13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13
Seoul, South Korea, May 26-30, 2019

Tightening the bound estimate of structural reliability under imprecise probability information

Cao Wang
School of Civil Engineering, The University of Sydney, Sydney, NSW, Australia. Email: cao.wang@sydney.edu.au

Hao Zhang
School of Civil Engineering, The University of Sydney, Sydney, NSW, Australia. Email: hao.zhang@sydney.edu.au

Michael Beer
Institute for Risk and Reliability, Leibniz Universität Hannover, Hannover, Germany; Institute for Risk and Uncertainty, University of Liverpool, Liverpool, UK; International Joint Research Center for Engineering Reliability and Stochastic Mechanics (ERSM), Tongji University, Shanghai, China. Email: beer@irz.uni-hannover.de

ABSTRACT: Structural reliability analysis is typically performed based on the identification of distribution types of random inputs. However, this is often not feasible in engineering practice due to limited available probabilistic information (e.g., limited observed samples or physics-based inference). In this paper, a linear programming-based approach is developed to perform structural reliability analysis subjected to incompletely informed random variables. The approach converts a reliability analysis into a standard linear programming problem, which can make full use of the probabilistic information of the variables. The proposed method can also be used to construct the best-possible distribution function bounds for a random variable with limited statistical information. Illustrative examples are presented to demonstrate the applicability and efficiency of the proposed method. It is shown that the proposed approach can provide a tighter estimate of structural reliability bounds compared with existing interval Monte Carlo methods which propagate probability boxes.

1. INTRODUCTION

The various sources of uncertainties arising from structural capacities and applied loads, as well as computational models, are responsible for the failure risk of civil structures and infrastructure. In an attempt to measure the safety of a structure, it is necessary to quantify and model these uncertainties with a probabilistic approach so as to further determine the failure probability (Melchers, 1999; Ellingwood, 2005; Li et al., 2015). In this procedure, the identification of the probability distributions of random variables is crucial. The uncertainty associated with a random variable is practically classified into either aleatory or epistemic (Der Kiureghian and Ditlevsen, 2009), with the former arising from inherent random nature of the quantity, and the latter due to knowledge-based factors such as imperfect modelling and simplifications. Statistical uncertainty is an important source of epistemic uncertainty, which accounts for the difference between the probability model of a random variable inferred from sampled data and the 'true' one. This uncertainty may be significant if the size of available data/observations is limited, and thus has gained much attention in the scientific community recently to better assess the safety of a structure.

Both statistical analysis-based and physical reasoning-based techniques are available in the literature to determine the probability distribution of a random variable (Tang and Ang, 2007). The quality of analysis results is sensitive to the selection of the probability distribution of random inputs. However, in many cases, the identification of a variable's distribution function is difficult or even impossible due to limited available information/data. Rather, only incomplete information such as the first- and the second- order moments (i.e., mean and variance) of the variable is available or predictable. In such a case, the incompletely-informed variable is expected to have a family of candidate probability distributions rather than a single known distribution function. As a result, the structural reliability in the presence of the incompletely-informed variable can no longer be uniquely determined. A practical way to represent an imprecise probability is to use a probability bounding approach by considering the lower and upper bounds of the imprecise probability. Under this context, approaches of interval estimate have been widely used to deal with the reliability problem with imprecise probabilistic information, including the probability-box (p-box for short) method (Ferson et al., 2003), random set and Dempster-Shafer evidence theory (Alvarez et al., 2018) and others. These methods are closely related to each other, and may be used as equivalent for the purpose of reliability assessment (Ferson et al., 2003; Zhang, 2012). However, the bounds of structural reliability estimated using a probability bounding approach may be overly conservative in some cases, due to the fact that it only considers the bounds of the distribution function, thus some useful information inside the bounds may be lost. This fact calls for an improved approach for reliability bound estimate which can take full use of the imprecise information of the variable(s).

