively for each instance of $\boldsymbol{\vartheta}$. This estimation is typically performed by the kernel density estimation (KDE), and

a tremendous number of model outputs is necessary for precise estimation of the PDF. Furthermore, the total number

of instances would also be significantly large throughout the Bayesian model updating; hence, in practical applications,

the repeated KDE for every instance of $\boldsymbol{\vartheta}$ can be rather time-consuming.

***Approximate Bayesian computation***

Recently, approximate Bayesian computation (ABC) methods (Turner and Zandt, 2012; Safta et al., 2015) have

been developed to circumvent the computationally expensive or even intractable likelihood evaluations by replacing

the full likelihood function with the kernel based on selected summary statistics:

$$P_L(\boldsymbol{D}|\boldsymbol{\vartheta}) \propto \frac{1}{\varepsilon} K\left(\frac{\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|}{\varepsilon}\right) \tag{1.25}$$

where $K(\cdot)$ denotes the kernel function, $s(\cdot)$ indicates the selected summary statistics based on the data, e.g., the means and quantiles, and $\varepsilon$ indicates the so-called width factor which controls the radius of the acceptable hypersphere around the target statistic $s(\boldsymbol{D})$. In other words, a large $\varepsilon$ leads to a flat posterior distribution and, inversely, a small $\varepsilon$ leads to a peaked posterior distribution. The peaked posterior implies the parameter is more likely to converge to the true value, while a too small $\varepsilon$ results in a necessity of significant calculation cost for convergence. The determination of $\varepsilon$ is thus the trade-off between the accuracy and efficiency, and depends on specific applications. A recommended interval of $\varepsilon$ is $[10^{-3}, 10^{-1}]$ (Patelli et al., 2015). The kernel $K$ will return a high value when the norm $\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|$ is small, but it will penalize the instance $\boldsymbol{\vartheta}$ when the norm is large. Therefore, the kernel based on the summary statistics can serve as an approximation of the likelihood function. Furthermore, in general, its computation is much affordable since estimation of the summary statistics requires a smaller number of the model outputs compared with that of the PDF. Various types of the kernel shape are investigated in the ABC literature, such as the Gaussian kernel (Rocchetta et al., 2018a):

$$\frac{1}{\varepsilon} K\left(\frac{\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|}{\varepsilon}\right) = \frac{1}{\varepsilon\sqrt{2\pi}} \exp\left(-\frac{\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|^2}{2\varepsilon^2}\right) \tag{1.26}$$

the Epanechnikov kernel (Beaumont et al., 2002):

$$\frac{1}{\varepsilon} K\left(\frac{\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|}{\varepsilon}\right) = \begin{cases} \frac{3}{4\varepsilon}\left(1 - \frac{\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|^2}{\varepsilon^2}\right) & \text{if } \|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\| \leq \varepsilon \\ 0 & \text{if } \|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\| > \varepsilon \end{cases} \tag{1.27}$$

and the sharp kernel (Safta et al., 2015):

$$\frac{1}{\varepsilon} K\left(\frac{\|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\|}{\varepsilon}\right) = \begin{cases} \frac{1}{2\varepsilon} & \text{if } \|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\| \leq \varepsilon \\ 0 & \text{if } \|s(\boldsymbol{D}) - s(\boldsymbol{z}(\boldsymbol{\vartheta}))\| > \varepsilon \end{cases} \tag{1.28}$$

Figure 1.5 shows the shapes of the above three kernel functions for the case that the width-factor is selected to be $\varepsilon = 0.01$. In this thesis, the Gaussian kernel is used since it has a long tail that results in the smooth transition from prior to posterior.
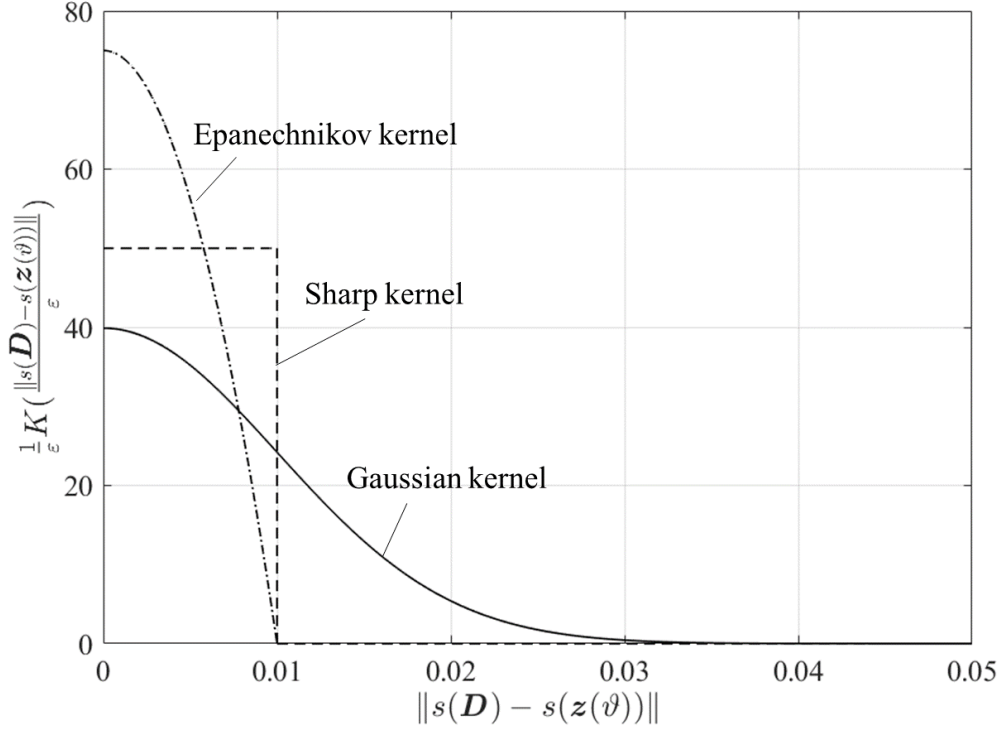
Figure 1.5 Illustration of kernel functions for $\varepsilon = 0.01$.

### Distance-based UQ metrics

In general, the norm $\left\|s(\boldsymbol{D}) - s\big(\boldsymbol{z}(\boldsymbol{\vartheta})\big)\right\|$ is computed using the $L_2$ norm as:

$$\left\|s(\boldsymbol{D}) - s\big(\boldsymbol{z}(\boldsymbol{\vartheta})\big)\right\| = \sqrt{\sum_{i=1}^{n}\big[s(\boldsymbol{D}_i) - s\big(\boldsymbol{z}_i(\boldsymbol{\vartheta})\big)\big]^2} \tag{1.29}$$

When the summary statistic employed is the mean as $s(\boldsymbol{D}) = \overline{\boldsymbol{D}}$ and $s\big(\boldsymbol{z}(\boldsymbol{\vartheta})\big) = \overline{\boldsymbol{z}}(\boldsymbol{\vartheta})$, Equation (1.29) is equivalent to the well-known Euclidian distance:

$$d_E\big(\boldsymbol{D}, \boldsymbol{z}(\boldsymbol{\vartheta})\big) = \sqrt{\sum_{i=1}^{n}[\overline{\boldsymbol{D}}_i - \overline{\boldsymbol{z}}_i(\boldsymbol{\vartheta})]^2} = \sqrt{\big(\overline{\boldsymbol{D}} - \overline{\boldsymbol{z}}(\boldsymbol{\vartheta})\big)\big(\overline{\boldsymbol{D}} - \overline{\boldsymbol{z}}(\boldsymbol{\vartheta})\big)^T} \tag{1.30}$$

The Euclidian distance has been without doubt the most widely used metric, in the field of model updating, to quantify the difference between the model predictions and observations. Nevertheless, as a point-to-point distance measure, the Euclidian distance quantifies the geometric distance between the centres of mass of two datasets, but it discards other useful information such as dispersion information of the datasets. As a matter of fact, in the Bayesian model updating,

the Euclidian distance-based approximate likelihood function is effective in calibrating the category II parameters and the mean parameters of the category IV parameters which are directly related to the centre of mass of the model outputs, however it is unsuitable in precisely inferring the higher-order moment parameters of the category IV parameters (Bi et al., 2019).

In the above context, it is natural to use statistical distances quantifying higher statistical information. For example, the Mahalanobis distance, more precisely pooled Mahalanobis distance, quantifies the correlation between two datasets by employing their covariance matrices as a weighting factor in the Euclidian distance. Its mathematical definition is given as:

$$d_M\big(\boldsymbol{D}, \boldsymbol{z}(\boldsymbol{\vartheta})\big) = \sqrt{\big(\overline{\boldsymbol{D}} - \overline{\boldsymbol{z}}(\boldsymbol{\vartheta})\big) C_{\text{pool}}^{-1}\big(\overline{\boldsymbol{D}} - \overline{\boldsymbol{z}}(\boldsymbol{\vartheta})\big)^T}, C_{\text{pool}} = \frac{(N_{\text{obs}} - 1)C_{\boldsymbol{D}} + (N_{\text{sim}} - 1)C_{\boldsymbol{z}}}{N_{\text{obs}} + N_{\text{sim}} - 2} \tag{1.31}$$

where $C_{\text{pool}}^{-1}$ denotes the pooled covariance matrix involving the covariance matrix of the observed data $C_{\boldsymbol{D}}$ and that of the model outputs $C_{\boldsymbol{z}}$. While the pooled Mahalanobis distance is a population-to-population distance, it is proven to be inappropriate for the purpose of calibrating the higher-order moment parameters of the category IV parameters, because covariance information merely measures the dependency between the two datasets but does not measure the similarity between them (Bi et al., 2017). Thus, the Bhattacharyya distance (Bhattacharyya, 1946), that is a more comprehensive statistical distance quantifying the difference between probability distributions of two datasets, is also investigated in Bi et al. (2017). Its original definition is given as:

$$d_B\big(\boldsymbol{D}, \boldsymbol{z}(\boldsymbol{\vartheta})\big) = -\log\left[\int_{\mathcal{D}_{\boldsymbol{Z}}} \sqrt{f_{\boldsymbol{D}}(\boldsymbol{z}) f_{\boldsymbol{z}(\boldsymbol{\vartheta})}(\boldsymbol{z})} d\boldsymbol{z}\right] \tag{1.32}$$

where $f_{\boldsymbol{D}}(\cdot)$ and $f_{\boldsymbol{z}(\boldsymbol{\vartheta})}(\cdot)$ denote the PDFs of the observed data and model outputs, respectively, and $\int_{\mathcal{D}_{\boldsymbol{Z}}}(\cdot)d\boldsymbol{z}$ indicates the integration performed over the whole support domain. The Bhattacharyya distance is essentially a measure of the overlap between two probability distributions; thus, not only mean information but whole statistical information of the two datasets can be considered. Nevertheless, because it requires PDF estimation, usually via KDE, the computational cost of the Bhattacharyya distance-based approximate likelihood is equivalent to that of the full likelihood function. In addition, a converged estimation in KDE is generally unavailable for the observed data due to the very limited number of observations, especially for applications where measurements are difficult or expensive. To tackle this issue, the so-called binning algorithm is proposed in Bi et al. (2019) to evaluate the probability mass function (PMF) of a discrete distribution, so that the discrete Bhattacharyya distance (Patra et al., 2015) is used instead:

$$d_B\big(\boldsymbol{D}, \boldsymbol{z}(\boldsymbol{\vartheta})\big) = -\log\left(\sum_{k=1}^{N_{\text{bin}}} \sqrt{P_{\boldsymbol{D}}^{(k)} P_{\boldsymbol{z}(\boldsymbol{\vartheta})}^{(k)}}\right) \tag{1.33}$$

where $P_{\boldsymbol{D}}^{(k)}$ and $P_{\boldsymbol{z}(\boldsymbol{\vartheta})}^{(k)}$ are the PMF values at the $k$th bin for the observed data and model outputs, respectively. In general, the PMF evaluation requires much fewer model outputs compared with PDF estimation; hence, the computational cost of the Bhattacharyya distance-based approximate likelihood function would be affordable. The binning algorithm for the PMF evaluation consists of the following steps:

1) Define a common interval $I_i$ of both the observed data $\boldsymbol{D}_i$ and model outputs $\boldsymbol{z}_i(\boldsymbol{\vartheta})$, for $i = 1, \cdots, n$, according to the $i$th response metric, by finding the general minimum and maximum in both $\boldsymbol{D}_i$ and $\boldsymbol{z}_i(\boldsymbol{\vartheta})$;

2) Divide each defined interval into an arbitrarily decided number of bins $n_{\text{bin}}$;

3) Count the joint PMF values at each bin, $P_{\boldsymbol{D}}^{(k)}$ and $P_{\boldsymbol{z}(\boldsymbol{\vartheta})}^{(k)}$. Note that the total number of bins employed is $N_{\text{bin}} = n_{\text{bin}}^{n}$.

The principle of $n_{\text{bin}}$ in Step 2) is that a smaller $n_{\text{bin}}$ results in a smaller value of the Bhattacharyya distance and larger $n_{\text{bin}}$ in a larger value of the Bhattacharyya distance. In the extreme case when $n_{\text{bin}} = 1$, Equation (1.33) always return the Bhattacharyya distance $d_B\big(\boldsymbol{D}, \boldsymbol{z}(\boldsymbol{\vartheta})\big) = 1$ for arbitrary $\boldsymbol{D}$ and $\boldsymbol{z}(\boldsymbol{\vartheta})$. In addition, too small as well as too large $n_{\text{bin}}$ results in a significantly biased PMF evaluation, as shown in Figure 1.6; thus, the appropriate choice of $n_{\text{bin}}$ is crucial to achieve reasonable updating results.
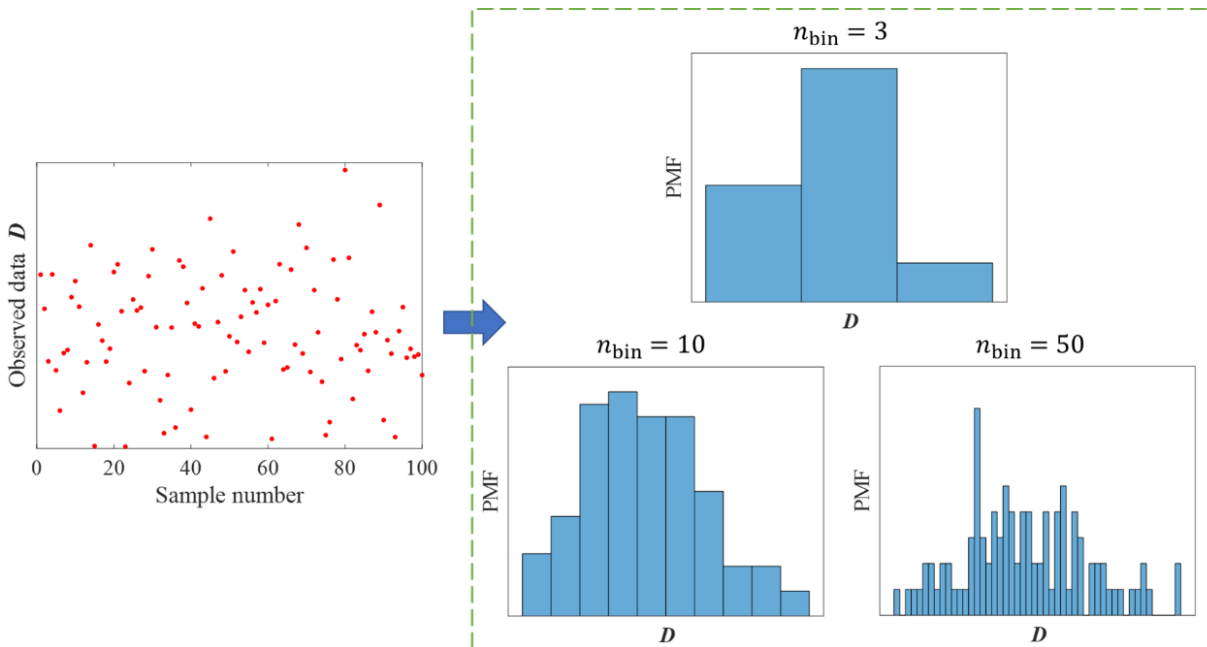


Figure 1.6 PMF of 100 data points for different number of bins $n_{\text{bin}}$.

***Two-step ABC updating with the Euclidian and Bhattacharyya distances***

Another important aspect of the Bhattacharyya distance is that, as long as two datasets have no overlap, it returns a value that is infinite and is thus insensitive to the relative position of the centres of mass of them. To this end, a two-step ABC updating framework is proposed in Bi et al. (2019), where the preliminary step is utilized to force an overlap between $\boldsymbol{D}$ and $\boldsymbol{z}(\boldsymbol{\vartheta})$ by using Euclidian distance-based likelihood, while comprehensive uncertainty characteristics of them are further quantified in the main step using the Bhattacharyya distance-based likelihood. Figure 1.7 shows the principle illustration of the two-step ABC updating framework. The feasibility of this framework is demonstrated on a subproblem related to uncertainty calibration (Subproblem A) of the NASA UQ challenge 2014. The high-dimensional epistemic space is reduced satisfactory according to the available observed data by updating both the mean and variance parameters. Moreover, it is reported that the computational burden is significantly reduced compared with the Bayesian approach using the full likelihood function (Safta et al., 2015) by a factor $10^4$ whereas achieving equivalent updating results.
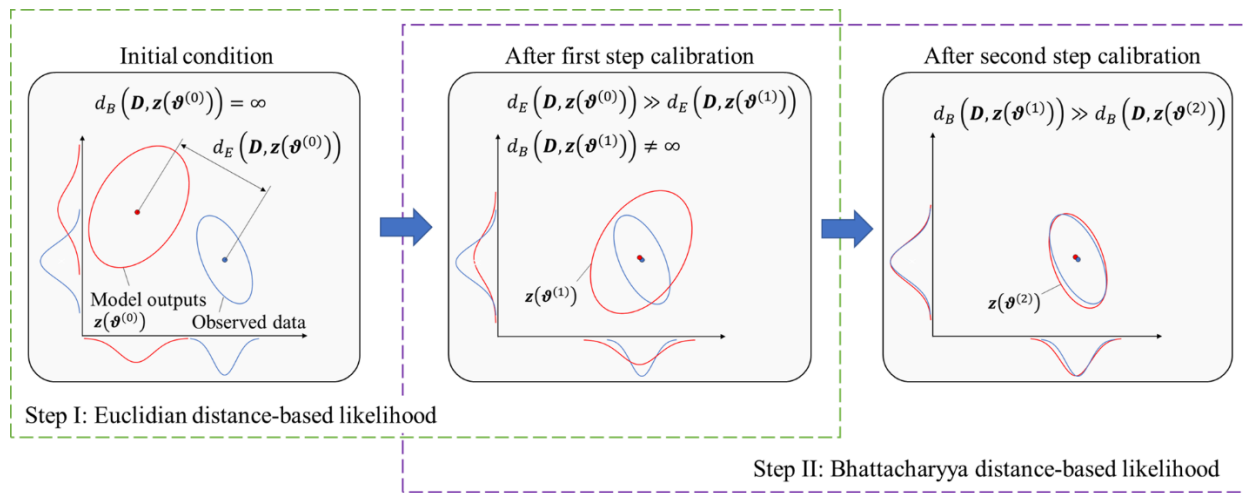


Figure 1.7 Illustration of two-step ABC updating framework.

The above two-step ABC updating framework with both the Euclidian and Bhattacharyya distances is a favorable general-purpose approach for uncertainty calibration. However, open problems still exist: (i) While the computational cost of the Bhattacharyya distance-based likelihood estimation is much lower than that of the full likelihood estimation, it is still not affordable for practical applications using high-fidelity models (e.g., FE models) because of the necessity of hundreds to thousands repeated model evaluations for each instance of the parameters $\boldsymbol{\vartheta}$; (ii) The original definition of the Bhattacharyya distance is only applicable to scalar quantities (e.g., system natural frequencies) since PMFs need

to be estimated based on a number of samples, and cannot be directly employed for time-domain sequences; (iii) The category IV parameters are characterized by the distributional p-boxes, however, the distribution families are often not known beforehand due to the scarce and/or incomplete available data for the parameters, i.e., the distribution families themselves can be epistemic uncertainty to be calibrated. Note that, instead of the Bhattacharyya distance, other types of stochastic metrics, such as the Wasserstein distance (Panaretos and Zemel, 2019) and the Jensen-Shannon divergence (Lin, 1991), have also been proposed in the ABC literature (see, e.g., Bernton et al., 2019; Yang et al., 2022), but they equally face the similar problems above in principle.

### *Transitional Markov chain Monte Carlo*

As mentioned above, in the Bayesian model updating, the posterior distribution is usually estimated by sampling methods. However, sampling from the posterior distribution is not a straightforward task because the posterior is only known implicitly, i.e., point-wise in terms of the outcome of the numerical model employed. Furthermore, the posterior distribution can be appeared as a broad range of distribution shapes from flat to very peaked one, and even as a multi-modal distribution because of the potential ill-posedness of the model updating problem. Therefore, the standard Monte Carlo method is inapplicable and MCMC methods constitute a popular class of methods to sample from the posterior distribution. One problem of MCMC methods is that an appropriate length of the burn-in period (i.e., the initial phase until the Markov chain reaches the stationary distribution) to ensure that the Markov chain produces samples from the posterior is often non-trivial. In addition, many MCMC methods become inefficient when the number of the parameters to be inferred is large.

The transitional Markov chain Monte Carlo (TMCMC) method (Ching and Cheng, 2007) belongs the class of the sequential particle filter methods (Chopin, 2002), and is based on MCMC sampling. The method tries to overcome the aforementioned problems in MCMC by gradually pushing the samples from the prior to posterior distribution by means of a sequence of the non-normalized intermediate distributions $\{P_j(\boldsymbol{\vartheta}): j = 0, \cdots, m\}$ that converge to the posterior. For this purpose, Equation (1.22) is modified to:

$$P_j(\boldsymbol{\vartheta}) \propto P_L(\boldsymbol{D}|\boldsymbol{\vartheta})^{q_j} P(\boldsymbol{\vartheta}) \tag{1.34}$$

where $q_j \in [0, 1]$ is chosen such that $q_0 = 0 < q_1 < \cdots < q_{m-1} < q_m = 1$. Consequently, the method starts sampling from the prior distribution (i.e., $P_0(\boldsymbol{\vartheta}) = P(\boldsymbol{\vartheta})$), thus initially populating all the space of parameters, and finally allows to sample from the posterior distribution up to a proportionality (i.e., $P_m(\boldsymbol{\vartheta}) \propto P(\boldsymbol{\vartheta}|\boldsymbol{D})$). In this context, the parameter $q_j$ can be interpreted as a measure of the amount of information employed in the $j$th intermediate step. The principle

behind the TMCMC method is to smoothly and gradually transit from the prior to posterior distribution by choosing a proper value of $q_j$. Consider the plausibility weight of the $k$th sample generated at the $j$th step, $w_{jk}$, defined as:

$$w_{jk} = \frac{P_{j+1}(\boldsymbol{\vartheta}_{jk})}{P_j(\boldsymbol{\vartheta}_{jk})} = P_L(\boldsymbol{D}|\boldsymbol{\vartheta}_{jk})^{q_{j+1}-q_j} \tag{1.35}$$

The degree of uniformity of the plausibility weights is an indicator to evaluate the similarity between two consecutive intermediate distributions $P_j(\boldsymbol{\vartheta})$ and $P_{j+1}(\boldsymbol{\vartheta})$. Therefore, the value of $q_{j+1}$ is chosen so that the coefficient of variation (COV) of the plausibility weights is close to one (Betz et al., 2016):

$$q_{j+1} = \arg\min_q \left( \left| CV_{w_j}(q) - 1 \right| \right) \tag{1.36}$$

where $CV_{w_j}(q)$ with $q \in [q_j, 1]$ is the COV of the set of the plausibility weights $\{w_{jk} = P_L(\boldsymbol{D}|\boldsymbol{\vartheta}_{jk})^{q-q_j} : k = 1, \cdots, N_s\}$, where $N_s$ indicates the number of samples generated at each step. Equation (1.36) can be solved, for example, by the bisection method. The smooth and gradual transition between two consecutive intermediate distributions results in the capability of the TMCMC method in inferring large number of parameters at one time (Ortiz et al., 2015; Patelli et al., 2017) and sampling from complex-shaped distributions (Patelli et al., 2015).

Once $q_{j+1}$ $(j = 0, \cdots, m-1)$ is estimated, MCMC sampling using the Metropolis-Hastings algorithm (Hastings, 1970) is employed for sampling from the $j + 1$th intermediate distribution. The seeds of the Markov chains are selected from samples at the $j$th step with a probability that is identical to their normalized plausibility weights:

$$\overline{w}_{jk} = \frac{w_{jk}}{\sum_{l=1}^{N_s} w_{jl}} \tag{1.37}$$

The normalized plausibility weight is a measure of the plausibility that the sample $\boldsymbol{\vartheta}_{jk}$ follow the distribution $P_{j+1}(\boldsymbol{\vartheta})$. As such, the selected seeds are asymptotically distributed as $P_{j+1}(\boldsymbol{\vartheta})$, indicating all Markov chains reach the stationary distribution from the very beginning, thus the burn-in period is unnecessary. Note that all Markov chains are perfectly parallel and the TMCMC method can be easily parallelized. In general, the proposal distribution is chosen as a Gaussian distribution centered at the seed sample. The covariance matrix of the Gaussian proposal distribution is expressed as:

$$\Sigma_j = \beta^2 \sum_{k=1}^{N_s} \overline{w}_{jk} (\boldsymbol{\vartheta}_{jk} - \overline{\boldsymbol{\vartheta}}_j)(\boldsymbol{\vartheta}_{jk} - \overline{\boldsymbol{\vartheta}}_j)^T, \overline{\boldsymbol{\vartheta}}_j = \sum_{k=1}^{N_s} \overline{w}_{jk} \boldsymbol{\vartheta}_{jk} \tag{1.38}$$

where $\beta^2$ indicates the scaling parameter to control the rejection rate in MCMC sampling. Furthermore, the TMCMC

method can estimate the evidence $P(\boldsymbol{D})$ as a by-product (Ching and Cheng, 2007):

$$P(\boldsymbol{D}) = \prod_{j=0}^{m-1} \left( \frac{1}{N_s} \sum_{k=1}^{N_s} w_{jk} \right) \tag{1.39}$$

The TMCMC algorithm can be summarized as follows: Note that the logarithmic of the likelihood function (log-likelihood) $\mathcal{L}(\boldsymbol{D}|\boldsymbol{\vartheta}) = \ln P_L(\boldsymbol{D}|\boldsymbol{\vartheta})$ is used instead to avoids numerical problems such as arithmetic underflow with the computation of the likelihood function.

1. Initialize $N_s$ and $\beta^2$;

2. Let $j = 0$ and $q_j = 0$. Generate $N_s$ samples of the calibrating parameters $\boldsymbol{\vartheta}_{jk}$ ($k = 1, \cdots, N_s$), according to the prior distribution $P(\boldsymbol{\vartheta})$;

3. Determine $q_{j+1}$ by solving Equation (1.36);

4. For all samples $\boldsymbol{\vartheta}_{jk}$ ($k = 1, \cdots, N_s$), compute the plausibility weight by $w_{jk} = \exp\{(q_{j+1} - q_j)\mathcal{L}(\boldsymbol{D}|\boldsymbol{\vartheta}_{jk})\}$, and the normalized plausibility weight $\bar{w}_{jk}$ using Equation (1.37);

5. Compute the covariance matrix of the Gaussian proposal distribution, $\Sigma_j$, using Equation (1.38);

6. Use the Metropolis-Hastings algorithm to generate $N_s$ samples from the $j + 1$th intermediate distribution $P_{j+1}(\boldsymbol{\vartheta})$. For $k = 1, \cdots, N_s$, do the following:

   • Select the $v$th Markov chain from the set $\{1, \cdots, N_s\}$ according to their probability equal to $\bar{w}_{jk}$ ($k = 1, \cdots, N_s$). Let the seed sample $\boldsymbol{\vartheta}^{\text{seed}} = \boldsymbol{\vartheta}_{jv}$ and compute the corresponding log-likelihood $\mathcal{L}^{\text{seed}} = \mathcal{L}(\boldsymbol{D}|\boldsymbol{\vartheta}_{jv})$;

   • Generate a candidate sample $\boldsymbol{\vartheta}^{\text{cand}}$ from the Gaussian proposal distribution and compute $\mathcal{L}^{\text{cand}} = \mathcal{L}(\boldsymbol{D}|\boldsymbol{\vartheta}^{\text{cand}})$;

   • Generate a sample $r$ from a uniform distribution on $[0, 1]$;

   • If $r \leq \min\left(1, \exp\left[q_{j+1}(\mathcal{L}^{\text{cand}} - \mathcal{L}^{\text{seed}})\frac{P(\boldsymbol{\vartheta}^{\text{cand}})}{P(\boldsymbol{\vartheta}^{\text{seed}})}\right]\right)$, let $\boldsymbol{\vartheta}_{j+1k} = \boldsymbol{\vartheta}^{\text{cand}}$. Otherwise, let $\boldsymbol{\vartheta}_{j+1k} = \boldsymbol{\vartheta}^{\text{seed}}$;

7. If $q_{j+1} < 1$, let $j = j + 1$ and return to Step 3. Otherwise, let $m = j + 1$ and end the algorithm.

### *Bayesian updating with structural reliability methods*

More recently, another class of methods to sample from the posterior distribution, called Bayesian updating with structural reliability methods (BUS) has been proposed by Straub and Papaioannou (2015). BUS converts the Bayesian updating problem into an equivalent reliability problem. In structural reliability analysis, estimation of probabilities of rare events (i.e., failure probabilities) are of one's main concern. By interpreting estimation of the posterior distribution

as rare event estimation, existing structural reliability methods can be utilized to perform the Bayesian model updating. The principle behind BUS is to add an auxiliary random variable that is uniformly distributed, $\pi \in \mathcal{D}_\pi = [0, 1]$, to the space of the calibrating parameters, $\mathcal{D}_\vartheta$. The Bayesian updating problem is then expressed as a reliability problem in the augmented space $\mathcal{D}_\vartheta \times \mathcal{D}_\pi$. The performance function $g(\vartheta, \pi)$ and the failure domain $Z$ of this reliability problem is defined as:

$$g(\vartheta, \pi) = \pi - c \cdot P_L(\boldsymbol{D}|\vartheta) \tag{1.40}$$

$$Z = \{\pi \le c \cdot P_L(\boldsymbol{D}|\vartheta)\} \tag{1.41}$$

where $c$ is chosen such that $c \cdot P_L(\boldsymbol{D}|\vartheta) \le 1$ is held for all $\vartheta$. In this context, the quantity $c \cdot P_L(\boldsymbol{D}|\vartheta)$ is expressed as (Betz et al., 2018):

$$c \cdot P_L(\boldsymbol{D}|\vartheta) = \int_{\pi \in Z} d\pi \tag{1.42}$$

Consequently, Equation (1.22) is expressed as:

$$P(\vartheta|\boldsymbol{D}) \propto P_L(\boldsymbol{D}|\vartheta)P(\vartheta) = c^{-1} \int_{\pi \in Z} P(\vartheta)d\pi \tag{1.43}$$

The integral $\int_{\pi \in Z} P(\vartheta)d\pi$ can be estimated by sampling in the failure domain, according to the prior distribution $P(\vartheta)$. As such, sampling from the posterior is converted to sampling from the failure domain of the aforementioned reliability problem. Furthermore, an estimate for the evidence $P(\boldsymbol{D})$ is obtained as a by-product of BUS (Straub and Papaioannou, 2015):

$$P(\boldsymbol{D}) = c^{-1} \int_{\mathcal{D}_\vartheta} \int_{\pi \in Z} P(\vartheta)d\pi d\vartheta = c^{-1}p_f \tag{1.44}$$

where $p_f$ denotes the failure probability.

The principle behind the choice of $c^{-1}$ is that it should not be smaller than the maximum value of the likelihood function while conservatively large $c^{-1}$ decreases the efficiency of the method. Naturally, its optimal choice would be $c^{-1} = \max P_L(\boldsymbol{D}|\vartheta)$. Nevertheless, the maximum value of the likelihood is not always known in advance, and in such cases, it is difficult to determine $c^{-1}$ appropriately. Two different strategies that avoid determining $c$ have been recently proposed. In DiazDelaO et al. (2017), the equivalent reliability problem is redefined such that the performance function depends on the likelihood function but not $c$. In Betz et al. (2018), an extended variant of BUS, termed aBUS (adaptive

BUS), is developed, in which the selection of $c$ is not required a priori while appropriate values for $c^{-1}$ are adaptively

selected during the simulation. Note that, in this thesis, the distance-based approximate likelihood functions are used;

thus, the appropriate value for $c$ is straightforwardly available (see, Figure 1.5).

Any structural reliability method can be used in the BUS framework. The simplest application of BUS is rejection

sampling that is equivalent to crude Monte Carlo simulation (MCS), however, it becomes inefficient when the number

of the calibrating parameters is large. Therefore, the combination of BUS with subset simulation (Au and Beck, 2001)

would be of particular interest (see, e.g., Straub and Papaioannou, 2015; Betz et al., 2018), since subset simulation is

efficient in estimating small failure probabilities which may arise in BUS and does not depend on the dimension of the

parameters. Subset simulation expresses the failure domain $Z$ by means of a sequence of intermediate nested domains

$\{Z_j : j = 1, \cdots, m\}$ which hold $Z_1 \supset \cdots \supset Z_{m-1} \supset Z_m = Z$. The intermediate domain is defined as $Z_j = \{g(\boldsymbol{\vartheta}, \pi) \leq b_j\}$,

where $b_j$ is the threshold which maintains $b_1 > \cdots > b_{m-1} > b_m = 0$. As such, the failure probability can be expressed

as the product $p_f = \prod_{j=1}^{m} p_j$ of relatively large conditional probabilities $p_j = \mathcal{P}(\{[\boldsymbol{\vartheta}_{j-1k}, \pi_{j-1k}] : k = 1, \cdots, N_s\} \in Z_j)$,

where $[\boldsymbol{\vartheta}_{0k}, \pi_{0k}]$ are initial samples which are populated in all the space of the parameters. The value of $b_j$ is adaptively

determined such that each conditional probability $p_j$ corresponds to a pre-defined value $p_0$. Samples conditional on $Z_j$,

$\{[\boldsymbol{\vartheta}_{jk}, \pi_{jk}] : k = 1, \cdots, N_s\}$, can be generated by MCMC sampling, in which the seeds of the Markov chains are selected

as samples at the $j - 1$th step which hold $g(\boldsymbol{\vartheta}_{j-1k}, \pi_{j-1k}) \leq b_j$, and the proposal distribution is typically chosen as an

independent Gaussian distribution centered at the seed sample and having the common variance $s_q$. A component-wise

variant of the Metropolis-Hastings algorithm (Au and Beck, 2001) has been broadly employed in the subset simulation

algorithms. Besides, the so-called conditional sampling has been recently proposed in Papaioannou et al. (2015), and

the similar algorithm has been proposed by independent researchers (Au and Patelli, 2016). It should be noted that, the

conditional sampling is capable of generating $N_s$ samples simultaneously per step, and hence its computation can be

easily parallelized.

An algorithm for the combination of BUS with subset simulation can be summarized as follows: It is noted that,

for convenience, the reliability problem is transformed to the underlying standard normal space by expressing $g(\boldsymbol{\vartheta}, \pi)$

equivalently as $g(\boldsymbol{\vartheta}, \pi) = H(\boldsymbol{u}) = \Phi(u_1) - c \cdot P_L(\boldsymbol{y}|T(\boldsymbol{u}^*))$, where $\boldsymbol{u}$ represents the $(n_\vartheta + 1)$-dimensional standard

normal random variables with $u_1 = \Phi^{-1}(\pi)$ and $\boldsymbol{u}^* = \mathrm{T}^{-1}(\boldsymbol{\vartheta})$; $\Phi(\cdot)$ is the standard normal CDF; $\mathrm{T}(\cdot)$ denotes one of

the classical transformations, e.g., the Rosenblatt transformation (Hohenbichler and Rackwitz, 1981) and the marginal

transformation by the Nataf model (Der Kiureghian and Liu, 1986).

1.  Initialize $N_s$, $c$, $p_0$, and $s_q$;

2.  Let $j = 0$. Drawn $N_s$ samples $\boldsymbol{u}_{jk}$ ($k = 1, \cdots, N_s$), according to the ($n_\vartheta + 1$)-variate standard normal distribution;

3.  For all samples $\boldsymbol{u}_{jk}$ ($k = 1, \cdots, N_s$), compute the performance function $H(\boldsymbol{u}_{jk})$;

4.  Define the intermediate domain $Z_{j+1} = \{H(\boldsymbol{u}) \le b_{j+1}\}$ by selecting $b_{j+1}$ as the $p_0$ percentile of $H(\boldsymbol{u}_{jk})$;

5.  Find conditional samples located in $Z_{j+1}$ as $\left\{\boldsymbol{u}_{jv}^{(Z_{j+1})}: v = 1, \cdots, \frac{N_s}{p_0}\right\} = \{\boldsymbol{u}_{jk}: v = 1, \cdots, N_s\} \in Z_{j+1}$;

6.  Use the conditional sampling to drawn $N_s$ conditional samples in $Z_{j+1}$. For $k = 1, \cdots, N_s$, do the following;

    - Choose the $v$th Markov chain from the set $\left\{1, \cdots, \frac{N_s}{p_0}\right\}$. Let the seed sample $\boldsymbol{u}^{\mathrm{seed}} = \boldsymbol{u}_{jv}^{(Z_{j+1})}$;

    - Generate a candidate sample $\boldsymbol{u}^{\mathrm{cand}}$ from the Gaussian proposal distribution and calculate the corresponding performance function value $H^{\mathrm{cand}} = H(\boldsymbol{u}^{\mathrm{cand}})$;

    - If $H^{\mathrm{cand}} \le b_{j+1}$, let $\boldsymbol{u}_{j+1k} = \boldsymbol{u}^{\mathrm{cand}}$. Otherwise, let $\boldsymbol{u}_{j+1k} = \boldsymbol{u}^{\mathrm{seed}}$.

7.  If $b_{j+1} > 0$, let $j = j + 1$ and return to Step 3. Otherwise, let $m = j + 1$ and obtain $\boldsymbol{\vartheta}_m = \mathrm{T}(\boldsymbol{u}_m^*)$.

Both the TMCMC and BUS methods have been proved to be very efficient in high-dimensional problems (Betz et al., 2016; Betz et al., 2018). This makes these methods particularly appropriate for stochastic model updating, where not the model parameters themselves but their hyper-parameters are considered as the calibrating parameters. In fact, the TMCMC method is employed to perform the aforementioned two-step ABC updating framework in Bi et al. (2019), where, it was reported, for the application to Subproblem A of the NASA UQ challenge 2014, the TMCMC algorithm was converged after total of $1.3 \times 10^4$ likelihood evaluations in the main step with the Bhattacharyya distance-based likelihood. However, because each computation of the Bhattacharyya distance utilized with a sample with $N_{\mathrm{sim}} = 10^3$, the procedure rendered totally $1.3 \times 10^7$ model evaluations. Such a significant number of required model evaluations might not be affordable in practical applications; thus, the development of a furthermore efficient inference method is necessary.

## 1.1.3   Uncertainty propagation

After the UM is calibrated based on available information on the system of interest, uncertainty propagation aims at propagating both types of uncertainty (i.e., aleatory and epistemic uncertainty) from the model parameters into the model predictions through the underlying numerical model, in order to quantify uncertainty in the model predictions. This propagation is typically attained based on the availability of efficient stochastic simulation methodologies.

Two main tasks of uncertainty propagation include the sensitivity analysis and reliability analysis. The sensitivity analysis aims at measuring the contribution of uncertainty in each parameter to uncertainty in the model predictions. The sensitivity indices can be used for ranking the importance of epistemic parameters, and are useful for planning the future data collection and/or model updating with the target of effectively reducing uncertainty in the model predictions. The available sensitivity analysis methods in the literature can be generally classified into two categories: the local and global sensitivity analysis method. One can refer to Wei et al. (2015) for comprehensive reviews and comparisons of the sensitivity analysis methods.

The reliability analysis, on the other hand, aims at estimating the failure probabilities, i.e., the probabilities that undesired failure events due to the intrinsic property of the system happen, and the failure event is typically expressed, using the performance function (also called the g-function), as $g \leq 0$. This implies that at least one aleatory parameter (i.e., the category II or IV parameter) needs to be involved in the g-function. Comparatively, the involvement of only the category III parameters (due to purely epistemic uncertainty) does not result in the intrinsic failure but makes the model predictions imprecise and might result in another type of the failure because of a poor or unsatisfactory design and management. The propagation of the category III parameters is typically involved the optimization algorithms (see, e.g., Jiang et al., 2013; Faes and Moens, 2020).

*Precise reliability analysis*

With only aleatory uncertainty being present (i.e., the involvement of only the category II parameters), the failure domain can be generally expressed as $Z = \{x: g(x) < 0\}$ with the aleatory parameters $x$ described by the probability models (i.e., the PDFs) $f_X(x)$. The failure probability is then expressed as:

$$p_f = \int_{x \in Z} f_X(x)dx = \int_{\mathbb{R}^{n_x}} I_Z(x)f_X(x)dx \tag{1.45}$$

where $I_Z(\cdot)$ denotes the indicator function of $Z$ that is formulated by $I_Z(x) = 1$ if $x \in Z$, and else $I_Z(x) = 0$. Plenty of methods found in the literature are applicable for estimating the failure probabilities according to the properties of the underlying performance function $g(x)$. These available methods can be typically categorized into three categories, i.e., the approximate analytical methods, stochastic simulation methods, and metamodel methods. Their combined methods are also investigated.

Approximate analytical methods include first-order reliability method (Hasofer and Lind, 1974) and second-order reliability method (Der Kiureghian and Stefano, 1991) and aim at analytically deriving the failure probability based on

statistical moments of the g-function values which are computed by, for example, the Taylor series expansion of the g-function around the so-called design point. This class of methods is extremely efficient but can be inaccurate when the performance function shows strong nonlinearity.

Stochastic simulation methods are, in principle, numerical integrations based on random sampling, and the most straightforward and simplest method is crude MCS. The MCS first generates a set of aleatory samples according to the distribution of interest, $f_X(x)$, and then obtain an estimator of the failure probability by the rate of the samples located in the failure domain. The MCS is well-known to be robust to the type and dimension of the performance function, but not suitable for estimating small failure probabilities (e.g., $p_f \leq 10^{-3}$) because the number of samples (i.e., the number of model evaluations) that is required to achieve a given accuracy is proportional to $1/p_f$. In general, advanced MCS methods thus aim at reducing the variance of estimators for efficiently estimating small failure probabilities while, at the same time, they lose application robustness. For instance, importance sampling (Au and Beck, 1999) aims to shift the distribution of interest towards the failure domain by the so-called importance sampling density, while its success is known to be dependent on the prior knowledge of the failure domain. Line sampling (Schuëller et al., 2004) searches the failure domain with a set of lines that are all parallel to the importance direction and is particularly efficient in the high-dimensional problems, while its performance is known to be highly dependent on the chosen importance direction. Comparatively to these methods, subset simulation (Au and Beck, 2001) is, as already mentioned in the last subsection, found to play a balance between the efficiency and robustness. In addition, the recently proposed conditional sampling techniques (Papaioannou et al., 2015; Au and Patelli, 2016) have a potential to further improve efficiency of the method in high-dimensional problems.

Metamodel (also called the surrogate model) methods aim at substituting the expensive-to-evaluate performance function $g$ with a surrogate $\hat{g}$, that is much faster to evaluate, in order to reduce the number of actual model evaluations. Various surrogates are available in the literature, including but not limited to the (linear or quadratic) response surfaces (Bucher and Bourgund, 1990), neural networks (Hurtado and Alvarez, 2001), support vector machines (Hurtado, 2004), kriging (Kaymaz, 2005) and polynomial chaos expansions (Blatman and Sudret, 2010). Regardless of the type of the surrogates, it is usually built by a set of training samples of the aleatory parameters and the corresponding performance function values. The training samples can be generated according to a design of experiment (DOE) using, for example, MC sampling or Latin hypercube sampling. However, conventional "static" DOEs require many samples in the DOE to establish a satisfactory surrogate for estimating small failure probabilities, because the samples are rarely generated in the vicinity of the failure surface.

***Adaptive DOE-based metamodels***

Comparatively, adaptive DOEs, that aim at enriching the training sample pool by dynamically adding samples in the vicinity of the failure surface, gain attention in recent years. Among various types of the metamodel based on the adaptive DOEs, adaptive Kriging is most widely used due to its significant accuracy and efficiency in estimating small failure probabilities. The well-known method in this class is AK-MCS (Echard et al., 2011), which combines Kriging and the adaptive DOE based on MC sampling, and a lot of its improved versions, such as AK-IS (Echard et al., 2013) and AK-SS (Huang et al., 2016), have been also proposed. The principles behind AK-MCS are two folds: (i) the method iteratively searches the optimal training sample to be added that results in the best improvement of the Kriging model; (ii) the method is automatically finished if the stopping criterion that ensure a given accuracy of the Kriging model is achieved. These properties are both based on the so-called learning function (also called U-function):

$$U(\boldsymbol{x}) = \frac{\mu_{\hat{g}}(\boldsymbol{x})}{\sigma_{\hat{g}}(\boldsymbol{x})} \tag{1.46}$$

where $\mu_{\hat{g}}(\cdot)$ means the mean estimator of $g$ by the Kriging model and $\sigma_{\hat{g}}(\cdot)$ means similarly the variance estimator. In general, the sample in the DOE that returns the minimum U-function value is considered as the optimal training sample in the vicinity of the failure surface, and the stopping criterion is defined by $\min[U(\boldsymbol{x})] \geq 2$ which corresponds to the case that the signs of the g-function at all the samples in the DOE are correctly classified with a probability of $\Phi(2) = 0.977$. The AK-MCS methods enable to drastically reduce the number of actual model evaluations, however they rely on "static" sampling techniques and require many samples in the DOE (while the model is evaluated at a small number of these samples).

More recently, a new adaptive Kriging method, called AK-MCMC, has been proposed by Wei et al. (2019a). The adaptive DOE strategy in this method is based on MCMC sampling, where samples in the vicinity of the sequence of intermediate surfaces, which is gradually converged to the failure surface, are dynamically added to the training sample pool. Figure 1.8 demonstrates the principle illustration of the AK-MCMC procedure for the case when $m = 2$. For the aforementioned purpose, the U-function in Equation (1.46) is modified as:

$$U_j(\boldsymbol{x}) = \frac{\mu_{\hat{g}}(\boldsymbol{x}) - b_j}{\sigma_{\hat{g}}(\boldsymbol{x})}, \forall j = 1, 2, \cdots, m \tag{1.47}$$

AK-MCMC can be iterpreted as the combination of AK-MCS and subset simulation. The method enables to drastically reduce the number of required samples in the DOE by efficiently pushing the samples forward to the failure domain, and thus especially efficient in estimating very small failure probabiities (e.g., $p_f \leq 10^{-6}$).
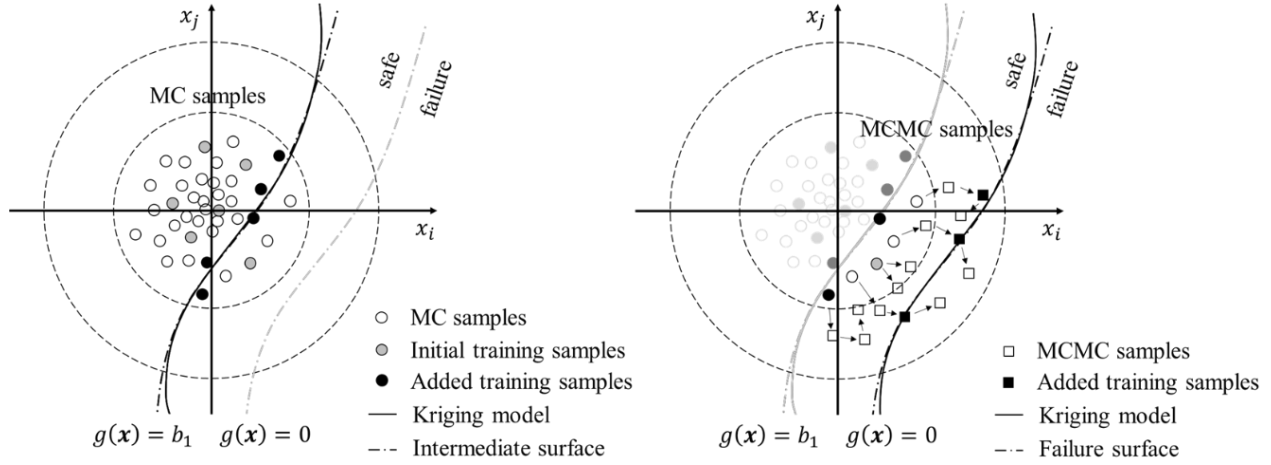
Figure 1.8 Illustration of AK-MCMC procedure.

*Imprecise reliability analysis*

   With both aleatory and epistemic uncertainty being present, on the other hand, the failure domain can be generally

represented as $Z = \{x, y: g(x, y) < 0\}$ with the aleatory parameters $x$ described by the imprecise probability models

(e.g., the p-boxes) $f_X(x, \theta)$ and the epistemic parameters $y$ described by the non-probabilistic models (e.g., the interval

models). The failure probability is then expressed as:

$$p_f(\theta, y) = \int_{x \in Z} f_X(x, \theta) dx = \int_{\mathbb{R}^{n_x}} I_Z(x, y) f_X(x) dx \qquad (1.48)$$

As such, the failure probability becomes a function of the epistemic parameters $y$ and $\theta$. In general, estimation of this

failure probability function involves a double-loop strategy to propagate aleatory and epistemic uncertainty separately.

Two distinct strategies can be developed while both aim at constructing the p-box of the g-function. In the first strategy,

the outer loop performs optimization in the epistemic space $\mathcal{D}_e$ to find two sets of the epistemic parameters that result

in the lower and upper bounds of g-function, while the inner loop propagates the aleatory parameters to the g-function,

conditioned on each pair of the epistemic parameters. The simplest method in this class is double-loop MCS (Rocchetta

et al., 2018b) but more sophisticated methods have been also developed, e.g., advanced line sampling (De Angelis et

al., 2015), which allows for using the same importance direction for the entire p-box analysis and reusing samples that

are generated in an inner loop for other iterations of the outer loop, significantly increasing the computational efficiency.

In addition, the AK-MCS technique is also employed for this starategy (Schöbi et al., 2017). In the second strategy, on

the other hand, the outer loop samples in the aleatory space $\mathcal{D}_a$ to retrieve the corresponding epistemic focal element,

while the inner loop performs optimization in each focal element to find the minimum and maximum g-function values.

The most well known method in this class is interval MCS (Zhang et al., 2010; Zhang et al., 2013). Furthremore, this strategy can be also combined with subset simulation (Alvarez et al., 2018) and the AK-MCS technique (Schöbi et al., 2017). Figure 1.9 shows difference between these double-loop strategies. It should be noted that, in the second strategy, the distributional p-box is converted into a distribution-free p-box having the same boundary CDFs.
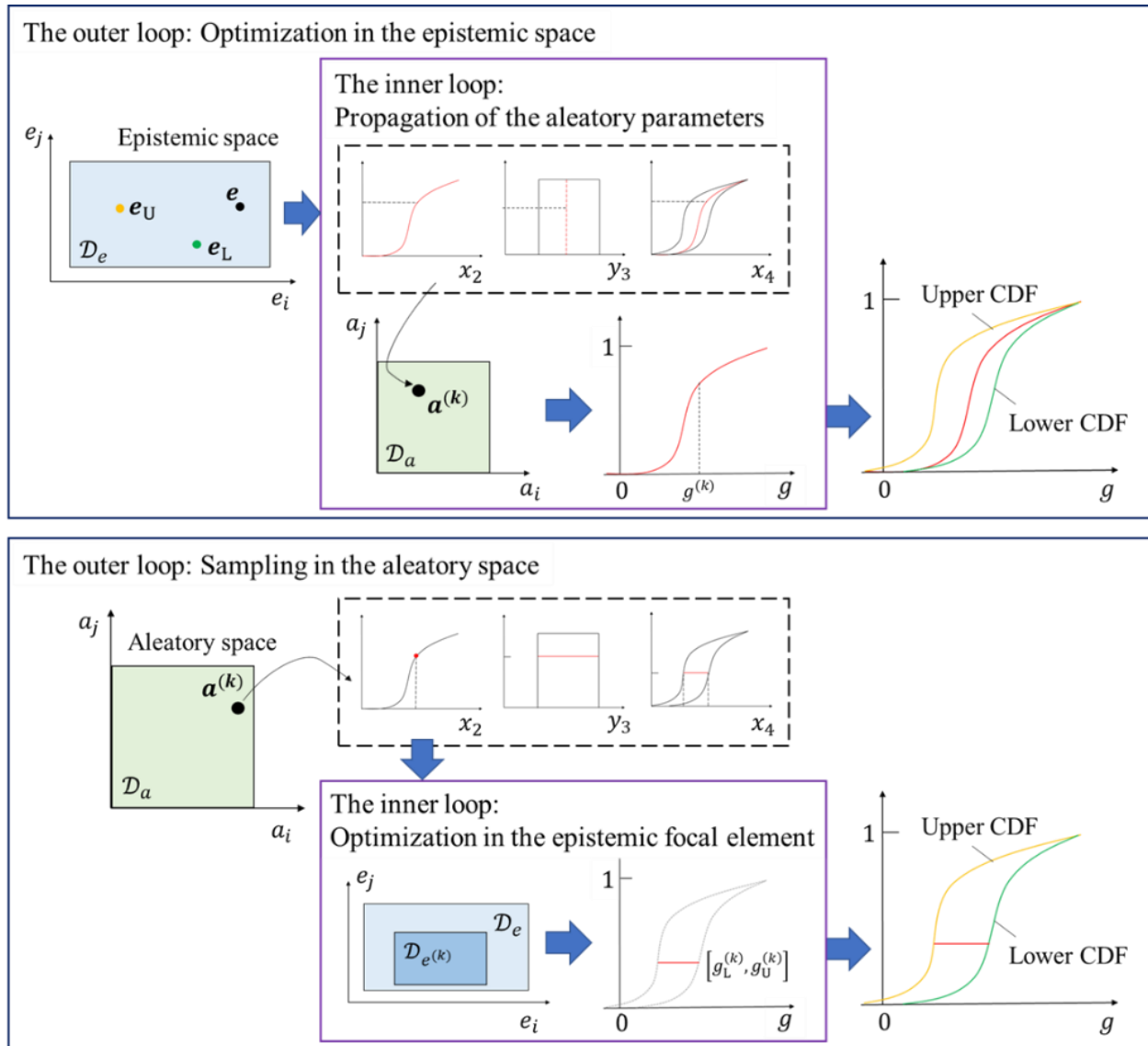


Figure 1.9 Illustration of two double-loop strategies.

### Non-intrusive imprecise stochastic simulation

Aimed at breaking the aforementioned double-loop, another class of methods for estimating the failure probability functions, called extended MCS (EMCS), has been developed by Wei et al. (2014), which is applicable for propagating distributional p-boxes. EMCS is based on importance sampling, where a single well-chosen realization of the p-box is

propagated and the obtained g-function values are reweighted to infer the failure probability function. The method is extremely efficient since only one stochastic simulation is necessary and thus its computational cost is identical to that of the conventional reliability analysis. In addition, the global version of the method has been also proposed to achieve a better global performance of the estimators in the epistemic space (Wei et al., 2019b). However, the variability of the EMCS estimators can be substantial when the dimension of the epistemic parameters is high.

More recently, the EMCS methods have been integrated with the high-dimensional model representation (HDMR) decompositions to compute the estimators in low-dimension and the sensitivity analyses to measure the importance of the epistemic parameters, so as to develop the so-called non-intrusive imprecise stochastic simulation (NISS) methods (Wei et al., 2019b; Wei et al., 2019c). Two types of NISS methods, i.e., the local and global NISS methods, have been developed based on the local and global versions of EMCS, respectively. The latter has been furthermore generalized for propagating distributional p-boxes and interval models simultaneously (Song et al., 2019). The generalized version of the global NISS method is employed and improved in this thesis to address the NASA UQ challenge 2019; thus, its brief review will be provided to close this section.

The global NISS method decomposes the failure probability function in Equation (1.48) by the random sampling (RS)-HDMR decomposition:

$$
p_f(\boldsymbol{\theta}, \boldsymbol{y}) = p_{f0} + \sum_{i=1}^{n_\theta} p_{f\theta_i}(\theta_i) + \sum_{i=1}^{n_y} p_{fy_i}(y_i) + \sum_{i=1}^{n_\theta-1}\sum_{j=i+1}^{n_\theta} p_{f\theta_{ij}}(\boldsymbol{\theta}_{ij}) + \sum_{i=1}^{n_y-1}\sum_{j=i+1}^{n_y} p_{fy_{ij}}(\boldsymbol{y}_{ij})
$$
$$
+ \sum_{i=1}^{n_\theta}\sum_{j=1}^{n_y} p_{f\theta_i y_j}(\theta_i, y_j) + \cdots + p_{f\boldsymbol{\theta}\boldsymbol{y}}(\boldsymbol{\theta}, \boldsymbol{y})
$$

(1.49)

where $\boldsymbol{\theta}_{ij} = [\theta_i, \theta_j]$ and $\boldsymbol{y}_{ij} = [y_i, y_j]$, and the functional components on its right-hand side can be expressed as:

$$
\begin{cases}
p_{f0} = \displaystyle\int p_f(\boldsymbol{\theta}, \boldsymbol{y}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f_Y(\boldsymbol{y}) d\boldsymbol{\theta} d\boldsymbol{y} \\
p_{f\theta_i}(\theta_i) = \displaystyle\int p_f(\boldsymbol{\theta}, \boldsymbol{y}) f_{\boldsymbol{\theta}_{-i}}(\boldsymbol{\theta}_{-i}) f_Y(\boldsymbol{y}) d\boldsymbol{\theta}_{-i} d\boldsymbol{y} - p_{f0} \\
p_{fy_i}(y_i) = \displaystyle\int p_f(\boldsymbol{\theta}, \boldsymbol{y}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f_{Y_{-i}}(\boldsymbol{y}_{-i}) d\boldsymbol{\theta} d\boldsymbol{y}_{-i} - p_{f0} \\
p_{f\theta_{ij}}(\boldsymbol{\theta}_{ij}) = \displaystyle\int p_f(\boldsymbol{\theta}, \boldsymbol{y}) f_{\boldsymbol{\theta}_{-ij}}(\boldsymbol{\theta}_{-ij}) f_Y(\boldsymbol{y}) d\boldsymbol{\theta}_{-ij} d\boldsymbol{y} - p_{f\theta_i}(\theta_i) - p_{f\theta_j}(\theta_j) - p_{f0} \\
p_{fy_{ij}}(\boldsymbol{y}_{ij}) = \displaystyle\int p_f(\boldsymbol{\theta}, \boldsymbol{y}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f_{Y_{-ij}}(\boldsymbol{y}_{-ij}) d\boldsymbol{\theta} d\boldsymbol{y}_{-ij} - p_{fy_i}(y_i) - p_{fy_j}(y_j) - p_{f0} \\
p_{f\theta_i y_j}(\theta_i, y_j) = \displaystyle\int p_f(\boldsymbol{\theta}, \boldsymbol{y}) f_{\boldsymbol{\theta}_{-i}}(\boldsymbol{\theta}_{-i}) f_{Y_{-j}}(\boldsymbol{y}_{-j}) d\boldsymbol{\theta}_{-i} d\boldsymbol{y}_{-j} - p_{f\theta_i}(\theta_i) - p_{fy_j}(y_j) - p_{f0}
\end{cases}
$$

(1.50)

where $\boldsymbol{\theta}_{-i} = [\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_{n_\theta}]$ and $\boldsymbol{\theta}_{-ij} = [\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_{j-1}, \theta_{j+1}, \cdots, \theta_{n_\theta}]$, and $\boldsymbol{y}_{-i}$ and $\boldsymbol{y}_{-ij}$ have

the similar structures as $\boldsymbol{\theta}_{-i}$ and $\boldsymbol{\theta}_{-ij}$ for $\boldsymbol{y}$. As such, the probability distributions of $\boldsymbol{\theta}$ and $\boldsymbol{y}$ are necessary to perform

the above integrations while these parameters are supposed to be interval parameters. This gap is usually addressed by

assigning purely instrumental PDFs to $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $f_Y(\boldsymbol{y})$ and the common choice for the instrumental PDF is an uniform

distribution within the interval of interest.

The principle behind the global NISS method is to first perform a stochastic simulation for estimating the constant

component $p_{f0}$, and reuse the same samples to estimate the remaining component functions in Equation (1.50) by the

reweighting procedure. Note that, in Equation (1.50), only the first- and second-order component functions are derived

while the higher-order component functions can be similarly derived. Nevertheless, it has been demonstrated that very

often the second-order HDMR decomposition provides a satisfactory approximation of the failure probability function

(see, e.g., Wei et al., 2019b; Wei et al., 2019c). Hence, the following NISS estimators are only derived in low-dimension,

which leads to a better estimation of the failure probability function in the high-dimensional epistemic space compared

with EMCS. The NISS estimators of Equation (1.50) can be expressed as:

$$
\begin{cases}
\hat{p}_{f0} = \dfrac{1}{N_s} \displaystyle\sum_{k=1}^{N_s} I_Z(\boldsymbol{x}_k, \boldsymbol{y}_k) \\
\hat{p}_{f\theta_i}(\theta_i) = \hat{p}_{f0} r_{\theta_i}(\boldsymbol{x}_k | \theta_i, \boldsymbol{\theta}_{-i,k}) \\
\hat{p}_{fy_i}(y_i) = \hat{p}_{f0} r_{y_i}(y_i | Z) \\
\hat{p}_{f\theta_{ij}}(\boldsymbol{\theta}_{ij}) = \hat{p}_{f0} r_{\theta_{ij}}(\boldsymbol{x}_k | \boldsymbol{\theta}_{ij}, \boldsymbol{\theta}_{-ij,k}) \\
\hat{p}_{f\boldsymbol{y}_{ij}}(\boldsymbol{y}_{ij}) = \hat{p}_{f0} r_{\boldsymbol{y}_{ij}}(\boldsymbol{y}_{ij} | Z) \\
\hat{p}_{f\theta_i y_j}(\theta_i, y_j) = [\hat{p}_{f\theta_i}(\theta_i) + \hat{p}_{f0}] r_{y_j}(y_j | Z) - \hat{p}_{fy_j}(y_j)
\end{cases}
\tag{1.51}
$$

with

$$
\begin{cases}
r_{\theta_i}(\boldsymbol{x}_k | \theta_i, \boldsymbol{\theta}_k) = \dfrac{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \theta_i, \boldsymbol{\theta}_{-i,k})}{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \boldsymbol{\theta}_k)} - 1 \\[2mm]
r_{y_i}(y_i | Z) = \dfrac{f_{Y_i}(y_i | Z)}{f_{Y_i}(y_i)} - 1 \\[2mm]
r_{\theta_{ij}}(\boldsymbol{x}_k | \boldsymbol{\theta}_{ij}, \boldsymbol{\theta}_k) = \dfrac{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \boldsymbol{\theta}_{ij}, \boldsymbol{\theta}_{-ij,k})}{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \boldsymbol{\theta}_k)} - \dfrac{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \theta_i, \boldsymbol{\theta}_{-i,k})}{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \boldsymbol{\theta}_k)} - \dfrac{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \theta_j, \boldsymbol{\theta}_{-j,k})}{f_{\boldsymbol{X}}(\boldsymbol{x}_k | \boldsymbol{\theta}_k)} + 1 \\[2mm]
r_{\boldsymbol{y}_{ij}}(\boldsymbol{y}_{ij} | Z) = \dfrac{f_{\boldsymbol{Y}_{ij}}(\boldsymbol{y}_{ij} | Z)}{f_{\boldsymbol{Y}_{ij}}(\boldsymbol{y}_{ij})} - \dfrac{f_{Y_i}(y_i | Z)}{f_{Y_i}(y_i)} - \dfrac{f_{Y_j}(y_j | Z)}{f_{Y_j}(y_j)}
\end{cases}
\tag{1.52}
$$

where $S = \{[\boldsymbol{x}_k, \boldsymbol{\theta}_k, \boldsymbol{y}_k] : k = 1, \cdots, N_s\}$ denotes $N_s$ sets of samples generated from the joint PDF $f_{\boldsymbol{X}}(\boldsymbol{x} | \boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) f_Y(\boldsymbol{y})$,

$f_{Y_i}(y_i | Z)$ and $f_{\boldsymbol{Y}_{ij}}(\boldsymbol{y}_{ij} | Z)$ are the conditional PDFs of $y_i$ and $\boldsymbol{y}_{ij}$, respectively, on the failure domain $Z$. The conditional

PDFs $f_{Y_i}(y_i | Z)$ and $f_{\boldsymbol{Y}_{ij}}(\boldsymbol{y}_{ij} | Z)$ can be estimated by KDE using the sample sets $S$. Theoretically, any kind of stochastic

simulation methods is applicable to the above framework, e.g., subset simulation (Wei et al., 2019c) and line simulation

(Song et al., 2020). In addition, the framework can be also combined with the metamodel methods (Wei et al., 2019c; Song et al., 2020).

The sensitivity indices can be then estimated as a by-product for each component function as:

$$
\begin{cases}
S_{\theta_i} = \dfrac{V\left[p_{f\theta_i}(\theta_i)\right]}{V\left[p_f(\boldsymbol{\theta}, \boldsymbol{y})\right]} \\[3mm]
S_{\boldsymbol{\theta}_{ij}} = \dfrac{V\left[p_{f\boldsymbol{\theta}_{ij}}(\boldsymbol{\theta}_{ij})\right]}{V\left[p_f(\boldsymbol{\theta}, \boldsymbol{y})\right]} \\[3mm]
S_{\theta_i y_j} = \dfrac{V\left[p_{f\theta_i y_j}(\theta_i, y_j)\right]}{V\left[p_f(\boldsymbol{\theta}, \boldsymbol{y})\right]}
\end{cases}
\tag{1.53}
$$

where $V[\cdot]$ indicates the variance operator, and $S_{y_i}$ and $S_{\boldsymbol{y}_{ij}}$ can be similarly defined. The above sensitivity indices are known as the Sobol' sensitivity indices (Sobol' et al., 2007). The sensitivity indices measure the relative importance of each epistemic parameter to the failure probability function, and thus can be used to identify the component functions that are negligible in estimating the failure probability function. Furthermore, they can also serve as a measure of the truncation error due to e.g., a HDMR decomposition with the second-order truncation. As such, the global NISS method can properly address numerical estimation errors.

The above generalized global NISS method provides a fascinating general framework for uncertainty propagation of all the category II, III, and IV parameters simultaneously. However, the open problem still exist. In the current NISS framework, the category IV parameters are restricted only to the parameterized imprecise probability models (e.g., the distributional p-boxes) that impose constraints on admissible distribution functions by defining a specific distribution family. As already mentioned in the last subsection, it is often the case that the distribution families cannot be properly determined beforehand due to the scarce and/or incomplete available data for the parameters. With such condition, the distribution families themselves should be treated as epistemic uncertainty to be propagated to the failure probability function; thus, the NISS method is necessary to be further improved so as to propagate any distribution family enclosed within the concerned p-boxes.

## 1.2   Aims and objectives

As can be shown in the previous section, the state-of-the-art developmentes in the field of imprecise probabilities have strengthened the subjective assumption-free probabilistic frameworks for uncertainty calibration and propagation, however, they still rely on the proper hypotheses on the underlying distribution families. Consequently, the aim of this

thesis is on the distribution-free uncertainty calibration and propagation under the presence of the hybrid uncertainties, and the focus will be on the approximate Bayesian computation (ABC) with the Bhattacharyya distance for uncertainty calibration and on the generalized version of the non-intrusive imprecise stochastic simulation (NISS) for uncertainty propagation. Herein, "distribution-free" indicates no limiting hypotheses on the distribution families are necessary for calibrating and propagating the category IV parameters. It is noted that this is not equivalent to the case that the category IV parameters are represented by non-parameterized imprecise probability models (e.g., the distribution-free p-boxes). In fact, it is still assumed in this thesis that the category IV parameters are characterized by the distributional p-boxes to calibrate them by reducing the hyper-parameters' epistemic space and to propagate them with respect to each hyper-parameter. Instead, a novel class of the distributional p-boxes based on staircase density functions (Crespo et al., 2018) is investigated. Staircase density functions are four parameter distributions that enable to flexibly approximate a wide range of distributions arbitrary close by selecting a proper quadruple of the hyper-parameters. As a consequence, hybrid uncertainties can be characterized by the hierarchical structure, in which the outer loop interval models of the hyper-parameters representing epistemic uncertainty while the inner loop probability distributions with the arbitrary formats characterizing aleatory uncertainty. Besides to the category IV parameters, it is assumted in this thesis that the category III parameters are represented by the interval models.

The objective of this thesis is first to improve the two-step ABC updating framework with both the Euclidian and Bhattacharyya distances in order to extend the scope of application to the case that the prior knowledge on the aleatory parameters is extremely limited as well as to the case that the system of interest is described by time-domain sequences. In addition, a highly efficient Bayesian inference algorithm is also developed to improve the computational efficiency of the framework. The second objective is to similarly improve the generalized NISS framework in order to extend the scope of application to the case that the distribution families of the aleatory parameters cannot be assumed. There are consequently five specific objectives:

(1) Develop a dimension reduction procedure for the time-domain sequences to deal with the curse of dimensionality in the evaluation of the Bhattacharyya distance;

(2) Develop a highly efficient Bayesian inference algorithm by combining BUS and AK-MCMC;

(3) Implement the staircase density functions into the two-step ABC updating framework to calibrate the probabilistic distributions of the category IV parameters whose distribution families are unknown;

(4) Further generalize the ABC updating framework with the Bhattacharyya distance to calibrate the joint probabilistic distribution by combining the staircase density functions and Gaussian copula function;

(5)  Implement the staircase density functions into the generalized NISS framework so as to propagate the category IV

parameters whose distribution families are unknown.

## 1.3   Original contributions

The contributions of this thesis are mainly on the development of distribution-free methodological frameworks

for stochastic model updating and uncertainty propagation of dynamic black-box systems under hybrid uncertainties.

This development furthermore leads to providing a solution to the NASA UQ challenge problem 2019. The five main

contributions are summarized as follows.

First, the two-step ABC updating framework with both the Euclidian and Bhattacharyya distances is extended to

the calibration of dynamic systems described by time-domain sequences. A dimension reduction procedure is proposed,

where the high-dimensional discrete time signals are evenly divided by a given window length and are degraded to a

series of scalars by taking root mean square (RMS) of the time signals within each window. The Bhattacharyya distance

is then evaluated for each RMS value and an approximate likelihood is defined by employing their RMS value as the

summary statistic. A highly efficient Bayesian inference algorithm is also developed based on the combination of BUS

and AK-MCMC. This algorithm is rooted in the conventional combination of BUS and subset simulation, however the

model evaluations at most of the samples are substituted for the Kriging surrogate. Since the adaptive DOE is designed

based on subset simulation, the algorithm can properly address very small failure probabilities which may arise in BUS.

Moreover, the use of common random numbers is also proposed to address the stochastic nature of the Bhattacharyya

distance in the construction of the Kriging surrogate for the Bhattacharyya distance-based performance function.

Second, a distribution-free stochastic model updating framework is developed by combining the staircase density

functions and Bhattacharyya distance to calibrate the category IV parameters whose distribution families are unknown.

The probability distributions of the category IV parameters are described by the staircase density functions. The prior

distributions of their hyper-parameters are derived based on the moment constraints and updated by the two-step ABC

updating procedure. The preliminary step with the Euclidian distance is served as a preconditioner to avoid non-unique

solutions while the main step with the Bhattacharyya distance quantifies the comprehensive uncertainty characteristics

of the observed features. By altering the problem setting of the model updating subproblem in the NASA UQ challenge

2014, the proposed model updating framework is demonstrated to be capable of calibrating the probability distributions

arbitrary close to the target distributions without limiting hypotheses on the distribution families.

Then, a solution to the latest edition (2019) of the NASA UQ challenge problem is provided. Four key techniques

are proposed to tackle the challenge: (i) a distribution-free Bayesian model updating framework for the calibration of the uncertainty model; (ii) an adaptive pinching approach to analyze and rank the relative sensitivity of the epistemic parameters; (iii) the probability bounds analysis to estimate failure probabilities; and (iv) a NISS approach to identify optimal design parameters. It should be noted that, among the above four techniques, the author mainly contributes to the techniques (i) and (iv). For calibrating the dynamic black-box system described in time-domain, the aforementioned distribution-free stochastic updating framework is further combined with the dimension reduction procedure developed in the first contribution. The NISS approach is developed for the reliability-based optimization subproblem, where the high-dimensional optimization is simplified to a set of one-dimensional searches by a first-order HDMR decomposition with respect to the design parameters. It is demonstrated that the approach can efficiently estimate the locally optimal design parameters.

Further, the above distribution-free stochastic model updating framework is generalized for calibrating the joint probabilistic distribution of the multivariate parameters accounting for the correlation structure. The staircase density functions are defined for univariate random variables and cannot account for the parameter dependencies. Hence, it is proposed to describe the joint probability distribution by a Gaussian copula function whose marginals are modeled by the staircase density functions. The prior distribution of the hyper-parameters of the staircase density functions and the correlation coefficients of the Gaussian copula is derived from the moment and correlation coefficient constraints, and are updated by the ABC updating procedure with the Bhattacharyya distance. This development reveals the importance of considering the correlation structure in stochastic model updating.

Finally, a distribution-free uncertainty propagation framework under the hybrid uncertainties is also proposed by incorporating the staircase density functions into the NISS framework in order to efficiently propagate the category IV parameters with no limiting assumptions on the distribution families. A hybrid NISS method is proposed, in which the local NISS method is used for propagating the p-boxes described by the staircase density functions while the global NISS method is used for propagating the interval models. By utilizing the proposed procedure, the reliability analysis subproblem in the NASA UQ challenge 2019 is solved with a satisfactory accuracy and efficiency compared with the aforementioned probability bounds analysis approach.

The above five original contributions are integrated as a distribution-free uncertainty calibration and propagation framework under hybrid uncertainties, and are successfully improved the scope of application and efficiency of hybrid uncertainty calibration and propagation in particular for the case that specific distribution families cannot be assumed due to the very scarce and/or incomplete available data for the parameters.

## 1.4   Structure of the thesis

This thesis is composed of seven chapters including five journal articles. Each article is aimed at demonstrating the five original contributions summarized in the previous section one by one. Specifically, this thesis is organized as follows.

Following this introduction, Chapter 2 introduces the first research article. In this article, the dimension reduction procedure aimed at evaluating the Bhattacharyya distance based on time-domain sequences and the Bayesian inference algorithm by the combination of BUS and AK-MCMC are developed for the two-step ABC updating framework with both the Euclidian and Bhattacharyya distances. The feasibility of the proposed approach is demonstrated on a seismic-isolated bridge pier model updating problem using simulated seismic response data. In addition, the proposed Bayesian inference algorithm is compared with the TMCMC algorithm to demonstrate its efficiency.

Chapter 3 brings up the second research article, in which the two-step ABC updating framework is improved to calibrate the category IV parameters whose distribution families are unknown by modeling the probability distributions to be calibrated as the staircase density functions. A well-known simple engineering example developed by Beck and Au (2002) is modified to demonstrate the principle of the proposed procedure. The procedure is further applied to the model updating subproblem in the NASA UQ challenge 2014 by modifying the problem setting and demonstrated to be capable of achieving satisfactory updating results without any prior knowledge on the distribution families of the parameters to be calibrated.

Chapter 4 presents the third research article which provides a solution to the NASA UQ challenge problem 2019. Five subproblems, including model calibration and UQ of the subsystem, uncertainty reduction, reliability analysis of baseline design, reliability-based design, and model update and design tuning, are successfully solved. These problems are aimed at representing the difficulties which are frequently encountered in the design of safety-critical systems based on the availability of very limited knowledge and data. The distribution-free Bayesian model updating framework with three key components, such as the staircase density functions for modeling the aleatory parameters, the Bhattacharyya distance for quantifying the uncertainty characteristics of the observations, and the dimension reduction procedure for evaluating the Bhattacharyya distance based on time-domain sequences, is demonstrated as a robust tool for uncertainty characterization and quantification throughout the challenge problem. Moreover, the NISS approach based on a HDMR decomposition of the failure probability function with respect to the design parameters is proposed so as to efficiently solve the reliability-based design optimization problem.

