Hybrid reliability analysis method for systems with random and non-parameterized p-boxes based on weight coefficients

Yan Ye*, Cenbo Xiong*, Biao Ma, Qinghua Zhou, Liang Yu

*School of Mechanical Engineering, Beijing Institute of Technology, China
bSchool of Aeronautics and Astronautics, Sichuan University, China

Abstract

This paper establishes a hybrid variable system failure probability optimization model based on sampling methods and weighting coefficients. By introducing auxiliary input variables, important sampling functions, and p-box, failure samples are mapped from the random variable space to the p-box variable space. The new weight coefficients are constructed, including important sampling weights and interval weights. Combining discretization methods and Monte Carlo simulation (MCS), the interval weights are transformed into variables, and constraints conforming to the p-box variable distribution are constructed. After calculating the weighting coefficients for all failure samples, the new failure probability optimization model is built. This model is independent of the performance functions and does not involve cyclic optimization, with computational complexity only related to the dimensions. Six cases are used for method comparison, validating that the new method exhibits higher efficiency and accuracy.

Keywords

hybrid reliability analysis, Non-Parameterized p-boxes; weight coefficients, optimizing equation

1. Introduction

In practical engineering, the factors influencing structural performance and security are frequently subject to inaccuracies. Many of these factors involve uncertainties, which introduce both epistemic uncertainty (lack of knowledge) and aleatory uncertainty (natural variability)[25, 34]. Therefore, conducting reliability analysis on systems with uncertain variables is a critical task in engineering.

The conventional approach to dealing with uncertain parameters is through the probability model, where the probability distributions of the uncertain parameters are explicitly defined using available information. In fact, obtaining complete information and accurately determining the probability distribution of the uncertain parameters is challenging. This limitation constitutes the primary drawback of these probability-based methods.

Furthermore, some available tools to describe the epistemic variables are often mentioned, such as Bayesian approaches[6, 12, 39], interval probabilities[22, 23, 38], fuzzy sets[40, 48], info-gap theory[2, 14], evidence theory[11, 32] and probability boxes[17, 18, 24]. Specifically, Bayesian approaches and fuzzy
sets focus on acquiring information to establish prior distributions and membership functions. Evidence theory and interval probabilities are effective in handling fragmented information. The interval model is part of the info-gap theory, allowing for the consideration of inadequately described structures. In this case, uncertain parameters are quantified using change bounds as opposed to exact probability distributions.

Probability boxes combine the advantages of probability and interval models, allowing for the consideration of both randomness and imprecision simultaneously. A p-box quantifies uncertainty using a pair of lower and upper cumulative distribution functions (CDFs). It restricts the distribution of epistemic variables to a specific range without requiring precise assumptions. In this paper, the variables are modeled by probability boxes that account for both aleatory and epistemic uncertainty. The p-box can be further classified into parameterized and non-parameterized (free) representations. Several approaches have been utilized to analyze the reliability of parametric p-boxes[18-20, 33]. Free p-boxes are a more general form, and various methods have been proposed for their analysis in recent years[9, 16, 29, 30, 37, 45, 46].

The conventional approach to handling p-box variables is through the Cartesian product method (CPM)[5]. However, the computational complexity of the CPM increases exponentially with the number of dimensions. Furthermore, a combination of line sampling and imprecise probabilities, which shows high efficiency in the cases with moderate non-linearities, was discussed by Angelis et al[7].

With the development of the sampling technique, the application of Monte Carlo simulation (MCS)[43, 44] become popular in hybrid reliability analysis under both random and p-box variables (HRA-RP)[31]. Double-Loop Sampling (DLS)[26] and Interval Monte Carlo (IMCS)[45] were proposed by Recuk et al and Zhang et al, respectively. DLS contains two layers of sampling, the outer layer samples the distribution parameters, and the inner layer samples the extracted probability distribution to calculate the failure probability. In IMCS, it gained samples of random variables and intervals of p-box variables by inverse transformation method, and calculated the extreme values over these samples and intervals. However, it was hard to deal with complex explicit performance functions and implicit performance function. Interval importance sampling[43] and interval quasi-MCS method[44] were proposed to solve the small failure problem and improve the efficiency of IMCS. On this basis, to decrease the number of performance function evaluations, surrogate models are considered, including polynomial response surface[10], support vector machine[28], neural networks[26], and Kriging[6, 35, 41, 42]. Kriging is widely used because of its advantages. Thus, the method named Active Learning Kriging And Optimization-based Interval Monte Carlo Simulation (ALK-OIMCS), and the method named Bounding-limit-state-surfaces-based Active Learning Kriging and IMCS (BLSK-IMCS) were proposed by Yang et al[38] and Zhang et al[46], respectively. They used some optimization algorithms[1, 3, 13, 15] to improve the accuracy of the model. The process of obtaining sample intervals in ALK-OIMCS and BLSK-IMCS is similar to IMCS, but the difference lies in the learning functions.

Diego et al[4] combined subset simulation with random set theory to calculate the lower and upper bounds of failure probability. It samples a series of points from random sets, calculates the boundary of failure probability and utilizes subset simulation to do further calculations. Liu et al[21] transform the original uncertain space into the standard normal space and establish two two-layer nesting optimization models to solve the extremum. Schöbi and Sudret[29] proposed imprecise structural reliability analysis (ISRA) based on multi-level metamodels[47]. The first level aims to approximate the performance function, and the second level included two independent Kriging metamodels are used to calculate the boundaries. Li et al[16] proposed an uncertainty propagation method based on the discretization of the cumulative distribution function, calculated the value range of the first four statistical moments of the response probability box, and used Johnson distribution fitting and percentage optimization to construct the boundary distribution of the response probability box. A collaborative interval quasi-Monte Carlo method (CIMCM) is presented by Xiao et al[49]. Rosen’s gradient projection method (RGPM) is utilized to solve the extreme values, and the number of repeated search iterations is reduced. A linear programming model, which is constructed by discrete CDF and used to solve the failure probability boundary, is proposed by Xie et al[36]. It solves the univariate problem by constructing linear constraints.
and decomposes the multivariable problem into a series of single-variable problems by using iterative techniques.

Accounting for all of them, the idea of finding extremums on intervals is similar. They assume that p-box variables have unrestricted distribution between the lower and upper CDFs, disregarding that the CDF of the p-box variable follows the rule of general CDF. Consequently, this perspective results in inaccurate estimations of the upper and lower bounds of failure probability. The method proposed by Xie et al[36] solves the reasonable constraint problem of univariate problems and compares it with IMCS. But it does not delve into the difference between itself and IMCS. At the same time, the method needs to call the function repeatedly to calculate the conditional failure probability in the process of iterative optimization, which reduces the computational efficiency.

In order to obtain more reasonable failure probability boundary accurately and simplify the solving process, this paper proposes a novel approach that utilizes a newly constructed weight coefficient. The main idea is to use auxiliary input functions, important sampling functions, and p-box probability density functions to map failure samples from the random space to the p-box variable space. Combining discretization methods to assign interval weights to the samples and treating them as variables to construct optimization equations. This approach not only aligns with the characteristics of p-box CDFs but also guarantees accuracy. The upper and lower bounds of failure probability are determined through the computation of an optimal solution. In addition, the performance of this method is tested by six case studies and compared with IMCS and the method proposed by Xie. The findings demonstrate that the proposed method is both accurate and effective in the domain of HRA-RP.

The rest of this article is outlined as follows. The characteristics of p-box variables, the IMCS, and its shortcomings are introduced in Section 2. Section 3 introduces the construction of the weight coefficient and optimization function of the proposed method. In Section 4, six cases studies are conducted to analyze the three methods. We make a conclusion in Section 5.

2. Hybrid Reliability Analysis with Both Random and P-box Variables

2.1. P-box Variables and Hybrid Reliability Analysis

The p-box variable is constructed by a pair of upper and lower limit cumulative distribution functions. When the cumulative distribution function of the P-box variable \( P \) is defined as \( F_P(p) \), the p-box variable is expressed as

\[
F_P(p) \leq F_P(p) \leq \bar{F}_P(p),
\]

where \( p \) is a realization of \( P \). \( F_p(p) \) and \( \bar{F}_p(p) \) are the lower and upper CDFs respectively.

