A NEW APPROACH TO UNCERTAINTY QUANTIFICATION AND INTEGRATION IN SINGLE AND MULTIDISCIPLINARY SYSTEMS

by

AMIRUL ISLAM SAIMON

A thesis

submitted to the

Department of Industrial and Production Engineering,

Bangladesh University of Engineering and Technology,

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Industrial and Production Engineering

Department of Industrial and Production Engineering (IPE)

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY (BUET)

January 2021

CERTIFICATE OF APPROVAL

The thesis titled "A NEW APPROACH TO UNCERTAINTY QUANTIFICATION AND INTEGRATION IN SINGLE AND MULTIDISCIPLINARY SYSTEMS" submitted by Amirul Islam Saimon, Student ID.: 1017082014, Session: October, 2017, has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science in Industrial and Production Engineering on January 31, 2021.

Board of Examiners

Dr. AKM Kais Bin Zaman Professor Department of IPE, BUET, Dhaka	Chairman (Supervisor)
Dr. Nikhil Ranjan Dhar Professor & Head Department of IPE, BUET, Dhaka	Member (Ex-officio)
Dr. Abdullahil Azeem Professor Department of IPE, BUET, Dhaka	Member
Dr. Ferdous Sarwar Associate Professor Department of IPE, BUET, Dhaka	Member
Dr. Mohammed Forhad Uddin Professor Department of Mathematics, BUET, Dhaka.	Member (External)

CANDIDATE'S DECLARATION

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Amirul Islam Saimon

Student ID: 1017082014

ABSTRACT

Representation and propagation of different types of uncertainties (aleatory and epistemic) are increasingly being acknowledged in the design and optimization of complex systems. Statistical moment-based existing uncertainty representation approaches quantify uncertainties in an input quantity through moment bounds, and for this purpose, estimate different moments of a random variable from different sample sets. However, being not independent of each other, all the moments of the same random variable should be calculated from the same sample dataset. Again, existing approaches compute the widest possible (most conservative) range for each of the moments which offer maximum possible alternative values of uncertain variable, and is regarded as worst case from the aspect of uncertainty quantification. Therefore, first part of this thesis proposes a new probabilistic uncertainty representation approach that bags all the advantages of probabilistic approaches and simultaneously eliminates the limitations of existing methods. Proposed method includes development of a function of interest, and optimization of this function yields effective estimation of all the moment bounds of a random variable described by either multiple interval data or a combination of multiple interval and sparse point data. Then, these moment bounds are used to fit uncertain data to bounded Johnson distribution through utilization of moment matching. Finally, proposed uncertainty representation method is demonstrated with four numerical problems, which include three challenge problems from Sandia Epistemic Uncertainty Workshop, and the results are compared with earlier studies. Again, uncertainty propagation through multidisciplinary system is difficult due to the presence of interdisciplinary coupling, and it becomes more difficult when uncertainty incorporates in the input quantity. Therefore, second part of this thesis presents a unified probabilistic framework for the representation and propagation of both aleatory and epistemic uncertainty through multidisciplinary system. Proposed framework exploits worst-case maximum likelihood estimation (WMLE) method to quantify uncertainty, and likelihood-based multidisciplinary analysis (LAMDA) method to propagate the uncertainty through multidisciplinary system. Finally, a numerical problem and an engineering problem (Fire detection satellite) are used to demonstrate our proposed framework.

ACKNOWLEDGEMENT

By the grace of Almighty Allah, the thesis titled "A NEW APPROACH TO UNCERTAINTY QUANTIFICATION AND INTEGRATION IN SINGLE AND MULTIDISCIPLINARY SYSTEMS" has been ended, and the report has been completed.

First and foremost, I would like to convey my sincere respect and heartfelt gratitude to my thesis supervisor Dr. AKM Kais Bin Zaman, Professor, Department of Industrial and Production Engineering (IPE), Bangladesh University of Engineering and Technology (BUET), for his whole-hearted guidance throughout my studies at BUET. His constant inspiration has been a great source of motivation both in my academic career and personal life. I am extremely grateful to him, for everything that I have learnt during this thesis work, and sincerely look forward to continuing this relationship for the rest of my professional career.

I would like to express my sincere gratitude to other committee members, Dr. Nikhil Ranjan Dhar, Professor and Head, Department of IPE, BUET, Dr. Abdullahil Azeem, Professor, Department of IPE, BUET, Dr. Ferdous Sarwar, Associate Professor, Department of IPE, BUET, and Dr. Mohammed Forhad Uddin, Professor, Department of Mathematics, BUET, for their insightful comments and kind assessment of this work.

Next, I would like to convey thankfulness to all of my well-wishers who encouraged me to continue this thesis. I am grateful to my family members for their support and inspiration.

Finally, I want to acknowledge all of those who have supported me in any respect during the study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

PDF	:	Probability Density Function
CDF	:	Cumulative Distribution Function
MDA	:	Multidisciplinary Analysis
MDO	:	Multidisciplinary Design Optimization
MCS	:	Monte Carlo Simulation
FORM	:	First-Order Reliability Method
SORM	:	Second-Order Reliability Method
PBO	:	Percentile-Based Optimization
EBO	:	Expectation-Based Optimization
WMLE	:	Worst-case Maximum Likelihood Estimation
SOFPI	:	Sampling outside Fixed Point Iteration
LAMDA	:	Likelihood-based Approach for Multidisciplinary Analysis
FPI	:	Fixed-Point Iteration
MBA	:	Moment Bounding Approach
FMBA	:	Feasible Moment Bounding Approach

CHAPTER 1
INTRODUCTION

1.1 Background of the Study

Design and Optimization of complex engineering systems or processes are the major concern of Industrial Engineering (IE). Multidisciplinary Design Analysis (MDA) and Optimization (MDO) are increasingly being recognized as a systematic approach to design and optimization of complex IE systems. However, presence of uncertainty in the system input quantities and the models in real-life engineering problems, such as supply chain management (Borodin et al., 2016), system reliability and risk analysis (Zaman and Mahadevan, 2017), additive manufacturing (Hu and Mahadevan, 2017), industrial robotics (Wu et al., 2019), etc., make single and multidisciplinary system design analysis computationally more challenging. Uncertainty appears in a system design analysis either in a form of aleatory (natural variability) or epistemic (i.e. intervals from experts).

Several probabilistic and non-probabilistic approaches are available in literature including Helton et al (2004), Ferson and Hajagos (2004), Zaman et al. (2011a), Sankararaman and Mahadevan (2011), Zaman and Dey (2017), Yin et al. (2018), Hu et al. (2019), Peng et al. (2020), etc., each approach having its own benefits and limitations in representation and propagation of both aleatory and epistemic uncertainties through single disciplinary system.

Among these, probabilistic approaches based on statistical moment bounds (e.g., Zaman et al. (2011a); Peng et al. (2020)) are gradually becoming popular among the researchers to represent epistemic uncertainty. Moment bounding approaches estimate different moments of a random variable by using different configurations of the interval data. However, moments are not independent of each other for an arbitrary random variable. Therefore, it is still an open problem to develop an approach that can estimate optimal (narrowest possible) bounds preserving the dependency among all the moments of a random variable, and finally propagate these input uncertainties through system models. This is one of the focuses of this thesis.

Again, all existing uncertainty representation and propagation approaches are not explicitly effective for multidisciplinary analysis problems due to the additional requirement of satisfying interdisciplinary compatibility. A likelihood-based uncertainty propagation method is proposed in Sankararaman and Mahadevan (2012) to efficiently propagate aleatory uncertainty through multidisciplinary system. Liang et al. (2015) has generalized this approach to deal with both aleatory and epistemic uncertainties. This generalized approach represents the epistemic uncertainty of an input variable as a nonparametric probability density function (PDF). Since propagating nonparametric PDF through single or multidisciplinary system is difficult and inefficient, a likelihood-based approach overcoming this difficulty is yet to be available. This thesis has intended to focus on this point.

1.2 Objectives with Specific Aims

The specific objectives of the research are:

- To develop a probabilistic framework to represent the epistemic uncertainty of the system input variables characterized by either multiple interval data or a mixture of multiple interval and discrete point data, which confirms the narrowest possible bound of the system response CDFs.
- To develop a unified probabilistic framework to represent both sparse and/or interval uncertainty and way of propagating this type of uncertainties through a feedback-coupled Multidisciplinary Analysis (MDA) system.

1.3 Contributions of the Present Study

This thesis proposes a probabilistic uncertainty representation method based on statistical moment bounds which is capable of overcoming the limitations of the existing uncertainty representation approaches. The proposed approach is demonstrated with four example problems including three from Sandia epistemic uncertainty workshop (Oberkampf et al., 2004) challenge problems to show its effectiveness.

This thesis also introduces a likelihood-based unified framework that can efficiently represent different types of epistemic uncertainties, and propagate these uncertainties through multidisciplinary systems. One numerical and one complex engineering problems have been solved using the proposed framework to show its efficacy in representing and propagating different types of uncertainties.

1.4 Organization of the Thesis

The rest of this thesis report is organized in the following manner:

Chapter 2 reviews the existing approaches in literature relevant to the scope of this thesis. Chapter 3 provides necessary theoretical backgrounds including definition of uncertainty, its sources and classification, different uncertainty representation approaches, uncertainty propagation approaches for single and multidisciplinary systems. Chapter 4 presents proposed probabilistic methodology for representation of epistemic uncertainty of the input variables described by either multiple intervals or a combination of multiple interval and point data based on statistical moments. Uncertainty quantified in Chapter 4 is propagated through the single disciplinary system in Chapter 5. Additionally, four numerical example problems including three challenge problems from Sandia epistemic uncertainty workshop are also solved in this chapter. Chapter 6 presents a unified likelihood-based uncertainty representation and propagation framework for multidisciplinary system. This chapter also illustrates the proposed framework with a mathematical MDA problem and an engineering MDA problem. Finally, Chapter 7 concludes the thesis report with overall conclusive remarks and recommendations for future work regarding uncertainty representation and propagation.

CHAPTER 2

LITERATURE REVIEW

Inclusion of different forms of uncertainty in system input variable is common in almost every model-based engineering design, and it has significant impact on the output of the design. Failure in identifying input uncertainty may influence the design badly. Therefore, proper uncertainty representation of each input quantity of the model is very important for effective propagation of the identified uncertainty through model to quantify output uncertainty, and finally for active communication with the decision makers. This chapter presents a brief review of the available methods in literature regarding uncertainty representation, and propagation through single and multidisciplinary systems.

Uncertainty usually arises in two classes – aleatorty and epistemic. Aleatory uncertainty is the inherent variability of the process (mostly irreducible), and epistemic uncertainty is the state-of-knowledge uncertainty about the system (mostly reducible by gathering more information). More detail discussions regarding aleatory and epistemic uncertainties are available in Section 3.1. Since it is frequently required to estimate the distribution of epistemic variable from insufficient amount of data like intervals from expert opinion, methods to represent interval data are undoubtedly of significant importance.

Sandia epistemic uncertainty project (Oberkampf et al., 2004) arranged a workshop to gather and compare the state of the art methods from the researchers and practitioners regarding representation and propagation of epistemic uncertainty including interval data. With an eye toward comparison of the collected methods, the workshop also introduced a set of challenge problems, and provided numerical values of input quantities of the problems. A number of uncertainty theories and ideas regarding representation and propagation of interval uncertainty are presented in the workshop, and the results and findings are summarized in Ferson et al. (2004). Some of the presented theories include polynomial chaos expansions (Red-Horse and Benjamin, 2004), Dempster-Shafer formulation (Klir, 2004; Helton et al., 2004), information-gap models (Ben-Haim, 2004), probability boxes (Ferson and Hajagos, 2004), probability and possibility distributions (Helton et al., 2004), random sets (Berleant and Zhang, 2004), sets of probability measures (Fetz and Oberguggenberger, 2004), fuzzy sets (Fetz and Oberguggenberger, 2004), random intervals (Fetz and Oberguggenberger, 2004), etc. The benefits and drawbacks of these interval uncertainty theories usually depend on the available information regarding uncertain variables.

In general, representation approaches of interval uncertainty can be classified into two broad categories - non-probabilistic and probabilistic. Beer et al. (2013) listed some non-probabilistic approaches popular in representing epistemic uncertainty including interval data. *Evidence theory* (Shafer, 1976) is one of the fundamental non-probabilistic approaches to represent interval data. It can represent interval uncertainty by using belief measure. This theory helps in aggregating evidences from a number of sources, and reaching at a degree of belief that considers all possible evidences. Agarwal et al. (2004), Guo and Du (2009), Zhang et al. (2015), and Yang et al. (2017) used evidence theory to represent epistemic uncertainty including interval data. Some researchers used *convex model* (Ben-Haim and Elishakoff, 1990) to represent interval uncertainty because of its capacity in interpreting the extreme system output when input variables are uncertain. Convex model is basically a set of functions, where each function is representative of an uncertain event. Most common examples of non-probabilistic convex model include interval and ellipsoid model (Ni et al., 2018). Wang et al. (2014), Dey et al. (2016) and Huang et al. (2018) have utilized *possibility* or *fuzzy set theory* to represent interval uncertainty, where fuzzy membership function is employed.

The non-probabilistic uncertainty representation approaches mentioned above can be adopted to represent epistemic uncertainty including interval data. However, in case of real-life engineering problem, these approaches have experienced some challenges. One of the major drawbacks of non-probabilistic uncertainty representation and propagation approaches is that, these approaches are computationally expensive because of the requirement of nested analysis of interval variables during uncertainty propagation. Again, if input variables of a model have both aleatory and epistemic uncertainties, non-probabilistic uncertainty representation approaches cannot represent both types of uncertainty in a unified framework. Usually, evidence theories or fuzzy theories are used to represent interval data, whereas probabilistic approaches are employed to deal with aleatory uncertainties.

Furthermore, according to Helton et al. (2008), decision makers are not clearly aware of these non-traditional non-probabilistic uncertainty representation approaches. Additional efforts are required to make the decision makers fully aware of the results of the uncertainty analysis performed by these non-probabilistic approaches, as communication of these results with decision makers is one of the fundamental stages of uncertainty analysis.

In a target of achieving computational competency, representing both aleatory and epistemic uncertainties in a unified framework, and making decision makers informed of the final results of uncertainty analysis from scratch, a number of probabilistic approaches are available in literature to deal with interval data.

Sankararaman and Mahadevan (2011) proposed a *likelihood-based nonparametric approach* (LBNA) to represent a variable described by a mixture of both discrete point and multiple interval data. Motivated by Meeker and Escobar (1995), they have utilized the probability density function for discrete point data and cumulative distribution function for interval data in constructing the likelihood function. In order to make the assumption of any specific distribution type redundant, LBA represented epistemic uncertainty through a nonparametric PDF. Additionally, this approach is applied to solve some challenge problems from Sandia epistemic uncertainty workshop to show its efficacy in representing interval uncertainty. However, propagating uncertainty described by a nonparametric PDF is not straightforward even through simple single disciplinary system. Another likelihood based epistemic uncertainty representation approach titled *Worst-case Maximum Likelihood Estimation* (WMLE) proposed by Zaman and Dey (2017) can eliminate this difficulty by constructing a parametric CDF of epistemic variable described by both discrete point and interval data.

Zhang and Shields (2018) introduced a probabilistic uncertainty representation and propagation approach based on information theory. Initially they list a set of candidate probability models by multimodel inference rule, in such a way that each of the candidate model can presumably represent the uncertain variable. Then, Bayesian inference is used to approximate the joint probability density of each of the models.

The use of statistical moments (i.e., mean, variance, skewness, kurtosis, etc.) is becoming popular gradually in representing multiple types of epistemic uncertainties. Zaman et al. (2011a) proposed a probabilistic uncertainty representation approach titled *Moment Bounding Approach* (MBA) to represent interval data based on lower and upper bounds of the first four statistical moments of the variable. This approach utilized the concept of moment matching in order to fit four parameter family of Johnson distributions to data. Peng et al. (2020) proposed another uncertainty representation approach based on moment bounds. This approach has employed cubic normal transformation to sample points from moment bounds of the first through fourth moment of the variable.

The moment bounding approaches including both Zaman et al. (2011a) and Peng et al. (2020) estimate the lower and upper bounds of the first four moments of a variable described by interval data by minimizing and maximizing corresponding moment expression. Since this technique estimate moments by independently optimizing respective moment expression, it unintentionally estimates different moments of a particular random variable from different sets of realizations. More specifically, realizations set obtained using optimization of first moment expression is used to estimate first moment bounds of a random variable, in the same way, realizations sets obtained by optimizing second, third, and fourth moment expressions are used to estimate the bounds on second, third, and fourth moments respectively of the same random variable. More specifically, realizations set obtained using optimization of first moment expression is used to estimate first moment bounds of a random variable, in the same way, realizations sets obtained by optimizing second, third, and fourth moment expressions are used to estimate the bounds on second, third, and fourth moments respectively of the same random variable. However, moments are not independent, and thus, all the moments of a random variable should be calculated from the same set of realizations. An approach to represent epistemic uncertainty based on moment bounds excluding this limitation is yet to be developed.

This thesis proposes an optimization-based epistemic uncertainty representation approach in Chapter 4 based on moment bounds. The proposed formulation introduces a new function of interest aggregating all the moment expressions in a single function. Optimization of this function yields a single set of realizations, and this set of realizations is used to estimate all the moments of the uncertain variable. Accordingly, proposed approach bags all the advantages of probabilistic uncertainty representation approaches as well as bypasses the limitation of the existing moment bounding approaches.

Uncertainty propagation is another important element of uncertainty analysis. After successfully representing different types of uncertainties in the input variable, it is required to propagate these uncertainties through single or multidisciplinary systems. Uncertainty propagation through single disciplinary system is quite straightforward, and a number of researches, available in literature, deal with the propagation of both aleatory and epistemic uncertainty through single disciplinary systems including Zadeh (2002), Ferson et al. (2003), Agarwal et al. (2004), Du et al. (2006), Helton et al. (2007), Zhang and Huang (2009), Sankararaman and Mahadevan (2011), Zaman et al. (2011b), Matsumura and Haftka (2013), Zaman and Dey (2017), etc.; but, in case of multidisciplinary system, individual disciplines are coupled with one or more coupling variables, which require performing multidisciplinary analysis (MDA). Multidisciplinary systems emerge in a number of practical engineering applications, e.g., fluid–structure interaction (Belytschko, 1980), thermal-structural analysis (Thornton, 1992, Culler and McNamara, 2010), topology optimization (Dunning et al., 2011), turbine-engine cycle analysis (Hearn et al, 2016), etc.

Individual disciplines of a multidisciplinary system are coupled by one or more coupling variables that make MDA methods computationally expensive. This coupling between two disciplines can be either *feedforward* (single-directional) or *feedback* (bidirectional). Propagation methods for aleatory uncertainty through feedforward system include Monte Carlo methods, first-order reliability method (FORM), and second-order reliability method (SORM), etc. (Haldar and Mahadevan, 2000). Handling feedforward coupling is relatively easy, while handling feedback coupling is challenging. The challenges multiply with the inclusion of aleatory and epistemic uncertainty in case of both feedforward and feedback couplings. Methods at hand regarding the propagation of uncertainty through feedback coupled MDA are few in number. Again, most of the available methods can only deal with aleatory uncertainty, such as Gu et al. (2000), Kokkolaras et al (2006), Du and Chen (2005), Mahadevan and Smith (2006), Sankararaman and Mahadevan (2012), etc.

Another broad classification of MDA methods is available in literature (Chaudhuri et al., 2017). It categorizes the methods into three categories. First category is *fixed point iteration* (FPI) based approach. Examples of this group of approaches involve Cramer et al (1994), Alexandrov and Kodiyalam (1998), and many more. In this case, coupled disciplinary analysis continues to be iterated as long as a feasible fixed point iteration solution is achieved. Habitually fixed point iteration is computationally expensive, and feedback coupled multidisciplinary analysis becomes more expensive when uncertainty is involved in the analysis. If uncertainty involves, sets of realizations need to be sampled (e.g., Monte Carlo sampling) over uncertain parameters of the input variable, and it is required to perform fixed point iteration for each set of realizations. Thus, inclusion of uncertainty in coupled multidisciplinary system makes the approaches of this category almost computationally prohibited.

Methods of the second category replace this computationally expensive fixed point iteration with relevant *surrogate models* to propagate uncertainty through coupled multidisciplinary system. Arnst et al. (2012), Arnst et al. (2013), Chen et al (2013), Arnst et al (2014), and Jiang et al (2015) are some of the examples of methods that use surrogate model instead of fixed point iteration to approximate coupling variables of a coupled MDA system. Chaudhuri et al. (2017) introduced another uncertainty propagation method based on surrogate model which includes approximation of the coupling variables of multidisciplinary system by a low fidelity surrogate model, and then refinement of this surrogates by using adaptive sampling. Methods of this category use surrogate in place of FPI, and the efficiency of these methods are highly dependent on the exact approximation of the coupling variables. However, in case of practical coupled MDA system, these surrogate based existing methods cannot promise exact convergence to fixed point feasible solution.

The third category approaches are based on the concept of *decoupling* the coupled disciplines to propagate uncertainty through feedback coupled multidisciplinary system. These methods include fully or partially decoupled approaches. Methods outlined in Du and Chen (2005), and Mahadevan and Smith (2006) are the examples of fully decoupled approach, while the method described in Sankararaman and Mahadevan (2012) is representative of a partially decoupled approach. Capacity to bypass coupled analysis as well as computationally expensive fixed point iteration in multidisciplinary system analysis are the fundamental advantages of the fully or partially decoupled approaches. However, all of these partially or fully decoupling approaches can only propagate aleatory uncertainty through multidisciplinary system. Zaman (2010) introduced a probabilistic method for the propagation of both aleatory and epistemic uncertainty through multidisciplinary systems, which has utilized the fully decoupled formulation developed by Mahadevan and Smith (2006). In case of multidisciplinary multilevel systems, multidisciplinary outputs become the inputs to higher level discipline, and therefore, it is required to maintain the functional dependence between the coupling variables, and accordingly between the multidisciplinary outputs. Sankararaman and Mahadevan (2012) argued that, fully decoupled approach cannot maintain this functional dependence, as a result, fully decoupled MDA is a good choice for the case when the target is to get only the statistics of the multidisciplinary outputs. Sankararaman and Mahadevan (2012) also claimed that, partially decoupled approach can successfully maintain this functional dependence. Therefore, this thesis is concerned about partially decoupled multidisciplinary analysis methods.

Liang et al. (2015) extended the partially decoupled approach proposed by Sankararaman and Mahadevan (2012), so that it can propagate aleatory and epistemic uncertainty through multidisciplinary system. However, this approach represents the epistemic uncertainty of system input variable as nonparametric PDF, and it is generally accepted that propagating a nonparametric PDF through multidisciplinary coupled system is not straightforward due to the anonymity regarding distribution type and parameter. Therefore, this thesis has proposed a likelihood-based multidisciplinary analysis method in Chapter 6, which waives the difficulty of the existing method regarding distribution anonymousness.

CHAPTER 3

THEORETICAL BACKGROUND

3.1 Uncertainty: Sources and Classification

Oberkampf and Roy (2010) defined uncertainty as imprecision or inaccuracy in a value. This imprecision or inaccuracy can be involved in a value from a number of sources in various forms. According to Haldar and Mahadevan (2000), uncertainty usually appears in any real-life engineering model from either cognitive or noncognitive sources.

Cognitive or qualitative sources of uncertainty are related to the ambiguousness of an engineering model while the basis of model formulation is solely intellectual beliefs and assumptions. Cognitive sources cover the uncertainties associated with interpretation of specific parameters (e.g., performance), definition of experience of the relevant personality (i.e., engineers, workers, etc.), assessment of external influence (i.e., environmental impact), evaluation of existing condition of the project, approximation of human interaction factors and many more (Ayyub, 1994).

Noncognitive or quantitative sources can be categorized into three different categories, such as (*i*) inherent variability of any observation, (*ii*) statistical uncertainty due to limited data, and (*iii*) modeling uncertainty (Haldar and Mahadevan, 2000).