In Chapter 5, the fourth research article is brought up, which strengthens the stochastic model updating framework developed in the second and third articles to calibrate the joint probability distribution of multivariate parameters with the consideration of the correlation structure. For this purpose, the joint distribution is modeled by the Gaussian copula function with the marginals described by the staircase density functions, and then updating the correlation coefficients as well as the hyper-parameters to the posterior. The proposed method is first demonstrated on the engineering example used in the second article and another simple engineering example, and then applied to the seismic-isolated bridge pier model updating problem using simulated seismic response data.

Chapter 6 presents the fifth research article, in which the generalized NISS framework is improved to propagate the category IV parameters the distribution families on which are unknown by characterizing them as the distributional p-boxes based on the staircase density functions. To achieve a good balance between the efficiency in propagating the staircase density-based p-boxes and the accuracy in propagating the interval models for the epistemic parameters, the hybrid NISS method is proposed, where the local NISS method is performed for the staircase density-based p-boxes while the global NISS method is used for the interval models. The proposed uncertainty propagation approach is first demonstrated by solving the reliability analysis subproblem in the NASA UQ challenge 2019.

Chapter 7 draws concluding remarks from the developed uncertainty calibration and propagation frameworks and the application to the real-world engineering problems including the NASA UQ challenge problems. The thesis closes giving an outlook for future work in the field of hybrid uncertainty quantification.

# Chapter 2

# Research article 1: Bayesian model updating in time domain with metamodel-based reliability method

This is the first phase in the main part of this thesis, that aims at extending the two-step ABC updating framework with both the Euclidian and Bhattacharyya distances to the calibration of dynamic systems described by time-domain sequences. As has been mentioned in the introduction chapter of this thesis, the two-step ABC updating framework is originally developed for calibrating modal properties of the system of interest. In general, the time-domain sequences are more representative of dynamic systems than the modal properties, and they can provide more information of the systems including their nonlinear properties. For instance, the seismic-isolated bridge is typically designed to dissipate the earthquake energy at rubber bearings by plastically deforming them earlier than the other members, so as to reduce the seismic force on e.g., reinforced concrete (RC) piers, where the plastic deformation is undesirable. Therefore, it is quite important to quantify the uncertainty characteristics of the nonlinear parameters of the rubber bearings (e.g., the post-yield stiffness) using seismic response data described by time-domain sequences, resulting in reliable estimate of the seismic performance of the entire bridge system. Nevertheless, the direct application of the two-step ABC updating framework to the calibration of dynamic systems described by time-domain sequences is intractable mainly due to the following two reasons: (i) The original definition of the Bhattacharyya distance-based approximate likelihood is only applicable to scalar quantities because PMFs of the quantities of interest need to be estimated by a number of samples, and cannot be employed for time-domain sequences; (ii) The computational cost to analyze response time histories is generally higher than that to analyze modal properties, and thus the computation of the Bhattacharyya distance becomes intractable because of the necessity of hundreds to thousands repeated model evaluations for each computation of the Bhattacharyya distance during the updating procedure.

To cope with the first issue, a dimension reduction procedure is proposed, where high-dimensional discrete time histories are evenly divided by a given window length and are degraded to a series of scalar values by taking RMS of them within each window. A novel Bhattacharyya distance-based approximate likelihood for time-domain sequences is then defined by utilizing the RMS value of the Bhattacharyya distances evaluated for each scalar quantity above as

the selected summary statistic. Moreover, to address the second issue, a highly efficient Bayesian inference algorithm is developed by combining BUS and a state-of-the-art adaptive Kriging metamodeling technique, called AK-MCMC, which is in particular efficient in estimation of extremely small failure probabiities (typically less than $10^{-6}$). The use of common random number is also proposed in the main step of the ABC updating framework to address the variability of the training samples for establishing the Kriging surrogate due to the stochastic nature of the Bhattacharyya distance.

The proposed procedure is demonstrated on the model updating problem of a seismic-isolated bridge pier model, where the probabilistic distribution of three stiffness parameters including the post-yield stiffness of the rubber bearings are calibrated based on the multiple sets of the simulated seismic response data. This application demonstrates that the novel Bhattacharyya distance-based approximate likelihood is capable to capture wholly the uncertainty characteristics of the observed time sequences, and the combination of BUS with AK-MCMC enables to accurately infer the posterior distribution with a significantly reduced computational burden compared with the TMCMC algorithm.

# Bayesian model updating in time domain with metamodel-based reliability method

Masaru Kitahara [a, *], Sifeng Bi [b], Matteo Broggi [a], Michael Beer [a, c, d]

[a] *Leibniz Universität Hannover, Institute for Risk and Reliability, Callinstrasse 34, Hannover, Germany*

[b] *Beijing Institute of Technology, School of Aerospace Engineering, Beijing, China*

[c] *University of Liverpool, Institute for Risk and Uncertainty, Peach Street L69 7ZF, Liverpool, United Kingdom*

[d] *Tongji University, International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Shanghai 200092, China*

\* Correspondence author. E-mail address: masaru.kitahara@irz.uni-hannover.de (M. Kitahara).

**Abstract:** In this study, a two-step approximate Bayesian computation (ABC) updating framework using dynamic response data is developed. In this framework, the Euclidian and Bhattacharyya distances are utilized as uncertainty quantification (UQ) metrics to define approximate likelihood functions in the first and second steps, respectively. A novel Bayesian inference algorithm combining Bayesian updating with structural reliability methods (BUS) with the adaptive Kriging model is then proposed to effectively execute the ABC updating framework. The performance of the proposed procedure is demonstrated with a seismic-isolated bridge model updating application using simulated seismic response data. This application denotes that the Bhattacharyya distance is a powerful UQ metric with the capability to recreate wholly the distribution of target observations and the proposed procedure can provide satisfactory results with much-reduced computational demand compared with other well-known methods, such as transitional Markov chain Monte Carlo (TMCMC).

**Keywords:** Stochastic model updating; Bayesian model updating; Bhattacharyya distance; Adaptive Kriging; Metamodeling; Bayesian updating with structural reliability methods.

## 2.1 Introduction

Bayesian model updating using observed dynamic response data has a broad range of applications in a number of engineering fields (Beck and Katafygiotis, 1998; Katafygiotis and Beck, 1998; Cheung and Beck, 2009; Jensen et al., 2013; Rocchetta et al., 2018). In the campaign of model updating, uncertainties in both modelling and observation procedures should be appropriately considered; hence, uncertainty quantification (UQ) metrics are significant so as to comprehensively and quantitatively measure the stochastic discrepancy between model predictions and observations.

In the context of UQ, parameters can be categorized according to the involvement of aleatory or/and epistemic uncertainties as (Kennedy and O'Hagan, 2001; Crespo et al., 2014):

I)    Parameters without any uncertainties, appearing as explicit constants;

II)   Parameters with only aleatory uncertainty, appearing as random variables with fully determined probability characteristics such as density functions and distribution coefficients;

III)  Parameters with only epistemic uncertainty, appearing as unknown-but-fixed constants bounded by given intervals;

IV)   Parameters with both aleatory and epistemic uncertainties, appearing as imprecise random variables with only vaguely determined probability characteristics.

Both Categories III and IV parameters are considered in Bayesian model updating, whose target is not a single set of the crisp values of the parameters but a reduced space of epistemic uncertainty that is achieved based on the posterior distribution, such as reduced intervals of Category III parameters and reduced bounds of the cumulative probability function (CDF) of Category IV parameters.

The geometric discrepancy between model predictions and observations caused by Category III parameters can be quantified using the classical Euclidian distance as the UQ metric. On the other hand, quantifying the stochastic discrepancy caused by Category IV parameters requires a more comprehensive UQ metric that is capable of capturing a higher level of statistical information. The Bhattacharyya distance (Bhattacharyya, 1946) has been recently proposed as such a potential UQ metric (Bi et al., 2017). The Bhattacharyya distance provides a stochastic measure between two sample sets (i.e., the model predictions and observations) that quantifies the degree of the overlap of their probability distributions.

Bi et al. (2019) has developed a Bayesian model updating framework, in which the Bhattacharyya distance is employed as the UQ metric to define an approximate but efficient likelihood function based on the approximate Bayesian computation (ABC) method (Turner and Van Zandt, 2012; Safta et al., 2015). This framework has been demonstrated upon a three degree of freedom (DOF) spring-mass system example and shown to have a potential to recreate wholly the target observations. While the target outputs in this example are scalar modal responses, the direct computation of the Bhattacharyya distance becomes infeasible if the target outputs are described as high-dimensional dynamic responses because of the so-called curse of dimensionality. A dimension reduction procedure is thus proposed in this study to calculate the Bhattacharyya distance based on dynamic response data.

On the other hand, Markov chain Monte Carlo (MCMC) algorithms are generally accepted as the most attractive Bayesian inference techniques (Beck and Au, 2002; Cheung and Beck, 2009). Of particular importance among these algorithms is transitional Markov chain Monte Carlo (TMCMC) (Ching and Chen, 2007; Betz et al., 2016) and Bi et al. (2019) has also utilized TMCMC to perform the ABC updating framework. Although TMCMC is quite flexible and general, it requires a large number of model evaluations for calculating the likelihood function. In the ABC updating framework, the approximate likelihood function is defined based on the Bhattacharyya distance, the computation of which requires random samples of model predictions generated by Monte Carlo (MC) sampling. Thus, the total number of model evaluations is extremely large compared with general model updating and the computational burden becomes excessive in cases of time-consuming model evaluations, which are often involved in predicting dynamic responses.

Straub and Papaioannou (2015) has recently developed a formulation called Bayesian updating with structural reliability methods (BUS). The key idea of this formulation is to transform the Bayesian updating problem into an equivalent reliability problem, allowing to obtain samples from the posterior distribution as conditional samples that are located in the failure domain of this reliability problem. By employing Subset simulation techniques (Au and Beck, 2001), BUS has shown great efficiency in estimating posterior distributions (Betz et al., 2018). Its efficiency depends on the choice of the so-called likelihood multiplier. While the optimal multiplier ensuring the best acceptance rate is generally not available, it can be defined a priori for the proposed approximate likelihood function. Hence, BUS has a great potential to be effectively integrated with the ABC updating framework.

At the same time, BUS can further improve its efficiency by applying metamodeling techniques (Giovanis et al., 2017). Among various types of the metamodels, the adaptive Kriging model has been shown to be one of the most accurate and efficient methods in solving reliability problems (Echard et al., 2011; Echard et al., 2013; Huang et al., 2016). However, the failure probability associated with the equivalent reliability problem in BUS is generally known to be extremely small. In such rare events, the adaptive Kriging model becomes significantly inefficient because the number of candidate samples should be extremely large to ensure that enough samples are contained in the failure domain. On the other hand, Wei et al. (2019) has recently proposed a new algorithm called AK-MCMC, in which the Kriging model is adaptively trained on dynamically updated MCMC populations. This algorithm is especially suitable for extremely rare event problems. The objective of this study is consequently to develop an efficient ABC updating framework using dynamic response data by combining BUS and the AK-MCMC algorithm.

The structure of this paper is as follows. In Section 2.2, we describe the dimension reduction procedure to evaluate the Bhattacharyya distance for high-dimensional dynamic response data, and the proposed ABC updating framework.

Section 2.3 outlines the novel Bayesian inference algorithm combining BUS and the adaptive Kriging model based on the AK-MCMC algorithm. The principle and illustrative applications are then detailed in Section 2.4, using a model updating problem of a seismic-isolated bridge pier using simulated seismic response data. The computational efficiency of the proposed scheme is also presented by comparing with the results using TMCMC. Finally, some conclusions are given in Section 2.5.

## 2.2    Approximate Bayesian computation using dynamic response data

### 2.2.1    Formulations of the Bhattacharyya distance for dynamic response data

In the context of Bayesian model updating, the system under investigation can be expressed as:

$$\mathbf{y} = h(\mathbf{x}) \tag{2.1}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_n]$ denotes a vector of $n$ input parameters; $\mathbf{y} = [y_1, y_2, \cdots, y_m]$ is a vector of the output features as $m$-dimensional dynamic responses; $h(\cdot)$ is the simulator (e.g., finite element model). The uncertainties of the system are first characterized by the input parameters defined as various categories (refer to Section 2.1) and then propagated through the simulator into the uncertain output features. In general, randomly sampled values of the input parameters and output features are used in Bayesian model updating. Suppose the required sample size is $N_{\text{sim}}$, the simulator $h$ is executed $N_{\text{sim}}$ times to generate the sample set of the simulated features $\mathbf{Y}_{\text{sim}} \in \mathbb{R}^{N_{\text{sim}} \times m}$:

$$\mathbf{Y}_{\text{sim}} = \left[\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_{N_{\text{sim}}}\right]^T, \text{with } \mathbf{y}_i = [y_{1i}, y_{2i}, \cdots, y_{mi}], \forall i = 1, 2, \cdots, N_{\text{sim}} \tag{2.2}$$

In addition to the simulated features, observed features are required as the target of model updating. Suppose the number of observations is $N_{\text{obs}}$, the sample set of the observed features has a similar structure as Equation (2.2), where only the number of rows is changed: $\mathbf{Y}_{\text{obs}} \in \mathbb{R}^{N_{\text{obs}} \times m}$. The objective of Bayesian model updating can be then expressed as to minimize the stochastic discrepancy between $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$ by inferring the uncertainty characteristics (i.e., the probability distributions), of the input parameters.

In the following, possible UQ metrics are defined for capturing the discrepancy between $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$. The most classical Euclidian distance metric is expressed as:

$$d_E(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{sim}}) = \sqrt{(\overline{\mathbf{Y}}_{\text{obs}} - \overline{\mathbf{Y}}_{\text{sim}})(\overline{\mathbf{Y}}_{\text{obs}} - \overline{\mathbf{Y}}_{\text{sim}})^T} \tag{2.3}$$

where $\bar{\mathbf{Y}}_{(\cdot)}$ denotes a row vector of means of the features. The Euclidian distance is a point-to-point distance capable to capture the geometric discrepancy caused by Category III parameters. On the other hand, in the presence of Category IV parameters, it is desirable to employ a more comprehensive UQ metric that is capable to consider a higher level of statistical information from the sample sets.

The Bhattacharyya distance is herein proposed as such a stochastic metric for robustly measuring the degree of overlap between distributions of two sample sets. Its original definition is given as:

$$d_B(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{sim}}) = -\log\left[\int_{\mathbb{y}} \sqrt{p_{\text{obs}}(y)p_{\text{sim}}(y)}\,\mathrm{d}y\right] \tag{2.4}$$

where $p_{(\cdot)}(y)$ is the probability density function (PDF) of each feature sample; $\mathbb{y}$ is the $m$-dimensional feature space; $\int_{\mathbb{y}}(\cdot)\,\mathrm{d}y$ indicates the integration performed over the whole feature space. Differently from the Euclidian distance, the Bhattacharyya distance takes not only the means but also the variances, covariances, and even the distribution shapes, of the sample sets into account. However, the direct evaluation of Equation (2.4) is not feasible since precise estimation of the PDF is generally unavailable, especially for applications where observations are difficult or expensive. Bi et al. (2019) hence proposed the so-called binning algorithm to evaluate the probability mass function (PMF) of a discrete distribution, so that the discrete Bhattacharyya distance is used instead. The PMF is a function which maps the possible values of a discrete random variable to the probabilities of its occurrence (Grimmett and Stirzaker, 2001). The discrete Bhattacharyya distance is defined as (Patra et al., 2015):

$$d_B(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{sim}}) = -\log\left\{\sum_{i_m=1}^{n_{bin}} \cdots \sum_{i_1=1}^{n_{bin}} \sqrt{p_{\text{obs}}(b_{i_1,i_2,\cdots,i_m})p_{\text{sim}}(b_{i_1,i_2,\cdots,i_m})}\right\} \tag{2.5}$$

where $p_{(\cdot)}(b_{i_1,i_2,\cdots,i_m})$ is the PMF value of the bin $b_{i_1,i_2,\cdots,i_m}$. The bin has $m$ subscripts because it is generated under a $m$-dimensional joint PMF space. More detailed information of the binning algorithm can be referred to Bi et al. (2019).

In this study, the output features are assumed as very high-dimensional dynamic responses. In such circumstances, the direct evaluation of Equation (2.5) becomes infeasible since the total number of bins is exponentially increasing with the dimension $m$ because of the so-called curse of dimensionality. To overcome this issue, a dimension reduction procedure consisting of the following steps is herein proposed (Kitahara et al., 2020).

1) Define the window length $L$ and divide $\mathbf{y}_i$ of $\mathbf{Y}_{\text{sim}}, \forall i = 1, 2, \cdots, N_{\text{sim}}$ into $\lceil m/L \rceil$ intervals, where $\lceil \cdot \rceil$ indicates the upper integer of the investigated values. The same procedure is also applied to $\mathbf{Y}_{\text{obs}}$;

2) Compute the root mean square (RMS) values of the discrete time histories within each interval to obtain a matrix $\mathbf{R} = [R_1, R_2, \cdots, R_{[m/L]}]$ and generate the sample set of the RMS values $\mathbf{R}_{\mathbf{Y}_{\text{sim}}} \in \mathbb{R}^{N_{\text{sim}} \times [m/L]}$ :

$$\mathbf{R}_{\mathbf{Y}_{\text{sim}}} = \left[ \mathbf{R}_{\mathbf{Y}_{\text{sim}}}^1, \mathbf{R}_{\mathbf{Y}_{\text{sim}}}^2, \cdots, \mathbf{R}_{\mathbf{Y}_{\text{sim}}}^{[m/L]} \right], \text{with } \mathbf{R}_{\mathbf{Y}_{\text{sim}}}^j = \left[ R_j^1, R_j^2, \cdots, R_j^{N_{\text{sim}}} \right]^T, \forall j = 1, 2, \cdots, [m/L] \tag{2.6}$$

and $\mathbf{R}_{\mathbf{Y}_{\text{obs}}} \in \mathbb{R}^{N_{\text{obs}} \times [m/L]}$. It is noted that, $\mathbf{R}_{\mathbf{Y}_{\text{obs}}}$ has a similar structure as Equation (2.6), where only the number of rows is changed;

3) Evaluate the Bhattacharyya distance $d_{Bj}$ between two sample sets of the RMS values $\mathbf{R}_{\mathbf{Y}_{\text{obs}}}^j$ and $\mathbf{R}_{\mathbf{Y}_{\text{sim}}}^j, \forall j = 1, 2, \cdots, [m/L]$;

4) Obtain the RMS value of the Bhattacharyya distances and employ it as a UQ metric.

The principle behind the window length $L$ is that a smaller $L$ leads to employing more detailed information of the target dynamic response data, whereas it leads to a larger computational demand at the same time. It is found that $L = 0.025 \cdot m$ is a reasonable choice in this study. This choice corresponds to the case where each RMS contains 2.5 % of the overall target signals.

## 2.2.2 Approximate Bayesian computation

The ABC updating framework with the distance-based UQ metrics is summarized here. Bayesian model updating is based on the Bayes' theorem (Beck and Katafygiotis, 1998):

$$P(\mathbf{x}|\mathbf{Y}_{\text{obs}}) = \frac{P_L(\mathbf{Y}_{\text{obs}}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{Y}_{\text{obs}})} \tag{2.7}$$

where $P(\mathbf{x})$ indicates the prior distribution of $\mathbf{x}$, representing the initial knowledge about the parameters $\mathbf{x}$; $P(\mathbf{x}|\mathbf{Y}_{obs})$ indicates the posterior distribution of $\mathbf{x}$, which represents the updated knowledge about the parameters $\mathbf{x}$ based on the availability of the observed data; $P(\mathbf{Y}_{obs})$ is the normalized factor (also called the evidence) ensuring that the posterior distribution $P(\mathbf{x}|\mathbf{Y}_{obs})$ integrates to one; $P_L(\mathbf{Y}_{obs}|\mathbf{x})$ is the likelihood function of $\mathbf{Y}_{obs}$ for an instance of the parameters $\mathbf{x}$.

The likelihood function is one of the key components in Bayesian model updating, because it quantifies the degree of relevance of a model with a given instance of the parameters to be calibrated, by representing the possibility of the observations. Under the assumption of independence between each observation, the likelihood function in Equation (2.7) is theoretically defined as:

$$P_L(\mathbf{Y}_{\text{obs}}|\mathbf{x}) = \prod_{k=1}^{N_{\text{obs}}} P(\mathbf{Y}_k|\mathbf{x}) \tag{2.8}$$

where $P(\mathbf{Y}_k|\mathbf{x})$ indicates the PDF value of the $k$th observed data $\mathbf{Y}_k$ conditional to the corresponding instance of the parameters $\mathbf{x}$. Note that, the precise estimation of the PDF requires a large number of simulated features. Consequently, an analytical formula of the likelihood in Equation (2.8) demands a huge number of model evaluations and it can be almost infeasible for complex simulators.

The ABC method (Turner and Van Zandt, 2012; Safta et al., 2015) is utilized to overcome the above obstacle by replacing the full likelihood with an approximate but efficient function that contains information of the observations and the instance of the parameters $\mathbf{x}$. In the approximate likelihood, any types of statistics can be used to measure the stochastic discrepancy between model predictions and observations (Turner and Van Zandt, 2012); hence, it is natural to define it employing the distance metrics. Various functional formulas have been investigated in the literature for the ABC method, including the Gaussian (Turner and Van Zandt, 2012), sharp (Rocchetta et al., 2018), and Epanechnikov (Safta et al., 2015) functions. Nevertheless, the basic principle of the approximate likelihood is that it should return a high value when the distance metric is small, whereas it penalizes the $\mathbf{x}$ instance when its corresponding distance metric is large. In this study, an approximate likelihood function based on the Gaussian function is proposed as:

$$P_L(\mathbf{Y}_{\text{obs}}|\mathbf{x}) \propto exp\left\{-\frac{d^2}{\varepsilon^2}\right\} \tag{2.9}$$

where $d$ is the distance metric; $\varepsilon$ denotes the so-called width factor, which is a pre-defined coefficient controlling the centralization degree of the posterior distribution. Based on a series of tests in various applications, $\varepsilon$ is determined to lie into the interval $[10^{-3}, 10^{-1}]$ (Patelli et al., 2017). A smaller $\varepsilon$ corresponds to a more peaked posterior distribution which is more likely to converge to the true value but requires more computational demand for convergence.

By employing the Bhattacharyya distance, the proposed approximate likelihood function is capable of capturing comprehensive uncertainty information from both model predictions and observations. However, the Bhattacharyya distance in Equation (2.5) will be infinite if the initial $\mathbf{Y}_{\text{sim}}$ is too far from $\mathbf{Y}_{\text{obs}}$, i.e., there is no overlap between the two sample sets, and thus cannot be directly employed in the likelihood. Hence, Bi et al. (2019) has proposed the two-step ABC updating framework, where a preliminary step with the Euclidian distance-based likelihood is employed to force an overlap between $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$. The comprehensive uncertainty characteristics of the parameters are then calibrated in the main step with the Bhattacharyya distance-based likelihood. This two-step strategy is also employed

in this study. The reader can refer to Bi et al. (2019) for detailed information of the two-step procedure.

## 2.3    Bayesian updating with adaptive Kriging model

### 2.3.1    Bayesian updating with structural reliability methods (BUS)

In this section, the BUS formulation (Straub and Papaioannou, 2015; DiazDelaO et al., 2017) is briefly reviewed. The BUS formulation is based on the conventional rejection principle. Let $c$ denotes the so-called likelihood multiplier such that the following inequality holds for all the parameters x:

$$cP_L(\mathbf{Y}_{\text{obs}}|\mathbf{x}) \leq 1 \tag{2.10}$$

In this context, a sample distributed as the posterior distribution $P(\mathbf{x}|\mathbf{Y}_{\text{obs}}) \propto P_L(\mathbf{Y}_{\text{obs}}|\mathbf{x})P(\mathbf{x})$ in Equation (2.7) can be generated by the following rejection principle:

1) Generate $u$ uniformly distributed on $[0, 1]$ and $\mathbf{x}$ distributed as the prior distribution $P(\mathbf{x})$;

2) If $u < cP_L(\mathbf{Y}_{\text{obs}}|\mathbf{x})$, return $\mathbf{x}$ as a posterior sample. Otherwise, go back to Step 1).

Although the rejection sampling is theoretically viable, it will become inefficient with increasing the number of observations due to the large rejection rate. Hence, BUS transforms the Bayesian updating problem into an equivalent reliability problem so as to maintain the advantage of the rejection principle but have much higher efficiency. Consider a reliability problem with uncertain parameters $(\mathbf{x}, u)$ according to the joint PDF $P(\mathbf{x})I(0 \leq u \leq 1)$, where $I(\cdot)$ means the indicator function that is equal to one if its argument is true and zero otherwise. The limit state function and failure domain of this reliability problem can be defined as:

$$G = u - cP_L(\mathbf{Y}_{\text{obs}}|\mathbf{x}) \tag{2.11}$$

$$F = \{G < 0\} \tag{2.12}$$

The PDF of the failure sample $(\mathbf{x}', u')$ can be then obtained as:

$$p_{\mathbf{x}',u'}(\mathbf{x}, u) = p_F^{-1}P(\mathbf{x})I(0 \leq u \leq 1)I(u < cP_L(\mathbf{Y}_{\text{obs}}|\mathbf{x})) \tag{2.13}$$

where

$$p_F = \iint P(\mathbf{x})I(0 \leq u \leq 1)I(u < cP_L(\mathbf{Y}_{\text{obs}}|\mathbf{x}))dud\mathbf{x}$$

denotes the failure probability of the reliability problem. In this formulation, the PDF of the failure sample $p_{x',u'}(\mathbf{x}, u)$ and the failure probability $p_F$ both correspond the posterior distribution $P(\mathbf{x}|\mathbf{Y}_{\text{obs}})$ and normalized factor $P(\mathbf{Y}_{\text{obs}})$ in Equation (2.7), respectively. Consequently, the samples for deriving the posterior distribution can be generated as the conditional samples falling into the failure domain by existing reliability analysis methods, such as Subset simulation (Au and Beck, 2001).

A key component in BUS is the likelihood multiplier, because the acceptance rate in BUS is directly proportional to it. Therefore, it should be selected as large as possible along with satisfying the inequality in Equation (2.10) for all the parameters $\mathbf{x}$ and its optimal choice is defined as $c = [\max P_L(\mathbf{Y}_{\text{obs}}|\mathbf{x})]^{-1}$. While the optimal multiplier is generally unknown a priori, it can be employed as $c = 1$ for the proposed approximate likelihood function, since the approximate likelihood function is maximized when the distance metric is minimized to be zero. Therefore, BUS can be efficiently utilized as the Bayesian inference tool in the two-step ABC updating framework.

### 2.3.2 Adaptive Kriging-based BUS algorithm

BUS has shown great efficiency in estimating the posterior distribution by utilizing Subset simulation techniques (Betz et al., 2018). However, the failure probability of the equivalent reliability problem in BUS becomes significantly small and can reach $10^{-6}$ or even smaller with increasing the number of observations. In such very rare events, a large number of limit state function evaluations is required for estimating the failure probabilities even for Subset simulation. Furthermore, in the main step of the proposed updating framework, the limit state function involves the Bhattacharyya distance which is evaluated based on random samples of model predictions. Consequently, BUS with Subset simulation demands a huge number of model evaluations and it can be almost infeasible for complex simulators.

BUS can further improve its efficiency by employing metamodeling techniques (Giovanis et al., 2017). Among various types of the metamodels, the adaptive Kriging model has been paid significant attention as one of the most accurate and efficient methods in solving reliability problems. It can be interpreted as the classification method for the failure domain by the Kriging model, also known as the Gaussian process model. In this model, the estimated responses follow a Gaussian distribution with the Kriging means and Kriging variances. The basic rationales of the kriging model can be found in Echard et al. (2011).

The key idea of the adaptive Kriging model is to adaptively identify samples close to the limit state function from the candidate MC samples based on the Kriging means and Kriging variances. The well-trained Kriging model by those samples enables to provide a precise classification for the failure domain and thus the failure probability can be

efficiently estimated by this model. Nevertheless, the failure probabilities of the equivalent reliability problem in BUS can be significantly small. In such rare events, the adaptive Kriging model becomes very inefficient since the candidate sample pool should be enlarged to ensure that enough samples are contained in the failure domain.

Meanwhile, Wei et al. (2019) proposed a new algorithm called AK-MCMC. In this algorithm, the classification for a seres of intermediate failure domains $F_i = \{G < b_i\}$ is performed, where $b_i$ is the intermediate failure thresholds $(b_1 > b_2 > \cdots > b_m = 0)$. An illustration of a two-dimensional case following the AK-MCMC algorithm is provided in Figure 2.1. Figure 2.1(a) presents its initial step as the classification for the initial intermediate failure domain $F_1 = \{G < b_1\}$ upon MC samples given by the plots. The grey and black plots indicate the arbitrary selected initial training samples and the additional training samples adaptively selected based on the Kriging means and Kriging variances, respectively. In addition, the dashed and solid lines show the initial intermediate failure surface and the Kriging model trained by the above samples, respectively. On the other hand, Figure 2.1(b) illustrates the classification for the failure domain $F_m = \{G < b_m(= 0)\}$ upon MCMC samples given as the squared points. Note that, this figure corresponds to the case where $m = 2$. As same as Figure 2.1(a), the black plots indicate the adaptively selected training samples and the dashed and solid lines show the failure surface and the Kriging model trained by all the training samples. As shown in these figures, this algorithm provides the classifications for a sequences of intermediate failure domains, which will finally converge to the classification for the true failure domain, and is much more efficient than the direct classification for the failure domain. As a consequence, this algorithm enables to efficiently employed for extremely rare events and thus it has a great potential to be combined with BUS.
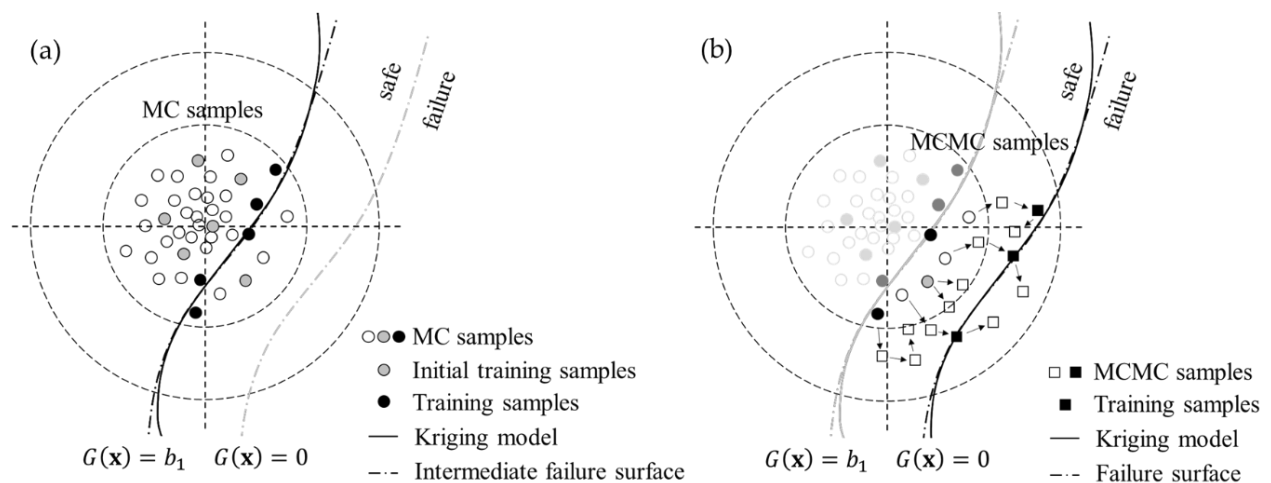


Figure 2.1 Illustration of the AK-MCMC algorithm: (a) Classification for the initial intermediate failure domain; (b) Classification for the failure domain.

In this study, a novel Bayesian inference algorithm is thus proposed by combining BUS with the adaptive Kriging model using the AK-MCMC algorithm. The flowchart of this algorithm is summarized in Figure 2.2 and the procedure is described in detail as below:

1) Let $i = 1$. Then, generate an $N$ MC samples population $\mathbf{W}_1$ of the parameters $(\mathbf{x}, u)$ according to the joint PDF $P(\mathbf{x})I(0 \leq u \leq 1)$;

2) Randomly select $N_0$ samples from $\mathbf{W}_1$ and evaluate the limit state function in Equation (2.11) on these samples. Attribute these $N_0$ samples to the training samples population $\mathbf{W}_t$;

3) Train or update the Kriging model $\hat{G}_i(\mathbf{x}, u)$ with $\mathbf{W}_t$;

4) Predict the limit state function value for each non-training sample contained in $\mathbf{W}_i$ using the Kriging model $\hat{G}_i(\mathbf{x}, u)$. Evaluate or update the intermediate failure threshold value $b_i$ based on the principle that $\lfloor p_0 N \rfloor$ samples in $\mathbf{W}_i$ is conditional on the intermediate failure domain $F_i$. Here, $p_0$ is the pre-defined target probability and $\lfloor \cdot \rfloor$ is the lower integer of the investigated values;

5) Compute the following learning function as:

$$U(\mathbf{x}, u) = |\mu_G(\mathbf{x}, u) - b_i|/\sigma_G(\mathbf{x}, u) \tag{2.14}$$

where $\mu_G(\mathbf{x}, u)$ indicates the Kriging mean and $\sigma_G(\mathbf{x}, u)$ is the Kriging standard deviation. If the stopping criterion as $\min(U(\mathbf{x}, u)) \geq 2$ is satisfied for all the $N$ samples, go to the next step. Otherwise, find the non-training sample in $\mathbf{W}_i$ that corresponds to the minimum value of the learning function in Equation (2.14) and evaluate the true limit state function. Attribute the sample to $\mathbf{W}_t$ and return to Step 3);

6) If $b_i \leq 0$, let $m = i$, save the Kriging model $\hat{G}_m(\mathbf{x}, u)$. Identify samples in $\mathbf{W}_m$ located into the failure domain $F$. Keep these samples as the seeds $\mathbf{W}_s$ and go to the next step. Otherwise, generate an $N$ MCMC samples population $\mathbf{W}_{i+1}$ of the parameters $(\mathbf{x}, u)$ conditional on the intermediate failure domain $F_i$ by calling the Kriging model $\hat{G}_i(\mathbf{x})$ by e.g., the modified Metropolis-Hastings algorithm (Au and Beck, 2001). Let $i = i + 1$ and $\hat{G}_i(\mathbf{x}) = \hat{G}_{i-1}(\mathbf{x})$, and return to Step 4).

7) Drawn $N_p$ posterior samples in $F$ with the seeds $\mathbf{W}_s$ by calling the Kriging model $\hat{G}_m(\mathbf{x}, u)$ based on the modified Metropolis-Hastings algorithm.

The learning function in Equation (2.14) has been proposed by Echard et al. (2011). Because the Kriging predictor follows a Gaussian distribution, $\Phi(U(\mathbf{x}, u))$ indicates the probability of making a wrong classification on the sign of

$\hat{G}(\mathbf{x}, u) - b_i$, where $\Phi$ denotes the standard normal cumulative distribution function. Therefore, the stopping criterion

$(\min(U(\mathbf{x})) \geq 2)$ corresponds to the case that the probability of making a wrong classification on the sign of $\hat{G}(\mathbf{x}) - $

$b_i$ is less than $\Phi(-2) = 0.023$.



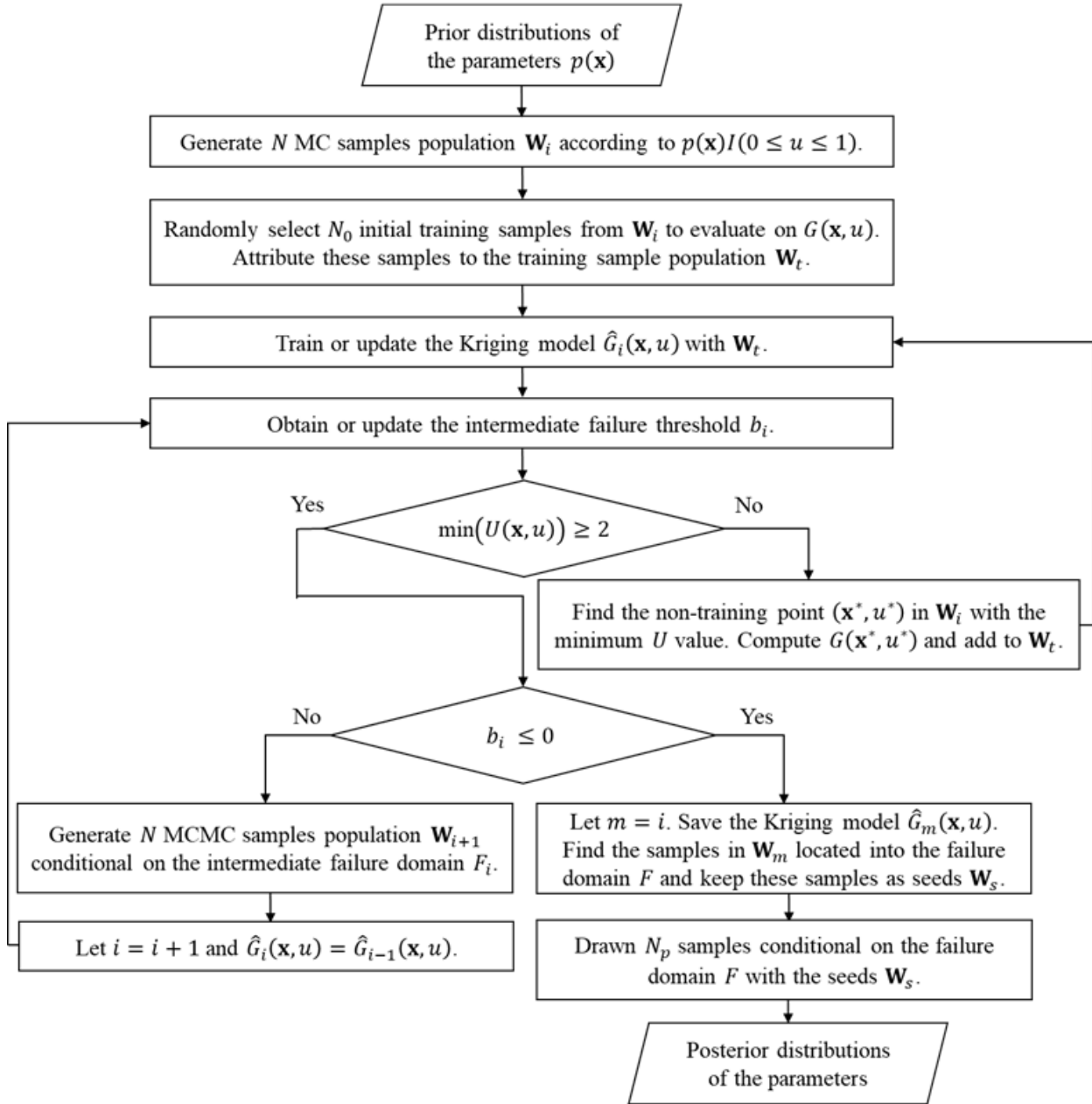Figure 2.2 Flowchart of the proposed Bayesian inference algorithm.

The advantage of the proposed Bayesian inference algorithm is that it only requires a small number of evaluations

to the computationally demanding limit state function for estimating the posterior distributions. Furthermore, no prior

information about the failure probability $p_F$ is required for implementing the algorithm, because the population size $N$

depends on the target probability $p_0$ which is defined by the analyst in advance. Nevertheless, in the main step of the ABC updating framework, the stochastic property of the Bhattacharyya distance might cause unsatisfactory inaccuracy in the classification of the failure surface by the Kriging model. Therefore, the use of common random numbers (CRN) (Kleinman et al., 1999) is also proposed in this step. CRN attempts to induce a positive correlation between stochastic outputs (i.e., Bhattacharyya distances) for different inputs and thereby reduces the variance in the difference between the stochastic outputs. Therefore, it works to avoid the inaccuracy in the establishment of the adaptive Kriging model of the Bhattacharyya distance-based limit state function.

## 2.4    Numerical example

### 2.4.1    Description of the Bayesian model updating problem

The proposed two-step ABC updating framework using dynamic response data is demonstrated through a model updating problem of a seismic-isolated bridge pier model based on simulated seismic response data. The target bridge is a seismic-isolated bridge, with lead rubber bearings, that is designed based on Japan Road Association (JRA) (2016). Descriptions of the target bridge are listed in Table 2.1. The reinforced concrete (RC) pier with the rubber bearings is modeled as a 2-DOF lumped mass system shown in Figure 2.3(a), where the superstructure and RC pier are represented as lumped masses and the rubber bearings and RC pier are characterized as nonlinear horizontal springs. The boundary condition at the surface is assumed to be fixed. The rubber bearings are idealized by a bi-linear model with the ratio of the yield stiffness $K_{B1}$ to the post-yield stiffness $K_{B2}$ as $6.5\!:\!1$ based on JRA (2004). On the other hand, the hysteresis and skeleton curves of the RC pier are idealized by a bi-linear model with the elasto-plastic characteristic and a stiffness degradation model, so-called Takeda model (Takeda et al., 1970), respectively. Rayleigh damping is assumed in which damping ratios of the rubber bearings and RC pier are given as 0 % and 2 %, respectively.

Table 2.1 Descriptions of the target bridge.

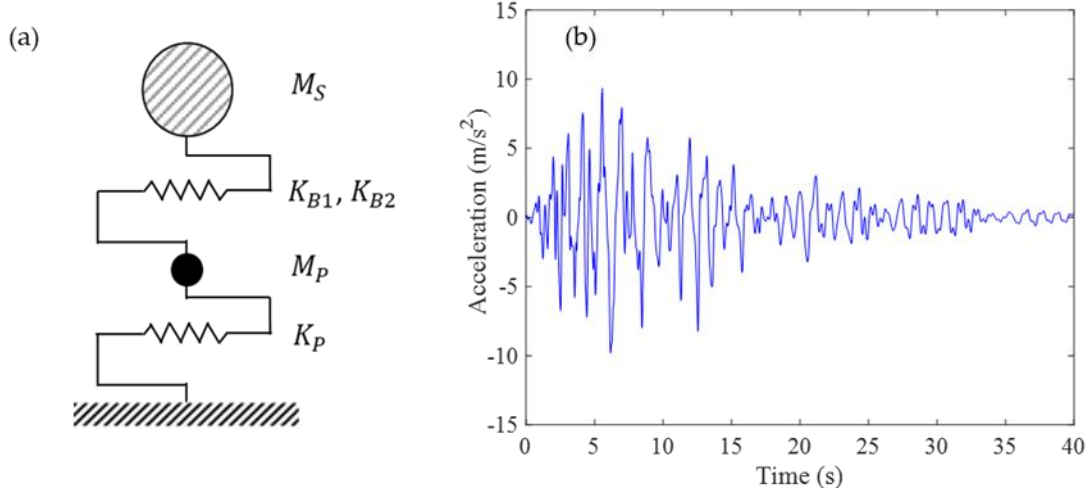| Member | Model parameter | Nominal value |
|---|---|---|
| Superstructure | Mass $M_S$ (ton) | 604.0 |
| Rubber bearings | Yield strength (kN) | 1118 |
| | Yield stiffness $K_{B1}$ (kN/m) | 40000 |
| | Post-yield stiffness $K_{B2}$ (kN/m) | 6000 |
| RC pier | Mass $M_P$ (ton) | 346.2 |
| | Yield strength (kN) | 3374 |
| | Yield stiffness $K_P$ (kN/m) | 110100 |
| | Yield displacement (m) | 0.0306 |
| | Ultimate displacement (m) | 0.251 |

Figure 2.3 (a) 2-DOF lumped mass system; (b) Time-history of the acceleration response at the superstructure.

The objective of the model updating problem is to capture the uncertainties in the post-yield stiffness of the rubber bearings $K_{B2}$, which characterize the nonlinear behavior of the target bridge under strong earthquakes, as well as in the other stiffness parameters $K_P$ and $K_{B1}$ using simulated seismic response data. The remaining parameters are assumed to be fixed constants with the nominal values, as shown in Table 2.1. The time-history of the acceleration response at the superstructure subjected to the level-2 type-II-II-2 earthquake, determined in JRA (2016), is employed as the output features whose uncertainties are driven by the three uncertain parameters $K_P$, $K_{B1}$, and $K_{B2}$. Dynamic response analysis of the 2-DOF system is carried out using Newmark $\beta$ method ($\gamma = 1/2$ and $\beta = 1/4$) with a time step $\Delta t = 0.001$ s. Figure 2.3(b) depicts a time-history of the acceleration response at the superstructure for the case where all parameters are considered as the nominal values in Table 2.1. The duration time of the time-history is 40 s with the time step $\Delta t = 0.001$ s; hence, the output features are in the 40,000 dimensional-space. Both aleatory and epistemic uncertainties are involved in this system and included by modeling $K_P$, $K_{B1}$, and $K_{B2}$ as independent Gaussian random variables, where the means and standard deviations are not fixed but unknown lying within given intervals. According to the parameter categories in Section 2.1, $K_P$, $K_{B1}$, and $K_{B2}$ are Category IV parameters while the remaining parameters are Category I parameters. The given intervals of the means $\boldsymbol{\mu}$ and standard deviations $\boldsymbol{\sigma}$ associated to $K_P$, $K_{B1}$, and $K_{B2}$ are detailed in Table 2.2.

Table 2.2 Uncertain characteristics and target epistemic inputs of the 2-DOF system.

| Parameter | Uncertainty characteristic | Target value of epistemic input |
|---|---|---|
| $K_P$ | Gaussian, $\mu_1 \in [0.5, 1.5]$, $\sigma_1 \in [0, 0.15]$ | $\mu_1 = 1.0$, $\sigma_1 = 0.07$ |
| $K_{B1}$ | Gaussian, $\mu_2 \in [0.5, 1.5]$, $\sigma_2 \in [0, 0.15]$ | $\mu_2 = 1.0$, $\sigma_2 = 0.07$ |
| $K_{B2}$ | Gaussian, $\mu_3 \in [0.5, 1.5]$, $\sigma_3 \in [0, 0.15]$ | $\mu_3 = 1.0$, $\sigma_3 = 0.07$ |

The target of the updating procedure, $\mathbf{Y}_{obs}$, is multiple sets of the output features obtained by assigning the target values of the epistemic inputs $\mu_1$, $\mu_2$, $\mu_3$, $\sigma_1$, $\sigma_2$, and $\sigma_3$ shown in Table 2.2. Those target values are set based on Adachi (2002). The sample size of the observed features is set to be $N_{obs} = 100$, generated by evaluating the model 100 times with the model parameters sampled from the assigned Gaussian distribution with the target epistemic inputs.

In addition to the target values in Table 2.2, a set of initial values of the epistemic inputs is arbitrary chosen within the pre-defined intervals but different from the target values. The sample size of the initial simulated features is set to be $N_{obs} = 500$, generated by evaluating the model 500 times with the model parameters sampled from their assigned Gaussian distribution with the initial epistemic inputs. Figure 2.4 illustrates the relative positions of the target observed features and initial simulated features. The RMS values of both the observed and simulated features for each interval divided based on the given window length $L = 0.025 \times 40000 = 1000$ are computed and five arbitrary selected RMS values $RMS_{ACC}^{j}$ are illustrated in this figure. The diagonal subfigures compare histograms of the observed and initial simulated features. Since the initial values of the epistemic inputs are assigned differently from there target values, the scatters and histograms of the initial simulated features are clearly apart from those of the target observed features. It is noted that, Bayesian model updating is not really started from these initial values, but from the prior distribution of the epistemic inputs, shown in the second column of Table 2.2.
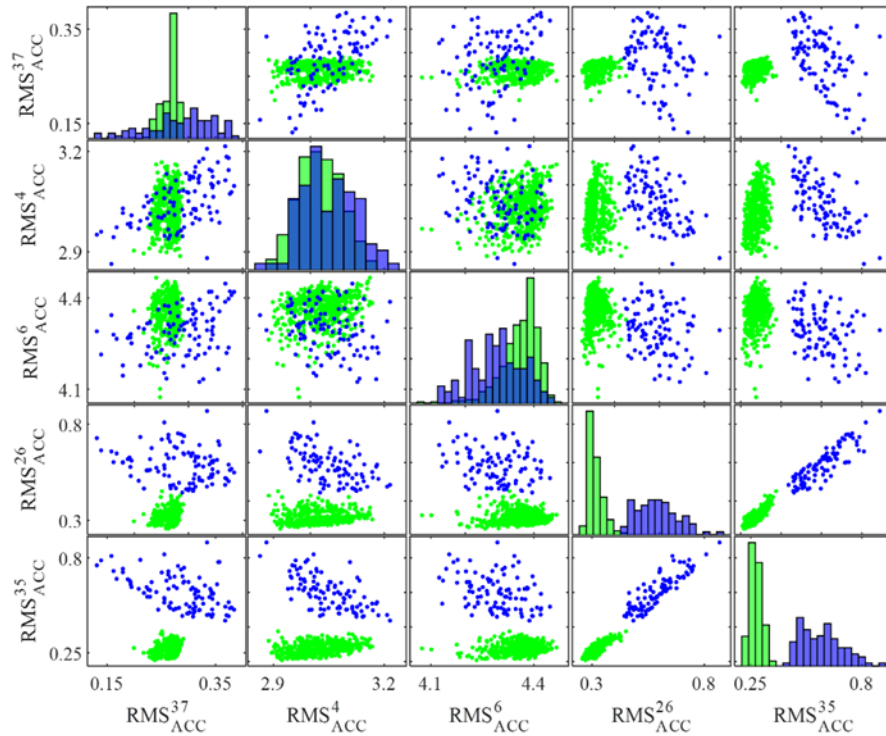


Figure 2.4 Target observed scatters (in blue) and initial simulated scatters (in green); unit: $\mathrm{m/s^2}$.

### 2.4.2   Updating results with the Euclidian distance

In the first step where the Euclidian distance is taken as the metric, the geometric distance between the centre of mass of the two sample sets is measured, whereas the dispersion and distribution information of the sample sets cannot be considered. Hence, only the mean parameters are calibrated, whose prior distribution is set to be uniform within the intervals shown in Table 2.1, by representing the model parameters by the mean parameters through this step.

The parameters of the proposed algorithm are set as $N = 3000$, $N_0 = 12$, $p_0 = 0.01$, and $N_p = 500$. The width factor in the approximate likelihood is set as $\varepsilon = 0.1$. Totally four intermediate failure surfaces are produced to finally provide the classification for the true failure domain. It implies that the failure probability of the equivalent reliability problem herein reaches around $10^{-8}$. Even for such a challenging problem, the number of the total training samples is 229, selected by evaluating the limit state function based on the Euclidian distance metric 229 times. The computation of the Euclidian distance needs a single model evaluation with the parameter means. Thus, only 229 model evaluations are required throughout this step.

As illustrated in Figure 2.5, the posterior distribution of the parameter means well converges to their target values shown in the red lines. The horizontal axes of the figure are set to correspond to their prior intervals listed in Table 2.2. Table 2.3 presents the updated values of the parameter means obtained by estimating means of the posterior distribution. Percentage errors compared with the target values are also provided in the parentheses after the updated values.
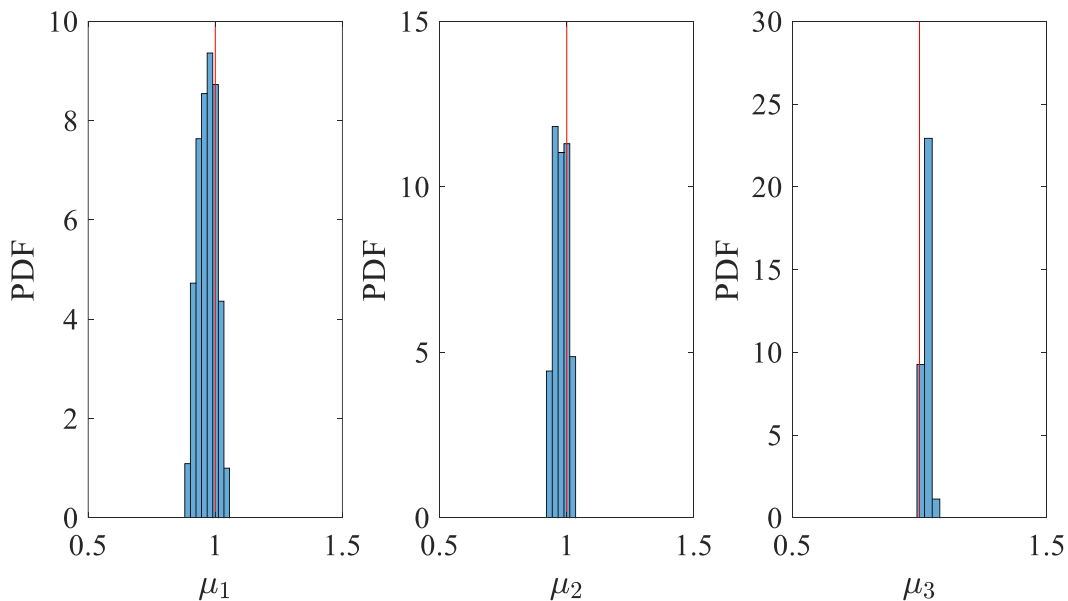


Figure 2.5 Posterior distribution of the parameter means after updating with the Euclidian distance.

Table 2.3 Updated epistemic inputs with both the Euclidian and Bhattacharyya distances.

| Input | Target value | Updated value | |
|---|---|---|---|
| | | With Euclidian distance | With Bhattacharyya distance |
| $\mu_1$ | 1.0 | 0.9682 (3.18 %) | 1.0125 (1.25 %) |
| $\mu_2$ | 1.0 | 0.9772 (2.28 %) | 1.0159 (1.59 %) |
| $\mu_3$ | 1.0 | 1.0276 (2.76 %) | 1.0024 (0.24 %) |
| $\sigma_1$ | 0.07 | – | 0.0604 (13.7 %) |
| $\sigma_2$ | 0.07 | – | 0.0572 (18.3 %) |
| $\sigma_3$ | 0.07 | – | 0.0813 (16.1 %) |

Moreover, Figure 2.6 shows the relative positions of the target observed features and updated simulated features. The updated simulated features are obtained by evaluating the model 500 times with the model parameters sampled from their assigned Gaussian distributions with the updated means shown in Table 2.3 and variances arbitrary selected from the prior intervals. It can be seen that the simulated features are progressively shifted toward the observed features as a result of minimizing the Euclidian distance metric, corresponding to the maximization of the likelihood. However, there are still some discrepancies between the observed and simulated features. These discrepancies are addressed in the next step using the Bhattacharyya distance as the metric.
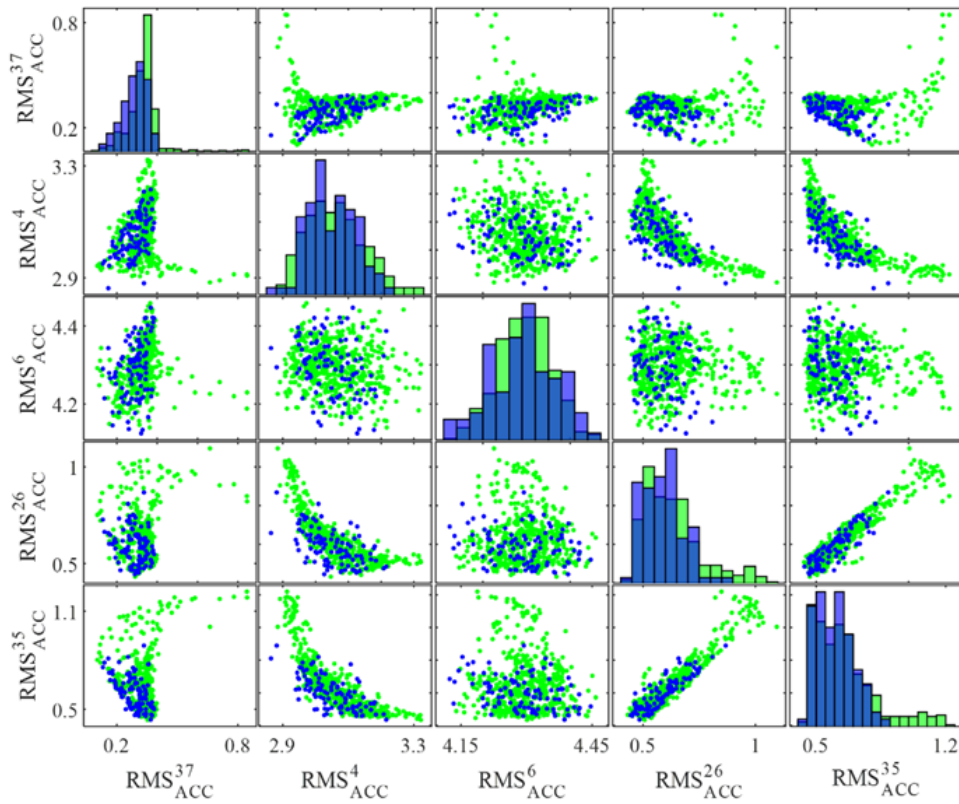


Figure 2.6 Target observed scatters (in blue) and simulated scatters after updating with the Euclidian distance (in green); unit: $m/s^2$.

### 2.4.3   Updating results with the Bhattacharyya distance

This section presents the second step where the Bhattacharyya distance is employed as the metric. The posterior distribution obtained in the first step is taken as the prior distribution of the parameter means in this step. At the same time, the prior distribution of the parameter variances is set to be uniform within their intervals listed in Table 2.1. The model parameters are represented as the assigned Gaussian distribution having the means and variances sampled from the prior distribution, so that both the parameter means and variances are updated simultaneously. In each computation of the Bhattacharyya distance, 100 random samples of the model parameters are generated and similarly 100 simulated features are obtained.

The parameters of the proposed algorithm are set to be same as those in the preliminary step. The width factor in the approximate likelihood is set to be $\varepsilon = 0.01$. After four intermediate failure surfaces are produced, the final Kriging model providing the classification for the true failure domain is obtained. It indicates that the failure probability of the equivalent reliability problem herein also reaches around $10^{-8}$. The number of the training samples is 521, selected by evaluating the limit state function based on the Bhattacharyya distance metric 521 times. Differently from the first step, the computation of the Bhattacharyya distance requires 100 model evaluations; hence, totally 52100 model evaluations are executed throughout this step.

Figure 2.7 presents the finally updated posterior distribution of all the epistemic inputs. The posterior distributions of the means are further updated to be more centralized to their target values compared with those shown in Figure 2.5. This is caused by introducing the posterior samples in the first step as the prior samples in this step. More attention is paid to the posterior distributions of the standard deviations, which almost well converge to their target values presented in the red lines. The estimated means of the posterior distributions are summarized in the last column of Table 2.3 as the updated values of the epistemic inputs. The parameter means indicate quite high updating precisions with predicted errors less than 2%, whereas the parameter standard deviations denote relatively large predicted errors more than 13%. This fulfils the general experience that dispersion information of the parameters is much more difficult to be precisely updated than mean information. Nevertheless, the finally updated simulated features obtained by evaluating the model 500 times with the model parameters sampled from their assigned Gaussian distributions with these updated epistemic inputs well coincide with the target observed features, as shown in Figure 2.8. This demonstrates that the Bhattacharyya distance is a powerful UQ metric with the capability to recreate wholly the distribution of the target observations that are described as the dynamic response data.
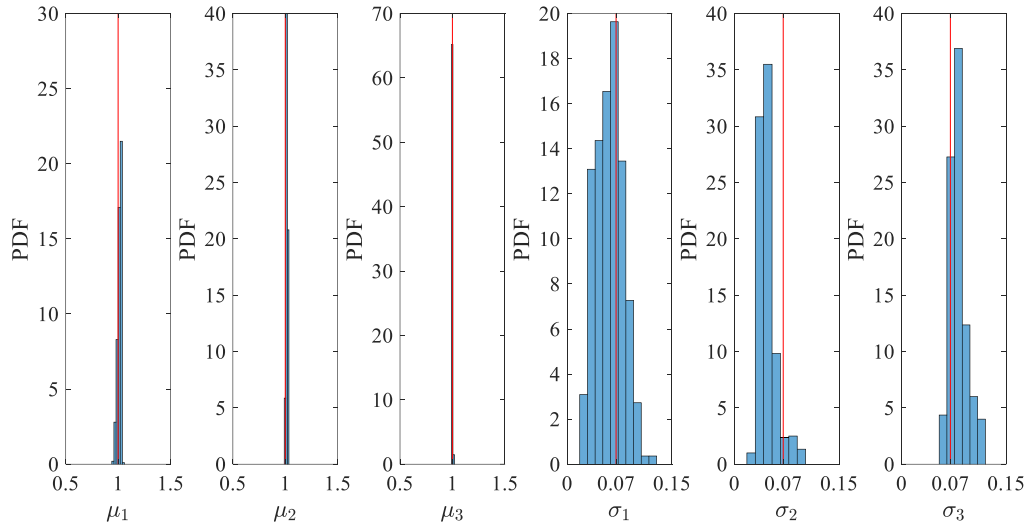
Figure 2.7 Posterior distribution of the epitemic inputs after updating with the Bhattacharyya distance.
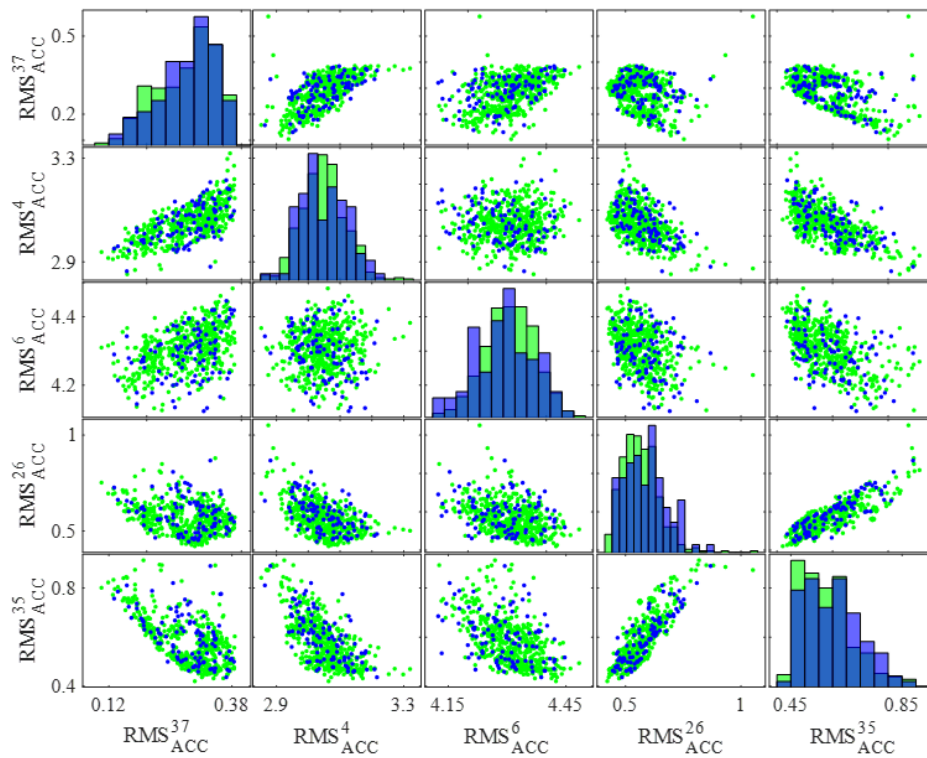


Figure 2.8 Target observed scatters (in blue) and simulated scatters after updating with the Bhattacharyya distance (in green); unit: $m/s^2$.

## 2.4.4   Computational efficiency

Finally, computational efficiency of the proposed procedure is demonstrated. For comparison, the two-step ABC updating procedure is also executed using the TMCMC algorithm. The number of samples generated from the posterior

distribution are set to be $N_p = 500$ as same as that in the proposed procedure. The width factors in the approximate

likelihoods are set to be also same as those in the proposed procedure in order to keep the same computational demand

for convergence as in the proposed procedure. All the computations are processed using a local parallelization on a 12

cores machine installing an Intel core 2.10 GHz processor.

Table 2.4 summarizes the total computational time (in minutes) to reach convergence for both the first and second

steps, where the Euclidian and Bhattacharyya distances are used as metrics, respectively. In this context, computational

efficiency is indicated as the ratio of the computational time by TMCMC and the proposed algorithm combining BUS

and the adaptive Kriging model, and is shown in the parentheses after the computational time of the proposed algorithm.

It can be seen that the second step with the Bhattacharyya distance needs much more computational demands than the

first step with the Euclidian distance. The computational time in the second step is more than 300 times of that in the

first step for TMCMC and is about 50 times of that in the first step for the proposed algorithm. It is obviously due to

the necessity of MC sampling for each computation of the Bhattacharyya distance. Nevertheless, the difference in the

computational time of those two steps is successfully decreased in the proposed procedure by confining the evaluation

of the likelihood function only for the Kriging approximation.

Table 2.4 Comparison of computational efficiency.

| Method | Computational time (minutes) | |
|---|---|---|
| | With Euclidian distance | With Bhattacharyya distance |
| TMCMC | 39.1 | 12437.5 |
| BUS with the adaptive Kriging | 7.9 (5.0) | 394.4 (31.5) |

In addition, it is noted that the proposed procedure reaches convergence with one-fifth of the computational time

in the first step and with less than one-thirty of that in the second step, compared with TMCMC. This is mainly because

the number of the model evaluations is significantly reduced by employing the adaptive Kriging model in the proposed

procedure. It should be noted that the adaptive Kriging model can be also implemented in TMCMC. However, several

modifications are necessary to employ the adaptive Kriging model in TMCMC as the approximation of the likelihood

function (Angelikopoulos et al., 2015; Jensen et al., 2017), because it has been originally developed as the classification

method in reliability problems. Meanwhile, the proposed algorithm transforms the Bayesian updating problem into the

equivalent reliability problem; thus, the adaptive Kriging model is naturally implemented as the classification method

for the associated limit state function. As a consequence, the proposed Bayesian inference algorithm combining BUS

and the adaptive Kriging model is capable to produce satisfactory results with the much-reduced computational demand

compared with TMCMC.

## 2.5 Conclusions

In this study, a novel Bayesian inference algorithm combining BUS and the adaptive Kriging model is developed so as to effectively carry out the two-step ABC updating framework using dynamic response data. The distance-based approximate likelihood function in this framework is capable of maximizing the acceptance rate in BUS, because the optimal likelihood multiplier is straightforwardly applicable. Furthermore, to cope with the tremendous computational demand in the Bhattacharyya distance evaluation, the adaptive Kriging model is implemented based on the AK-MCMC algorithm to provide the classification for the limit state function associated with the Bhattacharyya distance. The AK-MCMC algorithm provides the classifications for a series of intermediate failure domains, which will finally converge to the classification for the true failure domain, thus it is much efficient compared with the direct classification for the failure domain. The proposed procedure is demonstrated upon the seismic-isolated bridge pier model updating example using simulated seismic response data. This example demonstrates that the Bhattacharyya distance is a powerful UQ metric with the capability to recreate wholly the distribution of the target observations and that the proposed inference algorithm is enable to provide satisfactory results with much-reduced computational demand compared with TMCMC.

## References

Adachi, Y. (2002). *Reliability analysis and limit state design method of isolated bridges under extreme ground motions*. Doctoral thesis, Kyoto University, Kyoto. (In Japanese.)

Angelikopoulos, P., Papadimitriou, C., and Koumoutsakos, P. (2015). X-TMCMC: Adaptive kriging for Bayesian inverse modeling. *Computer Methods in Applied Mechanics and Engineering*, 289, 409-428.

Au, S. K., and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4), 263-277.

Beck, J. L., and Katafygiotis, L. S. (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechnics*, 124(4), 455-461.

Beck, J. L., and Au, S. K. (2002). Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *Journal of Engineering Mechanics*, 128(4), 380-391.

Betz, W., Papaioannou, I., and Straub, S. (2016). Transitional Markov chain Monte Carlo: Observations and improvements. *Journal of Engineering Mechanics*, 142(5), 04016016.

Betz, W., Papaioannou, I., Beck, J. L., and Straub, S. (2018). Bayesian inference with Subset Simulation: Strategies and improvements. *Computer Methods in Applied Mechanics and Engineering*, 331, 72-93.

Bhattacharyya, A. (1946). On a measure of divergence between two multinominal populations. *Indian Journal of Statistics*, 7(4), 401-406.

Bi, S., Prabhu, S., Cogan, S., and Atamturktur, S. (2017). Uncertainty quantification metrics with varying statistical information in model calibration and validation. *AIAA Journal*, 55(10), 3570-3583.

Bi, S., Broggi, M., and Beer, M. (2019). The role of the Bhattacharyya distance in stochastic model updating. *Mechanical System and Signal Processing*, 117, 437-452.

Cheung, S. H., and Beck, J. L. (2009). Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters. *Journal of Engineering Mechanics*, 1135(4), 243-255.

Ching, J., and Chen, Y. C. (2007). Transitional Markov chain Monte Carlo method for Bayesian updating, model class selection, and model averaging. *Journal of Engineering and Mechanics*, 133(7), 816-832.

Crespo, L. G., Kenny, S. P., and Giesy, D. P. (2014). The NASA Langley multidisciplinary uncertainty quantification challenge. In: *16th AIAA Non-Deterministic Approaches Conference*, National Harbor, Maryland.

DiazDelaO, E. A., Garbuno-Inigo, A., Au, S. K., and Yoshida, I. (2017). Bayesian updating and model class selection with Subset Simulation. *Computer Methods in Applied Mechanics and Engineering,* 317: 1102-1121.

Echard, B., Gayton, N., and Lemaire, M. (2011). AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(2), 145-154.

Echard, B., Gayton, N., Lemaire, M., and Relun, N. (2013). A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliability Engineering & System Safety*, 111, 232-240.

Grimmett, G. R., and Stirzaker, D. R. (2001). *Probability and random processes*. New York, Oxford University Press.

Huang, X., Chen, J., and Zhu, H. (2016). Assessing small failure probabilities by AK-SS: An active learning method combing Kriging and Subset Simulation. *Structural Safety*, 59, 86-95.

Japan Road Association. (2016). *Design specifications of highway bridges V: Seismic design.* Tokyo, Maruzen.

Japan Road Association. (2004). *Manual on bearings for highway bridges*. Tokyo, Maruzen. (In Japanese.)

Jensen, H. A., Vergara, C., Papadimitriou, C., and Millas, E. (2013). The use of updated robust reliability measures in stochastic dynamical systems. *Computer Methods in Applied Mechanics and Engineering,* 267: 293-317.

Jensen, H. A., Esse, C., Araya, V., and Papadimitriou, C. (2017). Implementation of an adaptive meta-model for Bayesian finite element model updating in time domain. *Reliability Engineering & System Safety,* 160, 174-190.

Kennedy, M. C., and O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3), 425-464.

Kitahara, M., Broggi, M, and Beer, M. (2020). Bayesian model updating for existing seismic-isolated bridges using observed acceleration response data. In: *XI international conference on structural dynamics*, Athens, Greece.

Kleinman, N. L., Spall, J. C., and Naiman, D. Q. (1999). Simulation-based optimization with stochastic approximation using common random numbers. *Management Science*, 45(11), 1570-1578.

Katafygiotis, L. S., and Beck, J. L. (1998). Updating models and their uncertainties. II: Model identifiability. *Journal of Engineering Mechanichs*, 124(4), 463-467.

Patelli, E., Govers, Y., Broggi, M., Gomes, H. M., Link, M., and Mottershead, J. E. (2017). Sensitivity or Bayesian model updating: a comparison of techniques using the DLR AIRMOD test data. *Archive of Applied Mechanics,* 87, 905-925.

Patra, B. K., Launonen, R., Ollikainen, V., and Nandi, S. (2015). A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. *Knowledge-Based Systems*, 82, 163-177.

Rocchetta, R., Broggi, M., Huchet, Q., and Patelli, E. (2018). On-line Bayesian model updating for structural health monitoring. *Mechanical Systems and Signal Processing*, 103, 174-195.

Safta, C., Sargsyan, K., Najm, H. N., Chowdhary, K., Debusschere, B., Swiler, L. P., and Eldred, M. S. (2015). Probabilistic methods for sensitivity analysis and calibration in the NASA challenge problem. *Journal of Aerospace Information Systems,* 12(1), 170-188.

Straub, D., and Papaioannou, I. (2015). Bayesian updating with structural reliability methods. *Journal of Engineering and Mechanics*, 141(3), 04014134.

Takeda, T., Sozen, M. A., and Nielsen, N. N. (1970). Reinforced concrete response to simulated earthquakes. *Journal of the Structural Division*, 96(12), 2557-2573.

Turner, B. M., and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2), 69-85.

Wei, P., Tang, C., and Yang, Y. (2019). Structural reliability and reliability sensitivity analysis of extremely rare failure events by combining sampling and surrogate model methods. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233(6), 943-957.

# Chapter 3

# Research article 2: Nonparametric Bayesian stochastic model updating with hybrid uncertainties

This is the second phase in the main part of this thesis, that aims at developing a distribution-free stochastic model updating framework for calibrating the category IV parameters distribution families of which are unknown a priori. In the current literature, stochastic model updating frameworks to calibrate the category IV parameters are generally based on the availability of prior information on the distribution families of the inferred parameters. This is also the case for the Bhattacharyya distance-based ABC updating framework, where, during the updating procedure, the Bhattacharyya distance is evaluated for each pair of the inferred parameters, i.e., the hyper-parameters of pre-determined distributions of the category IV parameters, for assigning the likelihood to it. In engineering applications, however, prior information on the parameters is often very limited due to scarce or/and incomplete available data for the parameters, and thus the distribution families of the parameters cannot be pre-determined with confidence. For instance, the latest (2019) edition of the NASA UQ challenge problem, which will be addressed in the next section, requires a model calibration task in an extremely challenging situation that no distribution information of the aleatory parameters is available other than a common bounded support domain. In such situation, not only the distribution hyper-parameters but also the distribution families themselves should be considered as epistemic uncertainty to be calibrated through the updating procedure, so that the calibrated category IV parameters follow arbitrary shaped distributions from which the model outputs identical to the observations can be obtained.

To achieve this challenging expectation in stochastic model updating, the present article proposes the use of the staircase density functions in conjunction with the Bhattacharyya distance-based ABC model updating. The staircase density function is a recently proposed novel class of distribution families, defined as a piecewise constant function on the bounded support set, based on the quadruple of the first four moments. It can flexibly approximate a wide range of distributions, including very skewed or/and multimodal distributions, arbitrary close by employing the corresponding quadruples of the first four moments. The category IV parameters can be straightforwardly described using the staircase density functions by defining the first four moments as interval valued hyper-parameters. Herein, the support set of the

first four moments can be derived based on moment constraints conditional on the given support set of the parameters. In this manner, by employing the staircase density functions, no detailed prior information on the parameters, such as their distribution families and the support set of their hyper-parameters, is necessary to perform model updating other than the support set of the parameters.

Consequently, a novel ABC updating framework is developed, in which the category IV parameters are described as the staircase density functions and their hyper-parameters are calibrated based on the Bhattacharyya distance-based approximate likelihood function. All the possible density shapes on the given support set are considered in its first step, while most of them are disregarded through the updating procedure such that only the density shapes which correspond to the hyper-parameters within the posterior support set are remained in the finally calibrated p-boxes. If the posterior distribution of the hyper-parameters is sharply converged to their true values, then the true distributions can be uniquely inferred. The proposed updating procedure is first demonstrated upon a well-known simple engineering example. The target distribution of the model parameters in this example shows bimodality and does not belong to the standard class of distribution families. In this example, a preliminary step with the Euclidian distance-based approximate likelihood function is also employed to serve as a preconditioner to avoid non-unique solutions due to the bimodality of the target distribution. This example demonstrates that the procedure has the capability to estimate arbitrary shaped distributions of the parameters identical to their target distributions.

Furthermore, the model updating subproblem of the NASA UQ challenge problem 2014 is addressed as a complex real-world application. Whereas the distribution families of all the three aleatory parameters are pre-determined in its original problem, an altered problem setting ignoring information on the distribution families is defined in this article, making the problem more challenging. This application demonstrates that the proposed procedure is capable to achieve a satisfactory narrow p-box of the model predictions containing the CDF of the observed data as a result of the reduction of epistemic uncertainty in both the distribution hyper-parameters and distribution families.

# Nonparametric Bayesian stochastic model updating with hybrid uncertainties

Masaru Kitahara [a, *], Sifeng Bi [b], Matteo Broggi [a], Michael Beer [a, c, d]

[a] *Leibniz Universität Hannover, Institute for Risk and Reliability, Callinstrasse 34, Hannover, Germany*

[b] *Beijing Institute of Technology, School of Aerospace Engineering, Beijing, China*

[c] *University of Liverpool, Institute for Risk and Uncertainty, Peach Street L69 7ZF, Liverpool, United Kingdom*

[d] *Tongji University, International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Shanghai 200092, China*

* Correspondence author. E-mail address: masaru.kitahara@irz.uni-hannover.de (M. Kitahara).