![Fig. 1. Boundary curves of free p-box and realizations of the true but unknown CDF.][1]

Free p-box is only defined by Eq. (1), and it means that the free p-box can have any shape in the restricted region as long as it conforms to the characteristics of the cumulative distribution function. Fig. 1 shows some possible shapes of free p-box, such as piecewise functions (Realization 1), irregular curves with zero and infinite slopes (Realization 2) and normal distributions (Realization 3). The parameter p-box has a certain distribution in the restricted region, but the distribution parameters are unknown. Both of them are introduced clearly in[29]. The type of p-box discussed in this article is free p-box, and it is the generalization of parametric p-box and called p-box in the rest of this article. In particular, the p-box variables degenerate into random variables, if \( E_p(p) = \bar{F}_P(p) \).

According to the reliability theory of system structure, when the system contains both random variable \( X \) and p-box variable \( P \), the expression of system failure probability is defined as

\[
P(G(X, P) \leq 0) = \int_{G(X, P) \leq 0} f_{X, P}(x, p) \, dx \, dp,
\]

where \( G(X, P) \) and \( f_{X, P}(x, p) \) respectively denote the limit state...
function and the joint probability density function (PDF). The failure probability is an interval value, which can be expressed as

$$P_f = \frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} I\{G(X^{(j)}) \leq 0\},$$

(4)

where $n_{mc}$ is the total number of simulations conducted, $X^{(j)}$ represents the $j$th randomly simulated vector of basic variables, and $j=1, 2, \ldots, n_{mc}$. $I[\cdot]$ is the indicator function, having the value 1 if $[\cdot]$ is ‘true’, and the value 0 if $[\cdot]$ is ‘false’.

For calculating the lower and upper bounds of the failure probability, the Monte Carlo simulation is extended to the case where the variables are p-box variables, the new process is called Interval Monte Carlo simulation (IMCS)[40].

For limit state function $G(X, P)$, where $X$ represents random variable, $P$ represents $n$-dimensional p-box variables, and the corresponding CDFs of each p-box variable is expressed as $F_{pi}(p)\in [E_{pi}(p), \overline{F}_{pi}(p)]$, $i=1,2,\ldots,n$, the main process of interval Monte Carlo method is as follows:

**Step 1:** According to the inverse transform method and the CDF of the random variable $X$, generate random sample $X^{(j)}$,

$$X^{(j)} = F_{X^{-1}}(r^{(j)}), \quad j = 1, 2, \ldots, n_{mc},$$

(5)

where $r^{(j)}$ is the $j$th sample by [0,1] uniform distribution;

**Step 2:** Generate p-box sample $P^{(j)}$. By the inverse transform method[27], the interval $[E_{P_i}(p^{(j)}), \overline{F}_{P_i}(p^{(j)})]$ of $i$th p-box sample $P_i^{(j)}$ can be gained by

$$P_i^{(j)} = E_{P_i}(p^{(j)})$$

$$E_{P_i}(p^{(j)}) = F_{X^{-1}}(G^{(j)})$$

(6)

where $p^{(j)}=[e^{(j)}_1,\ldots,e^{(j)}_n]$ is generated by [0,1] uniform distribution.

The space formed by Eq. (6) is

$$D_{p^{(j)}}=\{[E_{P_i}(p^{(j)})]_{i=1}^{n}\}.$$  

**Step 3:** To calculate the upper and lower limits of the failure probability, the maximum and minimum values of the limit state function on this interval should be obtained by

$$\hat{G}^{(j)} = \max_{p\in D_{p^{(j)}}} G(X^{(j)}, p^{(j)})$$

$$\underline{G}^{(j)} = \min_{p\in D_{p^{(j)}}} G(X^{(j)}, p^{(j)})$$

(7)

**Step 4:** Repeat Step 1 to 3 until the maximum number $n_{mc}$ of samples is met. The upper bound on the probability of failure is determined by $\hat{G}$, while the lower bound is determined by $\underline{G}$. According to the Monte Carlo simulation, the upper and lower limits of failure probability can be expressed as

$$P_f = \frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} I\{\hat{G}^{(j)} \leq 0\} \leq P_f \leq \frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} I\{\underline{G}^{(j)} \leq 0\},$$

(8)

MCS can effectively calculate the failure probability boundary of systems with p-box variables, which has become the cornerstone of the development of random sampling methods. However, when dealing with complex system functions, IMCS’s efficiency is compromised by the requirement for numerous function calls and complex optimization calculations. Moreover, when the limit state function is expressed implicitly, the process of obtaining the maximum and minimum values necessitates intricate optimization calculations using the interval finite element method. As the number of sampling instances increases, the time required by IMCS grows exponentially. Additionally, the accuracy of IMCS’s boundary estimation depends on the precision of extreme point calculations. Considering that it is unreasonable to regard the statistical result of extreme points on each interval as the final outcome, the accuracy of IMCS is compromised in scenarios involving complex non-monotonic problems where simultaneous attainment of extreme points across intervals is unattainable.

It is evident from the discussion on p-box theory in Section 2.1 that p-box variables are not entirely unconstrained within their upper and lower limits. Instead, they possess an unknown yet distinct distribution that adheres to the distribution principles governing general random variables. When represented on the cumulative distribution function graph, the curve does not exhibit a downward trajectory. This implies that the slope of the curve at any given point is non-negative.

The sampling process of IMCS is described by taking two
adjacent samples as an example. The steps are as follows:

Step 1: Suppose there is a two-dimensional p-box variable $P=[P_1,P_2]$, according to the uniform distribution of [0 1], two samples are successively generated. Suppose that the two samples $(c^{(1)}=[c^{(1)}_1,c^{(1)}_2]$ and $c^{(2)}=[c^{(2)}_1,c^{(2)}_2])$ are close enough.

Step 2: The maximum and minimum values ($G^{(1)}$ and $G^{(2)}$) of the limit function corresponding to the first sampling are obtained according to IMCS. The sample points on the p-box variable interval corresponding to the extreme value are $P^{(1)}_{max}$ and $P^{(1)}_{min}$, respectively, $P^{(2)}_{min}$ is shown in Fig. 2.

Step 3: As shown in Fig. 2, the value intervals of samples $P^{(2)}_1$ and $P^{(2)}_2$ obtained in the second sampling $c^{(2)}$ are divided into two regions by $P^{(1)}_{min}$ and $P^{(1)}_{max}$, respectively, called $a_1,b_1$ and $a_2,b_2$. When the function is complex and $P_1$ and $P_2$ are not completely monotonic, minimum point $P^{(2)}_{min}=[P^{(2)}_{min,1},P^{(2)}_{min,2}]$ corresponding to $c^{(2)}$ can occur anywhere along the entire range.

Step 4: If $P^{(2)}_{min}$ and $P^{(2)}_{max}$ are located in $a_1$ and $a_2$, respectively, the CDFs’ images don’t show a downward trend, it also means that the two minimum points can be obtained from the same distribution. When one or both of $P^{(2)}_{min}$ and $P^{(2)}_{max}$ are located in $b_1$ and $b_2$, the CDFs’ images tend to go down, so these two extreme points cannot be in the same distribution.

The process means that the true value intervals of $P^{(1)}_1$ and $P^{(2)}_2$ are $a_1$ and $a_2$, but not $[P^{(1)}_1,P^{(2)}_1]$ and $[P^{(2)}_2,P^{(2)}_2]$. Therefore, when the minimum of the limit state function solved by IMCS optimization is less than 0, the real minimum on the real value interval may be greater than 0. Because IMCS cannot identify the real value interval well, the upper bound of failure probability is overestimated. The process of solving function maximum by IMCS optimization is similar, leading to a smaller lower bound of failure probability.

For one-dimensional p-box variables or completely monotone cases, IMCS can accurately calculate the upper and lower limits of the failure probability, avoiding the previously mentioned phenomenon. However, when confronted with two-dimensional non-monotonic p-box variables or higher-dimensional scenarios, IMCS unavoidably encounters the challenges discussed in this section. These challenges highlight areas where further improvements are necessary for IMCS.