In general, uncertainty appears in engineering design in two forms – *aleatory* and *epistemic*. *Aleatory uncertainty* is the internal randomness of a quantity which is natural and irreducible. Stochastic uncertainty, irreducible uncertainty, variability, inherent uncertainty, etc. are used in literature to cite aleatory uncertainty. It appears due to the variations in production processes, operating conditions, quality control techniques, surrounding factors, etc. On the other hand, *epistemic uncertainty* arises due to shortage of knowledge about system, and if more knowledge can be gathered, this type of uncertainty can be minimized. In literature, a number of epistemic uncertainty classification approaches are available including Thunnissen (2003), and Zhuang and Pan (2012). Among these, one approach that covers almost rest of the approaches are outlined below according to Zaman and Mahadevan (2017).

i. *Statistical uncertainty*: Statistical uncertainty arises due to stochastic but inadequately known quantity. This happens when input random variables are provided as discrete point data or interval data, and it is required to estimate the probability distributions of the variable from this available data. Distribution parameters inferred from this available data are uncertain. If it is possible to collect more data about random variable, then epistemic uncertainty regarding distribution parameter can be minimized, but aleatory uncertainty still prevails. Probabilistic methods are more efficient to quantify statistical uncertainty.

ii. *Subjective uncertainty*: Subjective uncertainty arises due to fixed but inadequately known physical quantity. This happens when information regarding a quantity is available as an interval from experts, and the true value of the quantity lies within the lower and upper limits of the interval. If it is feasible to gather more information about the system, subjective uncertainty will disappear and the quantity turns into a fixed value. Non-probabilistic methods are more effective for quantification of subjective uncertainty.

iii. *Model uncertainty*: Model uncertainty arises due to system modeling errors. Rebba et al. (2006) classified model errors into two types, such as (1) model form error, and (2) solution approximation error. This type of uncertainty can be reduced by approximating model form more precisely or adopting more accurate numerical solution method. Bayesian methods can be utilized to quantify model uncertainty.

iv. *Method uncertainty*: Method uncertainty arises from the choice of different solution methods of a problem. When an engineering problem is solved by several methods, results may vary because of this uncertainty. Method uncertainty can be reduced by way of proper selection of the method based on their relative advantages and drawbacks.

When different types of uncertainty are incorporated in input variables, it needs to be represented for effective uncertainty propagation. Since this thesis is concerned about epistemic uncertainty, the following section discusses several such probabilistic representation approaches of epistemic uncertainty.

3.2 Representation of Epistemic Uncertainty

As discussed in Chapter 2, there are a number of probabilistic and non-probabilistic uncertainty representation approaches available in literature. This thesis is interested in only probabilistic approaches. Therefore, three probabilistic uncertainty representation approaches, such as moment bounding approach (Zaman et al., 2011a), likelihood-based approach (Sankararaman et al., 2011), and worst-case maximum likelihood estimation (Zaman and Dey, 2017) methods are discussed below.

3.2.1 Moment Bounding Approach (MBA)

If a random variable is described by interval data, it is not possible to calculate the accurate moments of the variable because there are infinite number of potential probability distributions that can represent the interval data. Therefore, moment bounding approach (MBA) developed by Zaman et al. (2011a) estimates the lower and upper bounds of first four moments of the variable, in such a way that, all the moments of potential distributions are bound to fall within respective moment bounds. MBA can represent uncertainty in a variable described by single interval data, multiple interval data, or a mixture of both multiple interval and sparse point data.

3.2.1.1 MBA for single interval data

The approach to estimate the lower and upper bounds of the moments of a random variable described by only single interval data are summarized in Table 3.1.

Table 3.1: Moment estimation technique for single interval data

Moment	Condition		Formula
	Lower bound	Upper bound	
1	PMF = 1 at lower end point = 0 elsewhere	PMF = 1 at upper end point = 0 elsewhere	$M_1 = E(x)$
2	PMF = 1 at any point = 0 elsewhere	PMF = 0.5 at each point	$M_2 = E(x^2) - (E(x))^2$
3	PMF = 0.2113 at lower end point = 0.7887 at upper end point	PMF = 0.7887 at lower end point = 0.2113 at upper end point	$M_3 = E(x^3) - 3E(x^2)E(x) + 2(E(x))^3$
4	PMF = 1 at any point = 0 elsewhere	PMF = 0.7887 at one of the end points = 0.2113 at the other end point	$M_4 = E(x^4) - 4E(x^3)E(x) + 6E(x^2)(E(x))^2 - 3(E(x))^4$

Note: $E(x^k) = \sum_{i=1}^{2} x_i^k p(x_i)$ for $k = 1$, or 2, or 3, or 4 and $p(x_i) =$ Probability Mass Function (PMF).

3.2.1.2 MBA for multiple interval data

If a random variable X is described by multiple interval data, the lower and upper bounds on first four moments of X can be estimated using the procedure illustrated in Table 3.2.

Table 3.2: Moment estimation technique for multiple interval data

<table>
<tr><th>Moment</th><th colspan="2">Formula</th></tr>
<tr><td rowspan="2">1</td><td>Lower Bound</td><td>$$M_1 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)$$ where, $x_i = lb_i, \quad i = \{1, 2, \ldots, n\}$</td></tr>
<tr><td>Upper Bound</td><td>$$M_1 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)$$ where, $x_i = ub_i, \quad i = \{1, 2, \ldots, n\}$</td></tr>
<tr><td>k</td><td colspan="2">$$\min_{x_1, x_2, \ldots, x_n}/\max M_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j\right)^k$$ $$s.t.\ lb_i \leq x_i \leq ub_i, \qquad i = \{1, 2, \ldots, n\}$$ where, n is the number of intervals and for $k = 2, 3,$ or 4 minimizing or maximizing the objective function yields the lower or upper bounds of kth moment.</td></tr>
</table>

3.2.1.3 MBA for a mixture of both multiple interval and sparse point data

Suppose an arbitrary random variable X is defined by m intervals and n discrete point data, collected from independent sources. First m observations are interval data, which are bounded to lie within corresponding interval $lb_i \leq x_i \leq ub_i$, for $i = \{1, 2, \ldots, m\}$. Following n observations are discrete point data fixed at their corresponding values c_i for $i = \{1, 2, \ldots, n\}$. Technique to estimate lower and upper bounds of the first four moments of random variable X is illustrated in Table 3.3 below.

Table 3.3: Moment estimation technique for multiple interval and sparse point data

<table>
<tr><th>Moment</th><th colspan="2">Formula</th></tr>
<tr><td rowspan="2">1</td><td>Lower Bound</td><td>$$M_1 = \frac{1}{m+n}\left(\sum_{i=1}^{m+n} x_i\right)$$ where, $x_i = lb_i, \quad i = \{1, 2, \ldots, m\}$
$x_{m+i} = c_i, \quad i = \{1, 2, \ldots, n\}$</td></tr>
<tr><td>Upper Bound</td><td>$$M_1 = \frac{1}{m+n}\left(\sum_{i=1}^{m+n} x_i\right)$$ where, $x_i = ub_i, \quad i = \{1, 2, \ldots, m\}$
$x_{m+i} = c_i, \quad i = \{1, 2, \ldots, n\}$</td></tr>
<tr><td>k</td><td colspan="2">$$\min_{x_1, x_2, \ldots, x_m} / \max M_k = \frac{1}{m+n}\sum_{i=1}^{m+n}\left(x_i - \frac{1}{m+n}\sum_{j=1}^{m+n} x_j\right)^k$$ $$s.t.\, lb_i \leq x_i \leq ub_i, \qquad i = \{1, 2, \ldots, m\}$$ where, $x_{m+i} = c_i$ for $i = \{1, 2, \ldots, n\}$ and minimized or maximized values of M_2, M_3 and M_4 are the lower or upper bounds on the second, third and fourth moments, respectively.</td></tr>
</table>

So far, moment bounding approach (MBA) of this section can quantify the uncertainty of an input variable described by single interval, multiple intervals, or a mixture of multiple interval and discrete points, in a form of moment lower and upper bounds. Once moment bounds are available, these can be utilized to fit interval data to any empirical probability distributions, such as Pearson, Beta, Lambda, or Johnson family of distributions. Zaman et al. (2011b) considered four-parameter Johnson distributions for this purpose because of its ability in representing normal, lognormal, bounded, or unbounded distributions, and flexibility in transforming a continuous random variable to standard normal space; and finally utilized moment matching approach to fit data to bounded Johnson distributions, which yields family of CDFs of the variable. Thus, MBA represents the uncertainty of a variable through multiple representative CDFs. Details of Johnson distributions and moment matching approach are available in Appendix A and B. The following subsection presents a nonparametric approach to represent epistemic uncertainty.

3.2.2 Likelihood-based Nonparametric Representation of Epistemic Uncertainty (LBNA)

Sankararaman and Mahadevan (2011) developed a likelihood-based nonparametric approach (LBNA) to represent epistemic uncertainty. Let a random variable X be described by n interval data and m sparse point data. The overall range (between minimum and maximum values of X) of these data are divided into a finite number of uniform divisions to get a set of discretization points $q_i(i = \{1, 2, \dots, Q\})$. If the PDF values at each of these Q points be $f_X(x_i = q_i) = p_i$ for $i = \{1, 2, \dots, Q\}$, then by using any of the interpolation techniques (i.e., linear, spline-based, Gaussian process interpolation, etc.) the PDF over the entire range of the random variable X can be approximated. Sankararaman and Mahadevan (2011) formulated the likelihood function $L(\boldsymbol{p})$ as shown in Eq. (3.1).

$$L(\boldsymbol{p}) \propto \left[\prod_{i=1}^{n} [F_X(b_i|\boldsymbol{p}) - F_X(a_i|\boldsymbol{p})]\right]\left[\prod_{j=1}^{m} f_X(x_i|\boldsymbol{p})\right] \quad (3.1)$$

Likelihood of Eq. (3.1) depends on (*i*) number of discretization points Q, (*ii*) corresponding PDF values p_i, and (*iii*) the choice of interpolation technique. By maximizing the likelihood function of Eq. (3.1) or solving the constrained optimization problem in Eq. (3.2), the value of PDF at different discretization points can be estimated.

$$\begin{gathered} \max_{\boldsymbol{p}} \quad L(\boldsymbol{p}) \\ s.t.\ p_i \geq 0 \text{ for } \forall p_i \in \boldsymbol{p} \\ f_x(x) \geq 0 \text{ for } \forall x \\ \int f_x(x)dx = 1 \end{gathered} \quad (3.2)$$

Interpolation of the PDF values obtained at different discretization points using Eq. (3.2) provides the overall PDF of the random variable X. Thus, uncertainty of a random variable characterized by interval and sparse point data can be represented through a nonparametric probability density function using LBNA method of this subsection. The following subsection presents another likelihood-based epistemic uncertainty representation approach.

3.2.3 Worst-case Maximum Likelihood Estimation (WMLE) Method

For a random variable X defined by n intervals and m discrete points, WMLE (Zaman and Dey, 2017) uses a likelihood function as in Eq. (3.3).

$$L(p) \propto \left[\prod_{j=1}^{m} f(x_i = c_i \mid p)\right]\left[\prod_{i=m+1}^{m+n} f(x_i \mid p)\right] \tag{3.3}$$

Where, first m observations are kept fixed at their corresponding point values c_i, and the subsequent n observations are from n intervals constrained to lie within their corresponding lower and upper bounds. However, likelihood function of Eq. (3.3) requires the expression of probability density function of the random variable. In this regard, WMLE uses four parameter bounded Johnson distribution because it can dismiss the probability of data falling outside their respective interval. Again, WMLE optimizes log-likelihood instead of direct likelihood function of Eq. (3.3) to enjoy computational convenience.

The generalized formulation of WMLE is given in Eq. (3.4). In this formulation, the inner loop optimization minimizes the likelihood and estimates the lower bound (worst) of the likelihood, and the outer loop optimization maximizes the minimized likelihood in inner loop.

$$\begin{gathered}\max_{p}\left(\min_{x}\left(f(x \mid p) = \log(L(p;x))\right)\right)\\ s.t.\, lb_i \le x_i \le ub_i, \qquad i = \{(m+1),(m+2),\ldots,(m+n)\}\end{gathered} \tag{3.4}$$

The optimization of Eq. (3.4) yields the four parameter of bounded Johnson distribution, and thus WMLE represents epistemic uncertainty of a variable through a parametric distribution.

Among abovementioned three uncertainty representation approaches, MBA method provides multiple CDFs, LBNA method provides single nonparametric PDF, and WMLE method provides single parametric CDF of uncertain variable embedded with epistemic uncertainty.

After effective uncertainty representation, another element of uncertainty analysis is uncertainty propagation. The following section discusses several approaches for uncertainty propagation through single disciplinary system.

3.3 Uncertainty Propagation through Single Disciplinary System

Representing interval data to a manageable and easy-to-handle format is the key to effective and efficient uncertainty propagation through system response models. Well-established uncertainty propagation methods, e.g., first-order reliability method (FORM), second-order reliability method (SORM), Monte Carlo simulations (MCS) can easily handle probabilistic format of epistemic uncertainty. Approach developed by Zaman et al. (2011b) can propagate both probabilistic and non-probabilistic types of uncertainty though system model. This thesis has utilized this propagation approach to clarify our proposed method in Chapter 4.

Zaman et al. (2011b) classified epistemic uncertainty related problems into two cases (Case 1: input variables defined by interval data, and Case 2: input variable distribution parameters defined by interval data), and suggested both (i) sampling and (ii) optimization-based methodologies for the propagation of uncertainties through a single disciplinary system for each of the cases. Although the detailed methodologies are available in the cited reference, a glimpse of it is outlined below.

3.3.1 Sampling-based uncertainty propagation method

Sampling-based approach suggests to propagate a family of input CDFs by following any probabilistic propagation methods (e.g., MCS, FORM, SORM, etc.) through the system equation to get a family of output response CDFs. This approach is computationally expensive.

3.3.2. Optimization-based uncertainty propagation method

Computationally expensive sampling approach can be replaced by efficient Percentile-Based Optimization (PBO), or Expectation-Based Optimization (EBO) methods.

PBO minimizes and maximizes the system response $g_{\alpha}(x|m)$ at different percentile values (α) to get the lower and upper bounds of the system response CDF, respectively, at that particular α level. To obtain the rigorous minimum and maximum system response CDF bounds by PBO, it is required to perform the optimization a number of times at different percentile points, which is computationally expensive and time consuming. On the other hand, EBO minimizes and maximizes the expectation of the system response $E\big(g(x|m)\big)$ and approximates the lower and upper bounds of the system response CDFs. EBO is less expensive than PBO because it requires single optimization for both the lower bound and the upper bound of output CDF.

PBO and EBO formulations are illustrated in Table 3.4 for Case 1 (input variable described by interval data) and Table 3.5 for Case 2 (input variable distribution parameters described by interval data), where the symbols carry similar meanings as in Zaman et al. (2011b).

Table 3.4: PBO and EBO for Case 1

PBO	EBO
$\min_m / \max \; g_\alpha(x \mid m)$ s.t. $m_i \geq a_i$ $m_i \leq b_i, \;\; \mathrm{i} = 1, 2, ..., 4$ $\beta_2 - \beta_1 - 1 \geq 0$ $\beta_2 - 2\beta_1 - 3 \leq 0$ where $\beta_1 = {m_3}^2 / {m_2}^3$ $\beta_2 = m_4 / {m_2}^2$	$\min_m / \max \; E(g(x \mid m))$ s.t. $m_i \geq a_i$ $m_i \leq b_i, \;\; \mathrm{i} = 1, 2, ..., 4$ $\beta_2 - \beta_1 - 1 \geq 0$ $\beta_2 - 2\beta_1 - 3 \leq 0$ where $\beta_1 = {m_3}^2 / {m_2}^3$ $\beta_2 = m_4 / {m_2}^2$

Table 3.5: PBO and EBO for Case 2

	Lower Bound	Upper Bound
PBO	$\min_{m}\left(\min_{\theta^D}\left(g_\alpha\left(x \mid \theta^D\right) \mid \theta^D \in \theta^B(m)\right)\right)$ s.t. $m_i \geq a_i$ $m_i \leq b_i, \ \mathrm{i}=1,2,\ldots,4$ $\beta_2-\beta_1-1 \geq 0$ $\beta_2-2\beta_1-3 \leq 0$ where $\beta_1={m_3}^2/{m_2}^3$ $\beta_2=m_4/{m_2}^2$	$\max_{m}\left(\max_{\theta^D}\left(g_\alpha\left(x \mid \theta^D\right) \mid \theta^D \in \theta^B(m)\right)\right)$ s.t. $m_i \geq a_i$ $m_i \leq b_i, \ \mathrm{i}=1,2,\ldots,4$ $\beta_2-\beta_1-1 \geq 0$ $\beta_2-2\beta_1-3 \leq 0$ where $\beta_1={m_3}^2/{m_2}^3$ $\beta_2=m_4/{m_2}^2$
EBO	$\min_{m}\left(\min_{\theta^D}\left(E\left(g\left(x \mid \theta^D\right) \mid \theta^D \in \theta^B(m)\right)\right)\right)$ s.t. $m_i \geq a_i$ $m_i \leq b_i, \ \mathrm{i}=1,2,\ldots,4$ $\beta_2-\beta_1-1 \geq 0$ $\beta_2-2\beta_1-3 \leq 0$ where $\beta_1={m_3}^2/{m_2}^3$ $\beta_2=m_4/{m_2}^2$	$\max_{m}\left(\max_{\theta^D}\left(E\left(g\left(x \mid \theta^D\right) \mid \theta^D \in \theta^B(m)\right)\right)\right)$ s.t. $m_i \geq a_i$ $m_i \leq b_i, \ \mathrm{i}=1,2,\ldots,4$ $\beta_2-\beta_1-1 \geq 0$ $\beta_2-2\beta_1-3 \leq 0$ where $\beta_1={m_3}^2/{m_2}^3$ $\beta_2=m_4/{m_2}^2$

Optimization constraints are exactly the same for all the cases of PBO and EBO in Tables 3.4 and 3.5. First two constraints ensure that decision variables $\boldsymbol{m}$ lie within the corresponding moment bounds $a_i \leq m_i \leq b_i$ for $i = 1, 2, \ldots, 4$ and the last two constraints give the confirmation that the selected moments agree with the bounded Johnson distribution fit.

Uncertainty propagation through single disciplinary system is quite straightforward. However, in case of multidisciplinary systems, it is complicated due to the presence of multidisciplinary coupling. The following section discusses uncertainty propagation through multidisciplinary system.

3.4 Multidisciplinary Analysis (MDA) Under Uncertainty

Consider a simple multidisciplinary system as shown in Figure 3.1. Here, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the local input variables to discipline 1 and 2, respectively, $\boldsymbol{x}_s$ are the shared variable, which are common to each of discipline 1 and 2. Again, discipline 1 and 2 are coupled by feedback coupling variables $\boldsymbol{u}_{12}$ and $\boldsymbol{u}_{21}$. g_1 and g_2 are the outputs of discipline 1 and 2, and same time, inputs to discipline 3, and finally f is the overall system output. Since coupled system output g_1 and g_2 is used for further analysis to compute f, the system in Figure 3.1 is not only a multidisciplinary but also a multilevel system. The comprehensive target in this problem includes calculation of subsystem and system response statistics through multidisciplinary analysis. Since "Analysis 1" and "Analysis 2" are coupled by coupling variables $\boldsymbol{u}_{12}$ and $\boldsymbol{u}_{21}$, without decoupling the analyses, it is difficult to perform subsystem analyses and subsequently system analysis.

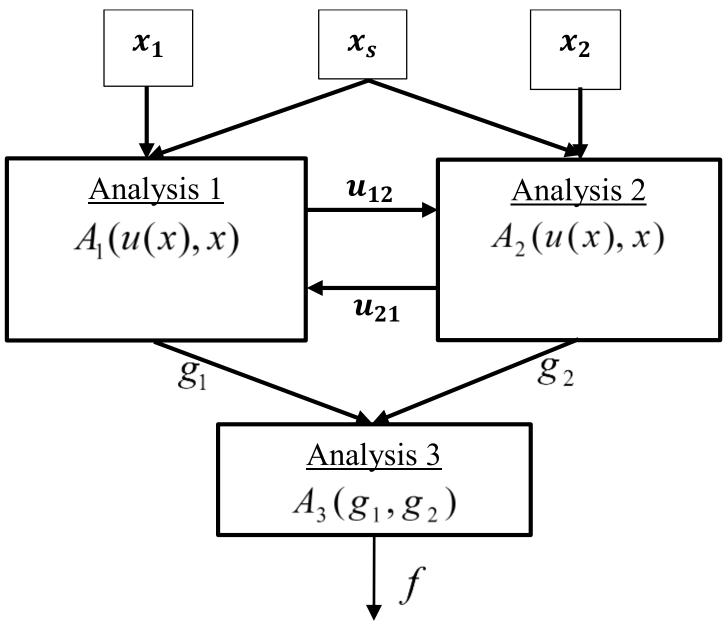


Figure 3.1: A multidisciplinary system

If the probability distributions of the input variables are provided, then, the target of the multidisciplinary system shown in Figure 3.1 is to evaluate the probability distributions of the subsystem outputs g_1 and g_2, and overall system output f. In this process, an important intermediary step is to estimate the probability distributions of the coupling variables u_{12} and u_{21}. For this, as mentioned in Chapter 2, this thesis utilized the concept of partially decoupling the coupling between discipline 1 and 2 following the approach illustrated in Sankararaman and Mahadevan (2012).

In Figure 3.1, starting with an initial guess of u_{12}, "Analysis 2" provides a value of u_{21}. Again, since u_{21} is an input to "Analysis 1", after performing individual disciplinary analysis in "Analysis 1", the output should be u_{12} itself. Sankararaman and Mahadevan (2012) introduced a function named G as shown in Figure 3.2 to illustrate this simple concept.

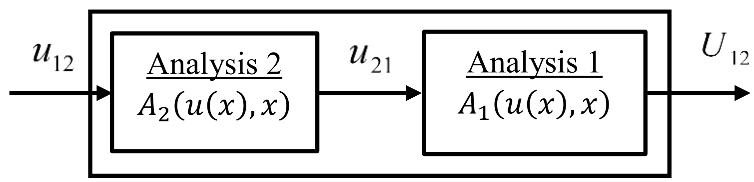


Figure 3.2: Definition of function G

From Figure 3.2, G function can be written as Eq. (3.5).

$$u_{12} = U_{12} = G(u_{12}, x) = A_1(u_{21}, x) = A_1(A_2(u_{12}, x), x) \tag{3.5}$$

Once the converged value of u_{12} can be estimated using Eq. (3.5), the feedback coupling of Figure 3.2 becomes easily manageable feedforward coupling as shown in Figure 3.3.

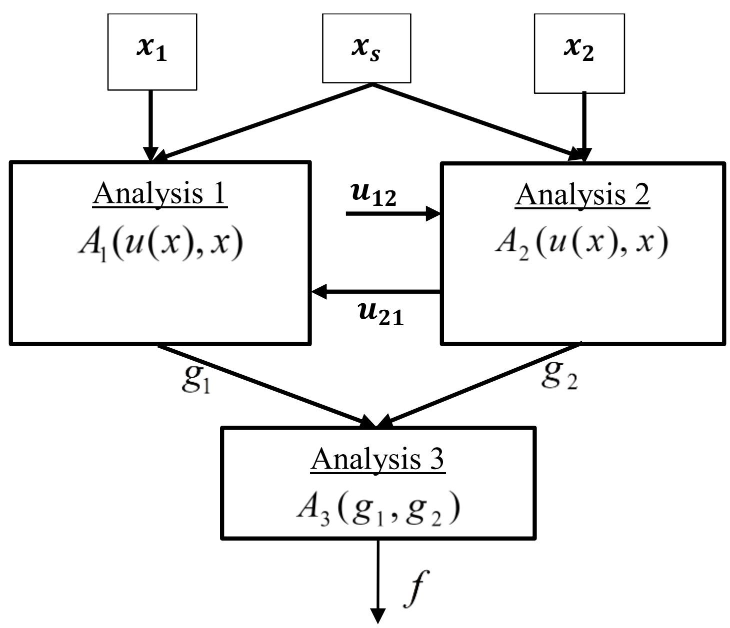


Figure 3.3: A partially decoupled multidisciplinary system

Based on Figure 3.2 and Eq. 3(.5), the following two subsection discuss two multidisciplinary analysis methods to estimate the probability density function of the coupling variables u_{12} and u_{21}.