The use of either the moment information or the probability distribution funciton of a random variable is exchangeable in structural reliability analysis, due to the fact that both of them can determine each other uniquely (Bisgaard and Sasvári, 2000). Many previous studies have actually conducted reliability analysis by making use of the moment information of random variables (Der Kiureghian et al., 1987; Zhao and Ang, 2003; Wang et al., 2016). The use of moment information is emphasized herein due to the fact that in many cases only limited observations/samples of a random variable are accessible and thus the calibration/prediction of the moments (typically low orders so as to avoid potentially biased estimate) based on the limited samples is relatively straightforward compared with that of the complete distribution function. In this paper, we consider the case of imprecise probability in which only the low-order moments of the random variable are known or predictable, while the distribution type is not known.

This paper proposes a linear programming-based method for solving reliability problems in the presence of imprecise probabilistic information. The estimate of reliability bounds is transformed into finding an optimized solution of a linear objective function, where the constraint equations are established by taking full use of the information of moments, and possibly the range information of the random variable. As a by-product, the proposed method can also be used to construct the best-possible cumulative density function (CDF) bounds for a random variable with limited statistic information.

2. Probability-box method in the presence of imprecise information

2.1. *Impact of imprecision on reliability assessment*

A typical structural reliability problem takes the form of

$$\mathscr{P} = \Pr(G(\mathbf{X}) \leq 0) = \int \cdots \int_{G(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} \quad (1)$$

where $\mathscr{P}$ represents the failure probability of the structure, Pr denotes the probability of the event in the bracket, G is the limit state function in the presence of m random inputs $\mathbf{X} = \{X_1, X_2, \ldots X_m\}$, which defines structural failure if $G < 0$ and the survival of the structure otherwise, and $f_{\mathbf{X}}(\mathbf{X})$ is the joint probability density function (PDF) of $\mathbf{X}$. The failure probability in Eq. (1) is often estimated by the well-known Monte Carlo method: $\mathscr{P} \approx \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}\left[G(\mathbf{x}_j) \leq 0\right]$, where N is the number of replications, $\mathbb{I}[\bullet]$ is an indicator function, which returns 1 if the statement in the bracket is true and 0 otherwise, and $\mathbf{x}_j$ is the jth simulated sample of $\mathbf{X}$. $\mathbf{x}_j$ can be generated using the inverse transform method according to $\mathbf{x}_j = F_{\mathrm{X}}^{-1}(\mathbf{r}_j)$ for $j = 1,2,\ldots,N$, with F_{X} being the CDF of $\mathbf{X}$, and $\mathbf{r}_j$ a sample of standard uniform random variable.

When the distribution function cannot be determined uniquely and one has to consider a family of all possible distribution functions, the probability of failure will vary in an interval $[\underline{\mathscr{P}}, \overline{\mathscr{P}}]$, which, theoretically, are given as follows for all possible F_X.

$$\underline{\mathscr{P}} = \min\left\{\frac{1}{N}\sum_{j=1}^{N}\mathbb{I}\left[G\left(F_X^{-1}(\mathbf{r}_j)\right) \le 0\right]\right\} \quad (2)$$

and

$$\overline{\mathscr{P}} = \max\left\{\frac{1}{N}\sum_{j=1}^{N}\mathbb{I}\left[G\left(F_X^{-1}(\mathbf{r}_j)\right) \le 0\right]\right\} \quad (3)$$

2.2. *Probability box approach*

A probability box describes a family of distribution functions by specifying the lower and upper bounds of the CDF, i.e.,

$$\underline{F}_X(x) \le F_X(x) \le \overline{F}_X(x), \quad x \in \mathbb{R} \quad (4)$$

where $F_X(x)$ is the (unknown) CDF of X, $\underline{F}_X$ and $\overline{F}_X$ are the lower and upper bounds of F_X respectively.