3. New Method Based on Weight Coefficients and Important Sampling

To enhance computational efficiency and address the challenges mentioned in the preceding section, a novel reliability analysis method for systems with p-box variables is proposed. The new approach constructs an optimization model of failure probability based on failure samples and weights, enabling the calculation of upper and lower bounds for failure probability. Firstly, the p-box variables are discretized into a series of sub-intervals, and each interval is assigned a different weight index. The failure probability is given by the ratio between the sum of the weights corresponding to the failure samples and the sum of the total samples. The main steps are as follows:

Step 1: Determine auxiliary input variables for each p-box variable and compute design points using the limit state functions. Then the initial failure probability is approximated by importance sampling and the number of failure sample points is obtained.

Step 2: Employ the auxiliary input variables to discretize the p-box variables and establish corresponding weight coefficients for each interval.

Step 3: Update the number of samples based on the
maximum interval weight, coefficient of variation, and initial number of failure samples to mitigate the adverse effects of excessive weight coefficients on the results. Obtain the failure samples used to construct the failure probability equation.

Step 4: Establish the optimization function for failure probability and solve for the upper and lower bounds of failure probability through optimization techniques.

By implementing these steps, the proposed method offers a more efficient approach to reliability analysis, effectively addressing the limitations of the conventional IMCS approach. The flowchart is shown in Fig. 3. The calculation of interval weight and the construction of the failure probability equation are the basis of the whole algorithm and the key ensuring the accuracy of the results. This paper divides the entire process into two modules for presentation.

![Flowchart of the proposed method](image)

1.1. Construction of weight coefficients

In the following analysis, we focus on the incorporation of P-box variables in reliability analysis, as the inclusion of random variables can be readily implemented. In other words, the system is considered to contain only \( n \)-dimensional independent p-box variable \( P \), and the limit state function is \( G(P) \). Firstly, the auxiliary input variable\([29]\) is introduced to approximate the p-box variable, whose expression is

\[
F_P(p) = \frac{1}{2}(F_P(p) + F_P(p)).
\]  

The joint probability density function (PDF) of auxiliary input variables is denoted as \( f_\tilde{p}(p) \). The auxiliary input distribution is selected to encompass the region where the majority of the p-box variables are located. This distribution aims to capture the essential characteristics of the p-box variables to the fullest extent possible. Additionally, it serves as the basis for importance sampling of the p-box variables and serves as a reference for dividing intervals.

![Failure sample weight construction diagram](image)

As shown in Fig. 4(a), the real CDF of the p-box is \( F(p) \), and its PDF is \( f(p) \). When \( G(p^*)=0 \), \( p=p^* \) is called the critical point. According to Eq. (4), MCS calculates the failure probability by collecting failure samples. In order to obtain sufficient failure samples to construct the optimization equation.
and obtain a more accurate interval distribution of failure samples, the important sampling method is used to sample the auxiliary input function. This paper uses the same distribution type as \( f_ρ(p) \) to construct the important sampling function \( h(p) \). For example, if \( f_ρ(p) \sim N(p, \sigma) \), \( h(p) \sim N(p^*, \sigma) \). The p-box variables are sampled for the first time according to \( h(p) \), and the number of samples is \( n_{m0} \). Eq. (4) is rewritten as

\[
P_f = \frac{1}{n_{m0}} \sum_{j=1}^{n_{m0}} I\{ G(p(j)) \leq 0 \} \cdot \frac{f_ρ(p(j))}{h(p(j))}
\]

where \( j=1, ..., n_{m0} \). According to Fig. 4(b), the weight of the sample \( p^{(j)} \) composed of important samples can be obtained and written as

\[
W_{IS}^{(j)} = \frac{f_ρ(p(j))}{h(p(j))}
\]

Inspired by important sampling, the failure sample on the auxiliary input variable can be projected into the p-box space by using the real p-box variable PDF, and the weight of the failure sample is

\[
W_{IN}^{(j)} = \frac{f_ρ(p(j))}{f_p(p(j))}
\]

As shown in Fig. 1, \( f_ρ(p) \) is unknown and could equal infinity. Therefore, the weight on the neighborhood of the failure sample is used to approximate the weight of this point. The nodes in the neighborhood interval are denoted as \([L, U]\), and Eq. (12) is rewritten as

\[
W_{IN}^{(j)} = \int_{L}^{U} \frac{f_ρ(p)}{f_p(p)} dp = \frac{Pr}{\int_{L}^{U} f_p(p) dp}
\]

When \( \Delta P \) is small enough, the physical meaning of \( Pr \) in Fig. 4(a) is related to the slope \( \theta^{(j)} \) of the p-box’s CDF over the neighborhood, which is approximately calculated as

\[
Pr = F_p(U) - F_p(L) = \theta^{(j)} (U - L).
\]

According to Eq. (11) and (13), each failure sample is assigned an important sample weight and an interval weight respectively. After the weights of all failure samples are calculated, the failure probability of this group of samples projected onto the p-box space can be obtained by the simultaneous Eq. (10)–(13) as

\[
Pr = \frac{1}{n_{m0}} \sum_{j=1}^{n_{m0}} I\{ G(p^{(j)}) \leq 0 \} \cdot \frac{f_ρ(p^{(j)})}{h(p^{(j)})} \cdot \frac{Pr}{\int_{L}^{U} f_p(p) dp}
\]

Since \( F_p(p) \) and \( f_ρ(p) \) are unknown on the p-box space, Eq. (15) cannot be solved directly. By introducing the interval discretization method, the interval weights \( W_{IS}^{(j)} \) of samples \( p^{(j)} \) are reconstructed. The specific steps are as follows:

**Step 1**: \( j=1, i=1, j \) denotes the \( j \)th sampling, and \( i \) denotes the \( i \)th p-box variable;

![Fig. 5. Construction diagram of interval weight constraints.](image)

**Step 2**: Interval partition is performed for the \( i \)th p-box variable. Since the interval length of each variable is different, uniform discretization is carried out on the vertical axis of \( F_p(p) \), and the discretization number is \( n_p \). As shown in Fig. 5, the \( k \)th and \( k+1 \)th interval nodes of \( i \)th p-box are \([L_i^{(k)}, U_i^{(k)}]\) and \([L_i^{(k+1)}, U_i^{(k+1)}]\), respectively. According to Eq. (13),

\[
\int_{L_i^{(k)}}^{U_i^{(k+1)}} f_ρ(p) dp = \int_{L_i^{(k+1)}}^{U_i^{(k+1)}} f_ρ(p) dp \int_{L}^{U} f_p(p) dp
\]

and the integrals of \( f_p(p) \) over the two intervals are denoted as \( Pr_i^{(k)} \) and \( Pr_i^{(k+1)} \), \( k=1, ..., n_p-1 \);

**Step 3**: Put constraints on \( Pr_i^{(k)} \). Due to the \( i \)th p-box has bounds \( E_p(p) \) and \( \bar{E}_p(p) \), \( Pr_i^{(k)} \) can be expressed as

\[
Pr_i^{(k)} \in [Pr_{i_{\text{min}}}^{(k)}, Pr_{i_{\text{max}}}^{(k)}]
\]

According to the discussion in Section 2.2 that \( F_p(p) \) should not have a downward trend and Eq. (14),

\[
Pr_{i_{\text{min}}}^{(k)} = 0 \quad \text{and} \quad Pr_{i_{\text{max}}}^{(k)} = \bar{F}_p(U_i^{(k)}) - E_p(L_i^{(k)})
\]