3.4.1 Sampling outside fixed point iteration (SOFPI)

Sampling outside Fixed Point Iteration (SOFPI) method (Sankararaman and Mahadevan (2012), Ghosh and Mavris (2020)) is based on the well-known fixed point iteration (FPI) algorithm. If the probability distributions of the input variables (local and shared) are available, then SOFPI method can be used to approximate the probability distributions of the coupling variables u_{12} and u_{21}. The steps of SOFPI propagation through multidisciplinary systems are discussed below.

Step 1 Generate *N* number of random samples for each of local and shared input variables of multidisciplinary system.

Step 2 Take a set of realizations of the input variables.

Step 3 Estimate the converged value of u_{12} of Eq. (3.6) using fixed point iteration.

$$(u_{12})_{n+1} = A_1(A_2((u_{12})_n, x), x) \tag{3.6}$$

Step 4 Repeat Step 2 and 3 with *N* different sets of realizations to get *N* number of converged u_{12} values.

Step 5 Propagate converged values of u_{12} obtained in Step 4 through $u_{21} = A_2(u_{12}, x)$ to get the same number of u_{21} values.

Since the input variables are generated randomly using distribution information, convergent values of the coupling variables will also be random. These convergent values are used to estimate the probability density function of the coupling variables using any nonparametric density estimator (e.g., kernel density estimator). Thus, SOFPI method estimates the entire distribution of the coupling variables.

3.4.2 Likelihood-based Approach for Multidisciplinary Analysis (LAMDA)

SOFPI method requires generating samples of input variables using distribution information. Likelihood-based MDA (LAMDA) approach (Sankararaman and Mahadevan, 2012) discards this necessity of computationally expensive sampling. Motivated by the maximum likelihood estimation (MLE) principle, LAMDA utilizes the whole likelihood function to attain not only the maximum estimate but also the entire probability density function of the parameter.

If likelihood of given data $L(p)$ is defined as the probability of observing the data given parameter p, then, up to a proportionality constant, the likelihood of a particular value of coupling variable u_{12} can be defined as the probability of satisfying interdisciplinary compatibility constraint ($U_{12} = u_{12}$) conditioned on u_{12} as shown in Eq. (3.7).

$$L(u_{12}) \propto P((U_{12} = u_{12}) \mid u_{12}) \tag{3.7}$$

Again, since probability of a particular point is zero, the calculation of probability in Eq. (3.8) requires a modification. In this regard, Pawitan (2001) suggested to integrate the conditional PDF over an infinitesimally small interval $\left[u_{12} - \frac{\varepsilon}{2}, u_{12} + \frac{\varepsilon}{2}\right]$ around u_{12} as shown in Eq. (3.8), where ε is the interval length.

$$P((U_{12} = u_{12}) \mid u_{12}) = \int_{u_{12}-\frac{\varepsilon}{2}}^{u_{12}+\frac{\varepsilon}{2}} f_{U_{12}}(U_{12} \mid u_{12}) dU_{12} \tag{3.8}$$

The integration in Eq. (3.8) can be evaluated by utilizing First Order Reliability Method (FORM) according to Sankararaman and Mahadevan (2012). Utilization of FORM includes evaluation of the probability of a limit state function $H \equiv h(x)$ to be less than or equal to h_c, that is $P(h(x) \leq h_c)$. Again, $P(h(x) \leq h_c)$ is the cumulative distribution function (CDF) of $H \equiv h(x)$ at the point $H = h_c$. By using this concept, the lower bound integral h_l and the upper bound integral h_u in Eq. (3.8) can be easily evaluated using Eqs. (3.9) and (3.10), respectively.

$$h_l = F_{U_{ij}}\left(U_{ij} < u_{ij} - \left(\varepsilon/2\right) \middle| u_{ij}\right) = P\left(U_{ij} < u_{ij} - \left(\varepsilon/2\right)\right) \tag{3.9}$$

$$h_u = F_{U_{ij}}\left(U_{ij} < u_{ij} + \left(\varepsilon/2\right) \middle| u_{ij}\right) = P\left(U_{ij} < u_{ij} + \left(\varepsilon/2\right)\right) \tag{3.10}$$

Once the lower bound and upper bound integrals of Eq. (3.8) are evaluated, Eq. (3.11) approximates the likelihood of u_{12}.

$$L(u_{12}) \propto \frac{h_u - h_l}{\delta} \tag{3.11}$$

After evaluating likelihood of u_{12} at few points, the PDFs of u_{12} can be approximated by Eq. (3.12).

$$f(u_{12}) = \frac{L(u_{12})}{\int L(u_{12}) du_{12}} \tag{3.12}$$

where, $L(u_{12}) \neq 0$. Using Eq. (3.12), someone can estimate PDF values at different u_{12} points, and interpolation of these calculated PDF values yields the entire PDF of the coupling variable u_{12}. Thus, LAMDA method estimates the entire PDF of the coupling variable u_{12}, and helps in decoupling the multidisciplinary analyses. PDF of u_{12} is used to estimate the PDF of u_{21}, and in this manner, coupling variables u_{12} and u_{21} becomes independent variables, similar to input variables x.

So far, if the probability distributions of the input variables of a multidisciplinary coupling system are at hand, we can estimate the probability distributions of the coupling variables and partially decouple the coupled disciplines. Once the probability distributions of the coupling variables be available, then uncertainty propagation through upper level subsystem and system analysis becomes similar to traditional uncertainty propagation through single disciplinary system.

Uncertainty propagation through single or multidisciplinary systems becomes more complicated when the uncertainty associated with input variables is of epistemic type. Therefore, effective representation of input uncertainty is a prerequisite for effective uncertainty propagation through single or multidisciplinary systems. The following chapter proposes a new approach to uncertainty representation of input variables characterized by multiple interval data or a mixture of both multiple interval and discrete point data.

CHAPTER 4

PROPOSED PROBABILISTIC APPROACH FOR REPRESENTATION OF EPISTEMIC UNCERTAINTY

In this chapter, this thesis proposes a probabilistic uncertainty representation approach to represent a random quantity described by either multiple interval data, or a mixture of sparse point and multiple interval data. As discussed in Chapter 2, a number of probabilistic and non-probabilistic uncertainty representation methods are ready for use in literature. Generally, non-probabilistic methods are computationally expensive due to the requirement of nested analysis in the presence of interval uncertainty. Additionally, absence of any non-probabilistic unified formulation that can represent both aleatory and epistemic uncertainty simultaneously pushes researchers and practitioners to adopt probabilistic uncertainty representation methods.

This thesis focuses on probabilistic uncertainty representation methods based on statistical moments (e.g., mean, variance, skewness, kurtosis, etc.). Statistical moment based probabilistic methods (e.g., Zaman et al., 2011a; Peng et al., 2020; etc.) are labelled as moment bounding approaches (MBA) in literature because these methods include estimation of lower and upper bounds of different moments of a random variable. To estimate different moment bounds, existing moment bounding approaches (outlined in Section 3.2.1) optimize individual moment expression. For example, estimation of bounds on the first, second, third and fourth moments of a random variable includes individual optimization of the first through fourth moment expressions, respectively. Individual optimization of the first four moment expressions provides four different sets of realizations, and these four different sets of realizations are used to estimate the first four moments of the same random variable. Since different statistical moments of a random variable are not independent of each other, all the moments of a variable should be calculated from the same realizations.

Again, existing moment bounding approaches estimate the maximum and minimum possible bounds of different moments of a random variable to include all possible combinations of different moments. Thus, these approaches provide widest possible moment interval of an epistemic variable. Consequently, this is the most conservative way of representing epistemic uncertainty. But, the probability of selecting the true value of a variable with reduced error increases in case of selection from narrower bound than from wider one. Therefore, this thesis proposes an optimization-based probabilistic framework titled *Feasible Moment Bounding Approach* (FMBA) for representation of different types of epistemic uncertainty, which can circumvent the limitations of the existing moment bounding approaches.

Proposed FMBA formulation includes development of a function of interest by adding the squared differences of the moments. Optimization of this function yields a single set of realizations of an input random variable, which are then used to estimate all the moments of the same random variable. By optimizing the single objective function to estimate all the moments of a random variable, the proposed formulation excludes the requirement of individual optimization of each moment expression independently, and thus eliminates the drawbacks of the existing moment bounding approaches.

Existing MBA minimizes and maximizes the individual moment expressions, and estimates the minimum and maximum possible moment bounds, consequently, bounds obtained by existing MBA are the most conservative or widest possible. On the other hand, proposed FMBA optimizes a single function and utilizes the optimized set of realizations to estimate all the moments instead of individually optimizing each moment expressions. Thus, resulting moment bounds of a random variable obtained by the proposed FMBA will always be narrower than the ones obtained by the existing MBA.

It can be obtained by using any empirical probability distribution system, such as Johnson family of distributions, Pearson family of distributions

Any empirical probability distribution system, such as Johnson family of distributions or Pearson family of distributions can be utilized for the purpose. This thesis uses a four-parameter family of Johnson distributions, because it can replicate the shape of a number of named probability distributions including normal distribution, lognormal distribution, etc. Additionally, Johnson distributions is useful in transforming a continuous random variable into standard normal space, which can be effectively used in Reliability-based design optimization (RBDO) approaches. However, details of Johnson distributions can be found in Appendix A.

After effective quantification of uncertainty of the random variable through first four moment bounds, it is required to approximate the probability distribution (distribution type and distribution parameters) of the variable. Any empirical probability distribution systems, such as Johnson family of distributions, Pearson family of distributions, Beta family of distributions, or Lambda family of distributions can be utilized for the purpose. This thesis uses a four-parameter family of Johnson distributions, because (*i*) it can replicate the shape of normal distribution, lognormal distribution, bounded distributions, or unbounded distributions; additionally, (*ii*) it is useful in transforming a continuous random variable into standard normal space, which can be effectively used in Reliability-based design optimization (RBDO) approaches. More specifically, bounded Johnson distribution is a reasonable choice for fitting interval data as it gives confirmation of not selecting any distribution parameter that results in realizations of the arbitrary random variable out of actual interval set (Zaman et al., 2011b). Details of Johnson distributions is provided in Appendix A. Again, for the purpose of fitting data to Johnson family of distributions, this thesis utilizes statistical moment matching approach as outlined in Appendix B. Thus, the proposed FMBA represents a random variable having epistemic uncertainty through a family of cumulative distribution functions (CDFs). Structured formulations of proposed FMBA is described in the following section.

4.1 Proposed Feasible Moment Bounding Approach (FMBA)

The Proposed FMBA can accommodate both multiple interval data, and a mixture of multiple interval and discrete point data. Estimation technique of lower and upper bounds of first four moments of an arbitrary random variable described by multiple interval data are outlined in Section 4.1.1, and variable described by a mixture of both multiple interval and discrete point data in Section 4.1.2.

4.1.1 FMBA for multiple interval data

Consider an arbitrary random variable X described by n number of intervals. Interval data are gathered from different sources (i.e., expert opinions), and sources are assumed independent of each other. The feasible bounds on the first four moments of the variable X having n intervals can be estimated in three steps as follows.

Step 1: Calculate the lower and upper bounds on first four moments of the random variable X described by multiple interval data using existing moment bounding approach (MBA) outlined in Section 3.2.1.2.

Step 2a: Solve the following optimization problem in Eq. (4.1) for the lower bounds of all the moments.

$$\min_{x_1,x_2,\ldots,x_n} f_{lb} = \left(\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) - m_{1lb\_mba}\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 - m_{2lb\_mba}\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^3 - m_{3lb\_mba}\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^4 - m_{4lb\_mba}\right)^2 \quad (4.1)$$

$$s.t. \quad lb_i \le x_i \le ub_i, \quad i = 1, 2, 3, \ldots, n$$

where, n is the number of intervals and m_{ilb_mba} is the lower bound of ith (i = 1, 2, 3, 4) moment of X, which is calculated in Step 1 using the existing MBA.

Step 2b: Solve the following optimization problem in Eq. (4.2) for the upper bounds of all the moments.

$$\min_{x_1,x_2,\ldots,x_n} f_{ub} = \left(\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) - m_{1ub\_mba}\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 - m_{2ub\_mba}\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^3 - m_{3ub\_mba}\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^4 - m_{4ub\_mba}\right)^2 \quad (4.2)$$

$$s.t. \quad lb_i \le x_i \le ub_i, \quad i = 1, 2, 3, \ldots, n$$

where, n is the number of interval and m_{iub_mba} is the upper bound of ith (i = 1, 2, 3, 4) moment of X, which is calculated in Step 1 using the existing MBA.

Step 3: Minimization of Eq. (4.1) in Step 2a and Eq. (4.2) in Step 2b yield two sets of realizations of the random variable *X*. Realizations set obtained in Step 2a is used to calculate the lower bounds of first four moments of the variable *X* by using moment expressions outlined in Section 3.2.1.2. Similarly, upper bounds of first four moments of the variable *X* can be calculated by utilizing the optimized realizations set of Step 2b and moment expressions of Section 3.2.1.2.

4.1.2 FMBA for a mixture of multiple interval and sparse point data

The technique to estimate the feasible bounds on the first four moments of a random variable *X* having *m* intervals and *n* sparse point data (c_j for $j = \{1, 2, \dots, n\}$) can be illustrated by following three steps.

Step 1: Calculate the bounds on the first through fourth moments of an arbitrary random variable *X* defined by both multiple interval data and discrete point data using the existing moment bounding approach (MBA) discussed in Section 3.2.1.3.

Step 2a: Solve the following optimization problem in Eq. (4.3) for the lower bounds of all the moments.

$$\min_{x_1, x_2, \dots, x_m} f_{lb} = \left(\frac{1}{m+n} \left(\sum_{i=1}^{m+n} x_i \right) - m_{1lb\_mba} \right)^2 + \left(\frac{1}{m+n} \sum_{i=1}^{m+n} \left(x_i - \frac{1}{m+n} \sum_{i=1}^{m+n} x_i \right)^2 - m_{2lb\_mba} \right)^2 + \left(\frac{1}{m+n} \sum_{i=1}^{m+n} \left(x_i - \frac{1}{m+n} \sum_{i=1}^{m+n} x_i \right)^3 - m_{3lb\_mba} \right)^2 + \left(\frac{1}{m+n} \sum_{i=1}^{m+n} \left(x_i - \frac{1}{m+n} \sum_{i=1}^{m+n} x_i \right)^4 - m_{4lb\_mba} \right)^2 \quad (4.3)$$

$$s.t. \; lb_i \le x_i \le ub_i \qquad i = \{1, 2, \dots, m\}$$

where, m_{klb_mba} is the lower bound of *k*th (*k* = 1, 2, 3, 4) moment, which is calculated in Step 1 using the existing MBA, and $x_{m+j} = c_j$ for $j = \{1, 2, \dots, n\}$.

Step 2b: Solve the following optimization problem in Eq. (4.4) for the upper bound of all the moments.

$$\begin{aligned}
\min_{x_1,x_2,\ldots,x_m} f_{ub} = & \left(\frac{1}{m+n} \left(\sum_{i=1}^{m+n} x_i \right) - m_{1ub_mba} \right)^2 \\
& + \left(\frac{1}{m+n} \sum_{i=1}^{m+n} \left(x_i - \frac{1}{m+n} \sum_{i=1}^{m+n} x_i \right)^2 - m_{2ub_mba} \right)^2 \\
& + \left(\frac{1}{m+n} \sum_{i=1}^{m+n} \left(x_i - \frac{1}{m+n} \sum_{i=1}^{m+n} x_i \right)^3 - m_{3ub_mba} \right)^2 \\
& + \left(\frac{1}{m+n} \sum_{i=1}^{m+n} \left(x_i - \frac{1}{m+n} \sum_{i=1}^{m+n} x_i \right)^4 - m_{4ub_mba} \right)^2 \\
& s.t. \; lb_i \le x_i \le ub_i \qquad i = \{1, 2, \ldots, m\}
\end{aligned} \tag{4.4}$$

where, m_{kub_mba} is the upper bound of kth (k = 1, 2, 3, 4) moment, which is calculated in step 1 using existing MBA, and $x_{m+j} = c_j$ for $j = \{1, 2, \ldots, n\}$.

Step 3: Step 2a and Step 2b yield two optimized realization sets of the random variable X, which are used in Step 3 to calculate the lower and upper bounds of the first four moments respectively through the utilization of moment expressions outlined in Section 3.2.1.3

Once the uncertainty of the random variable X (described by either multiple interval data or a mixture of multiple interval and discrete point data) are quantified in a form of lower and upper bounds on the first four moments following Section 4.1.1 and 4.1.2, the probability distribution of X can easily be approximated using any of the available empirical probability distributions. As mentioned earlier, this thesis makes proper use of four-parameter family of Johnson distributions for the purpose, and utilizes moment (statistical) matching to fit data to the cited distributions family.

The computational efforts of proposed FMBA is almost similar to existing MBA method except the requirements of two additional optimizations in Step 2a and 2b. Zaman et al. (2011a) argued with numerical examples for both overlapping and non-overlapping types of intervals that, if number of intervals increases, existing MBA is found to be scalable in polynomial time; and, this assertion holds true for our proposed FMBA.

The generalized formulations of the proposed FMBA, starting from arbitrary random variable available in a form of multiple interval data or a mixture of multiple interval and sparse point data to CDFs of the random variable is illustrated in Figure 4.1.

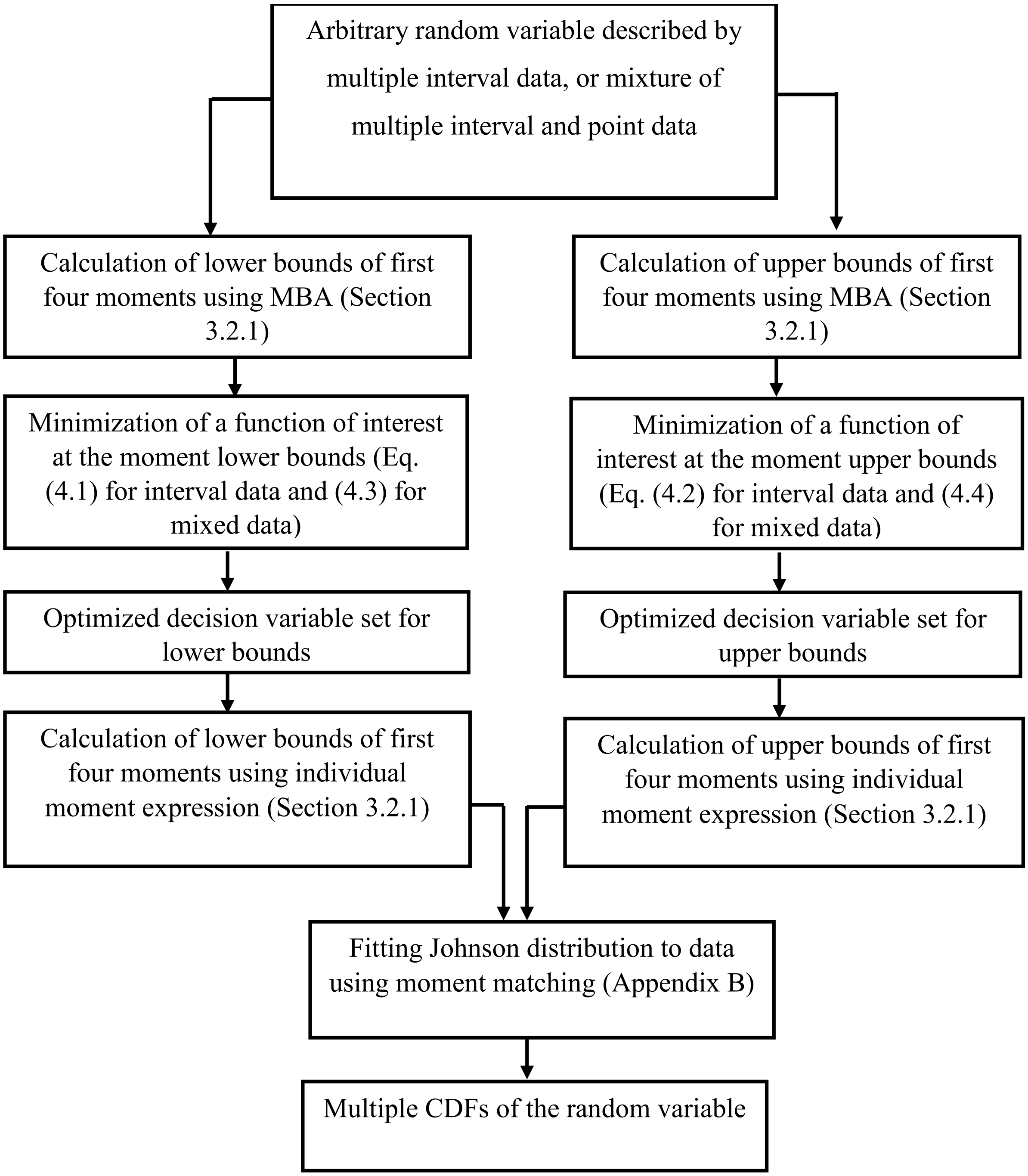


Figure 4.1: Proposed feasible moment bounding approach (FMBA) for the representation of epistemic uncertainty

In the following section, proposed FMBA is illustrated with two numerical example problems.

4.2 Numerical Examples

In this section, we demonstrate our proposed feasible moment bounding approach by two example problems. Since our proposed approach can represent a random variable defined by either multiple interval data or a mixture of multiple interval and sparse point data, this section presents two examples for these two types of problems. Example 4.1 consists of only multiple interval data, and Example 4.2 deals with both multiple interval and sparse point data.

4.2.1 Example 4a

Consider an arbitrary random variable *a* described by multiple interval data set ([0.5, 0.7], [0.3, 0.8], [0.1, 1.0]). It is required to quantify the uncertainty associated with variable *a*. For the purpose, lower and upper bounds of the first four statistical moments of uncertain variable *a* are estimated following existing MBA and our proposed FMBA. Existing MBA estimates different moment bounds from different realization sets, whereas proposed FMBA estimates all the moment bounds using single realizations set. Then, bounded Johnson distribution is fitted to data by means of moment matching approach, and resulting family of 150 sample CDFs are plotted in Figure 4.2 for each of the methods. Additionally, another uncertainty representation approach WMLE as outlined in Section 3.2.3 is also used to represent the uncertainty of variable *a*, and subsequently, resulting single representative CDF of the variable is superposed in the same figure (Figure 4.2).

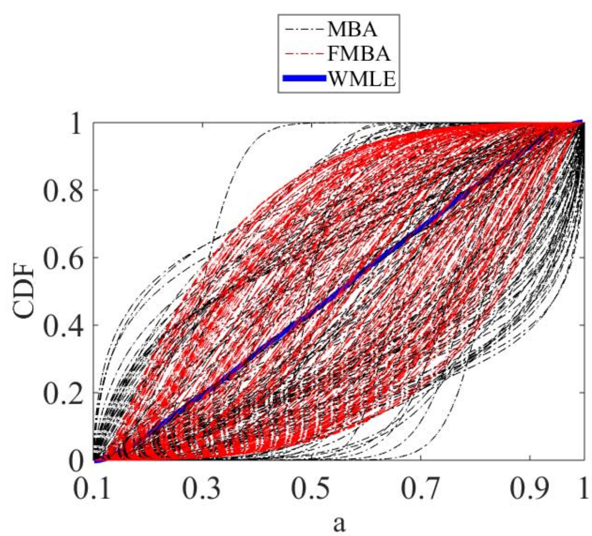


Figure 4.2: WMLE based CDF, MBA and FMBA based family of CDFs for variable *a* of Example 4a

Figure 4.2 clears up that CDFs (red) obtained using proposed FMBA are totally contained by CDFs (black) obtained using existing MBA. In other words, CDF bounds of the random variable *a* is narrower in case of proposed FMBA than that of existing MBA method. Again, single representative CDF (solid blue) of variable *a* found using WMLE approach is also successfully accommodated by CDFs (black) found using proposed FMBA method.

4.2.2 Example 4b

Random variable *b* is assumed to be described by three sets of interval ([0.4, 0.85], [0.2, 0.9], [0.0, 1.0]) and two sparse point data {0.55, 0.65}. After quantifying the uncertainty of variable *b* in a form of lower and upper bounds of first through fourth statistical moments using both MBA and FMBA, data is fitted to bounded Johnson distributions (Appendix A) following moment matching approach (Appendix B), and resulting family of sample CDFs are plotted in Figure 4.3 for both MBA and FMBA. Similar to Example 4.1, proposed FMBA based CDFs (red) of variable *b* are fully enveloped by existing MBA based CDFs (black), or CDF bounds following FMBA method is narrower than CDF bounds following MBA method.