For a number of cases of imprecise probability, methods are available in the literature to construct the corresponding probability boxes. If only the mean and standard deviation of X are known, denoted by μ_X and σ_X respectively, and the distribution type is unknown, the Chebyshev's inequality gives a lower and an upper bound of F_X (Oberguggenberger and Fellin, 2008), i.e.,

$$\underline{F}_X(x) = \begin{cases} 0, & x \le \mu_X + \sigma_X \\ 1 - \frac{\sigma_X^2}{(x-\mu_X)^2}, & x \ge \mu_X + \sigma_X \end{cases} \quad (5a)$$

$$\overline{F}_X(x) = \begin{cases} \frac{\sigma_X^2}{(x-\mu_X)^2}, & x \le \mu_X - \sigma_X \\ 1, & x \ge \mu_X - \sigma_X \end{cases} \quad (5b)$$

However, the CDF bounds as given in Eq. (5) are not the best-possible. As will be shown later in this paper, tighter CDF bounds can be constructed for this case.

In practice, the bounds of a random variable are often known, e.g., structural loads are non-negative. The range information can be utilized to tighten the bounds of F_X. Let $\underline{x}$ and $\bar{x}$ denote the minimum and maximum of X, respectively, Ferson et al. (2003) gave a tighter bounds of F_X, which are the best possible bounds in the sense that the bounds cannot be any tighter if one only knows the min, max, mean and variance of a random variable.

A distribution function with uncertain parameters represents another common case of imprecise probabilities. As the statistical parameters of a distribution function are usually estimated by statistical inference from sample observations, uncertainties arise in the estimation of the parameters when the available data is limited. A natural way to quantify the uncertainty of the parameters is to use the confidence intervals which define interval bounds of the distribution parameters. Zhang et al. (2010) and Zhang (2012) have considered the case in which the distribution type is known, but the distribution parameters are uncertain and modeled by intervals.

The present paper considers the imprecise probabilities in which the available information is limited to the mean and variance (either point estimates or interval estimates), and/or the range of the random variable. The distribution type is assumed to be unknown.

2.3. *Interval Monte Carlo methods to propagate p-boxes*

When the reliability analysis involves p-boxes, an interval Monte Carlo method can be used to propagate probability boxes and compute the bounds of probability of failure. The basic Monte Carlo simulation is extended to the case where the distribution function F_X is a p-box.

In the presence of the CDF envelope (c.f. Eq. (4)) for $\mathbf{X}$, for each simulation run, two samples can be generated from the lower and upper bounds of F_X, respectively, i.e.,

$$\underline{\mathbf{x}}_j = \overline{F}_X^{-1}(\mathbf{r}_j), \bar{\mathbf{x}}_j = \underline{F}_X^{-1}(\mathbf{r}_j), \quad j = 1,2,\ldots,N \quad (6)$$

The interval $[\underline{\mathbf{x}}_j, \overline{\mathbf{x}}_j]$ contains all possible simulated numbers from the family of distributions contained in the p-box for a given value of $\mathbf{r}_j$.

Let $\min G(\mathbf{x}_j)$ and $\max G(\mathbf{x}_j)$ respectively denote the minimum and maximum of the limit state function $G(\mathbf{X})$ when $\underline{\mathbf{x}}_j \leq \mathbf{X} \leq \overline{\mathbf{x}}_j$. It simply follows,

$$\mathbb{I}[\max G(\mathbf{x}_j) \leq 0] \leq \mathbb{I}[G(\mathbf{x}_j) \leq 0] \leq \mathbb{I}[\min G(\mathbf{x}_j) \leq 0] \quad (7)$$

which further gives

$$\frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\left[\max G\left(\mathbf{x}_j\right) \leq 0\right] \leq \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\left[G\left(\mathbf{x}_j\right) \leq 0\right] \leq \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\left[\min G\left(\mathbf{x}_j\right) \leq 0\right] \quad (8)$$

Thus, a lower and an upper bounds of $\mathscr{P}$, $\underline{\mathscr{P}}$ and $\overline{\mathscr{P}}$, are obtained respectively as follows (Zhang et al., 2010),

$$\underline{\mathscr{P}} = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\left[\max G\left(\mathbf{x}_j\right) \leq 0\right] \quad (9)$$

and

$$\overline{\mathscr{P}} = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}\left[\min G\left(\mathbf{x}_j\right) \leq 0\right] \quad (10)$$

Details about interval Monte Carlo method can be found elsewhere (Zhang et al., 2010; Zhang, 2012). Clearly, the reliability bounds as given by Eqs. (9) and (10) are more conservative than the true bounds of Eqs. (2) and (3). This issue is addressed in Section 4.