When \( F_p(p) \) enters the \( k+1 \)th interval along path 1, \( Pr_i^{(k+1)} \in [E_p(L_i^{(k+1)}) - \bar{F}_p(L_i^{(k+1)})], \bar{F}_p(U_i^{(k+1)}) - E_p(L_i^{(k+1)})] \). When \( F_p(p) \) enters the \( k+1 \)th interval along path 2, \( Pr_i^{(k+1)} \in [0, \bar{F}_p(U_i^{(k+1)}) - E_p(L_i^{(k+1)})] \). After extending the slope trend of the two intervals to all intervals, the constraint equation of the value of any interval \( Pr_i^{(k)} \) is

\[
\begin{cases}
Pr_{i_{\text{min}}}^{(k)} & \frac{\int_{E_p(L_i^{(k)}) - \bar{F}_p(L_i^{(k)})}^{\bar{F}_p(U_i^{(k)})}}{Pr_{i_{\text{max}}}^{(k)} - Pr_{i_{\text{min}}}^{(k)}} \\
Pr_{i_{\text{max}}}^{(k)} & \frac{\int_{0}^{\bar{F}_p(U_i^{(k)}) - E_p(L_i^{(k)})}}{Pr_{i_{\text{max}}}^{(k)} - Pr_{i_{\text{min}}}^{(k)}}
\end{cases}
\]
where \( \overline{P}_i(U_i^{(p)}) = \overline{P}_i(U_i^{(q)}) = 1 \) and \( \overline{P}_i(L_i^{(q)}) = \overline{P}_i(L_i^{(1)}) = 0 \) and 
\[ k = 1, \ldots, n_p. \]

### Step 4: Construct the sample \( p^{(0)} \)’s interval weights. Then

\[ i = i + 1, \] when \( i = n \), all p-box variables are discretized. The corresponding interval number of \( n \)-dimensional samples \( p^{(0)} \) is 
\[ k^{(0)} = \{1, \ldots, K^{(0)}\} \], the weight due to the discrete interval is denoted as \( W_{IN}^{(k)} \), represents the ratio of the two probabilities involved in Step 2. According to Eq. (13), the sample interval weight is rewritten as
\[ W_{IN}^{(k)} = (n_p)^{1/n} \prod_{j=1}^{n} P_{r_i}^{(k_j)} \cdot \] (17)

### Step 5: \( j = j + 1 \), when \( j = n_m \), the sampling ends.

To sum up, the sample weight index consists of \( W_{IS}^{(k)} \) and \( W_{IN}^{(k)} \):
\[ W^{(j)} = W_{IS}^{(j)} W_{IN}^{(j)} = \frac{\overline{P}_i(p^{(j)})}{\underline{P}_i(p^{(j)})} (n_p)^{1/n} \prod_{j=1}^{n} P_{r_i}^{(k_j)} \cdot \] (18)
where \( P_{r_i}^{(k_j)} \) is a particular combination of \( P_{r_i}^{(k)} \). When \( P_{r_i}^{(k)} \) is treated as a group of unknown parameters, the weight can be used to convert the failure probability into a function of \( P_{r_i}^{(k)} \), and the dimension is \( n \times n_p \). 

### 3.1. Calculation of upper and lower limit of failure probability

In order to prevent the increase of interval weight index \( W_{IN}^{(k)} \) caused by too large \( P_{r_i}^{(k)} \), thus affecting the accuracy of the solution, the proposed method sets two independent sampling. After the first sampling, the initial number of failed samples \( (n_{Fa}) \) and the coefficient of variation of failure probability \( (\text{Cov}(P_i)) \) are calculated according to the results of important sampling. After calculating \( P_{r_i}^{(k)} \), the sample number is updated to be
\[ n_{mc} = \max \{ \frac{n_{mc} \max (n_p P_{r_i}^{(k)})}{n_{Fa} \text{Cov}(P_i)^2} \}, 100 \times n_p \}. \] (19)

This means making it as large as possible without compromising accuracy. The purpose is to reduce the influence of excessive weight coefficient on the results and ensure the minimum sampling number. According to Eq. (18), Eq. (15) can be rewritten as
\[ P_f = \frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} I \{ G(p^{(j)}) \leq 0 \} \cdot W^{(j)} = \frac{1}{n_{mc}} \sum_{j=1}^{n_{mc}} I \{ G(p^{(j)}) \leq 0 \} \cdot \overline{P}_i(p^{(j)}) \cdot (n_p)^{1/n} \prod_{j=1}^{n} P_{r_i}^{(k_j)} \cdot \] (20)

Eq. (20) is the equation for the variable \( P_{r_i}^{(k)} \). \( P_{r_i}^{(k)} \) has a series of linear constraints, so that the problem can be transformed into a multivariable linear constrained optimization problem. Combined with Eq. (16), its constraints \( D_{pr} \) are constructed as follows
\[
\left\{ \sum_{k=1}^{K} P_{r_i}^{(k)} \in \left[ \overline{P}_i(U_i^{(k)}), \overline{P}_i(U_i^{(k)}) \right], K = 2, \ldots, n_p - 1 \right\}
\]
\[
\left\{ P_{r_i}^{(k)} \in \left[ P_{r_i}^{(k)\max}, \right], \sum_{k=1}^{K} P_{r_i}^{(k)} = 1 \right\}
\]
where \( i = 1, \ldots, n \).

Through the optimization solution of Eq. (20) under constraints \( D_{pr} \), upper and lower bounds of failure probability can be obtained. The expression of upper and lower bounds is
\[
\begin{align*}
\overline{P}_f &= \max_{P_{r_i}^{(k)} \in D_{pr}} P_f \\
\underline{P}_f &= \min_{P_{r_i}^{(k)} \in D_{pr}} P_f \\
\end{align*}
\] (22)

When Eq. (22) is used to solve the upper and lower bounds of the failure probability, the solutions corresponding to the upper and lower bounds satisfying the convergence condition of the optimization algorithm are the optimal solutions \( \overline{P}_f \) and \( \underline{P}_f \). Using Eq. (14) and (17), the weights \( \{ \overline{W}_{IN}^{(k)} \) and \( \overline{W}_{IN}^{(k)} \) of the optimal interval and the slopes \( \{ \theta^{(k)} \) of the curve on the corresponding interval can be calculated. Given the interval nodes, the boundary curves corresponding to the upper and lower bounds of the failure probability can be obtained by linear fitting. Moreover, because the solutions \( \overline{P}_f \) and \( \underline{P}_f \) satisfies constraint Eq. (16), the rationality of the curve trend can be guaranteed, and CDFs’ trend as shown in Section 2.2 will not appear.

From the above discussion, after comparing Eq. (22) and Eq. (7), it is found that the new method uses MCS to obtain limit state function \( G(P) \)’s failure samples at one time to build optimization equations, the process of solving the failure model does not need \( G(P) \). However, IMCS needs \( G(P) \) to calculate the extreme value of \( G(P) \) in each sampling. In other words, the optimization model complexity adopted by the new method is independent of \( G(P) \)’s complexity, but is only affected by the dimensions of the model itself. The complexity of the optimization model used in IMCS is proportional to the complexity of \( G(P) \).

The proposed p-box variable analysis method has the following advantages: (1) The upper and lower limits of the failure probability can be obtained by only two optimization
searches with linear constraints. (2) It ensures the inherent characteristics of p-box variable distribution, and avoids the problem of obtaining a larger failure probability interval when IMCS deals with complex problems; (3) The approximate cumulative distribution function of the p-box variable can be obtained at the same time as the failure probability boundary; (4) The performance function does not need to be called again during the calculation.

4. Case Study and Discussion

Six different cases are selected to verify that the proposed method has higher efficiency and accuracy than IMCS and the method proposed by Xie et al [36]. It should be noted that the six cases represent different models in typical computational examples and actual projects, and they have performance functions and variable structures of different complexity levels. The application of the proposed method in these cases also indicates its universality and can be extended.

4.1. Case Study 1: Single p-box variable problem

This case is used to verify the correctness of the proposed method. \( x \) is a p-box variable, and its mean value is bounded in the intervals [1.4, 1.6] and its standard deviation is 1. The performance function is defined as

\[
G(x) = 2 - \frac{x^2 + 4}{20} + \sin\left(\frac{5x}{2}\right).
\]  

The results are shown in

Table 1, which provides the number of partition intervals \( n_p \), the lower and upper bounds of failure probability (\( P_l \) and \( P_u \)), the relative errors and the Number of Function calls (NFc).