**Abstract:** This work proposes a novel methodology to fulfil the challenging expectation in stochastic model updating to calibrate the probabilistic distributions of parameters without prior knowledge about their distribution families. To achieve this expectation, an approximate Bayesian computation model updating framework is developed by employing staircase random variables and the Bhattacharyya distance. In this framework, parameters with aleatory and epistemic uncertainties are described as staircase random variables. The discrepancy between model predictions and observations are then quantified by the Bhattacharyya distance-based approximate likelihood. In addition, a Bayesian updating using the Euclidian distance is performed as preconditioner to avoid non-unique solutions. The performance of the proposed procedure is demonstrated on two exemplary applications, a simulated shear building model example and a challenging benchmark problem for uncertainty treatment. These examples demonstrate the feasibility of the combined use of the staircase random variables and Bhattacharyya distance in stochastic model updating and uncertainty characterization.

**Keywords:** Stochastic model updating; Approximate Bayesian computation; Nonparametric probability-box; Staircase random variable; Bhattacharyya distance.

## 3.1    Introduction

It has been widely acknowledged that uncertainties should be appropriately considered in the campaign of model updating. The uncertainties can be typically classified into the two categories, i.e., aleatory and epistemic uncertainties (Oberkampf et al., 2004; Roy and Oberkampf, 2011). Aleatory uncertainty is the inherent variation or randomness, and thus cannot be reduced, but it enables to be described as precise probability models. Conversely, epistemic uncertainty is due to lack of knowledge, and is not completely avoidable, although it can be reduced through model updating using available data.

The complexity of model updating depends on the presence of different levels of uncertainties. The deterministic model updating generally considers the case with the presence of only epistemic uncertainty, in which parameters are unknown-but-fixed constants and are represented as non-probabilistic models, such as interval/convex models (Moore et al., 2009) and fuzzy set theory (Faes and Moens, 2020). The deterministic model updating is aimed at a single set of parameter values and at generating a single model prediction with maximum fidelity with regard to the observation.

On the other hand, the presence of both aleatory and epistemic uncertainties simultaneously (i.e. hybrid or mixed uncertainties) is considered by the stochastic model updating. In this situation, parameters are represented as imprecise probability models, where parameters are indeed aleatory uncertainty but their distribution parameters, e.g., means and variances, are epistemic uncertainty. Commonly used imprecise probability models include evidence theory (Sentz and Ferson, 2002), probability-box (also known as p-box) (Ferson et al., 2003), and fuzzy probability model (Beer et al., 2013). In particular, those parameterized ones such as parametric p-box have attracted the most attentions due to their simplicity and ease of applications. The stochastic model updating aims at not the single set of parameter values but a reduced space of epistemic uncertainty and at generating stochastic model predictions capable to represent uncertainty characteristics of multiple sets of observations.

The uncertainties during the stochastic model updating procedure, i.e., the reasons for the stochastic discrepancy between the model predictions and observations, can be summarized as follows:

- Parameter uncertainty. The input parameters of the numerical model, such as material properties and boundary conditions, are imprecisely determined;

- Modelling uncertainty. The numerical model always contains inevitable simplifications and approximations of the physical system;

- Measurement uncertainty. The measurements are driven by hard-to-control randomnesses, e.g., environmental noises and measurement system errors.

A broad range of stochastic model updating methods has been investigated, such as perturbation method (Mares et al., 2006; Khodaparast et al., 2008), Monte Carlo method (Sairajan and Aglietti, 2012; Bi et al., 2017), and Bayesian inference (Goller et al., 2011; Patelli et al., 2017). No matter which method is utilized in the stochastic model updating, it is significant to define a comprehensive uncertainty quantification (UQ) metric capable of quantifying the statistical discrepancy between the model predictions and observations due to the above uncertainties. The Euclidian distance, Mahalanobis distance, and Bhattacharyya distance are different levels of distance metrics. These distances have been investigated as UQ metrics and the Bhattacharyya distance has been demonstrated to be able to capture a higher level

of statistical information from the investigated sample sets (Bi et al., 2017). Moreover, Bi et al. (2019) has developed a Bayesian model updating framework, in which the Bhattacharyya distance is utilized as the UQ metric to define an approximate likelihood function by the approximate Bayesian computation (ABC) framework (Beaumont et al., 2002; Turner and Van Zandt, 2012). This framework has been demonstrated to be a comprehensive updating procedure with the capability to recreate wholly the distribution of target observations.

However, stochastic model updating methods in the literature including Bi et al. (2019) are in general based on the parameterized imprecise probability models; hence, it relies upon the pre-hypothesis of the distribution format for propagating epistemic uncertainty into parameters. For instance, in the NASA UQ challenge problem 2014 (Crespo et al., 2014), which has gained attentions as the real-size practical uncertainty quantification problem, prior information about the distribution formats is fully provided to perform model updating. On the other hand, it is often the case that the distribution formats are unknown a priori due to scarce and incomplete available data for the parameters. Hence, it is desired to develop a nonparametric model updating framework, where epistemic uncertainty is propagated into the parameters without prior knowledge about the distribution formats.

Crespo et al. (2018) recently proposed a family of random variables having a bounded support set and prescribed values for the first four moments. The variables are called staircase random variable since the density function is given as piecewise constant functions on pre-defined subintervals partitioning the support sets. Moment constraints for the existence of such variables are obtained as a series of inequalities conditioned upon the support sets. As a consequence, the staircase random variable can represent a broad range of density shapes, including very skewed and/or multimodal distributions. The staircase random variable belongs to the precise probability models, however, its combination with the non-probabilistic models has the potential to provide a non-parameterized imprecise probability model, where the parameters are indeed aleatory uncertainty but the first four moments are epistemic uncertainty.

The objective of this work is consequently to propose a novel methodology that fulfils the challenging expectation in the stochastic model updating to calibrate the probabilistic distribution of parameters without prior knowledge about the distribution format. To achieve this objective, an ABC model updating framework is developed by employing the Bhattacharyya distance and staircase random variables. At the same time, a Bayesian updating based on the Euclidian distance is performed as preconditioner to avoid non-unique solutions. This updating framework is independent of the distribution format of investigated parameters; hence, it demonstrates clear advantages in calibrating parameters whose probabilistic distribution cannot be defined analytically. The proposed framework is demonstrated using a simple shear building model for illustration. Moreover, it is applied to the NASA UQ challenge problem 2014. We focus on solving

Sub-problem A (uncertainty characterization), where the distribution formats of the investigated parameters are given, however we redefine the problem ignoring them to demonstrate the performance of the proposed framework.

The rest part of this paper is organized as follows. In Section 3.2, we describe the theoretical and methodological bases of the Bhattacharyya distance metric and staircase random variables. Section 3.3 outlines the novel development of the Bayesian updating with the staircase random variables, and the proposed two-step ABC updating procedure. The principle and illustrative application are presented in Section 3.4, using a simple shear building model, and in Section 3.5, concentrating on the demonstration of the performance of the framework upon the highly challenging NASA UQ problem. Finally, some conclusions are given in Section 3.6.

## 3.2   Theories and methodologies

### 3.2.1   Bhattacharyya distance metric

In the context of stochastic model updating, the system under investigation is characterized as:

$$\mathbf{y} = h(\mathbf{x}) \tag{3.1}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_n]$ denotes a vector of $n$ input parameters; $\mathbf{y} = [y_1, y_2, \cdots, y_m]$ is a vector of $m$ output features; $h(\cdot)$ denotes the simulator. The simulator herein is usually presented as either a sophisticated numerical analysis code, e.g., finite element model, or a metamodel.

The uncertainties of the system are first characterized by the input parameters described as precise probability models, non-probabilistic models, and imprecise probability models depending on the presence of different levels of uncertainties (refer to Section 3.1). The uncertainties are then propagated through the simulator into the output features presenting various forms of uncertainty as well, such as probabilistic distributions, intervals, and fuzzy sets. In general, regardless of the form of uncertainty, randomly sampled values of the parameters and features are used in the stochastic model updating. Suppose the required sample size be $N_{\mathrm{sim}}$, the simulator $h$ is executed $N_{\mathrm{sim}}$ times for generating the sample set of the simulated features $\mathbf{Y}_{\mathrm{sim}} \in \mathbb{R}^{N_{\mathrm{sim}} \times m}$:

$$\mathbf{Y}_{\mathrm{sim}} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_m], \text{with } \mathbf{y}_i = \left[y_{1i}, y_{2i}, \cdots, y_{N_{\mathrm{sim}}i}\right]^T, \forall i = 1, 2, \cdots, m \tag{3.2}$$

In addition to the simulated features, observed features that are collected from the campaign of experiments or measurements are also required as the target of model updating. Suppose the number of observations be $N_{\mathrm{obs}}$, then the

sample set of the observed features has a similar structure as Equation (3.2), where only the number of rows is changed: $\mathbf{Y}_{\text{obs}} \in \mathbb{R}^{N_{\text{obs}} \times m}$. The objective of the stochastic model updating can be then expressed as to minimize the discrepancy between $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$ by updating uncertainty characteristics of the input parameters.

After the simulated and observed features are available, the UQ metric is defined for quantifying the discrepancy between $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$. The classical Euclidian distance metric is expressed as:

$$d_E(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{sim}}) = \sqrt{(\bar{\mathbf{Y}}_{\text{obs}} - \bar{\mathbf{Y}}_{\text{sim}})(\bar{\mathbf{Y}}_{\text{obs}} - \bar{\mathbf{Y}}_{\text{sim}})^T} \tag{3.3}$$

where $\bar{\mathbf{Y}}_{(\cdot)}$ refers to a row vector of means of the features. The Euclidian distance is a point-to-point distance between the centre of the mass of two sample sets and is generally used in the deterministic model updating. Comparatively, in the stochastic model updating, it is more desirable to employ a more comprehensive metric that is capable to consider not only the means but also a higher level of statistical information, e.g., variances, covariances, and even distribution shapes.

The Bhattacharyya distance is herein proposed as a stochastic metric measuring the degree of the overlap between distributions of two sample sets. Its original definition is given as:

$$d_B(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{sim}}) = -\log\left[\int_{\mathbb{y}} \sqrt{p_{\text{obs}}(y)p_{\text{sim}}(y)}\,\mathrm{d}y\right] \tag{3.4}$$

where $p_{(\cdot)}(y)$ is the probability density function (PDF) of each feature sample; $\mathbb{y}$ is the $m$-dimensional feature space; $\int_{\mathbb{y}}(\cdot)\,\mathrm{d}y$ indicates the integration performed over the whole feature space. Differently from the Euclidian distance, the Bhattacharyya distance takes not only the means but also the variances, covariances, and even the distribution shapes, of the sample sets into account. However, the direct evaluation of Equation (3.4) is not feasible since precise estimation of the PDF is generally unavailable due to the very limited number of observations. Bi et al. (2019) thus proposed the so-called binning algorithm for evaluating the probability mass function (PMF) of a discrete distribution, such that the discrete Bhattacharyya distance is used instead. The PMF is a function to map the possible values of a discrete random variable to the probabilities of their occurrences (Grimmett and Stirzaker, 2001). The discrete Bhattacharyya distance is evaluated as (Patra et al., 2015):

$$d_B(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{sim}}) = -\log\left\{\sum_{i_m=1}^{n_{bin}} \cdots \sum_{i_1=1}^{n_{bin}} \sqrt{p_{\text{obs}}(b_{i_1,i_2,\cdots,i_m})p_{\text{sim}}(b_{i_1,i_2,\cdots,i_m})}\right\} \tag{3.5}$$

where $p_{(\cdot)}\left(b_{i_1,i_2,\cdots,i_m}\right)$ is the PMF value of the bin $b_{i_1,i_2,\cdots,i_m}$. The bin has $m$ subscripts because it is generated under a $m$-dimensional joint PMF space.

The binning algorithm for the PMF calculation consists of the following steps:

1) Define a common interval $I_i$ of both $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$ according to the $i$th feature $\mathbf{y}_i$, $\forall i = 1, 2, \cdots, m$, by finding the maximum and minimum values of $\mathbf{y}_i$ in both $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{sim}}$;

2) Within the defined interval, arbitrary decide the number of bins $n_{\text{bin}}$;

3) Count the joint probability mass for each bin $p_{(\cdot)}\left(b_{i_1,i_2,\cdots,i_m}\right)$. It should be noted that, the total number of bins in the $m$-dimensional feature space is $N_{\text{bin}} = n_{\text{bin}}^m$.

The principle of $n_{\text{bin}}$ in Step 2) is that a larger $n_{\text{bin}}$ results in employing more detailed information of the distribution characteristics and to a larger value of the Bhattacharyya distance, while it also leads to a larger computational cost. In Bi et al. (2019), $n_{\text{bin}}$ is recommended to be $n_{\text{bin}} = \left\lceil \frac{\max(N_{\text{obs}}, N_{\text{sim}})}{10} \right\rceil$, where $\lceil \cdot \rceil$ is the upper integer.

## 3.2.2  Staircase random variables

In this study, parameters with hybrid uncertainties are characterized as staircase random variables. The staircase random variable $x$ is constrained to possess a bounded support set $[\underline{x}, \overline{x}]$ and a pair of variables $\boldsymbol{\theta}_x = [\mu, m_2, m_3, m_4]$ consisting of the mean $\mu$, variance $m_2$, third-order central moment $m_3$, and fourth-order central moment $m_4$. Moment constraints of the pair of variables $\boldsymbol{\theta}_x$ to realize the staircase random variable $x$ conditional on its support set $[\underline{x}, \overline{x}]$ are given as a series of inequalities: $\Theta = \{\boldsymbol{\theta}_x : g(\boldsymbol{\theta}_x) \leq 0\}$ (Sharma et al., 2009; Kumar, 2002), and are given in Table 3.1.

Table 3.1 Moment constraints of the pair of variables $\boldsymbol{\theta}_x$.

| | Moment constraints | |
|---|---|---|
| Mean $\mu$ | $g_1 = \underline{x} - \mu$ | $g_2 = \mu - \overline{x}$ |
| Variance $m_2$ | $g_3 = -m_2$ | $g_4 = m_2 - v^{\text{a}}$ |
| Third-order central moment $m_3$ | $g_5 = m_2^2 - m_2(\mu - \underline{x})^2 - m_3(\mu - \underline{x})$ | $g_6 = m_3(\overline{x} - \mu) - m_2(\overline{x} - \mu)^2 + m_2^2$ |
| | $g_7 = 4m_2^3 + m_3^2 - m_2^2(\overline{x} - \underline{x})^2$ | $g_8 = 6\sqrt{3}m_3 - (\overline{x} - \underline{x})^3$ |
| | $g_9 = -6\sqrt{3}m_3 - (\overline{x} - \underline{x})^3$ | |
| Fourth-order central moment $m_4$ | $g_{10} = -m_4$ | $g_{11} = 12m_4 - (\overline{x} - \underline{x})^4$ |
| | $g_{12} = (m_4 - vm_2 - u^{\text{a}}m_3)(v - m_2) + (m_3 - um_2)^2$ | $g_{13} = m_3^2 + m_2^3 - m_4 m_2$ |

$^{\text{a}}\ u = \underline{x} + \overline{x} - 2\mu$ and $v = (\mu - \underline{x})(\overline{x} - \mu)$.

Considering that the chosen support set $[\underline{x}, \overline{x}]$ is partitioned into $n_b$ subintervals of equal length $\kappa = (\overline{x} - \underline{x})/n_b$,

the staircase density function $f_x(x)$ is expressed as:

$$f_x(x) = \begin{cases} l_i & \forall x \in (x_i, x_{i+1}], \text{for } 1 \leq i \leq n_b \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

where $l_i$ indicates the staircase density height of the $i$th bin; $x_i = \underline{x} + (i-1)\kappa$ is the partitioning point of the $i$th bin. Note that, $l_i$ satisfies that $l_i \geq 0$ for all bins and $\kappa \sum_{i=1}^{n_b} l_i = 1$. The staircase density heights $l$ conditional on $\Theta$ can be determined by solving a following convex optimization problem:

$$\hat{l} = \underset{l \geq 0}{\text{argmin}} \left\{ J(l) : \sum_{i=1}^{n_b} \int_{x_i}^{x_{i+1}} x l_i dx = \mu, \sum_{i=1}^{n_b} \int_{x_i}^{x_{i+1}} (x-\mu)^r l_i dx = m_r, r = 2,3,4 \right\} \tag{3.7}$$

where $J$ denotes an arbitrary cost function. Equation (3.7) can be written as:

$$\hat{l} = \underset{l \geq 0}{\text{argmin}} \{ J(l) : \mathbf{A}(\boldsymbol{\theta}_x, n_b)l = \boldsymbol{b}(\boldsymbol{\theta}_x), \boldsymbol{\theta}_x \in \Theta \} \tag{3.8}$$

where

$$\mathbf{A} = \begin{bmatrix} \kappa\mathbf{e} \\ \kappa\mathbf{c} \\ \kappa\mathbf{c}^2 + \kappa^3/12 \\ \kappa\mathbf{c}^3 + \kappa^3\mathbf{c}/4 \\ \kappa\mathbf{c}^4 + \kappa^3\mathbf{c}^2/2 + \kappa^5/80 \end{bmatrix}, \text{and } \boldsymbol{b} = \begin{bmatrix} 1 \\ \mu \\ \mu^2 + m_2 \\ m_3 + 3\mu m_2 + \mu^3 \\ m_4 + 4m_3\mu + 6m_2\mu^2 + \mu^4 \end{bmatrix}$$

where $\mathbf{c}$ denotes a column vector of the centre of the bin $c_i = (x_i + x_{i+1})/2$; $\mathbf{c}^n$ is the component wise $n$th power of $\mathbf{c}$; $\mathbf{e}$ is a unit vector.

Regarding with the cost function $J$, several optimality criteria, including the maximal entropy, minimal squared amplitude, and maximal log-likelihood can be employed. The cost function used in this study is expressed as:

$$J(l) = l^T \mathbf{I} l \tag{3.9}$$

where $\mathbf{I}$ denotes the identity matrix. Employing this cost function yields a staircase random variable that minimizes the squared sum of the likelihood at the bins. Note that, while we do not investigate other cost functions, the choice of the cost function may affect the quality of model updating and thus it should be further investigated in the future work.

The convexity of the optimization problem in Equation (3.7) can very efficiently calculate the staircase densities and to solve for practically smooth probability densities. These features make the staircase random variables well suited for the stochastic model updating in which its repeated calculation is required. Moreover, the staircase random variables are independent of the distribution formats and can describe a broad range of density shapes, including very skewed

or/and multimodal distributions. This fulfills the expectation as a non-parameterized model for the proposed updating framework, where the epistemic uncertainty space is calibrated without prior knowledge about the distribution formats of the parameters.

## 3.3    Nonparametric approximate Bayesian computation

### 3.3.1    Bayesian model updating with staircase random variables

In this study, the well-known Bayesian inference is employed as the stochastic model updating methodology. The Bayesian inference is based on the Bayes' theorem (Beck and Katafygiotis, 1998):

$$P(\boldsymbol{\theta}|\mathbf{Y}_{\text{obs}}) = \frac{P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{Y}_{\text{obs}})} \tag{3.10}$$

where $P(\boldsymbol{\theta})$ indicates the prior distribution of the adjustable parameters $\boldsymbol{\theta}$, that is determined by prior information of the system and the expert knowledge; $P(\boldsymbol{\theta}|\mathbf{Y}_{\text{obs}})$ is the posterior distribution of $\boldsymbol{\theta}$, representing the updated knowledge of $\boldsymbol{\theta}$ based on the observations $\mathbf{Y}_{\text{obs}}$; $P(\mathbf{Y}_{\text{obs}})$ means the normalized factor (also known as the evidence) ensuring that the posterior distribution integrates to one; $P_L(\mathbf{Y}_{\text{obs}}|\mathbf{x})$ is the likelihood function of $\mathbf{Y}_{\text{obs}}$ for an instance of $\boldsymbol{\theta}$.

For the case where the parameters with hybrid uncertainties are represented as the staircase random variables, the first four moments $\boldsymbol{\theta}_x$ are considered as the adjustable parameters $\boldsymbol{\theta}$. Given the support set $[\underline{x}, \overline{x}]$, feasible intervals of $\boldsymbol{\theta}_x$ can be defined based on the moment constraints $\boldsymbol{\theta}_x \in \Theta$ as:

$$\mu \in [\underline{x}, \overline{x}], m_2 \in \left[0, \frac{(\overline{x} - \underline{x})^2}{4}\right], m_3 \in \left[-\frac{(\overline{x} - \underline{x})^3}{6\sqrt{3}}, \frac{(\overline{x} - \underline{x})^3}{6\sqrt{3}}\right], m_4 \in \left[0, \frac{(\overline{x} - \underline{x})^4}{12}\right] \tag{3.11}$$

The prior distribution of $\boldsymbol{\theta}_x$ is then known point-wise for any generated samples of $\boldsymbol{\theta}_x$ within the feasible intervals that satisfy the moment constraints. In this manner, only the support sets of the parameters are required as prior information of the system, however the feasible intervals in Equation (3.11) are possible to be narrower if more detailed information about the parameters is available. At the same time, the parameters with only epistemic uncertainty are also capable to be handled in Bayesian updating, where the parameters themselves are considered as the adjustable parameters. In this case, the prior distribution $P(\boldsymbol{\theta}) = P(\mathbf{x})$ is represented as an auxiliary uniform distribution on the given support set of the parameters.

One non-trivial component in Equation (3.10) is the evidence $P(\mathbf{Y}_{\text{obs}})$, since the direct evaluation of the posterior

PDF over the whole parameter space is quite difficult or even intractable especially for very peaked or/and multimodal distributions (Beck and Au, 2002). Thus, a well-known Bayesian inference algorithm, called transitional Markov chain Monte Carlo (TMCMC) (Ching and Cheng, 2007), is employed as an effective updating tool. TMCMC is essentially interpreted as an iterative approach sampling from a series of intermediate PDFs which will progressively converge to the true posterior distribution. The $j$th intermediate PDF is expressed as:

$$P_j \propto P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta})^{\beta_j} P(\boldsymbol{\theta}) \tag{3.12}$$

where $\beta_j$ means the so-called reduction coefficient. Its value starts from $\beta_0 = 0$ in the first iteration and progressively increases until reaching $\beta_m = 1$ in the last iteration. $\beta_j$ is adaptively computed using samples generated at the previous step. Markov chains with the Metropolis-Hasting algorithm (Hasting, 1970) propagate new samples starting from the ones with higher intermediate likelihood values, enabling to sample from the very complex posterior PDF. The readers can be referred to Ching and Cheng (2007) and Betz et al. (2016) for the details of TMCMC and to Rocchetta et al. (2018) for its application.

### 3.3.2 Two-step ABC updating framework

The likelihood function is one of the key components in Bayesian model updating, because it quantifies the degree of relevance of a model with a given instance of the adjustable parameters, by describing the possibility of the observed data. Under the assumption of independence between observations, the likelihood function in Equation (3.10) can be theoretically defined as:

$$P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta}) = \prod_{k=1}^{N_{\text{obs}}} P(\mathbf{Y}_k|\boldsymbol{\theta}) \tag{3.13}$$

where $P(\mathbf{Y}_k|\boldsymbol{\theta})$ indicates the PDF value of the $k$th observed data $\mathbf{Y}_k$ conditional to the corresponding instance of the adjustable parameters $\boldsymbol{\theta}$. Equation (3.13) requires to estimate the PDF for each of the $N_{\text{obs}}$ observations, that introduce considerable computation cost. Moreover, precise estimation of the PDFs is only achieved by a large number of model evaluations to generate a large number of simulated features. Therefore, the full likelihood evaluation could be almost infeasible for complex simulators.

The ABC method (Turner and Van Zandt, 2012; Safta et al., 2015) is utilized to overcome the above obstacle by replacing the above full likelihood function with an approximate likelihood function that contains information of both

the observations and the adjustable parameters $\boldsymbol{\theta}$. In the approximate likelihood, any types of statistics can be used to measure the discrepancy between the model predictions and observations, and hence, it is natural to define it employing the distance metrics. Various functional formulas have been investigated in the literature, such as the Gaussian (Patelli et al., 2015), the Epanechnikov (Safta et al., 2015), and the sharp (Rocchetta et al., 2018) functions. Regardless of the functional formula utilized, the basic principle of the approximate likelihood is that it should return a high value when the distance metric is small, while it penalizes the $\boldsymbol{\theta}$ instance when its corresponding distance metric is large. In this study, the approximate likelihood based on the Gaussian function is proposed as:

$$P_L(\mathbf{Y}_{\mathrm{obs}}|\boldsymbol{\theta}) \propto exp\left\{-\frac{d^2}{\varepsilon^2}\right\} \tag{3.14}$$

where $d$ is the distance metric; $\varepsilon$ denotes the so-called width factor, which is a pre-defined coefficient controlling the centralization degree of the posterior distribution. A smaller $\varepsilon$ corresponds to a more peaked posterior distribution, that is more likely to converge to the true value but requires more calculation for convergence. The choice of $\varepsilon$ is thus based on specific applications and is usually between $10^{-3}$ and $10^{-1}$ (Patelli et al., 2017). The distance-based approximate likelihood in Equation (3.14) is a convenient connection between the distance metrics and Bayesian model updating with significantly reduced calculation cost. Moreover, it provides a uniform framework for either the deterministic and stochastic updating, simply driven by the employed metric is the Euclidian or Bhattacharyya distances.

By employing the Bhattacharyya distance metric, the proposed approximate likelihood is capable of quantifying comprehensive uncertainty characteristics of both the model predictions and observations. Furthermore, thanks to the features of the staircase random variable, the stochastic updating procedure is theoretically applicable regardless of the distribution formats of the parameters, including the one cannot be defined analytically. However, the multimodality of the parameters, for instance, can lead to non-unique solutions (which will be further discussed in Section 3.4 through an illustrative example), and thus the direct application of the stochastic updating procedure cannot be utilized.

To cope with this issue, a two-step ABC updating framework is proposed as shown in Figure 3.1. This framework starts from performing an outer Bayesian updating using the Euclidian distance. Step I is equivalent to a deterministic updating procedure with the target for identifying all the possible solutions. The support sets of the parameters are then divided into appropriate sub-intervals, so that each divided sub-interval contain one possible solution. This preliminary procedure is necessary for avoiding to end with a local solution in the main step. After that, comprehensive uncertainty characteristics of the parameters are further updated in step II via stochastic updating of the epistemic parameters, i.e., the first four moments of the staircase random variables, using the Bhattacharyya distance.
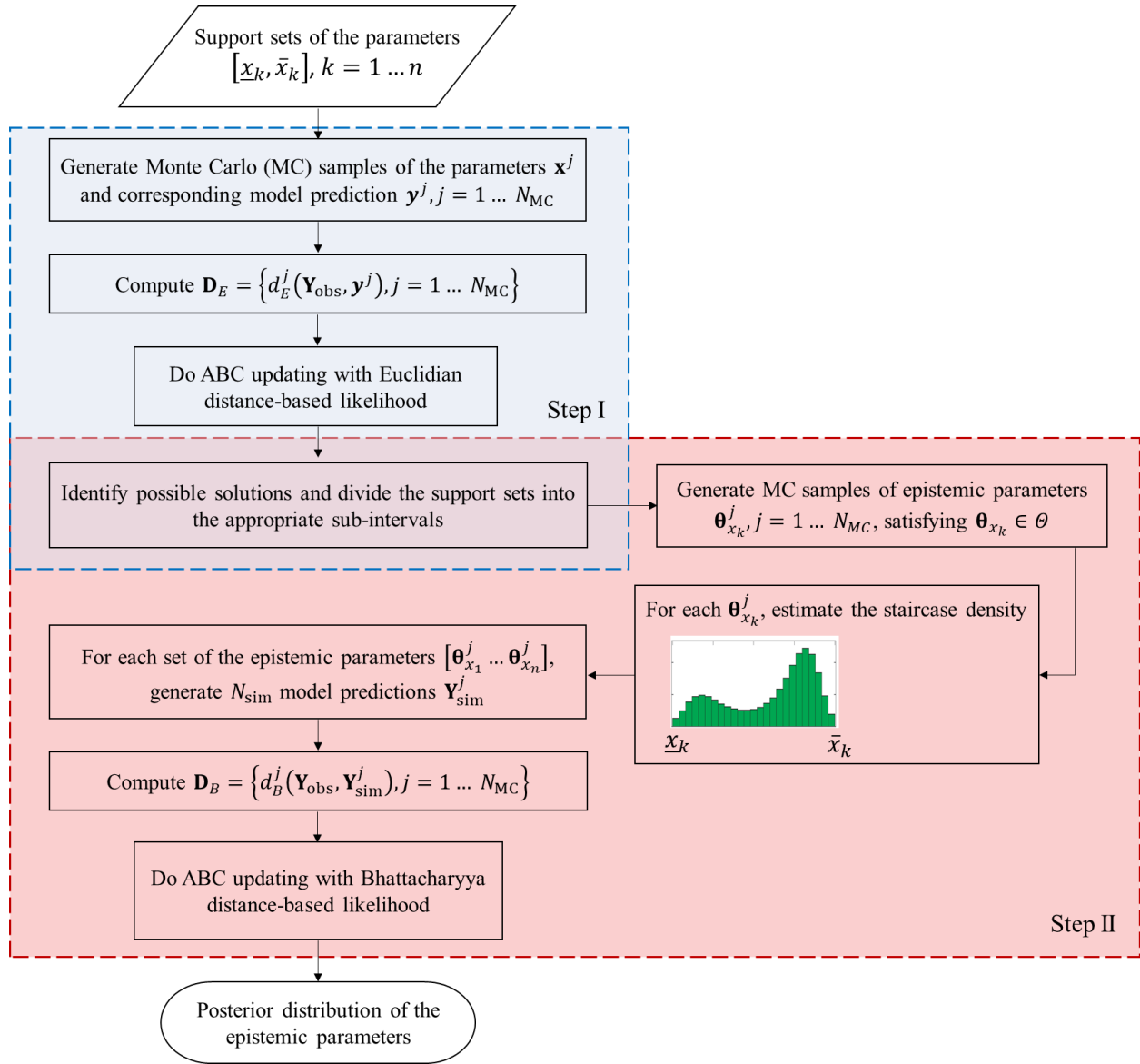
Figure 3.1 Schematic of the two-step ABC updating framework.

## 3.4 Principle and illustrative application of the ABC updating framework

### 3.4.1 Problem description

The proposed two-step ABC updating framework with both the Euclidian and Bhattacharyya distance metrics is demonstrated upon a two degree of freedom (DOF) shear building model shown in Figure 3.2(a). This model was first introduced by Beck and Au (2002). The first and second story masses are considered as deterministic values with $m_1 = 16.531 \times 10^3$ kg and $m_2 = 16.131 \times 10^3$ kg. The first and second interstory stiffnesses are parameterized to be $k_1 =$

$\bar{k}x_1$ and $k_2 = \bar{k}x_2$, where $\mathbf{x} = [x_1, x_2]$ are the adjustable parameters to be identified, and $\bar{k} = 29.7 \times 10^6$ N/m is the nominal value of the stiffnesses.

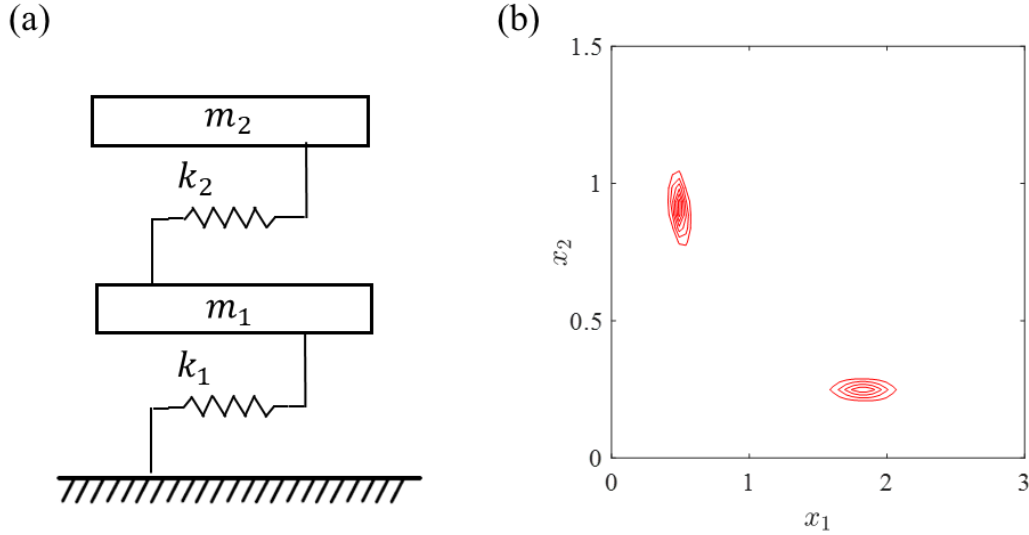(a)                                    (b)



Figure 3.2 (a) Two degree of freedom shear building model; (b) Posterior distribution in Beck and Au (2002).

In Beck and Au (2002), the prior PDF $P(\mathbf{x})$ is given by an uncorrelated lognormal distribution with most probable values 1.3 and 0.8 for $x_1$ and $x_2$, respectively, and unit standard deviations. Using the modal data, where the identified natural frequencies are $\tilde{f}_1 = 3.13$ Hz and $\tilde{f}_2 = 9.83$ Hz, the posterior PDF is formulated as:

$$P(\mathbf{x}|\mathbf{Y}_{\mathrm{obs}}) = exp\left[-\frac{J(\mathbf{x})}{2\sigma^2}\right] P(\mathbf{x}) \tag{3.15}$$

where $\sigma = 1/16$ is the standard deviation of the prediction error and $J(\mathbf{x})$ is a modal measure-of-fit function given by:

$$J(\mathbf{x}) = \sum_{j=1}^{2} \lambda^2 \left[\frac{f_j^2(\mathbf{x})}{\tilde{f}_j^2} - 1\right]^2 \tag{3.16}$$

where $\lambda = 1$ is the weight and $f_j(\mathbf{x})$ is the $j$th natural frequency predicted by the model with the adjustable parameters $\mathbf{x}$. Figure 3.2(b) depicts the posterior PDF in Equation (3.15), showing the bimodality. It has been already demonstrated in the literature that this bimodal distribution can be achieved using several sampling methods, including the TMCMC method.

This problem can be interpreted as a deterministic updating of the parameters themselves using the single set of observations. However, its uncertain characteristics and observation data are hereby altered to demonstrate capabilities of the proposed stochastic updating framework with the presence of hybrid uncertainties. Both aleatory and epistemic

uncertainties are involved in the model and are included by describing $x_1$ and $x_2$ as the staircase random variables with given support sets. The natural frequencies $f_1$ and $f_2$ are taken as investigated features whose uncertainty is driven by the uncertain parameters $x_1$ and $x_2$. Their target probability distribution is assumed to be the posterior PDF in Equation (3.15). In the altered problem setting, the identified natural frequencies $\tilde{f}_1$ and $\tilde{f}_2$ are not available, and hence the target probability distribution cannot be defined analytically. It is important to note that, in such case, the existing parametric stochastic updating procedures are not applicable. The support sets of $x_1$ and $x_2$ are detailed in Table 3.2.

Table 3.2 Uncertain parameters of 2-DOF model.

| Parameter | Uncertainty characteristic | Target distribution |
|---|---|---|
| $x_1$ | $x_1 \in [0, 3.0]$ | The marginal distribution of Equation (3.15) for $x_1$ |
| $x_2$ | $x_2 \in [0, 1.5]$ | The marginal distribution of Equation (3.15) for $x_2$ |

The target of the updating procedure $\mathbf{Y}_{\text{obs}}$ is multiple sets of the features $f_1$ and $f_2$ obtained by assigning the target probability distribution to $x_1$ and $x_2$, as detailed in the last column of Table 3.2. The number of observations is $N_{obs} = 100$, generated by evaluating the model 100 times with the parameters sampled from their assigned target distribution using the TMCMC method.

A single set of initial values of the first four moments $\boldsymbol{\theta}_{x_1}$ and $\boldsymbol{\theta}_{x_2}$ is set as their possible realizations within the feasible intervals in Equation (3.11), satisfying the moment constraints $\boldsymbol{\theta}_x \in \Theta$. These values are selected as different from the target values, which result in the staircase random variables matching with the target distribution, in order to illustrate how the imprecise model could produce outputs very different from the observations. Note that, these initial values are presented herein only for demonstration purpose as illustrated in Figures 3.3 and 3.9. The two-step updating procedure is not really started from these initial values, but from the initial support sets of the parameters, as shown in the second column of Table 3.2.

Suppose the sample size is $N_{\text{sim}} = 1000$, $N_{\text{sim}}$ parameter samples are generated from the staircase densities with the initial values of the first four moments. The corresponding initial simulated output samples of $f_1$ and $f_2$ are obtained and illustrated in Figure 3.3, together with the observed output samples. As shown in this figure, the objective of model updating herein is no longer a single updated point with the maximum fidelity to a single observation point, but updated distributions of the parameters which can represent the output samples as similar as the observed ones. Such parameter distributions can be estimated as the staircase densities with the updated first four moments. To achieve this objective, both the Euclidian and Bhattacharyya distances are employed as the UQ metrics to define approximate likelihoods in the ABC updating procedure.
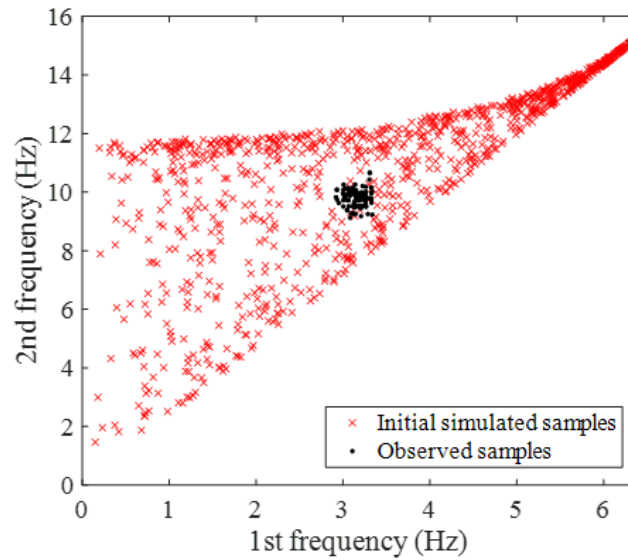
Figure 3.3 Observed and initial simulated output samples.

## 3.4.2    Step I: deterministic updating with the Euclidian distance metric

As shown in Figures 3.2(b) and 3.3, the target distribution of the parameters demonstrates the bimodality but the observed samples show the unimodality; thus, it leads to non-unique solutions. In such situation, the direct application of the stochastic updating procedure may end with a local solution. To avoid the local solution, a deterministic updating of the parameters themselves is herein performed with the Euclidian distance metric to identify the possible solutions. There are two parameters $x_1$ and $x_2$, whose prior distribution is set to be uniform within their support set in Table 3.2. When the Euclidian distance is taken as the metric, the geometric distance between a simulated sample and the centre of mass of the observed samples is measured, whereas the dispersion information of the observed samples cannot be considered.

In this section, the width factor in the distance-based likelihood is assumed as $\varepsilon = 0.1$, and totally eight TMCMC iterations are executed to reach convergence. Posterior samples generated are presented in Figure 3.4, along with the target distribution. As shown in this figure, the posterior samples clearly capture the bimodality of the target distribution as same as the result in Beck and Au (2002). It fulfils that the deterministic updating focuses only on the means of the parameters and they are obviously much easier to be properly calibrated compared with the higher moments. However, the orientation and dispersion of the posterior samples remain difference from the target distribution. Hence, a more comprehensive metric is required in the second step to further reduce the discrepancy between the samples in Figure 3.4 and the target distribution.
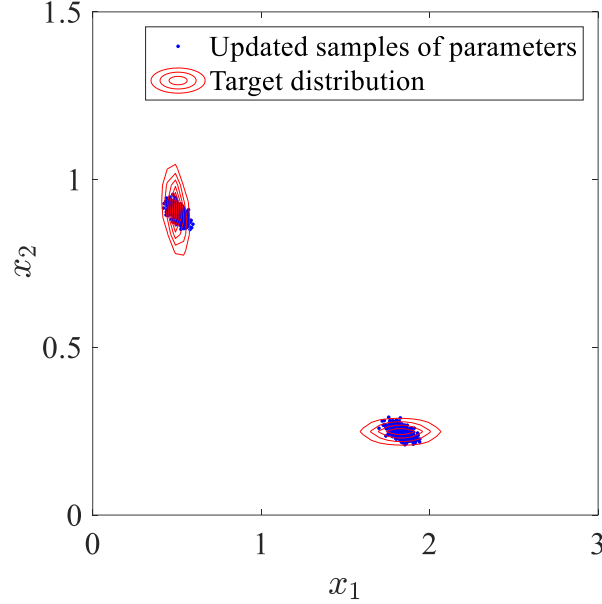
Figure 3.4 Updated samples of parameters with the Euclidian distance.

Based on the posterior distribution of the parameters, their support set is divided into two sub-intervals, such that each sub-interval contains single possible solution. Table 3.3 presents the defined sub-intervals of the parameters.

Table 3.3 Sub-intervals of parameters for stochastic updating.

| Sub-interval | Uncertain characteristic | Epistemic parameters |
|---|---|---|
| I | $x_1 \in [0, 1.0]$ | $\mu_1 \in [0, 1.0], m_{21} \in [0, 0.25], m_{31} \in \left[-\dfrac{1}{6\sqrt{3}}, \dfrac{1}{6\sqrt{3}}\right], m_{41} \in \left[0, \dfrac{1}{12}\right]$ |
| | $x_2 \in [0.5, 1.5]$ | $\mu_2 \in [0.5, 1.5], m_{22} \in [0, 0.25], m_{32} \in \left[-\dfrac{1}{6\sqrt{3}}, \dfrac{1}{6\sqrt{3}}\right], m_{42} \in \left[0, \dfrac{1}{12}\right]$ |
| II | $x_1 \in [1.0, 3.0]$ | $\mu_1 \in [1.0, 3.0], m_{21} \in [0, 1.0], m_{31} \in \left[-\dfrac{4}{3\sqrt{3}}, \dfrac{4}{3\sqrt{3}}\right], m_{41} \in \left[0, \dfrac{4}{3}\right]$ |
| | $x_2 \in [0, 0.5]$ | $\mu_2 \in [0, 0.5], m_{22} \in \left[0, \dfrac{1}{16}\right], m_{32} \in \left[-\dfrac{1}{48\sqrt{3}}, \dfrac{1}{48\sqrt{3}}\right], m_{42} \in \left[0, \dfrac{1}{192}\right]$ |

### 3.4.3 Step II: stochastic updating with the Bhattacharyya distance metric

This section presents the stochastic updating procedure with the Bhattacharyya distance metric. There are in total eight epistemic parameters in the updating procedure, i.e., $\boldsymbol{\theta}_{x_i} = \{\mu_i, m_{2i}, m_{3i}, m_{4i}\}$, for $i = 1, 2$, feasible intervals of which are computed by Equation (3.11) for each sub-interval, as summarized in the last column of Table 3.3. The prior distribution of $\boldsymbol{\theta}_{x_1}$ and $\boldsymbol{\theta}_{x_2}$ is discretely known by multiple sets of their realizations within the feasible intervals which satisfy the moment constraints $\boldsymbol{\theta}_x \in \Theta$.

The width factor in the likelihood is set as $\varepsilon = 0.01$ in this section. In addition, the numbers of bins in the binning algorithm and in staircase density estimation are set as $n_{bin} = 10$ and $n_b = 50$, respectively. After totally 16 TMCMC

iterations, the finally updated posterior histograms of the epistemic parameters are estimated for each sub-interval of the parameters, as illustrated in Figures 3.5 and 3.6. In the figures, the third and fourth central moments are normalized as $\tilde{m}_3 = m_3/m_2^{3/2}$ and $\tilde{m}_4 = m_4/m_2^2$ (also known as the skewness and kurtosis, respectively). In addition, the target and updated values of the epistemic parameters are given in these figures. The target values are computed by samples generated from the target distribution in Figure 3.2(b), and the updated values are obtained by estimating most probable values (MPVs) of the posterior distribution. As shown in these figures, the posterior histograms of all of the epistemic parameters are significantly updated compared with the prior feasible intervals. Furthermore, the updated values show almost good agreement with the target values, except for the kurtoses, implying that the Bhattacharyya distance metric is capable of capturing not only mean information but also dispersion and distribution information of both the model predictions and observations.

Tables 3.4 and 3.5 summarize the target and updated values of the epistemic parameters. The updated values of the means and variances are quite close to their targets and those of the skewnesses are also almost close to their targets, even though those of the kurtoses still remain differences compared with the targets. This fulfils the general experience in the stochastic updating that the higher level of statistical information is much more difficult to be precisely updated compared with the means. However, the proposed procedure is capable to quantify even the higher level of statistical information, such as the variances and skewnesses.
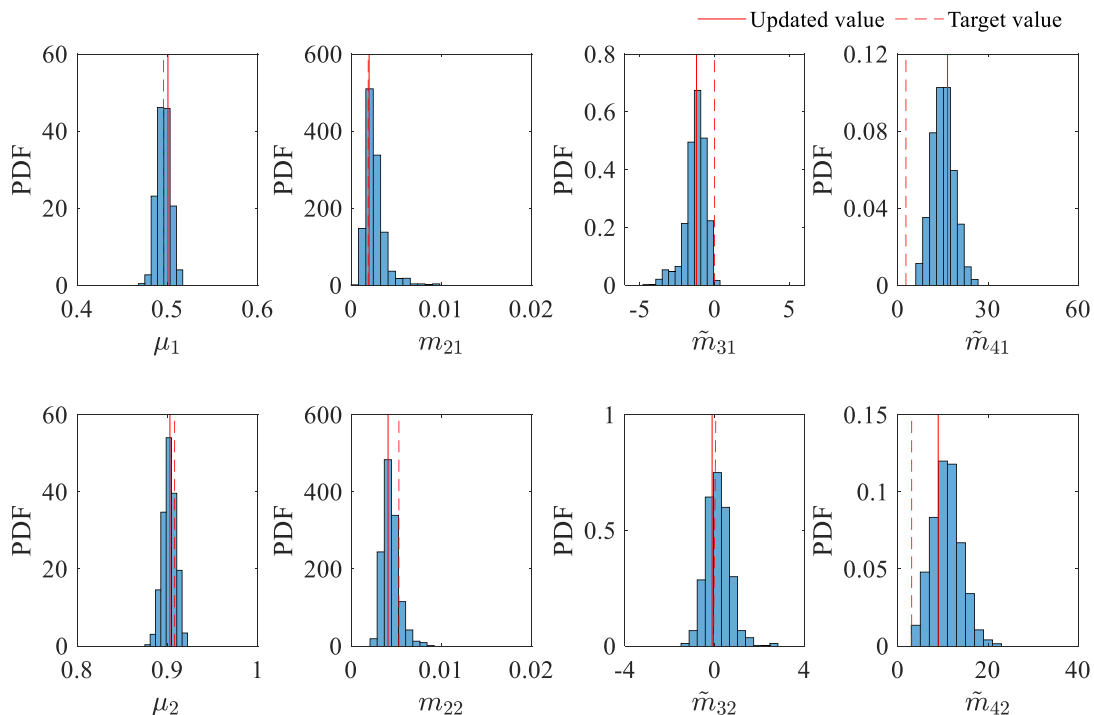


Figure 3.5 Posterior distributions of epistemic parameters for sub-interval I.
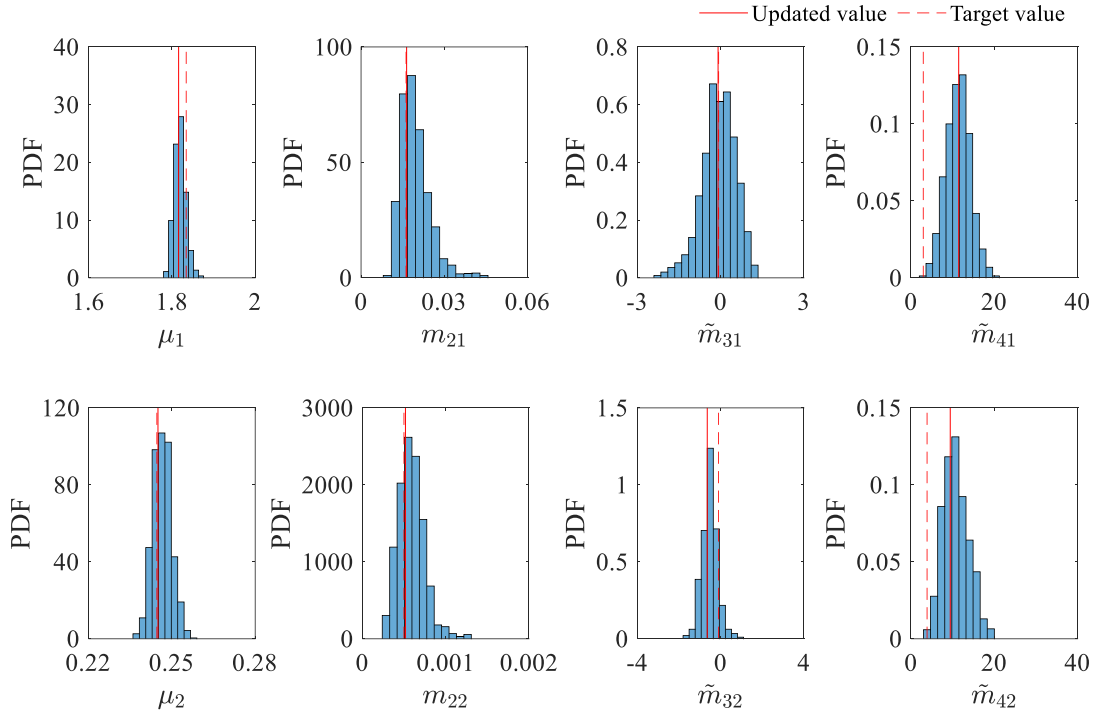
Figure 3.6 Posterior distributions of epistemic parameters for sub-interval II.

Table 3.4 Updated epistemic parameters for sub-interval I.

| Parameter | Mean $\mu$ | Variance $m_2$ | Skewness $\widetilde{m}_3$ | Kurtosis $\widetilde{m}_4$ |
|---|---|---|---|---|
| $x_1$ | | | | |
| Target value | 0.4957 | 0.0019 | -0.0086 | 2.7313 |
| Updated value | 0.5006 | 0.0020 | -1,2150 | 16,550 |
| $x_2$ | | | | |
| Target value | 0.9080 | 0.0053 | -0.0531 | 3.0972 |
| Updated value | 0.9028 | 0.0041 | -0,1150 | 8,9700 |

Table 3.5 Updated epistemic parameters for sub-interval II.

| Parameter | Mean $\mu$ | Variance $m_2$ | Skewness $\widetilde{m}_3$ | Kurtosis $\widetilde{m}_4$ |
|---|---|---|---|---|
| $x_1$ | | | | |
| Target value | 1.8344 | 0.0162 | 0.0695 | 2.9410 |
| Updated value | 1.8162 | 0.0164 | -0.0900 | 11.450 |
| $x_2$ | | | | |
| Target value | 0.2446 | 0.0005 | -0.0984 | 3.8560 |
| Updated value | 0.2450 | 0.0005 | -0.6400 | 9,4200 |

The updated samples of the parameters are generated from the staircase densities with the updated values of the first four moments and are illustrated in Figure 3.7. It can be seen that the updated epistemic parameters for each sub-interval lead to the staircase densities providing accurate estimation of each mode of the parameters. More attention is paid to the orientation and dispersion of the updated samples, which show good agreement with the target distribution, although some higher moment values are not precisely updated, as shown in Tables 3.4 and 3.5. It implies that estimated

errors in the higher moment values do not significantly affect the sample distributions of each mode of the parameters obtained via the staircase densities, compared with the means and variances. For comparison purpose, the cumulative distribution functions (CDFs) of $x_1$ and $x_2$ are plotted in Figure 3.8 for both the updated and target distributions. The CDFs of the updated distributions are determined by combining the updated samples for both sub-intervals. The CDFs of the updated distributions exhibit a good match with those for the target distributions.

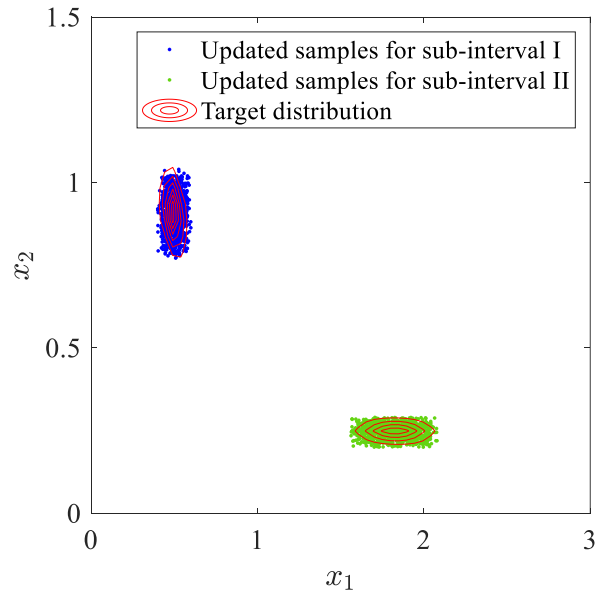

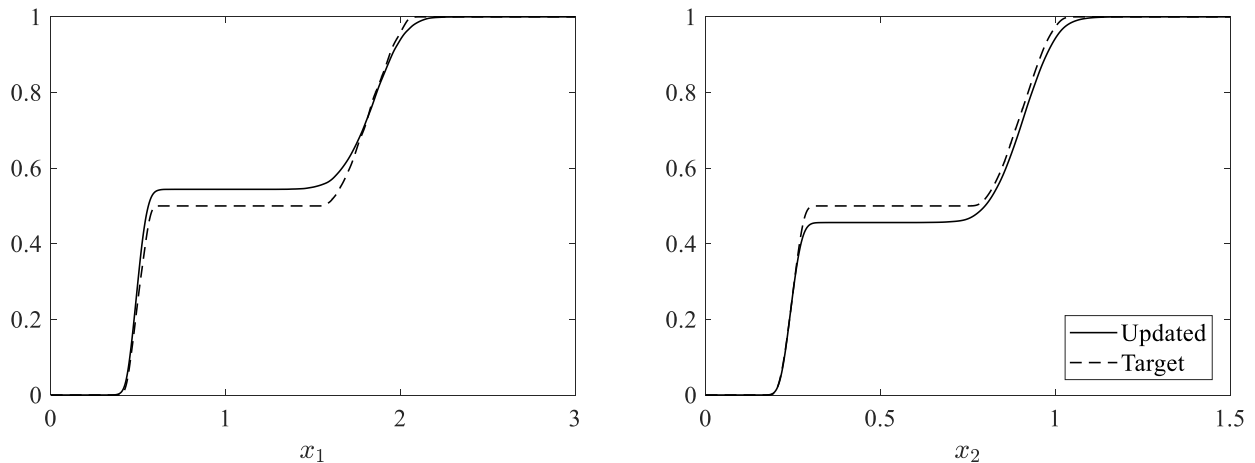Figure 3.7 Updated samples of parameters with the Bhattacharyya distance.



Figure 3.8 Updated cumulative distribution functions of parameters.

Finally, Figure 3.9 illustrates the final simulated output samples of $f_1$ and $f_2$, obtained by assigning the estimated staircase densities to $x_1$ and $x_2$, together with the initial simulated and observed output samples. The updated simulated

samples demonstrate a distribution identical to the target observed samples, implying that the Bhattacharyya distance

has the capability to recreate wholly the distribution of the target observations.
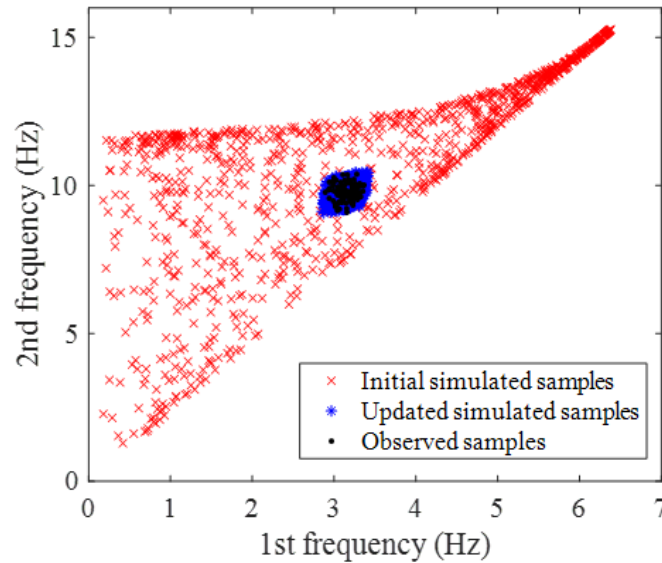


Figure 3.9 Updated simulated output samples.

### 3.4.4 Summary

This example presented the combined application of the Euclidian and Bhattacharyya distances as the metrics in

the two-step ABC updating procedure. The staircase random variables are employed in stochastic updating to calibrate

the epistemic uncertainty space without prior information about the distribution formats of parameters. The stochastic

model updating procedure with the Bhattacharyya distance metric is demonstrated to be capable to estimate the detailed

distributional properties of the parameters. However, a significant drawback of the updating procedure is revealed in

situations, e.g., in Figures 3.2(b) and 3.3, where the target distribution denotes the bimodality but the observed features

demonstrate the unimodality. In this situation, the direct application of the stochastic model updating procedure might

end with a local solution.

As a consequence, the two-step procedure is proposed to overcome this drawback by performing a deterministic

updating with the Euclidian distance metric in the first step to identify every modes of the target distribution, and then

to quantify their detailed distributional information in the second step. As a result, this example demonstrated that the

deterministic updating should be performed as a precondition of any stochastic updating procedure to avoid non-unique

solutions.

## 3.5    NASA UQ challenge problem 2014

### 3.5.1  Problem description

The NASA UQ challenge problem 2014 (Crespo et al., 2014) is investigated herein to demonstrate the capability

of the proposed framework for complex applications. The schematic in Figure 3.10 illustrates the general structure of

Sub-problem A, including the investigated parameters, outputs, and UQ metric. As shown in Figure 3.10, the simulator

is provided in a black-box, which evaluates a scalar output $y$ using five parameters: $p_i$, for $i = 1, \cdots, 5$.
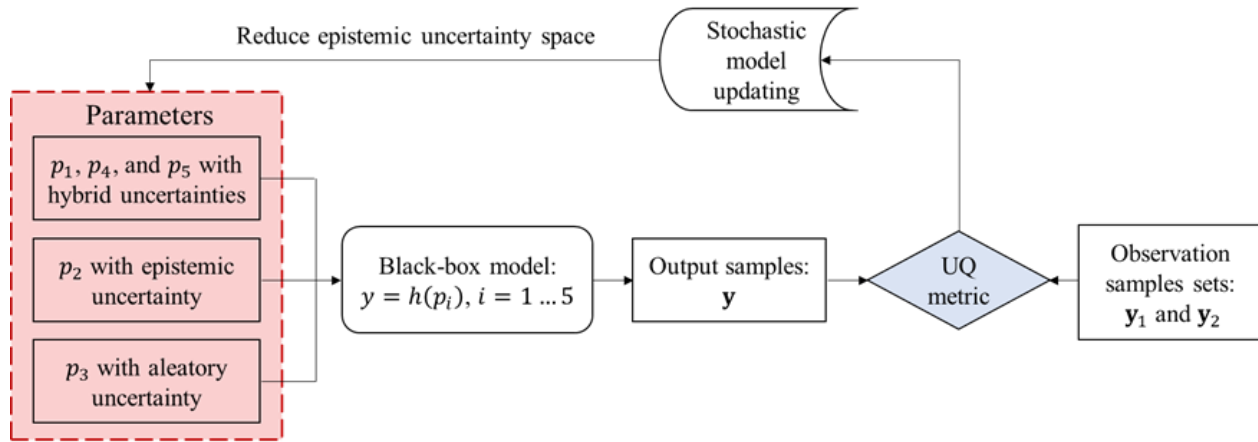


Figure 3.10 Schematic of the NASA UQ challenge 2014 Sub-problem A.

Table 3.6 summarizes the uncertainty characterizations of the parameters in the original problem setting. $p_1$, $p_4$,

and $p_5$ indicate parameters with hybrid uncertainties, $p_2$ refers to a parameter with only epistemic uncertainty, and $p_3$

is a parameter with only aleatory uncertainty represented by a fully prescribed uniform distribution with explicit mean

and variance. During model updating, only $p_1$, $p_2$, $p_4$, and $p_5$ are considered and $p_3$ is omitted because it involves only

irreducible aleatory uncertainty. More importantly, the distribution formats of $p_1$, $p_4$, and $p_5$ are fully provided, such

that $p_1$ follows a unimodal beta distribution and $p_4$ and $p_5$ follow Gaussian distributions. Recently, true values of the

epistemic parameters are released as shown in the last column of Table 3.6.

Table 3.6 Uncertainty parameters of Sub-problem A in the NASA UQ challenge problem.

| Parameter | Uncertainty characteristic | True value |
|---|---|---|
| $p_1$ | Unimodal beta, $\mu_1 \in [0.6, 0.8]$, $m_{21} \in [0.02, 0.04]$ | $\mu_1 = 0.6364$, $m_{21} = 0.0356$ |
| $p_2$ | Constant, $p_2 \in [0, 1.0]$ | $p_2 = 1$ |
| $p_3$ | Uniform, $\mu_3 = 0.5$, $m_{23} = 1/12$ | $-$ |
| $p_4, p_5$ | Gaussian, $\mu_i \in [-5.0, 5.0]$, $m_{2i} \in [0.0025, 4.0]$, $\rho \in [-1.0, 1.0]^a$, $i = 4,5$ | $\mu_4 = 4$, $\mu_5 = -1.5$, $m_{24} = 0.04$, $m_{25} = 0.36$, $\rho = 0.5$ |

$^a$ $\rho$ is the correlation coefficient.

On the contrary, the uncertainty characterizations of $p_1$, $p_4$, and $p_5$ are herein altered by ignoring their distribution formats to demonstrate the proposed nonparametric updating framework. Table 3.7 presents the redefined uncertainty characteristics of these parameters. The support set of $p_1$ is determined based on the definition of the beta distribution. While only the support set is necessary for the proposed nonparametric updating framework, prior information on the mean $\mu_1$ and variance $m_{21}$ is remained as the original setting to make the feasible interval in Equation (3.11) narrower. The support sets of $p_4$ and $p_5$ are set to cover more than 99.99 % confidence intervals of the true Gaussian distributions, and the feasible intervals of the means and variances for these support sets are already narrower than the given intervals in Table 3.6. Note that, while the support sets can be also set to cover the p-boxes defined by the original uncertainty characteristics, this leads too wide support sets to be precisely updated. Moreover, the correlation between $p_4$ and $p_5$ is ignored, since the staircase random variable is a univariate random variable and thus cannot consider the correlation. However, this assumption is still reasonable because the investigated output is insensitive to the correlation coefficient $\rho$, that has been investigated by several previously published works (Patelli et al., 2015; Safta et al., 2015; Ghanem et al., 2015) in other sub-tasks of the NASA UQ challenge problem 2014.

Table 3.7 Redefined uncertainty characteristics of $p_1$, $p_4$, and $p_5$

| Parameter | Uncertainty characteristic |
|---|---|
| $p_1$ | $p_1 \in [0, 1]$, $\mu_1 \in [0.6, 0.8]$, $m_{21} \in [0.02, 0.04]$ |
| $p_4, p_5$ | $p_4 \in [3, 5]$, $p_5 \in [-4, 1]$ |

There are two observation sets $\mathbf{y}_1$ and $\mathbf{y}_2$, both containing 25 values respectively. It is noted that, in the original NASA UQ challenge, there were different tasks in Sub-problem A, where the first observations ($\mathbf{y}_1$) are supposed to be used for model updating in Task 1, and the remaining ($\mathbf{y}_2$) for model validation in Task 2 and in Task 3, all the 50 observations are used for model updating so as to improve the result. However, in this work, only Task 3 is addressed, since the comparison of the results using 25 or 50 observations is not our focus. As a consequence, in total 13 epistemic parameters, such as $\boldsymbol{\theta}_{p_i} = \{\mu_i, m_{2i}, m_{3i}, m_{4i}\}$, for $i = 1,4,5$ and $p_2$ are calibrated using the 50 observations.

## 3.5.2  Results assessment

The stochastic updating procedure with the Bhattacharyya distance metric is executed. It is important to note that, the true distributions of all the parameters are unimodal; thus, an outer Bayesian updating with the Euclidian distance metric is not performed. The width factor in the likelihood is set to be $\varepsilon = 0.01$ and the numbers of bins in the binning algorithm and in staircase density estimation are set to be $n_{\text{bin}} = 25$ and $n_b = 50$, respectively.

Figure 3.11 presents the posterior histograms of all the epistemic parameters. Those are converted to distributions with kernel density estimation (KDE) and all the estimated distributions are also illustrated in the figure. It can be seen that all the epistemic parameters are successfully updated compared with their initial intervals. The updated values and intervals of $\boldsymbol{\theta}_{p_1}$, $\boldsymbol{\theta}_{p_4}$, $\boldsymbol{\theta}_{p_5}$, and $p_2$ are obtained from the posterior distributions, and their accuracy is assessed according to their true values as shown in Table 3.8. The posterior distributions are normalized so that their maximums are equal to one, as shown in Figure 3.12. With this procedure, the posterior distributions are interpreted as Fuzzy sets (Beer et al., 2013), such that different levels of confidence will result to interval values of increased width. The crisp updated values of $\boldsymbol{\theta}_{p_1}$, $\boldsymbol{\theta}_{p_4}$, $\boldsymbol{\theta}_{p_5}$, and $p_2$ are computed as the MPVs, i.e., the values corresponding to the case where the alpha-level is one.
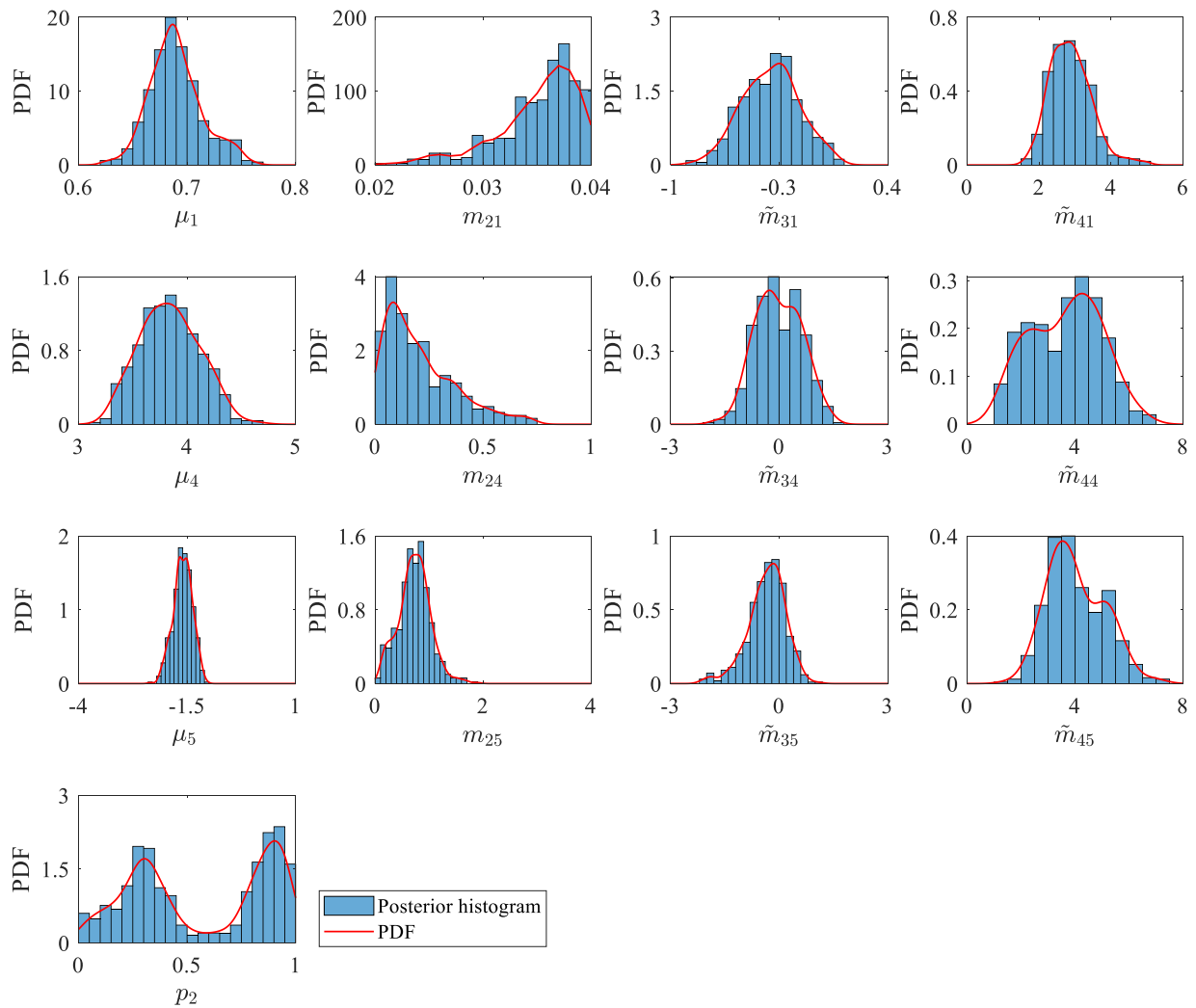


Figure 3.11 Posterior histograms and PDFs estimated via KDE.

Table 3.8 The updated results of the epistemic parameters.

| Epistemic parameter | Initial interval | True value | MPVs | 0.9-level intervals |
|---|---|---|---|---|
| $\mu_1$ | $[0.6, 0.8]$ | 0.6364 | 0.6824 | $[0.6793, 0.6933]$ |
| $m_{21}$ | $[0.02, 0.04]$ | 0.0356 | 0.0369 | $[0.0366, 0.0386]$ |
| $\widetilde{m}_{31}$ | $\left[-\dfrac{1}{6\sqrt{3}}, \dfrac{1}{6\sqrt{3}}\right]$ | -0.3840 | -0.3560 | $[-0.3963, -0.2400$ |
| $\widetilde{m}_{41}$ | $\left[0, \dfrac{1}{12}\right]$ | 2.4886 | 2.4360 | $[2.4007, 3.0787]$ |
| $\mu_4$ | $[3, 5]$ | 4 | 3.8780 | $[3.6517, 3.9517]$ |
| $m_{24}$ | $[0.0025, 1]$ | 0.04 | 0.0488 | $[0.0480, 0.1236]$ |
| $\widetilde{m}_{34}$ | $\left[-\dfrac{4}{3\sqrt{3}}, \dfrac{4}{3\sqrt{3}}\right]$ | 0.0068 | -0.2620 | $[-0.5241, 0.0499]$ |
| $\widetilde{m}_{44}$ | $\left[0, \dfrac{1}{12}\right]$ | 2.9780 | 4.0560 | $[3.7298, 4.7558]$ |
| $\mu_5$ | $[-4, 1]$ | -1.5 | -1.7000 | $[-1.7204, -1.4554$ |
| $m_{25}$ | $[0.0025, 4]$ | 0.36 | 0.7920 | $[0.6070, 0.8940]$ |
| $\widetilde{m}_{35}$ | $\left[-\dfrac{4}{3\sqrt{3}}, \dfrac{4}{3\sqrt{3}}\right]$ | 0.0068 | -0.0910 | $[-0.4172, 0.0208]$ |
| $\widetilde{m}_{45}$ | $\left[0, \dfrac{1}{12}\right]$ | 2.9780 | 3.6840 | $[3.6830, 3.9185]$ |
| $p_2$ | $[0, 1]$ | 1 | 0.9050 | $[0.8580, 0.9410]$ |
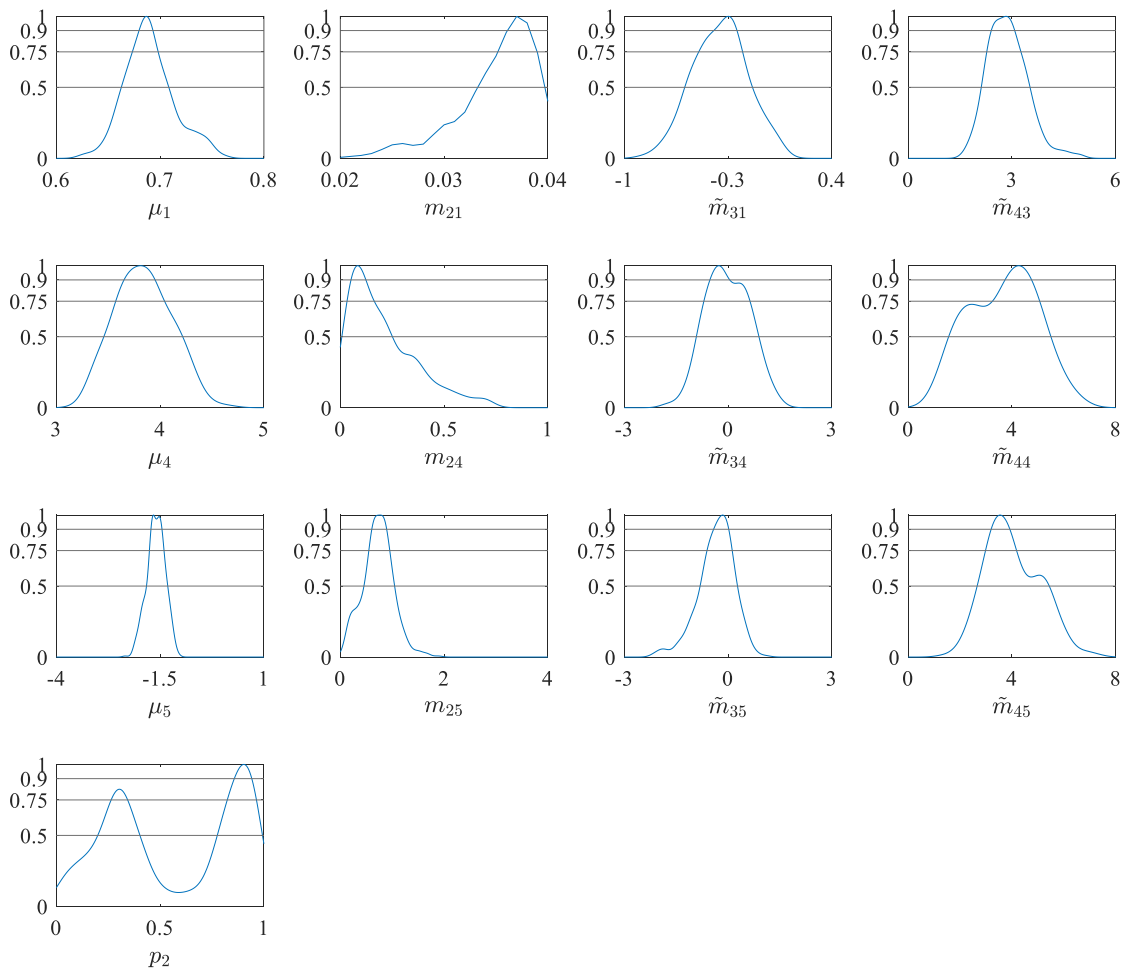


Figure 3.12 Three truncation levels of the normalized posterior distributions.