3. LINEAR PROGRAMMING-BASED PROBABILITY BOUNDS ANALYSIS

3.1. Formulation of the problem

Consider a reliability analysis problem which involves the random variables $[Q, \mathbf{S}]$, in which Q is a random variable with an imprecise distribution function, and $\mathbf{S} = [S_1, S_2, \ldots]$ is the remaining random vector with a known joint distribution function. Q and $\mathbf{S}$ are assumed to be statistically independent. The failure probability is given by

$$\mathscr{P} = \int_{G(\mathbf{S},Q)\leq 0} f_Q(q) f_{\mathbf{S}}(\mathbf{s})\mathrm{d}q\mathrm{d}\mathbf{s}, \quad (11)$$

in which $f_Q(q)$ and $f_{\mathbf{S}}(\mathbf{s})$ are the density functions of Q and $\mathbf{S}$, respectively. Eq. (11) can be rewritten as

$$\mathscr{P} = \int f_Q(q)\xi_Q(q)\mathrm{d}q \quad (12)$$

in which $\xi_Q(q)$ represents the conditional failure probability on $Q = q$, i.e.,

$$\xi_a(q) \triangleq \Pr(G(\mathbf{S}, Q = q) \leq 0) = \int_{G(\mathbf{S},Q=q)\leq 0} f_{\mathbf{S}}(\mathbf{s})\mathrm{d}\mathbf{s} \quad (13)$$

Note that the conditional failure probability $\xi_a(q)$ for a given value of $Q = q$ is referred to as the *fragility* in the risk analysis for natural hazards (Li and Ellingwood, 2008). The conditional failure probability $\xi_a(q)$ may be obtained analytically through the integration in Eq. (13), or numerically using the Monte Carlo methods.

To facilitate the derivation, Q is normalized into $[0,1]$ by introducing a reduced random variable $X = \dfrac{Q - Q_{\min}}{Q_{\max} - Q_{\min}}$, where $Q_{\max}$ and $Q_{\min}$ are the maximum and minimum of Q, respectively. With this, Eq. (12) becomes

$$\mathscr{P} = \int_0^1 f_X(x)\xi(x)\mathrm{d}x \quad (14)$$

where $f_X(x)$ is the PDF of X, and $\xi(x) = \xi_a((Q_{\max} - Q_{\min})x + Q_{\min})$. The computation of tight bounds of Eq. (14) is discussed next, employing the algorithms of linear programming.

Consider the case where the only information about the imprecise probability Q is its first two moments, i.e., the mean (μ_Q) and the standard deviation (σ_Q). To apply Eq. (14), the maximum and minimum of Q need to be estimated. In practice, they can be approximated as $\mu_Q \pm k\sigma_Q$, in which k is sufficiently large (e.g., $k = 5$). Note that the selection of k has a negligible impact on the aftermentioned optimization results. Clearly, the mean and standard deviation of the reduced variable X are

$$\mu_X = \frac{\mu_Q - Q_{\min}}{Q_{\max} - Q_{\min}}, \ \sigma_X = \frac{\sigma_Q}{Q_{\max} - Q_{\min}} \quad (15)$$

In Eq. (14), as the distribution type of X is unknown, the values of $f_X(x)$ for each x cannot be uniquely determined. We discretize the domain of X, $[0,1]$, into n identical sections, $[x_0 =$