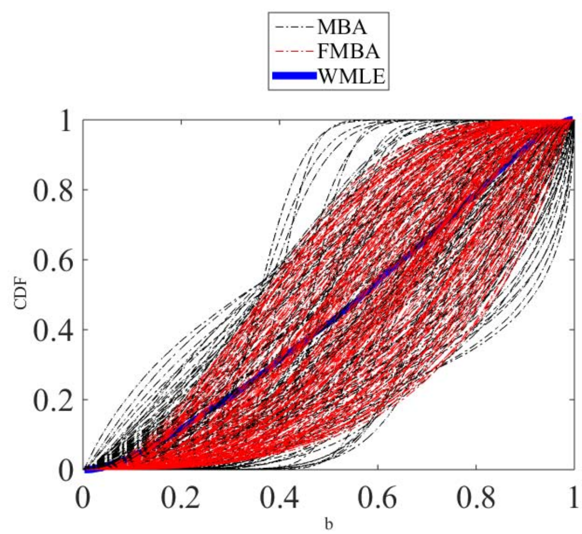


Figure 4.3: WMLE based CDF, MBA and FMBA based family of CDFs for variable *b* of Example 4b

The interpretation from the figures both in example 4a and 4b are well aligned to our claims regarding proposed feasible moment bounding approach (FMBA). Proposed FMBA can represent uncertainty of a random variable described by either multiple interval data or mixture of multiple interval and discrete point data in a form of multiple CDFs excluding the limitations of existing MBA. Now it is required to investigate either represented uncertainty using proposed FMBA can effectively and efficiently propagate through systems. The following chapter discusses the propagation of the uncertainty represented by our proposed FMBA through single disciplinary system.

Chapter 5

PROBABILISTIC PROPAGATION OF EPISTEMIC UNCERTAINTY THROUGH SINGLE DISCIPLINARY SYSTEM

Our proposed uncertainty representation method, *feasible moment bounding approach* (FMBA) can effectively represent different types of epistemic uncertainty as discussed in Chapter 4. Now, this uncertainty in the input variable needs to be propagated through the model. A number of uncertainty propagation approaches are available in literature. In this thesis, we have used a unified probabilistic approach (Zaman et al, 2011b) as outlined in Section 3.3 to propagate input variable uncertainty through single disciplinary system. In this chapter, first, we have represented the input variable uncertainty defined by either multiple interval or a mixture of multiple interval and sparse point data using the proposed FMBA, and then propagated this uncertainty through single disciplinary system by using sampling-based optimization, percentile-based optimization (PBO) and expectation-based optimization (EBO) methods. The following section illustrates the overall uncertainty representation and propagation strategy through four numerical examples, which includes three challenge problems, and compare the results with the earlier studies. The following section illustrates the overall uncertainty representation and propagation strategy through four numerical examples including three challenge problems, and compare the results with the earlier studies.

5.1 Numerical Examples

In this section, four numerical problems are solved to show how uncertainty of a random input variable propagates through a single disciplinary system. Among these, first three problems are adopted from Sandia epistemic uncertainty workshop (Oberkampf et al, 2004). Sandia workshop has proposed six challenge problems to show how to represent, aggregate and propagate epistemic, and a mixture of aleatory and epistemic uncertainty through simple single disciplinary model. All six challenge problems are designed based on either single interval or multiple interval data, and we solve three relevant challenge problems from the workshop. Additionally, since our proposed FMBA can also represent variable described by a mixture of multiple interval and sparse point data, we solve one extra problem outside the workshop to cover this extra feature of the proposed approach.

Consider an algebraic function of the form

$$y = (a + b)^a \tag{5.1}$$

where, y is the system response. Let a and b be two continuous input variables, which are independent to each other and assumed to be positive real numbers, and available information regarding a and b are summarized in Table 5.1. Our goal is to show how uncertainty in input variable a and b propagate through the model $y = (a + b)^a$ and to quantify the uncertainty in system response y assuming the model as perfect or zero model form error.

Table 5.1: Numerical values for four example problems

Problem	Input variable *a*	Input variable *b*
Example 5a	[0.1, 1.0]	([0.6, 0.8], [0.4, 0.85], [0.2, 0.9], [0, 1])
Example 5b	([0.5, 0.7], [0.3, 0.8], [0.1, 1.0])	([0.6, 0.6], [0.4, 0.85], [0.2, 0.9], [0, 1])
Example 5c	([0.5, 0.7], [0.3, 0.8], [0.1, 1.0])	Lognormally distributed with two parameter: • ([0.6, 0.8], [0.2, 0.90], [0.0, 1.0]) • ([0.3, 0.4], [0.2, 0.45], [0.1, 0.5])
Example 5d	([0.5, 0.7], [0.3, 0.8], [0.1, 1.0])	([0.4, 0.85], [0.2, 0.9], [0.0, 1.0]) and {0.55, 0.65}

5.1.1 Example 5a (Challenge Problem 2)

Example 5a is the second challenge problem from the Sandia workshop. In this problem, the uncertain input variable a is defined by single interval [0.1, 1.0], and b is described by multiple interval data ([0.6, 0.8], [0.4, 0.85], [0.2, 0.9], [0, 1]). We present the uncertainty representation of input variable a and b in the form of CDF obtained using proposed FMBA in Figure 5.1 and 5.2, respectively, and propagate these CDFs through the system model $y = (a + b)^a$ with the help of three uncertainty propagation methods – PBO, EBO and Sampling.

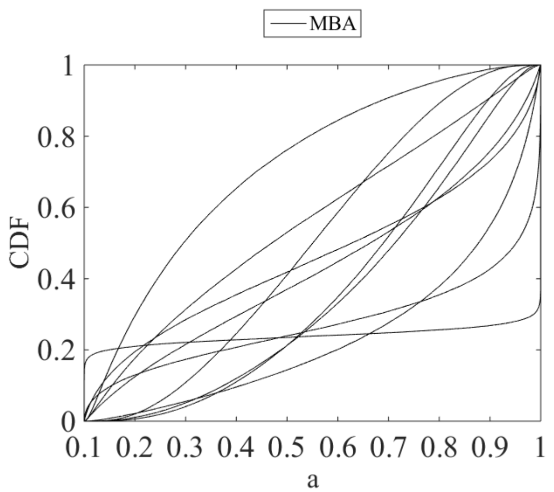


Figure 5.1: CDFs of input variable *a* of example 5a

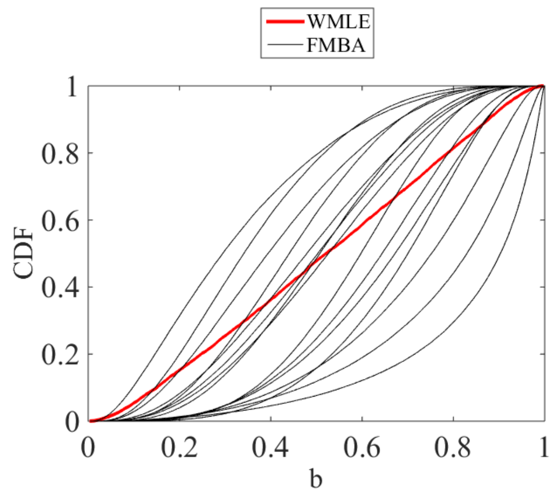


Figure 5.2: CDFs of input variable *b* of Example 5a

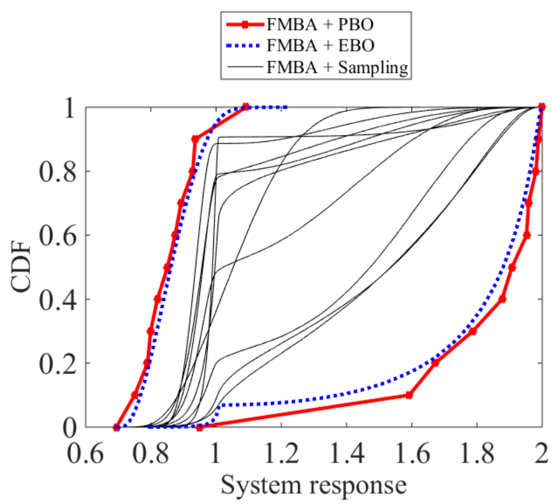


Figure 5.3: Bounds on system response CDFs of Example 5a

As discussed in Chapter 3, PBO minimizes or maximizes the system response values at different percentile points, whereas EBO minimizes or maximizes at the expected values, therefore, minimized PBO bound should always be lower than minimized EBO bound, and maximized PBO bound be higher than maximized EBO bound. For this example, this thesis has optimized the system response at 11 percentile points, and interpolated the points to get overall PBO bounds. Since in Figure 5.3, lower bound of CDFs obtained by PBO (red) is lower than that of EBO lower bound (blue), upper bound of CDFs obtained by PBO (red) is higher than that of EBO upper bound (blue), and sampling CDFs are within the bounds, it is clear that uncertainty represented by proposed FMBA can effectively propagate through systems. Now, in order to compare our proposed FMBA with the existing MBA, we plot PBO bounds in Figure 5.4 and EBO bounds in Figure 5.5 for both MBA and FMBA.

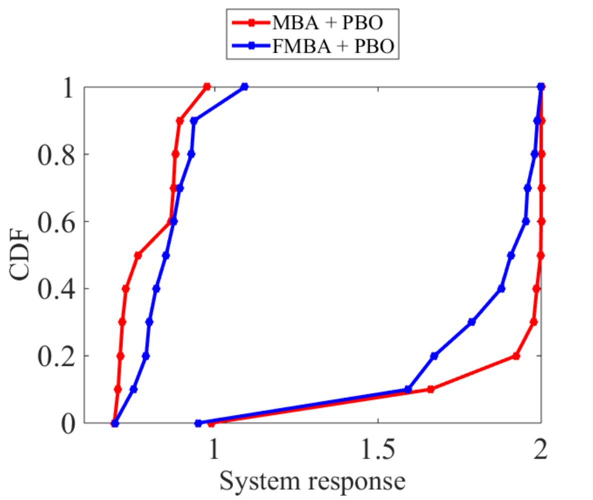


Figure 5.4: PBO based comparison of bounds on system response CDFs of Example 5a

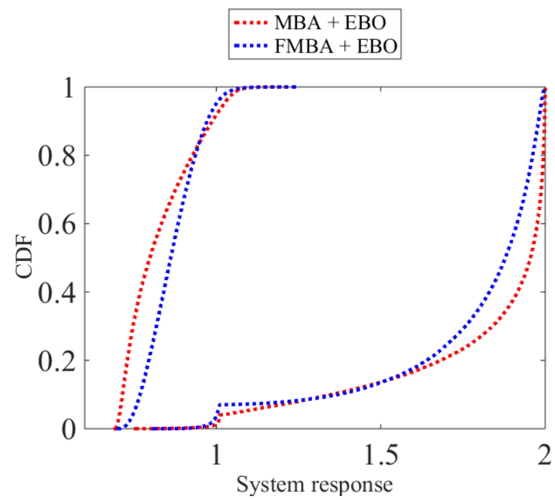


Figure 5.5: EBO based comparison of bounds on system response CDFs of Example 5a

Figure 5.4 and 5.5 shows that, proposed FMBA bounds of the system response is narrower than existing MBA bounds, which is one of the goals of our proposed FMBA. Thus, FMBA can effectively represent the interval uncertainty bypassing the limitations of existing MBA.

5.1.2 Example 5b (Challenge Problem 3)

Example 5b is the third challenge problem of the workshop. For this problem, the input variables *a* and *b* both are multiple interval data described by ([0.5, 0.7], [0.3, 0.8], [0.1, 1.0]), and ([0.6, 0.6], [0.4, 0.85], [0.2, 0.9], [0.0, 1.0]), respectively. Uncertainty in variable *a* is already represented in Figure 4.2 of Example 4a, and variable *b* has an uncertainty representation as shown in Figure 5.6. Similar to Example 5a, uncertainty in input variable *a* and *b*, represented using proposed FMBA, is propagated through Eq. (5.1) following PBO, EBO and Sampling propagation methods, and resulting bounds on system response CDFs are reported in Figure 5.7. One point is necessary to note that, this thesis has performed PBO at 11 percentile points, and interpolation of these points yield PBO bounds.

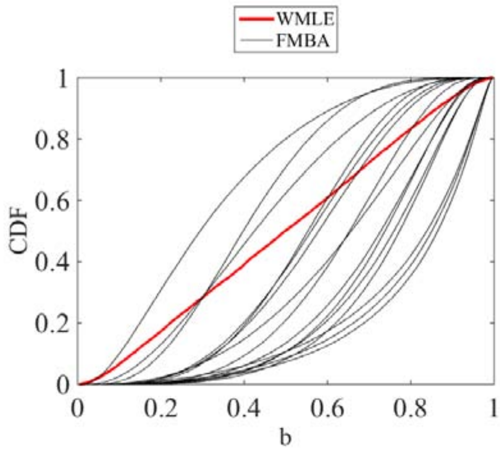


Figure 5.6: CDFs of input variable *b* of example 5b

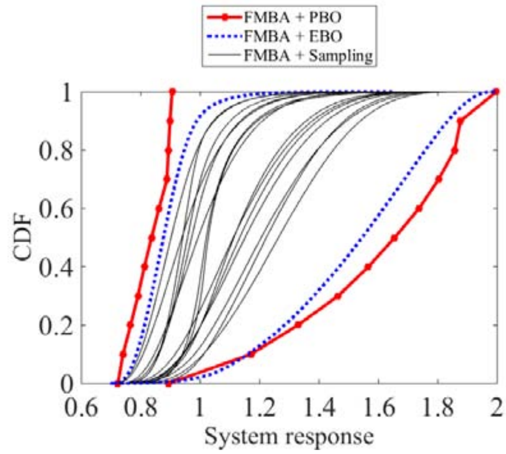


Figure 5.7: Bounds on system response CDFs of Example 5b

Again, PBO and EBO based bounds on system response *y* are plotted in Figure 5.8 and 5.9, respectively, for both proposed FMBA and existing MBA, and the figures show that, bounds obtained using FMBA (blue) is narrower than MBA (red) in each of the cases.

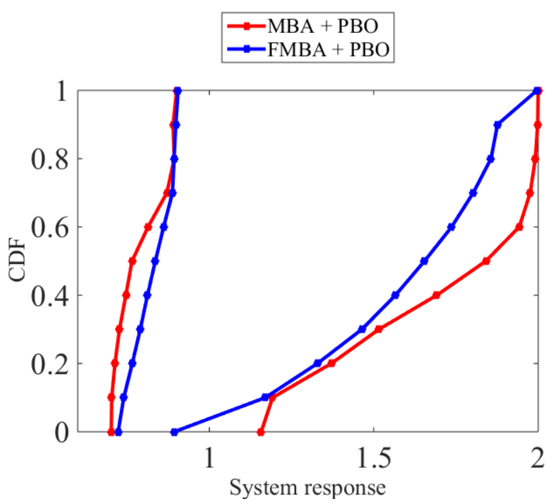


Figure 5.8: PBO based comparison of bounds on system response CDFs of Example 5b

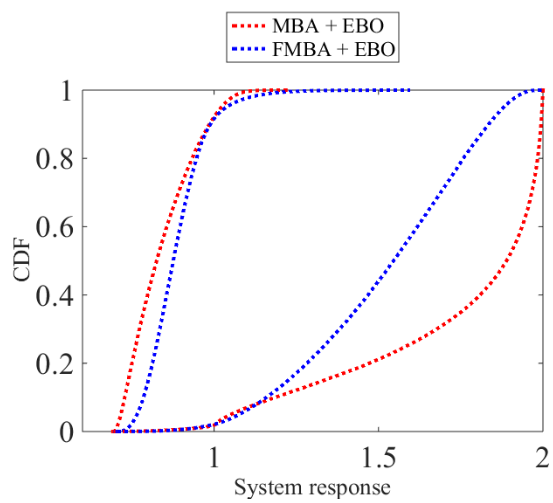


Figure 5.9: EBO based comparison of bounds on system response CDFs of Example 5b

5.1.3 Example 5c (Challenge Problem 5)

This is the fifth challenge problem of Sandia epistemic uncertainty workshop. For this problem, input variable *a* is multiple interval data ([0.5, 0.7], [0.3, 0.8], [0.1, 1.0]), and *b* is characterized by log-normal probability distribution with imprecise distribution parameters. Two parameters (mean and standard deviation) of log-normal distribution are specified by multiple intervals ([0.6, 0.8], [0.2, 0.9], [0.0, 1.0]) and ([0.3, 0.4], [0.2, 0.45], [0.1, 0.5]). After quantifying the uncertainty using proposed FMBA, sample CDFs for each of input variable *a* and two log-normal distribution parameters are shown in Figure 4.2 of Example 4a, Figure 5.6, and Figure 5.7, respectively with associated WMLE representation. Once the uncertainty of two log-normal distribution parameters are quantified, uncertainty in input variable *b* can easily be determined as shown in Figure 5.12. Then, available input variable uncertainty is propagated through system, and resulting bounds on system response CDFs are depicted in Figure 5.13.

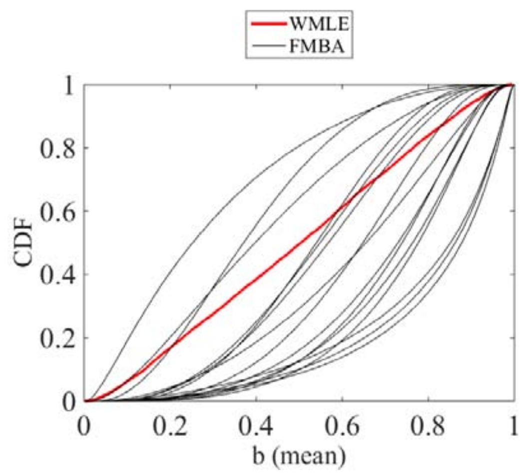


Figure 5.10: CDFs of mean of input variable *b* of example 5c

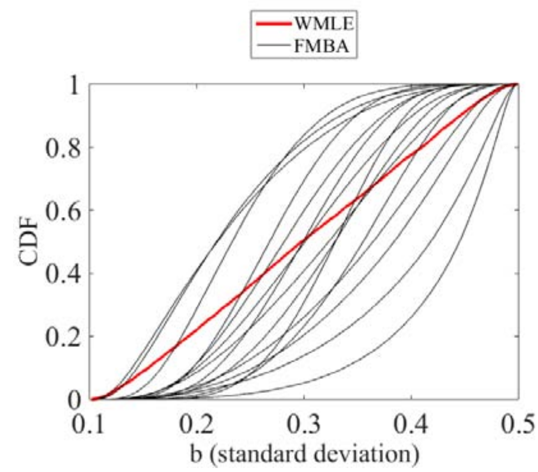


Figure 5.11: CDFs of standard deviation of input variable *b* of example 5c

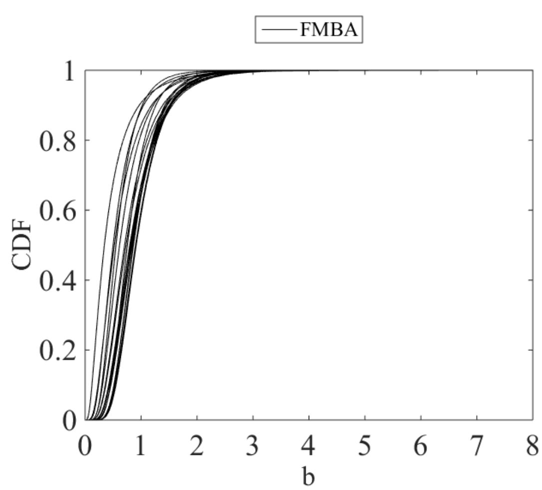


Figure 5.12: CDFs of input variable *b* of example 5c

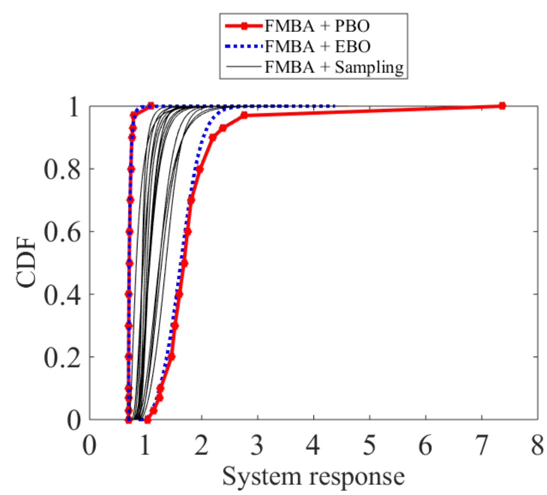


Figure 5.13: Bounds on system response CDFs of Example 5c

Figure 5.13 shows that, PBO bounds (red) are almost overlapped with EBO bounds (blue), therefore, this thesis has performed PBO at 15 percentile points instead of 11 points for this example. Since none of the PBO points in Figure 5.13 goes beyond EBO bounds, it can be said that proposed FMBA can effectively represent the uncertainty in input variable *a* and *b*. Bounds on system response CDFs obtained by both existing MBA and proposed FMBA are shown in Figure 5.14 and 5.15 for PBO and EBO based propagation. Both the figures show that, FMBA bound (blue) is narrower than MBA bound (red).

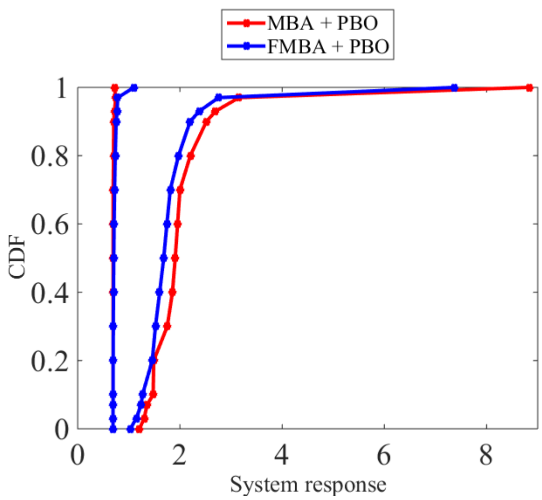


Figure 5.14: PBO based comparison of bounds on system response CDFs of Example 5c

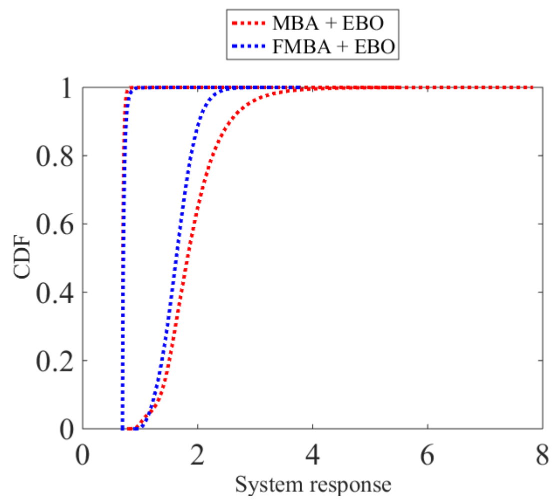


Figure 5.15: EBO based comparison of bounds on system response CDFs of Example 5c

5.1.4 Example 5d

Sandia workshop challenge problems are focused on system input variables only having single or multiple interval data. However, our proposed FMBA can also effectively represent uncertainty of input variables defined by a mixture of multiple interval and sparse point data as outlined in Section 4.1.2. Therefore, our fourth example problem is here to show the effectiveness of this additional feature of proposed FMBA.

For this problem, input variable *a* is defined by multiple interval data ([0.5, 0.7], [0.3, 0.8], [0.1, 1.0]) having uncertainty representation as shown in Figure 4.2. Input variable *b* is a random variable characterized by a mixture of three interval ([0.4, 0.85], [0.2, 0.9], [0.0, 1.0]) and two discrete points {0.55, 0.65}, and has an uncertainty representation as displayed in Figure 4.3.