By employing the crisp updated values of $\boldsymbol{\theta}_{p_1}$, $\boldsymbol{\theta}_{p_4}$, and $\boldsymbol{\theta}_{p_5}$, the corresponding crisp updated distributions of $p_1$, $p_4$, and $p_5$ are estimated as the staircase density functions. Figure 3.13 illustrates the CDFs of the updated distributions, together with those of their true distributions. It can be seen that the updated CDFs show almost good agreement with their true distributions, while some differences still remain for $p_4$ and $p_5$. These differences are mainly because of no consideration of the correlation between $p_4$ and $p_5$. Nevertheless, the proposed updating procedure is demonstrated to be capable to estimate the probabilistic distributions of unknown parameters regardless of their distribution formats.
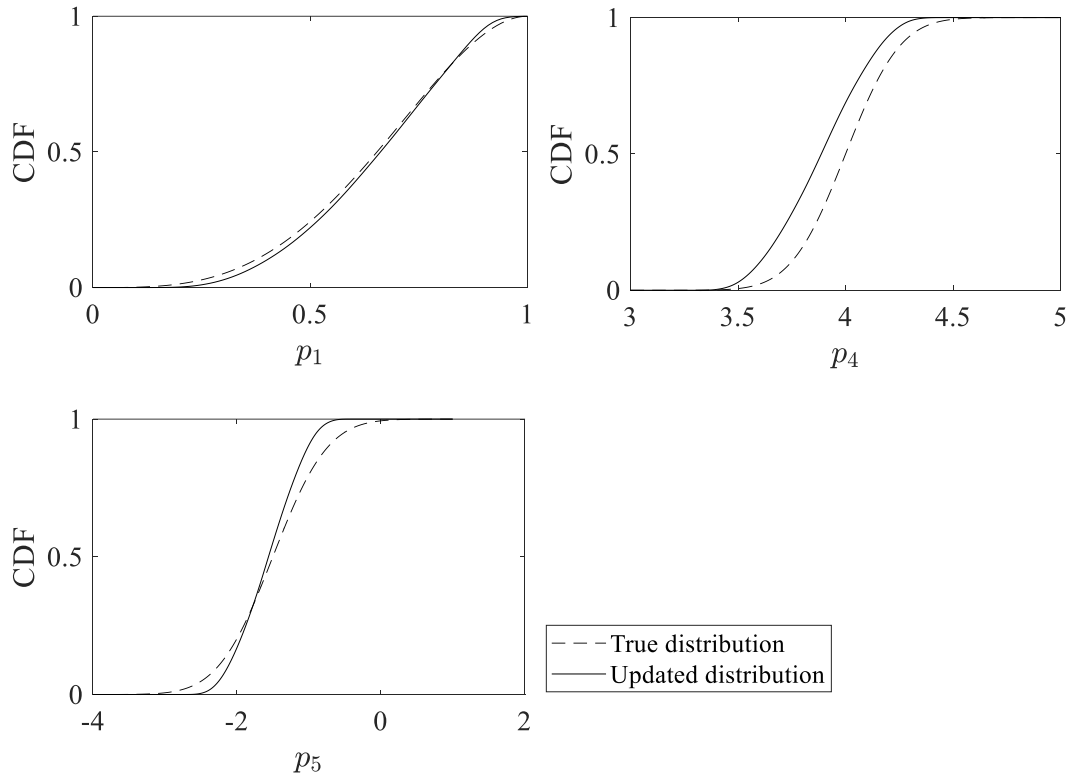


Figure 3.13 Updated cumulative distribution functions of $p_1$, $p_4$, and $p_5$.

A more comprehensive assessment of the result is performed by estimating the p-boxes of the output. The initial intervals of $\boldsymbol{\theta}_{p_1}$, $\boldsymbol{\theta}_{p_4}$, $\boldsymbol{\theta}_{p_5}$, and $p_2$ result in a large p-box of the output, representing a large epistemic uncertainty space. The objective of model updating in this problem is to reduce the epistemic uncertainty space, so that the p-box of the output is accordingly reduced. In the ideal case, when the true values of the epistemic parameters are achieved form a perfect updating process, the resulting p-box of the output would be reduced to a single CDF, which perfectly coincides with the CDF of the observations. Based on the above motivation, three alpha-levels, namely 0.5, 0.75, and 0.9, are set for the normalized PDFs as shown in Figure 3.12. The 0.9-level intervals are presented in the last column of Table 3.8, which are significantly reduced compared with the initial intervals. Together with the p-boxes, the updated crisp CDF

with 1-alpha level is also estimated by employing the updated distributions of $p_1$, $p_4$, and $p_5$ as shown in Figure 3.13 and the crisp updated value of $p_2$.

The p-boxes with the three alpha-levels and the updated CDF are illustrated in Figure 3.14. The initial p-box with the original epistemic uncertainty space is significantly reduced through the updating procedure with different alpha-levels. An integrative comparison of Figures 3.12 and 3.14 shows that the higher the alpha-level, the smaller the input epistemic intervals, and furthermore, the narrower the resulting p-box of the output, even though the differences in the three p-boxes are relatively small. In addition, the narrowest p-box with 0.9-alpha level still envelops the target CDF of the observations. More importantly, the updated CDF shows good agreement with the CDF of the target observations. This outcome clearly demonstrates the feasibility of the combination of the Bhattacharyya distance metric and staircase random variables in the stochastic model updating and uncertainty characterization.



Figure 3.14 Updated p-boxes with different alpha-levels.

## 3.6    Conclusions

The combined application of the Bhattacharyya distance metric and staircase random variable is demonstrated as key ingredients of the proposed ABC model updating framework. The application to the NASA UQ challenge problem reveals the feasibility of the combination of the Bhattacharyya distance and staircase random variable for the s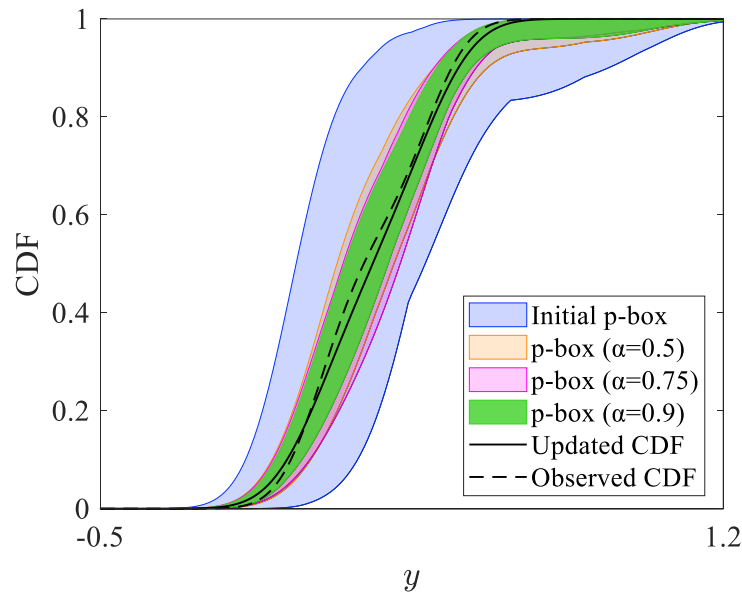tochastic model updating and uncertainty characterization. The staircase densities can act as a nonparametric connection between the epistemic uncertainty space and investigated outputs. In addition, the distance-based approximate likelihood serves

as a convenient connection between the Bayesian updating procedure and UQ metrics. By utilizing the Bhattacharyya distance, the proposed approximate likelihood enables to capture the comprehensive uncertainty characteristics of both the model predictions and observations. As a consequence, the proposed nonparametric updating framework fulfils the challenging expectation in stochastic updating to calibrate the probabilistic distributions of parameters without prior knowledge about their distribution formats.

In spite of the advantage on uncertainty characterization, the combined application of the Bhattacharyya distance and staircase random variable can be not appropriate in an exclusive manner as revealed in the shear building model example. It is required to be complemented by the Euclidian distance metric in a two-step scheme to avoid non-unique solutions in stochastic model updating. In other words, the Euclidian distance-based deterministic updating should be the precondition before performing the stochastic updating procedure.

One of the perspectives of the proposed nonparametric model updating framework is that it can be combined with structural health monitoring (SHM). In SHM, uncertainties not only in the input parameters but also in measurements need to be considered, because the measurements are always driven under hard-to-control randomnesses, such as the environmental noises and measurement system errors. Another challenging perspective focuses on uncertainties in the modeling. The numerical model always contains unavoidable simplifications and approximations, e.g., the linearized representation of the nonlinear behaviors. These uncertainties make the updating procedure more difficult, and thus it should aim at the robust calibration to achieve the maximum allowable uncertainty, while providing acceptable fidelity to the measurements. This extension of the proposed updating framework to be implemented in SHM will be addressed in the future work.

## CRediT authorship contribution statement

**Masaru Kitahara**: Conceptualization, Methodology, Software, Writing – original draft, Validation. **Sifeng Bi**: Conceptualization, Methodology, Writing – review & editing. **Matteo Broggi**: Conceptualization, Methodology, Software, Writing – review & editing. **Michael Beer**: Supervision, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

# References

Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162, 2025-2035.

Beck, J. L., and Katafygiotis, L. S. (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechnics*, 124(4), 455-461.

Beck, J. L., and Au, S. K. (2002). Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *Journal of Engineering Mechanics*, 128(4), 380-391.

Beer, M., Ferson, S., and Kreinovich, V. (2013). Imprecise probabilities in engineering analysis. *Mechanical System and Signal Processing*, 37, 4-29.

Betz, W., Papaioannou, I., and Straub, S. (2016). Transitional Markov chain Monte Carlo: Observations and improvements. *Journal of Engineering Mechanics*, 142(5), 04016016.

Bi, S., Prabhu, S., Cogan, S., and Atamturktur, S. (2017). Uncertainty quantification metrics with varying statistical information in model calibration and validation. *AIAA Journal*, 55, 3570-3583.

Bi, S., Broggi, M., and Beer, M. (2019). The role of the Bhattacharyya distance in stochastic model updating. *Mechanical System and Signal Processing*, 117, 437-452.

Ching, J., and Chen, Y. C. (2007). Transitional Markov chain Monte Carlo method for Bayesian updating, model class selection, and model averaging. *Journal of Engineering and Mechanics*, 133(7), 816-832. Crespo, L. G., Kenny, S. P., and Giesy, D. P. (2014). The NASA Langley multidisciplinary uncertainty quantification challenge. In: *16th AIAA Non-Deterministic Approaches Conference*, National Harbor, Maryland.

Crespo, L. G., Kenny, S. P., Giesy, D. P., and Stanford, B. K. (2018). Random variables with moment-matching staircase density function. *Applied Mathematical Modelling*, 64, 196-213.

Faes, M., and Moens, D. (2020). Recent trends in the modeling and quantification of non-probabilistic uncertainty. *Archives of Computational Methods in Engineering*, 27, 633-671.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., and Sentz, K. (2003). Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015, Sandia National Laboratories, California.

Ghanem, R., Yadegaran, I., Thimmisetty, C., Keshavarzzadeh, V., Masri, S., Red-Horse, J., Moser, R., Oliver, T., Spanos, P., and Aldraihem, O. J. (2015). Probabilistic approach to NASA Langley research center multidisciplinary uncertainty quantification challenge problem. *Journal of Aerospace Information Systems*, 12, 170-188.

Goller, B., Broggi, M., Calvi, A., and Schuëller, G. I. (2011). A stochastic model updating technique for complex aerospace structures. *Finite Elements in Analysis and Design*, 47, 739-752.

Grimmett, G. R., and Stirzaker, D. R. (2001). *Probability and random processes*. New York, Oxford University Press.

Hasting, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrica*, 57, 97-109.

Khodaparast, H. H., Mottershead, J. E., and Friswell, M. I. (2008). Perturbation methods for the estimation of parameter variability in stochastic model updating. *Mechanical System and Signal Processing*, 22, 1751-1773.

Kumar, P. (2002). Moment inequalities of a random variable defined over a finite interval. *Journal of Inequalities in Pure and Applied Mathematics*, 3, 1-11.

Mares, C., Mottershead, J. E., and Friswell, M. I. (2006). Stochastic model updating: Part 1—theory and simulated example. *Mechanical System and Signal Processing*, 20, 1674-1695.

Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to interval analysis*. SIMA Press, Philadelphia.

Oberkampf, W., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S. (2004). Challenge problems: uncertainty in system response given unertain parameters. *Reliability Engineering & System Safety*, 85, 11-19.

Patelli, E., Alvarez, D. A., Broggi, M., and de Angelis, M. (2015). Uncertanty management in multidisciplinary design of critical safety systems. *Journal of Aerospace Information Systems*, 12, 140-169.

Patelli, E., Govers, Y., Broggi, M., Gomes, H. M., Link, M., and Mottershead, J. E. (2017). Sensitivity or Bayesian model updating: a comparison of techniques using the DLR AIRMOD test data. *Archive of Applied Mechanics*, 87, 905-925.

Patra, B. K., Launonen, R., Ollikainen, V., and Nandi, S. (2015). A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. *Knowledge-Based Systems*, 82, 163-177.

Rocchetta, R., Broggi, M., Huchet, Q., and Patelli, E. (2018). On-line Bayesian model updating for structural health monitoring. *Mechanical Systems and Signal Processing*, 103, 174-195.

Roy, C., and Oberkampf, W. (2011). A comprehensive framework for verification, validation, and uncertainty quantification. *Computer Methods in Applied Mechanics and Engineering*, 200, 2131-2144.

Safta, C., Sargsyan, K., Najm, H. N., Chowdhary, K., Debusschere, B., Swiler, L. P., and Eldred, M. S. (2015). Probabilistic methods for sensitivity analysis and calibration in the NASA challenge problem. *Journal of Aerospace Information Systems,* 12(1), 170-188.

Sairajan, K. K., and Aglietti, G. S., (2012). Robustness of system equivalent reduction expansion process on spacecraft

structure model validation. *AIAA Journal*, 50. 2376-2388.

Sentz, K., and Ferson, S. (2002). Combination of evidence in Dempster-Shafer theory. Techical Report SAND2002-0835, Sandia National Laboratories, California.

Sharma, R., Devi, S., Kapoor, G., and Barnett, N. S. (2009). A brief note on some bounds connecting lower order moments for random variables defined on a finite interval. *Interval Journal of Theoretical & Applied Sciences*, 1, 83-85.

Turner, B. M., and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2), 69-85.

# Chapter 4

# Research article 3: Robust optimization of a dynamic Black-box system under severe uncertainty: A distribution-free framework

This article provides a solution to the NASA UQ challenge problem 2019. The challenge problem is proposed by researchers in the NASA Langley Research Center in 2019 to represent the difficulties which are often encountered in the design of critical safety systems under the availability of very limited data. In such circumstance, the system needs to be designed in order to cope with the unavoidable uncertainty. The design of systems under uncertainty requires the availability of robust and efficient tools for UQ, and the challenge problem 2019 is aimed at confirming the availability and applicability of discipline independent UQ tools. It follows from the success of the previous edition of the NASA Multidisciplinary Uncertainty Quantification (UQ) Challenge 2014. In the challenge problem 2019, a Black-box model of a physical system is used to evaluate and improve its operational reliability. Unlike the previous challenge, the prior knowledge of the UM for aleatory parameters is extremely limited and only a common boundary is given. In addition, the response of the system of interest is time-dependent. These specific properties of the challenge problem 2019 are closely related to the novel developments in the previous two articles, i.e., the extension of the Bhattacharyya distance-based ABC updating framework for calibrating dynamic systems described by time-domain sequences and the novel distribution-free stochastic model updating framework for calibrating aleatory parameters whose distribution families are unknown a priori.

Consequently, this article proposes four key techniques to address different tasks in the challenge problem. First, a distribution-free Bayesian model updating framework with three key components, i.e., the staircase density functions for characterizing aleatory parameters, the Bhattacharyya distance for quantifying the uncertainty characteristics of the observations, and a dimension-reduction procedure for evaluating the Bhattacharyya distance based on time-dependent data, is proposed. The advantages of the proposed distribution-free approach over the conventional distribution-based approach are discussed upon Sub-problem A (Model calibration and uncertainty quantification of the subsystem). The proposed approach is further employed as a robust tool for uncertainty characterization and quantification throughout the challenge problem (i.e., Tasks B3, E2, and E4). Second, an adaptive pinching approach is proposed to perform the

sensitivity analysis on epistemic parameters, providing an efficient way of identifying the largest possible reduction of the epistemic uncertainty space. This approach is demonstrated on Sub-problem B (Uncertainty reduction), and is also used for Tasks C3 and E3. Third, the probability bounds analysis by the double-loop Monte Carlo method is performed to obtain the bounds on the respective failure probabilities and the worst-case failure probabilities upon Sub-problem C (Reliability analysis of baseline design) and the following tasks, i.e., Tasks D2 and E5. Fourth, to identify an optimal design point of the system, a NISS approach is proposed providing a systematic way of exploring a hyper-rectangular space of the design point and identifying the local optimal values of each design point parameter minimizing the worst-case failure probabilities. The approach is demonstrated on Sub-problem D (Reliability-based design) and is also used for Task E4.

While the aforementioned four key techniques are successfully demonstrated as robust tools for UQ, the challenge problem 2019 also reveals some limitations of their applicability. For instance, the proposed distribution-free Bayesian model updating framework is only applicable for the case where aleatory parameters are independent with each other, since the staircase density functions are defined for univariate random variables and cannot account for the dependence structure among the parameters, though the results of Task E2 indicates the possibility of the presence of the correlation among some of the aleatory parameters and the link between ignoring the correlation and observed unstable behaviors of the system. In addition, the computational burden of the probability bounds analysis by the double-loop Monte Carlo method is relatively high and might not be acceptable for practical applications. These limitations will be addressed in the next two articles.

Finally, it is necessary to note that, this article is a product of the joint work with the group of Institute for Risk and Uncertainty at University of Liverpool. The author of this thesis has jointly contributed towards the solution of the challenge problem with Mr. Lye Adolphus, where the author has mainly contributed to the development of the first and fourth techniques above.

# Robust optimization of a dynamic Black-box system under severe uncertainty: A distribution-free framework

Adolphus Lye [a, 1], Masaru Kitahara [b, 1], Matteo Broggi [b], Edoardo Patelli [c, *]

[a] *Institute for Risk and Uncertainty, University of Liverpool, United Kingdom*

[b] *Insritute for Risk and Reliability, Leibniz Universität Hannover, Germany*

[c] *Centre for Intelligent Infrastructure, Department of Civil and Environmental Engineering, University of Strathclyde, United Kingdom*

[1] Both authors have equally contributed towards the solution of the problem and are equally addressed as the "first author" of the present work.

[*] Correspondence author. E-mail address: edoardo.patelli@strath.ac.uk (E. Patelli).

**Abstract:** In the real word, a significant challenge faced in designing critical systems is the lack of available data. This results in a large degree of uncertainty and the necessity for uncertainty quantification tools in order to make risk-informed decisions. The NASA-Langley UQ Challenge 2019 is proposed to provide such setting, requiring different discipline-independent approaches to address typical tasks for required for the design of critical systems. This paper addresses the NASA-Langley UQ Challenge by proposing four key techniques to provide the solution to the challenge: (1) a distribution-free Bayesian model updating framework for the calibration of the uncertainty model; (2) an adaptive pinching approach to analyze and rank the relative sensitivity of the epistemic parameters; (3) the probability bounds analysis to estimate failure probabilities; and (4) a non-intrusive stochastic simulation approach to identify an optimal design point.

**Keywords:** Uncertainty quantification; Model class selection; Sensitivity analysis; Staircase density function; Robust optimization; Non-intrusive imprecise stochastic simulation.

## 4.1    Introduction

### 4.1.1    Research context

The design of critical safety systems is frequently associated with the availability of limited data. Despite such challenge posed, the system needs to be designed in order to cope with the unavoidable uncertainty. Such uncertainty is generally classified as either aleatory or epistemic uncertainty (Oberkampf et al., 2004; Roy and Oberkampf, 2011). Aleatory uncertainty is often considered as the irreducible uncertainty that is caused by the inherent randomness of the system (Crespo and Kenny, 2021) and generally modelled as random variables according to some distribution function

(Der Kiureghian and Ditlevsen, 2009; Meng et al., 2020). On the other hand, epistemic uncertainty is caused by a lack of or limited knowledge which can be theoretically reduced or eliminated through, for instance, further data collection (Rocchetta et al., 2018). An epistemic parameter is generally represented by a fixed value within a bounded set whose intervals reflect the level of knowledge on the parameter (Crespo and Kenny, 2021). The lower the level of knowledge, the larger the interval of this bounded set. It is important to note that the aleatory and epistemic uncertainty can refer to the same physical quantity and, therefore, such classification becomes fuzzy. In fact, the aleatory uncertainty can be seen as the remaining uncertainty after a campaign of model updating, aimed at reducing the epistemic uncertainty, is performed.

The design of systems under uncertainty requires the availability of robust and efficient tools for the uncertainty characterization and quantification. In order to check the availability of discipline independent tools and applicability of such UQ tools, NASA Langley proposed a new UQ Challenge problem in 2019 (Crespo and Kenny, 2021) with the purpose of modelling the dynamic behaviour of a system, analyzing its operational reliability, and further devising an improved design configuration for the system under uncertainty. This UQ Challenge problem follows from the success of the previous edition in 2013 (Crespo et al., 2014).

In this challenge, a "Black-box" computational model of a physical system is used to evaluate and improve its reliability. Unlike the previous challenge (Crespo et al., 2014), the Uncertainty Model (UM) to the respective aleatory input parameters are completely unknown and they are to be derived by the participants. In addition, the response of the system is time-dependent providing a realistic setting under which different tasks will be addressed. This is because in the real-world, prior distributional knowledge to such models associated with the parameters of interest are usually unavailable.

This paper is part of a Special Issue providing the solution for the NASA-Langley UQ Challenge problem 2019. Therefore, for the sake of the content length, a detailed explanation to the problem is not shown. Instead, a summary of the challenge and the description of the notations is provided.

## 4.1.2   The NASA-Langley UQ challenge problem (2019)

The system is characterized by a design point $\theta$ with nine real components (i.e. $\theta \in \mathbb{R}^{n_\theta}$), and an uncertain model $\delta$ comprising of elements $a$ and $e$ (Crespo and Kenny, 2021). $a$ denotes the vector of five aleatory parameters real components while $e$ denotes the vector of four epistemic parameters.

The aleatory space $A$ is represented as $a \backsim f_a$ whereby $f_a$ is the joint density function. The initial aleatory space

is $A_0 = [0,2]^5$. The epistemic space $E$ is represented as $e \smallfrown E$. The initial epistemic space is $E_0 = [0,2]^4$. Hence, the UM for $\delta$ is fully characterized by: $\langle f_a, E \rangle$.

The system of interest consists of a set of interconnected subsystems for which $\delta$ is concentrated in one of these subsystems. This subsystem is modeled by a Black-box model function $\hat{y} = y(a,e,t)$ where $t \in [0,5]$ s indicates the time parameter. The output of this subsystem is given as a discrete time history: $y^l(t) = [y^l(0), y^l(dt), \cdots, y^l(5000 \cdot dt)]$, where $l = 1, \cdots, 100$, and $dt = 0.001 \ s$. This yields a total of 5001 data of $y^l(t)$ per given $l$ and the entire time history data is denoted as $D_1 = [y^l(t)]_{l=1,\cdots,100}$.

The goal of this challenge can be summarized as follows (Crespo and Kenny, 2021):

A. To create an UM for $\delta$;

B. To decide a limited number of refinements (up to four) on the epistemic variables;

C. To perform a reliability analysis on a given design point $\theta$;

D. To identify a new $\theta$ with improved reliability;

E. To improve the UM for $\delta$ and $\theta$ given observations of the integrated system.

It needs to be highlighted that Task F of the challenge is not addressed in this paper.

## 4.2 Task A: Model calibration and uncertainty quantification of the subsystem

### 4.2.1 Modelling strategy and hypothesis

The Bayesian model updating technique is adopted to calibrate the UM using the available data $D_1$. This provides a probabilistic approach through which the joint distribution function $f_a$ can be identified. Generally, Bayesian model updating is not adopted to reduce epistemic uncertainty when represented by intervals. However, the uncertainty of $e$ can be quantified by modelling the intervals as uniform distributions and then computing the posterior distribution. At this point, it is important to note that the posterior distributions are used to define new intervals as already successfully proposed in Patelli et al. (2015). Hence, the epistemic parameters $e_{i_e}$, for $i_e = 1, \cdots, 4$, are assumed to be independent between one another and their respective priors characterized by a non-informative uniform distribution with bounds defined by the epistemic space $E_0$ (i.e., see Section 4.1.2).

In this section, different strategies are adopted to represent the aleatory uncertainty and different metrics are employed to compare between distributions. As a result, we have obtained a conservative (i.e., low-risk) approach; and

a more aggressive (i.e., high-risk) approach aimed at reducing the uncertainty of the UM.

## 4.2.2 Bayesian model updating

Bayesian model updating is a probabilistic model updating approach whose mathematical formulation follows the Bayes' rule introduced by (Beck and Katafygiotis, 1998; Katafygiotis and Beck, 1998):

$$P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M}) = \frac{P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})P(\boldsymbol{\Theta}|\boldsymbol{M})}{P(\boldsymbol{D}|\boldsymbol{M})} \tag{4.1}$$

where $\boldsymbol{\Theta}$ is the vector of inferred parameters, $\boldsymbol{D}$ denotes the vector of observed data used for model updating, and $\boldsymbol{M} = \{\hat{y}, f_a\}$ is the model class which best represents the observed data $\boldsymbol{D}$. The different components in Equation (4.1) are the prior distribution (i.e., $P(\boldsymbol{\Theta}|\boldsymbol{M})$), the likelihood function (i.e., $P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})$), the evidence (i.e., $P(\boldsymbol{D}|\boldsymbol{M})$), and finally the posterior distribution (i.e., $P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M})$). The term of interest is $P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M})$ which describes our update knowledge of $\boldsymbol{\Theta}$ after observing the data $\boldsymbol{D}$ and it is usually expressed as a non-normalized distribution by neglecting the evidence, which serves as the normalization constant).

Advanced sampling techniques are used to sample from such non-normalized distributions (Lye et al., 2021). In this work, the transitional Markov chain Monte Carlo (TMCMC) sampler is implemented whose algorithm is based on the adaptive Metropolis-Hastings technique (Beck and Au, 2002) and uses "transitional" distributions $P^j$ from which samples are obtained sequentially. Details to the TMCMC algorithm can be found in Lye et al. (2021) and Ching and Cheng (2007). The motivations behind the use of TMCMC in this problem are attributed to the following: (1) it is able to sample from complex-shaped posteriors via "transitional" distributions $P^j$; (2) it can sample from high-dimensional posteriors (i.e., up to 24 dimensions) (Gray et al., 2020); and (3) it computes the evidence $P(\boldsymbol{D}|\boldsymbol{M})$ which makes the algorithm useful in model selection problems (Ching and Cheng, 2007).

Due to the large data set provided (i.e., $D_1$), it becomes computationally expensive to use a full likelihood function to perform an actual Bayesian computation (Bi et al., 2019). For this reason, approximate Bayesian computation (ABC) (Abdessalem et al., 2018; Turner and Van Zandt, 2012) is adopted and defined as (Bi et al., 2019):

$$P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M}) \propto exp\left(-\frac{d}{\epsilon}\right)^2 \tag{4.2}$$

where $d$ is the stochastic distance metric which quantifies the difference between the distribution of the observed data $\boldsymbol{D}$ and the model output of $\hat{y}$, while $\epsilon$ is the width factor of the approximate Gaussian function.

### 4.2.3 Proposed approach

To determine the UM for $a$, five possible distribution types for $f_a$ are identified and listed in Table 4.1. It needs

to be highlighted that the choice of the staircase density function (SDF) presents a distribution-free approach contrary

to the other choice of distributions presented in the table. A key strength of SDFs lies in its flexibility in describing a

wide range of density shapes, including highly-skewed and/or multi-modal distributions. This makes them particularly

applicable in modeling the marginal distributions of the aleatory variables whose density shapes are unknown a prior.

In this analysis, it has been assumed that: (1) the marginal distribution of all the aleatory uncertainties belong to the

same distribution class; (2) no dependency exists between all $a_{i_a}$ (i.e., no correlation matrix used).

Table 4.1 Distribution type with the non-informative uniform prior bounds of its corresponding parameters for each aleatory model $f_a$.

| Aleatory model | Distribution type | Prior distribution parameters |
|---|---|---|
| $f_a^1$ | Beta$(\alpha_{i_a}, \beta_{i_a})$ | $\alpha_{i_a}$ (Shape parameter 1): $U[0, 100]$ |
| | | $\beta_{i_a}$ (Shape parameter 2): $U[0, 100]$ |
| $f_a^2$ | Truncated normal$(\mu_{i_a}, \sigma_{i_a})$ | $\mu_{i_a}$ (Mean of $a_{i_a}$): $U[0, 2]$ |
| | $[TN(a_{i_a}: \mu_{i_a}, \sigma_{i_a})]$ | $\sigma_{i_a}$ (Standard deviation of $a_{i_a}$): $U[0.01, 2]$ |
| $f_a^3$ | Truncated lognormal | $\mu_{i_a}$ (Mean of $\log(a_{i_a})$): $U[-10, 10]$ |
| | $[TLN(a_{i_a}: \mu_{i_a}, \sigma_{i_a})]$ | $\sigma_{i_a}$ (Standard deviation of $\log(a_{i_a})$): $U[0.01, 5]$ |
| $f_a^4$ | Truncated Gamma | $\alpha_{i_a}$ (Shape parameter 1): $U[0, 10]$ |
| | $[TG(a_{i_a}: \alpha_{i_a}, \beta_{i_a})]$ | $\beta_{i_a}$ (Shape parameter 2): $U[0, 10]$ |
| $f_a^5$ | Staircase density function | $\mu_{i_a}$ (Mean of $a_{i_a}$): $U[0, 2]$ |
| | $[SDF(a_{i_a}: \mu_{i_a}, (m_2)_{i_a}, (m_3)_{i_a}, (m_4)_{i_a})]$ | $(m_2)_{i_a}$ (2nd central moment of $a_{i_a}$): $U[0, 1]$ |
| | | $(m_3)_{i_a}$ (3rd central moment of $a_{i_a}$): $U\left[-\frac{4}{3\sqrt{3}}, \frac{4}{3\sqrt{3}}\right]$ |
| | | $(m_4)_{i_a}$ (4th central moment of $a_{i_a}$): $U\left[0, \frac{4}{3}\right]$ |

For the case of $f_a^1$ to $f_a^4$, the distribution parameters add an additional 10 inferred parameters, while $f_a^5$ adds an

additional 20 inferred parameters. Each of these parameters are assigned a non-informative uniform prior with bounds,

stated in Table 4.1, chosen to ensure sufficient degrees of freedom in the model calibration procedure. It is also assumed

that these parameters are independent from one another. This brings the total number of inferred parameters to 24 for

the case of the SDF, and 14 for the rest of the distributions.

### 4.2.3.1 Distribution-based approach

For the case of $f_a^1$ to $f_a^4$, there is a need to reduce the size of the data to reduce the computational cost in evaluating

$P(\boldsymbol{D}|\boldsymbol{\Theta}, \boldsymbol{M})$. To achieve this, the fast Fourier transformation (FFT) procedure is performed on $D_1$ for each $l$ according

to Gray et al. (2020), Heideman et al. (1984), and Bai et al. (2020):

$$y^l(t) = \sum_{q=0}^{5000} C_q^l \cdot exp[-i \cdot q \cdot \omega_0 \cdot t] \tag{4.3}$$

where $\omega_0 = \frac{2\pi}{5001}$, and $C_q^l$ mean the numerical coefficient with real and imaginary components denoted as $Re(C_q^l)$ and

$Im(C_q^l)$ respectively. From which, the amplitude $A_q^l$ and phase angles $\phi_q^l$ are obtained as follows:

$$A_q^l = \sqrt{Re(C_q^l)^2 + Im(C_q^l)^2} \tag{4.4}$$

$$\phi_q^l = atan2\left[\frac{Im(C_q^l)}{Re(C_q^l)}\right] \tag{4.5}$$

To remove the periodicity associated with the values of $\phi_q^l$, we introduced a phase shift such that a factor of $2\pi$ rad is

added or subtracted whenever the jump between consecutive phase angles is greater than $\pi$ rad. This is achieved using

*unwrap* function in MATLAB to ensure the jump between any consecutive phase angles is always less than $\pi$ rad. In

doing so, it can ensure the monotonic behavior of $\phi_q^l$ and simplifies its subsequent computation for $P(D|\Theta, M)$. When

this is done, we obtain the frequency spectra of $A_q^l$ and $\phi_q^l$ as shown in Figure 4.1 where it can be observed that beyond

frequencies $\omega > 5.80$ Hz, the values of $A_q^l$ do not show any additional perturbations for all $l$, thereby allowing those

data to be discarded. Thus, only 30 values of $\omega$ between 0 Hz and 5.80 Hz are considered for both $A_q^l$ and $\phi_q^l$. Let this

set of values of $\omega$ be denoted as $\omega_n$, for $n = 1, \cdots, 30$. This effectively reduces the total number of data used for model

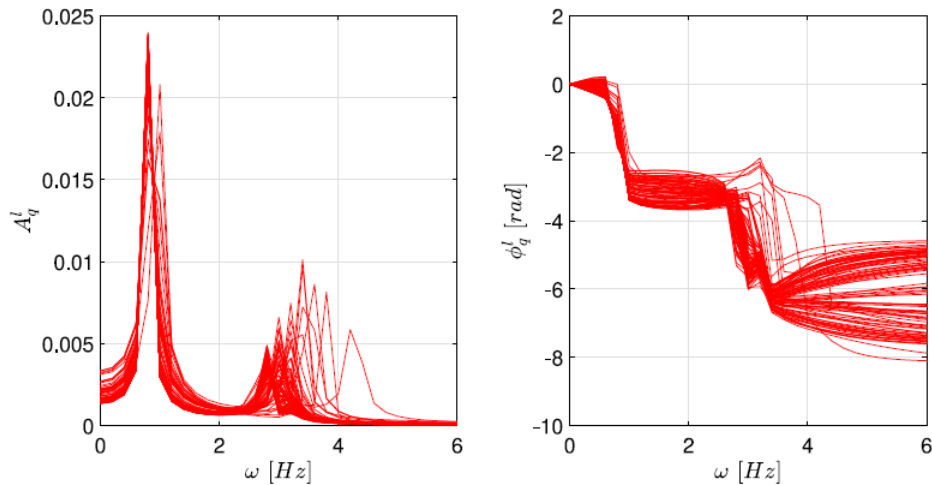calibration from 500100 to 6000 (i.e., 3000 for $A_q^l$ and 3000 for $\phi_q^l$).



Figure 4.1 Illustration of the frequency spectra obtained from $D_1$ via FFT.

In order to account for the variability of $A_q^l$ and $\phi_q^l$ at each $\omega_n$, the stochastic distance metric $d$ for $P(\mathbf{D}|\boldsymbol{\Theta},\mathbf{M})$ is the Wasserstein distance defined as (Panaretos and Zemel, 2019):

$$d_W = \int_{-\infty}^{\infty} \left| F_{\mathbf{D}}(x) - F_{\hat{y}}(x) \right| dx \tag{4.6}$$

whereby $F_{\mathbf{D}}(x)$ and $F_{\hat{y}}(x)$ are the respective empirical cumulative distribution functions (ECDFs) of the data (i.e., $A_q^l$ and $\phi_q^l$) and the stochastic model output of $\hat{y}$ at a given $\omega_n$, while $x$ is the variable denoting either $A_q^l$ or $\phi_q^l$. In essence, $d_W$ quantifies the enclosed area between both ECDFs. The smaller $d_W$ is, the higher the degree of similarity between the ECDFs of the data and the stochastic prediction by $\hat{y}$ (Ferson et al., 2008). Utilizing Equation (4.2), $P(\mathbf{D}|\boldsymbol{\Theta},\mathbf{M})$ is defined as:

$$P(\mathbf{D}|\boldsymbol{\Theta},\mathbf{M}) = \prod_{n=1}^{30} exp\left[ -\left(\frac{d_{W,n}^A}{\epsilon_n^A}\right)^2 - \left(\frac{d_{W,n}^\phi}{\epsilon_n^\phi}\right)^2 \right] \tag{4.7}$$

where the values of $\epsilon_n^A$ and $\epsilon_n^\phi$ are approximated by the standard deviations of $A_q^l$ or $\phi_q^l$ respectively at $\omega_n$. In addition, independence is assumed between data sets to reduce computational costs in computing $P(\mathbf{D}|\boldsymbol{\Theta},\mathbf{M})$. However, it needs to be highlighted that in reality, there exists dependencies between the identified $\omega_n$ for each $l$th sequence. To compute $P(\mathbf{D}|\boldsymbol{\Theta},\mathbf{M})$, 100 model evaluations by $\hat{y}$ per given set of model inputs comprising of $\{a, e\}$, are necessary to construct $F_{\hat{y}}(x)$.

### 4.2.3.2 Distribution-free approach

For the case of $f_a^5$, the distribution is defined by the SDF as (Crespo et al., 2018):

$$f_a = \begin{cases} h_{i_b} & \forall a \in ((i_b - 1) \cdot \kappa, i_b \cdot \kappa]^5, \text{for } 1 \leq i_b \leq N_b \\ 0 & , \text{otherwise} \end{cases} \tag{4.8}$$

where $N_b = 50$ indicates the number of bins, $h_{i_b}$ is the height of the SDF in the $i_b$th bin, and $\kappa = \frac{2}{N_b}$ is the length of each sub-interval. It needs to be noted that $h_{i_b} > 0$ for all $N_b$ bins and that their values are determined by solving the following convex optimization problem:

$$\hat{h}_{i_b} = \underset{h_{i_b} \geq 0}{\operatorname{argmin}} \left\{ J(\mathbf{h}): \sum_{i_b=1}^{N_b} \int_{(i_b-1)\cdot\kappa}^{i_b\cdot\kappa} z \cdot h_{i_b} dz = \mu_{i_a}, \sum_{i_b=1}^{N_b} \int_{(i_b-1)\cdot\kappa}^{i_b\cdot\kappa} (z - \mu_{i_a})^r \cdot h_{i_b} dz = (m_r)_{i_a}, r = 2, 3, 4 \right\} \tag{4.9}$$

where $J(h)$ denotes the arbitrary cost-function. Details to $J(h)$ and Equation (4.9) can be found in Crespo et al. (2018) and Kitahara et al. (2022).

To avoid a potential error in the employment of the likelihood function for the distribution-based approach being brought forward, a different set-up for $P(D|\theta, M)$ is used in this analysis, where a different stochastic distance is used and the data is analyzed in the time domain. Unlike in the distribution-based approach, the Bhattacharyya distance (Bi et al., 2019) is employed as the stochastic distance metric $d$:

$$d_B = -\log\left\{\sum_{i_{N_t}=1}^{n_b} \cdots \sum_{i_1=1}^{n_b} \sqrt{p_D\left(b_{i_{1,},\cdots,i_{N_t}}\right) p_{\hat{y}}\left(b_{i_{1,},\cdots,i_{N_t}}\right)}\right\} \tag{4.10}$$

where $p_D\left(b_{i_{1,},\cdots,i_{N_t}}\right)$ and $p_{\hat{y}}\left(b_{i_{1,},\cdots,i_{N_t}}\right)$ denote the probability mass function (PMF) values of the data from $D_1$ and the stochastic model output from $\hat{y}$ respectively within the bin $b_{i_{1,},\cdots,i_{N_t}}$, and $n_b = 20$ denotes the number of bins used to compute the Bhattacharyya distance. It is important to be highlighted that each bin has $N_t = 5001$ coordinates as it is generated within a $N_t$-dimensional joint PMF space. Because of this, the resulting joint PMF space has an excessive number of dimensions for a direct evaluation of $P(D|\theta, M)$. This brings the need for a dimension-reduction procedure which is employed through the following steps (Kitahara et al., 2021):

1) Define the window length $L_W$ and divide the data set $[y^l(t)]_{l=1,\cdots,100}$ into $\left\lceil \frac{N_t}{L_W} \right\rceil$ distinct intervals, where $\lceil \cdot \rceil$ is the ceil operator;

2) Compute the root mean square (RMS) values of each interval $\mathbf{R} = \left[R_1, \cdots, R_{\left\lceil \frac{N_t}{L_W} \right\rceil}\right]$ and generate the sample set of the RMS values $\mathbf{R}_D \in \mathbb{R}^{100 \times \left\lceil \frac{N_t}{L_W} \right\rceil}$ where:

$$\mathbf{R}_D = \left[\mathbf{R}_D^1, \cdots, \mathbf{R}_D^{\left\lceil \frac{N_t}{L_W} \right\rceil}\right], \text{with } \mathbf{R}_D^v = \left[R_{1,v}, \cdots, R_{100,v}\right]^T$$

For $v = 1, \cdots, \left\lceil \frac{N_t}{L_W} \right\rceil$ while $\mathbf{R}_{\hat{y}} \in \mathbb{R}^{N_{\text{sim}} \times \left\lceil \frac{N_t}{L_W} \right\rceil}$ where $N_{\text{sim}} = 1000$ denotes the number of model evaluations by $\hat{y}$ per given set of model inputs $\{a, e\}$. It needs to be highlighted that the matrix structure of $\mathbf{R}_{\hat{y}}$ is similar to that of $\mathbf{R}_D$ with the exception of the number of row elements;

3) Evaluate $d_B$ between two sample sets $\mathbf{R}_D^v$ and $\mathbf{R}_{\hat{y}}^v$ for all $v$;

4) Obtain the corresponding RMS value $R_{d_B}$ and use it as the distance metric.

Consequently, $P(D|\Theta, M)$ is re-expressed as:

$$P(D|\Theta, M) = exp\left(-\frac{R_{d_B}}{\epsilon_B}\right)^2 \qquad (4.11)$$

where $\epsilon_B = 0.01$.

## 4.2.4 Results

For all set-ups, $N_s = 500$ samples are obtained from the resulting $P(\Theta|D, M)$. Based on the analysis done for all

aleatory models $f_a$, two models are chosen on the basis of their quality of the results and for the subsequent purpose

of comparison: $f_a^1$ and $f_a^5$. It needs to be highlighted, that $f_a^1$ is chosen given its relatively higher value of the evidence

$P(D|M)$ compared to the other aleatory models employed in the distribution-based approach as summarized in Table

4.2.

Table 4.2 Results of the evidence computed via TMCMC for each choice of model for $f_a$.

| Aleatory model | $f_a^1$ | $f_a^1$ | $f_a^1$ | $f_a^1$ |
|---|---|---|---|---|
| $P(D|M)$ | $3.2229 \times 10^{-7}$ | $2.1952 \times 10^{-7}$ | $5.0815 \times 10^{-10}$ | $1.0180 \times 10^{-9}$ |

To create the UM based on the information from the Bayesian model updating results, the following procedure is

undertaken: For the aleatory space, the histograms of the distribution parameters of the given $f_a$ are obtained based on

$P(\Theta|D, M)$. These histograms are converted into probability distribution functions by Kernel density estimation with

a Gaussian kernel (Bowman an Azzalini, 1997) and are normalized such that the distribution peak equals to one. An

illustration is presented using the distribution parameters for $f_a^1$ as an example in Figure 4.2. From these results, the

posterior distributions are interpreted as fuzzy sets where different levels of confidence $L_c \in [0, 1]$ will yield intervals

of varying width (Beer et al., 2013). Here, intervals at $L_c = 0.5$ level of confidence are considered for both the $f_a^1$ and

$f_a^5$ distribution parameters. The resulting intervals obtained would serve as the shape parameter inputs of the respective

aleatory model $f_a$. This yields the probability-boxes (i.e., p-boxes) (Ferson et al., 2002; Beer et al., 2014) of $f_a^1$ and $f_a^5$

which are illustrated in Figure 4.3.

To define the epistemic space, the same procedure is done on the resulting histograms of $e_1$ to $e_4$ obtained through

$P(\Theta|D, M)$ given the respective $f_a$. These histograms are shown in Figure 4.4 for the respective set-up. $L_c = 0.5$ level

of confidence is considered in the case of $f_a^1$ while $L_c = 0.025$ level of confidence is considered in the case of $f_a^5$. The

resulting intervals constitute the updated hyper-rectangular set $E$ defined by each of the resulting UM to which results

are presented in Table 4.6. Let the UM determined from $P(\boldsymbol{\Theta}|\boldsymbol{D}, \boldsymbol{M})$ given $f_a^1$ be denoted as $UM_{y0}^1$, while that given $f_a^5$ be denoted as $UM_{y0}^2$.
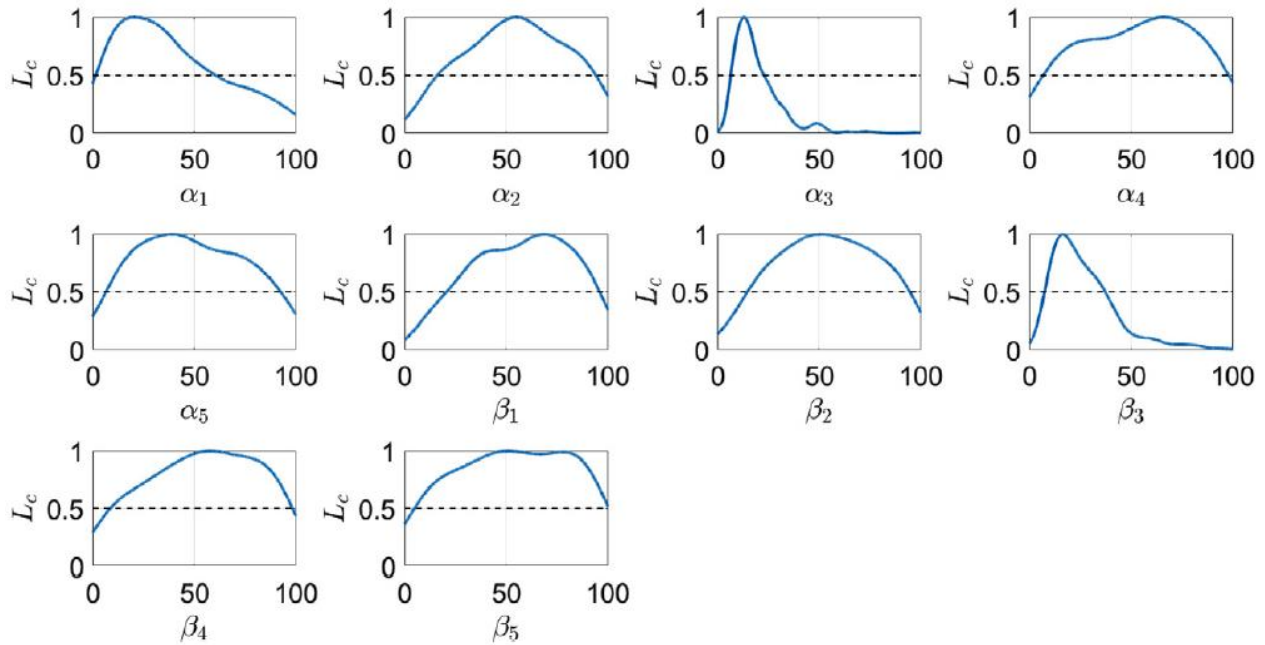


Figure 4.2 Illustration of the resulting distribution functions to the respective shape parameters of the joint beta distribution (i.e., $f_a^1$) obtained via Kernel density estimates.
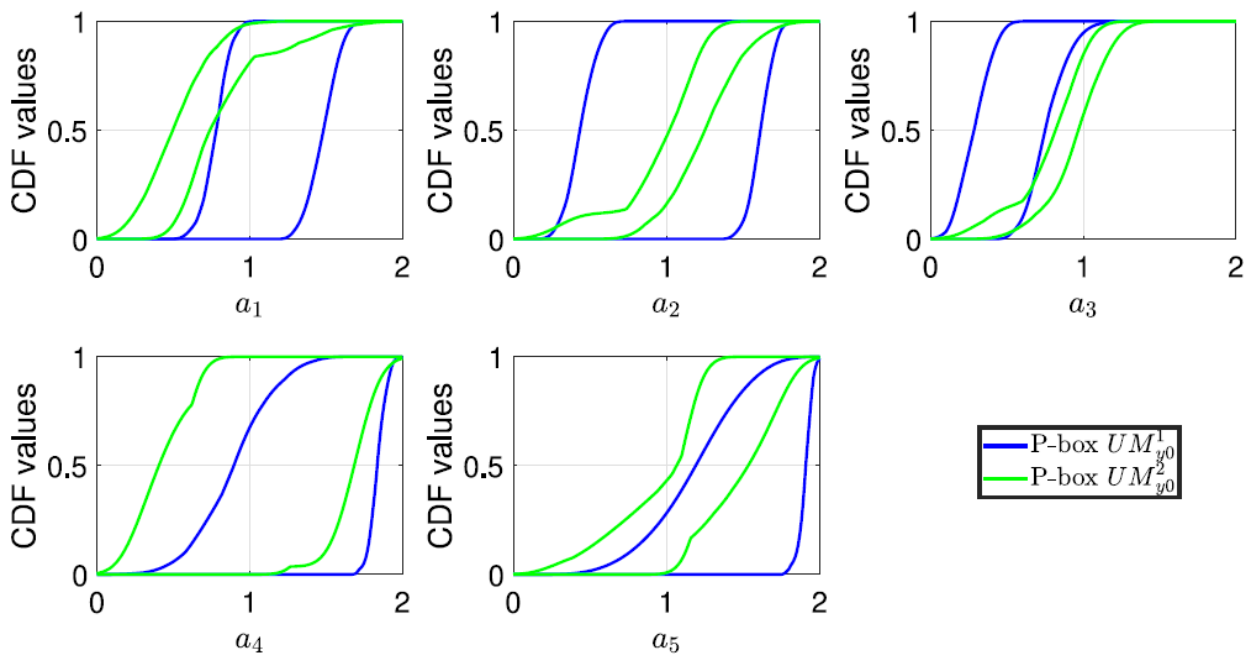


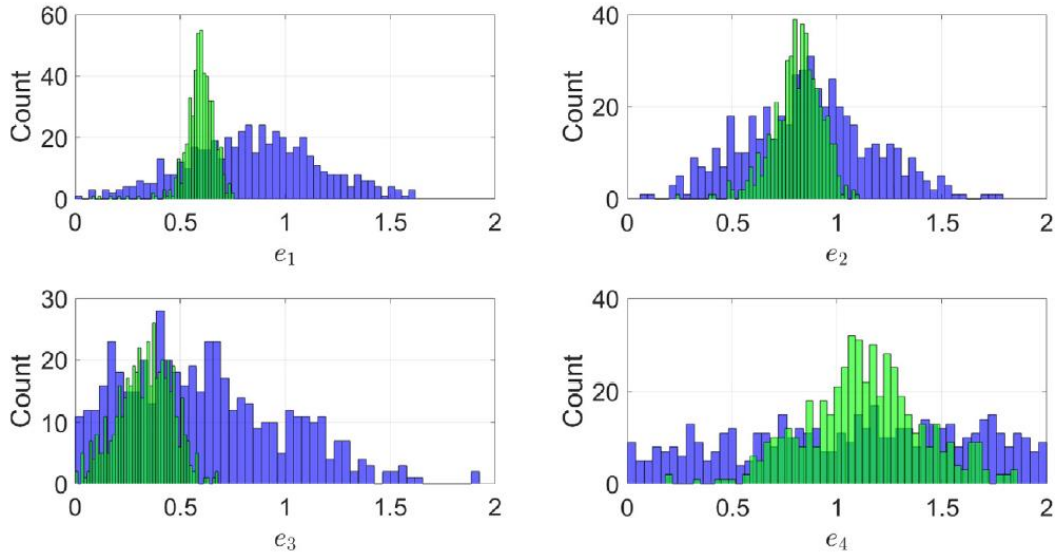Figure 4.3 P-box for $a_1$ to $a_5$ obtained from the respective UMs.

Figure 4.4 Histograms for $e_1$ to $e_4$ obtained from $P(\boldsymbol{\Theta}|\boldsymbol{D}, \boldsymbol{M})$ given $f_a^1$ (in blue) and $f_a^5$ (in green).

To verify the calibration results, the $N_s$ samples are generated from the hyper-rectangle defined by the bounds of the respective distribution parameters of $f_a$ and $e_{i_e}$ according to the respective UMs. For each sample realization from this hyper-rectangle, 100 model outputs of $\hat{y}$ is obtained for $t \in [0, 5]$ s. This is done by generating 100 realizations of $a$ from $f_a$, given the distribution parameters from the hyper-rectangle sample, whereas keeping $e$ fixed. This yields a $N_t \times 100 \times N_s$ array of data output of $\hat{y}$ for each UM whose results are plotted in Figure 4.5. From the figure, it can be seen that the model output bands of $UM_{y0}^1$ (in blue) and $UM_{y0}^2$ (in green) generally encompasses $D_1$ (in red) which indicates that the model calibration procedure, via Bayesian model updating, was done satisfactorily.
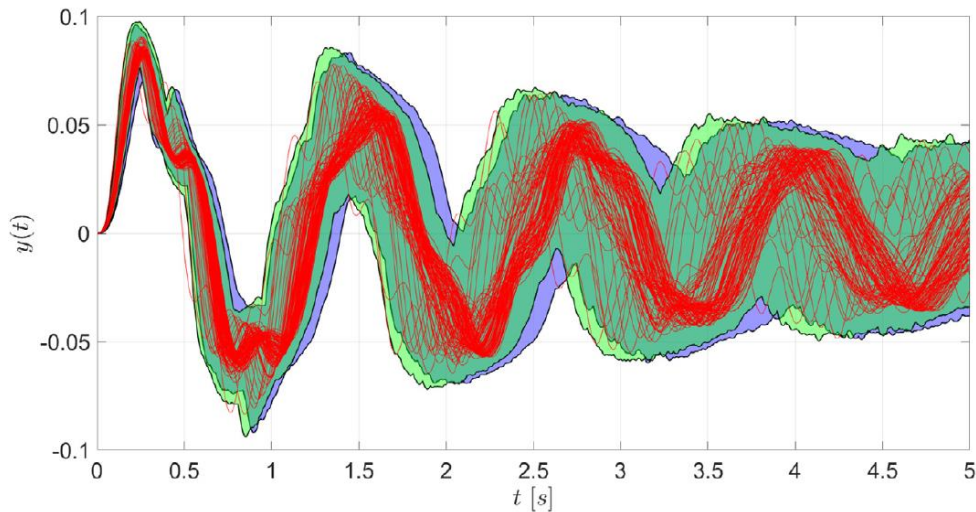


Figure 4.5 Output band from $\hat{y}$ according to $UM_{y0}^1$ (in blue) and $UM_{y0}^2$ (in green) along with the data sequence $D_1$ (in red) after calibration.

To further substantiate this, p-boxes of the calibrated model output of each UM are constructed at six points $t = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}\ s$. Each p-box describes the extreme bounds of the distribution of the $N_s$ ECDFs whereby each ECDF comprises of the 100 model output values at $t$. Figure 4.6 shows the resulting p-boxes from $UM_{y0}^1$ (in blue) and $UM_{y0}^2$ (in green) at each chosen $t$. From the plots, it can be seen that the ECDF of $D_1$ (in red) at any given $t$ is in general enclosed within the p-boxes. Furthermore, it can be observed that the shape of both p-boxes generally follows the shape profile of the ECDF of $D_1$ which indicates a good degree of fit by both UMs.
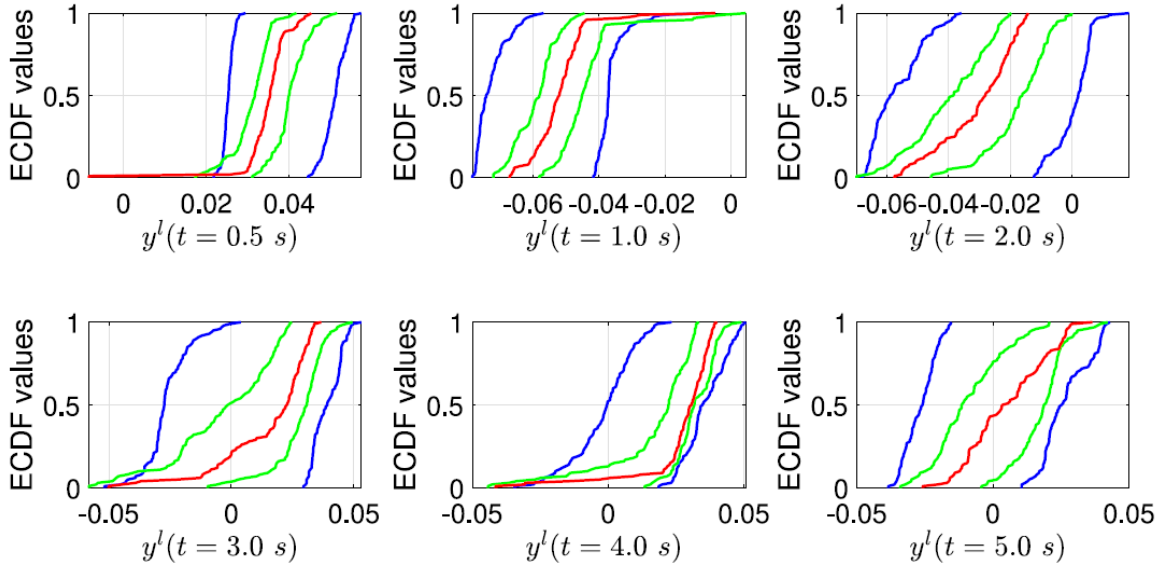


Figure 4.6 P-boxes of the model output from $\hat{y}$ according to $UM_{y0}^1$ (in blue) and $UM_{y0}^2$ (in green) at various time slices $t = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}\ s$. The red ECDF denotes the distribution of the data $D_1$.

## 4.2.5  Discussion

Figure 4.3 observes that the p-boxes obtained for $a_1$ to $a_5$ according to $UM_{y0}^1$ is generally wider than those using the second approach $UM_{y0}^2$. This implies a higher degree of uncertainty on the true distribution of all $a_{i_a}$ by $UM_{y0}^1$ that makes it less informative in identifying the true $f_a$ compared to $UM_{y0}^2$. In addition, the p-boxes for $a_1$ to $a_5$ obtained by $UM_{y0}^2$ are generally enclosed within those of $UM_{y0}^1$ which suggests that the true CDF defined by $f_a$ could lie within the p-box defined by $UM_{y0}^2$.

The intervals obtained from the posterior distributions for $e_1$ to $e_4$ through the model based on beta distribution is much wider compared to the UQ model obtained through the SDF based approach as presented in Figure 4.4. This further highlights the non-informative nature of $UM_{y0}^1$, especially for the case of $e_4$. In addition, the posteriors obtained through the SDF based approach show a much greater degree of update from the uniform prior and that it can identify

the epistemic parameters much more effectively as the peaks are more pronounced. This leads to the uncertainty bounds of $e$ according to $UM_{y0}^2$ being significantly narrower so that they are generally enclosed within that according to $UM_{y0}^1$ as seen from Table 4.6.

Figure 4.5 shows that the output bands of $\hat{y}$ obtained from both UMS follow the trend defined by $D_1$. However, from Figure 4.6, it is observed that the p-boxes obtained by $UM_{y0}^2$ have much tighter bounds compared to $UM_{y0}^1$ whilst still enclosing the ECDF of $D_1$. This is attributed to the p-box of the $a$, and the bounds on $e$ being narrower for $UM_{y0}^2$ than $UM_{y0}^1$ which resulted in the former yielding a significantly better degree of fit over $D_1$ than the latter. From the results, it can be concluded that $UM_{y0}^1$ is much more conservative compared to $UM_{y0}^2$ in modelling $f_a$ and $e$.

## 4.3    Task B: Uncertainty reduction

The objective of this task is to identify the epistemic parameters that have more predictive capability and improve the UM. This is achieved by performing a sensitivity analysis for the epistemic model parameters and the subsequent refinement of the epistemic space.

### 4.3.1   Sensitivity analysis

In this analysis, the epistemic parameters are ranked according to their ability to improve the predictive ability of the computational model of the subsystem. This predictive ability is quantified by the volume metric $\Omega$ defined as:

$$\Omega = \sum_{i_t=1}^{n_t} \rho_{i_t} \cdot \Delta_{i_t} \tag{4.12}$$

where $\rho_{i_t}$ is the area of the p-box at time-slice $i_t$, $\Delta_{i_t}$ denotes that time-step between time-slice $i_t - 1$ and $i_t$, and $n_t$ is the total of time-slices used for the computation. For the computation, we consider the six time-slices which were used for the illustration of the p-boxes in Figure 4.6.

To rank the epistemic parameters according to their respective sensitivity, an adaptive pinching method (Tucker and Ferson, 2006) is proposed for providing a non-empirical approach to determine the pinched bounds of a chosen epistemic parameter which yields that greatest reduction in the value of $\Omega$. The procedure is as follows: For a given $i_e$, the uncertainty space of $e_{i_e}$ is reduced by 90 %. This is performed whilst keeping the uncertainty space of the remaining three epistemic model parameters untouched. For a given $e_{i_e}$, its bounds would first be divided into 10 equally-spaced

units. Next, at iteration $j = 1$, a segment of bin length of one unit will be employed to isolate the region of the epistemic space defined by the lower and upper bounds of the first bin. This isolated region serves as the reduced (or "pinched") space. From there, the corresponding realizations of $\{a, e\}$ from the UM, whose $e_{i_e}$ value falls outside the bounds of the reduced epistemic space, is discarded. When this is done, the reduced volume $\Omega_p$ is computed again via Equation (4.12). After this is done, the segment shifts by one unit to the right as illustrated in Figure 4.7 and this initiates iteration $j = 2$ where the above procedure is repeated all the way to iteration $j = 10$. This approach is done for $e_1$ to $e_4$. As an illustrative example, the results of the reduced volume $\Omega_p$ for the respective iteration $j$ for each $e_{i_e}$ according to $UM_{y0}^1$ are shown in Figure 4.8. From this figure, the minimum value of $\Omega_p$ for each $e_{i_e}$ is obtained and the sensitivity index is computed:

$$S = 1 - \frac{\Omega_p}{\Omega_0} \tag{4.13}$$

where $\Omega_0$ denotes the initial volume before pinching. This sensitivity metric would then be utilized to rank $e_1$ to $e_4$ to which the results according to the respective UMs are shown in Table 4.3.
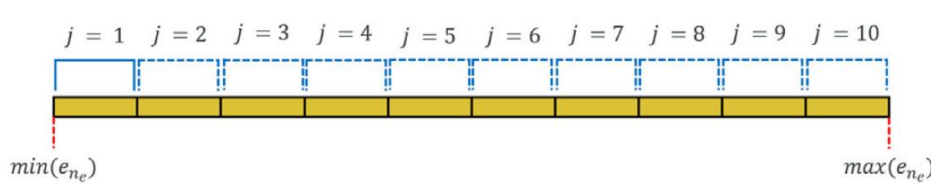


Figure 4.7 Illustration of the approach in identifying the maximum reduction of $\Omega$ from the pinching of $e_{i_e}$.
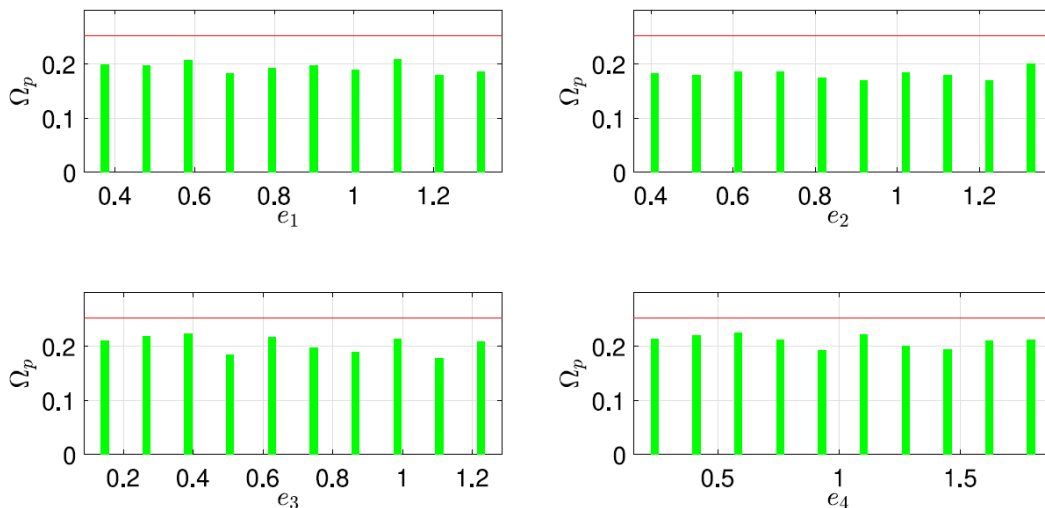


Figure 4.8 Results of $\Omega_p$ for different pinched intervals for $e_1$ to $e_4$ according to $UM_{y0}^1$. The red line denotes the initial volume $\Omega = 0.2521$, while the green bars represent the resulting $\Omega_p$.

Table 4.3 The ranking order of the epistemic model parameters based on their respective sensitivity index for the respective UMs.

| Rank | Uncertainty model $UM_{y0}^1$ | | | Uncertainty model $UM_{y0}^2$ | | |
|------|-----------|----------------|--------|-----------|----------------|--------|
|      | Parameter | Pinched bounds | $S$ | Parameter | Pinched bounds | $S$ |
| 1 | $e_2$ | $[1.1729, 1.2748]$ | 0.3300 | $e_2$ | $[0.8677, 0.9117]$ | 0.4542 |
| 2 | $e_3$ | $[1.0449, 1.1647]$ | 0.2972 | $e_3$ | $[0.3595, 0.4072]$ | 0.3240 |
| 3 | $e_1$ | $[1.1619, 1.2670]$ | 0.2882 | $e_1$ | $[0.5715, 0.5987]$ | 0.2654 |
| 4 | $e_4$ | $[0.8425, 1.0148]$ | 0.2272 | $e_4$ | $[0.8242, 0.9320]$ | 0.2556 |

Based on the results provided in Table 4.3, it is observed that $e_4$ is ranked the lowest in sensitivity according to both UMs. This implies that it is impossible to improve the knowledge on $e_4$ with the available model and data, making it impossible to extract or infer information on its true value, therefore contributing the highest degree of non-reducible epistemic uncertainty in the calibration of the UM. For this reason, the first refinement request to the challenge host was made for the lower bound of $e_4$, given the heavier left tail as seen from its histogram obtained via the distribution-free approach (in green) in Figure 4.4. Following this, a second round of sensitivity analysis was performed following the approach outlined above and accounting for the given refined bounds of $e_4$ and the results are detailed in Table 4.4. From the results, $e_3$ is consistently ranked within the bottom two according to both UMs which suggests that $e_3$ is the least informative parameter after $e_4$. Hence, the second refinement request was made for the lower bound of $e_3$ given the lack of such information according to both UMs as depicted in Figure 4.4. The resulting epistemic space, with the refined $e_3$ and $e_4$ bounds, constitutes the hyper-rectangle epistemic space denoted as $E_1$.

Table 4.4 The ranking order of the epistemic model parameters based on their respective sensitivity index for the respective UMs according for the refined bounds for $e_4$.

| Rank | Uncertainty model $UM_{y0}^1$ | | | Uncertainty model $UM_{y0}^2$ | | |
|------|-----------|----------------|--------|-----------|----------------|--------|
|      | Parameter | Pinched bounds | $S$ | Parameter | Pinched bounds | $S$ |
| 1 | $e_4$ | $[1.0224, 1.0575]$ | 0.5200 | $e_2$ | $[0.6036, 0.6476]$ | 0.5999 |
| 2 | $e_2$ | $[0.4652, 0.5625]$ | 0.5114 | $e_4$ | $[1.1276, 1.1632]$ | 0.4128 |
| 3 | $e_1$ | $[0.6333, 0.7367]$ | 0.5056 | $e_3$ | $[0.2166, 0.2642]$ | 0.4024 |
| 4 | $e_3$ | $[0.9227, 1.0393]$ | 0.4282 | $e_1$ | $[0.5715, 0.5988]$ | 0.3821 |

## 4.3.2 Updated uncertainty models

A second round of Bayesian model updating is performed with the bounds of the uniform priors for the respective epistemic model parameters defined by the hyper-rectangle $E_1$. The approach follows that outlines in Sections 4.2.3.1 and 4.2.3.2 from which $UM_{y1}^1$ and $UM_{y1}^2$ are obtained respectively. The corresponding numerical results of the updated bounds for each $e_{i_e}$ according to $UM_{y1}^1$ and $UM_{y1}^2$ are presented in Table 4.6.

A sensitivity analysis was done again following the methodology that is presented in Section 4.3.1 and the results

are given in Table 4.5. where the sensitivity ranking of each $e_{i_e}$ is the same as that in Table 4.3 for the respective UMs.

Table 4.5 The ranking order of the epistemic model parameters based on their respective sensitivity index for the respective UMs according for the refined space $E_1$.

| Rank | Uncertainty model $UM_{y0}^1$ | | | Uncertainty model $UM_{y0}^2$ | | |
|---|---|---|---|---|---|---|
| | Parameter | Pinched bounds | $S$ | Parameter | Pinched bounds | $S$ |
| 1 | $e_2$ | $[0.4447, 0.5340]$ | 0.4095 | $e_2$ | $[0.9514, 0.9726]$ | 0.3610 |
| 2 | $e_1$ | $[0.9500, 1.0413]$ | 0.2960 | $e_3$ | $[0.5675, 0.6173]$ | 0.3245 |
| 3 | $e_4$ | $[0.9274, 0.9585]$ | 0.2385 | $e_1$ | $[0.5550, 0.5725]$ | 0.2981 |
| 4 | $e_3$ | $[0.3333, 0.3590]$ | 0.2379 | $e_4$ | $[1.0225, 1.0584]$ | 0.2782 |

### 4.3.3  Results and discussion

The resulting model output of the response plot according to $UM_{y1}^1$ and $UM_{y1}^2$ are illustrated in Figure 4.9. From the figure, it can be observed that the response plots according to both UMs are well-fitted against $D_1$. However, such fitting is significantly tighter for the case of $UM_{y1}^2$ as seen in Figure 4.9 and this is supported by Figure 4.10 where it can also be seen that the resulting p-boxes of the response plot across all chosen time-slices are significantly narrower compared to $UM_{y1}^1$ whilst enclosing the ECDF for $D_1$. This observation is consistent to that discussed in Section 4.2.5 and concludes that the response plot according to $UM_{y1}^2$ is more representative of $D_1$.

Figure 4.11 illustrates the resulting p-boxes quantifying the uncertainty over the marginal distributions of $f_a$ by the respective UMs. From the figure, it can be seen that the p-boxes according to $UM_{y1}^2$ is significantly narrower and enclosed within that according to $UM_{y1}^1$ which verifies that the true marginal distributions of $f_a$ could lie within the p-boxes defined by $UM_{y1}^2$. The results by $UM_{y1}^1$ once again highlights its conservative nature in its uncertainty over $f_a$.
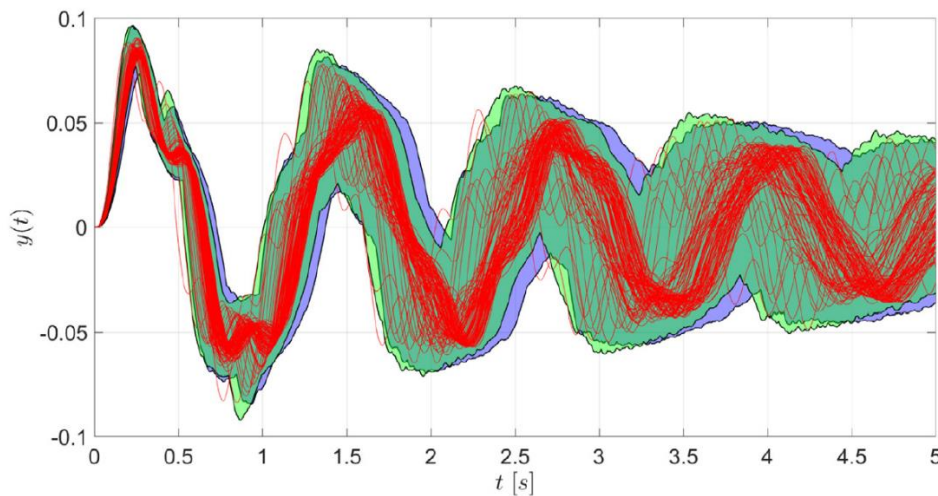


Figure 4.9 Output band from $\hat{y}$ according to $UM_{y1}^1$ (in blue) and $UM_{y1}^2$ (in green) along with the data sequence $D_1$ (in red) after calibration.
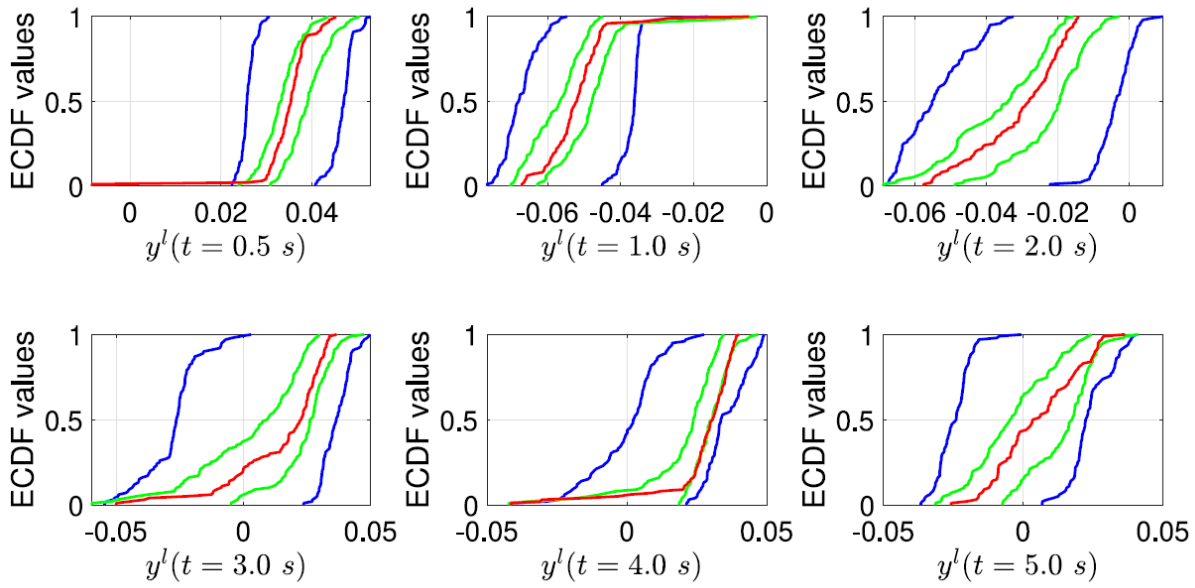
Figure 4.10 P-boxes of the model output from $\hat{y}$ according to $UM_{y1}^1$ (in blue) and $UM_{y1}^2$ (in green) at various time slices $t = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}\ s$. The red ECDF denotes the distribution of the data $D_1$.
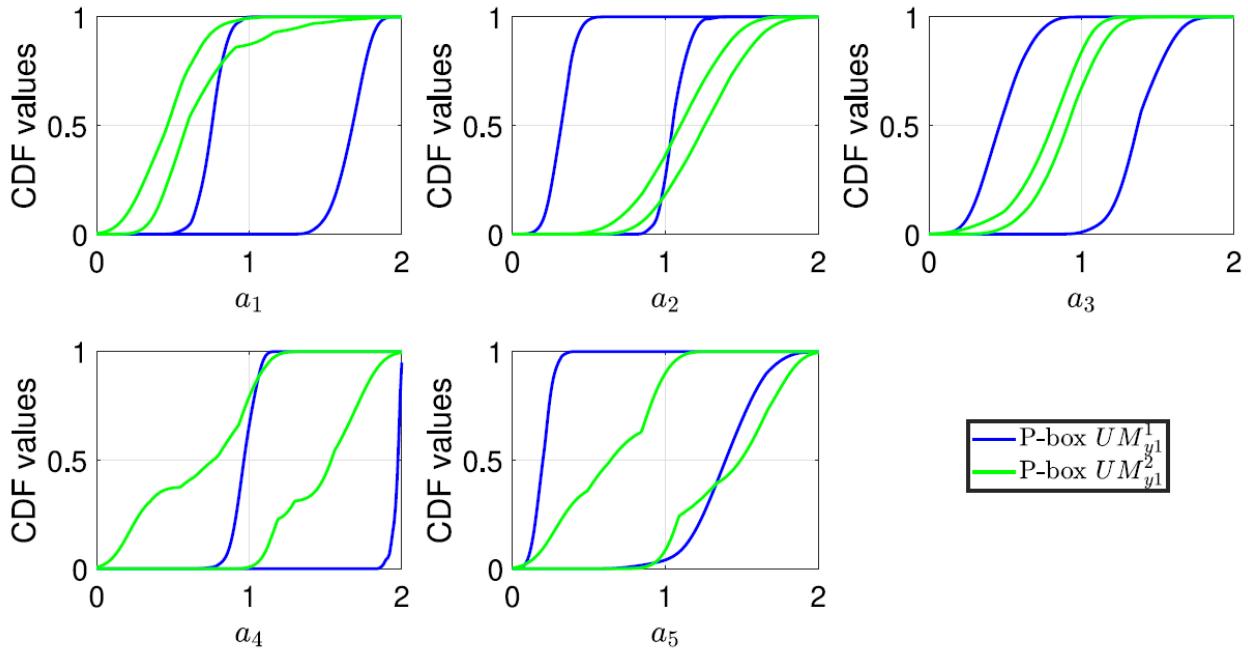


Figure 4.11 P-box for $a_1$ to $a_5$ obtained from the respective UMs.