<table>
<thead>
<tr>
<th>Method</th>
<th>( n_p )</th>
<th>( P_l ) (Error %)</th>
<th>( P_u ) (Error %)</th>
<th>NFc</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMCS</td>
<td>-</td>
<td>0.0015</td>
<td>0.0039</td>
<td>10^7</td>
</tr>
<tr>
<td>Xie’s</td>
<td>50</td>
<td>0.001(6.66)</td>
<td>0.0042(7.67)</td>
<td>200</td>
</tr>
<tr>
<td>The proposed method</td>
<td>10</td>
<td>0.0018(20)</td>
<td>0.0035(10.25)</td>
<td>4</td>
</tr>
<tr>
<td>The proposed method</td>
<td>20</td>
<td>0.0016(6.66)</td>
<td>0.0038(2.56)</td>
<td>4</td>
</tr>
<tr>
<td>The proposed method</td>
<td>30</td>
<td>0.00148(1.3)</td>
<td>0.00386(1.0)</td>
<td>4</td>
</tr>
<tr>
<td>The proposed method</td>
<td>40</td>
<td>0.00148(1.3)</td>
<td>0.00388(0.3)</td>
<td>4</td>
</tr>
</tbody>
</table>

The IMCS, Xie’s and proposed method are used for calculation, and the results are shown in Table 1. The upper and lower bounds of failure probability calculated by IMCS are 0.0015 and 0.0039, respectively. When \( n_p = 50 \), the upper and lower bounds of failure probabilities calculated by Xie’s are 0.0016 and 0.0042, and the relative errors are 6.66% and 7.67% respectively. When the proposed method is used, the results calculated at \( n_p = 40 \) are 0.00148 and 0.00388 respectively, with relative errors of 1.3% and 0.3%.

The extreme value points obtained by IMCS and the variable boundaries fitted by Xie’s and the proposed method are plotted, as shown in Fig. 6(a). It can be seen that there are obvious differences between the fitting curves and the distributions of extreme points, so the interval distribution of failure samples in the calculation process of the proposed method is further analyzed. When \( n_p = 40 \), the failure samples are concentrated in the interval 34-40. The failure point obtained by IMCS, and the variable boundary fitted by Xie’s and the proposed method in the interval 34-40 are plotted in Fig. 6(b). It can be found that the fitting curve of the proposed method is almost consistent with the distribution of IMCS failure sample points, while the fitting curve of Xie’s is significantly different. It shows that the proposed method can well approximate the variable distribution in the failure sensitive region and obtain the accurate failure probability boundary.
Fig. 6. Point distribution and fitting curves of Case 1.

The failure probability boundary curves and corresponding relative errors of the latter two methods under different discrete points are calculated and plotted, as shown in Fig. 7. It can be seen that with the increase of $n_p$, the lower and upper bounds of the failure probabilities yielded by the two methods gradually approach the reference solution, while the relative errors decrease. When $n_p<100$, the relative error of the proposed method is always below 20%, while the maximum error of Xie’s can reach 70%. According to Fig, when the relative errors of both limits are lower than 5%, the $n_p$ required by Xie's is 250, while the $n_p$ required by the proposed method is 25, which is much lower than the former. It means that the proposed method can converge to the exact boundary faster than Xie's as $n_p$ increases.

Fig. 7. Failure probability boundaries and relative errors of Case 1.

Count the number of functions calls (NFc) for each method when considering computational efficiency. IMCS uses the original function to optimize the solution in each loop, so the NFc is the number of sampling $10^5$. The NFc of Xie’s is $2 \times 2 \times 50$, concluding two optimizations, two iterations in each optimization and calculating conditional failure probabilities 50 times in each iteration. The total NFc of the proposed method is 4, of which the number of calls of the original function is 2, and the number of calls of the optimization function is 2. It is far lower than the first two methods. In the proposed method, the dimensionality of the optimization function keeps growing linearly with the increase of $n_p$. It means the increase in computation caused by $n_p$ is controllable, and the above results show that the proposed method has higher computational efficiency than the previous two methods.

4.2. Case Study 2: Nonlinear numerical problem

A nonlinear numerical problem is tested in this case. Compared with case study 1, this case considers the nonlinearity between two different types of variables. The performance function is defined as

$$G(x, y) = 2 - \frac{(x-1)(y^2+4)}{20} + \sin\left(\frac{5\pi}{2}y\right),$$

where $x$ is a normal variable and $y$ is a p-box variable, listed in Table 2.

Table 2. Distribution of random variables for case 2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Distribution type</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>Normal</td>
<td>2.5</td>
<td>1</td>
</tr>
<tr>
<td>$y$</td>
<td>Normal</td>
<td>[1.4,1.6]</td>
<td>1</td>
</tr>
</tbody>
</table>

The three methods are used to calculate and plot the failure probability curves, and the results are shown in Table 3 and Fig. 8, with three significant figures reserved. It can be seen that with the increase of $n_p$, the upper and lower bounds of the failure probabilities of the two methods gradually converge. Compared with Xie's, the proposed method has faster convergence and smaller errors. When $n_p=50$, the relative errors of the upper and lower bounds of the proposed method are 2.98% and 2.22%, respectively, which are lower than 9.29% and 4.56% of Xie’s. When $n_p=1000$, the lower and upper bounds proposed by Xie's converge to 0.204 and 0.429, which are slightly different from 0.0202 and 0.0436 calculated by IMCS. After calculating the relative error between the proposed method and the former, it is
found that the Errors (Xie 1000) are all lower than the Errors (IMCS). Both the lower bounds’ errors (Xie 1000) and upper bounds’ errors (Xie 1000) are lower than 1%. It shows that the results of the proposed method are closer to those of Xie’s. In order to explain this phenomenon, the IMCS extreme points’ distributions, the fitting variable boundaries of the other two methods, and the failure samples’ distributions of the proposed method are plotted.

Table 3. The result of case 2.

<table>
<thead>
<tr>
<th>Method</th>
<th>$n_p$</th>
<th>$P_f$</th>
<th>Lower Error (IMCS)</th>
<th>Lower Error (Xie 1000)</th>
<th>Upper Error (IMCS)</th>
<th>Upper Error (Xie 1000)</th>
<th>NFc</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.0202</td>
<td>-</td>
<td>0.0436</td>
<td>-</td>
<td>-</td>
<td>$10^5$</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.0221</td>
<td>9.29016</td>
<td></td>
<td>0.0456</td>
<td>4.56178</td>
<td></td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0204</td>
<td>1.16273</td>
<td></td>
<td>0.0430</td>
<td>1.31717</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.0204</td>
<td>0.75348</td>
<td></td>
<td>0.0429</td>
<td>1.64152</td>
<td>2000</td>
</tr>
<tr>
<td>Xie’s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.0244</td>
<td>20.75</td>
<td>19.56</td>
<td>0.0377</td>
<td>13.56</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.0226</td>
<td>11.70</td>
<td>10.6</td>
<td>0.0401</td>
<td>8.06</td>
<td>6.55</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.0213</td>
<td>5.30</td>
<td>4.25</td>
<td>0.0414</td>
<td>4.96</td>
<td>3.41</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.0209</td>
<td>3.44</td>
<td>2.43</td>
<td>0.0423</td>
<td>3.00</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.0208</td>
<td>2.98</td>
<td>1.97</td>
<td>0.0426</td>
<td>2.22</td>
<td>0.628</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.0206</td>
<td>1.91</td>
<td>0.907</td>
<td>0.0427</td>
<td>2.09</td>
<td>0.488</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>0.0206</td>
<td>1.82</td>
<td>0.822</td>
<td>0.0428</td>
<td>1.91</td>
<td>0.313</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.0206</td>
<td>1.73</td>
<td>0.729</td>
<td>0.0428</td>
<td>1.91</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.0205</td>
<td>1.62</td>
<td>0.620</td>
<td>0.0430</td>
<td>1.33</td>
<td>0.281</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.0205</td>
<td>1.36</td>
<td>0.369</td>
<td>0.0430</td>
<td>1.31</td>
<td>0.299</td>
</tr>
</tbody>
</table>

As shown in Fig. 9, it can be found that there are four distinct slope transition regions in both the extreme points’ distributions and the fitting boundaries. In these four regions, the extreme points and failure points in the two positions on the right are scattered, as described in Section 2.2. After local amplification, it can be seen that the trend of curves fitted by the proposed method is consistent with that fitted by Xie’s, and the unreasonable variables’ distributions can be avoided in both cases. Most importantly, in Case 2, the fluctuation of extreme points mainly occurs in the failure sensitive regions, so it can be considered that the difference between the results of IMCS and the other two methods is due to statistical errors caused by the distribution of extreme points. The above results show that the proposed method can approach the exact boundaries of p-box variables faster with small $n_p$, so as to obtain a tighter failure probability boundary.