Uncertainty in input variable *a* and *b*, represented using proposed FMBA, is permitted to propagate through the system $y = (a+b)^{a}$ according to sampling, PBO and EBO methods, and the resulting bounds on the system response CDFs are exhibited in Figure 5.16. Finally, to compare the efficacy of the proposed FMBA in representing input variable uncertainty, PBO bounds on system response CDFs for both MBA and FMBA are shown in Figure 5.17, and similarly EBO bounds on system response CDFs are shown in Figure 5.18. For both PBO and EBO cases, proposed FMBA bounds (blue) on system response CDFs are narrower than existing MBA bounds (red).

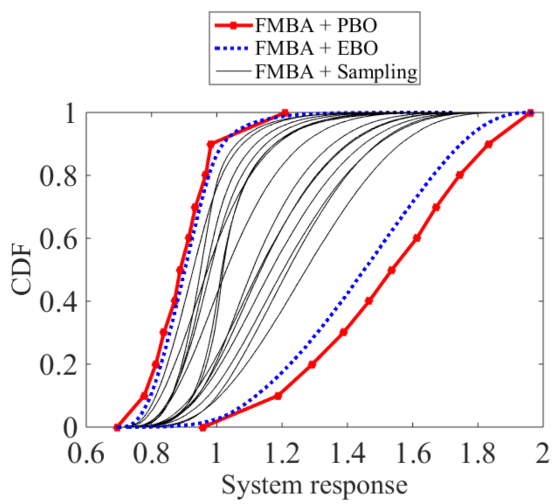


Figure 5.16: Bounds on system response CDFs of Example 5d

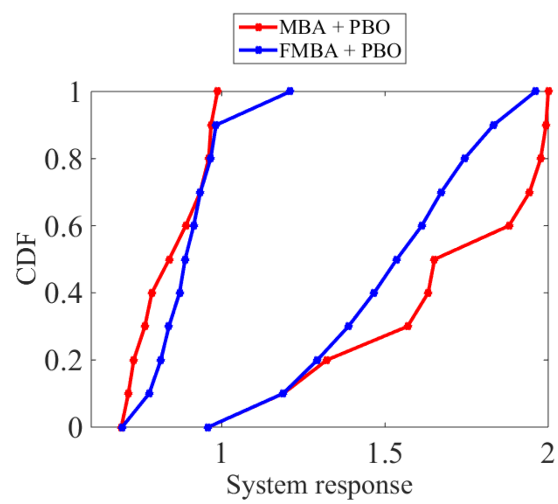


Figure 5.17: PBO based comparison of bounds on system response CDFs of Example 5d

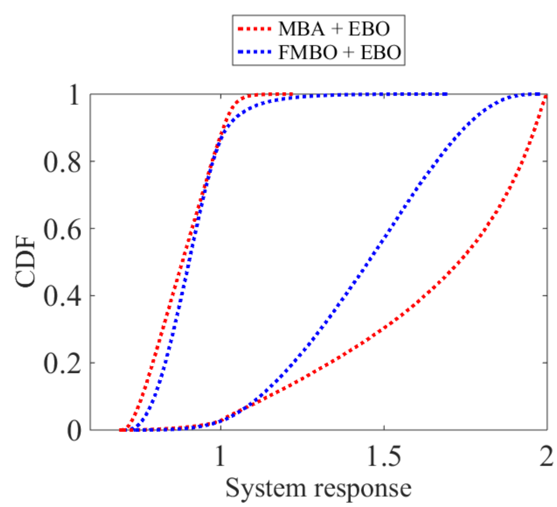


Figure 5.18: EBO based comparison of bounds on system response CDFs of Example 5d

For each of the four numerical examples, it is observed that, input variable uncertainty represented by proposed FMBA can effectively propagate through the system, and bounds on system response CDFs for both PBO and EBO are narrower in case of proposed FMBA comparing with existing MBA.

5.1.5 Computational cost of the example problems

Once the uncertainty of *a* and *b* be quantified in a form of CDFs using MBA and FMBA methods, successive propagation steps (PBO or EBO) are similar for both MBA and FMBA. In this thesis, we have observed the computational cost of this propagation step for each of the four numerical examples in terms of number of function evaluations and program run time, and summarized in Table 5.2.

Table 5.2: Computational cost (function evaluations and time) of the example problems

Problem	Percentile Points	Function evaluations		Time (s)		Function evaluations		Time (s)	
		MBA + PBO	FMBA + PBO	MBA + PBO	FMBA + PBO	MBA + EBO	FMBA + EBO	MBA + EBO	FMBA + EBO
Ex. 5a	11	25123	28660	13126	16407	531	619	182	303
Ex. 5b	11	26485	23940	16531	13237	2128	1917	707	590
Ex. 5c	15	57267	57140	31549	30809	6676	2005	2350	1315
Ex. 5d	11	27781	23502	16561	10728	874	586	455	234

Since both MBA and FMBA quantify uncertainty of the input variables in a form of family of CDFs, and these CDFs are propagated through system model, therefore, computational cost of the propagation step does not differ much due to the change in choice between existing MBA and proposed FMBA. Table 5.2 shows that, in case of Example 5a, function evaluations number and program execution times are higher for FMBA than MBA, whereas, in case of Example 5b, 5c and 5d, both number of function evaluations and program run times are smaller for proposed approach than existing one. This does not establishes the superiority of the proposed FMBA in computational competency, rather it is due to the fact that, the accuracy of PBO and EBO method depends on how accurately nonlinear optimization is being performed. To lessen the convergence related issues of nonlinear optimization problems, we have utilized genetic algorithm in this thesis. So, proposed FMBA is competitive in computational competency to existing MBA despite being able to circumvent the limitations of existing MBA.

Again, we have observed that, PBO and EBO bounds are very close for Example 5c (Figure 5.13), therefore, we have used 15 percentile points for this particular example and 11 percentile points for Example 5a, 5b and 5c, which leads larger number of function counts and longer program execution times for Example 5c than remaining three example problems.

5.1.6. Result analysis of the example problems

Researchers have come up with different approaches regarding representation, aggregation and propagation of aleatory and epistemic uncertainty through system model from the very beginning of the announcement of Sandia workshop challenge problems. Several such researchers reported the bounds on the expected values of the system response y of the challenge problems. Ferson et al. (2004) and Zaman et al. (2011b) have summarized these results in a tabular format.

Estimating moment bounds following our proposed FMBA, fitting Johnson distributions to data using matching statistical moments, and propagating the uncertainty through model by EBO methodology, this thesis has estimated the lower and upper bounds of the *expected values* (mean) of system response y for each of the four example problems, and the results are summarized in Table 5.3 with earlier studies. Example 5d, an additional example problem, introduced in this paper to show the validity of our proposed FMBA in representing uncertainty in input variables, which are described by both multiple interval and sparse point data, has also been included in the summary results in Table 5.3.

Table 5.3: Comparison of expected values of system response bounds for four example problems

Problem	Kozine and Utkin (2004)	de Cooman and Troffaes (2004)	Ferson and Hajagos (2004)	MBA based propagation (Zaman et al., 2011b)	**FMBA based propagation (this thesis)**
Ex. 5a	[0.93, 1.84]	[0.956196, 1.8]	[0.84, 1.89]	[0.6922, 2]	**[0.8596, 1.7941]**
Ex. 5b	[0.944, 1.473]	[1.04881, 1.2016]	[0.83, 1.56]	[0.6922, 2]	**[0.8932, 1.5806]**
Ex. 5c	[1.45, 2.824]	[1.54027, 2.19107]	[1.05, 3.79]	[0.6922, 8.4681]	**[0.7177, 1.664]**
Ex. 5d	-	-	-	-	**[0.8992, 1.4618]**

The results of four numerical example problems of Table 5.3 are in terms of bounds on the expected value of system output. Last column of the table is based on results found in this thesis. Though Sandia workshop challenge problems have no right or unique results (Oberkampf et al, 2004), a comparative agreement among different solutions of the same problem is noticed from the table. One additional comment can be made about proposed FMBA based bounds on expected values of system response that, these bounds can be considered as *optimal* or *best possible*, because bound widths could not be any narrower but still envelope all the sample CDFs of system output as shown in Figure 5.3 of Example 5a, 5.7 of Example 5b, 5.13 of Example 5c, and 5.16 of Example 5d.

So far, this chapter has described the propagation of different types of epistemic uncertainty through single disciplinary system. However, if the system consists of multiple disciplines with mutual coupling, uncertainty propagation through system becomes more complicated. The following chapter proposes a unified formulation for the representation and propagation of different types of uncertainty through multidisciplinary systems.

CHAPTER 6

PROPOSED UNIFIED LIKELIHOOD-BASED DECOUPLING APPROACH TO MULTIDISCIPLINARY ANALYSIS UNDER EPISTEMIC UNCERTAINTY

The success of any uncertainty representation approach is solely depended on the effective propagation of the represented uncertainty through system models. In literature, a number of methods are available to represent both aleatory and epistemic uncertainty, and to propagate those uncertainty through single or multidisciplinary system. Sankararaman and Mahadevan (2011) proposed a likelihood based epistemic uncertainty representation approach titled LBNA, which can represent a variable with epistemic uncertainty through a single nonparametric probability density function (PDF). Zaman et al. (2011a) suggested a moment bounding approach (MBA), which includes representing a variable having epistemic uncertainty through multiple representative cumulative distribution functions (CDFs). Zaman and Dey (2017) came up with another likelihood based worst-case maximum likelihood estimation (WMLE) approach to represent variable with epistemic uncertainty through a single CDF. All of these three uncertainty representation approaches are reviewed earlier in Section 3.2. Once uncertainty associated with all the input variables are in a manageable format, then this uncertainty can be propagated through multidisciplinary systems using any of the multidisciplinary analysis (MDA) methods. For this purpose, this thesis uses likelihood-based approach for multidisciplinary analysis (LAMDA) method which is discussed in section 3.4.

Figure 6.1 below shows a flowchart, to aggregate the representation and propagation of epistemic uncertainty through multidisciplinary systems. This thesis is interested in two highlighted boxes of Figure 6.1. The box in the left side is used to show uncertainty representation approaches, and the right side box is to present MDA methodologies. After performing multidisciplinary analysis in right side box, the PDFs of the coupling variables will be available, and can be utilized in subsequent subsystem and system level analyses.

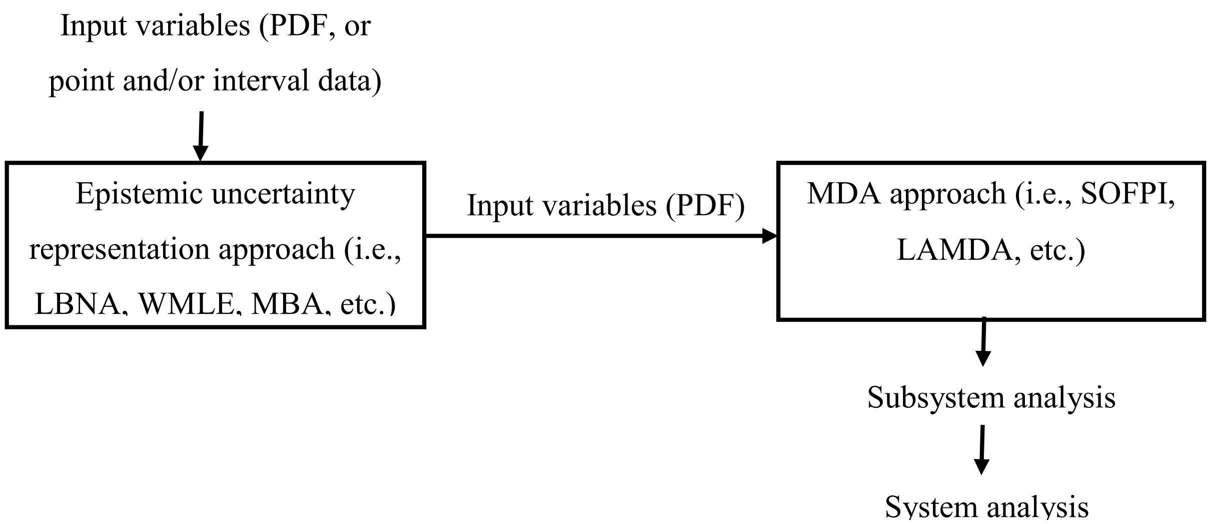
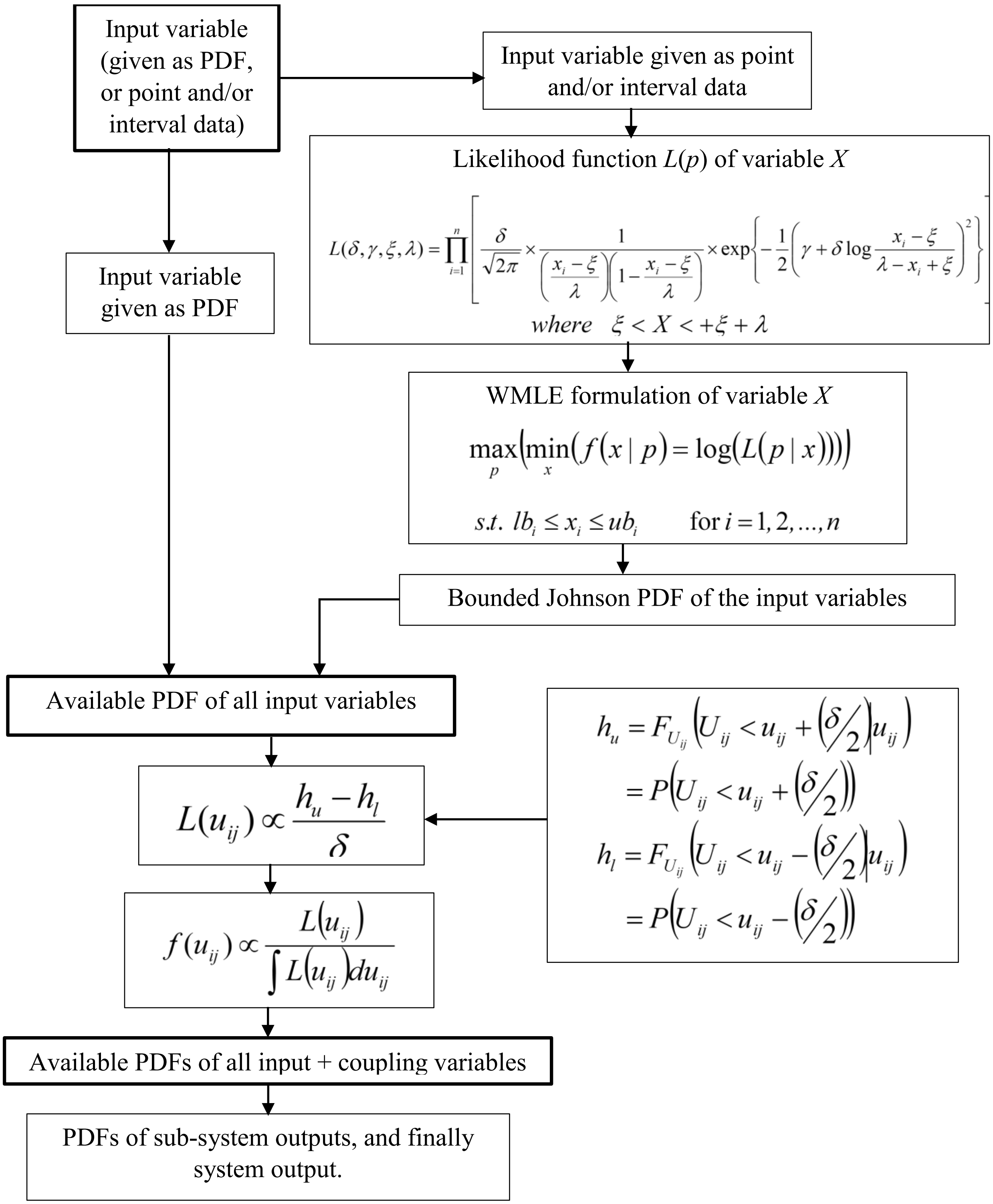


Figure 6.1: Representation and propagation of epistemic uncertainty through MDA systems

6.1 Proposed Approach for the Representation and Propagation of Epistemic Uncertainty through Multidisciplinary System

Liang et al. (2015) represented the epistemic uncertainty in the input variable using likelihood-based nonparametric approach (LBNA) developed by Sankararaman and Mahadevan (2011), and performed the multidisciplinary analysis (MDA) using likelihood-based approach for multidisciplinary analysis (LAMDA) method proposed by Sankararaman and Mahadevan (2012). As mentioned earlier, their likelihood-based LBNA estimates a nonparametric PDF of an uncertain variable, and propagation of this nonparametric PDF through systems is difficult and computationally expensive due to the shortage of information about distribution type and distribution parameters. Therefore, this thesis has proposed a likelihood based unified approach for representation and propagation of epistemic uncertainty in multidisciplinary system. The proposed approach as illustrated in Figure 6.2 uses the worst-case maximum likelihood estimation (WMLE) method to represent the input variable containing epistemic uncertainty through a parametric CDF, and LAMDA method to accomplish multidisciplinary analysis. Since proposed approach has utilized two already discussed methods (WMLE and LAMDA) in a single framework, the significance of the notations used in Figure 6.2 are same as the ones discussed in Sections 3.2.3 and 3.4.2 of Chapter 3.



$$L(\delta,\gamma,\xi,\lambda)=\prod_{i=1}^{n}\left[\frac{\delta}{\sqrt{2\pi}}\times\frac{1}{\left(\frac{x_i-\xi}{\lambda}\right)\left(1-\frac{x_i-\xi}{\lambda}\right)}\times\exp\left\{-\frac{1}{2}\left(\gamma+\delta\log\frac{x_i-\xi}{\lambda-x_i+\xi}\right)^2\right\}\right]$$

$$\max_p\left(\min_x\left(f(x\mid p)=\log(L(p\mid x))\right)\right)$$

$$L(u_{ij})\propto\frac{h_u-h_l}{\delta}$$

$$f(u_{ij})\propto\frac{L(u_{ij})}{\int L(u_{ij})du_{ij}}$$

$$h_u=F_{U_{ij}}\left(U_{ij}<u_{ij}+\left(\delta/2\right)\middle|u_{ij}\right)=P\left(U_{ij}<u_{ij}+\left(\delta/2\right)\right)$$

$$h_l=F_{U_{ij}}\left(U_{ij}<u_{ij}-\left(\delta/2\right)\middle|u_{ij}\right)=P\left(U_{ij}<u_{ij}-\left(\delta/2\right)\right)$$

Figure 6.2: Proposed unified decoupling approach to represent and propagate epistemic uncertainty through MDA system

The proposed formulation of Figure 6.2 has two sequential steps. The first step is to represent the uncertainty of the uncertain input variables using WMLE approach, and the second step is to propagate the uncertainty through multidisciplinary system model using LAMDA approach. The input variables of a multidisciplinary system may be available as probability density function (PDF), or discrete point data and/or interval data. Among these, input variables having probability distributions are ready to be used in uncertainty propagation step of the formulation, however, input variables described by multiple intervals or mixture of multiple intervals and discrete point data need to go through the first step. To represent the uncertainty of the input variables characterized by multiple intervals or mixture of multiple intervals and points data, WMLE formulates a likelihood function of uncertain variable X and maximizes the minimum (worst case) of the likelihood to get four parameter of the bounded Johnson distributions. Once the probability distributions of the uncertain input variables can be estimated, then all the input variables come to a same level and become qualified to be used in uncertainty propagation of the second step using LAMDA method, which is discussed in Section 3.4. LAMDA estimates the probability density function of the coupling variables, and to do so, it calculates the conditional likelihood at some discrete points of the coupling variable, and subsequently evaluates the PDF of the coupling variable. Once the PDF of the coupling variable be available, multidisciplinary coupled system turns into multidisciplinary partially decoupled system, and allows to propagate these PDFs of the coupling variables as well as input variables through higher level subsystem and system analyses to get the PDFs of the desired output.

By estimating a parametric distributions of the variable having different types of epistemic uncertainty, the proposed method can bypass the complications regarding nonparametric distributions of the existing likelihood-based MDA methodology. However, it is needed to make clear that, our unified proposed formulation utilizes WMLE method proposed by Zaman and Dey (2017) in place of LBNA method proposed by Sankararaman and Mahadevan (2011) to represent epistemic uncertainty, but multidisciplinary analysis LAMDA is common in both existing (developed by Liang et al., 2015) and proposed formulations. Therefore, to make the illustration easily understandable, we have used "WMLE + LAMDA" and "LBNA + LAMDA" to indicate the proposed and existing approaches, respectively. "WMLE + LAMDA" signifies that, epistemic uncertainty of the variable is represented by WMLE approach, and multidisciplinary analysis is performed according to LAMDA approach, and similar interpretation holds true for “LBNA + LAMDA”.

This following two sections have demonstrated the proposed formulation with a numerical MDA problem, and an engineering MDA problem. Both numerical and engineering problems have been used to compare the proposed “WMLE + LAMDA” approach with the existing “LBNA + LAMDA” approach.

6.2 Mathematical MDA Problem

This example problem is considered from Liang et al. (2015) with minor modification. Since model error uncertainty is out of scope of this thesis, it excludes model error related terms in the problem formulation as shown in Figure 6.3. There are three individual disciplinary analysis, among them “Analysis 1” and “Analysis 2” are mutually coupled by two feedback coupling variables u_{12} and u_{21}. Subsystem level output g_1 and g_2 are the inputs to “Analysis 3”, and f is the output of the same analysis. To calculate system level output f, it is required to calculate subsystem level outputs g_1 and g_2. Again, to calculate g_1 and g_2, coupled analyses need to be decoupled. To decouple the coupled analyses, estimation of at least one of the coupling variables is a must. In this example problem, this thesis has estimated the probability density function (PDF) of the coupling variable u_{12}. Once the PDF of u_{12} is estimated, the coupling between Analysis 1 and Analysis 2 can be solved easily, and subsequent subsystem and system level analyses can easily be converted to feedforward analysis from complex feedback coupling analysis as discussed in Chapter 3.

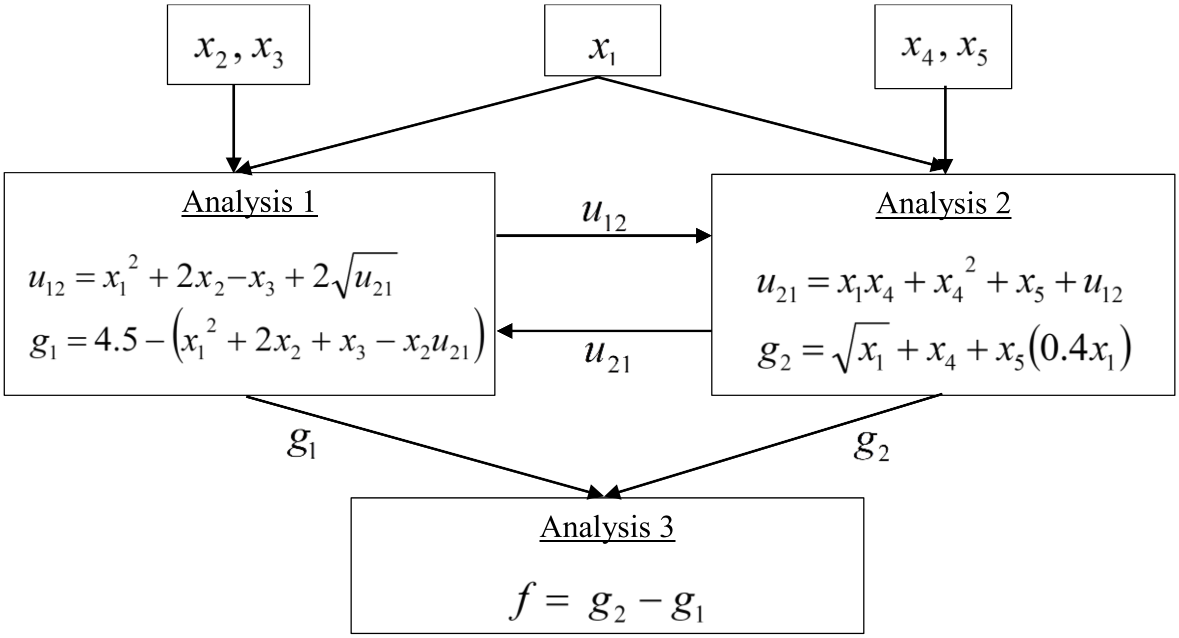


Figure 6.3: Mathematical Multidisciplinary Analysis (MDA) Problem

Among five input quantities, $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)$, x_1 is a shared input to both Analysis 1 and Analysis 2, while rest of the quantities are local to their respective analysis. For some of these variables, probability distributions are available. Input variables x_1, x_2 and x_3 are normally distributed, x_4 is log-normally distributed, and on the other hand, x_5 is a mixture of both sparse point and multiple interval data. All of the input variables are assumed to be uncertain, and they are not mutually correlated. Distribution information and data of all the input variables are listed in Table 6.1.