$0,x_1],[x_1,x_2],\dots[x_{n-1},x_n=1]$, where n is sufficiently large such that $\left|f_X(x)-f_X\left(\frac{x_{i-1}+x_i}{2}\right)\right|$ is negligible for $\forall i=1,2,\dots n$ and $\forall x\in[x_{i-1},x_i]$. The sequence $f_X\left(\frac{x_{i-1}+x_i}{2}\right), \forall i=1,2,\dots n$ is denoted by $\{f_1,f_2,\dots f_n\}$ for the purpose of simplicity. With this, Eq. (14) can be approximated by

$$\mathscr{P}=\int_0^1\xi(x)f_X(x)\mathrm{d}x=\lim_{n\to\infty}\sum_{i=1}^{n}\xi\left(\frac{i-0.5}{n}\right)\frac{1}{n}\cdot f_i \tag{16}$$

Note that the definition of the mean value and variance of X, as well as the basic characteristics of a distribution function simultaneously give

$$\begin{cases}\sum_{i=1}^n f_i\cdot\frac{1}{n}=1\\ \sum_{i=1}^n f_i\cdot\frac{1}{n}\cdot\frac{i}{n}=\mu_X\\ \sum_{i=1}^n f_i\cdot\frac{1}{n}\left(\frac{i}{n}\right)^2=\mu_X^2+\sigma_X^2\\ 0\le f_i\le n,\forall i=1,2,\dots n\end{cases} \tag{17}$$

Eqs. (16) and (17) indicate that the bound estimate of $\mathscr{P}$ can be converted into a classic linear programming problem, i.e., Eq. (16) is the objective function to be optimized, $\mathbf{f}=\{f_1,f_2,\dots f_n\}$ are the vector of variables to be determined, and Eq. (17) represents the constraints. The algorithms of linear programming-based optimization have been well studied and can be found elsewhere (Vanderbei, 2014).

Eqs. (16) and (17) represents a linear programming-based approach to compute the reliability bounds for imprecise probability distributions. Another useful application of Eqs. (16) and (17) is to construct the best-possible CDF bounds for a random variable with incomplete information. For an arbitrary value of τ, by setting

$$\xi(x)=\mathbb{I}(\tau\ge x)=\begin{cases}1, & x\le\tau\\ 0, & \text{otherwise}\end{cases} \tag{18}$$

Eq. (16) becomes

$$\int_0^1\xi(x)f_X(x)\mathrm{d}x=\int_0^\tau f_X(x)\mathrm{d}x=F_X(\tau) \tag{19}$$

Thus, by solving the linear programming problem defined by Eqs. (19) and (17), the best-possible bounds for $F_X(\tau)$ can be obtained.

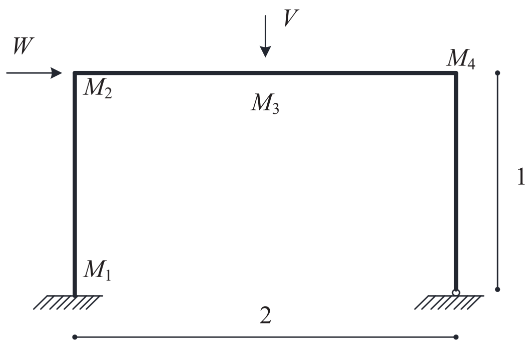


Figure 1: A rigid-plastic portal frame (after Melchers (1999)).

The constraints in Eq. (17) represent the case in which the only knowledge available are the point estimates of the mean and the standard deviation. The constraints can be easily modified for more generalized cases if additional information is provided. For example, if Q is known to be strictly defined in the range $[\underline{q},\overline{q}]$, where $0\le\underline{q}\le\overline{q}$, the introduction of a new variable $Q'=\frac{Q-\underline{q}}{\overline{q}-\underline{q}}$ enables the applicability of Eq. (17). Moreover, if the mean value of X is an interval estimate of $[\underline{\mu}_X,\overline{\mu}_X]$ rather than a point estimate, the second constraint equation in Eq. (17), $\sum_{i=1}^n f_i\cdot\frac{1}{n}\cdot\frac{i}{n}=\mu_X$, is modified as

$$\begin{cases}\sum_{i=1}^n f_i\cdot\frac{-1}{n}\cdot\frac{i}{n}\le-\underline{\mu}_X\\ \sum_{i=1}^n f_i\cdot\frac{1}{n}\cdot\frac{i}{n}\le\overline{\mu}_X\end{cases} \tag{20}$$