Finally, results detailed in Table 4.6 shows that the uncertainty bounds over $e$ according to $UM_{y1}^2$ is significantly narrower and generally enclosed within that according to $UM_{y1}^1$. This result is supported by Figure 4.12 where it can be seen that the resulting histograms of the epistemic parameters obtained from $P(\boldsymbol{\Theta}|\boldsymbol{D}, \boldsymbol{M})$ given $f_a^5$ are consistently

narrower than that obtained from $P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M})$ given $f_a^1$. This verifies the results obtained by $UM_{y1}^2$ further highlighting its informative nature over $UM_{y1}^1$. For this reason, $UM_{y1}^2$ is finally chosen to address the subsequent tasks presented in this challenge.
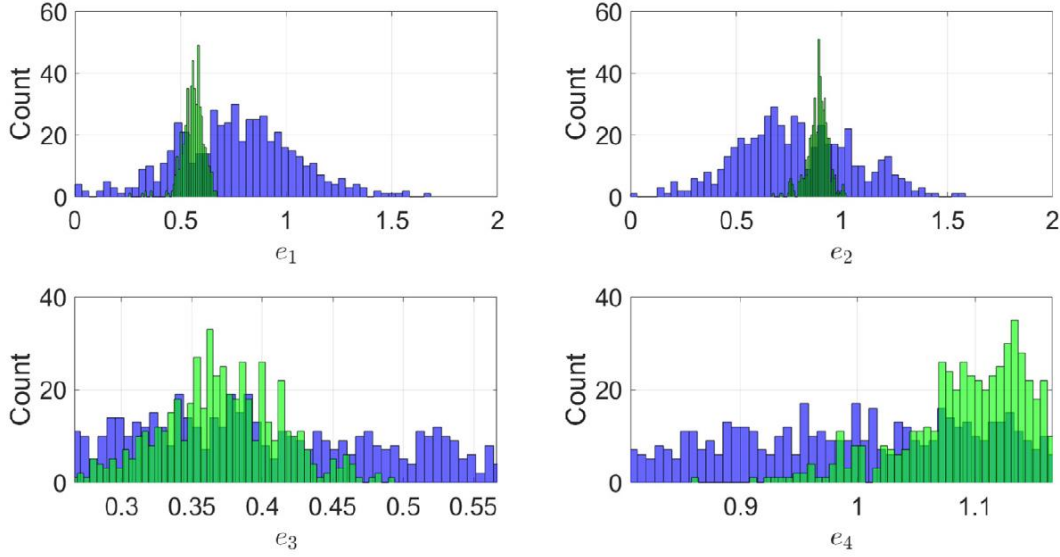


Figure 4.12 Histograms for $e_1$ to $e_4$ obtained from $P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M})$ given $f_a^1$ (in blue) and $f_a^5$ (in green).

Table 4.6 Updated epistemic space $E$ for $e_1$ to $e_4$ according to the respective UMs.

| Uncertainty model | $e_1$ | $e_2$ | $e_3$ | $e_4$ |
|---|---|---|---|---|
| $UM_{y0}^1$ | $[0.3182, 1.3787]$ | $[0.3574, 1.3771]$ | $[0.0827, 1.2870]$ | $[0.1486, 1.8828]$ |
| $UM_{y0}^2$ | $[0.4351, 0.7082]$ | $[0.5583, 1.0000]$ | $[0.0721, 0.5511]$ | $[0.6066, 1.6893]$ |
| $UM_{y1}^1$ | $[0.3097, 1.2306]$ | $[0.3522, 1.2487]$ | $[0.2819, 0.5400]$ | $[0.8337, 1.1461]$ |
| $UM_{y1}^2$ | $[0.4674, 0.6433]$ | $[0.7607, 0.9736]$ | $[0.2865, 0.4583]$ | $[0.9627, 1.1664]$ |

## 4.4    Task C: Reliability analysis of baseline design

The objective of this task is to perform a reliability analysis on the baseline design $\theta_{base}$ according to $UM_{y1}^2$ with respect to the individual requirements $g_{i_g}$, for $i_g = 1, 2, 3$. The requirements $g_2$ and $g_3$ are represented respectively as (Crespo and Kenny, 2021):

$$g_2(a, e, \theta) = \max_{t \in [2.5,5]\,\text{s}} |z_1(a, e, \theta, t)| - 0.02 \tag{4.14}$$

$$g_3(a, e, \theta) = \max_{t \in [0,5]\,\text{s}} |z_2(a, e, \theta, t)| - 4 \tag{4.15}$$

whose $z_1$ and $z_2$ denote the time-dependent response output of the integrated system associated with the given $\theta$. From

there, the worst-case performance function $w$ is defined:

$$w(a, e, \theta) = \max_{i_g = 1,2,3} g_{i_g}(a, e, \theta) \tag{4.16}$$

The system is defined to be system compliant for requirement $i_g$ when $g_{i_g} < 0$ whereby $g_1 < 0$ for the system to be stable; $g_2 < 0$ for the settling time of $z_1$ to be sufficiently fast; and $g_3 < 0$ finally for the energy consumption to be acceptable (Crespo and Kenny, 2021). Conversely, requirement $i_g$ is not satisfied when $g_{i_g} \geq 0$. Thus, for a fixed set of values of $\theta$ and $e$, the set of $a$ points where $g_{i_g} < 0$ is regarded as the "safe" domain, while the complement set is the "failure" domain. From this, the imprecise failure probability $R_{i_g}$ given requirement $g_{i_g}$ can then be defined as follows:

$$R_{i_g}(\theta) = \left[ \min_{e \in E} \mathbb{P}\left(g_{i_g} \geq 0\right), \max_{e \in E} \mathbb{P}\left(g_{i_g} \geq 0\right) \right] \tag{4.17}$$

through which the imprecise worst-case failure probability $R$ is defined:

$$R(\theta) = \left[ \min_{e \in E} \mathbb{P}(w \geq 0), \max_{e \in E} \mathbb{P}(w \geq 0) \right] \tag{4.18}$$

and finally, the severity of each requirement violation $s_{i_g}$ is defined:

$$s_{i_g}(\theta) = \max_{e \in E} \mathbb{E}\left[ g_{i_g} \middle| g_{i_g} \geq 0 \right] \cdot \mathbb{P}\left(g_{i_g} \geq 0\right) \tag{4.19}$$

where $\mathbb{P}(\cdot)$ indicates the probability operator, and $\mathbb{E}[\cdot | \cdot]$ denotes the conditional expectation.

### 4.4.1 Failure probability and severity computation

In this work, the computation and analysis of reliability metrics $R_{i_g}$, $R$, and $s_{i_g}$ are carried out through probability bounds analysis (PBA) with p-boxes (Ferson and Tucker, 2005).

To set up the p-box for each $g_{i_g}$, a double-loop Monte Carlo (Ali, 2012; Rocchetta et al., 2018) approach is used to generate $N_a \times N_e$ realizations of $g_{i_g}$ from inputs $\{a, e\}$ defined by $UM_{y1}^2$. In this approach, the outer-loop considers each of the $N_e$ realizations of $e$ obtained from the hyper-rectangle defined by $E$ according to the UM, while the inner-loop accounts for the $N_a$ realizations of $a$ from $f_a$. So as to ensure that the different failure domains are well-explored, in particular small failure regions, and that the epistemic uncertainties well represented, we set $N_a = 10000$ and $N_e = 500$. From there, the p-box is then constructed from the bounds of the distribution of $N_e$ ECDFs, in which each ECDF

comprising of $N_a$ values of $g_{i_g}$. As an illustration, the resulting p-boxes for $g_1$ to $g_3$ are illustrated in Figure 4.39.

From the p-box of a given $g_{i_g}$, its values computed at $g_{i_g} = 0$ has a lower and upper bound value denoted as $\underline{P}_{i_g}$ and $\overline{P}_{i_g}$ respectively. Following which, $R_{i_g}(\theta_{base})$ can be approximated according to:

$$R_{i_g}(\theta_{base}) = \left[1 - \overline{P}_{i_g}, 1 - \underline{P}_{i_g}\right] \tag{4.20}$$

To approximate $R(\theta_{base})$, the $N_a \times N_e$ matrix of $w$ is constructed by taking the element-wise maximum value between $g_1$, $g_2$, and $g_3$ as suggested in Equation (4.16). A p-box is then constructed for $w$ in similar fashion as $g_{i_g}$ from which the resulting lower and upper bound values, denoted as $\underline{W}$ and $\overline{W}$ respectively, are obtained at $w = 0$. $R(\theta_{base})$ can then be approximated according to:

$$R(\theta_{base}) = \left[1 - \overline{W}, 1 - \underline{W}\right] \tag{4.21}$$

To approximate $s_{i_g}$, the numerical value of $\mathbb{E}\left[g_{i_g} \mid g_{i_g} \geq 0\right]$ is necessary to be approximated first. This can be done as follows: Considering a realization of the ECDF for $g_{i_g}$ for a given $e$, a numerical integration is performed for obtaining the area of the region above the ECDF plot between $g_{i_g} = 0$ to its maximum value, denoted as $g_{i_g}^{max}$, for which the ECDF is defined. This is done according to:

$$\mathbb{E}\left[g_{i_g} \mid g_{i_g} \geq 0\right] \approx \sum_{k=1}^{N_a} \mathbf{1}_{\left(g_{i_g}\right)_k \geq 0} \cdot \left(g_{i_g}\right)_k \cdot \mathbb{P}\left(g_{i_g} \geq \left(g_{i_g}\right)_k\right) \tag{4.22}$$

where $\mathbf{1}_{\left(g_{i_g}\right)_k \geq 0}$ denotes the indicator function which gives the value 1 when $\left(g_{i_g}\right)_k \geq 0$ and 0 otherwise. This value of $\mathbb{E}\left[g_{i_g} \mid g_{i_g} \geq 0\right]$ is then multiplied by $\mathbb{P}\left(g_{i_g} \geq 0\right)$ determined from its ECDF that gives the nominal severity index associated with the given $e$. From which, the actual severity index $s_{i_g}$ can be finally computed according to Equation (4.19).

The numerical results to the aforementioned reliability metrics are presented in Table 4.14 where it is observed that the failure probability with the highest upper-bound value is $R_2$ and, thus, contributes the most of the worst-case failure probability $R$. In addition, it is noted that the severity $s_2$ is the highest among all $s_{i_g}$ which indicates that the failure $g_2 \geq 0$ is classified as a high-probability event with a large impact on the system relative to the failures $g_1 \geq 0$ and $g_3 \geq 0$. Further investigations to this are done in Section 4.4.3.

## 4.4.2 Sensitivity analysis

In this analysis, the epistemic uncertainties are ranked according to the construction of $R(\theta_{base})$ resulting from their reduction. This will be done via the adaptive pinching (Tucker and Ferson, 2006) approach shown in Section 4.3. The results of the reduced interval $R(\theta_{base})$ for the respective iteration $j$ for each $e_{i_e}$ are depicted in Figure 4.13. From the figure, the maximum reduction of $R(\theta_{base})$ for each $e_{i_e}$ is determined. Such information would then be utilized to rank $e_1$ to $e_4$ as shown in Table 4.7.
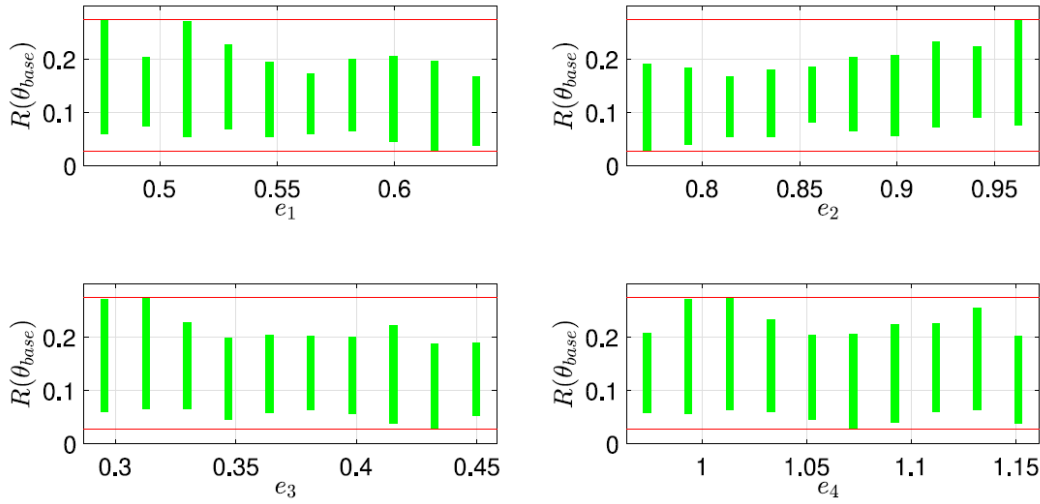


Figure 4.13 Results of the reduced $R(\theta_{base})$ bounds for different pinched intervals for $e_1$ to $e_4$. The red lines represent the initial bounds of $[0.0294, 0.2721]$, while the green bars represent the reduced bounds.

Table 4.7 The ranking order of the epistemic model parameters based on the maximum possible reduction in $R(\theta_{base})$ interval according to $UM_{y1}^2$.

| Rank | Parameter | Pinched bounds | $R(\theta_{base})$ Before pinching | After pinching |
|---|---|---|---|---|
| 1 | $e_2$ | $[0.8458, 0.8690]$ | $[0.0294, 0.2721]$ | $[0.0812, 0.1851]$ |
| 2 | $e_1$ | $[0.5555, 0.5731]$ | $[0.0294, 0.2721]$ | $[0.0588, 0.1717]$ |
| 3 | $e_3$ | $[0.4411, 0.4582]$ | $[0.0294, 0.2721]$ | $[0.0533, 0.1888]$ |
| 4 | $e_4$ | $[0.9639, 0.9836]$ | $[0.0294, 0.2721]$ | $[0.0588, 0.2067]$ |

## 4.4.3 Identifying different transitions to failure

From the $N_a \times N_e$ matrix of all $g_{i_g}$, the realizations of $\{a, e\}$ are grouped into seven distinct categories of failure: (1) $g_1 \geq 0$; (2) $g_2 \geq 0$; (3) $g_3 \geq 0$; (4) $g_1, g_2 \geq 0$; (5) $g_1, g_3 \geq 0$; (6) $g_2, g_3 \geq 0$; and (7) $g_1, g_2, g_3 \geq 0$. The statistics summarizing the number of $\{a, e\}$ realizations in each failure category is presented in Table 4.12.

From the table, it can be seen that the most common failure type for $\theta_{base}$ according to $UM_{y1}^2$ is $g_2 \geq 0$ while

the least likely failure type would be $g_1, g_2, g_3 \geq 0$. No failure of type $g_1, g_3 \geq 0$ has occurred. In order to provide a quantitative understanding of the characteristic of each failure and its severity, the response curves $z_1(t)$ and $z_2(t)$ are plotted for 25 representative sample sets of $\{a, e\}$ in each failure category in Figures 4.14 and 4.15.
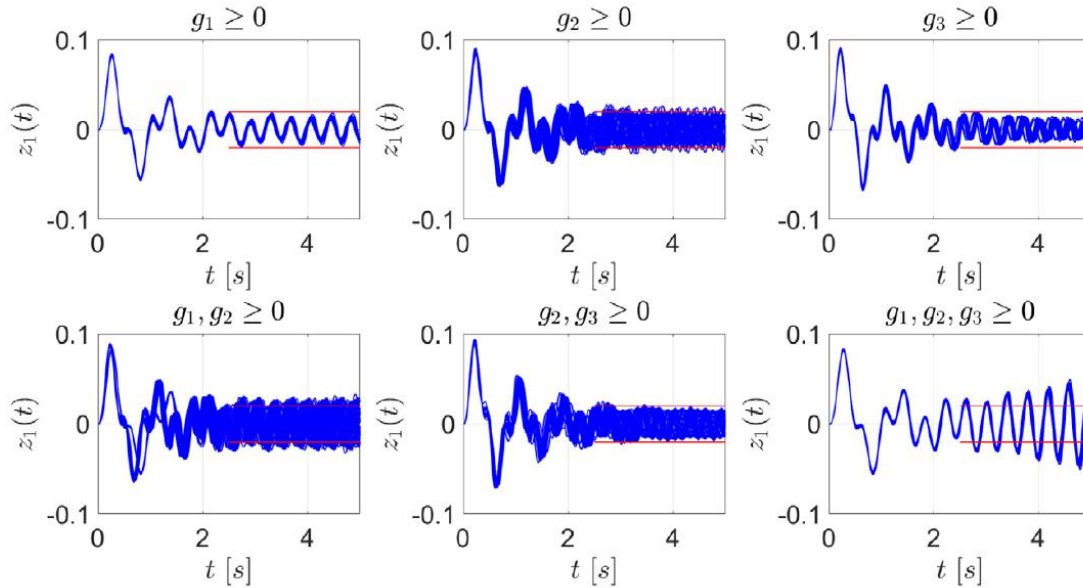


Figure 4.14 Response plot of $z_1(t)$ corresponding to 25 representative realizations of $\{a, e\}$ for each failure type. The red lines denote the safety limits.
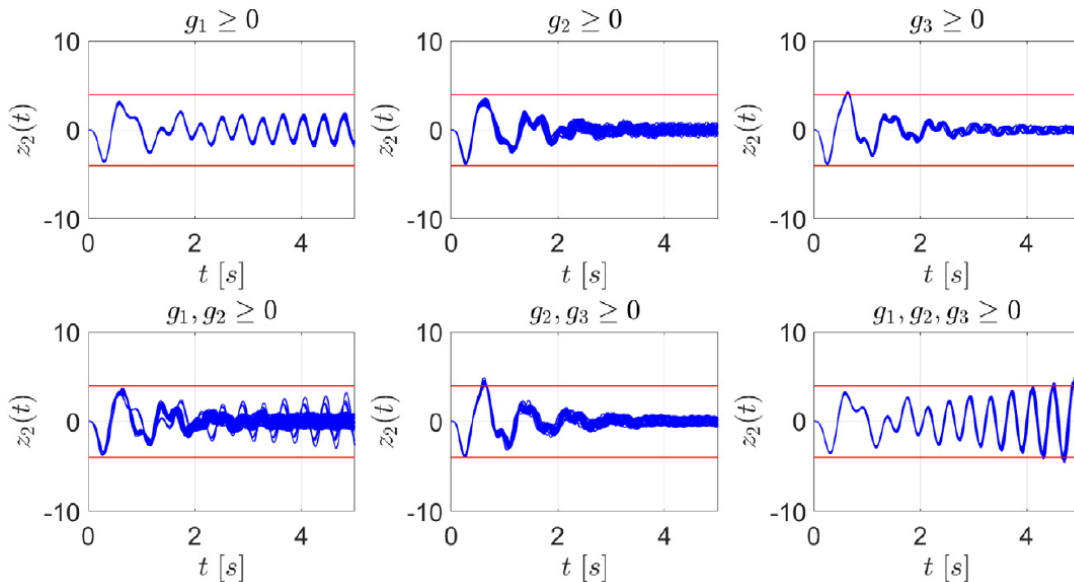


Figure 4.15 Response plot of $z_2(t)$ corresponding to 25 representative realizations of $\{a, e\}$ for each failure type. The red lines denote the safety limits.

From the figures, it can be seen that the failure type of the worst severity is that of $g_1, g_2, g_3 \geq 0$ where it can be

seen that the plots for $z_1(t)$ and $z_2(t)$ present an unstable behavior with an increasing amplitude as time $t$ increases. This results in the largest degree of deviation from the safety limits which causes the aforementioned failure type to contribute the most towards the severity values of $s_1$, $s_2$, and $s_3$. In addition, it can also be observed that the $z_2(t)$ plot for failure type $g_1, g_2 \geq 0$ also exhibit an increasing amplitude with time, although still within the safety limits and not as pronounced as that for $g_1, g_2, g_3 \geq 0$. Such unstable behavior is due to the common failure of $g_1$ which concerns the stability of the system's behavior. Hence, it is important to identify a new design point $\theta_{new}$ in Section 4.5 so that the likelihood of occurrence of failure types $g_1, g_2 \geq 0$ and $g_1, g_2, g_3 \geq 0$ is as close to zero as possible.

To identify the representative realizations of $\delta \in A \times E$ having a comparatively large likelihood near the failure domain, the methodology is as follows: For each $i_g$, realizations of $\{a, e\}$ from $UM_{y1}^2$ which correspond to the top 500 numerically least negative matrix elements of $g_{i_g}$ are identified. These realizations will be classified as those "near" the failure domain. Following this, the likelihood values of the identified $\{a, e\}$ are computed by calculating the PDF value of the SDF for the corresponding realization of $a$. This generates 500 values of likelihood from which the sample set $\{a, e\}$ having the top five-percentile likelihood values are identified. Let these sample sets be denoted as $\{a, e\}_{nf}^{g_{i_g}}$. This yields 25 sets of $\{a, e\}_{nf}^{g_{i_g}}$ remaining that will constitute the realizations with comparatively large likelihood near the failure domain of $g_{i_g}$. This procedure is implemented for $g_1$ to $g_3$ and the resulting 25 sets of $\{a, e\}_{nf}^{g_{i_g}}$ identified for each $g_{i_g}$ are presented as parallel plots in Figure 4.16 and whose corresponding response plots of $z_1(t)$ and $z_2(t)$ are illustrated in Figure 4.17.
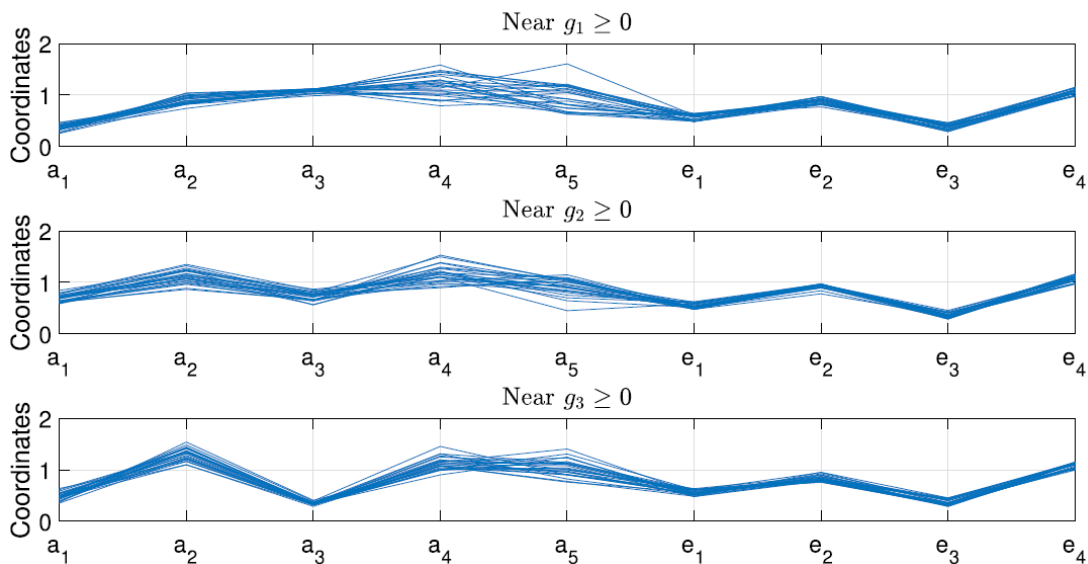


Figure 4.16 Parallel plots of $\{a, e\}_{nf}^{g_1}$, $\{a, e\}_{nf}^{g_2}$, and $\{a, e\}_{nf}^{g_3}$ from $UM_{y1}^2$.
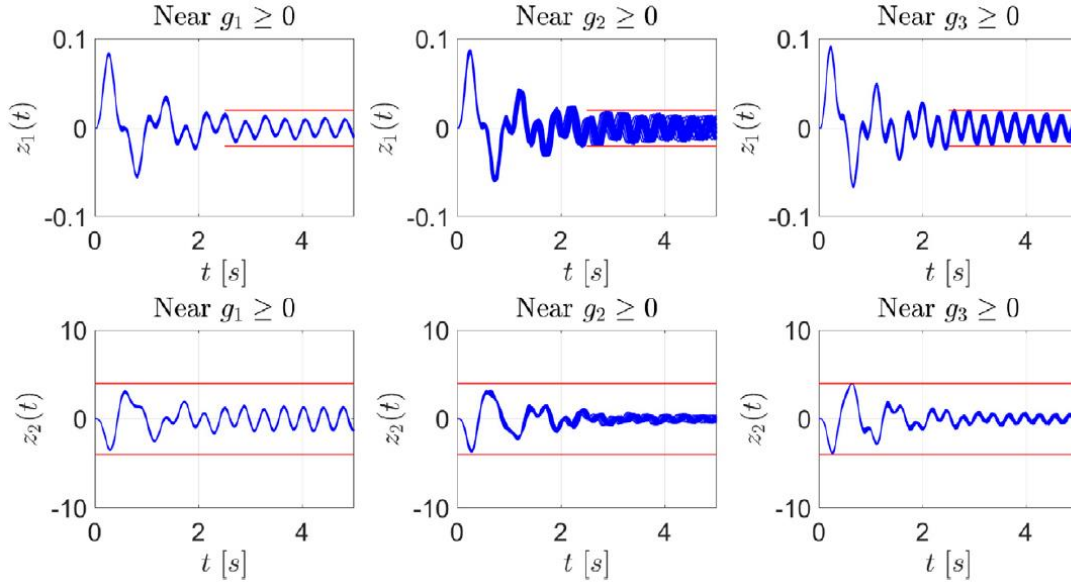
Figure 4.17 Response plot of $z_1(t)$ and $z_2(t)$ of $\{a, e\}_{nf}^{g_1}$, $\{a, e\}_{nf}^{g_2}$, and $\{a, e\}_{nf}^{g_3}$ from $UM_{y1}^2$ near the respective failure domains. The red lines denote the safety limits.

From Figure 4.17, the following observations are made: Near the failure domain of $g_1$, the response plots $z_1(t)$ and $z_2(t)$ are all well-within the safe limits which indicates that the sample sets $\{a, e\}_{nf}^{g_1}$ all lie within the safe domain near the boundary of $g_1 \geq 0$ domain. Near the failure domain of $g_2$, the response plots $z_1(t)$ and $z_2(t)$ are all within safe limits as well. Based on the fast-decaying characteristics of the response plot for $z_2(t)$ and the stable behavior of the plot for $z_1(t)$, it can be inferred that the requirement of $g_1$ is at the same time satisfied. However, it is noted that at approximately $t \in [2.7, 2.9]$ s, the plots are extremely close to the lower and upper boundaries respectively, verifying that the realizations of $\{a, e\}_{nf}^{g_2}$ are in the safe domain near the boundary of $g_2 \geq 0$. Finally, near the failure domain of $g_3$, it can be seen from the plot of $z_1(t)$ that the curves exceeded the safety limits at approximately $t = \{2.6, 2.9, 3.0\}$ s which indicates the failure of $g_2$. On the other hand, the plot of $z_2(t)$ are all within the safety limits although it can be observed that the plots at approximately $t = 6$ s are extremely close to the upper boundary. Due to the stable behavior of the plots of $z_1(t)$ and $z_2(t)$, it can be inferred that the requirement of $g_1$ is satisfied. Hence, it can be concluded that the realizations of $\{a, e\}_{nf}^{g_3}$ lie within the domain of $g_2 \geq 0$ near the boundary of $g_2, g_3 \geq 0$.

## 4.5    Task D: Reliability-based design identification

The objective of this task is to identify a new design point $\theta_{new}$ such that the likelihood of failure types $g_2, g_3 \geq 0$ and $g_1, g_2, g_3 \geq 0$ occurring is reduced to as close to zero as possible given that such failures are responsible for the

unstable behavior of the system. To achieve this, $\theta_{new}$ has to be optimized so that it satisfies the following conditions:

(1)  Minimize the upper-bound of $R$;

(2)  Reduce the worst-case severity metric $\bar{s}$ defined as (Gray et al, 2020):

$$\bar{s}(\theta) = \max_{e \in E} \mathbb{E}[w|w \geq 0] \cdot \mathbb{P}(w \geq 0) \tag{4.23}$$

To perform the optimization procedure, the generalized non-intrusive imprecise stochastic simulation (NISS) method is employed to approximate a solution to $\theta_{new}$ (Song et al., 2019). For the benefit of the readers, a description to the generalized NISS is provided in Section 4.5.1.

## 4.5.1   Generalized non-intrusive imprecise stochastic simulation

The generalized NISS approach provides a surrogate model for computing $R$ through the random sampling high-dimensional model representation (RS-HDMR) decomposition defined as:

$$
\begin{aligned}
R(e,\theta) = R_0 &+ \sum_{i_e}^{4} R_{E_{i_e}}(e_{i_e}) + \sum_{1 \leq i_e < j_e \leq 4} R_{E_{i_e j_e}}(e_{i_e j_e}) + \sum_{i_\theta}^{9} R_{\widetilde{\Theta}_{i_\theta}}(\theta_{i_\theta}) + \sum_{1 \leq i_\theta < j_\theta \leq 9} R_{\widetilde{\Theta}_{i_\theta j_\theta}}(\theta_{i_\theta j_\theta}) \\
&+ \sum_{1 \leq i_e \leq 4, 1 \leq i_\theta \leq 9} R_{E_{i_e}\widetilde{\Theta}_{j_\theta}}(e_{i_e}\theta_{j_\theta}) + \cdots + R_{E\widetilde{\Theta}}(e,\theta)
\end{aligned}
\tag{4.24}
$$

where the constant terms and the first two order component functions are defined respectively as:

$$R_0 = \mathbb{E}_{E\widetilde{\Theta}}[R(e,\theta)]$$

$$R_{E_{i_e}}(e_{i_e}) = \mathbb{E}_{E_{-i_e}\widetilde{\Theta}}[R(e,\theta)] - R_0$$

$$R_{E_{i_e j_e}}(e_{i_e j_e}) = \mathbb{E}_{E_{-i_e j_e}\widetilde{\Theta}}[R(e,\theta)] - R_{E_{i_e}} - R_{E_{j_e}} - R_0$$

$$R_{\widetilde{\Theta}_{i_\theta}}(\theta_{i_\theta}) = \mathbb{E}_{E\widetilde{\Theta}_{-i_\theta}}[R(e,\theta)] - R_0$$

$$R_{\widetilde{\Theta}_{i_\theta j_\theta}}(\theta_{i_\theta j_\theta}) = \mathbb{E}_{E\widetilde{\Theta}_{-i_\theta j_\theta}}[R(e,\theta)] - R_{\widetilde{\Theta}_{i_\theta}} - R_{\widetilde{\Theta}_{j_\theta}} - R_0$$

$$R_{E_{i_e}\widetilde{\Theta}_{j_\theta}}(e_{i_e}\theta_{j_\theta}) = \mathbb{E}_{E_{-i_e}\widetilde{\Theta}_{-i_\theta}}[R(e,\theta)] - R_{E_{i_e}} - R_{\widetilde{\Theta}_{j_\theta}} - R_0$$

Here, $\mathbb{E}_{E\widetilde{\Theta}}[\cdot]$ denotes the expectation operator as a function of $e$ and $\theta$, $\mathbb{E}_{E_{-i_e}\widetilde{\Theta}}[\cdot]$ denotes the expectation operator as a

function of $\theta$ and the three-dimensional vector $e_{-i_e}$ which contains all elements of $e$ except those of component $e_{i_e}$, $\mathbb{E}_{E\tilde{\Theta}_{-i_\theta}}[\cdot]$ denotes the expectation operator as a function of $e$ and the $(n_\theta - 1)$-dimensional vector $\theta_{-i_\theta}$ which contains all elements of $\theta$ except those of component $\theta_{i_\theta}$, $\mathbb{E}_{E_{-i_e j_e}\tilde{\Theta}}[\cdot]$ means the expectation operator as a function of $\theta$ and all elements of $e$ except $e_{i_e j_e}$, $\mathbb{E}_{E\tilde{\Theta}_{-i_\theta j_\theta}}[\cdot]$ indicates the expectation operator as a function of $\theta$ and all elements of $e$ except $\theta_{i_\theta j_\theta}$, and $\mathbb{E}_{E_{-i_e}\tilde{\Theta}_{-i_\theta}}[\cdot]$ denotes the expectation operator as a function of $e_{-i_e}$ and $\theta_{-i_\theta}$.

The above component functions are approximated numerically via the extended Monte Carlo simulation (EMCS) (Wei et al., 2014; Rezaie et al., 2007) of $N_{EMCS} = 50000$ realizations of the joint sample sets of $\{a_i, e_i, \theta_i\}_{i=1,\cdots,N_{EMCS}}$. Realizations of $a_i$ and $e_i$ can be generated from $UM_{y1}^2$ while those of $\theta_i$ are generated from its hyper-rectangular space $\tilde{\Theta}$ which will be discussed later in this section. From these realizations, a bootstrap scheme is implemented to compute the variance of each estimator, with the number of bootstrap replications set to be 20 in this work. This allows for the computation of the Sobol' sensitivity index (Sobol', 2001; Saltelli et al., 2008) for each of these component function as a by-product of the generalized NISS technique (Song et al., 2019). Given that the Sobol' sensitivity indices measure the relative importance of each component function, the component functions with relatively small Sobol' index can be neglected (Sobol', 2001). As such, based on initial analysis, it was found that we can just consider the first-order component functions in the optimization procedure. The readers can refer to Song et al. (2019) for more details to the generalized NISS technique.

It should be noted that the generalized NISS has been originally introduced for the efficient propagation of hybrid uncertainties avoiding double-loop Monte Carlo method. A key aspect of the generalized NISS is that it allows for an explicit formulation of the functional dependence of the probability of failure with respect to the epistemic parameters, one parameter at a time. In this work, the design variables are treated like epistemic parameters in the NISS framework, simplifying the optimization problem to a set of one-dimensional searches. However, no bounds are given to the design parameters limiting the optimization, thus the optimization procedure is carried out for $\theta_{new}$ in an iterative approach such that the first-order component function $R_{\tilde{\Theta}_{i_\theta}}$ is minimized which is assumed to also minimize the upper bound of $R$ within the target design space. Based on this assumption, the procedure is undertaken as follows:

(1) At iteration $j = 1$, the initial nine-dimensional hyper-rectangle of the design space $\tilde{\Theta}^j$ is given so that component $i_\theta$ has bounds $\pm N_b^j = 5\%$ of its nominal value: $[0.95, 1.05] \times (\theta_{base})_{i_\theta}$;

(2) Identify the candidate optimal design point $\theta_c^j$ such that $\theta_c^j \in \tilde{\Theta}^j$;

(3) If a unique local minimum value of $R_{\widetilde{\Theta}_{i_\theta}}$ exists for a particular design parameter $\left(\theta_c^j\right)_{i_\theta}$, the parameter will then be assigned that fixed value and will not be updated in subsequent iterations. Let the total number of such optimized design parameters up to iteration $j$ be denoted by $N_{op}^j$, for $N_b^j = 0, \cdots, 9$;

(4) Define the new design space $\widetilde{\Theta}^{j+1}$ whose $\left(9 - N_{op}^j\right)$-dimensional hyper-rectangle is defined such that component $i_\theta$, which is not optimized, has increased bounds of $\pm N_b^{j+1} = (j+1) \times N_b^{j=1}$ % of its nominal value;

(5) Set $j = j + 1$ and repeat steps 2 to 4 until the local minimum of $R_{\widetilde{\Theta}_{i_\theta}}$ is identified for all possible $\theta_{i_\theta}$ for which an optimal point exists.

The entire recursive optimization procedure involved six iterations and as a representative graphical illustration, Figure 4.18 presents the optimization result obtained at iteration $j = 3$. From the figure, it is observed that the local optimal points corresponding to the local minimum of $R_{\widetilde{\Theta}_{i_\theta}}$ are identified for $\theta_3$ and $\theta_7$. In addition, it can also be seen that the remaining components of $\theta$ tend to be increasing or decreasing monotonically in general. For these components, their bounds will be increased according to step 4 of the NIIS optimization procedure, and subsequently step 5, until the local optimal points are identified for all nice components of $\theta$.
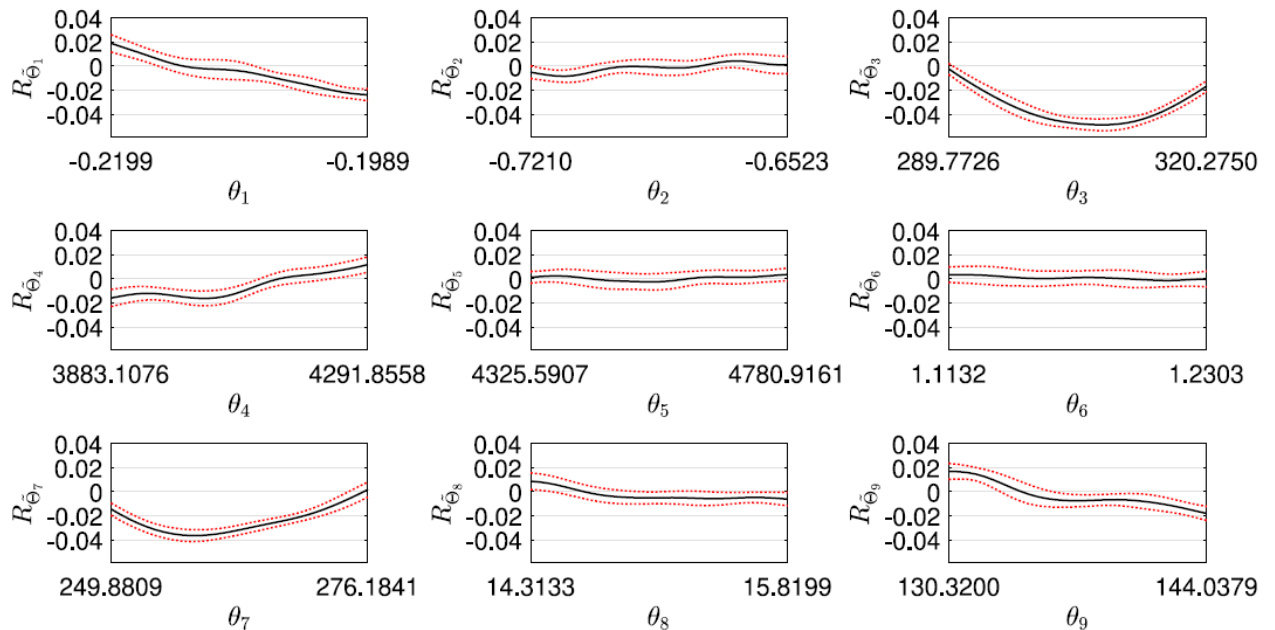


Figure 4.18 Results of the optimization of each design parameter for $\theta_{new}$ from the generalized NISS method at iteration $j = 3$. The red dotted lines denote the 95 % confidence interval bounds.

## 4.5.2   Reliability analysis of new design

The failure probability and severity computation on $\theta_{new}$ is done according to $UM^2_{y1}$ following the approach that is outlined in Section 4.4.1. The numerical results to the reliability metrics $R_{i_g}$, $R$, and $s_{i_g}$ are presented in Table 4.14. From the table, it can be seen that the upper-bounds of the failure probabilities $R_1$ to $R_3$ as well as the severities $s_1$ to $s_3$ have been reduced tremendously which validates the improvement of the design $\theta_{new}$ over $\theta_{base}$. This can be seen from the three-fold reduction in the upper-bound of $R_1$, two-fold reduction in $R_2$ and $R$, and a ten-fold reduction in $\bar{s}$ between $\theta_{base}$ and $\theta_{new}$. Such results highlight the effectiveness of the optimization procedure and the identified $\theta_{new}$.

Next, the sensitivity analysis is performed on the epistemic parameters $e_1$ to $e_4$ where they are ranked according to the maximum possible reduction on $R(\theta_{new})$ bounds. The methodology follows that outlined in Section 4.4.2 and the resulting illustrative plots from the analysis is shown in Figure 4.19. From which, the sensitivity ranking of $e_1$ to $e_4$, together with the corresponding reduced bounds of $R(\theta_{new})$, are presented in Table 4.8. From the results, it can be observed that the ranking order is consistent with the results obtained in Section 4.4.2 (i.e., see Table 4.7).
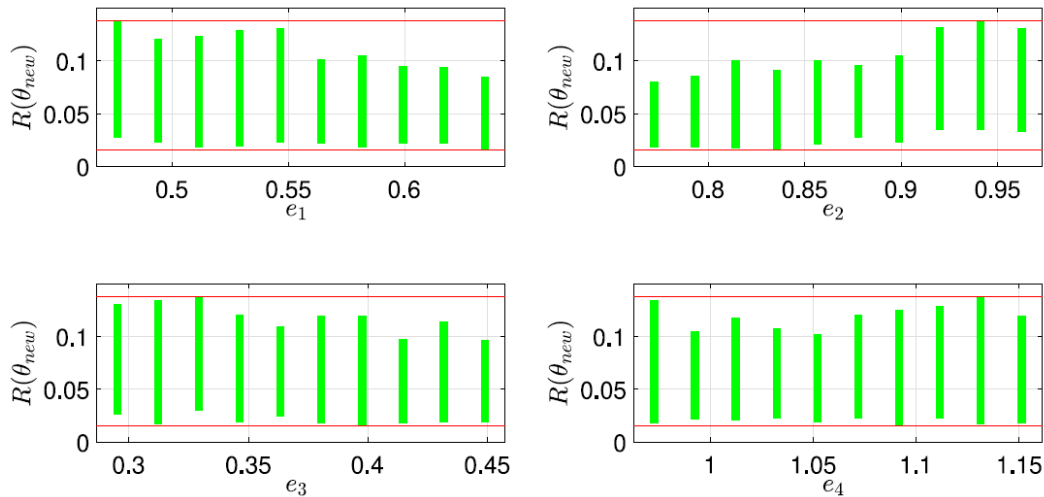


Figure 4.19 Results of the reduced $R(\theta_{new})$ bounds for different pinched intervals for $e_1$ to $e_4$. The red lines represent the initial bounds of $[0.0058, 0.1418]$, while the green bars represent the reduced bounds.

Table 4.8 The ranking order of the epistemic model parameters based on the maximum possible reduction in $R(\theta_{new})$ interval according to $UM^2_{y1}$.

| Rank | Parameter | Pinched bounds | $R(\theta_{base})$ | |
|---|---|---|---|---|
| | | | Before pinching | After pinching |
| 1 | $e_2$ | $[0.7613, 0.7825]$ | $[0.0058, 0.1418]$ | $[0.0183, 0.0798]$ |
| 2 | $e_1$ | $[0.6257, 0.6433]$ | $[0.0058, 0.1418]$ | $[0.0153, 0.0850]$ |
| 3 | $e_3$ | $[0.4404, 0.4575]$ | $[0.0058, 0.1418]$ | $[0.0189, 0.0960]$ |
| 4 | $e_4$ | $[0.9826, 1.0025]$ | $[0.0058, 0.1418]$ | $[0.0216, 0.1049]$ |

Finally, the failure analysis is performed on $\theta_{new}$ following the methodology that is outlined in Section 4.4.3. The resulting statistics summarizing the number of realizations of $\{a, e\}$ in each failure category in presented in Table 4.12 where it can be seen that likelihood of occurrences for all failure types have been reduced significantly, most notably for failure types $g_3 \geq 0$, $g_1, g_3 \geq 0$, and $g_1, g_2, g_3 \geq 0$ which have been reduced to zero. Failure type $g_2 \geq 0$ still has the highest likelihood because it has the highest number of realizations among the different failure types as per $\theta_{base}$. Following which, the sample sets $\{a, e\}_{nf}^{g_1}$, $\{a, e\}_{nf}^{g_2}$, and $\{a, e\}_{nf}^{g_3}$ are identified and the resulting parallel plots are given in Figure 4.20 while the response plots of $z_1(t)$ and $z_2(t)$ are presented in Figure 4.21.
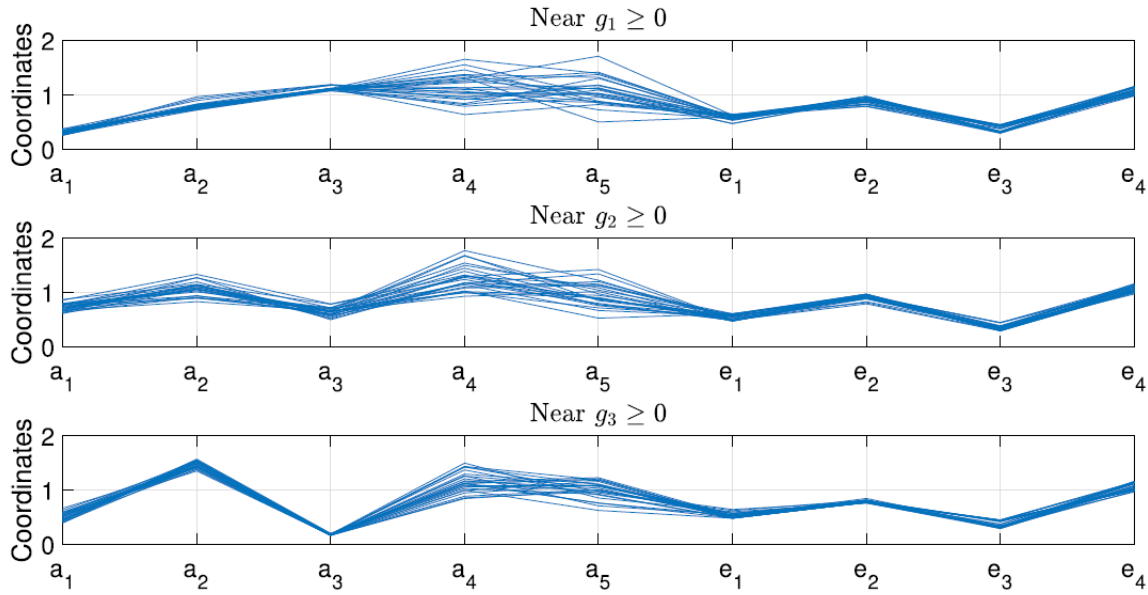


Figure 4.20 Parallel plots of $\{a, e\}_{nf}^{g_1}$, $\{a, e\}_{nf}^{g_2}$, and $\{a, e\}_{nf}^{g_3}$ from $UM_{y1}^2$.

As shown in Figure 4.21, the response plots of $z_1(t)$ for $\{a, e\}_{nf}^{g_1}$ exceeds the upper-bound of the safe boundary at approximately $t = 2.6$ s which indicates the presence of failure $g_2 \geq 0$. The plots for $z_2(t)$, on the other hand, are well-within the safety limits. This means that the sample set $\{a, e\}_{nf}^{g_1}$ lie within the domain of $g_2 \geq 0$ near the boundary of $g_1 \geq 0$. For $\{a, e\}_{nf}^{g_2}$, the response plots of $z_1(t)$ and $z_2(t)$ are within the safety limits although it can also be seen that the $z_1(t)$ plots are extremely close to both the upper- and lower-bounds of the safe boundary which verifies that the identified $\{a, e\}_{nf}^{g_2}$ are near $g_2 \geq 0$. As seen from their stable behavior, it can be inferred that the requirement $g_1$ is satisfied. This indicates that the sample set $\{a, e\}_{nf}^{g_2}$ lie within the safe domain near the boundary of $g_2 \geq 0$. Moreover, for $\{a, e\}_{nf}^{g_3}$, the response plots of $z_1(t)$ exceeds the upper- and lower-bounds of the safe boundary at approximately $t = \{2.5, 2.7, 2.8, 3.1, 3.4\}$ s which indicates the response of failure $g_2 \geq 0$. The plots for $z_2(t)$, on the other hand, are

well-within the safety limits. Given the stable behavior of the plots, it can be inferred that the requirement of $g_1$ is at the same time satisfied. This implies that the sample set $\{a, e\}_{nf}^{g_3}$ lie within the domain of $g_2 \geq 0$ near the boundary of $g_2, g_3 \geq 0$.
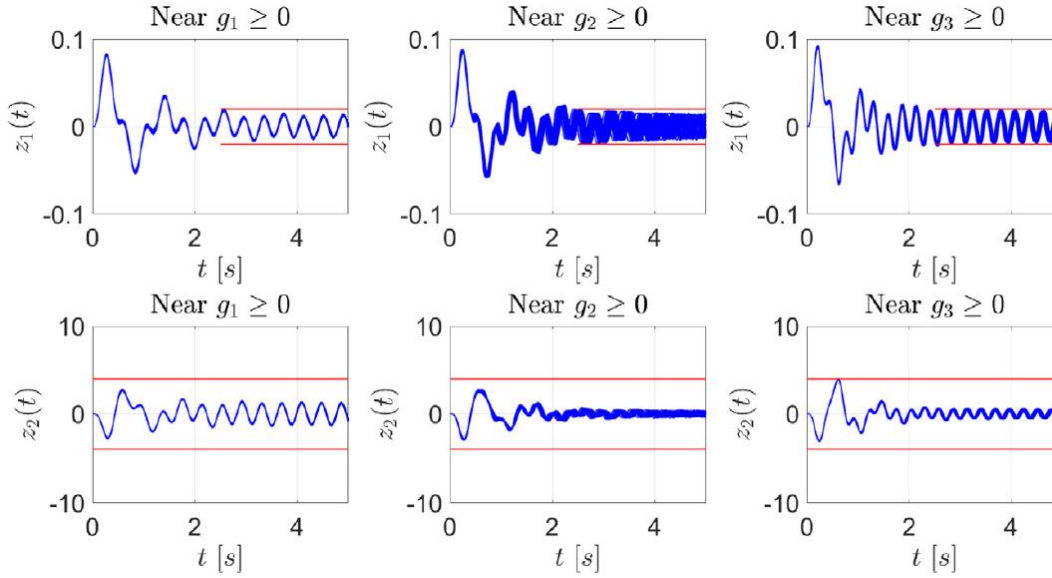


Figure 4.21 Response plot of $z_1(t)$ and $z_2(t)$ of $\{a, e\}_{nf}^{g_1}$, $\{a, e\}_{nf}^{g_2}$, and $\{a, e\}_{nf}^{g_3}$ from $UM_{y1}^2$ near the respective failure domains. The red lines denote the safety limits.

A quantitative study is also done for the different failure types with non-zero likelihood. As per what was done in Section 4.4.3, the response plots of $z_1(t)$ and $z_2(t)$ for the 25 representative sample sets of $\{a, e\}$ for each failure type (five for failure type $g_2, g_3 \geq 0$) are presented in Figures 4.22 and 4.23 respectively. From Figure 4.23, it can be seen that the unstable behavior in $z_2(t)$ for failure type $g_1, g_3 \geq 0$ is still present along with those showing a stable behavior. In addition, Figure 4.22 shows that the $z_1(t)$ response plots for failure type $g_1, g_3 \geq 0$ exceeds the safety limits to the largest extent relative to the other failure types which indicates that the failure type contributes the most to the severity $s_1$, $s_2$, and $\bar{s}$. This motivates the further necessity for identifying $\theta_{final}$ such that the number of realizations of $\{a, e\}$ corresponding to such unstable behavior is minimized as much as possible. The identification of $\theta_{final}$ will be done in Section 4.6.3.

## 4.6    Task E: Model update and design tuning

The objective of this task is to improve the current UM and identify an improved design point $\theta_{final}$ based on the observations of $z_1(t)$ and $z_2(t)$ from the integrated system corresponding to $\theta_{new}$.
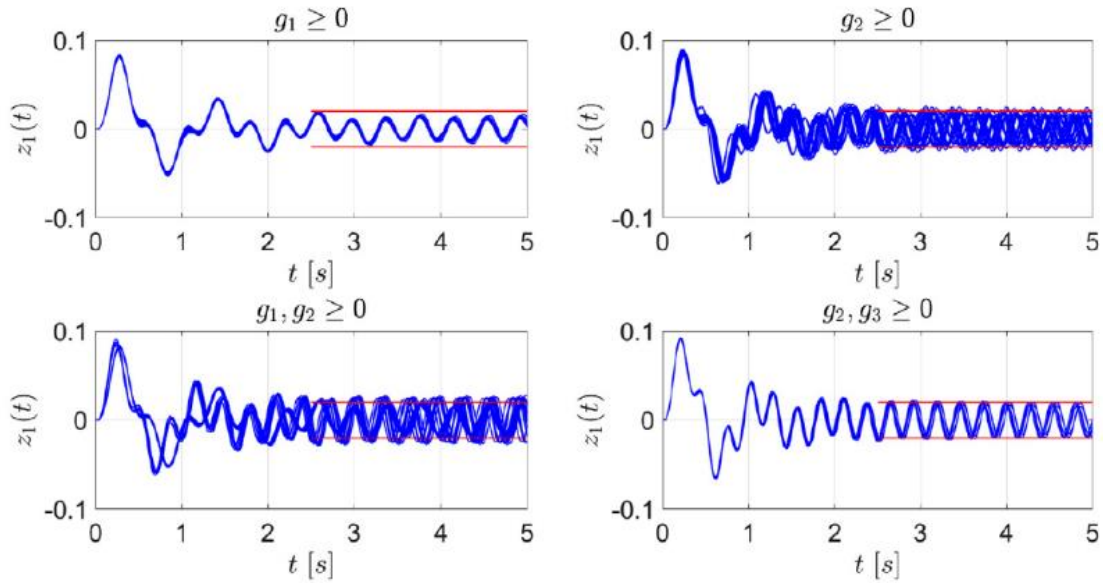
Figure 4.22 Response plot of $z_1(t)$ corresponding to 25 representative realizations of $\{a, e\}$ for each failure type (five for failure type $g_2, g_3 \geq 0$). The red lines denote the safety limits.
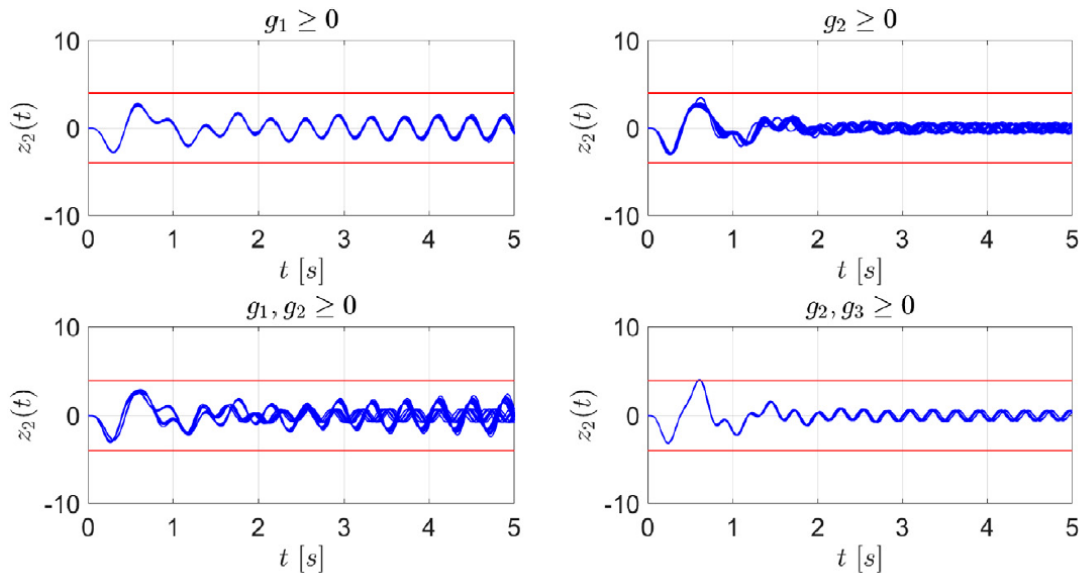


Figure 4.23 Response plot of $z_2(t)$ corresponding to 25 representative realizations of $\{a, e\}$ for each failure type (five for failure type $g_2, g_3 \geq 0$). The red lines denote the safety limits.

## 4.6.1 Model calibration

The model calibration is performed with the Black-box model function $\hat{z} = z(a, e, \theta_{new}, t)$ using the new data sequence $D_2 = \{z^l(t)\}_{l=1,\cdots,100}$. This is done following the approach that is outlined in Section 4.2.3.2, but with bounds

of the uniform priors for the respective epistemic parameters defined by the hyper-rectangle $E_1$ and $P(D|\boldsymbol{\theta}, M)$ having configurations $n_b = 5$ and $N_{sim} = 500$ to reduce the computational burdens. Let the calibrated UM be denoted $UM_{z0}$ whose resulting p-box representation of the marginal distributions of $f_a$ is shown in Figure 4.32 and whose uncertainty bounds over $e$ are presented in Table 4.10. To validate the model calibration results, the resulting model output bands of $\hat{z}$ from $UM_{z0}$ (in blue) are illustrated in Figures 4.28 and 4.30.

### 4.6.1.1 Results and discussion

In both figures, it is observed that the model output bands of $UM_{z0}$ (in blue) generally encloses the data $D_2$ (in red) which indicates that the model calibration procedure, via Bayesian model updating, was done satisfactorily. From Figure 4.30, however, some unstable behavior is observed in the response of $z_2(t)$. For investigating this, a scatterplot matrix of the aleatory samples is shown in Figure 4.24 where it is seen that the samples responsible for such unstable behavior are predominantly located into the corners of the aleatory space as seen from the following two-dimensional space: (1) $a_1$ vs $a_2$; (2) $a_1$ vs $a_3$; (3) $a_2$ vs $a_2$. There is a total of 165 of such aleatory samples and in order to reduce the number of samples from those regions, correlations need to be introduced such that a negative correlation exists between $a_1$ and $a_2$ (i.e., $C_{1,2}$), while a positive correlation exists between $a_1$ and $a_3$ (i.e., $C_{1,3}$) as well as between $a_2$ and $a_3$ (i.e., $C_{2,3}$). These correlations will be modelled using a Gaussian copula function (Nelsen, 2006).
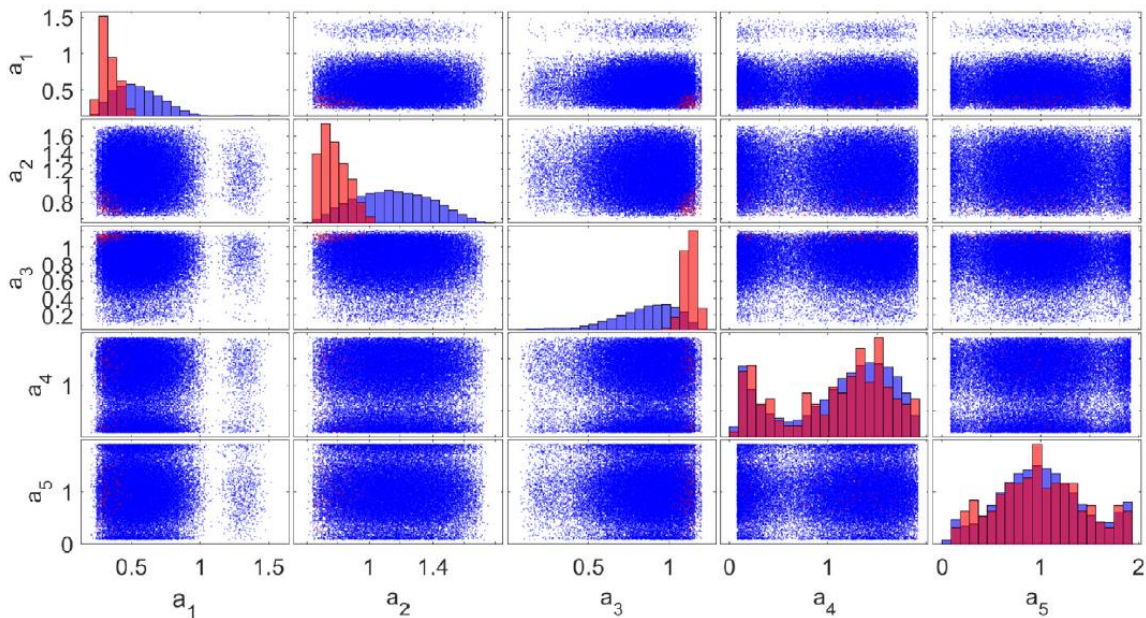


Figure 4.24 Scatterplot matrix of the aleatory samples from $UM_{z0}$ without considering correlations. The plots in blue contribute to the stable realizations while those in red contribute to the unstable realizations of $z_2(t)$.

To identify the correlation parameters $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$, a second round of Bayesian model updating is carried out on $UM_{z0}$ with the inferred parameters being the aforementioned correlation parameters. The prior distributions and bounds for $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$ are given in Table 4.9. This procedure is done keeping the p-box of $f_a$ and the uncertainty bounds of $e$ determined in the previous round of Bayesian model updating as fixed models for $a$ and $e$. The set-up of $P(D|\theta, M)$ follows that used to calibrate $UM_{z0}$ initially. The resulting histograms for $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$ are illustrated in Figure 4.25 from which the most probable value (MPV) for the respective correlation parameters are obtained and presented in Table 4.9.

Table 4.9 Summary of the prior distributions and the resulting MPVs for the respective correlation parameters.

| Parameter | $C_{1,2}$ | $C_{1,3}$ | $C_{2,3}$ |
|---|---|---|---|
| Prior distribution parameters | $U[-1, 0]$ | $U[0, 1]$ | $U[0, 1]$ |
| MPV | -0.0427 | 0.2064 | 0.0316 |
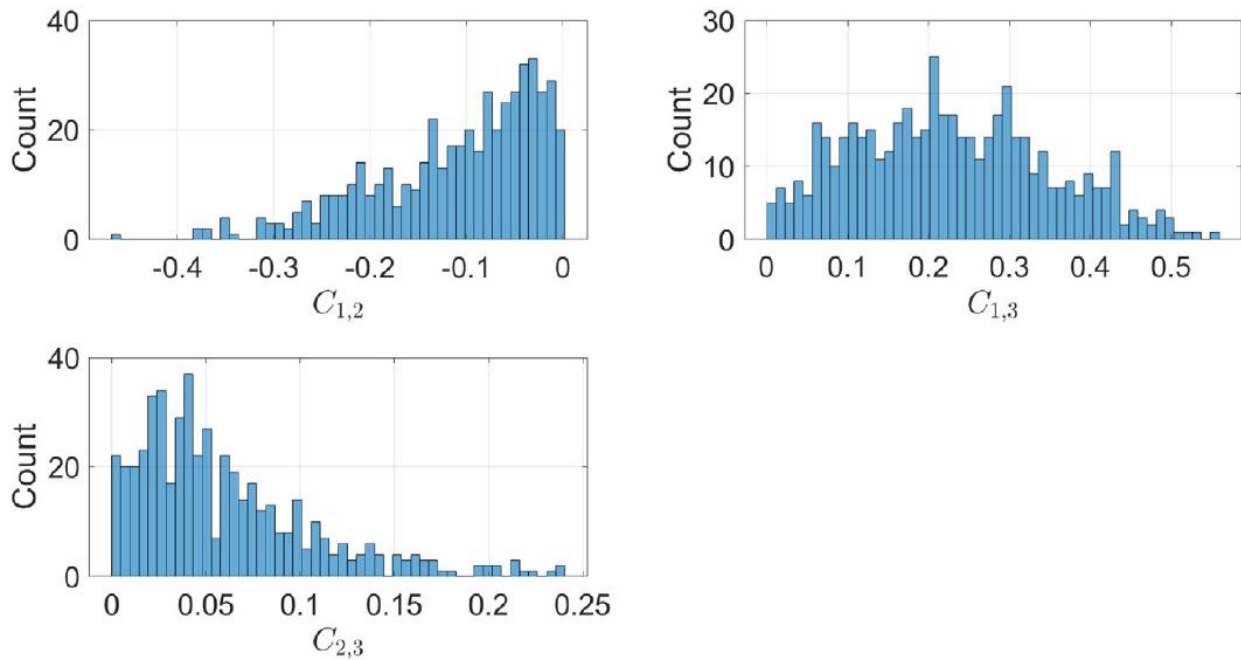


Figure 4.25 Histograms for $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$ obtained from $P(\theta|D, M)$.

To illustrate the effectiveness of introducing correlations, $UM_{z0}$ is updated to consider $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$ whose values correspond to the MPVs as shown in Table 4.9. This is done following the set-up which was utilized to calibrate $UM_{z0}$ prior to the introduction of correlations. The resulting scatterplot matrix is presented in Figure 4.26 which shows that the number of aleatory samples corresponding to the unstable realizations of $z_2(t)$ has been reduced substantially from 165 to 39.
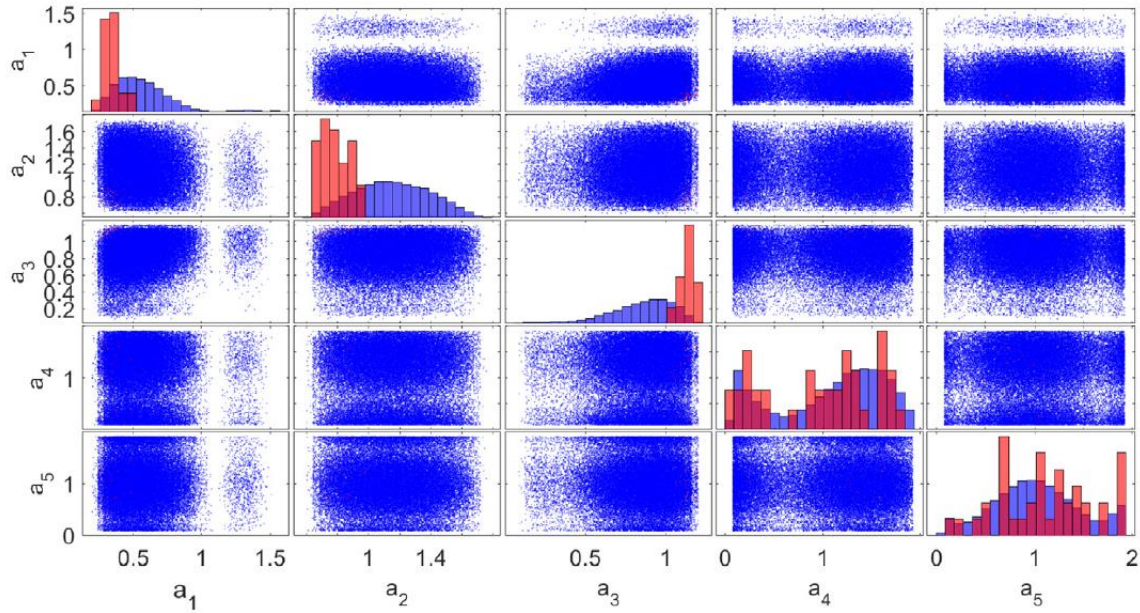
Figure 4.26 Scatterplot matrix of the aleatory samples from $UM_{z0}$ after considering correlations. The plots in blue contribute to the stable realizations while those in red contribute to the unstable realizations of $z_2(t)$.

## 4.6.2   Uncertainty reduction

From the resulting bounds of the epistemic space according to $UM_{z0}$ as shown in Table 4.10, it is observed that the uncertainty bounds of $e_4$ is the largest among the epistemic parameters. This is attributed to the lack of information provided by $e_4$ to which evidence is given by the sensitivity analysis done in Section 4.3.3 (i.e., see Table 4.5) which shows that $e_4$ is still the least sensitive parameter even after refinement. Because of substantial knowledge of its upper-bound, as indicated by the histograms of $e_4$ in Figures 4.12 and 4.33 which are truncated at the upper-bound, a request is made for the refinement of the lower-bound of $e_4$.

Table 4.10 Updated epistemic space $E$ for $e_1$ to $e_4$ according to the respective UMs.

| Updated model | $e_1$ | $e_2$ | $e_3$ | $e_4$ |
|---|---|---|---|---|
| $UM_{z0}$ | [0.5961, 0.7319] | [0.7790, 0.9337] | [0.4777, 0.5670] | [0.8521, 1.1664] |
| $UM_{z1}$ | [0.4384, 0.5795] | [0.5350, 0.5704] | [0.3353, 0.5670] | [0.9027, 0.9497] |

Based on the sensitivity analysis done in Section 4.3.3, 4.4.2, and 4.5.2, the results show that $e_2$ has consistently been ranked the most sensitive epistemic parameter. In addition to this, Figure 4.27 illustrates the parallel plots of the samples according to $UM_{z0}$ from which it can be observed that the lower-bound of the interval for $e_2$ is increased after introducing correlation with the identified values of $C_{1,2}$, $C_{1,3}$, and $C_{2,3}$ (in blue) compared to the absence of correlation (in red). In fact, the change in interval of $e_2$ is the most substantial compared to the other epistemic parameters after

correlation is introduced to the $UM_{z0}$. This can further substantiate the evidence that $e_2$ is the most sensitive epistemic parameter whose refinement could furthermore reduce the unstable realizations of $z_2(t)$. For these reasons, the final refinement request is made for the upper-bound of $e_2$ given that there is substantial knowledge in its lower-bound as observed from Figure 4.27. The resulting epistemic space, with the refined $e_2$ and $e_4$ bounds, constitutes the hyper-rectangle epistemic space denoted as $E_2$.
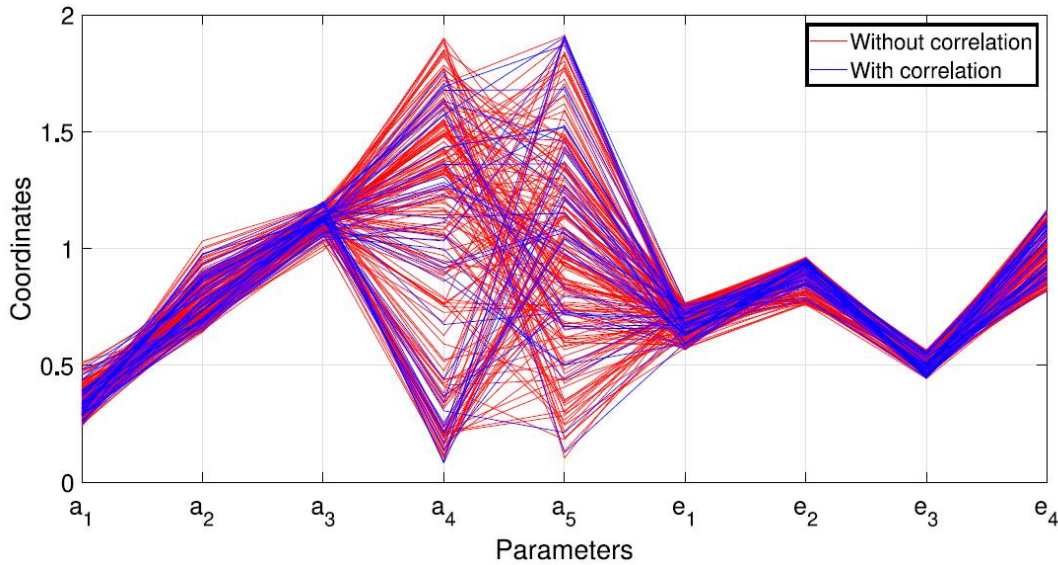


Figure 4.27 Parallel plots of the samples from $UM_{z0}$ corresponding to the unstable realizations before considering correlations (in red) and after considering correlations (in blue).

Following this, a third round of Bayesian model updating is performed with bounds of the uniform priors for the respective epistemic parameters defined by the hyper-rectangle $E_2$ and the correlation parameters taking fixed values as defined in Table 4.9. The approach follows that outlined in Section 4.6.1 from which $UM_{z0}$ was obtained, however, this time, accounting for the identified correlation parameters in the aleatory space. Let the refined UM be denoted as $UM_{z1}$ whose uncertainty bounds over $e$ are presented in Table 4.10. In addition, due to substantial information on the parameters of the SDF, their respective MPV values are utilized to define the final CDF representation of the marginal distributions of $f_a$ which are illustrated in Figure 4.32. Finally, to validate the model calibration results, the resulting model output bands of $\hat{z}$ from $UM_{z1}$ (in green) are illustrated in Figures. 4.28 and 4.30.

## 4.6.2.1  Results and discussion

In both figures, it is observed that the model output bands of $UM_{z1}$ (in green) generally encloses the data $D_2$ (in red) which indicates that the model calibration procedure was dene satisfactorily. From Figure 4.30, it can also be seen

that there is no unstable response behavior of $z_2(t)$ which indicates that the refinement procedure has removed all the

remaining unstable realizations. Figures 4.29 and 4.31 illustrate the p-boxes of the model output of $UM_{z0}$ (i.e., before

accounting for correlations) and $UM_{z1}$ at time-slices $t = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}$ $s$ for $z_1(t)$ and $z_2(t)$ respectively.

From the figures, it can be seen that the p-boxes obtained from the respective UMs generally enclose and follow the

trend of the ECDF defined by $D_2$ in general. This further highlights the effectiveness of the refinement procedure that

is undertaken in this task and validates the results obtained by $UM_{z1}$.