Table 6.1: Data of five input variables for mathematical MDA problem

Input variable	Type	Data
x_1	Normal distribution	Mean: 1, Std.: 0.1
x_2	Normal distribution	Mean: 1, Std.: 0.1
x_3	Normal distribution	Mean: 1, Std.: 0.1
x_4	Lognormal distribution	Mean: 1, Std.: 0.1
x_5	Mixed	Mixture of three intervals ($[0, 2], [0.02, 1.97], [0.14, 1.89]$) and two point data $\{0.99, 1.02\}$

As shown in Table 6.1, the first four input variables have probabilistic information about their respective probability distribution, but the last input variable x_5 is represented only by three intervals and two specific point data. As discussed earlier, first four input variables having probabilistic information can be used directly in the uncertainty propagation through multidisciplinary system. Since input variable x_5 is characterized by a mixture of three intervals and two point data, uncertainty associated with x_5 needs to be represented first. To represent the epistemic uncertainty in variable x_5, proposed formulation has used WMLE uncertainty representation approach as outlined in Section 3.2 of Chapter 3.

Existing formulation proposed by Liang et al. (2015) has used LBNA method to represent the uncertainty of the input variable. However, as mentioned before, this approach provides a nonparametric probability distribution of variable x_5, and propagating this nonparametric PDF through multidisciplinary system is difficult and computationally expensive. Therefore, the proposed unified likelihood-based approach has utilized the worst-case maximum likelihood estimation (WMLE) method to represent the uncertainty of variable x_5; this method can represent an uncertain variable through a parametric CDF, and thus allow efficient uncertainty propagation.

Uncertainty representation of variable x_5 is displayed in Figure 6.4 as cumulative distribution functions (CDFs) following LBNA, WMLA and moment bounding approach (MBA). LBNA method provides a nonparametric PDF of x_5, and nonparametric PDF is not furnished with information regarding distribution type and distribution parameter. Therefore, as in Sankararaman and Mahadevan (2011), this thesis has used first-order reliability method (FORM) to convert nonparametric PDF to nonparametric CDF. FORM method calculates discrete CDF values at some sampled points. Through interpolation of those discrete CDF values, the overall CDF of the variable x_5 can be approximated, and for that reason, the CDF in Figure 6.4 by LBNA method is little bit irregular. However, it is clear from Figure 6.4 that, CDF obtained using WMLE method is well aligned with the CDF found by LBNA method, and also falls within MBA multiple CDFs.

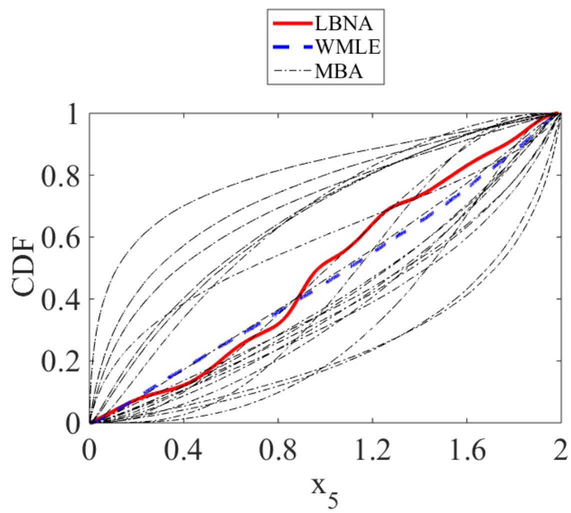


Figure 6.4: CDFs of input variable x_5

After representing the epistemic uncertainty in variable x_5 by above mentioned methods, it is time to propagate the uncertainty through multidisciplinary system by performing multidisciplinary analysis (MDA). This thesis has used LAMDA approach for this purpose.

While existing formluation (LBNA + LAMDA) represents the uncertainty in x_5 by LBNA method (nonparametric probability distribution) and propagates the uncertainty of the input variables through mustidisciplinary system using LAMDA approach, proposed formulation (WMLE + LAMDA) uses WMLE method (parametric probability distribution) to represent the uncertainty in x_5 and LAMDA approach to propagate the uncertainty through system, and resulting PDFs of coupling variable u_{12} using both existing and proposed formulations are depicted in Figure 6.5.

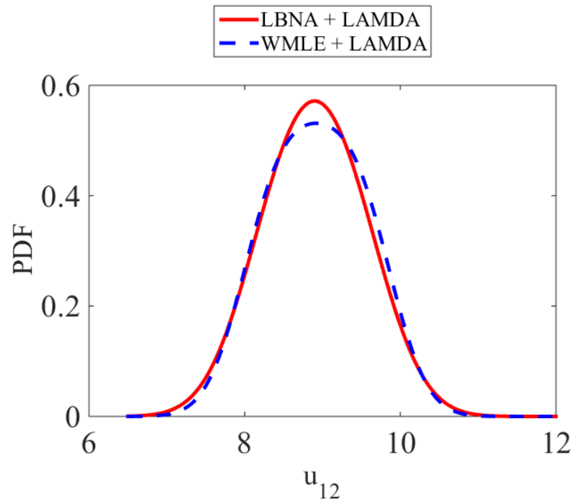


Figure 6.5: PDFs of coupling variable u_{12}

It is clear form Figure 6.5 that PDFs of the coupling variable u_{12} obtained by the existing "LBNA + LAMDA" and proposed "WMLE + LAMDA" formulations are almost similar. Since the PDF of one of the coupling variables, u_{12} is available as shown in Figure 6.5, the PDF of another coupling variable u_{21} can easily be found by simple one-directional propagation as discussed in Section 3.4, and the results are displayed in Figure 6.6.

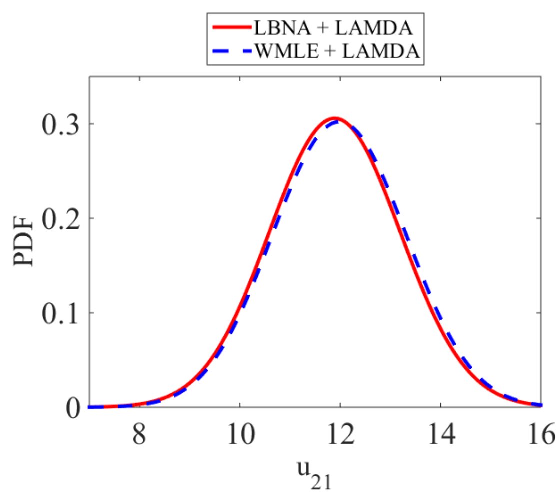


Figure 6.6: PDFs of coupling variable u_{21}

So far, the PDFs of two coupling variable u_{12} and u_{21} are available as shown in Figures 6.5 and 6.6. Then the PDFs of the coupling variables as well as the PDFs of the input variables are propagated through subsystem and system equation to obtain subsystem level output g_1 and system level output f, and resulting PDFs of g_1 and f are plotted in Figures 6.7 and 6.8, respectively.

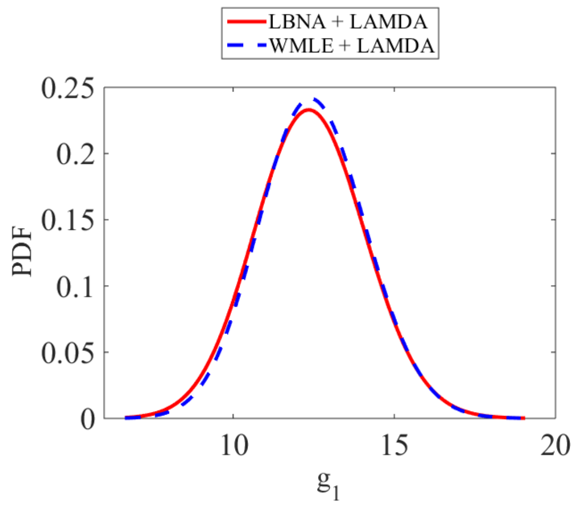


Figure 6.7: PDFs of subsystem output g_1

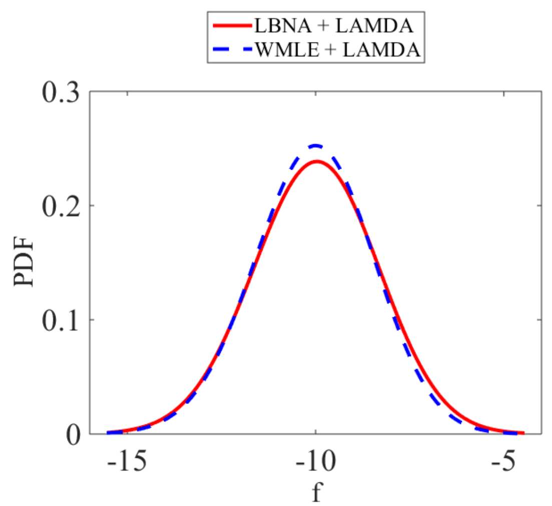
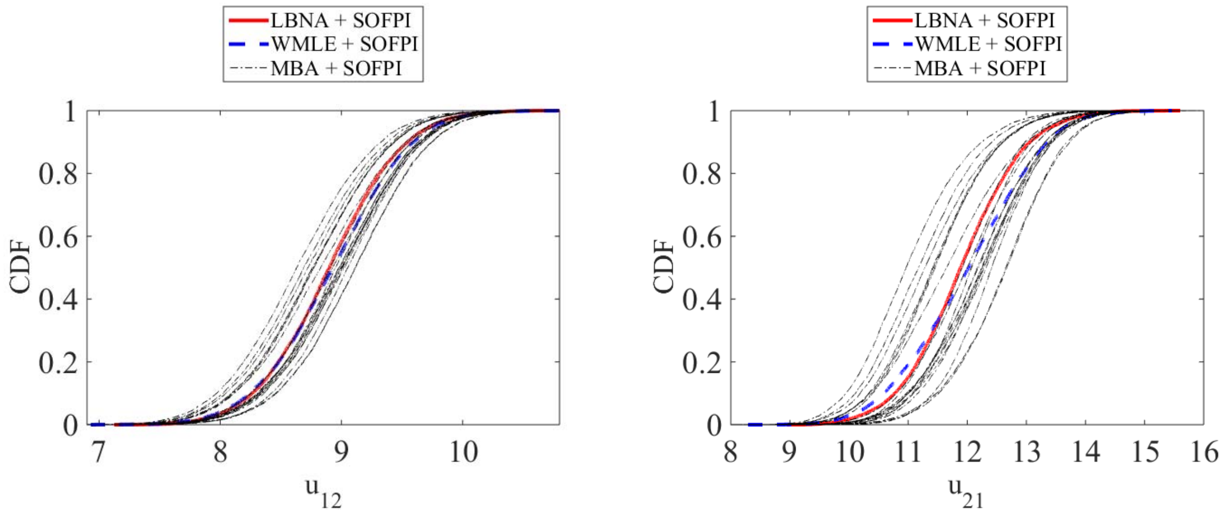


Figure 6.8: PDFs of system output f

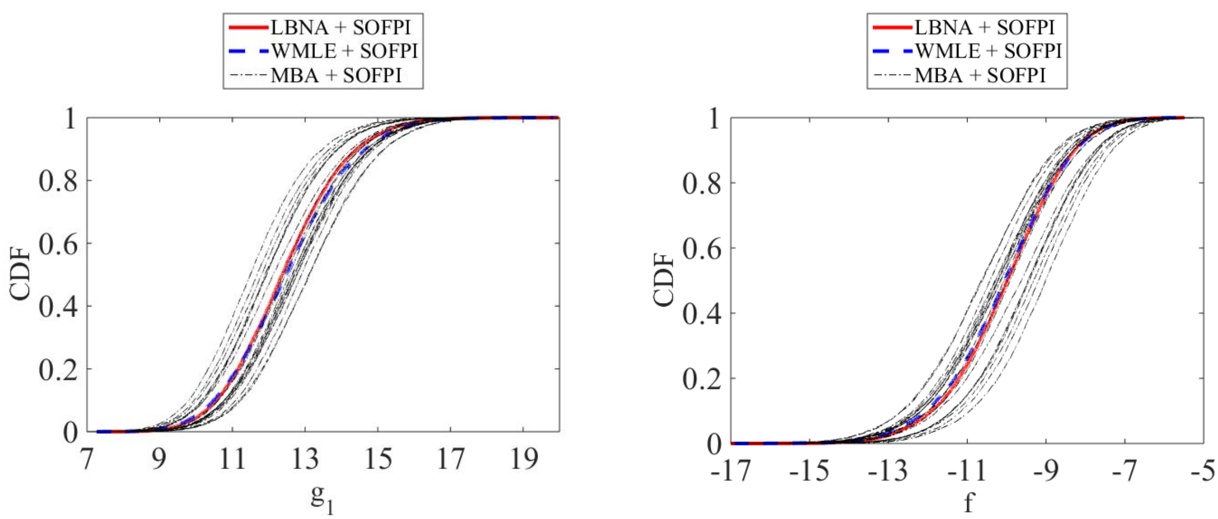
The PDFs of the two coupling variables u_{12} and u_{21}, subsystem output g1 and overall system output f obtained using existing formulation (LBNA + LAMDA) and proposed formulation (WMLE + LAMDA) are already compared in earlier figures. Additionally, we compare the CDFs of u_{12}, u_{21}, g_1 and f obtained by representing the uncertainty in input variable x_5 using LBNA, WMLE and MBA, and propagating the uncertainty through multidisciplinary systems using sampling-based

SOFPI method for each of the representation method, and the resulting CDFs are depicted in Figure 6.9 and 6.10.



(*a*) (*b*)

Figure 6.9: Uncertainty propagation-based comparison of CDFs of coupling variable (*a*) u_{12}, and (*b*) u_{21} for different uncertainty representation approaches



(*a*) (*b*)

Figure 6.10: Uncertainty propagation-based comparison of CDFs of (*a*) system level output g_1, and (*b*) overall system output f for different uncertainty representation approaches

The Statistics (Mean and Standard deviation) of two coupling variables u_{12} and u_{21} are summarized in Table 6.2 obtained using existing and proposed formulations.

Table 6.2: Mean and Standard deviation of coupling variables u_{12} and u_{21}

Uncertainty representation method	Uncertainty propagation method	Coupling variables			
		u_{12}		u_{21}	
		Mean	Standard deviation	Mean	Standard deviation
LBNA	LAMDA	8.9047	0.4519	11.8965	0.6987
WMLE	LAMDA	8.9192	0.5312	11.9828	0.8432

Similarly, the mean and standard deviation of subsystem output g_1 and overall system output f are listed in Table 6.3.

Table 6.3: Mean and Standard deviation of sub-system output g_1 and overall system output f

Uncertainty representation method	Uncertainty propagation method	Sub-system output g_1		System output f	
		Mean	Standard deviation	Mean	Standard deviation
LBNA	LAMDA	12.4007	1.2271	-10.0052	1.1681
WMLE	LAMDA	12.4861	1.3156	-10.0658	1.2243

The results from Table 6.2 and 6.3 indicate that, sinstead of being fundamentally dissimilar, existing and proposed formulations yield almost the same statistics for the coupling variables, subsystem and system level outputs.

For this mathematical MDA problem, number of individual disciplinary analysis (DA) and execution time for existing and proposed formulations are listed in Table 6.4.

Table 6.4: Computational effort of different uncertainty representation and propagation methods to estimate the PDF of the coupling variable u_{12}.

Uncertainty representation method	Uncertainty propagation method	Number of disciplinary analysis (DA)	Execution time (s)
LBNA	LAMDA	1,098	27.89
WMLE	LAMDA	1,030	21.69

The following remarks can be made from Table 6.4.

- Existing formulation (LBNA + LAMDA) uses LBNA method to represent the epistemic uncertainty in the input variable x_5, which estimates the nonparametric probability density function of x_5. Again, proposed formulation (WMLE + LAMDA) uses WMLE method to represent x_5, which evaluates the parametric CDF of x_5. For uncertainty propagation through multidisciplinary system, both existing and proposed formulations use LAMDA method. For this mathematical example problem, the proposed formulation requires less number of disciplinary analyses than existing formulation by 6.19%, and also proposed formulation requires lower program execution time than existing formulation by 22.23%.
- LAMDA method is common in both existing and proposed formulations. It estimates the PDF values at some discrete points of the coupling variable u_{12}, and interpolation of these PDF values approximates the overall PDF of u_{12}. Therefore, number of discrete points is a major issue in the accuracy of PDF estimation in case of LAMDA. More discrete points confirms more accurate and smoother PDF. However, increasing number of discrete points increases the computational effort (number of disciplinary analysis required and program run time). 15 uniformly distributed discrete points are considered in this thesis to approximate the PDF of the coupling variable u_{12} for both existing and proposed formulations.

Similar to the mathematical MDA problem of this section, the following section illustrates an engineering MDA problem to demonstrate our proposed formulation.

6.3 Engineering MDA Problem (Hypothetical Fire Detection Satellite Model)

This problem was primarily illustrated by Wertz and Larson (1999). Fire Detection Satellite (FireSat) is a hypothetical but realistic spacecraft. It consists of huge number of subsystems, and subsystems are coupled by complex feed-forward and feedback coupling variables. Detecting, identifying, and monitoring forest fires in near real time are the fundamental goal of FireSat. It is aimed to carry a large sized accurate optical sensor (length: 3.2 m, weight: 720 kg and angular resolution: 8.8e-07 radians). Ferson et al. (2009) and Zaman (2010) considered the modified version of the Firesat problem.

This thesis have used a simplified subset of FireSat subsystems considered by Zaman (2010) based on Ferson et al. (2009). Simplified FireSat spacecraft has three subsystems: *i*) Orbit Analysis, *ii*) Attitude Control, and *iii*) Power subsystem, as shown in Figure 6.11.

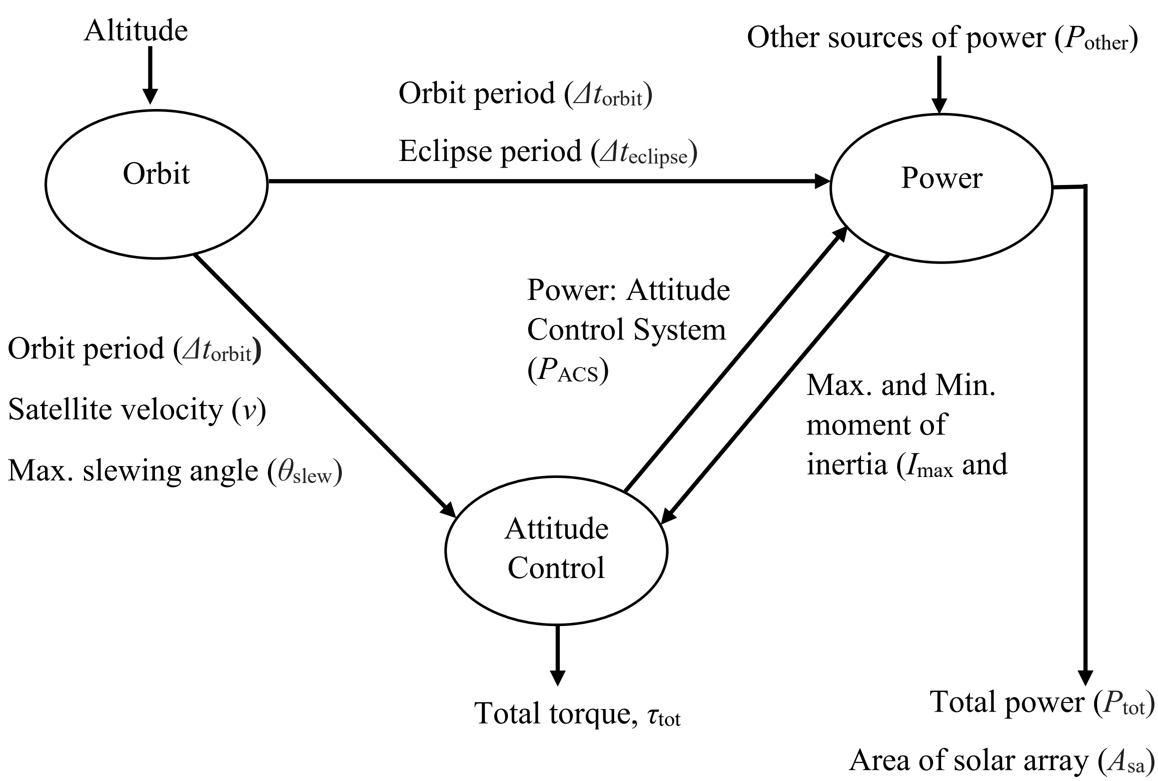


Figure 6.11: Three-discipline fire detection satellite

Figure 6.11 indicates that, the Orbit subsystem is connected to both Power and Attitude control subsystems in forward direction, but Power and Attitude control subsystems are mutually coupled through three coupling variables P_{ACS}, I_{max} and I_{min}, which made FireSat problem a complex multidisciplinary analysis (MDA) problem.

This thesis has considered a particular satellite structure holding two solar panels extended out from the main spacecraft body. The length and width of each of these panel are L and W respectively, and panel inner edge is constrained to maintain a distance L from the centerline of the satellite main body as sketched in Figure 6.12 according to Ferson et al (2009).

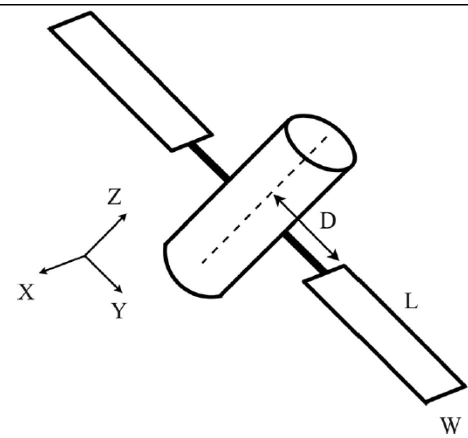


Figure 6.12: Schematic diagram for the spacecraft solar array

The functional relationships among three individual disciplines of FireSat are shortly described below.

Subsystem 1: The Orbit Subsystem

The inputs to the orbit subsystem are: orbit altitude (H), earth's radius (R_E), target diameter (ϕ_{target}), and earth's gravity constant (μ). The outputs of this subsystem are illustrated through Eqs. (6.1) to (6.4).

The satellite velocity:
$$v = \sqrt{\frac{\mu}{R_E + H}} \tag{6.1}$$

The orbit period:
$$\Delta t_{orbit} = 2\pi\sqrt{\frac{(R_E + H)^3}{\mu}} = \frac{2\pi(R_E + H)}{v} \tag{6.2}$$

The maximum eclipse time:

$$\Delta t_{eclipse} = \frac{\Delta t_{orbit}}{\pi} \arcsin\left(\frac{R_E}{R_E + H}\right) \tag{6.3}$$

The maximum slewing angle:

$$\theta_{slew} = \arctan\left(\frac{\sin\left(\frac{\phi_{target}}{R_E}\right)}{1 - \cos\left(\frac{\phi_{target}}{R_E}\right) + \frac{H}{R_E}}\right) \tag{6.4}$$

Subsystem 2: The Attitude Control Subsystem

The Attitude Control Subsystem has 23 inputs listed in Table 6.5.

Table 6.5: List of 23 inputs to the Attitude Control Subsystem of FireSat

• orbit altitude (H), • earth's radius (R_E), • minimum and maximum moment of inertia of the spacecraft obtained in Power subsystem (I_{max} and I_{min}), • the deviation of major moment axis from local vertical (θ), • moment arm for the solar radiation torque which is the distance between the center of gravity of the spacecraft and the center of the solar pressure (L_{sp}), • average solar flux (F_s), • Max. slewing angle (θ_{slew}),	• surface reflectivity or reflectance factor (q), • surface area off which solar radiation is reflected (A_s), • Slewing time (Δt_{slew}), magnetic moment of the Earth in Am^2 (M), • residual dipole of the spacecraft (R_D), • moment arm for aerodynamic drag torque or the distance between the spacecraft's center of gravity and the center of the aerodynamic pressure (L_a), • atmospheric density (ρ), • light speed (c),	• sun incidence angle or the angle at which the sun radiation hits the spacecraft surface (i), • drag coefficient (C_d), • cross sectional surface area in the direction of flight (A), • satellite velocity (v), • rotation velocity of reaction wheel (ω_{max}), • number of reaction wheels (n), and • holding power (P_{hold}), i.e., the power required to maintain the constant velocity (ω_{max}).