A similar modification can be made to the third constraint equation in Eq. (17) if the standard deviation of X is known to have a predefined range. It should be noted that the probability-box obtained by the proposed linear programming method will be identical to the probability-box given by Ferson et al. (2003) if one knows the min, max, mean and variance of a random variable. However, the proposed linear programming-based approach represents a more general approach for constructing the best-possible probability-boxes.

4. Example: A portal frame

In this section, an illustrative example is presented to demonstrate the applicability and efficiency of the proposed method. The reliability analysis of

Table 1: Statistics of the variables of a portal frame.

Variable	Dist. type	Mean	Std. Dev.
M_1, M_2, M_3, M_4	Normal	1.0	0.3
V	Normal	1.5	0.3

a rigid-plastic portal frame as shown in Fig. 1 is considered. The frame is subjected to a horizontal wind load W and a vertical load V. The layout and member geometry of the structure are adopted from Melchers (1999). The structure may fail due to one of the following three limit states,

$$\begin{aligned} G_1(\mathbf{X}) &= M_1 + 2M_3 + 2M_4 - W - V \\ G_2(\mathbf{X}) &= M_2 + 2M_3 + M_4 - V \\ G_3(\mathbf{X}) &= M_1 + M_2 + M_4 - W \end{aligned} \tag{21}$$

in which $M_1, \ldots, M_4$ are the plastic moment capacities at the joints as shown in the figure. Since the structure is a series system, the system fails if $G < 0$, where $G(\mathbf{X}) = \min\{G_1(\mathbf{X}), G_2(\mathbf{X}), G_3(\mathbf{X})\}$. The random variables considered include $\{M_1, M_2, M_3, M_4, V, W\}$. All random variables are assumed to be statistically independent with each other. The distributions of the moment capacities and the vertical load are fully known, and summarized in Table 1. However, only limited statistical information is available for the horizontal wind load W. For illustration purpose, we consider the following three representative cases of the imprecise probability information of W.

Case (1) W has a mean of 1.9 and a standard deviation of 0.45, with its distribution type unknown;

Case (2) W has a mean of 1.9 and a standard deviation of 0.45, and is strictly defined within $[1.0, 3.0]$, with its distribution type unknown;

Case (3) W has a mean within $[1.87, 1.93]$ and a standard deviation of 0.45, with its distribution type unknown.

Note that in Case 1 and 3, the wind load may take negative values for the purpose of comparison only.

4.1. Constructing the p-box for W

We first examine the CDF bounds of W constructed from different methods. For all three cases, the p-boxes for W are determined using the proposed linear programming method. As a comparison, the p-box in Case (1) is also constructed using the Chebyshev's inequality (Eq. (5)), and the equations by Ferson et al. (2003) for Case (3).

Fig. 2(a) compares the p-boxes for Case (1) obtained from the proposed method and the Chebyshev's inequality. It can be seen that the p-box from the Chebyshev's inequality is significantly wider than the p-box from linear programming. This confirms that the Chebyshev's inequality does not give the best-possible bounds, thus if it is used in reliability analysis, the obtained reliability bounds may be overly conservative. Fig. 2(b) plots the p-boxes for Case (2), obtained from the proposed linear programming, and also from the equations by Ferson et al. (2003). It is observed that the CDF bounds from the proposed method are identical to those by Ferson et al. (2003). Note that it has been proved that the equations by Ferson et al. (2003) give the best-possible CDF bounds for this case; this comparison implies that the proposed linear programming method also yields the best-possible CDF bounds.