In this case, the NFc of IMCS is $10^5$. When $n_p$ is 50, 500, and 1000 respectively, the NFc of Xie’s is 200, 2000, and 4000 in turn, increasing in proportion to $n_p$ and the number of variables. The proposed method always calls the function 4 times, and the optimization function dimension increases with the increase of $n_p$ and the number of variables. When $n_p=100$,
the dimension of the optimization function is 100.

4.3. Case Study 3: A mathematical problem

This Case is quoted from Xie's [36], and the performance function is

\[ G(Y_1, Y_2) = Y_2 + 0.1 \times (Y_1 - 2)^2 + Y_1 - 3 \]  \hspace{1cm} (25)

where \( Y_1 \) and \( Y_2 \) are all non-parameterized P-box variables. Their details are listed in Table 4.

Table 4. Distribution of p-box variables for case 3.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Distribution type</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_1 )</td>
<td>Normal</td>
<td>[1.76, 1.84]</td>
<td>0.3</td>
</tr>
<tr>
<td>( Y_2 )</td>
<td>Normal</td>
<td>[1.67, 1.73]</td>
<td>0.25</td>
</tr>
</tbody>
</table>

In Xie's [36], IMCS and Xie's are used to calculate Case 3. The sample number of IMCS is \( 10^6 \), and the failure probability boundary calculated by IMCS is [0.0619, 0.1196]. When the discrete number is 200, the failure probability boundary calculated by Xie's is [0.0606, 0.1171], and the relative errors are 2.03% and 2.08%, respectively. By comparing the proposed method with the two methods in Case 3, the results are shown in Table 5 and Fig. 10. As shown in Fig. 11 and Fig. 12, it can be seen that the failure extreme points, in this case, show a monotonous trend, so IMCS can obtain an accurate failure probability boundary. The weight fitting curve obtained by the proposed method can fully reflect the trend and can be regarded as the boundary curve. As shown in Table 5 and Fig. 10, with the increase of \( n_p \), the accuracy of the proposed method is gradually improved. When \( n_p > 20 \), the errors of the upper and lower failure probability calculated by the proposed method are less than 1%. Compared with the results of Case 2, it further shows that the difference between the two methods mainly comes from the difference between the extreme point and the weight fitting boundary curve. Compared with the upper and lower bound errors of Xie's, it can be seen that the proposed method can obtain more accurate results at a smaller discrete cost. Further comparing the NFe required by different methods, the proposed method has higher computational efficiency.

Table 5. The result of case 3.

<table>
<thead>
<tr>
<th>Method</th>
<th>( n_p )</th>
<th>( P_f(\text{Error} %) )</th>
<th>( P_f'(\text{Error} %) )</th>
<th>NFe</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMCS</td>
<td>-</td>
<td>0.06( \text{19} )</td>
<td>0.1( \text{198} )</td>
<td>2.09 ( \times 10^7 )</td>
</tr>
<tr>
<td>Xie's</td>
<td>50</td>
<td>0.06( \text{06}(2.03) )</td>
<td>0.18( \text{97}(1.07) )</td>
<td>7826</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.06( \text{38}(3.13) )</td>
<td>0.11( \text{8}(1.34) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.06( \text{21}(0.36) )</td>
<td>0.11( \text{86}(0.84) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.06( \text{21}(4.39) )</td>
<td>0.11( \text{98}(0.17) )</td>
<td></td>
</tr>
<tr>
<td>The proposed method</td>
<td>40</td>
<td>0.06( \text{2}(0.16) )</td>
<td>0.12( \text{0}(0.50) )</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.06( \text{16}(0.47) )</td>
<td>0.12( \text{0}(0.50) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.06( \text{17}(0.37) )</td>
<td>0.12( \text{0}(0.42) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>0.06( \text{16}(0.49) )</td>
<td>0.12( \text{0}(0.42) )</td>
<td></td>
</tr>
</tbody>
</table>
and three p-box variables. Its schematic view is shown in Fig. 13, and the details about these variables are listed in Table 6.

The performance function is given as

\[ G = 3r - \frac{2F_1}{m\omega_0^2} \sin\left(\frac{\omega_0 t_1}{2}\right) \]

where \( r \) is the yield displacement of the springs, and \( \omega_0 \) is calculated by

\[ \omega_0 = \sqrt{\frac{c_1 + c_2}{m}}. \]

Three methods are used to calculate, and IMCS extremum points and fitting curves are plotted. As shown in Fig. 14-16, the extreme value points of IMCS show a reasonable distribution trend, which means that IMCS can obtain accurate calculation results. For \( r \) and \( t_1 \), both the proposed method and Xie's can better approximate the distribution curve. But when according to \( F_1 \), the slope of IMCS maximum points’ distribution and minimum points’ distribution appear to be non-existent and zero respectively in the interval \([-0.4, 0]\). The slope variation interval of the curves fitted by the proposed method is \([-0.44, -0.08]\), which captures the change trend of the distribution curve well and achieves a better fitting effect in the whole interval. The slope variation positions of the fitting curves yielded by Xie’s are the interval \([-0.6, -0.18]\) and \( F_1=0.616 \) respectively, which has a large deviation from the distribution curves. It shows that with the increase of variables and their complexity, the proposed

### 4.4. Case Study 4: A single-degree-of-freedom oscillator

Single degree of freedom oscillator is a common structure in engineering, and the oscillator problem is analyzed with multiple variables. It is a highly nonlinear undamped single-degree-of-freedom oscillator, including three random variables

![Schematic diagram of a single-degree-of-freedom oscillator](image)

**Table 6. Distribution of variables for a single-degree-of-freedom oscillator**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Distribution type</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_1 )</td>
<td>Normal</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>( c_2 )</td>
<td>Normal</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>( m )</td>
<td>Normal</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>( r )</td>
<td>Normal</td>
<td>[0.49, 0.51]</td>
<td>0.05</td>
</tr>
<tr>
<td>( F_1 )</td>
<td>Normal</td>
<td>[-0.2, 0.2]</td>
<td>0.5</td>
</tr>
<tr>
<td>( t_1 )</td>
<td>Normal</td>
<td>[0.95, 1.05]</td>
<td>0.2</td>
</tr>
</tbody>
</table>
method can approach the variables’ boundaries more accurately.

As Table 7 listed, the upper and lower bounds of failure probability calculated by IMCS are 0.00065 and 0.0159. Calculate and plot the failure probability bounds and relative errors of the proposed method and Xie’s at different $n_p$. As shown in Fig. 17, with the increase of $n_p$, the results of the proposed method approach the IMCS results gradually, and the relative errors decrease continuously. When $n_p$ exceeds 30, the results of the proposed method converge, and the relative errors of the upper and lower bounds are less than 5%. Although the lower bound of Xie’s presents a decreasing trend with the increase of $n_p$, the results are unstable, and the relative errors remain at about 10%. The relative error level of Xie’s upper bound is lower than that of the lower bound, but the relative errors increase when $n_p$ exceeds 150. The above results show that the proposed method can obtain more accurate upper and lower bounds of failure probabilities with less discrete cost.

When $n_p=120$, the upper and lower bounds of the failure probabilities obtained by the proposed method are $6.5\times10^{-4}$ and 0.157, respectively, while the relative errors are 0.13% and 1.36%.