Subsystem output total torque (τ_{tot}) is the maximum of slewing torque (τ_{slew}) and disturbance torque (τ_{dist}), and its value is calculated using below mentioned equations.

$$\tau_{tot} = \max(\tau_{slew}, \tau_{dist}) \tag{6.5}$$

$$\tau_{slew} = \frac{4\theta_{slew}}{(\Delta t_{slew})^2} I_{\max} \tag{6.6}$$

$$\tau_{dist} = \sqrt{\tau_g^2 + \tau_{sp}^2 + \tau_m^2 + \tau_a^2} \tag{6.7}$$

$$\tau_g = \frac{3\mu}{2(R_E + H)^3} | I_{\max} - I_{\min} | \sin(2\theta) \tag{6.8}$$

$$\tau_{sp} = L_{sp} \frac{F_s}{c} A_s (1+q) \cos(i) \tag{6.9}$$

$$\tau_m = \frac{2MR_D}{(R_E + H)^3} \tag{6.10}$$

$$\tau_a = \frac{1}{2} L_a \rho C_d A v^2 \tag{6.11}$$

Where, τ_g, τ_{sp}, τ_m and τ_a are the torques due to gravity gradient, solar radiation, magnetic field interaction, and aerodynamic drag, respectively. Two inputs I_{max} and I_{min} of attitude control subsystem are the output of power subsystem. Therefore, as long as the values of I_{max} and I_{min} are not available from power subsystem, it is not possible to perform individual disciplinary analysis of attitude control subsystem. Again, attitude control subsystem produces another variable P_{ACS} as in Eq. (6.12), which is the input to power subsystem, and thus attitude control and power subsystems are coupled by three coupling variables (I_{max}, I_{min} and P_{ACS}).

$$P_{ACS} = \tau_{tot} \omega_{\max} + nP_{hold} \tag{6.12}$$

Subsystem 3: The Power Subsystem

16 inputs to the power subsystem are summarized in Table 6.6.

Table 6.6: List of 16 inputs to the Power Subsystem of FireSat

• attitude control power (P_{ACS}), • other sources of power (P_{other}), • orbit period (**Δt_{orbit}**), • eclipse period (**$\Delta t_{eclipse}$**), • sun incidence angle (i), • inherent degradation of the array (I_d), • average solar flux (F_s),	• power efficiency (η), • lifetime of the spacecraft (L_T), • degradation in power production capability in % per year (ε_{deg}), • length to width ratio of solar array (r_{lw}), • number of solar arrays (n_{sa}),	• average mass density of solar arrays (ρ_{sa}), • thickness of solar panels (t), • distance between the panels (D), and • moments of inertia of the main body of the spacecraft (I_{bodyX}, I_{bodyY} and I_{bodyZ}).

There are two major outputs of this subsystem. These are the total power (P_{tot}) and the area of the solar array (A_{sa}). In simplified FireSat formulation, attitude control subsystem is the only consumer that consumes power explicitly, and all other sources of power consumption are aggregated into term P_{other}. Power from attitude control subsystem and power from other sources are arithmetically summed to obtain the total power P_{tot} as in Eq. (6.13).

$$P_{tot} = P_{ACS} + P_{other} \tag{6.13}$$

Now consider, the required power for the spacecraft during eclipse and daylight be P_e and P_d, and the length of time period spent per orbit in eclipse and daylight be T_e and T_d, respectively. In this thesis, it is assumed that, spacecraft's required power during both eclipse and daylight are exactly equal to the total power $P_e = P_d = P_{tot}$. This thesis also assumes that $T_e = \Delta t_{eclipse}$ and $T_d = \Delta t_{orbit} - T_e$. Based on these assumptions required power output P_{sa} is calculated as Eq. (6.14).

$$P_{sa} = \frac{\left(\frac{P_e T_e}{0.6} + \frac{P_d T_d}{0.8}\right)}{T_d} \qquad (6.14)$$

At the beginning and end of life, the array's power production capacities can be computed according to Eqs. (6.15) and (6.16).

$$P_{BOL} = \eta F_s I_d \cos(i) \qquad (6.15)$$

$$P_{EOL} = P_{BOL}\left(1 - \varepsilon_{\deg}\right)^{LT} \qquad (6.16)$$

Eq. (6.17) calculates the area of the solar array A_{sa}, which is needed to confirm the power requirements of the spacecraft.

$$A_{sa} = \frac{P_{sa}}{P_{EOL}} \qquad (6.17)$$

As mentioned before, two outputs I_{max} and I_{min} of power subsystem are taken as inputs by attitude control subsystem. Hence, to perform individual disciplinary analysis in attitude control subsystem, the expression of I_{max} and I_{min} are crucial. Again, since the requirement is to estimate the critical value (maximum and minimum) of the moment of inertia other than average value, it is essential to calculate moment of inertia in all three directions (X, Y, and Z). Equations to calculate length (L), width (W), mass (m_{sa}), and finally maximum and minimum moment of inertia (I_{max} and I_{min}) are provided below.

$$L = \sqrt{\frac{A_{sa} r_{lw}}{n_{sa}}} \qquad (6.18)$$

$$W = \sqrt{\frac{A_{sa}}{r_{lw} n_{sa}}} \qquad (6.19)$$

$$m_{sa} = 2\rho L W t \qquad (6.20)$$

$$I_{saX} = m_{sa}\left[\frac{1}{12}\left(L^2 + t^2\right) + \left(D + \frac{L}{2}\right)^2\right] \quad (6.21)$$

$$I_{saY} = \frac{m_{sa}}{12}\left(t^2 + W^2\right) \quad (6.22)$$

$$I_{saZ} = m_{sa}\left[\frac{1}{12}\left(L^2 + W^2\right) + \left(D + \frac{L}{2}\right)^2\right] \quad (6.23)$$

$$I_{tot} = I_{sa} + I_{body} \quad (6.24)$$

$$I_{\max} = \max\left(I_{totX}, I_{totY}, I_{totZ}\right) \quad (6.25)$$

$$I_{\min} = \min\left(I_{totX}, I_{totY}, I_{totZ}\right) \quad (6.26)$$

FireSat is a complex MDA problem including a large number of input quantities. Sankararaman and Mahadevan (2012), Chaudhuri et al (2017), Baptista et al (2018), Isaac et al (2018), Friedman et al (2018) had solved this multidisciplinary problem considering all the input variables having aleatory uncertainty only. Again, Zaman and Mahadevan (2017), and Kibria (2018) had regarded the input variables having epistemic uncertainty, but described by only single interval data. To the best of knowledge, no one in literature had solved this complex MDA problem where input variables are described by multiple interval data, or a mixture of both multiple interval and discrete point data. Therefore, this thesis has considered this research gap, and assumed the input variables of FireSat problem as having either aleatory uncertainty or epistemic uncertainty described by mixture of multiple interval and discrete point data.

However, several input quantities are considered to be deterministic, and numerical details of these quantities are listed in Table 6.7.

Table 6.7: Numerical details of deterministic variables for FireSat problem

No.	Variable	Symbol	Unit	Data
1	Earth's radius	R_E	m	6,378,140
2	Gravitational parameter	μ	$m^3\ s^{-2}$	3.986×10^{14}
3	Target diameter	ϕ_{target}	M	235,000
4	Light speed	c	$m\ s^{-1}$	2.9979×10^{8}
5	Area reflecting radiation	A_s	m^2	13.85
6	Sun incidence angle	i	deg	0
7	Slewing time period	Δt_{slew}	s	760
8	Magnetic moment of earth	M	$A\ m^2$	7.96×10^{15}
9	Atmospheric density	ρ	$kg\ m^{-3}$	5.1480×10^{-11}
10	Cross-sectional in flight direction	A	m^2	13.85
11	No. of reaction wheels	n	-	3
12	Maximum velocity of a wheel	$\omega_{\max}$	rpm	6000
13	Holding power	P_{hold}	W	20
14	Inherent degradation of array	I_d	-	0.77
15	Power efficiency	η	-	0.22
16	Lifetime of spacecraft	LT	Years	15
17	Degradation in power production capability	$\varepsilon_{\deg}$	% per year	0.0375
18	Length to width ratio of solar array	r_{lw}	-	3
19	Number of solar arrays	n_{sa}	-	3
20	Average mass density to arrays	ρ_{sa}	$kg\ m^3$	700
21	Thickness of solar panels	t	m	0.005
22	Distance between panels	D	m	2
23	Moments of inertia of spacecraft body	I_{body}	$kg\ m^2$	$I_{body,X} = I_{body,Y}=6200$; $I_{body,Z}=4700$

Rest of the input quantities are assumed to have either aleatory or epistemic uncertainties, and numerical details of these quantities are summarized in Table 6.8.

Table 6.8: Numerical details of stochastic and epistemic variables for Firesat problem

No.	**Variable**	**Symbol**	**Unit**	**Data**	
				Mean	**Standard Deviation**
1	Altitude	H	M	18,000,000	1,000,000
2	Power other than ACS	P_{other}	W	1,000	50
3	Average solar flux	F_s	W/m^2	1,400	20
4	Deviation of moment axis	θ	deg	15	1
5	Moment arm for radiation torque	L_{sp}	M	2	0.4
6	Reflectance factor	q	-	0.5	1
7	Residual dipole of spacecraft	R_D	Am^2	5	1
8	Moment arm for aerodynamic torque	L_a	M	2	0.4
9	Drag coefficient	C_d	-	Mixture of three intervals ([0.2, 2.2], [0.3,2.4], [0.4,2.3]) and two point data {0.9, 1.15}	

All the variables listed in Table 6.8 have aleatory uncertainty except for drag coefficient C_d. It is a variable having epistemic uncertainty described by both multiple interval and sparse point data. Representing the uncertainty in the variable C_d is the first step of the proposed unified formulation.

Existing formulation (LBNA + LAMDA) estimates the nonparametric probability distribution of the epistemic variable drag coefficient C_d using LBNA method, and proposed formulation (WMLE + LAMDA) estimates the parametric probability distribution of Cd using WMLE method. The CDFs obtained by these two methods are compared with sample CDFs found by implementing MBA approach in Figure 6.13.

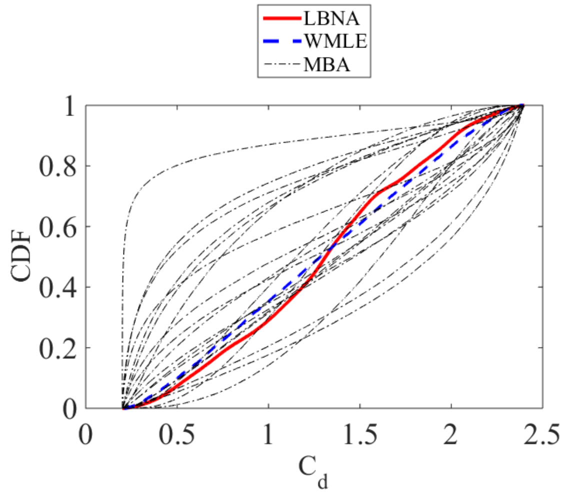


Figure 6.13: CDFs of input variable C_d

After representing the uncertainty of variable C_d using both LBNA and WMLE methods as shown in Figure 6.13, uncertainty is allowed to propagate through multidisciplinary system following LAMDA method, and the resulting PDFs of three coupling variables P_{ACS}, I_{max}, and I_{min} of the FireSat problem are reported in Figures 6.14, 6.15, and 6.16, respectively.

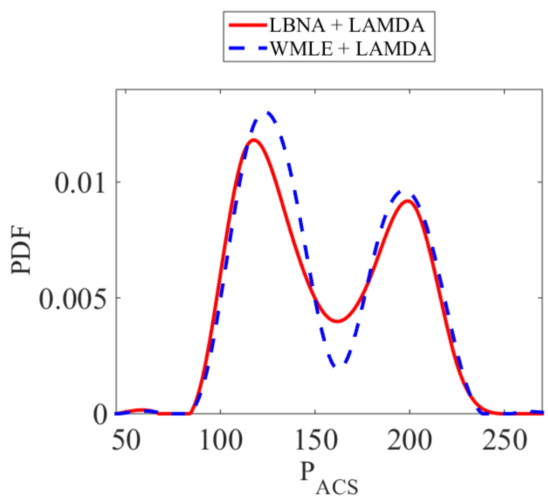


Figure 6.14: PDFs of coupling variable P_{ACS}

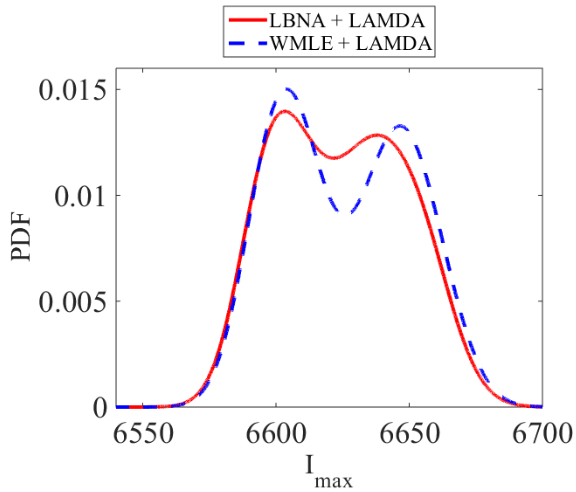


Figure 6.15: PDFs of coupling variable $I_{\max}$

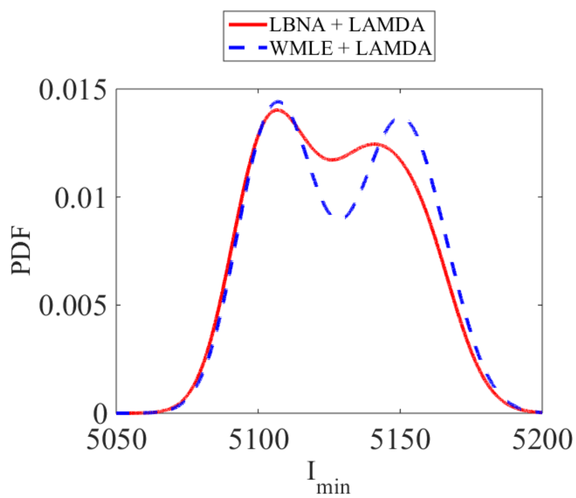


Figure 6.16: PDFs of coupling variable $I_{\min}$

Once the PDFs of the three coupling variables are available, the PDFs of three system outputs such as Total Output Power P_{tot}, Area of Solar Array A_{sa}, and Total Torque τ_{tot} can be approximated by simple uncertainty propagation through Eqs. (6.13), (6.17), and (6.5), and approximated PDFs are reported in Figures 6.17, 6.18 and 6.19, respectively.

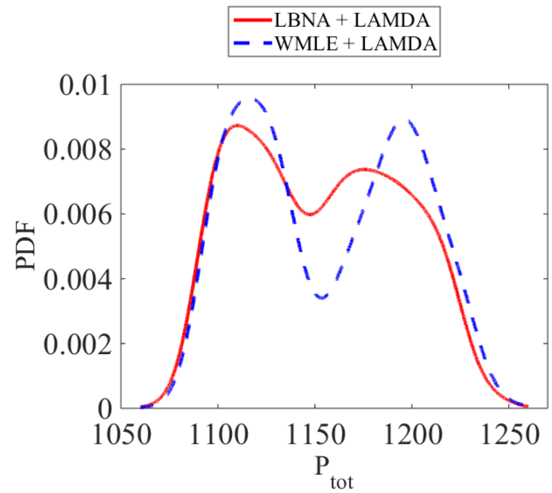


Figure 6.17: PDFs of system output: Total Output Power P_{tot}

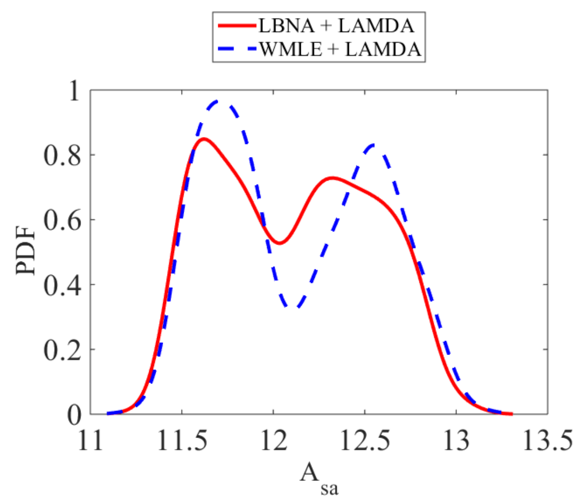


Figure 6.18: PDFs of system output: Area of Solar Array A_{sa}

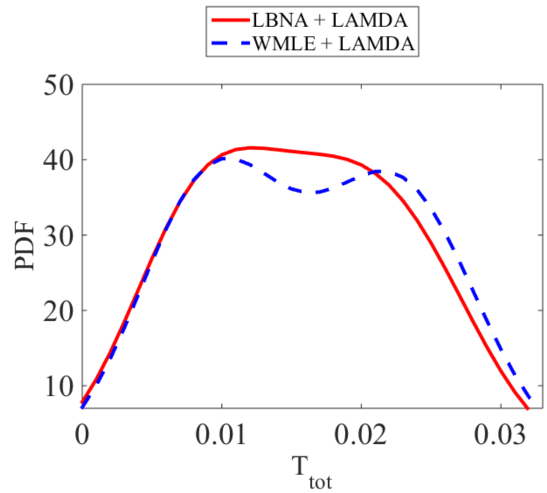


Figure 6.19: PDFs of system output: Total Torque τ_{tot}

Additionally, this thesis has also compared the CDFs of three coupling variables of the FireSat problem P_{ACS}, I_{max}, and I_{min} yielded from sampling-based SOFPI method in Figure 6.20, while uncertainty of input variable C_d is represented by LBNA, WMLE and MBA uncertainty representation methods.

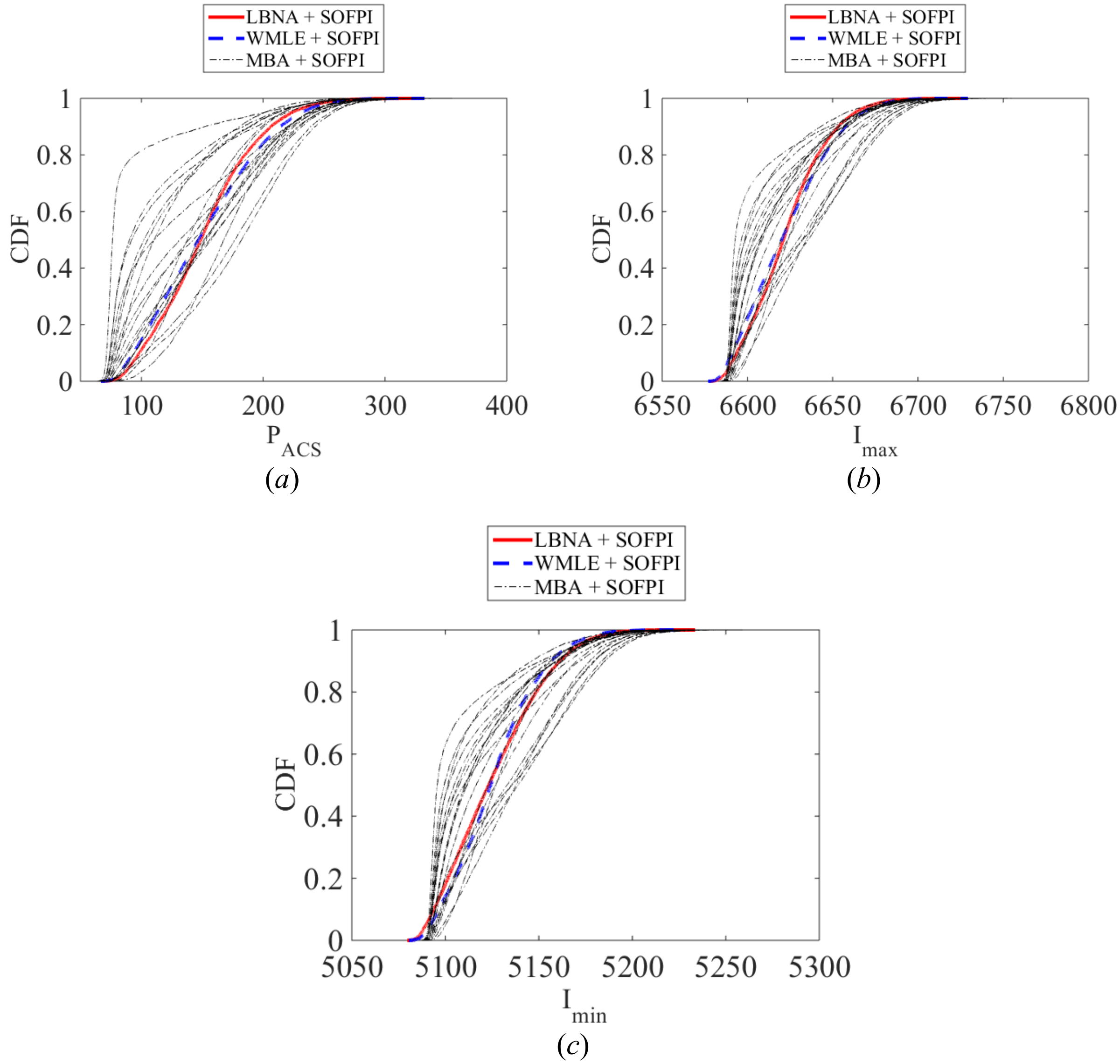


Figure 6.20: Uncertainty propagation-based comparison of CDFs of three coupling variables (*a*) Power: Attitude Control Subsystem P_{ACS}, (*b*) Maximum Moment of Inertia I_{max}, and (*c*) Minimum Moment of Inertia I_{min} for different uncertainty representation approaches

Tables 6.9 and 6.10 are used to compare the statistics (mean and standard deviation) of three coupling variables and three system outputs of this problem calculated following existing and proposed formulations.

Table 6.9: Statistics (Mean and Standard deviation) of three coupling variables P_{ACS}, I_{max}, and I_{min}

Uncertainty representation method	Uncertainty propagation method	`Coupling variable					
		P_{ACS}		I_{max}		I_{min}	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
LBNA	LAMDA	154.55	39.46	6623.51	22.52	5126.68	22.74
WMLE	LAMDA	155.70	39.42	6625.85	23.66	5129.04	23.89

Table 6.10: Statistics (Mean and Standard deviation) of three system level output P_{tot}, A_{sa}, and T_{tot}

Uncertainty representation method	Uncertainty propagation method	`System output					
		P_{tot}		A_{sa}		T_{tot}	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
LBNA	LAMDA	1153.20	40.8917	12.1009	0.42820	0.01544	0.00680
WMLE	LAMDA	1156.86	42.7892	12.1450	0.44921	0.01614	0.00713

From Tables 6.9 and 6.10, it can be said that, means and standard deviations of three coupling variables (P_{ACS}, I_{max}, and I_{min}) and three system outputs (P_{tot}, A_{sa}, and τ_{tot}) of FireSat problem are almost similar for both existing and proposed formulations.

Finally, the number of disciplinary analyses (DA) and program execution time to estimate the PDF of one of the coupling variable P_{ACS} are reported in Table 6.11 to compare the computational effort of existing (LBNA + LAMDA) and proposed (WMLE + LAMDA) formulations. To compare the computational costs of Table 6.11, an additional piece of information needs to keep in mind is that LAMDA method have used 15 discrete points to construct the PDF of the coupling variable PACS.

Table 6.11: Computational effort of different uncertainty representation and propagation methods to estimate the PDF of the coupling variable P_{ACS}

Uncertainty representation method	Uncertainty propagation method	Number of disciplinary analysis (DA)	Execution time (*s*)
LBNA	LAMDA	1,548	28.78
WMLE	LAMDA	1,338	23.99

Table 6.11 indicates that, similar to the mathematical problem, both the number of disciplinary analyses and program execution times are smaller for the proposed "WMLE + LAMDA" formulation than the existing "LBNA + LAMDA" formulation. To be specific, the proposed "WMLE + LAMDA" reduces the number of disciplinary analyses by 13.56% than the existing "LBNA + LAMDA" formulation.