Figure 4.28 Output band from $\hat{z}$ according to $UM_{z0}$ without correlations (in blue) and $UM_{z1}$ (in green) along with the
data sequence $D_2$ (in red) after calibration. The black lines denote the safety limits.
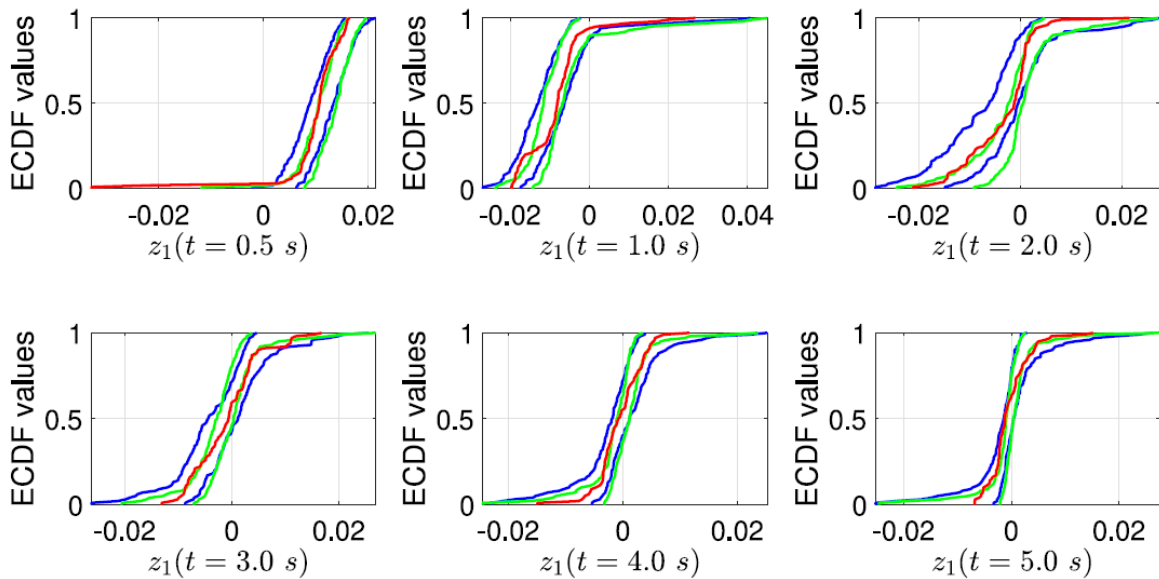


Figure 4.29 P-boxes of the model output from $\hat{z}$ obtained from $UM_{z0}$ without correlations (in blue) and $UM_{z1}$ (in
green) at various time slices $t = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}$ $s$. The red ECDF denotes the distribution of the data $D_2$.
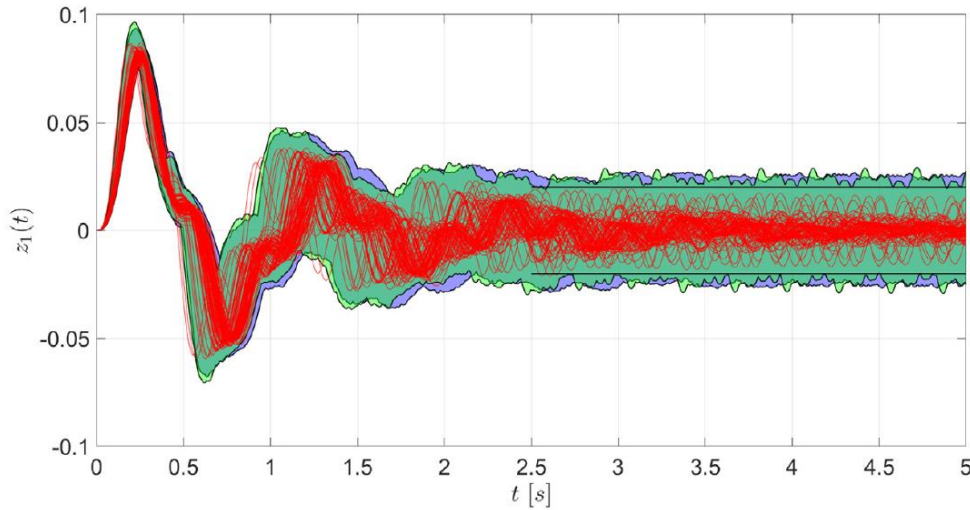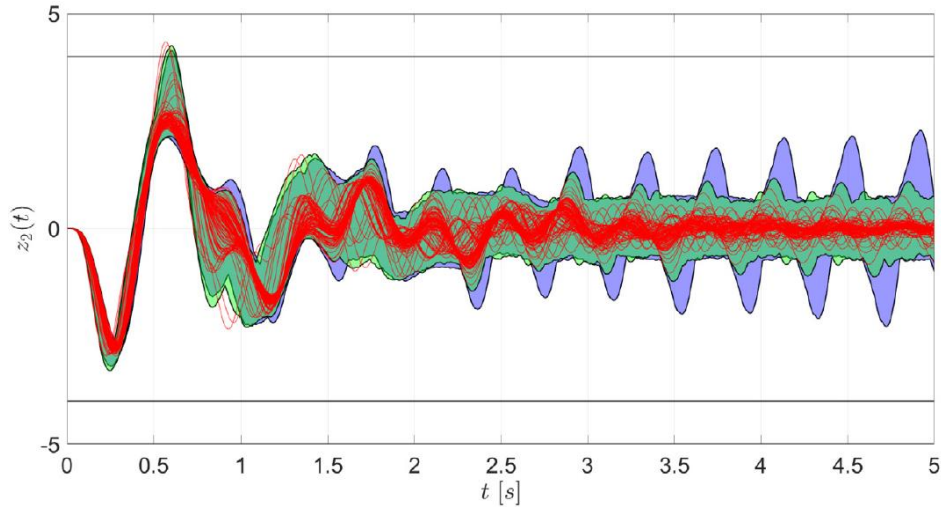
Figure 4.30 Output band from $\hat{z}$ according to $UM_{z0}$ without correlations (in blue) and $UM_{z1}$ (in green) along with the data sequence $D_2$ (in red) after calibration. The black lines denote the safety limits.
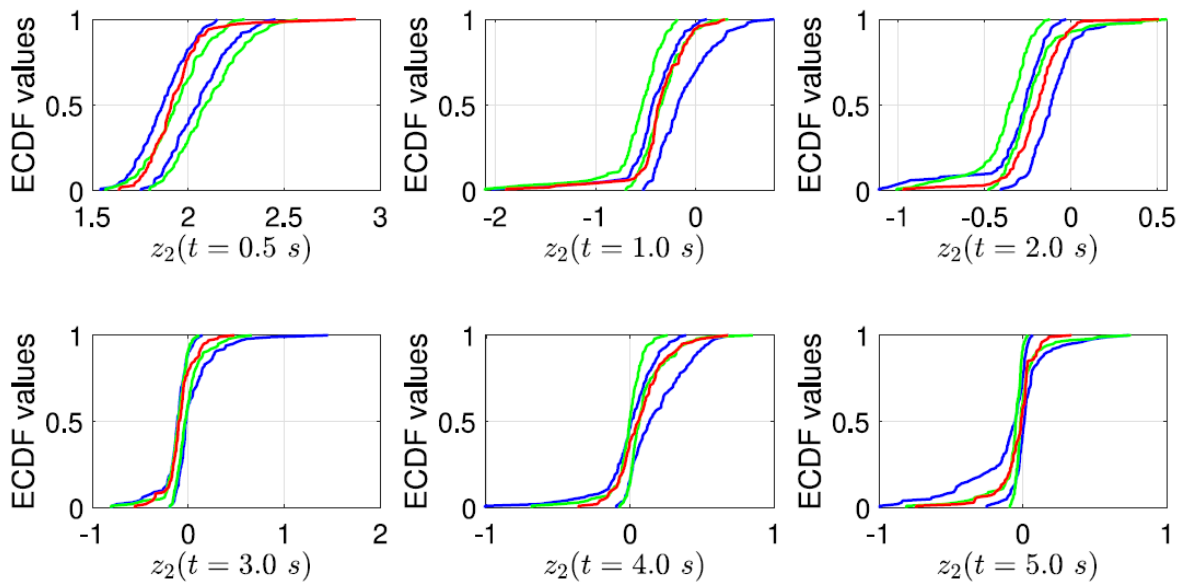


Figure 4.31 P-boxes of the model output from $\hat{z}$ obtained from $UM_{z0}$ without correlations (in blue) and $UM_{z1}$ (in green) at various time slices $t = \{0.5, 1.0, 2.0, 3.0, 4.0, 5.0\}\ s$. The red ECDF denotes the distribution of the data $D_2$.

Figure 4.32 shows the resulting p-box quantifying the uncertainty over the marginal distributions of $f_a$ by $UM_{z0}$ as well as the final CDF for $f_a$ by $UM_{z1}$. From the figure, the p-boxes (in blue) generally encloses and reveal a good degree of agreement with the final CDF (in green) which verifies the results of $UM_{z1}$. Such observation also provides good evidence that the final CDF illustrated in Figure 4.32 is a good representation of the true CDF of $f_a$.

Finally results from Table 4.10 presents that the uncertainty bounds over $e$ according to $UM_{z1}$, with the exception of $e_3$, is greatly narrower than those according to $UM_{z0}$. For the case of $e_1$ and $e_2$, their bounds according to $UM_{z1}$ are

not enclosed within that defined by $UM_{z0}$. Such observation is supported by Figure 4.33 where it can be seen that there

is little to no overlap between the histograms for $e_1$ and $e_2$ obtained before and after refinement. This is due to the fact

that the initial bounds of $e_2$ defined by $UM_{z0}$ lie entirely outside the refined bounds provided by the challenge hosts.

As such, when the calibration was performed once again with such new information, the effect of such refinement is

significant on $e_1$ which results in the epistemic parameter also having reduced bounds which largely outside the initial

bounds defined by $UM_{z0}$. This observation also suggests a significant correlation between $e_1$ and $e_2$.



Figure 4.32 P-box for $a_1$ to $a_5$ obtained from the respective UMs.
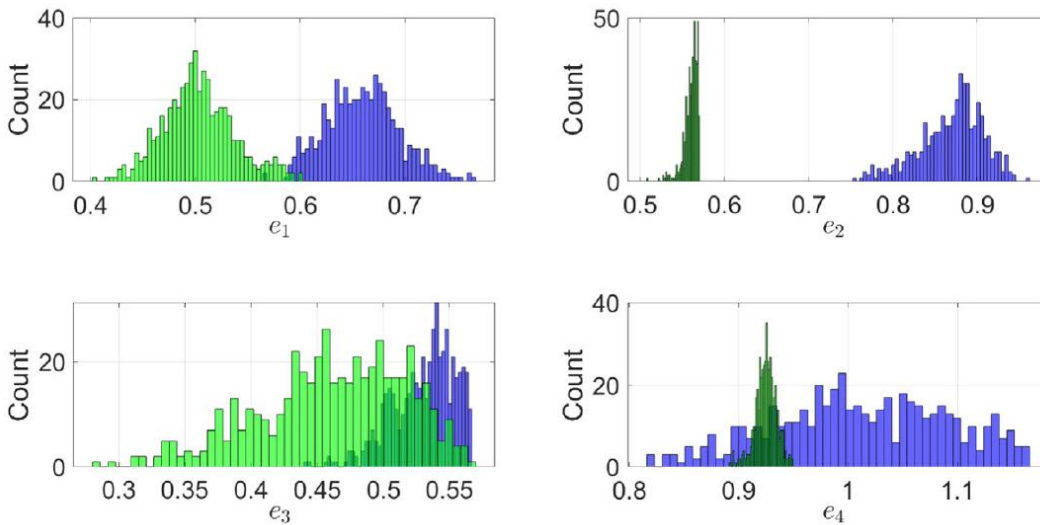


Figure 4.33 Histograms for $e_1$ to $e_4$ obtained before refinement (in blue) and after refinement (in green).

### 4.6.3 Identification and reliability analysis of final design

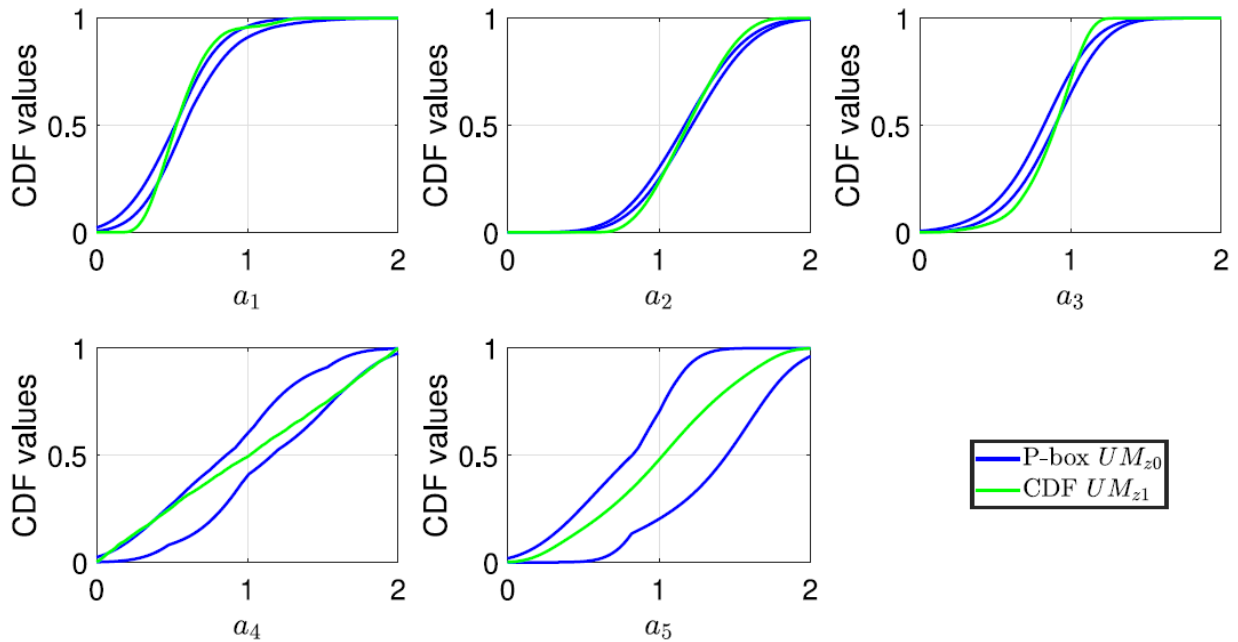The objectives of this task are: (1) to identify the final design point $\theta_{final}$; and (2) perform the reliability analysis on $\theta_{final}$ according to $UM_{z1}$.

The optimization procedure to identify $\theta_{final}$ follows the methodology shown in Section 4.5.1 with $N_b^{j=1} = 3\%$.

The reliability analysis was performed following the methodology outlined in Section 4.4.1. The results for $R_{i_g}$, $R$, and $s_{i_g}$ are presented in Table 4.14. From the table, it can be seen that the upper-bounds of the failure probabilities $R_1$ and $R_3$ have been reduced compared to $\theta_{new}$ according to $UM_{y1}^2$, and that also the severities $s_1$ to $s_3$ have all been reduced to almost zero. Although the upper-bound of $R_2$ shows a small increase from $\theta_{new}$ according to $UM_{z1}$. Such results and observations highlight the effectiveness of the optimization procedure and the identified $\theta_{final}$.

Next, the sensitivity analysis is performed on the epistemic parameters $e_1$ to $e_4$ where they are ranked according to the maximum possible reduction on $R(\theta_{final})$ bounds. The methodology follows that outlined in Section 4.4.2 and the resulting illustrative plots from the analysis is depicted in Figure 4.34. From which, the sensitivity ranking of $e_1$ to $e_4$, along with the corresponding reduced bounds of $R(\theta_{final})$, are presented in Table 4.11. From the results, it can be seen that $e_1$ and $e_2$ have two of the lowest sensitivities which implies that the bounds obtained for these two parameters are already sufficiently narrow such that no further information can be obtained on them by reducing their bounds any further.



Figure 4.34 Results of the reduced $R(\theta_{final})$ bounds for different pinched intervals for $e_1$ to $e_4$. The red lines represent the initial bounds of $[0.0028, 0.0144]$, while the green bars represent the reduced bounds.

Table 4.11 The ranking order of the epistemic model parameters based on the maximum possible reduction in $R(\theta_{final})$ interval according to $UM_{z1}$.

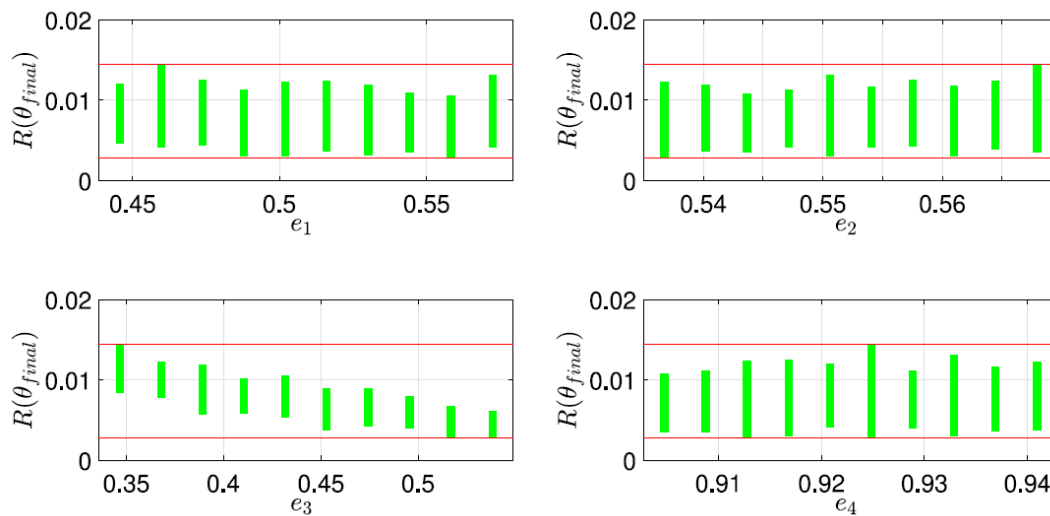| Rank | Parameter | Pinched bounds | $R(\theta_{base})$ Before pinching | After pinching |
|------|-----------|----------------|-----------------------------------|----------------|
| 1 | $e_3$ | $[0.5274, 0.5487]$ | $[0.0028, 0.0144]$ | $[0.0029, 0.0061]$ |
| 2 | $e_4$ | $[0.9269, 0.9309]$ | $[0.0028, 0.0144]$ | $[0.0040, 0.0111]$ |
| 3 | $e_2$ | $[0.5419, 0.5454]$ | $[0.0028, 0.0144]$ | $[0.0036, 0.0108]$ |
| 4 | $e_1$ | $[0.5372, 0.5512]$ | $[0.0028, 0.0144]$ | $[0.0036, 0.0109]$ |

Finally, the failure analysis is done on $\theta_{final}$ following the methodology outlined in Section 4.4.3. The resulting statistics summarizing the number of realizations of $\{a, e\}$ in each failure category in presented in Table 4.12 where it can be seen that likelihood of occurrences for all failure types have been reduced significantly from $\theta_{new}$ according to $UM_{y1}^2$, except for failure types $g_3 \geq 0$ and $g_2, g_3 \geq 0$ where there is an increase in the number of sample realizations in these failure domains. Failure type $g_2 \geq 0$ still has the highest likelihood as it has the highest number of realizations among the different failure types as per $\theta_{base}$ and $\theta_{new}$. The sample sets $\{a, e\}_{nf}^{g_1}$, $\{a, e\}_{nf}^{g_2}$, and $\{a, e\}_{nf}^{g_3}$ are identified, and the parallel plots are shown in Figure 4.35 while the response plots of $z_1(t)$ and $z_2(t)$ are shown in Figure 4.36.

Table 4.12 Statistics of the different failures based on analysis for different design points and UMs.

| Failure type | $\theta_{base}$ $UM_{y1}^2$ | $UM_{z1}$ | $\theta_{new}$ $UM_{y1}^2$ | $UM_{z1}$ | $\theta_{final}$ $UM_{y1}^2$ | $UM_{z1}$ |
|--------------|------------|-----------|------------|-----------|------------|-----------|
| No failure | 4357960 | 4735066 | 4696943 | 4955517 | 4713788 | 4963836 |
| $g_1 \geq 0$ | 54827 | 4002 | 3529 | 22 | 2485 | 36 |
| $g_2 \geq 0$ | 490718 | 75237 | 290863 | 30231 | 275785 | 34004 |
| $g_3 \geq 0$ | 16239 | 126310 | 0 | 7391 | 0 | 657 |
| $g_1, g_2 \geq 0$ | 45802 | 1888 | 8660 | 187 | 7940 | 81 |
| $g_1, g_3 \geq 0$ | 0 | 0 | 0 | 0 | 0 | 0 |
| $g_2, g_3 \geq 0$ | 34256 | 57471 | 5 | 6652 | 0 | 1386 |
| $g_1, g_2, g_3 \geq 0$ | 198 | 26 | 0 | 0 | 2 | 0 |
| Total samples | | | $5 \times 10^6$ | | | |

From Figure 4.36, the response plots of $z_1(t)$ and $z_2(t)$ corresponding to $\{a, e\}_{nf}^{g_1}$ are all within the safety limits. This indicates that the sample set $\{a, e\}_{nf}^{g_1}$ lie within the safe domain near the boundary of $g_1 \geq 0$. For $\{a, e\}_{nf}^{g_2}$, the response plots of $z_1(t)$ and $z_2(t)$ are all within the safety limits as well which indicates that the sample set $\{a, e\}_{nf}^{g_2}$ lie within the safe domain near the boundary of $g_2 \geq 0$. For $\{a, e\}_{nf}^{g_3}$, the response plots of $z_1(t)$ exceeds the upper-bounds of the safe boundary at approximately $t = \{2.5, 2.6, 2.9, 3.0, 3.2, 3.3, 3.4\}$ s which indicates the presence of failure of $g_2 \geq 0$. The plots for $z_2(t)$, on the other hand, are well-within the safety limits. Given the stable behavior of the plots, it can be inferred that the requirement of $g_1$ is satisfied that indicates that the sample set $\{a, e\}_{nf}^{g_3}$ lie within the domain of $g_2 \geq 0$ near the boundary of $g_2, g_3 \geq 0$.
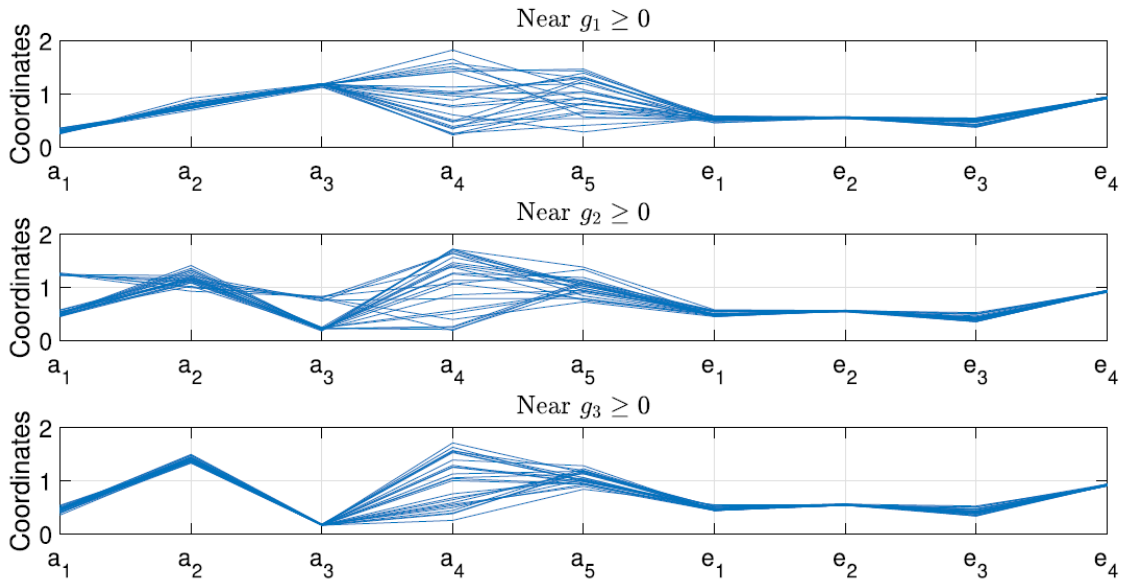
Figure 4.35 Parallel plots of $\{a,e\}_{nf}^{g_1}$, $\{a,e\}_{nf}^{g_2}$, and $\{a,e\}_{nf}^{g_3}$ from $UM_{z1}$
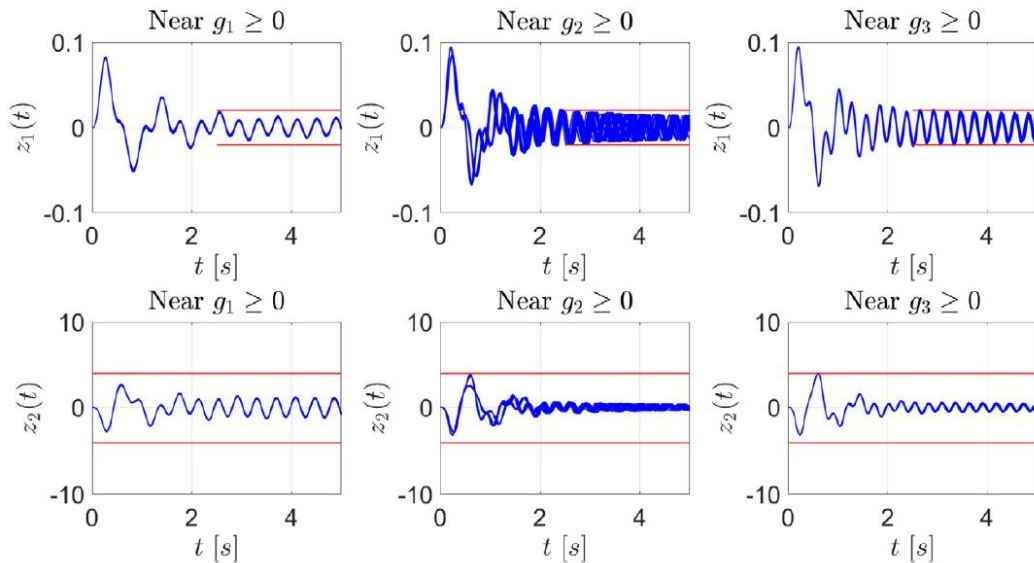


Figure 4.36 Response plot of $z_1(t)$ and $z_2(t)$ of $\{a,e\}_{nf}^{g_1}$, $\{a,e\}_{nf}^{g_2}$, and $\{a,e\}_{nf}^{g_3}$ from $UM_{z1}$ near the respective failure domains. The red lines denote the safety limits

Figures 4.37 and 4.38 present the response plots of $z_1(t)$ and $z_2(t)$ respectively for the 25 representative sample sets of $\{a,e\}$ for each failure type with non-zero likelihood. From Figure 4.38, it can be seen that the unstable behavior in $z_2(t)$ that was previously presented for failure type $g_2, g_3 \geq 0$ are no longer presented as a result of the refinement procedure that was done in Section 4.6.2. In addition, Figure 4.37 shows that the $z_1(t)$ response plots for failure type $g_2, g_3 \geq 0$ still exceeds the safety limits to the largest extent compared to the other failure types.
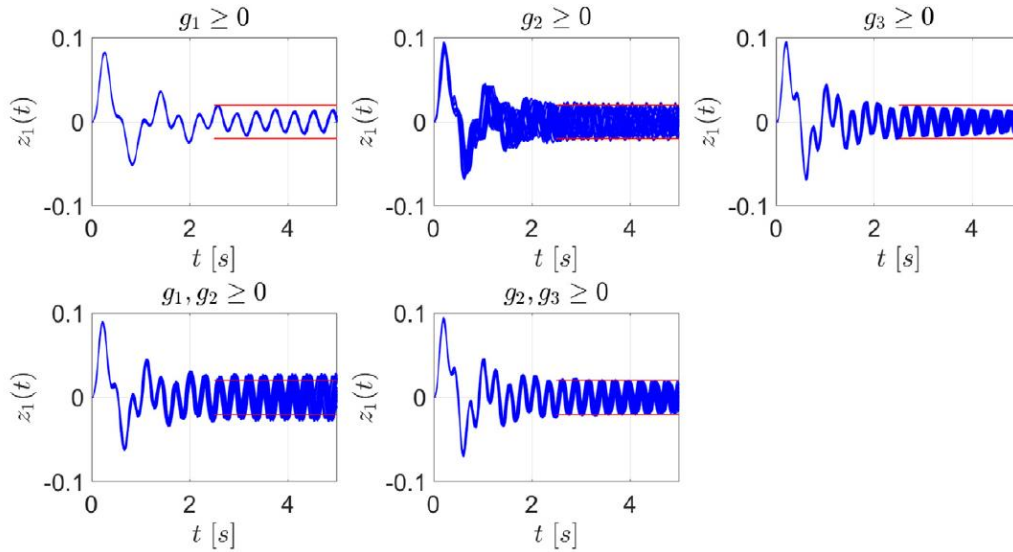
Figure 4.37 Response plot of $z_1(t)$ corresponding to 25 representative realizations of $\{a, e\}$ for each failure type. The red lines denote the safety limits
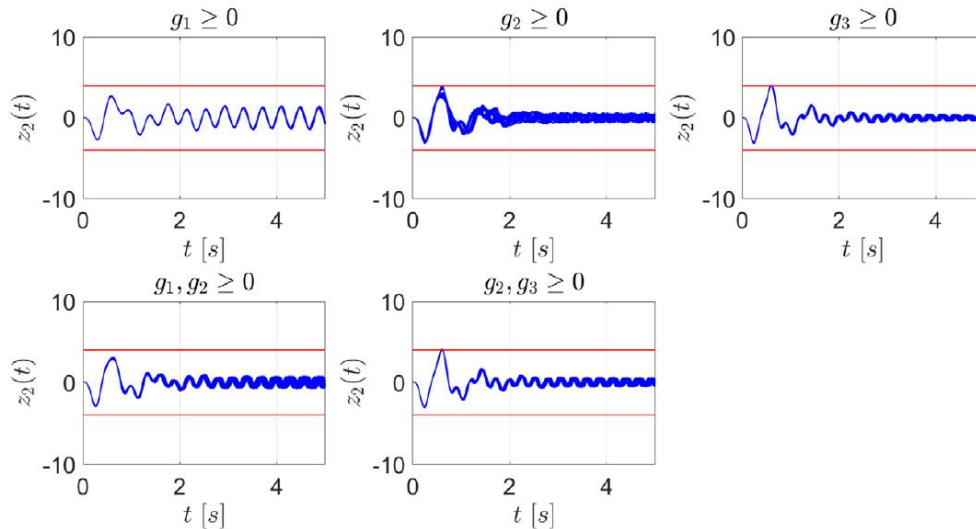


Figure 4.38 Response plot of $z_2(t)$ corresponding to 25 representative realizations of $\{a, e\}$ for each failure type. The red lines denote the safety limits

### 4.6.4 Comparison of design points

The objective of this task is to present a quantitative comparison between the three design points $\theta_{base}$, $\theta_{new}$, and $\theta_{final}$ on the basis of the reliability metrics $R_{i_g}$, $R$, and $s_{i_g}$.

In Table 4.14, it can be observed that the reliability analysis results for the different $\theta$ suggests $\theta_{final}$ is the most

optimal design point. This is due to $\theta_{final}$ having the lowest upper-bound failure probabilities and severities in general compared to $\theta_{base}$ and $\theta_{new}$ based on the reliability results according to $UM_{y1}^2$ and $UM_{z1}$. Such results are supported by Figures. 4.39 and 4.40 which illustrate the resulting p-boxes for $g_1$ to $g_3$ according to the analysis for the different $\theta$ by $UM_{y1}^2$ and $UM_{z1}$ respectively. In both figures, it can be seen that the p-boxes for $\theta_{final}$ are so that they are mainly lie within the safe domain of the respective requirement and that the tails of the p-boxes do not extend as far into the failure regions compared to $\theta_{base}$ and $\theta_{new}$.
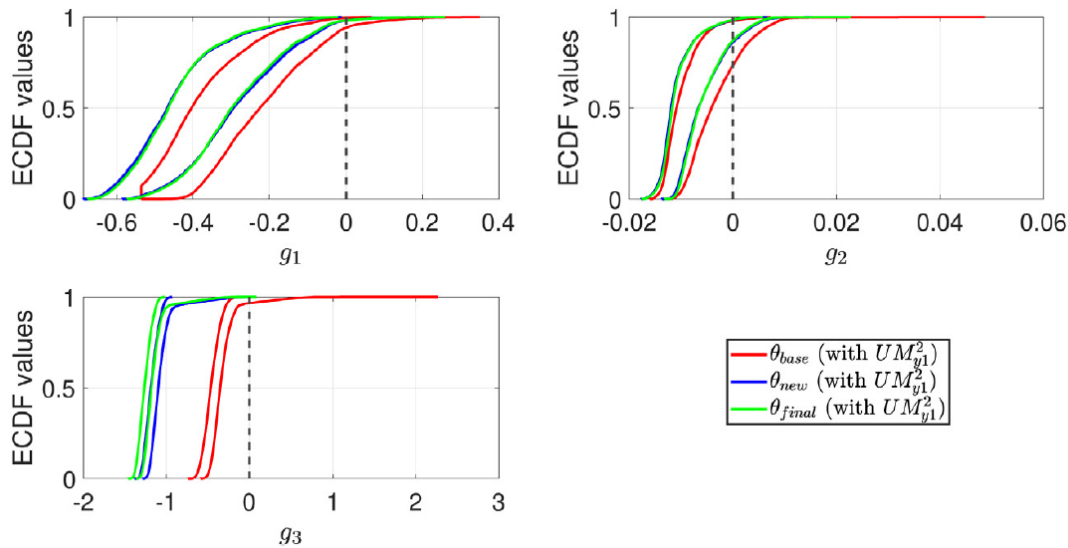
Figure 4.39 P-boxes obtained for $g_1$, $g_2$, and $g_3$ for different $\theta$ according to $UM_{y1}^2$
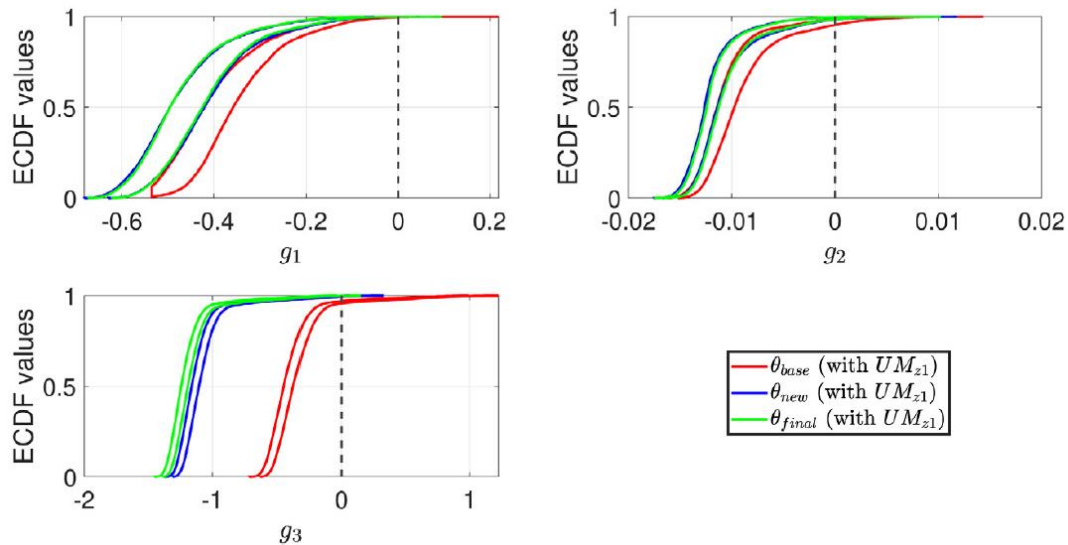
Figure 4.40 P-boxes obtained for $g_1$, $g_2$, and $g_3$ for different $\theta$ according to $UM_{z1}$

A further analysis was also done in order to compare the number of realizations of $\{a, e\}$ for the different failure

types according to $UM_{y1}^2$ and $UM_{z1}$ and the resulting statistics are provided in Table 4.12. For both $UM_{y1}^2$ and $UM_{z1}$, the number of safe (i.e., no failure) realizations is the highest for $\theta_{final}$ which further substantiates it being the most optimal design point for the system. While this is achieved, a drawback of $\theta_{final}$ is the increase in realizations of failure types $g_1 \geq 0$ and $g_2 \geq 0$ from $\theta_{new}$ according to $UM_{z1}$ while for the case of $UM_{y1}^2$, such drawback comes in the form of a slight increase in realizations of failure type $g_1, g_2, g_3 \geq 0$ from $\theta_{new}$.

### 4.6.5 Numerical implementation and computational time

In addressing the tasks presented in this challenge, the adopted algorithms are mainly based on random sampling and stochastic algorithms. Thus, the execution time fluctuates significantly due to the inherent randomness depending on the uncontrolled conditions such as starting samples, the evolution of the sample, etc. However, it needs to be noted that the random seed has not been fixed to allow for generality of the implementation and solution. Therefore, we only provide the approximate timing as the performance indicator: (1) the TMCMC which takes between five to eight hours of sampling time; (2) the adaptive pinching approach which involves less than a minute of computation time; (3) the double-loop Monte Carlo simulation which involves between 1.5 to 2 hours of simulation time; and (4) the NISS which takes only three minutes of simulation time.

It needs to be highlighted that the above computational times are also dependent on the computational efficiency of the high-performance CPUs. These estimated timings may differ between different CPUs of different specifications.

### 4.7 Conclusions

Different techniques have been presented for solving the NASA UQ challenge problem. Bayesian model updating technique has been used to calibrate the uncertainty model by performing a stochastic update on both the distribution parameters as well as the epistemic parameters. Two different uncertainty models have been analyzed, each adopting a different selection of joint distribution function for the aleatory space: (1) beta distribution; and (2) staircase density functions. The use of the staircase density function provided more informative results of the distribution parameters and the epistemic parameters from their respective posteriors and used in the subsequent problems.

An adaptive pinching analysis based on Tucker and Ferson (2006) was utilized to perform the sensitivity analysis on the epistemic parameters, providing an efficient way of identifying the largest possible reduction of the proposed metric by the single pinched component of the epistemic space. In doing so, it allows for a systematic, non-empirical

way to justify the pinched bounds and ensure that all regions of the individual components of the epistemic space are taken into account in the investigation of the respective effect of each epistemic parameter on the two aforementioned quantities.

To perform the reliability analysis and compute the reliability metrics, probability-boxes are constructed through the double-loop Monte Carlo approach. From which, the probability bounds analysis was performed on the resulting probability-boxes to evaluate necessary bounds on the respective failure probabilities as well as the worst-case failure probabilities. The approach does not assume a fixed distribution on each performance and considers only the extreme bounds on the probability values obtained.

Finally, to identify an optimal design point of the system, the non-intrusive imprecise stochastic simulation (Song et al, 2019) was utilized. The approach provides a systematic way to explore a defined hyper-rectangular space of the design point and identify the values of each design point parameter corresponding to the local minimum of the first-order component function. This can allow for the optimized design point to be identified in accordance to the defined criteria. A summary of key results is provided in Tables 4.13-4.15.

Table 4.13 Results to the epistemic spaces defined by the respective UMs

| Updated model | $e_1$ | $e_2$ | $e_3$ | $e_4$ |
|---|---|---|---|---|
| $UM_{y0}^1$ | [0.4149, 1.5271] | [0.2407, 1.5673] | [0.1506, 1.8001] | [0.1601, 1.9236] |
| $UM_{y0}^2$ | [0.4351, 0.7082] | [0.5583, 1.0000] | [0.0721, 0.5511] | [0.6066, 1.6893] |
| $UM_{y1}^1$ | [0.3730, 1.3457] | [0.1869, 1.1529] | [0.2997, 0.5570] | [0.8295, 1.1664] |
| $UM_{y1}^2$ | [0.4674, 0.6433] | [0.7607, 0.9736] | [0.2865, 0.4583] | [0.9627, 1.1664] |
| $UM_{z0}$ | [0.5961, 0.7319] | [0.7790, 0.9337] | [0.4777, 0.5670] | [0.8521, 1.1664] |
| $UM_{z1}$ | [0.4384, 0.5795] | [0.5350, 0.5704] | [0.3353, 0.5670] | [0.9027, 0.9497] |

Table 4.14 Reliability analysis results for the different design points $\theta$ with $UM_{y1}^2$ and $UM_{z1}$

| Design point | $R_1(\theta)$ | $R_2(\theta)$ | $R_3(\theta)$ | $R(\theta)$ |
|---|---|---|---|---|
| $\theta_{base}$ (with $UM_{y1}^1$) | [0.0028, 0.0580] | [0.0190, 0.2644] | [0.0000, 0.0343] | [0.0270, 0.2746] |
| $\theta_{base}$ (with $UM_{z1}$) | [0.0003, 0.0023] | [0.0137, 0.0445] | [0.0291, 0.0436] | [0.0386, 0.0699] |
| $\theta_{new}$ (with $UM_{y1}^1$) | [0.0000, 0.0176] | [0.0150, 0.1378] | [0.0000, 0.0001] | [0.0153, 0.1379] |
| $\theta_{new}$ (with $UM_{z1}$) | [0.0000, 0.0004] | [0.0025, 0.0141] | [0.0011, 0.0059] | [0.0040, 0.0154] |
| $\theta_{final}$ (with $UM_{y1}^1$) | [0.0000, 0.0171] | [0.0137, 0.1305] | [0.0000, 0.0001] | [0.0142, 0.1306] |
| $\theta_{final}$ (with $UM_{z1}$) | [0.0000, 0.0002] | [0.0024, 0.0143] | [0.0000, 0.0013] | [0.0028, 0.0144] |

| Design point | $s_1(\theta)$ | $s_2(\theta)$ | $s_3(\theta)$ | $s(\theta)$ |
|---|---|---|---|---|
| $\theta_{base}$ (with $UM_{y1}^1$) | 0.0413 | 0.1981 | 0.0411 | 0.3779 |
| $\theta_{base}$ (with $UM_{z1}$) | $2 \times 10^{-6}$ | 0.0007 | 0.1082 | 0.1669 |
| $\theta_{new}$ (with $UM_{y1}^1$) | 0.0009 | 0.0253 | $1 \times 10^{-9}$ | 0.0277 |
| $\theta_{new}$ (with $UM_{z1}$) | $3 \times 10^{-9}$ | $2 \times 10^{-5}$ | 0.0001 | 0.0001 |
| $\theta_{final}$ (with $UM_{y1}^1$) | 0.0010 | 0.0202 | $1 \times 10^{-9}$ | 0.0220 |
| $\theta_{final}$ (with $UM_{z1}$) | $2 \times 10^{-9}$ | $1 \times 10^{-5}$ | $4 \times 10^{-7}$ | $2 \times 10^{-5}$ |

Table 4.15 Summary of the type of refinements requested for the respective refinement round. The superscripts "$+$" and "$-$" denote the refinement of the upper- and lower-bound of the epistemic parameter respectively

| Refinement round | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Refinement type | $\{e_4^-\}$ | $\{e_3^-\}$ | $\{e_4^-\}$ | $\{e_2^+\}$ |

## CRediT authorship contribution statement

**Adolphus Lye**: Methodology, Investigation, Software, Writing – original draft. **Masaru Kitahara**: Methodology, Investigation, Software, Writing – original draft. **Matteo Broggi**: Supervision, Conceptualization, Writing – review & editing. **Edoardo Patelli**: Supervision, Conceptualization, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

## References

Abdessalem, A. B., Dervilis, N., Wagg, D., and Worden, K. (2018). Model selection and parameter estimation in structural dynamics usiing approximate Bayesian computation. *Mechanical System and Signal Processing*, 99, 306-325.

Ali, T. (2012). Modeling uncertainty in risk assessment using double Monte Carlo method. *International Journal of Engineering and Innovative Technology*, 1, 114-118.

Bai, Y., Huang, Z., and Lam, H. (2020). A distributionally robust optimization approach to the NASA Langley uncertainty quantification challenge. In: *30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*.Venice, Italy.

Beck, J. L., and Au, S. K. (2002). Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *Journal of Engineering Mechanics*, 128(4), 380-391.

Beck, J. L., and Katafygiotis, L. S. (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechnics*, 124, 455-461.

Beer, M., Ferson, S., and Kreinovich, V. (2013). Imprecise probabilities in engineering analysis. *Mechanical System and Signal Processing*, 37, 4-29.

Beer, M., Kougioumtzoglou, I. A., and Patelli, E. (2014). Emerging concepts and approaches for efficient and realistic uncertainty quantification. *Maintenance and Safety of Aging Infrastructure*, 121-162.

Bi, S., Broggi, M., and Beer, M. (2019). The role of the Bhattacharyya distance in stochastic model updating. *Mechanical System and Signal Processing*, 117, 437-452.

Bowman, A. W., and Azzalini, A. (1997). *Applied smoothing techniques for data analysis*, OUP Oxford.

Ching, J., and Chen, Y. C. (2007). Transitional Markov chain Monte Carlo method for Bayesian updating, model class selection, and model averaging. *Journal of Engineering and Mechanics*, 133(7), 816-832.

Crespo, L. G., Kenny, S. P., and Giesy, D. P. (2014). The NASA Langley multidisciplinary uncertainty quantification challenge. In: *16th AIAA Non-Deterministic Approaches Conference*, National Harbor, Maryland.

Crespo, L. G., Kenny, S. P., Giesy, D. P., and Stanford, B. K. (2018). Random variables with moment-matching staircase density function. *Applied Mathematical Modelling*, 64, 196-213.

Crespo, L. G., and Kenny, S. P. (2021). The NASA Langley Challenge on Optimization Under Uncertainty. *Mechanical System and Signal Processing*, 152, 107405.

Der Kiureghian, A., and Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, 31, 105-112.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., and Sentz, K. (2003). Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015, Sandia National Laboratories, California.

Ferson, S., and Tucker, W. T. (2008). Sensitivity analysis using probability bounding. *Reliability Engineering & System Safety*, 91, 10-11.

Ferson, S., Oberkampf, W., and Ginzburg, L. (2008). Model validation and preictive capability for the thermal challenge problem. *Computer Methods in Applied Mechanics and Engineering*, 197, 2408-2430.

Gray A., Wimbush, A., de Angelis, M., Hristov, P. O., Miralles-Dolz, E., Calleja, D., and Rocchetta, R. (2020). Bayesian calibration and probability bounds analysis solution to the NASA 2020 UQ challenge on optimization under uncertainty. In: *30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*.Venice, Italy.

Heideman, M., Johnson, D., and Burrus, C. (1984). Gauss and the history of the fast fourier transform. *IEEE ASSP*

*Magazine*, 1, 14-21.

Katafygiotis, L. S., and Beck, J. L. (1998). Updating models and their uncertainties. II: Model identifiability. *Journal of Engineering Mechnics*, 124, 463-467.

Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2021). Bayesian model updating in time domain with metamodel-based reliability method. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7, 04021030.

Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2022). Nonparametric Bayesian stochastic model updating with hybrid uncertainties. *Mechanical System and Signal Processing*, 163, 108165.

Lye, A., Cicirello, A., and Patelli, E. (2021). Sampling methods for solving Bayesian model updating problems: A tutorial. *Mechanical System and Signal Processing*, 159, 107760.

Meng, Z., Pang, Y., Pu, Y., and Wang. X. (2020). New hybrid reliability-based topology optimization method combining fuzzy and probabilistic models for handling epistemic and aleatory uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 363, 112886.

Nelsen, R. B. (2006). *An introduction to copulas*, Second edition, Springer-Verlag, New York.

Oberkampf, W., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S. (2004). Challenge problems: uncertainty in system response given unertain parameters. *Reliability Engineering & System Safety*, 85, 11-19.

Panaretos, V. M., and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annual Review of Statistic and its Application*, 6, 405-431.

Patelli, E., Alvarez, D. A., Broggi, M., and de Angelis, M. (2015). Uncertanty management in multidisciplinary design of critical safety systems. *Journal of Aerospace Information Systems*, 12, 140-169.

Rezaie, K., Amalnik, M., Gereie, A., Ostadi, B., and Shakhseniaee, M. (2007). Using extended Monte Carlo simulation method for the improvement of risk management: Consideration of relationships between uncertainties. *Applied Mathematics and Computation*, 190, 1492-1501.

Rocchetta, R., Broggi, M., and Patelli, E. (2018). Do we have enough data? Robust reliability via uncertainty quantification. *Applied Mathematical Modelling*, 54, 710-721.

Roy, C., and Oberkampf, W. (2011). A comprehensive framework for verification, validation, and uncertainty quantification. *Computer Methods in Applied Mechanics and Engineering*, 200, 2131-2144.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008).

*Global sensitivity analysis*, The Primer, Wiley.

Sobol' I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55, 271-280.

Song, J., Wei, P., Valdebenito, M., Bi, S., Broggi, M., Beer, M., and Lei, Z. (2019). Generalization of non-intrusive imprecise stochastic simulation for mixed uncertain variables. *Mechanical System and Signal Processing*, 134, 106316.

Tucker, W. T., and Ferson, S. (2006). Sensitivity in risk analysis with uncertain numbers. Technical Report SAND2006-2801, Sandia National Laboratories, California.

Turner, B. M., and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2), 69-85.

Wei, P., Lu, Z., and Song, J. (2014). Extended Monte Carlo simulation for parametric global sensitivity analysis and optimization. *AIAA Journal*, 52, 867-878.

# Chapter 5

# Research article 4: Distribution-free stochastic model updating of dynamic systems with parameter dependencies

As has been shown in the last chapter, the distribution-free Bayesian model updating framework combining the staircase density functions for characterizing aleatory parameters with the Bhattacharyya distance for quantifying the uncertainty characteristics of the observations is a robust tool for uncertainty calibration which is effective even on the very challenging condition where the prior knowledge about aleatory parameters is extremely limited. Nevertheless, it is, in its current form, only applicable to the case where aleatory parameters are independent with each other, since the staircase density functions are defined for univariate random variables and cannot account for the dependence structure among the parameters.

To overcome this limitation, the present article strengthens the above distribution-free stochastic model updating framework for the calibration of the joint probabilistic distribution of multivariate parameters by representing it as a Gaussian copula function. The Gaussian copula is well-known to be capable of providing an effective way to describe the dependence structure among multivariate parameters using the correlation matrix, and this property is in particular attractive for the stochastic updating problem where very large number of parameters needs to be considered as random variables. Moreover, for maintaining the flexibility in the description of aleatory parameters, the marginal distributions of the Gaussian copula function are modeled by the staircase density functions. To this end, both the hyper-parameters of the staircase density functions and the correlation coefficients of the correlation matrix are taken into account as the parameters to be inferred, and their prior distribution is defined by the moment constraints and correlation coefficient constraint. These parameters are then updated to the posterior by the ABC updating procedure with the Bhattacharyya distance.

The proposed procedure is first demonstrated on the well-known simple engineering example utilized in the third chapter, where another pair of the natural frequencies are employed as the observed features to result in the correlated target joint distribution of the model parameters. This example clearly demonstrates how the stochastic model updating fails by ignoring the parameter dependency while the proposed procedure is capable of calibrating the correlated joint

distribution precisely. The proposed procedure is also demonstrated upon another simple engineering example with the purpose of calibrating the multi-variate model parameters that show all of the no, negative, and positive correlations. Finally, it is applied to the seismic-isolated bridge pier model updating problem using simulated seismic response data to demonstrate its capability to recreate very complicated nonlinear structure of the observed time signals.

# Distribution-free stochastic model updating of dynamic systems with parameter dependencies

Masaru Kitahara [a, *], Sifeng Bi [b], Matteo Broggi [a], Michael Beer [a, c, d]

[a] *Leibniz Universität Hannover, Institute for Risk and Reliability, Callinstrasse 34, Hannover, Germany*

[b] *University of Strathclyde, School of Mechanical & Aerospace Engineering, Aerospace Centre of Excellence, Glasgow, United Kingdom*

[c] *University of Liverpool, Institute for Risk and Uncertainty, Peach Street L69 7ZF, Liverpool, United Kingdom*

[d] *Tongji University, International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Shanghai 200092, China*

* Correspondence author. E-mail address: masaru.kitahara@irz.uni-hannover.de (M. Kitahara).

**Abstract:** This work proposes a distribution-free stochastic model updating framework for the calibration of the joint probabilistic distribution of the multivariate correlated parameters. In this framework, the marginal distributions are defined as the staircase density functions and the correlation structure is represented by the Gaussian copula function. The first four moments of the staircase density functions and the correlation coefficients of the correlation matrix are calibrated by an approximate Bayesian computation, in which the Bhattacharyya distance-based metric is proposed to define an approximate likelihood that is capable of quantifying the stochastic discrepancy between model outputs and observations. The feasibility of the framework is demonstrated on two illustrative examples and a followed engineering application to the updating of a nonlinear dynamic system using observed time histories. The results demonstrate the capability of the proposed updating procedure in the very challenging condition where the prior knowledge about the distribution of the parameters is extremely limited (i.e., no information on the marginal distribution families as well as the correlation structure is available).

**Keywords:** Uncertainty quantification; Bayesian model updating; Staircase density function; Gaussian copula function; Bhattacharyya distance.

## 5.1   Introduction

The model updating has been developed as a fascinating technique to mitigate the discrepancy between model outputs and experimental measurements (Mottershead et al., 2011; Patelli et al., 2017). The causes of the discrepancy during the model updating can be generally classified into following three categories:

- Parameter uncertainty. Model parameters, such as the geometry dimensions, boundary conditions, and material

properties, often cannot be exactly determined;

- Modelling uncertainty. Simplifications or approximations, such as the linearization and frictionless mechanical joints, have to be made to numerically represent the physical system;

- Measurement uncertainty. Measured quantities are always contaminated by the hard-to-control randomnesses, e.g., environmental noises and measurement system errors.

The deterministic model updating, especially for the sensitivity method (Mottershead et al., 2011), might be one of the most successfully applied model updating techniques. It aims at calibrating the model parameters to obtain their optimal values from a single set of measurements. It has been employed in a wide range of practical applications e.g., the calibration of large-scale finite element (FE) models (Mottershead and Friswell, 1993; Shan et al., 2015). However, it considers measurement data as an exactly determined values/signals, ignoring the measurement uncertainty.

Comparatively, the stochastic model updating, e.g., the perturbation method (Mares et al., 2006; Khodaparast et al., 2008), Monte Carlo method (Sairajan and Aglietti, 2012; Bi et al., 2017), and Bayesian method (Goller et al., 2011; Patelli et al., 2017), can be interpreted as the techniques to calibrate not the parameters themselves but the uncertainty characteristics, i.e., probabilistic distributions, so that the model outputs are committed not to the maximum fidelity to a single set of the measurements but to the uncertainty characteristics of the multiple sets of the measurements. In the stochastic model updating, uncertainty quantification (UQ) metrics are important to quantify the statistical discrepancy between the model outputs and measurements because of the above three sources of uncertainty. A series of distances, such as the Euclidian distance, Mahalanobis distance, and Bhattacharyya distance has been successfully proposed to define the UQ metrics in the stochastic model updating (Bi et al., 2017). In addition, the Frobenius norm has been also utilized to define the UQ metric to quantify the difference between the covariance matrices of the model outputs and measurements (Govers and Link, 2010). Bi et al. (2019) has recently developed a Bayesian updating framework which employs the approximate Bayesian computation (ABC) technique (Beaumont et al., 2002; Turner and Van Zandt, 2012), where the Bhattacharyya distance-based approximate likelihood is used. This framework has been demonstrated to be capable of calibrating numerical models such that the model outputs recreate wholly the uncertainty characteristics of target measurements. The framework has been furthermore extended to the calibration of dynamic systems, so that the procedure enables to quantify wholly the uncertainty characteristics of the measured time signals (Kitahara et al., 2021).

In the stochastic model updating, distribution families of the parameters commonly need to be assigned a priori, then the prior distribution of the hyper-parameters such as means and variances is updated to the posterior distribution using the measurement data. The distribution families, however, are often unknown beforehand due to the scarce and/or

incomplete available data for the parameters. The recently released NASA UQ challenge problem 2019 (Crespo and Kenny, 2021), for instance, requires a model calibration task in an extremely challenging condition that no distribution information of the aleatory parameters is provided other than a common bounded support domain. In such situation, the assumption on the distribution formats might significantly affect the model updating results. Therefore, Kitahara et al. (2022) has recently developed a distribution-free Bayesian updating framework, where staircase density functions (Crespo et al., 2018) are assigned to the underlying distribution families of the parameters. Staircase density functions enable to flexibly approximate a broad range of distributions arbitrary close, such as highly skewed and/or multi-modal distributions, hence in particular appropriate to characterize the parameters whose density formats cannot be specified. The framework has been demonstrated to be capable to calibrate the probabilistic distribution of the parameters without limiting hypotheses on the distribution families.

Nevertheless, the aforementioned distribution-free model updating framework still has open questions. First, the framework has been currently only demonstrated on the updating using scalar-valued modal responses. Hence, in this study, it is extended to the updating of dynamic systems by measured time signals. Second, staircase density functions are provided for univariate random variables, and thus cannot consider the parameter dependencies, which might lead to inaccurate updating results in the presence of strong correlation among parameters. Copula functions are well-known to be capable of providing an effective way to characterize the dependence structure among parameters, and have been broadly utilized for reliability problems (Tang et al., 2013; Li et al., 2015; Tang et al., 2015). Among various types of copula functions, the Gaussian copula function is most widely used since it can be easily generalized to the multivariate case, and this property is particularly attractive for the stochastic model updating problem, in which very large number of parameters is considered as random variables.

The objective of this work is consequently to develop a stochastic model updating framework for calibrating the joint probabilistic distribution of the correlated parameters without prior knowledge on the distribution families of the marginal distributions. In order to achieve this task, it is assumed that the joint probability distribution of the parameters is characterized by a combination of the Gaussian copula function and staircase density functions. Moment constraints for the existence of the staircase density functions and a correlation coefficient constraint for the existence of the copula function are then derived. Furthermore, the Bhattacharyya distance is used to define an approximate likelihood function quantifying the stochastic discrepancy between the model outputs and measurements, such that the hyper-parameters of the staircase density functions as well as the correlation coefficients of the correlation matrix are calibrated through an ABC updating approach. The proposed framework is first demonstrated on both bi-variate and multi-variate cases

using two simple illustrative examples, and then applied to a model updating problem of a seismic-isolated bridge pier model using the simulated seismic response data, so as to demonstrate the feasibility of the framework in the updating of nonlinear dynamic systems.

The structure of this paper is as follows. Section 5.2 first describes theoretical and methodological bases of the three key ingredients of the proposed framework, i.e., the Bhattacharyya distance-based UQ metrics, staircase density functions, and Gaussian copula function. In Section 5.3, we then outline the formulation of the Bayesian updating for the combination of the Gaussian copula function with the staircase density functions, and the proposed ABC updating framework. Illustrative applications are provided in Section 5.4, employing a simple shear building model and a spring-mass system, and the feasibility of the proposed framework in the updating of nonlinear dynamic systems based on the measured time signals is further demonstrated in Section 5.5. Finally, Section 5.6 gives conclusions to this paper.

## 5.2　Theories and methodologies

### 5.2.1　Bhattacharyya distance-based UQ metrics

The system under investigation in the stochastic model updating is described as:

$$y = h(x) \tag{5.1}$$

where $x = [x_1, x_2, \cdots, x_n]$ indicates a row vector of $n$ input parameters; $y = [y_1, y_2, \cdots, y_m]$ indicates a row vector of $m$ output features; $h(\cdot)$ is the simulator. The output features herein can be either scalar-valued modal responses or time signals. In the latter case, $y$ is replaced to be $y = [y_1, y_2, \cdots, y_m]$, with $y_i = [y_i(0), y_i(1), \cdots, y_i(t)]^T$, $\forall i = 1, 2, \cdots, m$, where $t$ is the time parameter. The simulator $h(\cdot)$ can be either high-fidelity models, e.g., FE models, or surrogates.

Uncertainties involved in the system are first described by representing the input parameters as random variables, and are then propagated through the simulator to the output features. This can be achieved by randomly generating the multiple sets of the parameters and corresponding output features. Suppose the sample size be $N_{\text{sim}}$, the simulator $h$ is evaluated $N_{\text{sim}}$ times for obtaining the sample set of the simulated features $\mathbf{Y}_{\text{sim}} \in \mathbb{R}^{N_{\text{sim}} \times m}$:

$$\mathbf{Y}_{\text{sim}} = \left[ \mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \cdots, \mathbf{y}^{(N_{\text{sim}})} \right]^T, \text{with } \mathbf{y}^{(k)} = \left[ y_1^{(k)}, y_2^{(k)}, \cdots, y_m^{(k)} \right], \forall k = 1, 2, \cdots, N_{\text{sim}} \tag{5.2}$$

in the case where the output features are represented as the modal responses. $\mathbf{Y}_{\text{sim}}$ can be simply extended to be $\mathbf{Y}_{\text{sim}} \in \mathbb{R}^{N_{\text{sim}} \times m \times (t+1)}$ for the time signals case.

In addition to the simulated features, corresponding observed features are also necessary in the model updating. Suppose the number of observations be $N_{\text{obs}}$, the sample set of the observed features $\mathbf{Y}_{\text{obs}}$ possesses a same structure as Equation (5.2), but only the number of rows is changed from $N_{\text{sim}}$ to $N_{\text{obs}}$. The stochastic model updating is aimed at minimizing the stochastic discrepancy between $\mathbf{Y}_{\text{sim}}$ and $\mathbf{Y}_{\text{obs}}$ by calibrating the joint distribution of the parameters.

To quantify the discrepancy between $\mathbf{Y}_{\text{sim}}$ and $\mathbf{Y}_{\text{obs}}$, the Bhattacharyya distance-based UQ metric is employed in this study. The original definition of the Bhattacharyya distance is given as (Bhattacharyya, 1964):

$$d_B(\mathbf{Y}_{\text{sim}}, \mathbf{Y}_{\text{obs}}) = -\log\left[\int_{\mathbb{y}} \sqrt{f_{\mathbf{Y}_{\text{sim}}}(\boldsymbol{y}) f_{\mathbf{Y}_{\text{obs}}}(\boldsymbol{y})}\, \mathrm{d}\boldsymbol{y}\right] \tag{5.3}$$

where $f_{(\cdot)}(\boldsymbol{y})$ indicates the probability density function (PDF) of the output features $\boldsymbol{y}$; $\mathbb{y}$ is the support domain of the output features that comprises $m$-dimensional space for the modal responses but the $\{m \times (t+1)\}$-dimensional space for the time signals. Equation (5.3) indicates that the Bhattacharyya distance is a measure of the overlap between the two probability distributions. Thus, it is capable to consider not only mean information but whole statistical information of two different sample sets. However, the direct evaluation of Equation (5.3) is usually impractical because precisely estimating the joint PDF of the output features is non-trivial due to the necessity of time-consuming repeated model evaluations or the very limited number of available measurement data. To overcome this issue, Bi et al. (2019) proposed the so-called binning algorithm to evaluate the probability mass function (PMF) of the given sample sets, such that the discrete Bhattacharyya distance is utilized instead (Patra et al., 2015):

$$d_B(\mathbf{Y}_{\text{sim}}, \mathbf{Y}_{\text{obs}}) = -\log\left\{\sum_{j=1}^{N_{\text{bin}}} \sqrt{P_{\mathbf{Y}_{\text{sim}}}^{(j)} P_{\mathbf{Y}_{\text{obs}}}^{(j)}}\right\} \tag{5.4}$$

where $N_{\text{bin}}$ denotes the total number of bins; $P_{(\cdot)}^{(j)}$ indicates the PMF value of the output features at the $j$th bin. In the binning algorithm, a grid is created in the whole support domain of the output features, and hence the total number of bins would be $N_{\text{bin}} = n_{\text{bin}}^m$ for the modal responses and $N_{\text{bin}} = n_{\text{bin}}^{m \times (t+1)}$ for the time signals, where $n_{\text{bin}}$ indicates the number of bins for each output feature. The readers can refer to Bi et al. (2019) for the detailed procedure of the binning algorithm. The discrete Bhattacharyya distance-based UQ metric has been demonstrated to be effective in the relatively low-dimensional problems (e.g., the dimension is less than six).

Contrary to that, even the evaluation of Equation (5.4) is still impractical for the very high-dimensional problems where the output features comprise time signals, since the number of bins is exponentially increasing with the number

of dimensions because of the so-called curse of dimensionality. To tackle this issue, Kitahara et al. (2021) has proposed a dimension reduction procedure to utilize the Bhattacharyya distance for the comparison of two different time signals, consisting of the following steps:

1) Define a common window length L for both $\mathbf{Y}_{\text{sim}}$ and $\mathbf{Y}_{\text{obs}}$. Divide them into three-dimensional sub-arrays $\mathbf{Y}_{\text{sim}}^s \in \mathbb{R}^{N_{\text{sim}} \times m \times L}$ and $\mathbf{Y}_{\text{obs}}^s \in \mathbb{R}^{N_{\text{obs}} \times m \times L}$, $\forall s = 1, \cdots, \lfloor(t+1)/L\rfloor$, where $\lfloor \cdot \rfloor$ means the lower integer of the investigated values;

2) Compute the root mean square (RMS) matrices $\mathbf{R}_{\mathbf{Y}_{\text{sim}}}^s \in \mathbb{R}^{N_{\text{sim}} \times m}$ of each sub-array $\mathbf{Y}_{\text{sim}}^s$ along its third dimension and obtain the sample set of the RMS values $\mathbf{R}_{\mathbf{Y}_{\text{sim}}} \in \mathbb{R}^{N_{\text{sim}} \times m \times \lfloor(t+1)/L\rfloor}$. Do similar procedure for the observed features and obtain $\mathbf{R}_{\mathbf{Y}_{\text{obs}}} \in \mathbb{R}^{N_{\text{obs}} \times m \times \lfloor(t+1)/L\rfloor}$;

3) Evaluate in total $\lfloor(t+1)/L\rfloor$ Bhattacharyya distances $d_B^s$ between two sample sets $\mathbf{R}_{\mathbf{Y}_{\text{sim}}}^s$ and $\mathbf{R}_{\mathbf{Y}_{\text{obs}}}^s$ using Eq. (4);

4) Employ the RMS value of the set of the Bhattacharyya distances, $R_{d_B}$, as the UQ metric.

The authors' experience shows that $L = (0.02 \sim 0.03) \cdot t$ is the reasonable choice for the window length L, and such choice indicates that each window contains $2 \sim 3$ % of the time signals. As such, the time signals are degraded to a series of RMS values. The above defined Bhattacharyya distance-based UQ metric has been demonstrated to be able to quantify the uncertainty characteristics of the entire time signals (Kitahara et al., 2021).

## 5.2.2  Staircase density functions

Let the input parameter $x_i$, $\forall i = 1, \cdots, n$, be a random variable having the support domain $[\underline{x}_i, \overline{x}_i]$ and a quadruple of the hyper-parameters $\boldsymbol{\theta}_{x_i} = [\mu_i, m_{2i}, \widetilde{m}_{3i}, \widetilde{m}_{4i}]$ consisting of the mean $\mu_i$, variance $m_{2i}$, skewness $\widetilde{m}_{3i}$, and kurtosis $\widetilde{m}_{4i}$. The skewness $\widetilde{m}_{3i}$ and kurtosis $\widetilde{m}_{4i}$ are defined as ratios of the variance to the third and fourth central moments by $\widetilde{m}_{3i} = m_{3i}/m_{2i}^{3/2}$ and $\widetilde{m}_{4i} = m_{4i}/m_{2i}^2$, respectively. The feasibility condition for the existence of $x_i$ can be defined as moment constraints given by a series of inequalities $\Theta_i = \{\boldsymbol{\theta}_{x_i}: g(\boldsymbol{\theta}_{x_i}) \leq 0\}$, and their components are summarized in Table 5.1 (Sharma et al., 2009; Kumar, 2002).

Let the support domain $[\underline{x}_i, \overline{x}_i]$ equally partitioned into $n_b$ subintervals with the length $\kappa = (\overline{x}_i - \underline{x}_i)/n_b$, $x_i$ can be considered as a staircase random variable, and then its PDF $f_{x_i}(x)$ can be expressed as (Crespo et al., 2018):

$$f_{x_i}(x) = \begin{cases} l^j & \forall x \in (x_i^j, x_i^{j+1}], \forall j = 1, 2, \cdots, n_b \\ 0 & \text{otherwise} \end{cases} \tag{5.5}$$

where $l^j$ is the PDF value of the $j$th bin; $x_i^j = \underline{x}_i + (j-1)\kappa$ is the left partitioning point of the $j$th bin. It is noted that $l^j$ holds that $l^j \geq 0$ for all the bins and $\kappa \sum_{j=1}^{n_b} l^j = 1$.

Table 5.1 Moment constraints for the existence of staircase density functions.

| Hyper-parameters | Moment constraints |
|---|---|
| Mean $\mu_i$ | $g_1 = \underline{x}_i - \mu_i$ |
| | $g_2 = \mu_i - \overline{x}_i$ |
| Variance $m_{2i}$ | $g_3 = -m_{2i}$ |
| | $g_4 = m_{2i} - v_i$ |
| Skewness $\widetilde{m}_{3i}$ | $g_5 = m_{2i}^2 - m_{2i}(\mu_i - \underline{x}_i)^2 - \widetilde{m}_{3i}m_{2i}^{3/2}(\mu_i - \underline{x}_i)$ |
| | $g_6 = \widetilde{m}_{3i}m_{2i}^{3/2}(\overline{x}_i - \mu_i) - m_{2i}(\overline{x}_i - \mu_i)^2 + m_{2i}^2$ |
| | $g_7 = 4m_{2i}^2 + \widetilde{m}_{3i}^2 m_{2i}^3 - m_{2i}^2(\overline{x}_i - \underline{x}_i)^2$ |
| | $g_8 = 6\sqrt{3}\widetilde{m}_{3i}m_{2i}^{3/2} - (\overline{x}_i - \underline{x}_i)^3$ |
| | $g_9 = -6\sqrt{3}\widetilde{m}_{3i}m_{2i}^{3/2} - (\overline{x}_i - \underline{x}_i)^3$ |
| Kurtosis $\widetilde{m}_{4i}$ | $g_{10} = -\widetilde{m}_{4i}m_{2i}^2$ |
| | $g_{11} = 12\widetilde{m}_{4i}m_{2i}^2 - (\overline{x}_i - \underline{x}_i)^4$ |
| | $g_{12} = (\widetilde{m}_{4i}m_{2i}^2 - v_i m_{2i} - u_i\widetilde{m}_{3i}m_{2i}^{3/2})(v_i - m_{2i}) + (\widetilde{m}_{3i}m_{2i}^{3/2} - \mu_i m_{2i})^2$ |
| | $g_{13} = \widetilde{m}_{3i}^2 m_{2i}^3 + m_{2i}^3 - \widetilde{m}_{4i}m_{2i}^3$ |

$^a$ $u_i = \underline{x}_i + \overline{x}_i - 2\mu_i$ and $v_i = (\mu_i - \underline{x}_i)(\overline{x}_i - \mu_i)$.

The PDF values $l_i$ are derived by solving the following optimization problem (Crespo et al., 2018):

$$\hat{l}_i = \underset{l \geq 0}{\operatorname{argmin}} \left\{ J(l) : \sum_{j=1}^{n_b} \int_{x_i^j}^{x_i^{j+1}} x l^j dx = \mu_i, \sum_{j=1}^{n_b} \int_{x_i^j}^{x_i^{j+1}} (x - \mu_i)^r l^j dx = m_{ri}, r = 2, 3, 4 \right\} \tag{5.6}$$

where $J(\cdot)$ is an arbitrary selected cost function expressed as:

$$J(l) = l^T I l \tag{5.7}$$

where $\mathbf{I}$ denotes the identity matrix. This cost function leads to the resultant staircase random variables minimizing the squared sum of the likelihood at each bin. Based on the moment matching constraints, Equation (5.6) can be rewritten as (Crespo et al., 2018):

$$\hat{l}_i = \underset{l \geq 0}{\operatorname{argmin}} \{ J(l) : \mathbf{A}(\boldsymbol{\theta}_{x_i}, n_b) l = \boldsymbol{b}(\boldsymbol{\theta}_{x_i}), \boldsymbol{\theta}_{x_i} \in \Theta_i \} \tag{5.8}$$

where

$$\mathbf{A} = \begin{bmatrix} \kappa \boldsymbol{e} \\ \kappa \boldsymbol{c} \\ \kappa \boldsymbol{c}^2 + \kappa^3/12 \\ \kappa \boldsymbol{c}^3 + \kappa^3 \boldsymbol{c}/4 \\ \kappa \boldsymbol{c}^4 + \kappa^3 \boldsymbol{c}^2/2 + \kappa^5/80 \end{bmatrix}, \text{and } \boldsymbol{b} = \begin{bmatrix} 1 \\ \mu_i \\ \mu_i^2 + m_{2i} \\ \widetilde{m}_{3i}m_{2i}^{3/2} + 3\mu_i m_{2i} + \mu_i^3 \\ m_{4i}m_{2i}^2 + 4\widetilde{m}_{3i}m_{2i}^{3/2}\mu_i + 6m_{2i}\mu_i^2 + \mu_i^4 \end{bmatrix}$$

where $\mathbf{c}$ indicates a column vector of the centre of the bin $c_j = \left(x_i^j + x_i^{j+1}\right)/2$; $\mathbf{c}^n$ indicates the component wise $n$th power of $\mathbf{c}$; $\mathbf{e}$ refers to a vector of ones.

The convexity of the optimization problem in Equation (5.8) enables the fast computation of the staircase density heights. In addition, a relatively small value of $n_b$ (e.g., $n_b = 25\sim50$) is enough for representing practically smooth distribution shapes for the PMF evaluation, which makes the computation further faster. These features are particularly appropriate for the stochastic updating, where the tremendous number of computations of the probability distributions is necessary. Furthermore, the staircase density functions enable to flexibly approximate a broad range of distributions arbitrary close, such as highly skewed and/or multi-modal distributions. Therefore, they can serve as a distribution-free uncertainty characterization model of the parameters whose distribution families cannot be determined.

## 5.2.3   Gaussian copula function

Copula functions couple the multivariate joint cumulative distribution function (CDF) with its one-dimensional marginal CDFs. Conversely, copula functions can be also seen as the multivariate joint CDFs whose one-dimensional marginal CDFs follow a uniform distribution on the interval of [0, 1]. According to Sklar's theorem (Nelsen, 2006), the bivariate joint CDF of two random variables $x_1$ and $x_2$ can be expressed as:

$$F_{\mathbf{x}}(x_1, x_2) = C\left(F_{x_1}(x_1), F_{x_2}(x_2)\right) \tag{5.9}$$

where $F_{x_1}(x_1)$ and $F_{x_2}(x_2)$ indicate the marginal CDFs of $x_1$ and $x_2$, respectively; $C\left(F_{x_1}(x_1), F_{x_2}(x_2)\right)$ is the copula function. From Equation (5.9), the bivariate joint PDF of $x_1$ and $x_2$ is then written as:

$$f_{\mathbf{x}}(x_1, x_2) = c\left(F_{x_1}(x_1), F_{x_2}(x_2)\right) f_{x_1}(x_1) f_{x_2}(x_2) \tag{5.10}$$

where $c\left(F_{x_1}(x_1), F_{x_2}(x_2)\right)$ denotes the copula density function given as:

$$c\left(F_{x_1}(x_1), F_{x_2}(x_2)\right) = c(u_1, u_2) = \frac{\partial^2 c(u_1, u_2)}{\partial u_1 \partial u_2} \tag{5.11}$$

Theoretically, the joint distribution of $x_1$ and $x_2$ can be fully and uniquely represented by Equations (5.9) and (5.10) if the marginal distributions of $x_1$ and $x_2$, and the copula function are given.

There are a lot of copula function types in the literature, including the Gaussian, t, Frank, Gumbel, and Clayton copula functions. They are characterized by their own dependence structures. The latter three types of copula function

can be referred to as Archimedean copulas. The Archimedean copulas have only a single parameter, and hence cannot provide the general dependence structure among multivariate random variables. Alternatively, the general dependence structure is often modeled using the pair-copula decomposition introduced as a canonical vine copula (Haff, 2013). Conversely, the Gaussian and t copulas, which belong to elliptical copulas, can be straightforwardly generalized to the multivariate case. Particularly, the Gaussian copula function is most widely used because it only needs the correlation matrix to determine the dependence structure.

In this study, the joint probability distribution of the input parameters $\boldsymbol{x}$ is finally characterized by the combination of the Gaussian copula function and staircase density functions as:

$$
\begin{aligned}
F_{\mathbf{x}}(\boldsymbol{x}) &= C_G\big(F_{x_1}(x_1), F_{x_2}(x_2), \cdots, F_{x_n}(x_n); \boldsymbol{\rho}\big) \\
&= \Phi_\rho\left(\Phi^{-1}\left(F_{x_1}(x_1)\right), \Phi^{-1}\left(F_{x_2}(x_2)\right), \cdots, \Phi^{-1}\left(F_{x_n}(x_n)\right)\right)
\end{aligned}
\tag{5.12}
$$

where $C_G$ indicates the Gaussian copula function; $\boldsymbol{\rho}$ denotes the correlation matrix; $\Phi_\rho$ means the multivariate standard normal CDF with $\boldsymbol{\rho}$; $\Phi^{-1}$ denotes the inverse function of the standard normal CDF. It is noted that each marginal CDF $F_{x_i}(x_i)$ can be described by the empirical CDF of the staircase density function $f_{x_i}(\cdot)$, for $i = 1,2,\cdots,n$. The correlation matrix $\boldsymbol{\rho}$ can be expressed as:

$$
\boldsymbol{\rho} = \begin{bmatrix}
1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1n} \\
\rho_{12} & 1 & \rho_{23} & \cdots & \rho_{2n} \\
\rho_{13} & \rho_{23} & \ddots & \cdots & \vdots \\
\vdots & \vdots & \vdots & 1 & \rho_{n-1n} \\
\rho_{1n} & \rho_{2n} & \cdots & \rho_{n-1n} & 1
\end{bmatrix}
\tag{5.13}
$$

where $\rho_{ij}$, for $i = 1,2,\cdots,n-1$ and $j = i+1,\cdots,n$, means the correlation coefficient. The range of each correlation coefficient can reach [-1, 1]. The correlation matrix $\boldsymbol{\rho}$ should be the symmetric and positive semi-definite matrix. Thus, the feasibility condition for the existence of the Gaussian copula function can be defined by the correlation coefficient constraint $\mathcal{P} = \{\boldsymbol{\rho}: \text{chol}(\boldsymbol{\rho}) \neq \emptyset\}$, where $\text{chol}(\cdot)$ means the Cholesky factorization.

## 5.3 Distribution-free stochastic model updating

### 5.3.1 Bayesian model updating of the joint probabilistic distribution

In the proposed stochastic updating framework, the well-known Bayesian inference is utilized. It is based on the Bayes' theorem (Beck and Katafygiotis, 1998):

$$P(\boldsymbol{\vartheta}|\mathbf{Y}_{\text{obs}}) = \frac{P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\vartheta})P(\boldsymbol{\vartheta})}{P(\mathbf{Y}_{\text{obs}})} \tag{5.14}$$

where $P(\boldsymbol{\vartheta})$ denotes the prior PDF of the parameters to be inferred $\boldsymbol{\vartheta}$ that is determined by the initial knowledge of the system and expert judgement; $P(\boldsymbol{\vartheta}|\mathbf{Y}_{\text{obs}})$ means the posterior PDF of $\boldsymbol{\vartheta}$ conditional to the measurements, representing the updated knowledge of $\boldsymbol{\vartheta}$; $P(\mathbf{Y}_{\text{obs}})$ denotes the normalization factor (evidence) ensuring the integral of the posterior distribution equal to one; $P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta})$ indicates the likelihood function of $\mathbf{Y}_{\text{obs}}$ that is defined as the PDF values of the measurements conditional to each instance of $\boldsymbol{\vartheta}$.