Table 7. The results of case 4

<table>
<thead>
<tr>
<th>Method</th>
<th>$n_p$</th>
<th>$P_f$ (Error %)</th>
<th>$\bar{P}_f$ (Error %)</th>
<th>NFe</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMCS</td>
<td>-</td>
<td>$6.5\times10^{-4}$</td>
<td>0.0159</td>
<td>10^6</td>
</tr>
<tr>
<td>Xie’s</td>
<td>50</td>
<td>8.16$\times10^{-4}$ (25.5)</td>
<td>0.0169 (6.23)</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>7.10$\times10^{-4}$ (9.18)</td>
<td>0.0155 (2.58)</td>
<td>6000</td>
</tr>
<tr>
<td></td>
<td>750</td>
<td>7.08$\times10^{-4}$ (8.91)</td>
<td>0.0154 (3.12)</td>
<td>12000</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>7.43$\times10^{-4}$ (14.3)</td>
<td>0.0142 (10.7)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>6.89$\times10^{-4}$ (5.94)</td>
<td>0.0151 (5.36)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>6.79$\times10^{-4}$ (4.50)</td>
<td>0.0152 (4.20)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>6.77$\times10^{-4}$ (4.09)</td>
<td>0.0153 (3.55)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.73$\times10^{-4}$ (3.55)</td>
<td>0.0154 (3.31)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>6.69$\times10^{-4}$ (2.98)</td>
<td>0.0154 (2.94)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>6.67$\times10^{-4}$ (2.69)</td>
<td>0.0154 (2.92)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>6.63$\times10^{-4}$ (2.05)</td>
<td>0.0155 (2.68)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>6.56$\times10^{-4}$ (0.99)</td>
<td>0.0155 (2.47)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>6.51$\times10^{-4}$ (0.12)</td>
<td>0.0156 (2.21)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>6.52$\times10^{-4}$ (0.29)</td>
<td>0.0157 (1.40)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>$6.5\times10^{-4}$ (0.13)</td>
<td>0.157 (1.36)</td>
<td>4</td>
</tr>
</tbody>
</table>

In this case, the proposed method calls the function four times. When $n_p=120$, the optimization function dimension is 3 $\times$ 120, and the calculation time is about 4.28s. The number of
IMCS calculations is $10^6$, and the calculation time is much longer than the proposed method. When $n_p=50$, the NFe of Xie’s is $2 \times 2 \times 3 \times 50 = 600$, representing two optimizations, two iterations, the number of p-box variables and $n_p$ in turn. Because

the conditional failure probability needs to be calculated repeatedly, the calculation time is also higher than that of the proposed method.

![Fig. 17. Failure probability boundaries and relative errors of Case 4.](image)

**4.5. Case Study 5: A cantilever tube**

The cantilever tube, which is used to consider the influence of variable complexity, is another typical application in engineering. Its schematic diagram is shown in Fig. 18.

There are three external forces $F_1$, $F_2$, $P$, and a torsion $T$ exerted on the tube. Because the maximum Von Mises stress $\sigma_{\text{max}}$ in the tube should be less than 180mpa, the performance function is constructed by

$$G = 180 - \sigma_{\text{max}}$$  \hspace{1cm} (28)

and the maximum Von Mises stress $\sigma_{\text{max}}$ is computed by

$$\sigma_{\text{max}} = \sqrt{\sigma_x^2 + 3\tau_{zx}^2}$$  \hspace{1cm} (29)

where is normal stress $\sigma_{\text{max}}$ and is torsional stress $\tau_{zx}$, and they are calculated by

$$\begin{cases}
\sigma_x = \frac{P + F_1 \sin(\theta_1) + F_2 \sin(\theta_2)}{A} + \frac{M d}{2l}, \\
\tau_{zx} = \frac{T d}{4l}
\end{cases}$$  \hspace{1cm} (30)

where $M$, $A$ and $I$ are, respectively, calculated by

$$\begin{cases}
M = F_1 L_1 \cos(\theta_1) + F_2 L_2 \cos(\theta_2) \\
A = \frac{\pi}{4} [d^2 - (d - 2t)^2] \\
I = \frac{\pi}{64} [d^4 - (d - 2t)^4]
\end{cases}$$  \hspace{1cm} (31)

The details regarding the random variables and p-box variables can be found in Table 8. The upper and lower bounds of IMCS are 0.0056 and 0.0167, respectively. Set $n_p=500$, the upper and lower bounds calculated by Xie’s are 0.00572 and 0.0165. The errors between the two methods are 2.14% and 1.44%. The failure boundaries and relative errors obtained by the proposed method under different $n_p$ are calculated and plotted, as shown in Fig. 19 and Table 11. When $n_p=70$, the results are 0.0052 and 0.0166. With the increase of $n_p$, the upper and lower bound of the failure probability gradually converges, and the relative errors of the results tend to decrease. When $n_p$ exceeds 30, the errors between the proposed method and the other two methods are less than 4%. The above results can support the view that the proposed method can rapidly converge to the exact boundary at small discrete cost.

---

![Fig. 18. Schematic diagram of a cantilever tube.](image)
Fig. 19. Failure probability boundaries and relative errors of Case 5.

Table 8. Distribution of variables for a cantilever tube.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Distribution type</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>Normal</td>
<td>5</td>
<td>0.1</td>
</tr>
<tr>
<td>$d$</td>
<td>Normal</td>
<td>42</td>
<td>0.5</td>
</tr>
<tr>
<td>$P(N)$</td>
<td>Normal</td>
<td>12000</td>
<td>1200</td>
</tr>
<tr>
<td>$T(N\cdot m)$</td>
<td>Normal</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>$L_2$</td>
<td>Uniform</td>
<td>[59.75,60.25]</td>
<td>-</td>
</tr>
<tr>
<td>$F_1(N)$</td>
<td>Normal</td>
<td>5000</td>
<td>500</td>
</tr>
<tr>
<td>$F_2(N)$</td>
<td>Normal</td>
<td>5000</td>
<td>500</td>
</tr>
<tr>
<td>$L_1$</td>
<td>Normal</td>
<td>[119.5,120.5]</td>
<td>1</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>Normal</td>
<td>[4.9,5.1]</td>
<td>0.9</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>Normal</td>
<td>[9.9,10.1]</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 9. The results of case 5.

<table>
<thead>
<tr>
<th>Method</th>
<th>$n_p$</th>
<th>$P_f$</th>
<th>Lower Error (IMCS)</th>
<th>Lower Error (Xie 500)</th>
<th>$\overline{P_f}$</th>
<th>Upper Error (IMCS)</th>
<th>Upper Error (Xie 500)</th>
<th>NFe</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMCS</td>
<td>-</td>
<td>0.0056</td>
<td>-</td>
<td>-</td>
<td><strong>0.0167</strong></td>
<td>-</td>
<td>-</td>
<td>$10^5$</td>
</tr>
<tr>
<td>Xie’s</td>
<td>500</td>
<td><strong>0.00572</strong></td>
<td><strong>2.14</strong></td>
<td><strong>14.1</strong></td>
<td><strong>12.1</strong></td>
<td><strong>0.0151</strong></td>
<td><strong>9.57</strong></td>
<td><strong>8.48</strong></td>
</tr>
<tr>
<td>The proposed method</td>
<td>10</td>
<td>0.00639</td>
<td>14.1</td>
<td>12.1</td>
<td>0.0151</td>
<td>9.57</td>
<td>8.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.00589</td>
<td>5.24</td>
<td>3.39</td>
<td>0.0161</td>
<td>3.57</td>
<td>2.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.00563</td>
<td>0.611</td>
<td>1.15</td>
<td>0.01633</td>
<td>2.22</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.0056</td>
<td>0.0674</td>
<td>1.82</td>
<td>0.0164</td>
<td>1.80</td>
<td>0.606</td>
<td></td>
</tr>
<tr>
<td>The proposed method</td>
<td>50</td>
<td>0.00554</td>
<td>1.09</td>
<td>2.83</td>
<td>0.0165</td>
<td>1.20</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.00554</td>
<td>1.12</td>
<td>2.85</td>
<td>0.01658</td>
<td>0.718</td>
<td>0.485</td>
<td></td>
</tr>
<tr>
<td><strong>70</strong></td>
<td><strong>0.00552</strong></td>
<td><strong>1.38</strong></td>
<td><strong>3.11</strong></td>
<td><strong>0.0166</strong></td>
<td><strong>0.599</strong></td>
<td><strong>0.606</strong></td>
<td><strong>-</strong></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 20. Extreme points distribution of Case 5.