Therefore, based on the results and computational expenses of both mathematical MDA problem and FireSat problem, we can conclude that the proposed "WMLE + LAMDA" formulation can perform as accurately as existing "LBNA + LAMDA" formulation, at the same time, it represents the uncertainty in the input variable through a parametric CDF, which can eliminates the complexity of propagating a nonparametric PDF through multidisciplinary system models used by existing formulation.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

This thesis proposes a new probabilistic feasible moment bounding approach (FMBA) for the representation of epistemic uncertainty arising from either multiple interval data or a mixture of both multiple interval and discrete point data. The proposed FMBA methodology includes calculation of statistical moment bounds, and then use of these bounds to approximate the probability distribution of the uncertain variable. Data can be fitted to any four-parameter family of distributions including Beta, Pearson, Johnson, and Lambda distributions. This thesis has used moment matching approach to fit data of uncertain variable to four-parameter family of Johnson distributions, as it can effectively map a number of common probability distributions, such as normal, lognormal, bounded or unbounded. Thus, the proposed FMBA represents variable with epistemic uncertainty through a family of parametric CDFs. The advantages offered by the proposed approach can be summarized as follows:

(a) Conventional moment bounding approaches estimate the first, second, third and fourth moments of a random variable by optimizing respective moment equation, and thus individual moments are calculated from individual sample sets. However, all moments of a variable should be calculated from the same sample set. The proposed FMBA eliminates this limitation through the development of a single function aggregating first to fourth moment equation, and optimization of this function facilitates the estimation of all moments of a random variable from the same sample set.

(b) Existing approaches calculate the widest possible interval for each of the moments, which is the most conservative way of representing epistemic uncertainty. Conservative representation makes decision regarding design of a system more complicated by providing more alternatives. In this case, the proposed FMBA provides narrower moment bounds than the existing methods.

(c) Traditional uncertainty representation approaches usually represent different types of uncertainty in different formats. However, the proposed FMBA can represent uncertainty in a random variable defined by either multiple interval data, or a mixture of multiple interval and discrete point data through parametric CDFs.

(d) By representing the uncertainty in a variable through parametric CDF, the proposed FMBA facilitates effective implementation of computationally more competent uncertainty propagation methods such as FORM, SORM, MCS, etc.

(e) Additionally, the solutions of Sandia epistemic uncertainty workshop challenge problems obtained by the proposed FMBA approach are appeared to be compatible with the results reported by the existing approaches.

Thus, proposed FMBA has the potential to effectively represent epistemic uncertainty emerging from multiple interval data, or a mixture of discrete point and multiple interval data.

Again, this thesis has proposed a unified formulation for the representation and propagation of epistemic uncertainty through multidisciplinary systems. The proposed formulation employs the worst-case maximum likelihood estimation (WMLE) method for the purpose of representation of epistemic uncertainty, and the likelihood-based multidisciplinary analysis (LAMDA) method to propagate uncertainty through the coupled multidisciplinary system. The proposed formulation (WMLE + LAMDA) is demonstrated with one numerical and one practical multidisciplinary analysis problems, and the results are compared with existing formulation (LBNA + LAMDA). Existing formulation represents the uncertainty in the input variable as nonparametric PDF, and propagates this PDF through multidisciplinary system. However, propagating a nonparametric PDF through systems is difficult due to the absence of information regarding distribution type and distribution parameters. Proposed formulation bypasses this limitation by representing the input variable uncertainty as a parametric CDF.

7.2 Future Work

The proposed methodologies consider input quantities as independent of each other. However, it may not be the case for the real-world problems. Future research may concentrate on representation and propagation of epistemic uncertainty based on statistical moments considering correlation among input quantities. Again, this thesis only focuses on multidisciplinary analysis (MDA), it can be extended to multidisciplinary design optimization (MDO). Additionally, this thesis considers only Johnson family of distributions to fit data, but other family of distributions (i.e., Pearson distribution) can also be explored.

REFERENCES

[1] Agarwal, H., Renaud, J. E., Preston, E. L., & Padmanabhan, D. (2004). Uncertainty quantification using evidence theory in multidisciplinary design optimization. *Reliability Engineering & System Safety, 85*(1-3), 281-294.

[2] Alexandrov, N., & Kodiyalam, S. (1998). *Initial results of an MDO method evaluation study.* Paper presented at the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization.

[3] Arnst, M., Soize, C., & Ghanem, R. (2013). Hybrid sampling/spectral method for solving stochastic coupled problems. *SIAM/ASA Journal on Uncertainty Quantification, 1*(1), 218-243.

[4] Ayyub, B. (1994). The nature of uncertainty in structural engineering. *Machine Intelligence and Pattern Recognition, 17*, 195-195.

[5] Baptista, R., Marzouk, Y., Willcox, K., & Peherstorfer, B. (2018). Optimal approximations of coupling in multidisciplinary models. *AIAA journal, 56*(6), 2412-2428.

[6] Beer, M., Ferson, S., & Kreinovich, V. (2013). Imprecise probabilities in engineering analyses. *Mechanical systems and signal processing, 37*(1-2), 4-29.

[7] Belytschko, T. (1980). Fluid-structure interaction. *Computers & Structures, 12*(4), 459-469.

[8] Ben-Haim, Y. (2004). Uncertainty, probability and information-gaps. *Reliability Engineering & System Safety, 85*(1-3), 249-266.

[9] Ben-Haim, Y., & Elishakoff, I. (2013). *Convex models of uncertainty in applied mechanics*: Elsevier.

[10] Berleant, D., & Zhang, J. (2004). Representation and problem solving with distribution envelope determination (DEnv). *Reliability Engineering & System Safety, 85*(1-3), 153-168.

[11] Borodin, V., Bourtembourg, J., Hnaien, F., & Labadie, N. (2016). Handling uncertainty in agricultural supply chain management: A state of the art. *European Journal of Operational Research, 254*(2), 348-359.

[12] Chaudhuri, A., Lam, R., & Willcox, K. (2018). Multifidelity uncertainty propagation via adaptive surrogates in coupled multidisciplinary systems. *AIAA journal, 56*(1), 235-249.

[13] Chen, X., Ng, B., Sun, Y., & Tong, C. (2013). A flexible uncertainty quantification method for linearly coupled multi-physics systems. *Journal of Computational Physics, 248*, 383-401.

[14] Cramer, E. J., Dennis, J., John E, Frank, P. D., Lewis, R. M., & Shubin, G. R. (1994). Problem formulation for multidisciplinary optimization. *SIAM Journal on Optimization, 4*(4), 754-776.

[15] Culler, A. J., & McNamara, J. J. (2010). Studies on fluid-thermal-structural coupling for aerothermoelasticity in hypersonic flow. *AIAA journal, 48*(8), 1721-1738.

[16] De Cooman, G., & Troffaes, M. C. (2004). Coherent lower previsions in systems modelling: products and aggregation rules. *Reliability Engineering & System Safety, 85*(1-3), 113-134.

[17] DeBrota, D. J., Roberts, S. D., Dittus, R., Wilson, J., Swain, J., & Venkatraman, S. (1988). *Input modeling with the Johnson system of distributions.* Paper presented at the 1988 Winter Simulation Conference Proceedings.

[18] Dey, S., Mukhopadhyay, T., Khodaparast, H. H., & Adhikari, S. (2016). Fuzzy uncertainty propagation in composites using Gram–Schmidt polynomial chaos expansion. *Applied Mathematical Modelling, 40*(7-8), 4412-4428.

[19] Du, L., Choi, K. K., Youn, B. D., & Gorsich, D. (2006). Possibility-based design optimization method for design problems with both statistical and fuzzy input data.

[20] Du, X., & Chen, W. (2005). Collaborative reliability analysis under the framework of multidisciplinary systems design. *Optimization and Engineering, 6*(1), 63-84.

[21] Dunning, P. D., Kim, H. A., & Mullineux, G. (2011). Introducing loading uncertainty in topology optimization. *AIAA journal, 49*(4), 760-768.

[22] Ferson, S., & Hajagos, J. G. (2004). Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliability Engineering & System Safety, 85*(1-3), 135-152.

[23] Ferson, S., Joslyn, C. A., Helton, J. C., Oberkampf, W. L., & Sentz, K. (2004). Summary from the epistemic uncertainty workshop: consensus amid diversity. *Reliability Engineering & System Safety, 85*(1-3), 355-369.

[24] Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D., & Sentz, K. (2003). Constructing probability boxes and Dempster-Shafer structures. Sandia National Laboratories: Technical report SANDD2002-4015.

[25] Ferson, S., Tucker, W., Paredis, C., Bulgak, Y., & Kreinovich, V. (2009). Accounting for epistemic and aleatory uncertainty in early system design. *NASA SBIR Phase, 2*.

[26] Fetz, T., & Oberguggenberger, M. (2004). Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering & System Safety, 85*(1-3), 73-87.

[27] Friedman, S., Isaac, B., Ghoreishi, S. F., & Allaire, D. L. (2018). *Efficient Decoupling of Multiphysics Systems for Uncertainty Propagation.* Paper presented at the 2018 AIAA Non-Deterministic Approaches Conference.

[28] Ghosh, S., & Mavris, D. N. (2020). Methodology for Probabilistic Analysis of Distributed Multidisciplinary Architecture. *Journal of Aircraft*, 1-13.

[29] Gu, X., Renaud, J. E., Batill, S. M., Brach, R. M., & Budhiraja, A. S. (2000). Worst case propagated uncertainty of multidisciplinary systems in robust design optimization. *Structural and Multidisciplinary Optimization, 20*(3), 190-213.

[30] Haldar, A., & Mahadevan, S. (2000). *Probability, reliability, and statistical methods in engineering design*: J. Wiley & Sons, Incorporated.

[31] Hearn, T. A., Hendricks, E., Chin, J., & Gray, J. S. (2016). *Optimization of turbine engine cycle analysis with analytic derivatives.* Paper presented at the 17th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.

[32] Helton, J., Johnson, J., Oberkampf, W., & Storlie, C. B. (2007). A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory. *Computer Methods in Applied Mechanics and Engineering, 196*(37-40), 3980-3998.

[33] Helton, J. C., Johnson, J. D., & Oberkampf, W. L. (2004). An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering & System Safety, 85*(1-3), 39-71.

[34] Helton, J. C., Johnson, J. D., Oberkampf, W. L., & Sallaberry, C. J. (2010). Representation of analysis results involving aleatory and epistemic uncertainty. *International Journal of General Systems, 39*(6), 605-646.

[35] Hu, X., Chen, X., Parks, G. T., & Tong, F. (2019). Uncertainty-based design optimization approach based on cumulative distribution matching. *Structural and Multidisciplinary Optimization, 60*(4), 1571-1582.

[36] Hu, Z., & Mahadevan, S. (2017). Uncertainty quantification in prediction of material properties during additive manufacturing. *Scripta materialia, 135*, 135-140.

[37] Huang, Z., Jiang, W., & Tang, Y. (2018). A new method to evaluate risk in failure mode and effects analysis under fuzzy information. *Soft Computing, 22*(14), 4779-4787.

[38] Isaac, B., Friedman, S., & Allaire, D. L. (2018). *Efficient approximation of coupling variable fixed point sets for decoupling multidisciplinary systems.* Paper presented at the 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference.

[39] Jiang, Z., Li, W., Apley, D. W., & Chen, W. (2015). A spatial-random-process based multidisciplinary system uncertainty propagation approach with model uncertainty. *Journal of Mechanical Design, 137*(10).

[40] Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika, 36*(1/2), 149-176.

[41] Kibria, M. (2018). Decoupled approach to Multidisciplinary system analysis and Design under epistemic uncertainty.

[42] Klir, G. J. (2004). Generalized information theory: aims, results, and open problems. *Reliability Engineering & System Safety, 85*(1-3), 21-38.

[43] Kokkolaras, M., Mourelatos, Z. P., & Papalambros, P. Y. (2004). *Design optimization of hierarchically decomposed multilevel systems under uncertainty.* Paper presented at the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference.

[44] Kozine, I. O., & Utkin, L. V. (2004). An approach to combining unreliable pieces of evidence and their propagation in a system response analysis. *Reliability Engineering & System Safety, 85*(1-3), 103-112.

[45] Liang, C., Mahadevan, S., & Sankararaman, S. (2015). Stochastic multidisciplinary analysis under epistemic uncertainty. *Journal of Mechanical Design, 137*(2).

[46] Mahadevan, S., & Rebba, R. (2006). Inclusion of model errors in reliability-based optimization.

[47] Mahadevan, S., & Smith, N. (2006). Efficient first-order reliability analysis of multidisciplinary systems. *International Journal of Reliability and Safety, 1*(1-2), 137-154.

[48] Marhadi, K. (2007). Quantifying reliability estimate due to statistical uncertainty. *Independent Study Report, School of Mathematics, Claremont Graduate University, Claremont, California*.

[49] Matsumura, T., & Haftka, R. T. (2013). Reliability based design optimization modeling future redesign with different epistemic uncertainty treatments. *Journal of Mechanical Design, 135*(9).

[50] Meeker, W. Q., & Escobar, L. A. (1995). Teaching about approximate confidence regions based on maximum likelihood estimation. *The American Statistician, 49*(1), 48-53.

[51] Ni, B., Jiang, C., & Huang, Z. (2018). Discussions on non-probabilistic convex modelling for uncertain problems. *Applied Mathematical Modelling, 59*, 54-85.

[52] Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., & Ferson, S. (2004). Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering & System Safety, 85*(1-3), 11-19.

[53] Oberkampf, W. L., & Roy, C. J. (2010). *Verification and validation in scientific computing*: Cambridge University Press.

[54] Pawitan, Y. (2001). *In all likelihood: statistical modelling and inference using likelihood*: Oxford University Press.

[55] Peng, X., Gao, Q., Li, J., Liu, Z., Yi, B., & Jiang, S. (2020). Probabilistic Representation Approach for Multiple Types of Epistemic Uncertainties Based on Cubic Normal Transformation. *Applied Sciences, 10*(14), 4698.

[56] Red-Horse, J., & Benjamin, A. S. (2004). A probabilistic approach to uncertainty quantification with limited information. *Reliability Engineering & System Safety, 85*(1-3), 183-190.

[57] Sankararaman, S., & Mahadevan, S. (2011). Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. *Reliability Engineering & System Safety, 96*(7), 814-824.

[58] Sankararaman, S., & Mahadevan, S. (2012). Likelihood-based approach to multidisciplinary analysis under uncertainty. *Journal of Mechanical Design, 134*(3).

[59] Shafer, G. (1976). *A mathematical theory of evidence* (Vol. 42): Princeton university press.

[60] Thornton, E. A. (1992). Thermal structures-four decades of progress. *Journal of Aircraft, 29*(3), 485-498.

[61] Thunnissen, D. P. (2003). *Uncertainty classification for the design and development of complex systems.* Paper presented at the 3rd annual predictive methods conference.

[62] Venkatraman, S., & Wilson, J. R. (1988). *Modeling Univariate Populations with Johnson's Translation System: Description of the FITTR1 Software*: School of Industrial Engineering, Purdue University.

[63] Wang, X., Wang, L., & Qiu, Z. (2014). A feasible implementation procedure for interval analysis method from measurement data. *Applied Mathematical Modelling, 38*(9-10), 2377-2397.

[64] Wertz, J. R., & Larson, W. J. (1999). Space mission analysis and design:[SMAD III]: Microcosm Press, Torrance.

[65] Wu, J., Han, X., & Tao, Y. (2019). Kinematic response of industrial robot with uncertain-but-bounded parameters using interval analysis method. *Journal of Mechanical Science and Technology, 33*(1), 333-340.

[66] Yang, X., Liu, Y., & Ma, P. (2017). Structural reliability analysis under evidence theory using the active learning kriging model. *Engineering Optimization, 49*(11), 1922-1938.

[67] Yin, S., Yu, D., Luo, Z., & Xia, B. (2018). An arbitrary polynomial chaos expansion approach for response analysis of acoustic systems with epistemic uncertainty. *Computer Methods in Applied Mechanics and Engineering, 332*, 280-302.

[68] Zadeh, L. A. (2002). Toward a perception-based theory of probabilistic reasoning with imprecise probabilities *Soft Methods in Probability, Statistics and Data Analysis* (pp. 27-61): Springer.

[69] Zaman, K. (2010). *Modeling and management of epistemic uncertainty for multidisciplinary system analysis and design.* Vanderbilt University.

[70] Zaman, K., & Dey, P. R. (2017). Likelihood-based representation of epistemic uncertainty and its application in robustness-based design optimization. *Structural and Multidisciplinary Optimization, 56*(4), 767-780.

[71] Zaman, K., & Mahadevan, S. (2017). Reliability-based design optimization of multidisciplinary system under aleatory and epistemic uncertainty. *Structural and Multidisciplinary Optimization, 55*(2), 681-699.

[72] Zaman, K., McDonald, M., & Mahadevan, S. (2011b). Probabilistic framework for uncertainty propagation with both probabilistic and interval variables. *Journal of Mechanical Design, 133*(2).

[73] Zaman, K., Rangavajhala, S., McDonald, M. P., & Mahadevan, S. (2011a). A probabilistic approach for representation of interval uncertainty. *Reliability Engineering & System Safety, 96*(1), 117-130.

[74] Zhang, J., & Shields, M. D. (2018). On the quantification and efficient propagation of imprecise probabilities resulting from small datasets. *Mechanical systems and signal processing, 98*, 465-483.

[75] Zhang, X., & Huang, H.-Z. (2010). Sequential optimization and reliability assessment for multidisciplinary design optimization under aleatory and epistemic uncertainties. *Structural and Multidisciplinary Optimization, 40*(1-6), 165.

[76] Zhang, Z., Jiang, C., Wang, G., & Han, X. (2015). First and second order approximate reliability analysis methods using evidence theory. *Reliability Engineering & System Safety, 137*, 40-49.

[77] Zhuang, X., & Pan, R. (2012). *Epistemic uncertainty in reliability-based design optimization.* Paper presented at the 2012 Proceedings Annual Reliability and Maintainability Symposium.

APPENDICES

Appendix A: Johnson Family of Distribution

Johnson distribution is a four parameter (δ, γ, ξ and λ) family of distributions. Researchers are interested in this distribution because it can be fitted to sample data set from almost any probability distribution type. The first step of fitting data to Johnson distribution includes transformation of these data to standard normal space. Johnson, 1949 proposed a generalized equation (Eq. (A1)) to transform a continuous random variable, *x* into standard normal variable, *z*.

$$z = \gamma + \delta f\left(\frac{x-\xi}{\lambda}\right) \quad \text{(A1)}$$

Where, $f(\ )$ - transformation functions

z - standard normal variable, $z \sim N(0,1)$

δ and γ - shape parameters

ξ - location parameter

λ - scale parameter

Again, transformation function $f(y)$ in Eq. (A1) can be of four different forms (Table A.1) to map normal, lognormal, bounded, and unbounded Johnson distributions as shown, where $y = \frac{x-\xi}{\lambda}$.

Table A1: Johnson's four distribution functions

Distribution	$f(y)$
Normal distribution, S_N	$f(y) = y$
Log-normal distribution, S_L	$f(y) = \ln(y)$
Bounded Johnson distribution, S_B	$f(y) = \ln(\frac{y}{1-y})$
Unbounded Johnson distribution, S_U	$f(y) = \ln(y + \sqrt{y^2+1})$

Since Johnson family of distributions can represent sample data in probabilistic format without knowing distribution information, it cancels out the necessity of checking a number of candidate named distributions (i.e. normal, lognormal etc.) in the race of being best fit to that particular sample. Therefore, Johnson distribution is a convenient choice for the transformation of data having no distribution information to a probabilistic format.

Appendix B: Fitting Johnson Distributions to Data

Parameter estimation of Johnson distribution is one of the major task in fitting Johnson distribution to data. For this purpose, DeBrota et al (1988) proposed four approaches including moment matching and percentile matching. Among these, this thesis has used moment matching approach to fit Johnson distributions to data.

B.1 Fitting Johnson distributions to point data

Available point data may fit to any one of normal, lognormal, bounded and unbounded Johnson distribution types. To find out the best one, a standard three step approach (Venkatraman and Wilson, 1987) is available in literature as discussed below.

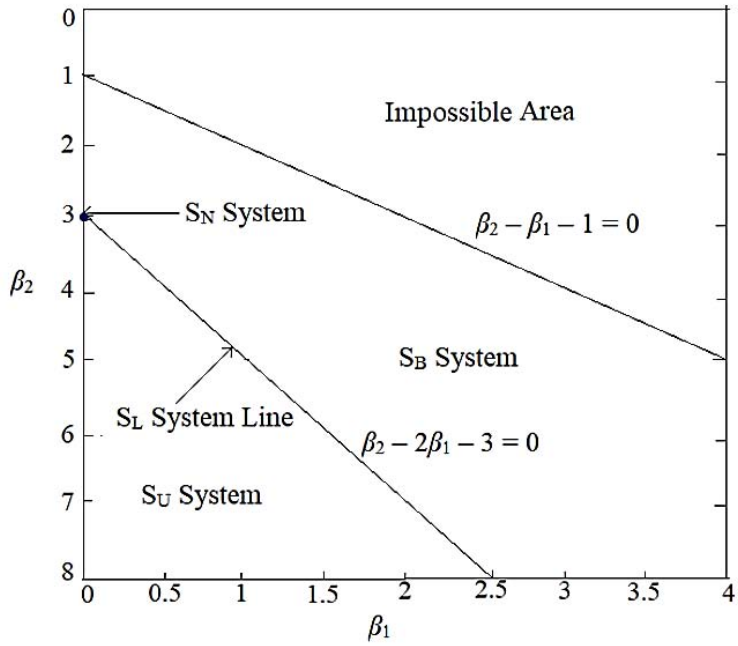
Step 1: From the point data set $\boldsymbol{x}$, calculate second (m_2), third (m_3) and fourth (m_4) moments.

Step 2: Calculate $\beta_1 \equiv {m_3}^2 / {m_2}^3$ and $\beta_2 \equiv m_4 / {m_2}^2$

Step 3: Find out the suitable distribution family by using chart in Figure A.1.

If information regarding distribution type is available, then anyone can calculate distribution parameters from sample point data. Once the values of distribution parameters are ready for use, it is easy to return x space from standard normal space by using inverse translation of z as Eq. (B1), where $z \sim N(0,1)$.

$$x = \xi + \lambda f^{-1}\left(\frac{z-\gamma}{\delta}\right) \tag{B1}$$



Here,

$$\beta_1 \equiv {m_3}^2 / {m_2}^3$$

$$\beta_2 \equiv m_4 / {m_2}^2$$

where,

m2 = second moment

m3 = third moment

m4 = fourth moment

Figure A.1: Chart for the identification of Johnson distribution family (Marhadi, 2007).

B.2 Fitting Johnson distributions to interval data

Fitting Johnson distribution to interval data involves estimation of four Johnson parameters δ, γ, ξ, and λ. If variable described by interval data are available as moment bounds, moment matching approach can be utilized to fit Johnson distribution to these bounds. In response of this, Zaman et al (2011) proposed a *sampling-based* approach, which includes generating random sample of moments from already estimated moment bounds, and fitting single Johnson distribution to each sampled moments set. Detail steps of sampling-based approach are outlined below.

Step 1: Estimate lower and upper bounds on the first through fourth moments of a random variable defined by single, or multiple interval, or mixture of both multiple interval and sparse point data.

Step 2: Pick a set of first four moments (m_1, m_2, m_3, and m_4) from their respective bounds. This selection can be done by using any discretization method. Choice of sampling approach may affect final results. This thesis uses *uniform distribution* in this regard.

Step 3: Since true value must prevail within the overall bounds [*min*(*lower bounds*) *max*(*upper bounds*)] of the data, we can assume the distribution type as bounded Johnson distribution.

Step 4: Estimate four parameters of the bounded Johnson distribution by using moment matching approach.

Step 5: Step 2, 3 and 4 yield a single Johnson distribution, and in the same way desired number of distribution can be obtained to form family of distributions.