To calibrate the joint probabilistic distribution of the parameters $\boldsymbol{x}$ given by Equation (5.13), the hyper-parameters of the staircase density functions $\boldsymbol{\theta}_{x_i}$, for $i = 1, \cdots, n$, and the correlation coefficients of the Gaussian copula function $\rho_{ij}$, for $i = 1, \cdots, n-1$ and $j = i+1, \cdots, n$, are accounted for as the parameters to be inferred $\boldsymbol{\vartheta}$. Based on the moment constraints $\Theta_i$, the support domains of $\boldsymbol{\theta}_{x_i}$ can be determined as:

$$\mu_i \in [\underline{x}_i, \overline{x}_i], m_{2i} \in \left[0, \frac{(\overline{x}_i - \underline{x}_i)^2}{4}\right], m_{3i} \in \left[-\frac{(\overline{x}_i - \underline{x}_i)^3}{6\sqrt{3}}, \frac{(\overline{x}_i - \underline{x}_i)^3}{6\sqrt{3}}\right], m_{4i} \in \left[0, \frac{(\overline{x}_i - \underline{x}_i)^4}{12}\right] \tag{5.15}$$

Note that the support domains are defined not for the skewness and kurtosis but the third and fourth central moments, since the skewness and kurtosis are conditional on the variance. On the other hand, the support domain of $\rho_{ij}$ is simply defined as [-1, 1]. In this study, it is assumed that all parameters to be inferred are independent. Based on these support domains with the moment constraints $\Theta_i$ and correlation coefficient constraint P, the prior PDF $P(\boldsymbol{\vartheta})$ is expressed as:

$$P(\boldsymbol{\vartheta}) = \prod_{i=1}^{n} P(\boldsymbol{\theta}_{x_i})I_{\Theta_i}(\boldsymbol{\theta}_{x_i}) \cdot \prod_{i=1}^{n-1}\prod_{j=i+1}^{n} P(\rho_{ij})I_{\mathcal{P}}(\rho_{ij}) \tag{5.16}$$

where $P(\boldsymbol{\theta}_{x_i})$ and $P(\rho_{ij})$ indicate the prior PDFs of the hyper-parameters and correlation coefficients that are chosen as the uniform distributions on their respective support domains; $I_{\Theta_i}(\boldsymbol{\theta}_{x_i})$ denotes the indicator function of $\boldsymbol{\theta}_{x_i}$, which equals to one if $\Theta_i$ is satisfied and otherwise equals to zero; $I_{\mathcal{P}}(\rho_{ij})$ is similarly the indicator function of $\rho_{ij}$. As such, the proposed updating framework requires only assumptions on the support domains of the input parameters $\boldsymbol{x}$.

This brings totally $4n + n(n-1)/2$ parameters to be inferred. However, it is widely recognized that the direct evaluation of the posterior PDF over such a high-dimensional parameter space is not trivial (Beck and Au, 2002). Thus, the well-known advanced sampling technique, termed transitional Markov chain Monte Carlo (TMCMC) (Ching and Cheng, 2007), is employed in this study. TMCMC is a sequential procedure sampling from a series of transitional PDFs

that will gradually converge to the actual posterior PDF, thus it enables to generate samples from the very complex posterior PDF. The readers can refer to Ching and Cheng (2007) and Betz et al. (2016) for more details of the TMCMC algorithm.

## 5.3.2 Approximate Bayesian computation

The likelihood function plays a key role in the Bayesian model updating. Utilizing the Bayesian inference in the stochastic updating results in the following theoretical likelihood function:

$$P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\vartheta}) = \prod_{k=1}^{N_{\text{obs}}} P\big(\mathbf{Y}_{\text{obs}}^{(k)}|\boldsymbol{\vartheta}\big) \tag{5.17}$$

where $P\big(\mathbf{Y}_{\text{obs}}^{(k)}|\boldsymbol{\vartheta}\big)$ is the PDF value of the $k$th observations $\mathbf{Y}_{\text{obs}}^{(k)}$ conditional to each instance of the inferred parameters $\boldsymbol{\vartheta}$. The direct evaluation of Equation (5.17), however, is often impractical since it requires the significant number of model evaluations so as to precisely estimate the PDFs of the corresponding model outputs.

The ABC method (Beaumont et al., 2002; Turner and Van Zandt, 2012) has been successfully used to overcome this obstacle by replacing the above full likelihood function with an approximate but efficient likelihood function that contains information of both the measurements and inferred parameters $\boldsymbol{\vartheta}$. Various forms of the approximate likelihood functions have been investigated in the literature, such as the Gaussian (Patelli et al., 2015), Epanechnikov (Safta et al., 2015), and inverse squared error (Rocchetta et al., 2018) functions. Regardless of the function form, it is essential to utilize the comprehensive UQ metric which can serve as an effective connection between the measurements and inferred parameters. In this study, an approximate likelihood function by the Gaussian function is defined by utilizing the Bhattacharyya distance-based UQ metric as:

$$P_L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\vartheta}) \propto exp\left\{-\frac{d_B^2}{\varepsilon^2}\right\} \tag{5.18}$$

where $\varepsilon$ means the pre-defined width factor, which controls the centralization degree of the posterior PDF. A smaller $\varepsilon$ provides a more peaked posterior PDF, which is more likely to converge to its true value but needs more computation cost for convergence. Hence, its choice is based on specific applications, whereas it is recommended to be within the interval of $[10^{-3}, 10^{-1}]$ (Patelli et al., 2015). By utilizing the Bhattacharyya distance (or the RMS value $R_{d_B}$ for the time signals case), the proposed likelihood enables to quantify the comprehensive uncertainty characteristics of the model outputs and measurements.

The schematic in Figure 5.1 illustrates overall framework of the proposed distribution-free stochastic updating procedure. As already mentioned, only the support domains of the parameters are required to perform this framework. Sampling from the prior PDF in Equation (5.16) can be achieved by the rejection sampling. TMCMC is then utilized to update the inferred parameters to the posterior PDF using the Bhattacharyya distance-based approximate likelihood. By assigning most probable values (MPVs) of the posterior PDF, the joint distribution of the input parameters is finally calibrated such that the stochastic model outputs generated from the joint distribution is capable to recreate wholly the uncertainty characteristics of the target measurements. Note that, the calibrated distribution of the parameters is not necessarily applicable to the reliability analysis in order to estimate the probabilities of rare events, since the parameters are finitely bounded due to the definition of the staircase density functions and a domain where the rare event happens might be excluded.
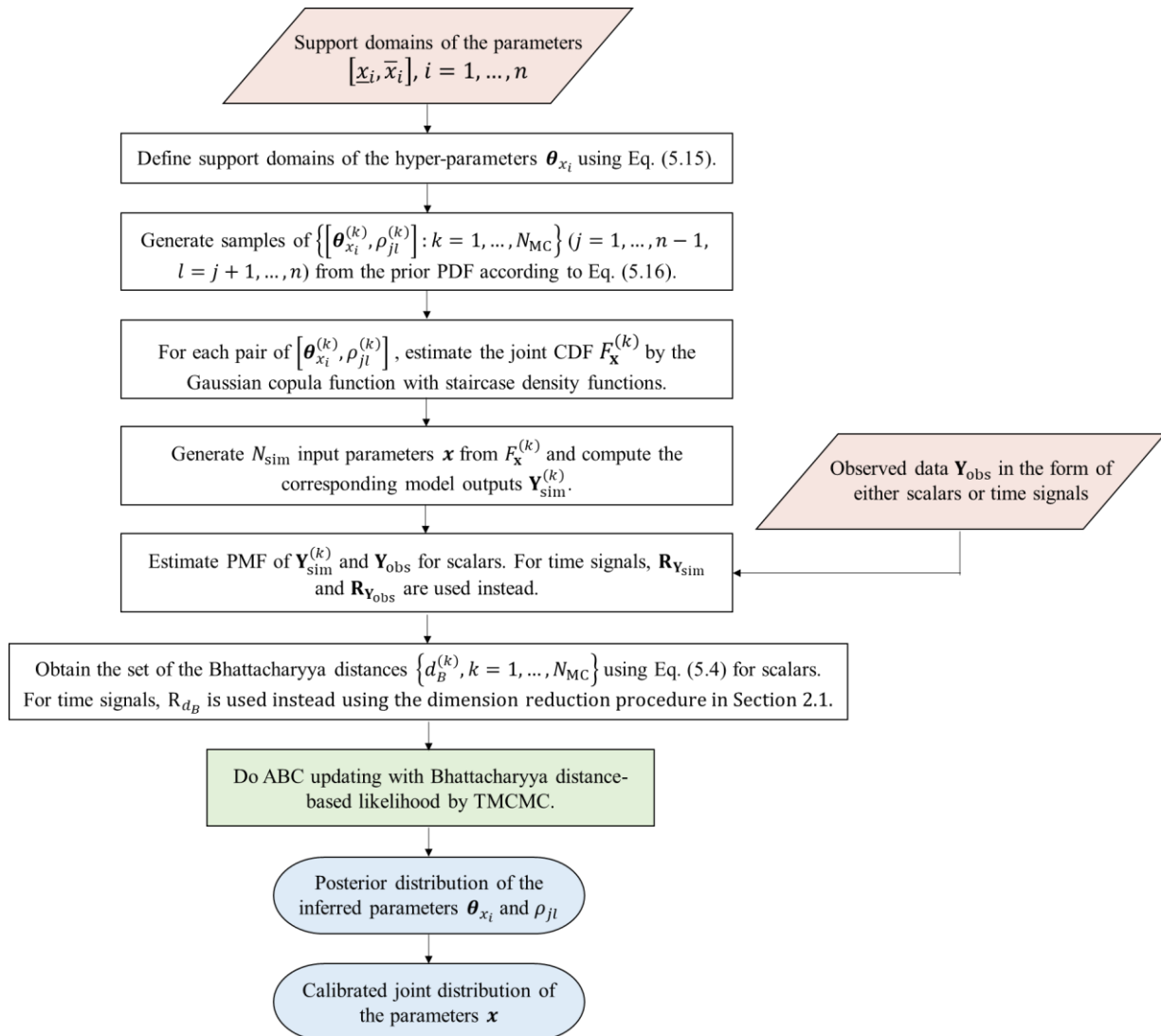


Figure 5.1 Schematic of the proposed stochastic model updating framework.

## 5.4   Principle and illustrative applications

### 5.4.1  Case study I: The two degree of freedom shear building model

The first case study is performed on a two degree of freedom (DOF) shear building model given in Figure 5.2(a). This case study aims at demonstrating the feasibility of the proposed updating procedure for illustrative bivariate case, and how the stochastic model updating fails when ignoring the parameter dependency. This model has been initially introduced by Beck and Au (2002), where the first and second story masses are considered to be the fixed values with $m_1 = 16.531 \times 10^3$ kg and $m_2 = 16.131 \times 10^3$ kg. On the other hand, the first and second interstory stiffnesses are characterized as $k_1 = \bar{k}x_1$ and $k_2 = \bar{k}x_2$, where $x = [x_1, x_2]$ is the inferred parameters, and $\bar{k} = 29.7 \times 10^6$ N/m is the nominal value.
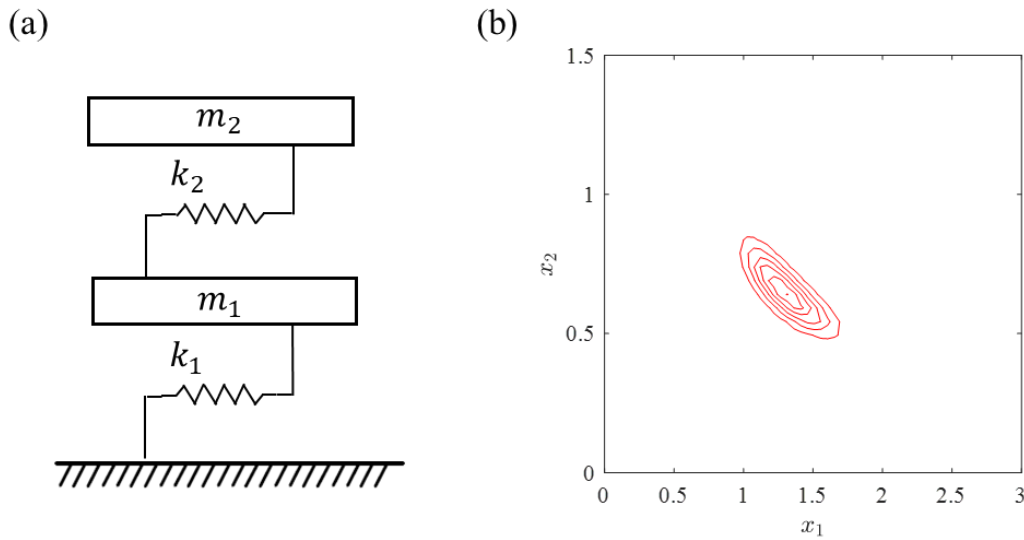


Figure 5.2 (a) 2-DOF shear building model; (b) Posterior distribution in Equation (5.19).

In Beck and Au (2002), the prior PDF $P(x)$ is presented by the pair of uncorrelated lognormal distributions with the MPVs 1.3 and 0.8 for $x_1$ and $x_2$, respectively, and the unit standard deviations. By employing the first two natural frequencies $\tilde{f}_1 = 4.31$ Hz and $\tilde{f}_2 = 9.83$ Hz as the observed features, the posterior PDF $P(x|Y_{\text{obs}})$ is expressed as:

$$P\left(x|[\tilde{f}_1, \tilde{f}_2]\right) \propto exp\left[-\frac{M(x)}{2\sigma^2}\right] P(x) \tag{5.19}$$

where $\sigma = 1/16$ denotes the standard deviation of the prediction error; $M(\cdot)$ means the modal measure-of-fit function expressed as:

$$M(x) = \sum_{j=1}^{2} \lambda^2 \left[ \frac{f_j^2(x)}{\tilde{f}_j^2} - 1 \right]^2 \qquad (5.20)$$

where $\lambda = 1$ is the weight; $f_j(x)$ means the $j$th natural frequency obtained as the model output. Figure 5.2(b) illustrates the posterior distribution in Equation (5.19). It can be seen that the posterior distribution demonstrates a clear negative correlation.

The aforementioned original problem can be interpreted to aim at finding the set of plausible values of the input parameters $x$ using the single set of observed features $[\tilde{f}_1, \tilde{f}_2]$ by the Bayesian scheme. Comparatively, the uncertainty characteristics of the input parameters and observed features are altered hereafter so as to perform the stochastic model updating where the joint probability distribution of the input parameters, $F_x(x)$, is calibrated using the multiple sets of the observed features. The target joint distribution is defined to be identical to the posterior PDF in Equation (5.19). The number of observed features is set to be $N_{\text{obs}} = 100$; thus, $N_{\text{obs}}$ sample sets of the input parameters are generated from the target distribution $P(x|[\tilde{f}_1, \tilde{f}_2])$ using TMCMC, and the corresponding observed features $Y_{\text{obs}}$ are collected by evaluating the model with these sample sets. Note that the target distribution of the input parameters is unknown beforehand in actual. As such, the altered problem is aimed at calibrating the joint distribution of the input parameters to recreate wholly the uncertainty characteristics of $Y_{\text{obs}}$ by the model outputs generated based on the joint probability distribution.

The bounded support domains of $x_1$ and $x_2$ are determined as presented in Table 5.2. The support domain of the hyper-parameters $\theta_{x_1}$ and $\theta_{x_2}$ can be computed using Equation (5.15). Let the sample size be $N_{\text{MC}} = 1000$, $N_{\text{MC}}$ sets of the initial values of $\theta_{x_1}$ and $\theta_{x_2}$, maintaining the moment constraints $\Theta_1$ and $\Theta_2$, are generated using the rejection sampling while $N_{\text{MC}}$ initial values of the correlation coefficient $\rho_{12}$ are arbitrary generated from its support of $[-1, 1]$. For each set of $[\theta_{x_1}, \theta_{x_2}, \rho_{12}]$, the joint probability distribution of the input parameters $x$ described by the Gaussian copula function with the marginal staircase density functions is determined. The number of bins in staircase density estimation is chosen as $n_b = 25$. At the same time, the number of simulated features is set to be $N_{\text{sim}} = 1000$; hence, in total $N_{\text{sim}}$ sample sets of the input parameters $x$ are generated from each joint distribution $\left\{ F_x^{(k)} : k = 1, \cdots, N_{\text{MC}} \right\}$. The corresponding initial simulated features $Y_{\text{sim}}^{(k)}$ are then collected by evaluating the model for each sample sets of the input parameters. Arbitrary selected initial simulated features are illustrated in Figure 5.3, together with the target observed features. The figure clearly demonstrates the presence of significant discrepancy between the simulated and

observed features, implying the necessity of the stochastic model updating for better representation of the uncertainty characteristics of the observed features by means of the model outputs.

Table 5.2 Uncertainty characteristics of the 2-DOF model.

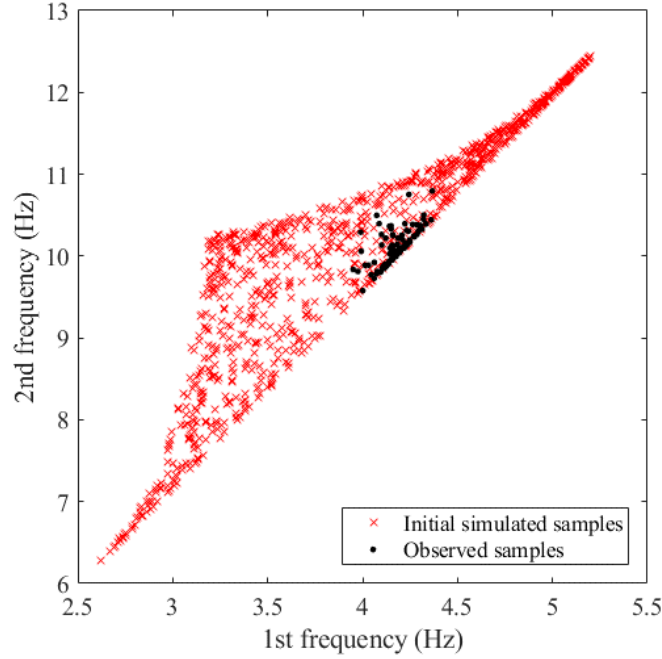| Parameter | Support domain | Target distribution |
|---|---|---|
| $x_1$ | $x_1 \in [0, 3.0]$ | The marginal distribution of Equation (5.19) |
| $x_2$ | $x_2 \in [0, 1.5]$ | The marginal distribution of Equation (5.19) |



Figure 5.3 Observed and initial simulated features.

The Bhattacharyya distance is estimated for each set of the simulated features $\mathbf{Y}_{\text{sim}}^{(k)}$. The number of bins in the binning algorithm is chosen to be $n_{\text{bin}} = 5$. Then, the Bayesian model updating of totally nine inferred parameters (i.e., the hyper-parameters $\boldsymbol{\theta}_{x_i} = \{\mu_i, m_{2i}, \widetilde{m}_{3i}, \widetilde{m}_{4i}\}$, for $i = 1, 2$, as well as the correlation coefficient $\rho_{12}$) is performed using the Bhattacharyya distance-based likelihood function. The width factor in the likelihood function is set to be $\varepsilon = 0.02$.

Figure 5.4 shows the posterior PDFs of all the inferred parameters obtained after totally ten TMCMC iterations, together with their target and calibrated values. The target values are estimated based on samples generated from the target joint distribution, while the calibrated values are estimated as the MPVs of the posterior PDFs. These values are summarized in Table 5.3. It can be seen that the posterior PDFs of all the inferred parameters are significantly updated compared with their uniform priors that are identical to ranges of the horizontal axes of Figure 5.4. Compared with the

means and variances, the posterior supports of the skewnesses and kurtoses are, however, not reduced much from their prior supports, fulfilling the general experience that the higher order moments are difficult to be precisely updated compared with the means (Li et al., 2021; Liao et al., 2022). Nevertheless, the calibrated values including the kurtoses and the correlation coefficient are in good agreement with their target values with the largest relative error less than 6 %, except for the variance $m_{21}$ and two skewnesses $\tilde{m}_{31}$ and $\tilde{m}_{32}$. The relative errors are provided as percentages in the parentheses after the calibrated values in Table 5.3. It is noted that, the large relative error in $m_{21}$ can be explained to be caused by its quite small target value, and it is within an allowable limit to evaluate the staircase density function approximating the target distribution as in Figure 5.5. On the other hand, the large errors in $\tilde{m}_{31}$ and $\tilde{m}_{32}$ are apparently caused by the wrongly identified signs, while their absolute values are close to those of the targets. It should be noted that, the signs of the skewnesses do not strongly affect the uncertainty characteristics of the output features as long as their absolute values are small such that the resultant distributions are almost symmetric. To this end, it is demonstrated that the Bhattacharyya distance can quantify not only mean information but also higher statistical information, i.e., the variances, skewnesses, and kurtoses as well as correlation coefficient.
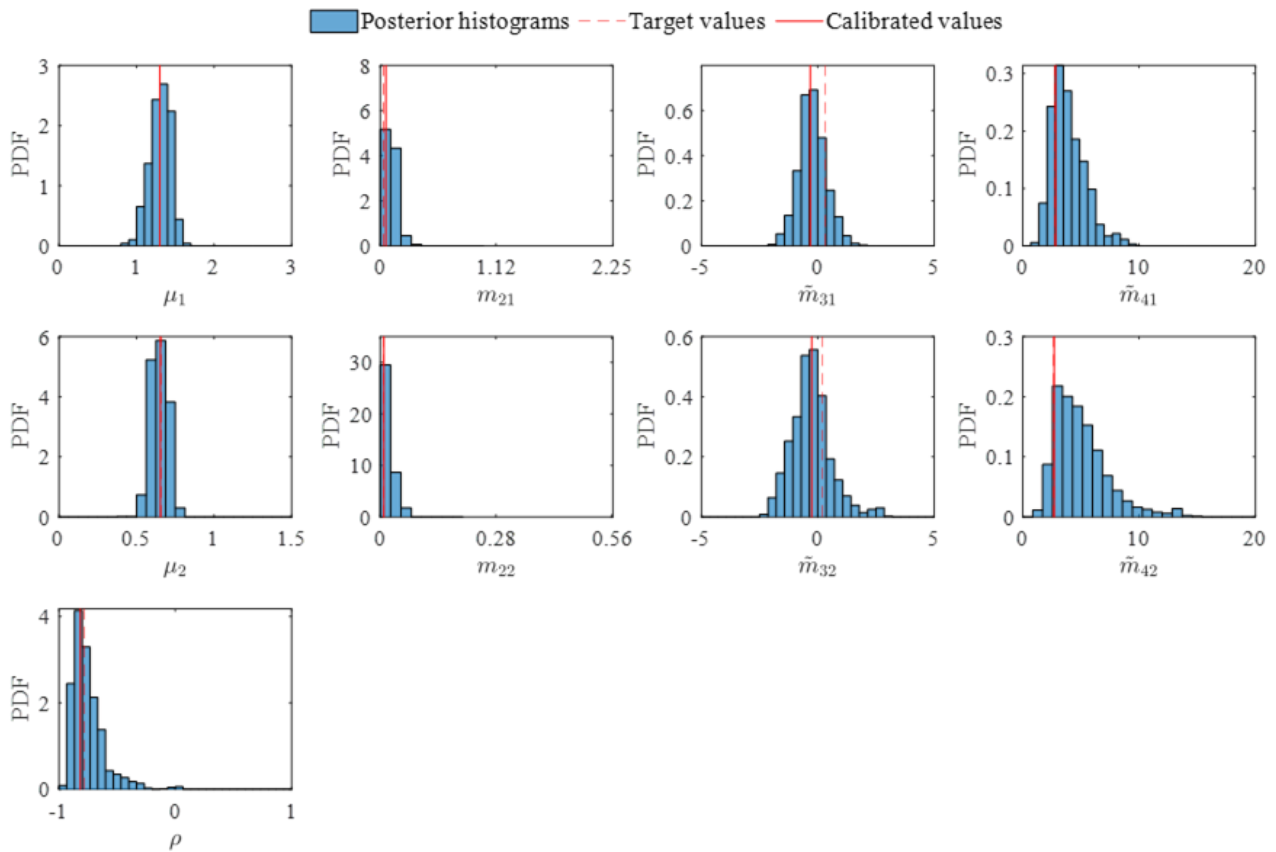


Figure 5.4 Posterior PDFs of the inferred parameters.

Table 5.3 Calibrated parameters of the 2-DOF model.

| Parameter | Target value | Calibrated value | Calibrated value without correlation |
|---|---|---|---|
| $\mu_1$ | 1.3007 | 1.3000 (-0.05 %) | 1.1214 (-13.78 %) |
| $m_{21}$ | 0.0348 | 0.0550 (48.85 %) | 0.0157 (-54.89 %) |
| $\widetilde{m}_{31}$ | 0.3102 | -0.3200 (-203.16 %) | 0.0022 (-99.29 %) |
| $\widetilde{m}_{41}$ | 2.7503 | 2.8300 (2.90 %) | 2.2912 (-16.69 %) |
| $\mu_2$ | 0.6568 | 0.6540 (-0.43 %) | 0.7474 (13.79 %) |
| $m_{22}$ | 0.0085 | 0.0090 (5.88 %) | 0.0031 (63.53 %) |
| $\widetilde{m}_{32}$ | 0.1866 | -0.1780 (-195.39 %) | -0.0849 (-145.50 %) |
| $\widetilde{m}_{42}$ | 2.6679 | 2.7000 (1.20 %) | 2.4355 (-8.71 %) |
| $\rho_{12}$ | -0.7858 | -0.8172 (-4.00 %) | – |

For further demonstrating the results, the sample sets of the input parameters $x$ are generated from the calibrated joint distribution (i.e., the Gaussian copula function with the marginal staircase density functions) and are illustrated in Figure 5.5. It can be seen that the samples generated from the calibrated distribution show good agreement with the target distribution. Meanwhile, the Bayesian updating of only the marginal staircase density functions is also performed to demonstrate how the stochastic model updating fails by ignoring the parameter dependency. The calibrated values of all the hyper-parameters are listed in the last column of Table 5.3. Most of the parameters denote quite large relative errors when compared to those estimated by taking int account the parameter dependency. The sample sets of the input parameters $x$ generated from the calibrated uncorrelated staircase density functions are also plotted in the figure, which are only distributed in a part of the target probability distribution, implying the importance of considering the parameter dependency in stochastic model updating.

Finally, Figure 5.6 illustrates the updated simulated features of $f_1$ and $f_2$ obtained by assigning the calibrated joint distribution to $x$, together with the initial simulated and target observed features. Moreover, the simulated features are also computed for the case ignoring the parameter dependency, and are provided in the figure. It can be seen that the updated simulated features show a distribution equivalent to the observed features for both cases with and without the parameter dependency, while the former provides better results. It implies that the Bhattacharyya distance metric has a potential to recreate wholly the probability distribution of the target observed features regardless of the consideration of the parameter dependency. Nonetheless, these results emphasize the importance of consideration of the parameter dependency in the stochastic model updating, because even though the observed features can be ideally quantified, the incompletely calibrated joint distribution of the input parameters can result in an inaccurate prediction of other quantity of interests, which might be important for other UQ tasks such as the risk assessment and design optimization of the target structure.
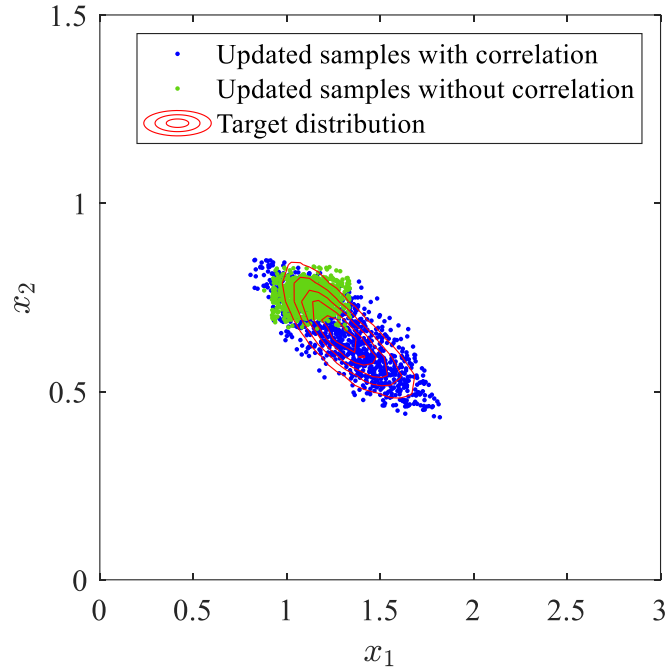
Figure 5.5 Updated samples of the input parameters.
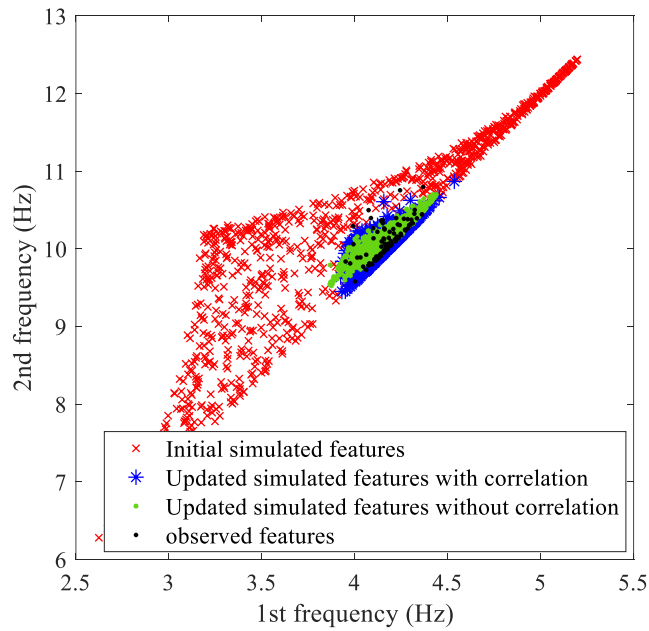


Figure 5.6 Updated simulated features.

## 5.4.2 Case study II: The three degree of freedom spring-mass system

The next case study is performed on a 3-DOF spring-mass system illustrated in Figure 5.7. This case study aims at demonstrating the capability of the proposed updating procedure for multivariate case. This numerical system has

been employed for demonstrating various stochastic updating techniques (Mares et al., 2006; Bi et al., 2019), however, the uncertainty characteristics of the system are altered in this study to demonstrate the proposed method. The stiffness coefficients $k_1$, $k_2$, and $k_3$ are considered to be the uncertain input parameters to be calibrated, whereas the remaining parameters (i.e. $k_4$ to $k_6$ and the three masses $m_1$ to $m_3$) are treated to be the deterministic values: $k_{4-6} = 5.0$ N/m, $m_1 = 0.7$ kg, $m_2 = 0.5$ kg, and $m_3 = 0.3$ kg. The first three natural frequencies $f_1$, $f_2$, and $f_3$ are accounted for as the target output features the uncertainty characteristics of which are driven by the joint probabilistic distribution of $k_1$, $k_2$, and $k_3$ that is assumed to be a correlated tri-variate Gaussian distribution. The given support domains of $k_1$, $k_2$, and $k_3$ and the target values of both the hyper-parameters and correlation coefficients are provided in Table 5.4. Note that the support domains are set to cover more than 99.99 % confidence intervals of the target marginal distributions. Such notification is important because the support domain of the target joint distribution is not bounded, differently from the initial case study.
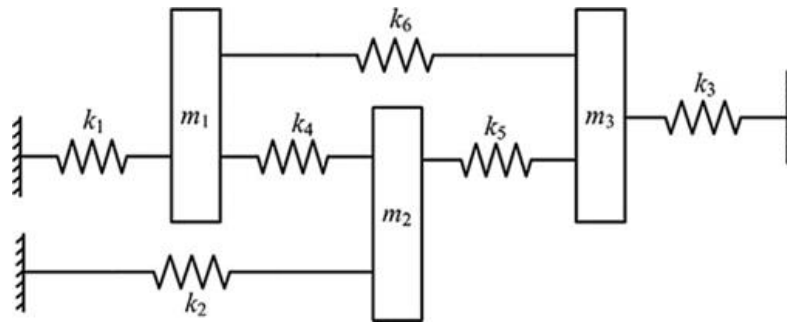


Figure 5.7 3-DOF spring-mass system.

Table 5.4 Uncertainty characteristics of the 3-DOF system.

| Parameter | Support domain | Target distribution |
|---|---|---|
| $k_1$, $k_2$, $k_3$ | $k_1 \in [2.5, 5.5]$, $k_2 \in [4.5, 5.5]$, $k_3 \in [5, 7]$ | Gaussian, $\mu_1 = 4.0$, $\mu_2 = 5.0$, $\mu_3 = 6.0$, $m_{21} = 0.09$, $m_{22} = 0.01$, $m_{23} = 0.04$, $\rho_{12} = 0$, $\rho_{13} = -0.6$, $\rho_{23} = 0.6$ |
| $k_4$–$k_6$, $m_1$–$m_3$ | Deterministic | — |

Consider the number of observed features be $N_{\text{obs}} = 500$, $N_{\text{obs}}$ sample sets of $k_1$, $k_2$, and $k_3$ are generated from the target joint distribution and then the corresponding observed features $\mathbf{Y}_{\text{obs}}$, comprising $f_1$, $f_2$, and $f_3$, are collected by evaluating the model with these sample sets.

On the other hand, let the sample size be $N_{\text{MC}} = 1000$, $N_{\text{MC}}$ sets of the initial values of the hyper-parameters $\boldsymbol{\theta}_{k_1}$, $\boldsymbol{\theta}_{k_2}$, and $\boldsymbol{\theta}_{k_3}$ and the correlation coefficients $\rho_{12}$, $\rho_{13}$, and $\rho_{23}$, satisfying the moment constraints $\Theta_1$, $\Theta_2$, and $\Theta_3$ and the correlation coefficient constraint $\mathcal{P}$, are generated by the rejection sampling in the support domains. For each set

of the hyper-parameters and correlation coefficients, the joint distribution of the three stiffness parameters described by the Gaussian copula function with the marginal staircase density functions is determined. The number of bins $n_b$ is set as the same value as that in the first case study. The number of simulated features is set to be $N_{\text{sim}} = 1000$, so that totally $N_{\text{sim}}$ sample sets of $k_1$, $k_2$, and $k_3$ are generated from each joint distribution. Then, the corresponding initial simulated features are collected by evaluating the model for each sample sets. Figure 5.8 compares the histograms and scatters between the observed features and arbitrary chosen simulated features.
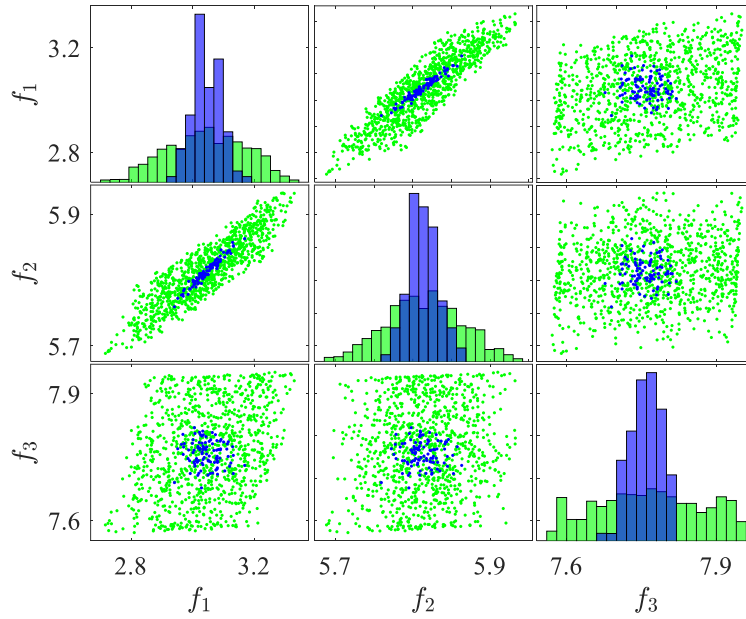


Figure 5.8 Observed features in blue and initial simulated features in green, with the unit in Hz.

The Bhattacharyya distance is estimated for each set of the initial simulated features. The number of bins $n_{\text{bin}}$ is chosen as the same value as that in the first case study. Then, the Bayesian updating of in total 15 inferred parameters, i.e., $\boldsymbol{\theta}_{k_i} = \{\mu_i, m_{2i}, \widetilde{m}_{3i}, \widetilde{m}_{4i}\}$, for $i = 1, 2, 3$, and $\rho_{ij}$, for $i = 1, 2$ and $j = i + 1, 3$, is done utilizing the Bhattacharyya distance-based likelihood function. The width factor $\varepsilon$ is set to be $\varepsilon = 0.01$.

By employing in total 17 TMCMC iterations, all the inferred parameters are well updated to the posterior PDFs. The calibrated values (i.e., the MPVs of the posterior PDFs) of all inferred parameters are shown in Table 5.5, together with the corresponding target values. The relative estimation errors are also given in the parentheses after the calibrated values. Note that the errors are not provided for the skewnesses and the correlation coefficient $\rho_{12}$ because their true values are zero. It can be seen that the calibrated values of the mean and variance parameters are in good agreement with the target values with the largest relative error less than 5 %. On the contrary, the skewness parameters exhibit

differences compared with the target values, especially for $\widetilde{m}_{31}$. However, these calibrated values are small enough for resulting in almost symmetric distributions as similar as the target distributions as depicted in Figure 5.9. The kurtosis parameters also exhibit large errors when compared to their target values, whereas these errors are also permissible to obtain reasonable distributions when compared to the target distributions as shown in Figure 5.9. It implies that both the skewness and kurtosis parameters are relatively insensitive to the uncertainty characteristics of the target output features when compared to the means and variances. It is noted that, it does not mean that such higher order moment parameters are always insensitive to the target output features. In fact, in the previous example, the kurtosis parameters are precisely updated, implying that they are sensitive to the target output features in that example. More importantly, all correlation coefficients are in good agreement with their target values with the largest relative error around 10 %, implying that the proposed procedure is capable to properly capture the correlation structure regardless of no, negative, and positive correlations.

Table 5.5 Calibrated parameters of the 3-DOF system.

| Parameter | Target value | Calibrated value |
|---|---|---|
| $\mu_1$ | 4.0 | 4.0286 (0.72 %) |
| $m_{21}$ | 0.09 | 0.0890 (-1.11 %) |
| $\widetilde{m}_{31}$ | 0 | 0.2220 |
| $\widetilde{m}_{41}$ | 3.0 | 4.4400 (48.00 %) |
| $\mu_2$ | 5.0 | 5.0035 (0.07 %) |
| $m_{22}$ | 0.01 | 0.0104 (4.00 %) |
| $\widetilde{m}_{32}$ | 0 | -0.0500 |
| $\widetilde{m}_{42}$ | 3.0 | 3.8100 (27.00 %) |
| $\mu_3$ | 6.0 | 6.0030 (0.05 %) |
| $m_{23}$ | 0.04 | 0.0410 (2.50 %) |
| $\widetilde{m}_{33}$ | 0 | -0.0214 |
| $\widetilde{m}_{43}$ | 3.0 | 3.9400 (31.33 %) |
| $\rho_{12}$ | 0 | -0.0058 |
| $\rho_{13}$ | -0.6 | -0.5728 (4.53 %) |
| $\rho_{23}$ | 0.6 | 0.5398 (10.03 %) |

The joint distribution of $k_1$, $k_2$, and $k_3$ is then derived as the Gaussian copula function with the marginal staircase density functions, assigned the calibrated values of the hyper-parameters and correlation coefficients. Figure 5.9 shows the marginal CDF of each stiffness, together with the corresponding target marginal distribution. It is noted that, since the marginal distributions are obtained as staircase (i.e., discrete) density functions, the CDFs are estimated from the samples generated according to the staircase density functions, via the kernel density estimation. As can be seen, the estimated staircase density functions are in good agreement with their target marginal distribution, which supports that the above estimate errors in the skewness and kurtosis parameters are still permissible for the purpose of calibrating the joint distribution of the parameters. Nevertheless, it is noted that, in the tail regions, the calibrated distributions, in

particular for $k_1$, remain discrepancy from the target distributions due to the estimate errors in the kurtosis parameters, while the discrepancy in the tail regions do not affect greatly the model outputs as depicted in Figure 5.10. Furthermore, the obtained distributions have bounded support domains which are identical to ranges of the horizontal axes of Figure 5.9 because of the definition of the staircase density functions, while the target Gaussian distributions do not have the bounded supports. It implies that the proposed updating procedure does not limit its applicability to the case where the investigated parameters have bounded supports. However, at the same time, it also indicates that the calibrated joint distribution cannot be employed for reliability analysis, where the target is to estimate the probabilities of rare events that can be occurred out of the support domains. It is noted that, this limitation does not prevent the use of the proposed procedure in the stochastic model updating, since the main motivation of the stochastic model updating is to obtain the model that is capable to describe the system of interest conditioned on the observed data whereas reliability analysis is only one of the potential usages of the calibrated model.



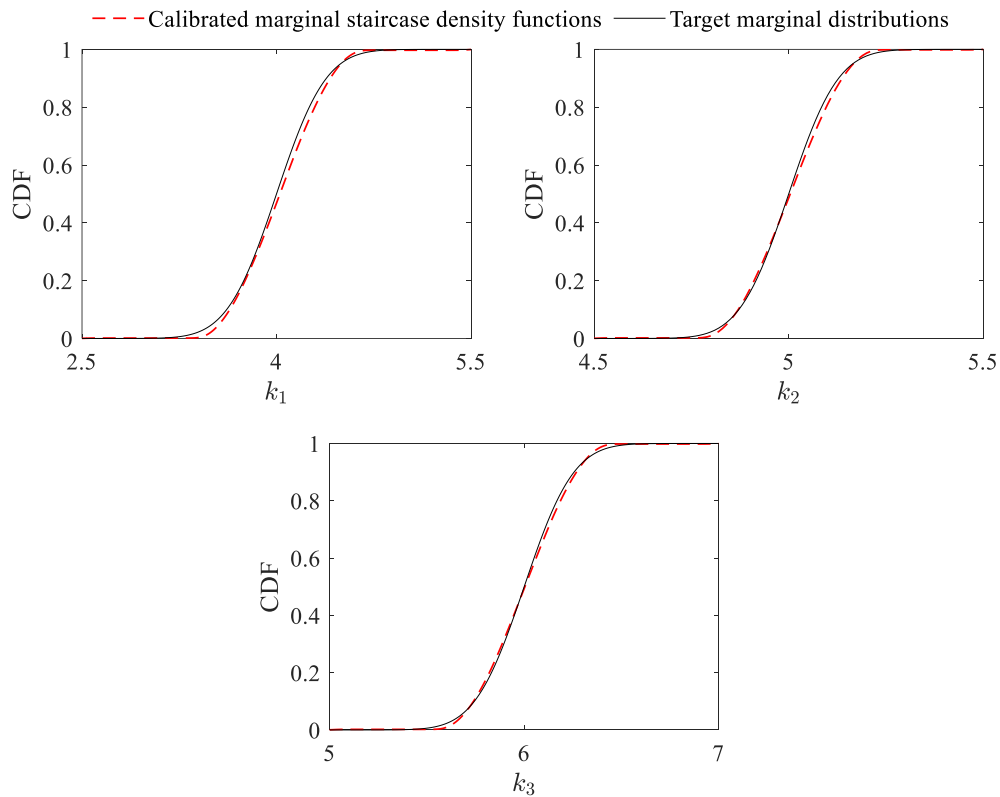Figure 5.9 Calibrated marginal distributions of the input parameters.

Finally, Figure 5.10 compares both the histograms and scatters between the target and updated simulated features. Compared with the initial simulated features in Figure 5.8, It can be seen that the uncertainties in the three stiffness parameters are correctly calibrated by the proposed procedure and the calibrated model is capable to recreate wholly

the uncertainty characteristics of the observed features. As such, the relatively large errors in the higher order moment parameters can be considered to be permissible, because the proposed procedure aims at not estimating the individual moment parameters precisely but rather obtaining model outputs which are identical to the observations to which some higher moment parameters might be relatively insensitive.
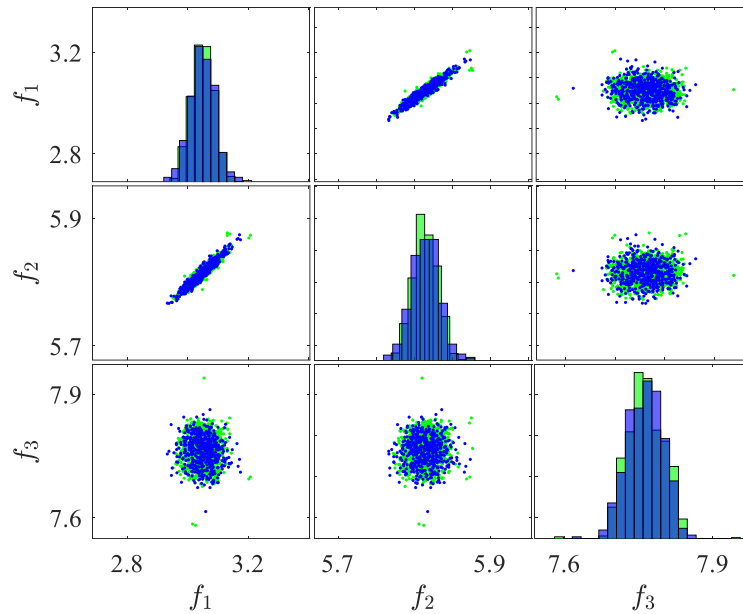


Figure 5.10 Observed features in blue and updated simulated features in green, with the unit in Hz.

## 5.5 Nonlinear dynamic system updating

### 5.5.1 Problem description

The proposed approach is further demonstrated on the updating of nonlinear dynamic systems using the measured time signals. For this purpose, a model updating problem of a reinforced concrete (RC) bridge pier using simulated seismic response data is investigated. The target bridge is a seismic-isolated bridge with lead rubber bearings, designed based on the specifications for highway bridges in Japan (Japan Road Association (JRA), 2016). Its descriptions are detailed in Table 5.6. Figure 5.11 shows the 2-DOF lumped mass model as the numerical model of the target structure, in which the two lumped masses represent the superstructure and RC pier, and the two horizontal springs represent the rubber bearings and RC pier. The boundary condition at the surface is assumed as fixed. The nonlinearity of the rubber bearings is characterized by a bilinear model with the ratio of the yield stiffness $K_{B1}$ to the post-yield stiffness $K_{B2}$ of 6.5:1 based on the manual on bearings for highway bridges in Japan (JRA, 2004). Meanwhile, that of the RC pier is

represented by a bilinear model with the elastoplastic characteristic and the stiffness degradation model, termed Takeda model (Takeda et al., 1970). The well-known Rayleigh damping model is utilized as the damping model in which the damping ratios of the rubber bearings and RC pier are set to be 0 % and 2 %, respectively.

Table 5.6 Descriptions of the target bridge pier.

|  | Structural parameter | Nominal value |
|---|---|---|
| Superstructure | Mass $M_S$ (ton) | 604.0 |
| Rubber bearings | Yield strength (kN) | 1118 |
|  | Yield stiffness $K_{B1}$ (kN/m) | 40000 |
|  | Post-yield stiffness $K_{B2}$ (kN/m) | 6000 |
| RC pier | Mass $M_P$ (ton) | 346.2 |
|  | Yield strength (kN) | 3374 |
|  | Yield stiffness $K_P$ (kN/m) | 110100 |
|  | Yield displacement (m) | 0.0306 |



Figure 5.11 Numerical modeling of the target bridge pier.

The aim of this updating problem is quantifying the uncertainty characteristics of the post-yield stiffness of the rubber bearings, $K_{B2}$, which governs the nonlinear behaviour of the target bridge pier under strong earthquakes, as well as the remaining stiffnesses $K_{B1}$ and $K_P$. These three stiffnesses are parameterized as $K_{B1} = \bar{K}_{B1}x_1$, $K_{B2} = \bar{K}_{B2}x_2$, and $K_P = \bar{K}_P x_3$, where $x = [x_1, x_2, x_3]$ are uncertain input parameters, and $\bar{K}_{B1}$, $\bar{K}_{B2}$, and $\bar{K}_P$ are the nominal values shown in Table 5.6. The other parameters are assumed to be fixed constants with the nominal values. The time-history of the acceleration response at the superstructure subjected to the level-2 type-II-II-1 earthquake introduced in JRA (2016) is used as the target output features. The duration time of the input ground motion is 40 s. Time history analysis of the 2-DOF model is performed by the Newmark $\beta$ method ($\beta = 1/4$ and $\gamma = 1/2$) with the time step $\Delta t = 0.001$ s. Figure 5.12 illustrates the time-history of the acceleration response at the superstructure and the force-displacement responses of the rubber bearings and pier for the case when all parameters are fixed to the nominal values shown in Table 5.6.
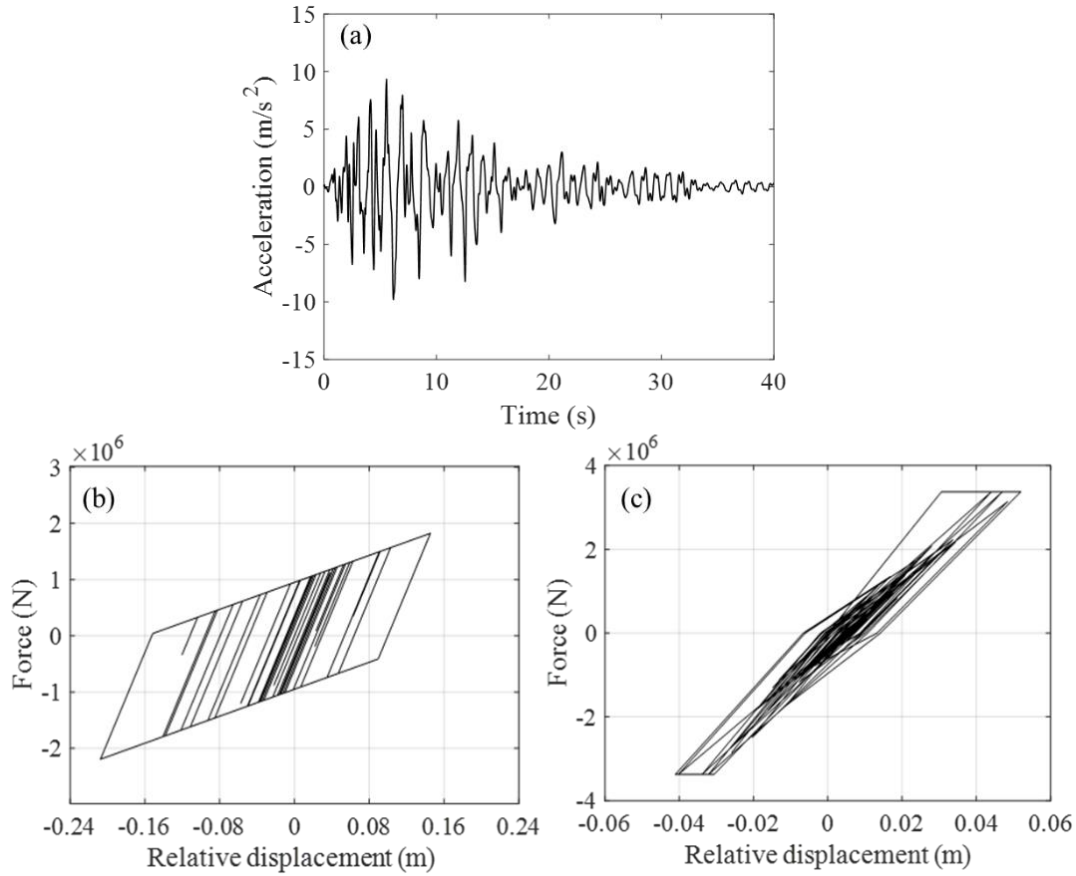
Figure 5.12 (a) Time-history of the acceleration response at superstructure; (b) Force-displacement response of rubber bearings; (c) Force-displacement response of RC pier.

The target joint probability distribution of the input parameters $x$ is considered as a correlated tri-variate Gaussian distribution, where a positive correlation between $x_1$ and $x_2$ (i.e., between the initial and post stiffnesses of the rubber bearings) is induced. The pre-defined support domains of $x_1, x_2$, and $x_3$ and the target values of the hyper-parameters and correlation coefficients are summarized in Table 5.7. Similar to the previous case study, the support domains are determined to cover more than 99.99 % confidence intervals of their target marginal distributions. Suppose the number of observed features be $N_{obs} = 100$, $N_{obs}$ sample sets of the input parameters $x$ are generated according to the target joint distribution and then the corresponding observed features $\mathbf{Y}_{obs}$ of the time-history of the acceleration response at the superstructure, are collected by evaluating the model with these sample sets.

Table 5.7 Uncertainty characteristics of the target bridge pier.

| Parameter | Support set | Target distribution |
|---|---|---|
| $x_1, x_2, x_3$ | $x_1 \in [0.7, 1.3]$, $x_2 \in [0.7, 1.3]$, $x_3 \in [0.7, 1.3]$ | Gaussian, $\mu_1 = 1.0$, $\mu_2 = 1.0$, $\mu_3 = 1.0$, $m_{21} = 0.0049$, $m_{22} = 0.0049$, $m_{23} = 0.0049$, $\rho_{12} = 0.8$, $\rho_{13} = 0$, $\rho_{23} = 0$ |
| $M_S, M_P$ | Deterministic | — |

In this example, totally 13 inferred parameters, i.e., the hyper-parameters $\boldsymbol{\theta}_{x_i} = \{\mu_i, m_{2i}, m_{3i}, m_{4i}\}$, for $i = 1, 2, 3$, and correlation coefficient $\rho_{12}$ is considered. It is noted that the remaining correlation coefficients are assumed as zero in advance and ignored in the updating procedure. Let the sample size be $N_{\mathrm{MC}} = 100$, $N_{\mathrm{MC}}$ sets of the hyper-parameters maintaining the moment constraints $\Theta_1$ $\Theta_2$, and $\Theta_3$, are generated by the rejection sampling whereas $N_{\mathrm{MC}}$ sets of the correlation coefficient are arbitrary generated from the support of [-1, 1]. For each set of $[\boldsymbol{\theta}_{x_1}, \boldsymbol{\theta}_{x_2}, \boldsymbol{\theta}_{x_3}, \rho_{12}]$, the joint probability distribution of the input parameters $\boldsymbol{x}$ is defined as the Gaussian copula function with the marginal staircase density functions. The number of bins $n_b$ is chosen as $n_b = 50$. The number of simulated features, on the other hand, is set as $N_{\mathrm{sim}} = 500$; hence, $N_{\mathrm{sim}}$ sample sets of $\boldsymbol{x}$ are generated from each joint distribution $\left\{F_{\mathbf{x}}^{(k)}: k = 1, \cdots, N_{\mathrm{MC}}\right\}$, and then the corresponding initial simulated features $\mathbf{Y}_{\mathrm{sim}}^{(k)}$ are obtained by evaluating the model with these samples. The window length in the dimension reduction procedure introduced in Section 5.2.1 is set to be $\mathrm{L} = 0.025(t + 1)$, with $t = 40/0.001 = 40000$. Hence, the RMS matrices of both the simulated and observed features, $\mathbf{R}_{\mathbf{Y}_{\mathrm{sim}}^{(k)}}^s$ and $\mathbf{R}_{\mathbf{Y}_{\mathrm{sim}}}^s$, for $\forall s = 1, \cdots, 40$, are defined. Figure 5.13 compares the histograms and scatters between the observed and simulated features by utilizing five arbitrary selected RMS matrices at the time $s = 30, 7, 6, 9, 37$. The figure demonstrates that the target features show strong nonlinearity, making the updating problem significantly challenging. The Bhattacharyya distance is obtained for each pair of the simulated and observed RMS matrices, and the RMS value of the Bhattacharyya distances, $R_{d_B}$, is used as the UQ metric in the approximate likelihood. The number of bins $n_{\mathrm{bin}}$ is chosen as $n_{\mathrm{bin}} = 10$ while the width factor $\varepsilon$ is set as $\varepsilon = 0.01$.

## 5.5.2 Results assessment

By employing totally 13 TMCMC iterations, all the inferred parameters are well updated to the posterior PDFs. The calibrated values of the inferred parameters are detailed in Table 5.8, together with the corresponding target values. The relative estimation errors are also shown in the parentheses after the calibrated values. The calibrated values of all mean parameters and the variance parameter $m_{21}$ are in good agreement with the target values, whereas the remaining variance parameters exhibit large estimate errors because of their quite small target values. In spite of the large relative errors, the calibrated values of the variance parameters are close to the target values compared with the prior supports and these errors are permissible to result in the model outputs close to the observations. Moreover, the skewnesses and kurtoses exhibit differences compared with their target values, whereas these errors are also within allowable limits to achieve the model outputs close to the target observations, as similar as the previous example. Furthermore, the positive

correlation induced is also captured by the proposed procedure though a certain error is still remained when compared to the target value. By assigning these calibrated values, the joint probability distribution of $x$ is tuned identical to the target distribution.
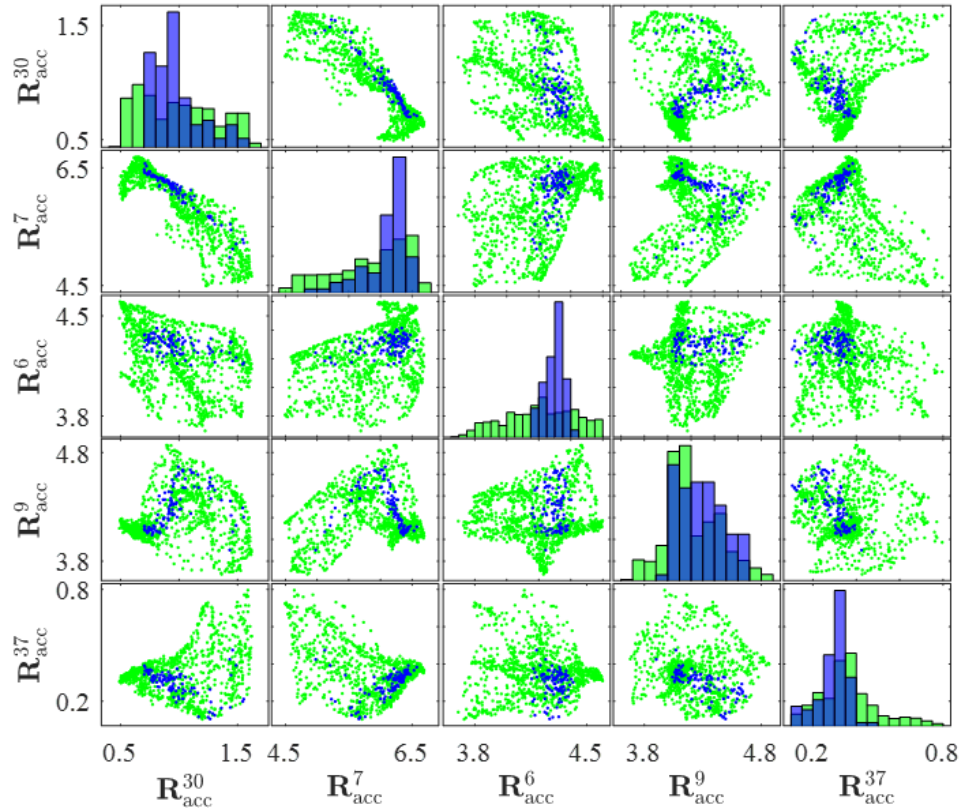


Figure 5.13 Observed features in blue and initial simulated features in green, with the unit in $m/s^2$.

Table 5.8 Calibrated parameters of the target bridge pier model.

| Parameter | Target value | Calibrated value |
|---|---|---|
| $\mu_1$ | 1.0 | 0.9958 (-0.42 %) |
| $m_{21}$ | 0.0049 | 0.0045 (-4.44 %) |
| $\tilde{m}_{31}$ | 0 | -0.1688 |
| $\tilde{m}_{41}$ | 3.0 | 4.3450 (44.83 %) |
| $\mu_2$ | 1.0 | 0.9992 (-0.08 %) |
| $m_{22}$ | 0.0049 | 0.0065 (32.65 %) |
| $\tilde{m}_{32}$ | 0 | 0.4050 |
| $\tilde{m}_{42}$ | 3.0 | 3.9500 (31.67 %) |
| $\mu_3$ | 1.0 | 0.9996 (-0.04 %) |
| $m_{23}$ | 0.0049 | 0.0068 (38.78 %) |
| $\tilde{m}_{33}$ | 0 | -0.3025 |
| $\tilde{m}_{43}$ | 3.0 | 4.3960 (46.53 %) |
| $\rho_{12}$ | 0.8 | 0.6736 (-15.80 %) |

Finally, Figure 5.14 compares the histograms and scatters between the target and simulated features for the five arbitrary selected RMS matrices of $s = 30, 7, 6, 9, 37$. Compared with the initial simulated features in Figure 5.13, It

can be seen the updated simulated features are identical to the observed features capturing the complicated nonlinear structure. This indicates the feasibility of the proposed approach in the stochastic model updating of nonlinear dynamic systems for recreating wholly the uncertainty characteristics of the target measured time signals, even though the prior knowledge about the joint distribution of the parameters is extremely limited.
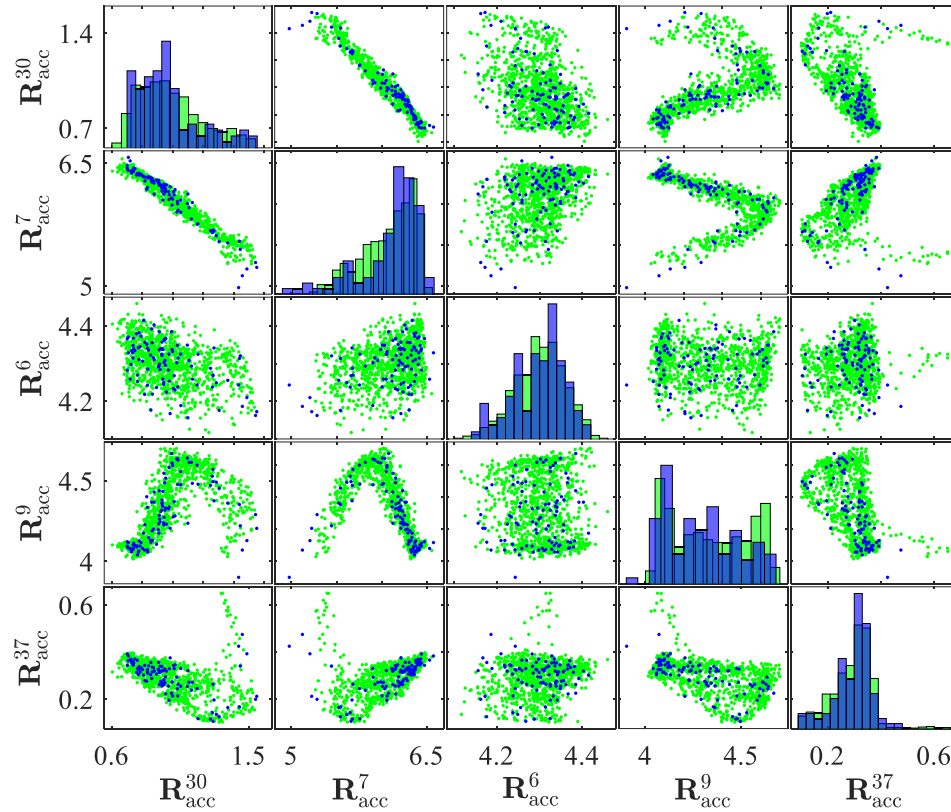


Figure 5.14 Observed features in blue and updated simulated features in green, with the unit in $m/s^2$.

## 5.6    Conclusions

This paper presents three contributions for the calibration of the joint probabilistic distribution of the correlated parameters through the stochastic model updating based on a limited number of measurement data. First, each marginal distribution is characterized by staircase density functions and their hyper-parameters are subjected to be updated. The staircase density functions can flexibly describe a broad range of distributions; thus, no limiting hypotheses about the distribution families is required differently from the most of the available stochastic model updating frameworks. Next, the dependence structure among the parameters are described by the Gaussian copula. The correlation coefficients are also subjected to be updated; thus, even the prior knowledge on the presence of parameter dependencies is not required. Finally, the Bhattacharyya distance-based UQ metric is proposed to define an approximate likelihood which is capable

of quantifying the stochastic discrepancy between the model outputs and measurements. As such, the parameters to be inferred, i.e., the hyper-parameters and correlation coefficients are successfully updated through the Bayesian model updating. Two exemplary applications and the followed nonlinear dynamic system updating problem demonstrate the feasibility of the proposed model updating procedure and the importance of considering the parameter dependency in the stochastic model updating.

However, open problems still exist. First, the cost function in the optimization problem solved for estimating the staircase density function is solely chosen in this study. The staircase density that attains other optimality criteria, such as the maximal entropy, can be similarly formulated, but further studies are necessary to investigate the most suitable choice of the cost function for model updating. Second, the Gaussian copula might not be suitable if the parameters demonstrate a strong nonlinear dependency. The assumption on the copula function type introduces another source of uncertainty, i.e., the model bias, and such uncertainty should be quantified by, for instance, the Bayesian model class selection. These two challenges will be addressed in the future work.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

## References

Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162, 2025-2035.

Bhattacharyya, A. (1946). On a measure of divergence between two multinominal populations. *Indian Journal of Statistics*, 7(4), 401-406.

Beck, J. L., and Au, S. K. (2002). Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *Journal of Engineering Mechanics*, 128(4), 380-391.

Beck, J. L., and Katafygiotis, L. S. (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechnics*, 124, 455-461.

Betz, W., Papaioannou, I., and Straub, S. (2016). Transitional Markov chain Monte Carlo: Observations and improvements. *Journal of Engineering Mechanics*, 142(5), 04016016.

Bi, S., Prabhu, S., Cogan, S., and Atamturktur, S. (2017). Uncertainty quantification metrics with varying statistical information in model calibration and validation. *AIAA Journal*, 55, 3570-3583.

Bi, S., Broggi, M., and Beer, M. (2019). The role of the Bhattacharyya distance in stochastic model updating. *Mechanical System and Signal Processing*, 117, 437-452.

Ching, J., and Chen, Y. C. (2007). Transitional Markov chain Monte Carlo method for Bayesian updating, model class selection, and model averaging. *Journal of Engineering and Mechanics*, 133(7), 816-832.

Crespo, L. G., Kenny, S. P., Giesy, D. P., and Stanford, B. K. (2018). Random variables with moment-matching staircase density function. *Applied Mathematical Modelling*, 64, 196-213.

Crespo, L. G., and Kenny, S. P. (2021). The NASA Langley Challenge on Optimization Under Uncertainty. *Mechanical System and Signal Processing*, 152, 107405.

Goller, B., Broggi, M., Calvi, A., and Schuëller, G. I. (2011). A stochastic model updating technique for complex aerospace structures. *Finite Elements in Analysis and Design*, 47, 739-752.

Haff, L. H. (2013). Parameter estimation for pair-copula constructions. *Bernoulli*, 19, 462-491.

Japan Road Association. (2016). *Design specifications of highway bridges V: Seismic design.* Tokyo, Maruzen.

Japan Road Association. (2004). *Manual on bearings for highway bridges*. Tokyo, Maruzen. (In Japanese.)

Khodaparast, H. H., Mottershead, J. E., and Friswell, M. I. (2008). Perturbation methods for the estimation of parameter variability in stochastic model updating. *Mechanical System and Signal Processing*, 22, 1751-1773.

Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2021). Bayesian model updating in time domain with metamodel-based reliability method. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7, 04021030.

Kitahara, M., Bi, S., Broggi, M., and Beer, M. (2022). Nonparametric Bayesian stochastic model updating with hybrid uncertainties. *Mechanical System and Signal Processing*, 163, 108165.

Kumar, P. (2002). Moment inequalities of a random variable defined over a finite interval. *Journal of Inequalities in Pure and Applied Mathematics*, 3, 1-11.

Li, D. Q., Zhang, L., Tang, X. S., Zhou, W., Li, J. H., Zhou, C. B., and Phoon, K. K. (2015). Bivariate distribution of shear strength parameters using copulas and its impact on geotechnical system reliability. *Computers and Geotechnics*, 68, 184-195.

Li, P., Lu, Z., and Zhao, Y. (2021). Bayesian updating of time-dependent structural reliability using the method of moment. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7, 04021066.

Liao, B., Zhao, R., Yu, K., and Liu, C. (2022). Stochastic model updating method for estimates of arbitrary distributed parameters using resampling technique. *Applied Mathematical Modelling*, 105, 387-405.

Mares, C., Mottershead, J. E., and Friswell, M. I. (2006). Stochastic model updating: Part 1—theory and simulated example. *Mechanical System and Signal Processing*, 20, 1674-1695.

Mottershead, J. E., and Friswell, M. I. (1993). Model Updating In Structural Dynamics. *Journal of Sound and Vibration*, 167, 347-375.

Mottershead, J. E., Ling, M., and Friswell, M. I. (2011). The sensitivity method in finite element model updating: A tutorial. *Mechanical System and Signal Processing*, 25, 2275-2296.

Nelsen, R. B. (2006). *An introduction to copulas*, Second edition, Springer-Verlag, New York.

Patelli, E., Alvarez, D. A., Broggi, M., and de Angelis, M. (2015). Uncertanty management in multidisciplinary design of critical safety systems. *Journal of Aerospace Information Systems*, 12, 140-169.

Patelli, E., Govers, Y., Broggi, M., Gomes, H. M., Link, M., and Mottershead, J. E. (2017). Sensitivity or Bayesian model updating: a comparison of techniques using the DLR AIRMOD test data. *Archive of Applied Mechanics*, 87, 905-925.

Patra, B. K., Launonen, R., Ollikainen, V., and Nandi, S. (2015). A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. *Knowledge-Based Systems*, 82, 163-177.

Rocchetta, R., Broggi, M., Huchet, Q., and Patelli, E. (2018). On-line Bayesian model updating for structural health monitoring. *Mechanical Systems and Signal Processing*, 103, 174-195.

Safta, C., Sargsyan, K., Najm, H. N., Chowdhary, K., Debusschere, B., Swiler, L. P., and Eldred, M. S. (2015). Probabilistic methods for sensitivity analysis and calibration in the NASA challenge problem. *Journal of Aerospace Information Systems,* 12(1), 170-188.

Sairajan, K. K., and Aglietti, G. S., (2012). Robustness of system equivalent reduction expansion process on spacecraft structure model validation. *AIAA Journal*, 50. 2376-2388.

Shan, D., Li, Q., Khan, I., and Zhou, X. (2015). A novel finite element model updating method based on substructure and response surface model. *Engineering Structures*, 103, 147-156.

Sharma, R., Devi, S., Kapoor, G., and Barnett, N. S. (2009). A brief note on some bounds connecting lower order moments for random variables defined on a finite interval. *Interval Journal of Theoretical & Applied Sciences*, 1, 83-85.

Takeda, T., Sozen, M. A., and Nielsen, N. N. (1970). Reinforced concrete response to simulated earthquakes. *Journal of the Structural Division*, 96(12), 2557-2573.

Tang, X. S., Li, D. Q., Zhou, C. B., Phoon, K. K., and Zhang, L. M. (2013). Impact of copulas for modeling bivariate distributions on system reliability. *Structural Saftey*, 44, 80-90.

Tang, X. S., Li, D. Q., Zhou, C. B., and Phoon, K. K. (2015). Copula-based approaches for evaluating slope reliability under incomplete probability information. *Structural Saftey*, 52, 90-99.

Turner, B. M., and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2), 69-85.

# Chapter 6

# Research article 5: A distributionally robust approach for mixed aleatory and epistemic uncertainties propagation

This is the last phase of the main part of this thesis, that aims at extending the generalized NISS framework so as to propagate the category IV parameters whose distribution families are unknown a priori, by representing them as the distributional p-boxes based on the staircase density functions. This is motivated by the NASA UQ challenge problem 2019 addressed in Chapter 4, where it is revealed that the computational burden of the probability bounds analysis for the p-boxes constructed based on the staircase density functions by the double-loop Monte Carlo method might not be permissible for practical applications. As such, a more efficient method to propagate the p-boxes based on the staircase density functions is required as a robust tool for uncertainty propagation that is effective even on the very challenging condition where the prior knowledge on aleatory parameters is extremely limited.

As has been mentioned in the introduction chapter of this thesis, the generalized NISS framework is designed to propagate the category IV parameters described as the parameterized p-boxes and category III parameters represented as the interval models simultaneously. Thus, it is theoretically applicable to the propagation of the p-boxes constructed based on the staircase density functions. However, it requires, in its current form based on the global NISS method, to parameterize the distribution function for a significant number of the hyper-parameters sets to derive NISS estimators. The staircase density functions are parameterized by solving optimization problems; hence, their excessively repeated parameterizations can be computationally prohibitive.

To overcome this limitation, the present article develops a novel hybrid NISS method. In this method, the p-boxes based on the staircase density functions are propagated using the local NISS method in order to significantly suppress the computational cost to derive the NISS estimators over their hyper-parameters by performing the parameterizations of the staircase density functions only at a single well-chosen point (e.g., the mid-point of the support domain in this study) of the hyper-parameters. On the contrary, this method maintains to utilize the global NISS method to propagate the interval models in order to ensure the global accuracy of the estimators of the corresponding component functions. As such, the proposed method can achieve a good balance between the efficiency in deriving the NISS estimators for

the hyper-parameters of the p-boxes and the global accuracy of those for the interval parameters. The proposed method is applied for solving the reliability analysis subproblem (Sub-problem C) in the NASA UQ challenge 2019. The results demonstrate that the proposed method is capable to estimate the failure probability bounds accurately when compared to the results in Chapter 4 based on the probability bounds analysis by the double-loop Monte Carlo method, and the computational burden is reduced ten times.

# A distributionally robust approach for mixed aleatory and epistemic uncertainties propagation

Masaru Kitahara [a, *], Jingwen Song [b], Pengfei Wei [c], Matteo Broggi [a], Michael Beer [a, d, e]

[a] *Leibniz Universität Hannover, Institute for Risk and Reliability, Callinstrasse 34, Hannover, Germany*

[b] *Northwestern Polytechnical University, State IJR center of Aerospace Design and Additive Manufacturing, Xi'an 710072, China*

[c] *Northwestern Polytechnical University, School of Mechanics, Civil Engineering and Architecture, Xi'an 710072, China*

[d] *University of Liverpool, Institute for Risk and Uncertainty, Peach Street L69 7ZF, Liverpool, United Kingdom*

[e] *Tongji University, International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Shanghai 200092, China*

\* Correspondence author. E-mail address: masaru.kitahara@irz.uni-hannover.de (M. Kitahara).

## 6.1    Introduction

Uncertainties are typically grouped into aleatory and epistemic uncertainty (Der Kiureghian and Ditlevsen, 2009), and uncertainty characterization models can be then categorized into the following three groups:

- Probability model is the most classical one, and is usually used to represent aleatory uncertainty;

- Non-probabilistic models (Faes and Moens, 2020) are set-theoretical models, and are usually used to characterize epistemic uncertainty;

- Imprecise probability models (Beer et al., 2013) are accounted for as a combination of the former two models, and can separately characterize aleatory and epistemic uncertainties.

Among the above three models, the effective propagation of the imprecise probability models has been intensively investigated in the past decades. The extended Monte Carlo simulation (EMCS) (Wei et al., 2014) is an importance sampling-based method relying on a single stochastic simulation. Therefore, its computational cost is the same as that for the conventional reliability analysis. Moreover, the method has been integrated with the high-dimensional model representation (HDMR) decomposition as the metamodel strategy and sensitivity analysis to measure the importance of the epistemic parameters, to establish a general methodology framework, called non-intrusive imprecise stochastic simulation (NISS) (Wei et al., 2019a; Wei et al., 2019b), and it has been also generalized to propagate the imprecise

probability models and non-probabilistic models simultaneously (Song et al., 2019).

However, the main drawback of the NISS method is that it is restricted to the parameterized imprecise probability models, e.g., distributional p-boxes (Faes et al., 2021), that impose constraints on admissible distribution functions by assuming a specific distribution family. Comparatively, if the distribution families of the aleatory parameters cannot be determined a priori, it becomes necessary to propagate all the possible distributions of arbitral distribution families enclosed within a concerned p-box, so as to accurately estimate the failure probability bounds. Crespo et al. (2018) has recently developed a novel distribution family, called staircase distribution, that enables to approximate a broad range of distributions arbitrary close. Whereas its applications in imprecise stochastic simulation are quite limited in current literatures, it has a potential to define a parametric p-box approximately containing any distributions within its bounds. The aim of this work is consequently to generalize the staircase distribution-based p-boxes and integrate them with the NISS method to develop a novel framework to propagate the imprecise probability models without limiting hypotheses on the distribution family.