4.2. Bounds of probability of failure

In this section, we examine the bounds of failure probability for the three cases. Table 2 presents the intervals of failure probability obtained from different methods. The second column of Table 2 gives the failure probability bounds computed by the interval Monte Carlo simulation. In this method, the probability-box of W was first constructed using the existing methods (i.e., using the equations of Oberguggenberger and Fellin (2008) for Case 1 and Ferson et al. (2003) for Case 2), and then the failure probability bounds were computed using the interval Monte Carlo method. This method is referred to as IMC1 in the following discussions. The results presented in the third column of Table 2 were also computed using the interval Monte Carlo method; however, the probability-boxes for W were constructed using the proposed linear programming method. This method is referred to as IMC2. The

Table 2: Interval bounds of failure probability.

Case No.	Interval MC (IMC1)	Interval MC (IMC2)	Direct optimization
(1)	$[0.0090, 0.3678]$	$[0.0184, 0.2593]$	$[0.0597, 0.1057]$
(2)	$[0.0223, 0.2490]$	$[0.0223, 0.2490]$	$[0.0831, 0.1106]$
(3)	–	$[0.0097, 0.4233]$	$[0.0523, 0.1918]$

fourth column of Table 2 lists the results computed by the proposed linear programming method. In this method, it is not required to construct the probability-box of W; instead, the failure probability bounds were determined directly solving the linear programming problem. For this reason, the method is referred to as "Direct Optimization". In applying the linear programming method, the conditional failure probability function, $\xi_W(w)$, was approximated first based on 10^6 Monte Carlo simulations. Then, substituting the expression of $\xi_W(w)$ into Eq. (16) yields the estimate of lower and upper bounds of $\mathscr{P}$ without the need to consider the CDF envelope of W.

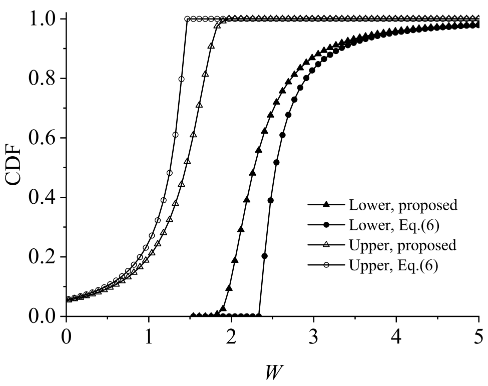
The results from IMC1 and IMC2 are firstly compared. From Table 2, it can be seen that for Case 1, the failure probability bounds from IMC2 is narrower than those from IMC1. This is to be expected, as the p-box for W from linear programming is tighter than that from the Chebyshev's inequality. For Case 2, IMC1 and IMC2 yielded the identical results, since the p-box for W is the same in both methods. For Case 3, since there is no analytical solution in the literature for constructing the CDF bounds of W, the failure probability bounds were not computed in IMC1. Next, the failure probability bounds from IMC2 and Direction Optimization method are compared. It is observed that the failure probability intervals obtained with the direct optimization method are significantly narrower than those based on interval Monte Carlo method with p-boxes. For example, the upper bound of failure probability for Case 1 is 0.1057 from direct optimization, as compared to 0.2593 from IMC2. The latter is more than twice than the former. Similar observations are also made in Case 2 and Case 3. This comparison shows that the proposed linear programming method can better utilize the available information, and yields more informative results than the interval Monte Carlo method with p-boxes.

5. Conclusions

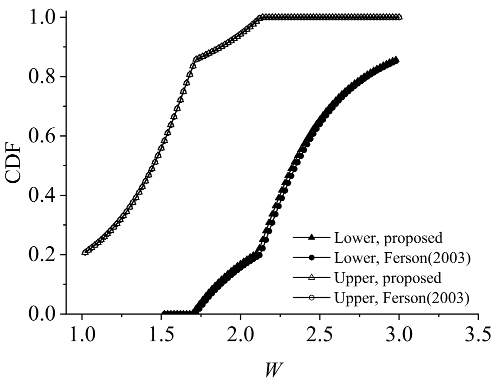
A linear programming-based method has been proposed to handle the imprecise probability problem. The proposed method has two separate but related applications: (1) to construct the best-possible CDF bounds for a random variable with limited statistical information, and (2) to estimate the lower and upper bounds of structural failure probability when one of the random variables is described with imprecise probability. The proposed method does not require the assumption of a distribution type; it considers all possible distribution types which are compatible with the data. The proposed method makes full use of the available information, without introducing additional assumptions.