The distributions of IMCS extreme points are plotted in Fig. 20. It can be seen that the boundaries of the three variables all appear scattered distribution, but the calculation results show that this phenomenon does not affect the calculation results of Case 4. Plot the failure point distribution and the fitting curve in Fig. 21-23. As can be seen from the figure, the failure boundaries of the three variables present a monotonous trend and do not fluctuate. It is concluded that the distribution of failure points is the main factor affecting the calculation result of IMCS. After further observation of the fitting curves, it is found that the proposed method can better reflect the changing trend of the variable failure boundaries. As shown in Błąd! Nie można odnaleźć źródła odwołania.(a), the slope variation interval of the failure points distribution is [9.46,9.67]. The slope variation interval of the curves fitted by the proposed method is [9.51,9.67], and the relative errors are 0.05 and 0. The slope variation interval of the curves fitted by Xie’s is [9.352,9.50] with 0.108 and 0.17 errors. The above results show that the fitting curve calculated by the proposed method can describe the characteristics of the variables’ failure boundaries.
more accurately, so as to obtain a more accurate upper and lower bound of the failure probability.

In addition, in this case, the proposed method has a significant efficiency improvement compared to the other two methods because the function accesses are always 4 times. The NFc of IMCS is $10^5$, and that of Xie’s is $2 \times 3 \times 3 \times 500 = 9000$, both of which far exceed the proposed method.

Fig. 21. Failure points distribution and fitting curves of $L_1$.

Fig. 22. Failure points distribution and fitting curves of $\theta_1$.
4.6. Case Study 6: A truss structure

A truss structure is a typical engineering structure, which can be analyzed by the finite element method. As shown in Fig. 24, the truss structure contains 23 bars. Eleven bars are horizontal with random cross section $A_1$ and Young’s moduli $E_1$, and the others are sloping with random cross section $A_2$ and Young’s moduli $E_2$. Six loads are applied on nodes of horizontal bars. The distribution information is listed in Table 10. When the deflection of node $s(x)$ is larger than a given threshold value of 0.113m, this structure is treated as failed. Therefore, the performance function of the structure is defined as

$$G(x) = 0.113 - |s(x)|.$$  \hspace{1cm} (32)

![Fig. 24. Schematic diagram of a truss structure.](image)

Table 10. Distribution of variables for a truss structure.

<table>
<thead>
<tr>
<th>Variables $P_1$-$P_6$</th>
<th>Distribution type</th>
<th>Mean $[5\times10^4,5.5\times10^4]$</th>
<th>Standard deviation $5\times10^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>Normal</td>
<td>$2\times10^{-3}$</td>
<td>$2\times10^{-4}$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>Normal</td>
<td>$1\times10^{-3}$</td>
<td>$1\times10^{-4}$</td>
</tr>
<tr>
<td>$E_1$</td>
<td>Normal</td>
<td>$2.1\times10^{11}$</td>
<td>$2.1\times10^{10}$</td>
</tr>
<tr>
<td>$E_2$</td>
<td>Normal</td>
<td>$2.1\times10^{11}$</td>
<td>$2.1\times10^{10}$</td>
</tr>
</tbody>
</table>

The sampling number of IMCS is set as $10^5$, and the results are 0.0128 and 0.0523. As shown in Fig. 25, it can be found that the distribution of extreme points shows a monotonous trend. The $n_p$ of Xie’s is set as 100, and the results are 0.0127 and 0.0525. As shown in Fig. 26 and Table 11, the failure probability boundaries and relative errors of the proposed method with different $n_p$ are calculated and plotted. The upper and lower bounds of the failure probability calculated by the proposed method are 0.01266 and 0.05227 at $n_p=50$. When $n_p$ exceeds 20, the relative errors are less than 1.5%. Therefore, it can be considered that the results of the proposed method are convergent and can achieve sufficient precision. It can also be concluded from the results that the proposed method has lower discrete costs and higher computational efficiency than Xie’s.

![Fig. 25. Points distribution and fitting curves of Case 6.](image)

![Fig. 26. Failure probability boundaries and relative errors of Case 6.](image)

Table 11. The results of case 6.

<table>
<thead>
<tr>
<th>Method</th>
<th>$n_p$</th>
<th>$P_f$(Error %)</th>
<th>$\overline{P_f}$(Error %)</th>
<th>NFe</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMCS</td>
<td>-</td>
<td>0.0128</td>
<td>0.0523</td>
<td>$10^5$</td>
</tr>
<tr>
<td>Xie’s</td>
<td>100</td>
<td>0.0127</td>
<td>0.0525</td>
<td>2400</td>
</tr>
<tr>
<td>The proposed method</td>
<td>10</td>
<td>0.01321</td>
<td>0.05281</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.01280</td>
<td>0.05266</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.0127</td>
<td>0.05249</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.01268</td>
<td>0.05249</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.01266</td>
<td>0.05227</td>
<td></td>
</tr>
</tbody>
</table>

4.7. Discussion

The three methods are compared and analyzed according to six cases, and the following conclusions can be drawn:

(1) The results of six cases indicate that the proposed method can obtain sufficiently accurate results at small dispersion levels. When $n_p=40$–50, the errors between the results of the proposed
method and the results of the IMCS can be kept below 5%. Xie’s usually requires more discrete numbers to achieve the same accuracy.

(2) The results of six cases also show that the weight-fitting curves generated by the proposed method accurately depict the main distribution characteristics of IMCS failure points. In Case 4 and Case 5, the proposed method can accurately capture the slope variation trend of the failure boundaries with a relative error of less than 0.1.

(3) According to Case 2 and Case 5, the results show that IMCS exists in the problems described in section 2.2, resulting in a bias in the results. Failure points are the main factor affecting the accuracy of IMCS results. IMCS results tend to be conservative, while the proposed method yields results with higher accuracy.

(4) Compared with Xie’s and IMCS, the computational efficiency of the proposed method has been significantly improved in the six cases. This is because the proposed method calls the function and optimizes fewer throughout the calculation process, avoiding the complicated optimization and iteration process.

5. Conclusions
In this paper, a Random-P-box various system failure probability optimization model based on MCS and weight coefficients is constructed, and the reliability of the system is evaluated by the proposed optimization model. The failure samples are mapped from the general random space to the p-box variable space by using the PDFs of the auxiliary input variables and the p-box variable, as well as the important sampling function. During the mapping process, two types of weights are generated: important sampling weights and interval weights. After reconstructing the interval weights using interval discretization methods, these weights are transformed into variables subject to linear constraints, and a failure probability optimization model is established. The proposed optimization model not only avoids cyclic optimization during the sampling process but also has a complexity independent of the system’s performance functions. By solving the optimization equations twice under linear constraints, upper and lower bounds of failure probabilities can be respectively calculated.

By comparing with IMCS and Xie’s, the accuracy and efficiency of the proposed method are validated across six different variable dimensions and function complexity cases. The results from each case demonstrate that the proposed method can achieve more accurate results at a lower discretization cost. The weight fitting curves obtained through the proposed method can reflect the main features of the true boundaries of the p-box. Therefore, when dealing with a single p-box variable and monotonic problems, it can ensure sufficient accuracy. For complex monotonicity issues, this method can obtain a more compact failure probability interval, reducing the overestimation of the results. In terms of computational efficiency, the proposed optimization model requires fewer optimization iterations and does not involve calls to the functional functions. Case study results also demonstrate the higher efficiency of the proposed method. In summary, the proposed method is an accurate and effective approach for p-box system reliability assessment, applicable to various computational models. This study has established the foundational model and compared it with the basic model IMCS. To further enhance the computational efficiency and applicability of the model, future work will integrate other optimization algorithms to improve the model solving process and combine it with surrogate models for engineering applications.

References


Eksplotacja i Niezawodność – Maintenance and Reliability Vol. 26, No. 4, 2024