The present work particularly focuses on the generalized global NISS method (Song et al., 2019), because it can propagate the imprecise probability models and non-probabilistic models at the same time. The staircase distributions are theoretically ready to be utilized in this method by constructing parametric p-boxes defining their hyper-parameters as interval values. However, it requires to parameterize the distribution function for a significant number of the hyper-parameters sets to derive NISS estimators. This can be computationally prohibitive for the staircase distributions whose density functions are parameterized by solving optimization problems. To overcome this obstacle, a novel hybrid NISS method is proposed, where the staircase distribution-based p-boxes are propagated by the local NISS method (Wei et al., 2019a) while the non-probabilistic models, i.e., interval models, are propagated using the global NISS method (Wei et al., 2019b). The feasibility of the proposed hybrid NISS method is demonstrated by solving the reliability analysis subproblem of the NASA UQ challenge problem 2019 (Crespo and Kenny, 2021).

## 6.2    Parametric p-boxes bases on staircase distributions

Staircase distributions (Crespo et al., 2018) are functions of their hyper-parameters $\boldsymbol{\theta} = [\mu, m_2, \widetilde{m}_3, \widetilde{m}_4]$ which consists of the mean $\mu$, variance $m_2$, skewness $\widetilde{m}_3$, and kurtosis $\widetilde{m}_4$. The PDF of a staircase random variable $x$ on its support domain $[\underline{x}, \overline{x}]$ can be expressed as:

$$f_{\mathrm{x}}(x) = \begin{cases} l_i & \forall x \in (x^i, x^{i+1}], \text{for } 1 \le i \le n_b \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

where $l_i(\geq 0)$ indicates the PDF value of the $i$th bin; $x^i = \underline{x} + (i-1)\kappa$, with the length $\kappa = (\overline{x} - \underline{x})/n_b$, denotes the $i$th left partitioning point; $n_b$ is the number of bins. The PDF values $l_i$, for $1 \leq i \leq n_b$, can be obtained by solving an optimization problem based on the moment matching constraints, and the readers can refer to Crespo et al. (2018) for their detailed derivation. The staircase distribution can define a parametric p-box by CDF families the hyper-parameters of which are known in intervals:

$$F_{\mathrm{x}}(x) = F_{\mathrm{x}}(x|\boldsymbol{\theta}), \text{for } \boldsymbol{\theta} \in \{D_{\boldsymbol{\theta}} \cap \Theta\} \tag{6.2}$$

where $D_{\boldsymbol{\theta}}$ is the interval domain of $\boldsymbol{\theta}$; $\Theta$ denotes the feasible domain of $\boldsymbol{\theta}$ as moment constraints for the existence of $x$ [9]. Without loss of generality, we assume that each interval is independent and $D_{\boldsymbol{\theta}} = \left[\underline{\mu}, \overline{\mu}\right] \times \left[\underline{m_2}, \overline{m_2}\right] \times \left[\underline{\widetilde{m}_3}, \overline{\widetilde{m}_3}\right] \times \left[\underline{\widetilde{m}_4}, \overline{\widetilde{m}_4}\right]$ denotes a hyper-rectangular domain.

Figure 6.1 illustrates an example of a parametric p-box consisting of a staircase distribution family with a support set $x \in [-5, 5]$, mean $\mu \in [-1, 1]$, variance $m_2 \in [0.8, 1.2]$, skewness $\widetilde{m}_3 \in [-0.75, 0.75]$, and kurtosis $\widetilde{m}_4 \in [2, 4]$, as well as a parametric p-box that consists of a Gaussian distribution family with the same intervals for the mean and variance as above. Moreover, four possible CDF realizations for each type of the p-box are shown in the figure. The Gaussian distribution-based p-box naturally only contains Gaussian distributions, while the staircase distribution-based p-box contains a broad range of distributions, including skewed and bi-modal distributions.
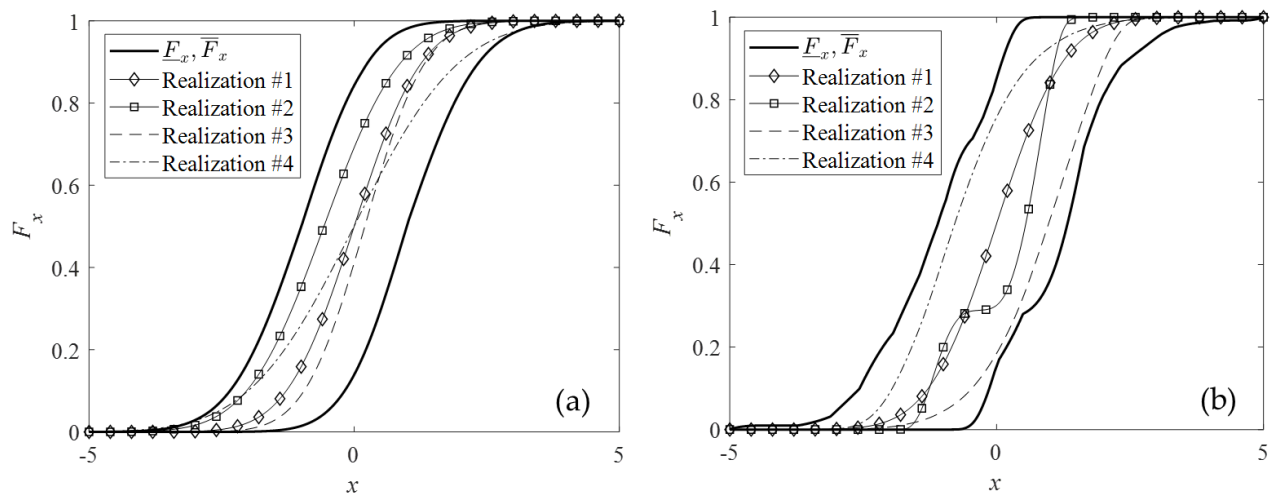


Figure 6.1 Illustration of Gaussian distribution-based and staircase distribution-based p-boxes.

The staircase distribution-based p-box is capable of realizing arbitral distribution functions the hyper-parameters of which are in $\{D_{\boldsymbol{\theta}} \cap \Theta\}$, while it allows a clear separation of aleatory uncertainty, represented by distribution families, and epistemic uncertainty, described by given intervals of the hyper-parameters. These properties fulfill the expectation

as a parameterized imprecise probability model with no limiting hypothesis about the distributions of the investigated aleatory parameters.

## 6.3    Hybrid NISS method

Suppose $g(x, y)$ be the limit state function, where $x = (x_1, x_2, \cdots, x_n)^T$ is the $n$-dimensional staircase random variables and $y = (y_1, y_2, \cdots, y_m)^T \in D_y$ indicates the $m$-dimensional independent interval parameters with the hyper-rectangular domain $D_y$. Without loss of generality, we assume that $x$ are independent to each other, so that the joint PDF is expressed as $f_x(x) = \prod_{i=1}^n f_{x_i}(x_i | \mu_i, m_{2i}, \widetilde{m}_{3i}, \widetilde{m}_{4i})$, for $\mu \in D_\mu$, $m_2 \in D_{m_2}$, $\widetilde{m}_3 \in D_{\widetilde{m}_3}$, and $\widetilde{m}_4 \in D_{\widetilde{m}_4}$, where $\mu, m_2, \widetilde{m}_3$ and $\widetilde{m}_4$ are columns of the means, variances, skewnesses and kurtoses, respectively; $D_\mu$, $D_{m_2}$, $D_{\widetilde{m}_3}$ and $D_{\widetilde{m}_4}$ mean the hyper-rectangular domains of $\mu, m_2, \widetilde{m}_3$ and $\widetilde{m}_4$. Noted that, the independence assumption on $x$ is not crucial for the proposed method. In fact, the dependence structure among the staircase random variables enables to be uniquely defined by a copula function (Nelsen, 2006)), thus the following steps to derive the NISS estimators do almost not affected by the presence of dependent inputs. Such definition further brings $4n$-dimensional epistemic parameters $\vartheta = (\mu_1, \cdots, \mu_n, m_{21}, \cdots, m_{2n}, \widetilde{m}_{31}, \cdots, \widetilde{m}_{3n}, \widetilde{m}_{41}, \cdots, \widetilde{m}_{4n})^T$, and their support set is defined as the hyper-rectangular $D_\vartheta = D_\mu \times D_{m_2} \times D_{\widetilde{m}_3} \times D_{\widetilde{m}_4}$. For convenience in notation, let $\vartheta = (\vartheta_1, \cdots, \vartheta_{4n})^T$, where $\vartheta_i = \mu_i$, $\vartheta_{2i} = m_{2i}$, $\vartheta_{3i} = \widetilde{m}_{3i}$, and , $\vartheta_{4i} = \widetilde{m}_{4i}$, for $i = 1, \cdots, n$.

We assume that the failure event happens when $g(x, y) < 0$ and the failure domain can be represented as $F = \{x, y : g(x, y) < 0\}$. The indicator function of $F$ is then formulated by $I_F(x, y) = 1$ if $\{x, y\} \in F$, and else $I_F(x, y) = 0$. The failure probability function can be then expressed as:

$$P_f(\vartheta, y) = \int_{\mathbb{R}^n} I_F(x, y) f_x(x | \vartheta) dx \tag{6.3}$$

The HDMR decomposition of the failure probability function expresses $P_f(\vartheta, y)$ as the sum of a series of component functions:

$$P_f(\vartheta, y) \approx P_{f0} + \sum_{i=1}^{4n} P_{f\vartheta_i}(\vartheta_i) + \sum_{i=1}^{m} P_{fy_i}(y_i) + \sum_{i=1}^{4n-1} \sum_{j=i+1}^{4n} P_{f\vartheta_{ij}}(\vartheta_{ij}) + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} P_{fy_{ij}}(y_{ij})$$

$$+ \sum_{i=1}^{4n} \sum_{j=1}^{m} P_{f\vartheta_i y_j}(\vartheta_i, y_j) + \cdots + P_{f\vartheta y}(\vartheta, y) \tag{6.4}$$

where $P_{f0}$ denotes the constant component; $P_{f\vartheta_i}$ and $P_{fy_i}$ refer to the first-order component functions; $P_{f\vartheta_{ij}}$, $P_{fy_{ij}}$, and

$P_{f\vartheta_iy_j}$ mean the second-order component functions; $\boldsymbol{\vartheta}_{ij}$ is the two-dimensional vector consisting of $\vartheta_i$ and $\vartheta_j$, and $\boldsymbol{y}_{ij}$

possesses a similar structure for $\boldsymbol{y}$. It has been demonstrated the HDMR decomposition with second-order truncation

commonly results in a satisfactory approximation of the failure probability function (Wei et al., 2019b; Song et al.,

2019). Thus, we only consider up to second-order component functions and do not go for the higher-order component

functions in the rest part of this paper.

A hybrid NISS method is herein developed, where the staircase distribution-based p-boxes are propagated by the

local NISS method (Wei et al., 2019b) to significantly suppress the computational burden for estimating the component

functions over the hyper-parameters, by performing the parameterizations of the joint PDFs at a single well-chosen

point of $\boldsymbol{\vartheta}$. On the other hand, the interval models are propagated using the global NISS method (Song et al., 2019) for

ensuring the global accuracy of the estimators of the corresponding component functions. In this context, the HDMR

component functions can be defined as:

$$
\begin{cases}
P_{f0} = \int P_f(\boldsymbol{\vartheta}^*, \boldsymbol{y})f_{\boldsymbol{y}}(\boldsymbol{y})d\boldsymbol{y} \\
P_{f\vartheta_i}(\vartheta_i) = \int P_f(\vartheta_i, \boldsymbol{\vartheta}^*_{-i}, \boldsymbol{y})f_{\boldsymbol{y}}(\boldsymbol{y})d\boldsymbol{y} - P_{f0} \\
P_{fy_i}(y_i) = \int P_f(\boldsymbol{\vartheta}^*, \boldsymbol{y})f_{\boldsymbol{y}_{-i}}(\boldsymbol{y}_{-i})d\boldsymbol{y}_{-i} - P_{f0} \\
P_{f\vartheta_{ij}}(\boldsymbol{\vartheta}_{ij}) = \int P_f(\boldsymbol{\vartheta}_{ij}, \boldsymbol{\vartheta}^*_{-ij}, \boldsymbol{y})f_{\boldsymbol{y}}(\boldsymbol{y})d\boldsymbol{\vartheta}_{-ij}d\boldsymbol{y} - P_{f\vartheta_i}(\vartheta_i) - P_{f\vartheta_j}(\vartheta_j) - P_{f0} \\
P_{fy_{ij}}(\boldsymbol{y}_{ij}) = \int P_f(\boldsymbol{\vartheta}^*, \boldsymbol{y})f_{\boldsymbol{y}_{-ij}}(\boldsymbol{y}_{-ij})d\boldsymbol{y}_{-ij} - P_{fy_i}(y_i) - P_{fy_j}(y_j) - P_{f0} \\
P_{f\vartheta_iy_j}(\vartheta_i, y_j) = \int P_f(\vartheta_i, \boldsymbol{\vartheta}^*_{-i}, \boldsymbol{y})f_{\boldsymbol{y}_{-j}}(\boldsymbol{y}_{-j})d\boldsymbol{y}_{-j} - P_{f\vartheta_i}(\vartheta_i) - P_{fy_j}(y_j) - P_{f0}
\end{cases}
\tag{6.5}
$$

where $\boldsymbol{\vartheta}^*$ indicates the well-chosen fixed point of $\boldsymbol{\vartheta}$ chosen as the mid-point of $D_{\boldsymbol{\vartheta}}$ in this study. Unbiased estimators

of each component function in Equation (6.5) are then derived using the joint sample set $W = \{\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}\}$, for $k =$

$1,2,\cdots,N$, as:

$$
\begin{cases}
\hat{P}_{f0} = \frac{1}{N}\sum_{k=1}^{N} I_F(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}) \\
\hat{P}_{f\vartheta_i}(\vartheta_i) = \hat{P}_{f0}r_{\vartheta_i}(\boldsymbol{x}^{(k)}|\vartheta_i, \boldsymbol{\vartheta}^*_{-i}) \\
\hat{P}_{fy_i}(y_i) = \hat{P}_{f0}r_{y_i}(y_i|F, W) \\
\hat{P}_{f\vartheta_{ij}}(\boldsymbol{\vartheta}_{ij}) = \hat{P}_{f0}r_{\vartheta_{ij}}(\boldsymbol{x}^{(k)}|\boldsymbol{\vartheta}_{ij}, \boldsymbol{\vartheta}^*_{-ij}) \\
\hat{P}_{fy_{ij}}(\boldsymbol{y}_{ij}) = \hat{P}_{f0}r_{y_{ij}}(\boldsymbol{y}_{ij}|F, W) \\
\hat{P}_{f\vartheta_iy_j}(\vartheta_i, y_j) = [\hat{P}_{f\vartheta_i}(\vartheta_i) + \hat{P}_{f0}]r_{y_j}(y_j|F, W) - \hat{P}_{fy_j}(y_j)
\end{cases}
\tag{6.6}
$$

with

$$
\begin{cases}
r_{\vartheta_i}\left(\boldsymbol{x}^{(k)}\middle|\vartheta_i,\boldsymbol{\vartheta}_{-i}^*\right)=\dfrac{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\vartheta_i,\boldsymbol{\vartheta}_{-i}^*\right)}{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\boldsymbol{\vartheta}^*\right)}-1 \\[2ex]
r_{y_i}\left(y_i\middle|F,W'\right)=\dfrac{\hat{f}_{\mathbf{y}_i}\left(y_i\middle|F,W\right)}{f_{\mathbf{y}_i}\left(y_i\right)}-1 \\[2ex]
r_{\vartheta_{ij}}\left(\boldsymbol{x}^{(k)}\middle|\boldsymbol{\vartheta}_{ij},\boldsymbol{\vartheta}_{-ij}^*\right)=\dfrac{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\boldsymbol{\vartheta}_{ij},\boldsymbol{\vartheta}_{-ij}^*\right)}{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\boldsymbol{\vartheta}^*\right)}-\dfrac{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\vartheta_i,\boldsymbol{\vartheta}_{-i}^*\right)}{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\boldsymbol{\vartheta}^*\right)}-\dfrac{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\vartheta_j,\boldsymbol{\vartheta}_{-j}^*\right)}{f_{\mathbf{x}}\left(\boldsymbol{x}^{(k)}\middle|\boldsymbol{\vartheta}^*\right)}+1 \\[2ex]
r_{\mathbf{y}_{ij}}\left(\boldsymbol{y}_{ij}\middle|F,W'\right)=\dfrac{\hat{f}_{\mathbf{y}_{ij}}\left(\boldsymbol{y}_{ij}\middle|F,W\right)}{f_{\mathbf{y}_{ij}}\left(\boldsymbol{y}_{ij}\right)}-\dfrac{\hat{f}_{\mathbf{y}_i}\left(y_i\middle|F,W\right)}{f_{\mathbf{y}_i}\left(y_i\right)}-\dfrac{\hat{f}_{\mathbf{y}_j}\left(y_j\middle|F,W\right)}{f_{\mathbf{y}_j}\left(y_j\right)}+1
\end{cases}
\tag{6.7}
$$

are regarded as weight coefficients, where $\hat{f}_{\mathbf{y}_i}(y_i|F,W)$ and $\hat{f}_{\mathbf{y}_{ij}}(\boldsymbol{y}_{ij}|F,W)$ denote the conditional PDFs of $y_i$ and $\boldsymbol{y}_{ij}$, respectively, on the failure domain $F$ estimated based on the sample set $W$. One can refer to Song et al. (2019) for the detailed derivations of these conditional PDFs. It is noted that, to generate the joint sample set $W$, an auxiliary PDF of $\boldsymbol{y}$, $f_{\mathbf{y}}(\boldsymbol{y}) = \prod_{i=1}^m f_{\mathbf{y}_i}(y_i)$, is necessary. Without loss of generality, we assume that each $y_i$ follows a uniform distribution on its relaxed interval domain $\left[\underline{y}_i - \delta\Delta y_i, \overline{y}_i + \delta\Delta y_i\right]$, where $\Delta y_i$ denotes the difference of the original interval and $\delta$ is a given value (e.g., $\delta = 0.2$), to improve the estimation performance around the original bounds (Song et al., 2019). Finally, sensitivity indices are proposed as follows for measuring the relative importance of the component functions:

$$
S_{(\cdot)} = \frac{\mathrm{V}\left[P_{f(\cdot)}(\cdot)\right]}{\mathrm{V}\left[P_f(\boldsymbol{\vartheta},\boldsymbol{y})\right]}
\tag{6.8}
$$

with

$$
\mathrm{V}\left[P_f(\boldsymbol{\vartheta},\boldsymbol{y})\right] = \sum_{i=1}^{4n}\mathrm{V}\left[P_{f\vartheta_i}(\vartheta_i)\right] + \sum_{i=1}^{m}\mathrm{V}\left[P_{fy_i}(y_i)\right] + \sum_{i=1}^{4n-1}\sum_{j=i+1}^{2n}\mathrm{V}\left[P_{f\vartheta_{ij}}(\boldsymbol{\vartheta}_{ij})\right] + \sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\mathrm{V}\left[P_{fy_{ij}}(\boldsymbol{y}_{ij})\right]
$$

$$
+ \sum_{i=1}^{4n}\sum_{j=1}^{m}\mathrm{V}\left[P_{f\vartheta_i y_j}(\vartheta_i,y_j)\right] + \cdots + \mathrm{V}\left[P_{f\vartheta y}(\boldsymbol{\vartheta},\boldsymbol{y})\right]
$$

where V indicates the variance operator. The sensitivity indices measure the average $\mathrm{L}^2$ distance of the components to the fixed point $\boldsymbol{\vartheta}^*$. The smaller the distance is, the less influential the component is.

The detailed procedure of the proposed method is shown in Figure 6.2. The statistical error of the NISS estimators is assessed by the bootstrap scheme. Let $n_{boot}$ indicate the number of total bootstrap replications, we can obtain $n_{boot}$ estimates of each component function and sensitivity index, and can calculate the confidence intervals $\left[\underline{\hat{P}}_{f(\cdot)}, \overline{\hat{P}}_{f(\cdot)}\right]$, e.g., $\left[\mathrm{E}\left[\hat{P}_{f(\cdot)}\right] - 2\left(\mathrm{V}\left[\hat{P}_{f(\cdot)}\right]\right)^{1/2}, \mathrm{E}\left[\hat{P}_{f(\cdot)}\right] + 2\left(\mathrm{V}\left[\hat{P}_{f(\cdot)}\right]\right)^{1/2}\right]$, where E is the mean operator, since the NISS estimators

follow Gaussian distributions. We propose to estimate two coefficients of variations (CV), i.e., CVs at the points where $\hat{\underline{P}}_{f(\cdot)}$ returns the minimum and maximum values, $\text{CV}_{\min(\hat{\underline{P}}_{f(\cdot)})}$ and $\text{CV}_{\max(\overline{\hat{P}}_{f(\cdot)})}$. If their larger value is less than a given tolerance $\varepsilon$, the statistical error is acceptable, and if not, one should enrich the size of the joint sample set $N$. The truncation error on the other hand is quantified by the sensitivity indices. The components sensitivity indices of which are less than a threshold $S_{\text{thr}}$ are ignored, and the resultant truncation error can be accepted if the summation of the sensitivity indices for all the components used is larger than a given threshold $\epsilon$. Otherwise, one should decrease $S_{\text{thr}}$. Finally, the failure probability function $P_f(\boldsymbol{\vartheta}, \boldsymbol{y})$ is approximated as synthetic of all the influential component functions. One can also estimate the failure probability bounds by sampling methods, where not only the mean estimators but also the variance estimators can be evaluated within the bootstrap scheme.
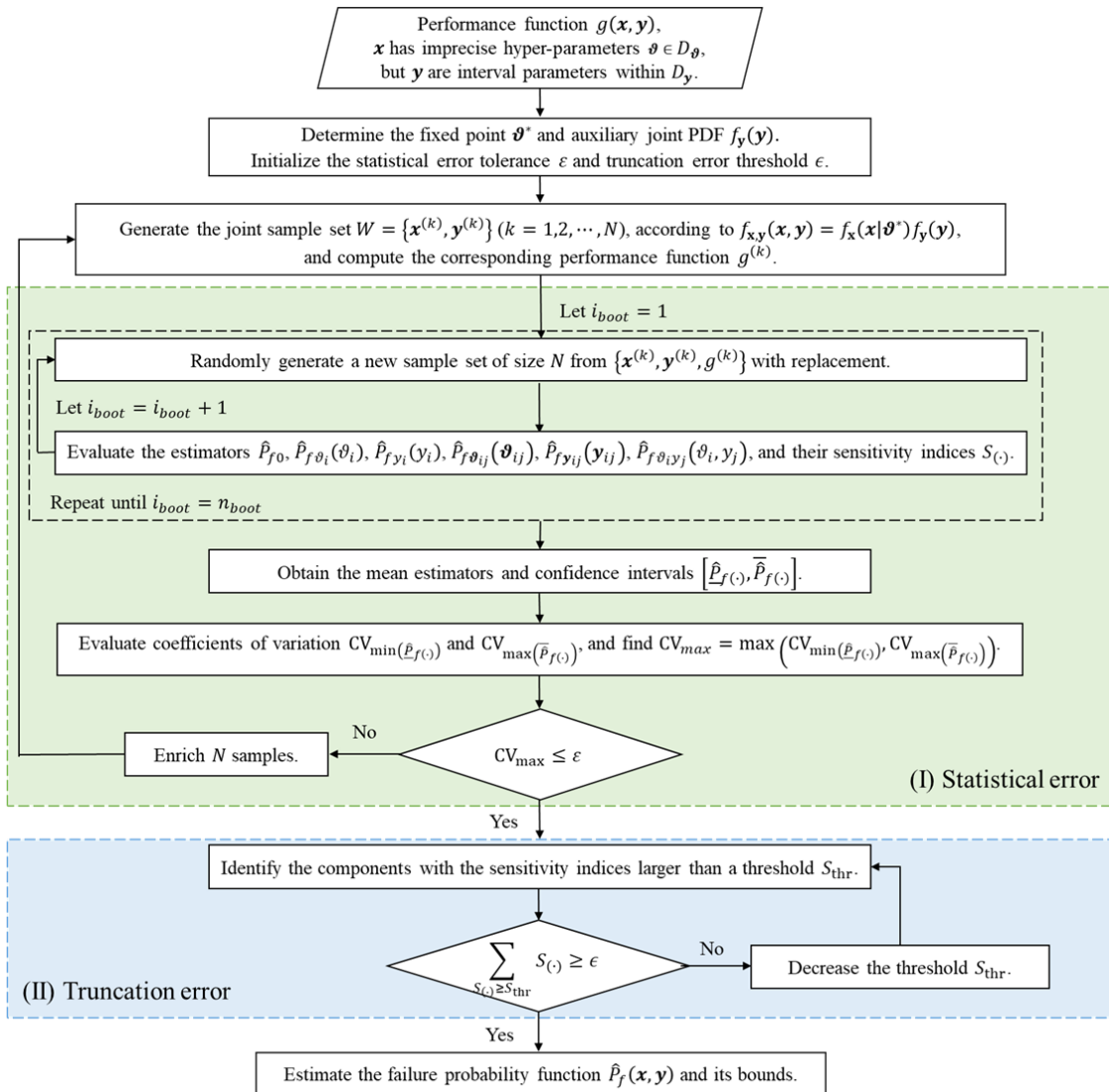


Figure 6.2 Flowchart of the hybrid NISS method.

## 6.4    NASA UQ challenge problem 2019

The NASA UQ challenge problem 2019 (Crespo and Kenny, 2021) is investigated to demonstrate the capabilities of the proposed hybrid NISS method. Figure 6.3 shows the overall structure of Sub-problem C (Reliability analysis of baseline design). The model inputs consist of five aleatory parameters $\boldsymbol{a} = (a_1, a_2, \cdots, a_5)^T$, four epistemic parameters $\boldsymbol{e} = (e_1, e_2, \cdots, e_4)^T$, and pre-specified design variable $\theta_{baseline}$ with nine components. The aleatory parameters $\boldsymbol{a}$ are modeled by p-boxes while the epistemic parameters $\boldsymbol{e}$ are modeled by intervals, based on the results of the first two subtasks, i.e., Sub-problem A (Model calibration & UQ of the subsystem) and Sub-problem B (Uncertainty reduction). It is important to note that, the distribution families of each aleatory parameter are completely unknown a priori. The reliability requirements of interest are represented by following three limit state functions, i.e., a black-box function $g_1(\boldsymbol{a}, \boldsymbol{e}, \theta_{baseline}) < 0$,

$$g_2 = \max_{t \in [T/2, T]} |z_1(\boldsymbol{a}, \boldsymbol{e}, \theta_{baseline}, t)| - 0.02 < 0 \tag{6.9}$$

with a black-box time-independent output $z_1$, and

$$g_3 = \max_{t \in [0, T]} |z_2(\boldsymbol{a}, \boldsymbol{e}, \theta_{baseline}, t)| - 4 < 0 \tag{6.10}$$

with a black-box time-independent output $z_2$. The worst-case limit state function is then defined as:

$$\omega(\boldsymbol{a}, \boldsymbol{e}, \theta_{baseline}) = \max_{i=1,2,3} g_i(\boldsymbol{a}, \boldsymbol{e}, \theta_{baseline}) \tag{6.11}$$

The safe domain of the system is determined by the $\boldsymbol{a}$ points where $\omega(\boldsymbol{a}, \boldsymbol{e}, \theta_{baseline}) < 0$, while its complement set is accounted for as the failure domain.

Some of the authors have addressed the first two subproblems and have represented the aleatory parameters by the staircase distribution-based p-boxes (Lye et al., 2022). We herein use the results in Lye et al. (2022) for uncertainty characterization of $\boldsymbol{a}$ and $\boldsymbol{e}$, as summarized in Table 5.1. Under this assumption, 20 hyper-parameters of the staircase distributions are additionally taken into account as epistemic parameters, and thus totally 24 epistemic parameters (i.e., $\vartheta_i = \{\mu_i, m_{2i}, \tilde{m}_{3i}, \tilde{m}_{4i}\}$, for $i = 1, \cdots, 5$, and $\boldsymbol{e}$) are considered in the reliability analysis. Each auxiliary PDF $f_{e_i}(e_i)$ is assumed as a uniform distribution on its relaxed intervals as shown in parentheses after the true intervals in Table 5.1. The parameters of the proposed hybrid NISS method are set as $N = 5 \times 10^5$ and $n_{boot} = 20$, $\varepsilon = 0.15$ and, $\epsilon = 0.9$, respectively.
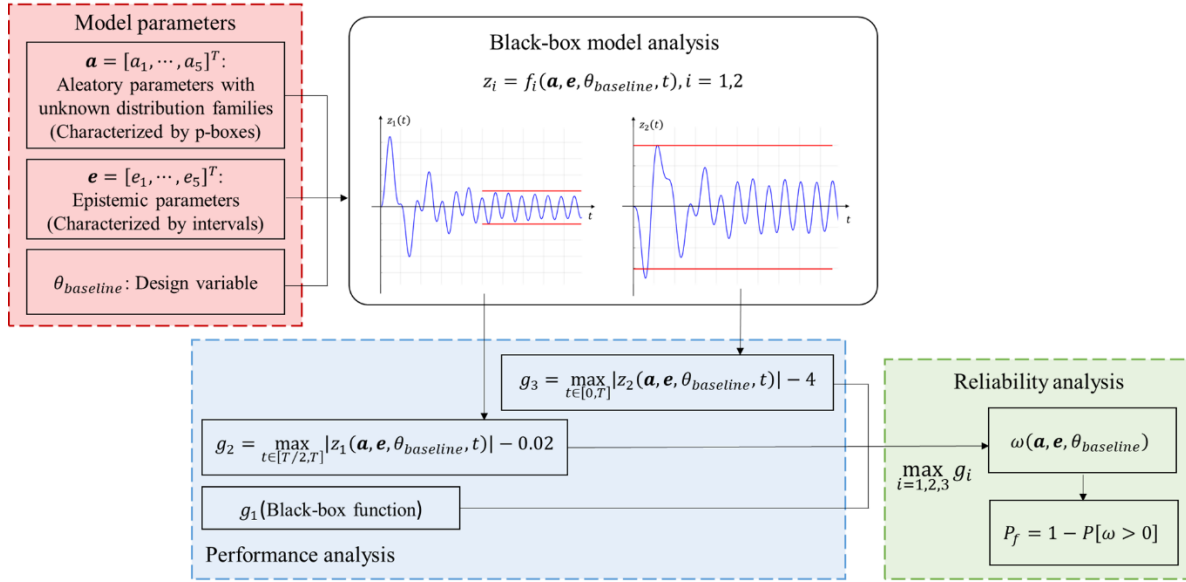
Figure 6.3 Schematic of the NASA UQ challenge 2019 Sub-problem C.

Table 6.1 Model parameters of the NASA UQ challenge problem.

| Parameter | Uncertainty characteristic |
|---|---|
| $a_1$ | Staircase distribution, $\mu_1 \in [0.5050, 0.5980]$, $m_{21} \in [0.0200, 0.0750]$, $\widetilde{m}_{31} \in [0.9800, 1.4550]$, $\widetilde{m}_{41} \in [4.0790, 6.3690]$ |
| $a_2$ | Staircase distribution, $\mu_1 \in [1.1110, 1.2290]$, $m_{21} \in [0.0660, 0.0670]$, $\widetilde{m}_{31} \in [-0.6640, -0.2440]$, $\widetilde{m}_{41} \in [3.7760, 4.9680]$ |
| $a_3$ | Staircase distribution, $\mu_1 \in [0.8040, 0.8720]$, $m_{21} \in [0.0300, 0.0440]$, $\widetilde{m}_{31} \in [-0.9620, -0.6080]$, $\widetilde{m}_{41} \in [3.7140, 3.7150]$ |
| $a_4$ | Staircase distribution, $\mu_1 \in [0.7870, 1.2050]$, $m_{21} \in [0.3520, 0.3530]$, $\widetilde{m}_{31} \in [-0.7430, 0.2340]$, $\widetilde{m}_{41} \in [1.4030, 2.5000]$ |
| $a_5$ | Staircase distribution, $\mu_1 \in [0.8510, 1.2240]$, $m_{21} \in [0.2390, 0.3690]$, $\widetilde{m}_{31} \in [-0.5430, 0.4370]$, $\widetilde{m}_{41} \in [1.3040, 3.0480]$ |
| $e_1$ | Interval, $e_1 \in [0.4674, 0.6433]$ (Relaxed interval $[0.2674, 0.8433]$) |
| $e_2$ | Interval, $e_2 \in [0.7607, 0.9736]$ (Relaxed interval $[0.5607, 1.1736]$) |
| $e_3$ | Interval, $e_3 \in [0.2865, 0.4583]$ (Relaxed interval $[0.0865, 0.6583]$) |
| $e_4$ | Interval, $e_4 \in [0.9627, 1.1664]$ (Relaxed interval $[0.7627, 1.3664]$) |

The mean value and standard deviation of the constant component estimator $\hat{P}_{f0}$ are evaluated as 0.1646 and $4.4 \times 10^{-4}$, respectively. The mean estimates of the first-order components and their 95.45 % confidence intervals are shown in Figure 6.4 for the hyper-parameters of the aleatory parameters $\boldsymbol{a}$ and in Figure 6.5 for the interval parameters $\boldsymbol{e}$. It can be seen that the confidence intervals of each component function are narrow enough, indicating all the 24 component functions are accurately estimated by the proposed method. Similarly, among the second-order component functions, the mean estimators of the three most influential component functions are shown in Figure 6.6, along with their 95.45 % confidence intervals. As can be seen, these three second-order component functions are also effectively estimated with narrow confidence intervals.
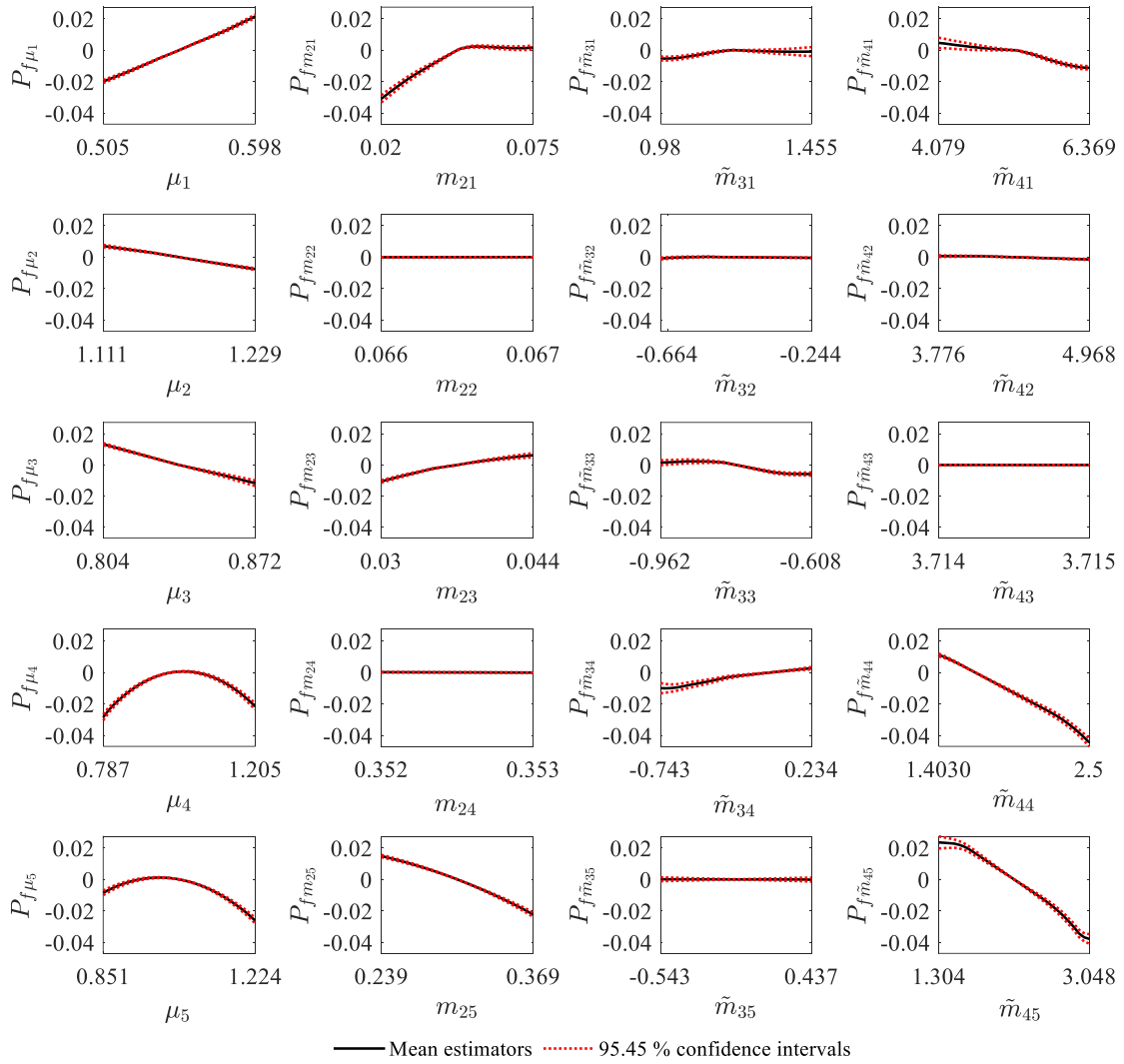
Figure 6.4 First-order component functions of the hyper-parameters of $\boldsymbol{a}$.

Table 6.2 Sensitivity indices of the NASA UQ challenge problem.

| | Mean estimate | Standard deviation | | Mean estimate | Standard deviation |
|---|---|---|---|---|---|
| $S_{\mu_1}$ | 0.0107 | $2.3 \times 10^{-4}$ | $S_{\tilde m_{34}}$ | 0.0018 | $3.0 \times 10^{-4}$ |
| $S_{m_{21}}$ | 0.0121 | $3.2 \times 10^{-4}$ | $S_{\tilde m_{44}}$ | 0.0331 | $1.2 \times 10^{-3}$ |
| $S_{\tilde m_{31}}$ | 0.0006 | $4.9 \times 10^{-5}$ | $S_{\mu_5}$ | 0.0062 | $1.9 \times 10^{-4}$ |
| $S_{\tilde m_{41}}$ | 0.0027 | $1.6 \times 10^{-4}$ | $S_{m_{25}}$ | 0.0087 | $2.3 \times 10^{-4}$ |
| $S_{\mu_2}$ | 0.0017 | $7.3 \times 10^{-5}$ | $S_{\tilde m_{35}}$ | 0.0000 | $1.8 \times 10^{-6}$ |
| $S_{m_{22}}$ | 0.0000 | $9.3 \times 10^{-9}$ | $S_{\tilde m_{45}}$ | 0.0273 | $1.0 \times 10^{-3}$ |
| $S_{\tilde m_{32}}$ | 0.0000 | $2.2 \times 10^{-6}$ | $S_{e_1}$ | 0.2202 | $3.8 \times 10^{-3}$ |
| $S_{\tilde m_{42}}$ | 0.0001 | $6.1 \times 10^{-6}$ | $S_{e_2}$ | 0.4112 | $4.8 \times 10^{-3}$ |
| $S_{\mu_3}$ | 0.0047 | $5.0 \times 10^{-4}$ | $S_{e_3}$ | 0.1044 | $2.0 \times 10^{-3}$ |
| $S_{m_{23}}$ | 0.0019 | $2.0 \times 10^{-4}$ | $S_{e_4}$ | 0.0213 | $5.7 \times 10^{-4}$ |
| $S_{\tilde m_{33}}$ | 0.0010 | $1.9 \times 10^{-4}$ | $S_{e_1 e_2}$ | 0.0415 | $5.9 \times 10^{-4}$ |
| $S_{\tilde m_{43}}$ | 0.0000 | $2.1 \times 10^{-10}$ | $S_{e_1 e_3}$ | 0.0105 | $2.9 \times 10^{-4}$ |
| $S_{\mu_4}$ | 0.0093 | $2.9 \times 10^{-4}$ | $S_{e_2 e_3}$ | 0.0197 | $5.4 \times 10^{-4}$ |
| $S_{m_{24}}$ | 0.0000 | $6.7 \times 10^{-9}$ | | | |
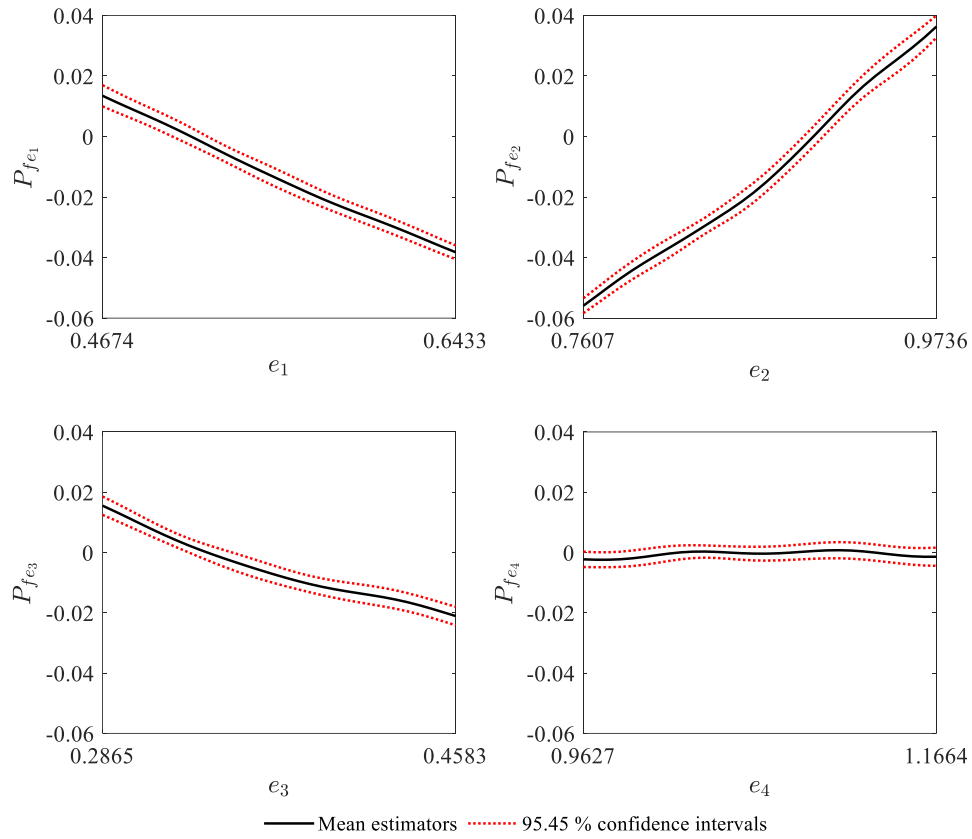
Figure 6.5 First-order component functions of the hyper-parameters of $\boldsymbol{a}$.
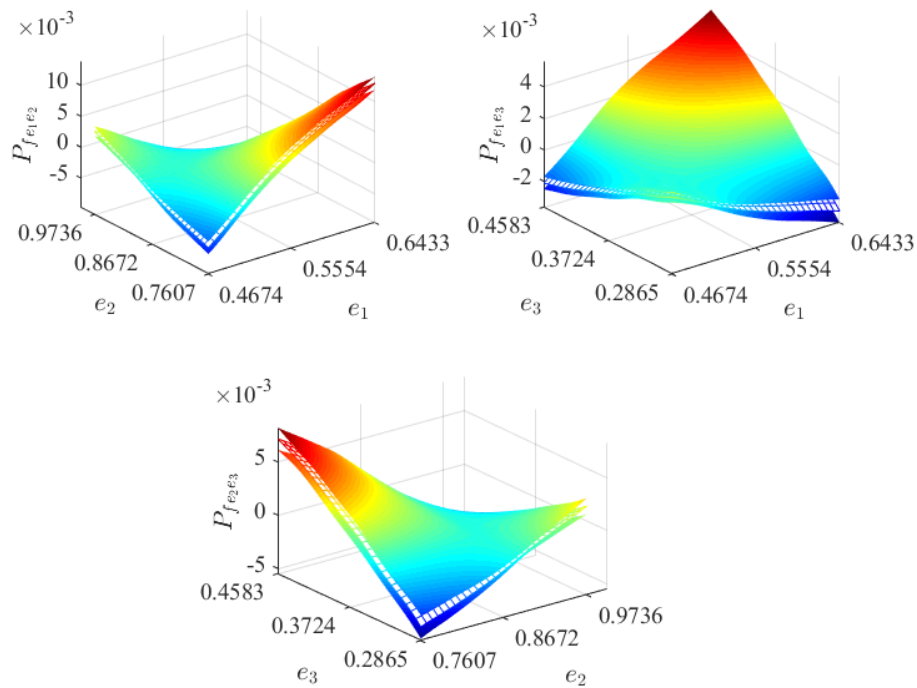


Figure 6.6 The three most influential second-order component functions, where the in-between surfaces indicate the mean estimators and the other two surfaces indicate the 95.45 % confidence intervals.

We compute the sensitivity indices for all the first- and second-order component functions. The mean estimates as well as standard deviations are presented in Table 5.2, for all the first-order and three most influential second-order component functions. It can be seen that all the sensitivity indices are accurately derived with small standard deviations. We assume that the component functions with the sensitivity indices larger than 0.01 are influential. Among the total 24 first-order component functions, the eight components $\hat{P}_{f\mu_1}$, $\hat{P}_{fm_{21}}$, $\hat{P}_{f\tilde{m}_{44}}$, $\hat{P}_{f\tilde{m}_{45}}$, $\hat{P}_{fe_1}$, $\hat{P}_{fe_2}$, $\hat{P}_{fe_3}$, and $\hat{P}_{fe_4}$, and the three most influential second-order component functions $\hat{P}_{fe_1e_2}$, $\hat{P}_{fe_1e_3}$, and $\hat{P}_{fe_2e_3}$ are thus employed. The summation of the sensitivity indices of all these 11 components is larger than the threshold $\epsilon$, implying the truncation error due to the truncation of the remaining components is acceptable. Finally, the failure probability function can be approximated using the 11 influential components, and the failure probability bounds are estimated. The mean estimates and standard deviations are listed in Table 5.3. The results are compared with the reference bounds in Lye et al. (2022) by the double-loop MCS (Rocchetta et al., 2018) based on the same parameter settings. It can be seen that both the upper and lower bounds show good agreement with the reference bounds. In addition, the total number of model evaluations of the hybrid NISS method is $N = 5 \times 10^5$, whereas that of the double-loop MCS is $5 \times 10^6$ (Lye et al., 2022). Hence, the hybrid NISS method is ten times more efficient than the double-loop MCS. These outcomes demonstrate the feasibility of the proposed method in the propagation of mixed aleatory and epistemic uncertainties for the case where distribution families of the aleatory parameters are unknown.

Table 6.3 Failure probability bounds of the NASA UQ challenge problem.

| Parameter | Double-loop MC in Lye et al. (2022) | Hybrid NISS method Mean estimate | Standard deviation |
|---|---|---|---|
| Lower bound of $P_f$ | 0.0270 | 0.0299 | 0.0024 |
| Upper bound of $P_f$ | 0.2746 | 0.2564 | 0.0030 |

## 6.5    Conclusions

This paper presents two contributions to effectively propagate the imprecise probability models without limiting hypotheses on the distribution family. First, the staircase distribution-based p-boxes are defined as a novel class of the parametric p-box. They are capable to explicitly account for the imprecision not only in the hyper-parameters but also in the distribution families. Thus, they are especially suitable to characterize the true-but-unknown CDFs of the random variables whose distribution families are unknown. Second, the novel hybrid NISS method is developed, in which the staircase distribution-based p-boxes are propagated by the locally expanded HDMR decomposition while the interval models are propagated based on the globally expanded HDMR decomposition. This method can achieve a good balance

between the efficiency in deriving the NISS estimators for the hyper-parameters of the p-boxes and the global accuracy of those for the interval parameters. The NASA UQ challenge 2019 has demonstrated the effectiveness of the proposed method.

## References

Beer, M., Ferson, S., and Kreinovich, V. (2013). Imprecise probabilities in engineering analysis. *Mechanical System and Signal Processing*, 37, 4-29.

Crespo, L. G., Kenny, S. P., Giesy, D. P., and Stanford, B. K. (2018). Random variables with moment-matching staircase density function. *Applied Mathematical Modelling*, 64, 196-213.

Crespo, L. G., and Kenny, S. P. (2021). The NASA Langley Challenge on Optimization Under Uncertainty. *Mechanical System and Signal Processing*, 152, 107405.

Der Kiureghian, A., and Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, 31, 105-112.

Faes, M., and Moens, D. (2020). Recent trends in the modeling and quantification of non-probabilistic uncertainty. *Archives of Computational Methods in Engineering*, 27, 633-671.

Faes, M., Daub, M., Marelli, S., Patelli, E., and Beer, M. (2021). Engineering analysis with probability boxes: A review on computational methods. *Structural Safety*, 93, 102092.

Lye, A., Kitahara, M., Broggi, M., and Patelli, E. (2022). Robust optimization of a dynamic system under severe uncertainty: A distribution-free framework. *Mechanical System and Signal Processing*, 167, 108522.

Nelsen, R. B. (2006). *An introduction to copulas*, Second edition, Springer-Verlag, New York.

Rocchetta, R., Broggi, M., and Patelli, E. (2018). Do we have enough data? Robust reliability via uncertainty quantification. *Applied Mathematical Modelling*, 54, 710-721.

Song, J., Wei, P., Valdebenito, M., Bi, S., Broggi, M., Beer, M., and Lu, Z. (2019). Generalization of non-intrusive imprecise stochastic simulation for mixed uncertain variables. *Mechanical Systems and Signal Processing*, 134, 106316.

Wei, P., Lu, Z., and Song, J. (2014). Extended Monte Carlo simulation for parametric global sensitivity analysis and optimization. *AIAA Journal*, 52, 867-878.

Wei, P., Song, J., Bi, S., Broggi, M., Beer, M., Yue, Z., and Lu, Z. (2019a). Non-intrusive stochastic analysis with parameterized imprecise probability models: I. Performance estimation. *Mechanical Systems and Signal Processing*, 124, 349-368.

Wei, P., Song, J., Bi, S., Broggi, M., Beer, M., Yue, Z., and Lu, Z. (2019b). Non-intrusive stochastic analysis with parameterized imprecise probability models: II. Reliability and rare events analysis. *Mechanical Systems and Signal Processing*, 126, 227-247.

# Chapter 7

# Conclusions and Prospects

## 7.1  Conclusions

The robust and efficient tools for uncertainty quantification (UQ) have been nowadays widely recognized as the necessary developments for quantitative characterization and/or reduction of non-determinism in both simulations and measurements, so as to derive the most realistic estimates of the behavior of structures. The probabilistic approach has been conventionally used for this purpose by considering non-determinism as the likelihood that a quantity of interest (e.g., inputs of the numerical model and responses of the actual structure) assumes a certain value within a given range. Nevertheless, this approach often involves a large number of subjective assumptions due to the scarcity, incompleteness, imprecision, vagueness, etc., of available information or data on the quantity of interest. To avoid including subjectivity, the concepts of imprecise probabilities have recently gain attentions, which combine the probabilistic and possibilistic approaches providing the bounds on the probabilities that model the epistemic uncertainty involved. The state-of-the-art developments in this field, including the Bhattacharyya distance-based ABC model updating framework to calibrate the parameters with mixed aleatory and epistemic uncertainties and the NISS framework to very efficiently propagate such parameters, have strengthened the subjective assumption-free UQ approach. Nevertheless, these frameworks rely on the proper hypotheses about the distribution families of the parameters; thus, their scope of application is still limited and further developments are necessary to enhance their robustness such that they are even applicable into cases where prior information on the parameters is extremely limited.

Within this context, this thesis has further strengthened the subjective assumption-free UQ approach by improving the Bhattacharyya distance-based ABC model updating framework and the NISS framework, and promoted the real-world engineering applications of those frameworks. Five main contributions have been made in the following journal articles respectively. The first article concerns the extension of the Bhattacharyya distance-based ABC model updating framework to update dynamic systems using observed time signals. Two key components, i.e., the dimension reduction procedure to define the novel Bhattacharyya distance-based UQ metric for discrete time signals and the novel efficient Bayesian inference algorithm for the combination of BUS with AK-MCMC to address the high computational burden

in each nonlinear dynamic analysis, have been developed. As has been shown, the novel Bhattacharyya distance-based metric is capable of quantifying the uncertainty characteristics of the overall observed time signals, whereby the high-dimensional time signals are degraded to a relatively low-dimensional series of scalar values. The proposed approach has been demonstrated on the updating of a seismic-isolated bridge pier model using observed seismic response data, where the results with the newly developed Bayesian inference algorithm for the combination of BUS with AK-MCMC are compared with those by TMCMC, showing that the proposed algorithm enables satisfactory updating results with much-reduced computational demand to be delivered.

The second article involves the development of a distribution-free stochastic model updating framework to update the parameters with mixed aleatory and epistemic uncertainties without limiting hypotheses on the distribution families of the parameters. The probability distributions of the parameters to be inferred are represented as the staircase density functions, which are capable of describing a broad range of distributions arbitrary close. Their hyper-parameters (i.e., the means, variances, skewnesses, and kurtoses) are considered as the inferred parameters, whose prior is defined based on the moment constraints, and are updated using the Bhattacharyya distance-based approximate likelihood. Moreover, an outer Bayesian updating by the Euclidian distance-based approximate likelihood is maintained as the preconditioner of the proposed updating procedure to avoid non-unique solutions. The proposed distribution-free stochastic updating framework has been first implemented into the well-known updating problem of the shear building model. The results demonstrate the capability of the proposed framework to calibrate the distributions of the parameters which cannot be defined analytically and to identify all the modes of the target distributions thanks to the outer Bayesian updating. Then, the model updating subproblem of the NASA UQ challenge problem 2014 has been correctly solved by the proposed updating procedure, wherein its problem setting is altered by ignoring information on the distribution families of the parameters, and the results demonstrate that the proposed procedure works well for this real-world engineering problem even under more challenging conditions than that of the original problem setting.

The developments in the above two articles have been then employed in the third article to provide a solution to the NASA UQ challenge problem 2019 that is designed to represent the difficulties which are often encountered in the design of critical safety systems under the availability of very limited data. Differently from the previous challenge in 2014, the system of interest is described by time-domain sequences and the prior knowledge on the aleatory parameters is extremely limited other than a common boundary. This challenge starts from the model updating subproblem and is solved with the distribution-free stochastic model updating framework developed in the second article by introducing the dimension reduction procedure in the first article into the framework. The updating results are compared with those

that use the distribution-based approach, wherein a specific distribution family is assumed for the aleatory parameters, showing that the subjective assumption-free approach is crucial in achieving more informative updating results. The followed subproblems of sensitivity analysis, reliability analysis, and reliability-based design optimization have been addressed based on the updating results by the proposed procedure. In particular, to solve the reliability-based design optimization subproblem with mixed aleatory and epistemic uncertainties, a NISS approach is proposed, that simplifies the high-dimensional optimization to a set of one-dimensional searches using a first-order HDMR decomposition with respect to each design parameter. This type of reliability-based optimization generally requires triple-loop approach to explore all the aleatory, epistemic, and design parameter spaces, whereas the proposed NISS approach only requires a few iterations of the single stochastic simulation and is thus significantly efficient.

The results of the third article have, at the same time, revealed room for further improvements with regards to the Bhattacharyya distance-based ABC model updating framework and the NISS framework, as has been addressed in the fourth and fifth articles. The fourth article concerns the improvement of the distribution-free stochastic model updating framework for updating the parameters with mixed aleatory and epistemic uncertainties whilst taking into account their correlation structure. The joint probability distribution of the parameters to be inferred are represented as the Gaussian copula function the marginal distributions of which being described by the staircase density functions. The correlation coefficients of the correlation matrix that defines the correlation structure of the Gaussian copula function, as well as the hyper-parameters of the staircase density functions are accounted for as the parameters to be inferred, whose prior distributions are defined based on the correlation coefficient constraint and moment constraints, and these parameters are updated using the Bhattacharya distance-based approximate likelihood. The proposed updating procedure has been first demonstrated upon the updating problem of the shear building model used in the second article by utilizing another pair of the natural frequencies as the observed features for yielding the correlated target distribution of the inputs. This example clearly exhibits the capability of the proposed procedure to update the joint distribution among the parameters, which cannot be obtained analytically, with unknown correlation structure. Another simple engineering problem with multi-variate inputs has been also analyzed to indicate that the proposed procedure is capable of calibrating the general correlation structure regardless of the presence of no, negative, and positive correlations. Then, the proposed procedure has been applied to the updating problem of the seismic-isolated bridge pier model, showing its capability of recreating very complicated nonlinear uncertainty structure of the observed seismic response data.

The fifth article addresses the improvement of the generalized NISS framework to propagate the parameters with mixed aleatory and epistemic uncertainties without limiting hypotheses on the distribution families of the parameters.

The novel hybrid NISS method has been developed, in which the p-boxes based on the staircase density functions that represent the parameters with mixed aleatory and epistemic uncertainties are propagated using the local NISS method while the interval models that represent the parameters with only epistemic uncertainty are propagated using the global NISS method. The reliability analysis subproblem of the NASA UQ challenge problem 2019 has been correctly solved using the proposed hybrid NISS method, and the results demonstrate that the proposed method works well for this real-world engineering problem whereas the computational burden is eased a lot when compared to the results in the third article, which uses the probability bounds analysis by the double-loop Monte Carlo method.

In summary, the five developments in this thesis have made contributions to strengthen the subjective assumption-free UQ approach, in terms of suitability and efficiency, whereby dealing with both aleatory and epistemic uncertainties. All the developments, except for the first one, are based on the staircase density functions, that are capable of discretely approximating a broad range of distributions arbitrary close, to model the aleatory parameters the distribution families of which are not known. These developments provide their application in the field of both uncertainty calibration and propagation in which they play a role as key components in the developed distribution-free UQ approach.

## 7.2   Open problems and prospects

Concerning the realistic consideration and numerically efficient quantitative evaluation of aleatory and epistemic uncertainty for real-world applications, several open problems on uncertainty calibration and propagation have yet to be addressed.

Regarding the field of uncertainty calibration, this thesis has developed an efficient Bayesian inference algorithm for the combination of BUS with AK-MCMC and it has been successfully used to perform the Bhattacharyya distance-based ABC updating procedure. The algorithm is capable of drastically reducing the number of model evaluations for inferring the posterior by substituting the model evaluations in most of the samples for evaluating the adaptively trained Kriging surrogate. However, in this algorithm, the learning function employed to adaptively train the Kriging surrogate can only identify the single best point to be added in DOE at each iteration, hinding the use of ever-increasing parallel-computing facilities. To overcome this obstacle, paralell active learning strategies mainly based on applying clustering algorithm such as the $k$-means clustering have been investigated in the field of reliability problems, and such strategies can also be combined with each other within the BUS framework to provide a furthermore efficient Bayesian inference algorithm.

Furthermore, in this thesis, the application of the Gaussian copula function to the stochastic model updating has

been presented to calibrate the joint distribution of aleatory parameters that are correlated with each other. While this development has strengthened the subjective assumption-free stochastic model upating framework, it still relies on the assumption that the correlation structure to be calibrated can be uniquely described by means of the correlation matrix. Nevertheless, parameter dependencies found in practice can be very complex, e.g., nonlinear and asymmetric; hence, a more flexible strategy that is capable of describing a broad range of correlation structures is required to further extend the suitability of the current developments in the field of the stochastic model updating. The multi-variate copulas such as hierarchical copulas and pair-copulas have been studied in the field of reliablity problems to characterize the general dependence structure among multi-variate parameters, however the acceptability of such multi-variate copulas strongly depends on the selection of an appropriate family of copulas. Alternatively, the sliced-normal class of distributions has been recently proposed to characterize complex parameter dependencies and this is suggested to be more versatile than most copula families. Still, it has been only demonstrated to model multi-variate data and further studies are necessary about its applicability to the calibraiton of the dependence structure through model updating procedure.

Meanwhile, regarding the field of uncertainty propagation, this thesis has strengthened the subjective assumption-free approach by developing the hybrid NISS method to efficiently propagate the p-boxes based on the staircase density functions. However, one of the main issues of the NISS class of imprecise uncertainty propagation methods is that, the variations of NISS estimators can be very large due to the large variations of the weight coefficients of the estimators in the case where epistemic uncertainty is substantial due to the extreme lack of knowledge. This phenomena could be remarkable in the propagation of the p-boxes based on the staircase density functions, because they potentially consist of a wide range of distributions while the standard class of distributional p-boxes consists of a single distribution family. Recently, a novel class of method for the propagation of the distributional p-boxes has been proposed, where the failure probability function is inferred by means of the Bayesian quadrature, and shown to be effective for the problems with large epistemic uncertainty. Implementing this method for efficiently propagating p-boxes which are constructed based on the staircase density functions is one of author's themes in my future work.

Finally, most of the developments in this thesis are relevant to the applications of the staircase density functions. The staircase density functions are capable of providing the approximation of a broad range of distributions by solving an optimization problem based on the moment matching constraints. By their nature, the resultant approximations can differ for different cost functions employed in the optimization problem, and several cost functions can be considered, including the maximal-entropy and minimum-likelihood. Nonetheless the minimum-likelihood cost function has been employed in this thesis, further studies are necessary to identify the most appropriate choice of the cost function. In

particular, its most appropriate choice may vary for each task in the UQ process, because each task has its own aim; the model updating problem aims at calibrating the distributions of the parameters to quantify the general uncertainty characteristics of the target output features, whereas the reliability problems aim at estimating the failure probabilities that are relevant to the tail regions of the distributions of the parameters.

# Bibliography

Alvarez, D. A., Uribe, F., Hurtado, J. E. (2018). Estimation of the lower and upper bounds on the probability of failure using subset simulation and random set theory. *Mechanical Systems and Signal Processing*, 100, 782-801.

Au, S. K., and Beck, J. L. (1999). A new adaptive important sampling scheme. *Structural Safety*, 21, 135-158.

Au, S. K., and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4), 263-277.

Au, S. K., and Patelli, E. (2016). Rare event simulation in finite-infinite dimensional space. *Reliability Engineering & System Safety*, 148, 67-77.

Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162, 2025-2035.

Beck, J. L., and Katafygiotis, L. S. (1998). Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechnics*, 124(4), 455-461.

Beck, J. L., and Yuen, K. V. (2004). Model Selection Using Response Measurements: Bayesian Probabilistic Approach. *Journal of Engineering Mechanics*, 130(2), 192-203.

Beer, M., Ferson, S., and Kreinovich, V. (2013). Imprecise probabilities in engineering analysis. *Mechanical System and Signal Processing*, 37, 4-29.

Bernton, E., Jacob, P. E., Gerber, M., and Robert, C. R. (2019). Approximate Bayesian computation with the Wasserstein distance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81(2), 235-269.

Betz, W., Papaioannou, I., and Straub, S. (2016). Transitional Markov chain Monte Carlo: Observations and improvements. *Journal of Engineering Mechanics*, 142(5), 04016016.

Betz, W., Papaioannou, I., Beck, J. L., and Straub, S. (2018). Bayesian inference with Subset Simulation: Strategies and improvements. *Computer Methods in Applied Mechanics and Engineering*, 331, 72-93.

Bhattacharyya, A. (1946). On a measure of divergence between two multinominal populations. *Indian Journal of Statistics*, 7(4), 401-406.

Blatman, G., and Sudret, B. (2010). An adaptive algorithm to built up sparse polunomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, 25(2), 183-197.

Bucher, C. G., and Bourgund, U. (1990). A fast and efficient response surface approach for structural reliability problems. *Structural Safety*, 7(1), 57-66.

Bi, S., Deng, Z., and Chen, Z. (2013). Stochastic validation of structural FE-models based on hierarchical cluster analysis and advanced Monte Carlo simulation. *Finite Elements in Analysis and Design*, 67, 22-33.

Bi, S., Prabhu, S., Cogan, S., and Atamturktur, S. (2017). Uncertainty quantification metrics with varying statistical information in model calibration and validation. *AIAA Journal*, 55, 3570-3583.

Bi, S., Broggi, M., and Beer, M. (2019). The role of the Bhattacharyya distance in stochastic model updating. *Mechanical System and Signal Processing*, 117, 437-452.

Cheung, S. H., and Beck, J. L. (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.

Ching, J., and Chen, Y. C. (2007). Transitional Markov chain Monte Carlo method for Bayesian updating, model class selection, and model averaging. *Journal of Engineering and Mechanics*, 133(7), 816-832.

Chopin, N. (2002). A sequential particle filter method for static models. *Biometrica*, 89(3), 539-552.

Crespo, L. G., Kenny, S. P., and Giesy, D. P. (2014). The NASA Langley multidisciplinary uncertainty quantification challenge. In: *16th AIAA Non-Deterministic Approaches Conference*, National Harbor, Maryland.

Crespo, L. G., Kenny, S. P., Giesy, D. P., and Stanford, B. K. (2018). Random variables with moment-matching staircase density function. *Applied Mathematical Modelling*, 64, 196-213.

Crespo, L. G., and Kenny, S. P. (2021). The NASA Langley Challenge on Optimization Under Uncertainty. *Mechanical System and Signal Processing*, 152, 107405.

De Angelis, M., Patelli, E., and Beer, M. (2015). Advanced line sampling for efficient robust reliability analysis, *Structural Safety*, 52, 170-182.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38, 325-339.

Der Kiureghian, A., and Liu, P. L. (1986). Structural reliability under incomplete probability information. *Journal of Engineering Mechanics*, 112(1), 85-104.

Der Kiureghian, A., and De Stefano, M. (1991). Efficient algorithm for second-order reliability analysis. *Journal of Engineering Mechanics*, 117(12), 2906-2923.

Der Kiureghian, A., and Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2), 105-

112.

DiazDelaO, E. A., Garbuno-Inigo, A., Au, S. K., and Yoshida, I. (2017). Bayesian updating and model class selection with Subset Simulation. *Computer Methods in Applied Mechanics and Engineering*, 317, 1102-1121.

Echard, B., Gayton, N., and Lemaire, M. (2011). AK-MCS: An active learning reliability method conbining Kriging and Monte Carlo simulation. *Structural Safety*, 32, 145-154.

Echard, B., Gayton, N., Lemarie, M., and Relun, N. (2013). A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliability Engineering & System Safety*, 111, 232-240.

Faes, M., and Moens, D. (2020). Recent trends in the modeling and quantification of non-probabilistic uncertainty. *Archives of Computational Methods in Engineering*, 27(3), 633-671.

Faes, M., Daub, M., Marelli, S., Patelli, E., and Beer, M. (2021). Engineering analysis with probability boxes: A review on computational methods. *Structural Safety*, 93, 102092.

Govers, Y., Khodaparast, H. H., Link, M., and Mottershead, J. E. (2015). A comparison of two stochastic model updating methods using the DLR AIRMOD test structure. *Mechanical Systems and Signal Processing*, 52-53, 105-114.

Hasofer, A. M., and Lind, N. C. (1974). Exact and invariant second-moment code format. *Journal of the Engineering Mechanics Division*, 100(1), 111-121.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97-109.

Helton, J. C., Johnson, J. D., Oberkampf, W. L., and Sallaberry, C. J. (2010). Representation of analysis results involving aleatory and epistemic uncertainty. *International Journal of General Systems*, 39(6), 605-646.

Hohenbichler, M., and Rackwitz, R. (1981). Non-normal dependent vectors in structural safety. *Journal of Engineering Mechanics Division*, 107(6), 1227-1238.

Huang, X., Chen, J., and Zhu, H. (2016). Assessing small failure probabilities by AK–SS: An active learning method combining Kriging and Subset Simulation. *Structural Safety*, 59, 86-95.

Hurtado, J. E., and Diego, A. A. (2001). Neural-network-based reliability analysis: a comparative study. *Computer Methods in Applied Mechanics and Engineering*, 191, 113-132.

Hurtado, J. E. (2004). An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. *Structural Safety*, 26(3), 271-293.

Jiang, C., Bi, R. G., Lu, G. Y., and Han, X. (2013). Structural reliability analysis using non-probabilistic convex model. *Computer Methods in Applied Mechanics and Engineering*, 254, 83-98.

Kaymaz, I. (2005). Application of Kriging method to structural reliability problems. *Structural Safety*, 27, 133-151.

Katafygiotis, L. S., and Beck, J. L. (1998). Updating models and their uncertainties. II: Model identifiability. *Journal of Engineering Mechnics*, 124(4), 463-467.

Khodaparast, H. H., Mottershead, J. E., and Friswell, M. I. (2008). Perturbation methods for the estimation of parameter variability in stochastic model updating. *Mechanical System and Signal Processing*, 22, 1751-1773.

Khodaparast, H. H., Mottershead, J. E., and Badcock, K. J. (2011). Interval model updating with irreducible uncertainty using the Kriging predictor. *Mechanical Systems and Signal Processing*, 25(4), 1204-1226.

Lin, J. (1994). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1), 145-151.

Lye, A., Kitahara, M., Broggi, M., and Patelli, E. (2022). Robust optimization of a dynamic system under severe uncertainty: A distribution-free framework. *Mechanical System and Signal Processing*, 167, 108522.

Mares, C., Mottershead, J. E., and Friswell, M. I. (2006). Stochastic model updating: Part 1—theory and simulated example. *Mechanical System and Signal Processing*, 20(7), 1674-1695.

Mottershead, J. E., and Friswell, M. I. (1993). Model Updating In Structural Dynamics. *Journal of Sound and Vibration*, 167(2), 347-375.

Mottershead, J. E., Mares, C., James, S., and Friswell, M. I. (2006). Stochastic model updating: Part 2—application to a set of physical structures. *Mechanical System and Signal Processing*, 20(8), 2171-2185.

Mottershead, J. E., Ling, M., and Friswell, M. I. (2011). The sensitivity method in finite element model updating: A tutorial. *Mechanical System and Signal Processing*, 25(7), 2275-2296.

Nelsen, R. B. (2006). *An introduction to copulas*, Second edition, Springer-Verlag, New York.

Ortiz, G. A., Alvarez, D. A., and Bedoya-Ruíz, D. (2015). Identification of Bouc–Wen type models using the Transitional Markov Chain Monte Carlo method. *Computers & Structures*, 146, 252-269.

Panaretos, V. M., and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annual Review of Statistic and its Application*, 6, 405-431.

Papaioannou, I., Betz, W., Zwirglmaier, K., and Straub, D. (2015). MCMC algorithms for Subset Simulation. *Probabilistic Engineering Mechanics*, 41, 89-103.

Patelli, E., Alvarez, D. A., Broggi, M., and de Angelis, M. (2015). Uncertanty management in multidisciplinary design

of critical safety systems. *Journal of Aerospace Information Systems*, 12(1), 140-169.

Patelli, E., Govers, Y., Broggi, M., Gomes, H. M., Link, M., and Mottershead, J. E. (2017). Sensitivity or Bayesian model updating: a comparison of techniques using the DLR AIRMOD test data. *Archive of Applied Mechanics*, 87, 905-925.

Patra, B. K., Launonen, R., Ollikainen, V., and Nandi, S. (2015). A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data. *Knowledge-Based Systems*, 82, 163-177.

Rocchetta, R., Broggi, M., Huchet, Q., and Patelli, E. (2018a). On-line Bayesian model updating for structural health monitoring. *Mechanical Systems and Signal Processing*, 103, 174-195.

Rocchetta, R., Broggi, M., and Patelli, E. (2018b). Do we have enough data? Robust reliability via uncertainty quantification. *Applied Mathematical Modelling*, 54, 710-721.

Safta, C., Sargsyan, K., Najm, H. N., Chowdhary, K., Debusschere, B., Swiler, L. P., and Eldred, M. S. (2015). Probabilistic methods for sensitivity analysis and calibration in the NASA challenge problem. *Journal of Aerospace Information Systems,* 12(1), 170-188.

Sairajan, K. K., and Aglietti, G. S., (2012). Robustness of system equivalent reduction expansion process on spacecraft structure model validation. *AIAA Journal*, 50(11), 2376-2388.

Shafer, G. (1976). *A mathematical theory of evidence*, Princeton University Press, Princeton.

Schöbi, R., and Sudret, B. (2017). Structural reliability analysis for p-boxes using multi-level meta-models. *Probabilistic Engineering Mechanics*, 48, 27-38.

Schuëller, G. I., Pradlwarter, H. J., and Koutsourelakis, P. S. (2004). A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics*, 19, 463-473.

Simoen, E., De Roeck, G., and Lombaert, G. (2015). Dealing with uncertainty in model updating for damage assessment: A review. *Mechanical System and Signal Processing*, 56-57, 123-149.

Sobol', L. M., Tarantola, S., Gatelli, D., Kucherenko, S. S., and Mauntz, W. (2007). Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety*, 957-960.

Song, J., Wei, P., Valdebenito, M., Bi, S., Broggi, M., Beer, M., and Lu, Z. (2019). Generalization of non-intrusive imprecise stochastic simulation for mixed uncertain variables. *Mechanical Systems and Signal Processing*, 134, 106316.

Song, J., Wei, P., Valdebenito, M., and Beer, M. (2020). Adaptive reliability analysis for rare events evaluation with global imprecise line sampling. *Computer Methods in Applied Mechanics and Engineering*, 372, 113344.

Stein, M., Beer, M., and Kreinovich, V. (2013). Bayesian approach for inconsistent information. *Information Sciencees*, 245, 96-111.

Straub, D., and Papaioannou, I. (2015). Bayesian updating with structural reliability methods. *Journal of Engineering and Mechanics*, 141(3), 04014134.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22(4), 1701-1728.

Turner, B. M., and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2), 69-85.

Wei, P., Lu, Z., Song, J. (2014). Extended Monte Carlo simulation for parametric global sensitivity analysis and optimization. *AIAA Journal*, 52, 867-878.

Wei, P., Lu, Z., and Song, J. (2015). Variable importance analysis: A comprehensive review. *Reliability Engineering & System Safety*, 142, 399-432.

Wei, P., Tang, C., and Yang, Y. (2019a). Structural reliability and reliability sensitivity analysis of extremely rare failure events by combining sampling and surrogate model methods. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233(6), 943-957.

Wei, P., Song, J., Bi, S., Broggi, M., Beer, M., Yue, Z., Lu, Z. (2019b). Non-intrusive stochastic analysis with parameterized imprecise probability models: I. Performance estimation. *Mechanical Systems and Signal Processing*, 124, 349-368.

Wei, P., Song, J., Bi, S., Broggi, M., Beer, M., Yue, Z., and Lu, Z. (2019c). Non-intrusive stochastic analysis with parameterized imprecise probability models: II. Reliability and rare events analysis. *Mechanical Systems and Signal Processing*, 126, 227-247.

Yang, L., Bi, S., Faes, M., Broggi, M., and Beer, M. (2022). Bayesian inversion for imprecise probabilistic models using a novel entropy-based uncertainty quantification metric. *Mechanical Systems and Signal Processing*, 162, 107954.

Zhang, H., Mullen, R. L., and Muhanna, R. L. (2010). Interval Monte Carlo methods for structural reliability. *Structural Safety*, 32(3), 183-190.

Zhang, H., Dai, H., Beer, M., and Wang, W. (2013). Structural reliability analysis on the basis of small samples: An interval quasi-Monte Carlo method. *Mechanical Systems and Signal Processing*, 37, 137-151.

# Curriculum Vitae

## Personal Details

| | |
|---|---|
| Name | Masaru Kitahara |
| Date & place of birth | October 28, 1993 in Chiba, Japan |
| Nationality | Japanese |

## Education

| | |
|---|---|
| 04.2019 – current | Ph.D. student |
| | Institute for Risk and Reliability, Faculty of Civil Engineering and Geodetic Science, |
| | Leibniz Universität Hannover, Hannover, Germany |
| 04.2017 – 03.2019 | M.Eng. in Civil Engineering |
| | Department of Civil Engineering, University of Tokyo, Tokyo, Japan |
| | Master thesis: "Seismic load estimation for wind turbine support structures by |
| | response spectrum method considering non-classical damping and reliability level" |
| 04.2013 – 03.2017 | B.Eng. in Civil Engineering |
| | Faculty of Engineering Science, Yokohama National University, Kanagawa, Japan |
| | Bachelar thesis: "A study of efficient seismic capacity evaluation of existing bridge |
| | system using the response surface" |

## Work experience

| | |
|---|---|
| 04.2019 – current | Research associate at Institute for Risk and Reliability, Leibniz Universität |
| | Hannover, Hannover, Germany |
| | Research project: SFB 1463 Offshore-Metastrukturen (01.2021 – 12.2024) |
| | "Robuste lebensdauerumfassende Monitoring Konzepte für Offshore- |
| | Windenergieanlagen" |

## Awards

| | |
|---|---|
| 2019 | Dean Prize of Department of Civil Engineering, University of Tokyo (*FURUICHI* |
| | Award) |