It has been shown that the proposed method can yield tighter CDF bounds than the Chebyshev's inequality when only the mean and variance of the random variable are known. In the case where the min, max, mean and variance of a random variable are known, the CDF bounds from the proposed method are the same as the best-possible bounds provided in Ferson et al. (2003). The proposed method can also handle other general cases of imprecise probability, without assuming the type of distribution.

The bounds on the failure probability obtained from the direct optimization are significantly tighter than those from the interval Monte Carlo method, suggesting that more information is provided by the proposed method. The reason is that the interval Monte Carlo method only considers the CDF bounds of an imprecise probability, thus useful information "inside" the bounds may be lost in the procedure. The proposed method, on the other hand, makes full use of available information.



(a) Case (1)



(b) Case (2)

Figure 2: Lower and upper bounds of the CDF of W for Cases (1) and (2).

Acknowledgements

This research has been supported by the Faculty of Engineering and IT PhD Research Scholarship (SC1911) from the University of Sydney. This support is gratefully acknowledged.

References

Alvarez, D. A., Uribe, F., and Hurtado, J. E. (2018). "Estimation of the lower and upper bounds on the probability of failure using subset simulation and random set theory." *Mechanical Systems and Signal Processing*, 100, 782–801.

Bisgaard, T. M. and Sasvári, Z. (2000). *Characteristic Functions and Moment Sequences: Positive Definiteness in Probability*. Nova Publishers.

Der Kiureghian, A. and Ditlevsen, O. (2009). "Aleatory or epistemic? Does it matter?." *Structural Safety*, 31(2), 105–112.

Der Kiureghian, A., Lin, H., and Hwang, S. (1987). "Second-order reliability approximations." *Journal of Engineering Mechanics, ASCE*, 113(8), 1208–1225.

Ellingwood, B. R. (2005). "Risk-informed condition assessment of civil infrastructure: State of practice and research issues." *Structure and Infrastructure Engineering*, 1, 7–18.

Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D. S., and Sentz, K. (2003). "Constructing probability boxes and dempster-shafer structures." *Report No. SAND2002–4015*, Sandia National Laboratories.

Li, Q. and Ellingwood, B. R. (2008). "Damage inspection and vulnerability analysis of existing buildings with steel moment-resisting frame." *Engineering Structures*, 30, 338–351.

Li, Q., Wang, C., and Ellingwood, B. R. (2015). "Time-dependent reliability of aging structures in the presence of non-stationary loads and degradation." *Structural Safety*, 52, 132–141.

Melchers, R. E. (1999). *Structural reliability analysis and prediction*. John Wiley & Son Ltd.

Oberguggenberger, M. and Fellin, W. (2008). "Reliability bounds through random sets: Non-parametric methods and geotechnical applications." *Computers & Structures*, 86(10), 1093–1101.

Tang, W. H. and Ang, A. (2007). *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering. 2nd Ed.* Wiley, New York.

Vanderbei, R. J. (2014). *Linear Programming: Foundations and Extensions*. Springer US.

Wang, C., Li, Q., Pang, L., and Zou, A. (2016). "Estimating the time-dependent reliability of aging structures in the presence of incomplete deterioration information." *Journal of Zhejiang University Science A*, 17, 677–688.

Zhang, H. (2012). "Interval importance sampling method for finite element-based structural reliability assessment under parameter uncertainties." *Structural Safety*, 38, 1–10.

Zhang, H., Mullen, R. L., and Muhanna, R. L. (2010). "Interval Monte Carlo methods for structural reliability." *Structural Safety*, 32(3), 183–190.

Zhao, Y. and Ang, A. H. (2003). "System reliability assessment by method of moments." *Journal of Structural Engineering*, 129(10), 1341–1349.