# A hybrid uncertainty analysis method based on random set theory

Yongsheng Duan[1, a], Jiguang Zhao[1,b], Peng Chen[1,c]

1Equipment Academy, No. 1 Bayi Road, Huairou Beijing 101416, China;

[a]duanys_vip@yeah.net, [b]jiguangzhao@yeah.net, [c]pengchen@yeah.net

**Keywords:** Hybrid uncertainty; Risk assessment; Random set theory; D-S evidence theory; Extension principle

**Abstract.** In view of the fact that hybrid uncertainty presentation and propagation with dissonance or imprecision information in risk assessment, a variable steps size discrete random set method is proposed. Various types of incomplete and inaccurate information were transformed to random set formation by the proposed framework. A hybrid uncertainty propagation model is built based on random extension principle, the lower and upper CDF of outputs are calculated. In order to decrease dissonance between uncertainty parameters, D-S evidence combination principle is used to merge the collision information. Variable step size parameter discretization strategy was proposed based on ADM and ODM for reducing the truncated tail relative errors. In conclusion, take the response of a mass–spring–damper system acted on by a harmonic forcing an example, the effectiveness and feasibility of the proposed method is validated.

## Introduction

Probability theory, as a commonly used uncertainty describing and propagation tool, its effectiveness has been suspected by many scholars [1-4]. Especially under the circumstance of inaccurate data and incomplete knowledge, epistemic uncertainty can hardly be effectively described. Moens D et. al. [5] proposed to adopt interval theory to describe specific to the epistemic uncertainty; Rohmer J et. al. [6] adopted possibility distribution to describe seismic risk model and data uncertainty to establish a seismic risk uncertainty model on the basis of fuzzy interval analysis; Agarwal H [7] put forward a uncertainty quantitative method based on evidence theory for multiple attribute optimization design. Shah H et. Al [8]. adopted evidence theory for modeling and propagation of epistemic uncertainty. In case of random and epistemic uncertainty existing in a risk model, Guyonnet D et. Al [9] proposed a hybrid uncertainty analysis framework combining with probability and possibility. Tonon et. Al [10] proposed the theory of random sets to propagate epistemic uncertainty, and compared the influence of relative error on uncertainty propagation under different discrete precisions. The above method only considered the combination of one or two theories to describe propagation stochastic and epistemic uncertainty. It doesn't establish a unified theoretical framework to study the hybrid uncertainty propagation for heterogeneous uncertainty information.

Discrete random set is a generalized random variable. Discrete random set describes probability assignment with intervals. And the interval value has the capability to describe the uncertainty. Consequently, random set theory framework has the ability to describe stochastic and epistemic uncertainty. Therefore, specific to the issues ① and ③ proposed by Ferson [11], the paper put forward a hybrid uncertainty propagation method based on variable-step-size discrete random set. In addition, it specified a uniform random set framework of heterogeneous uncertainty information, designed a random focus elements allocation strategy of a variable-step-size uncertainty information, compared and analyzed two discrete policies. D-S evidence theory integrated multi-source conflict information specific to inconsistent information. Finally, it verified the effectiveness and feasibility of the method via numerical examples.

# 1 Random set presentation of hybrid uncertainty

## 1.1 Random set presentation of probability distribution

Given a probability density function $PDF(x)$, $x$ is the variable and $x \in [m, n]$. Divide $[m, n]$ to $k$ subintervals, representing as $\Psi = \{\varphi_i = [m_i, n_i], i \in [1, 2, \cdots, k]\}$. In addition, assume discernment frame as $\Theta$, and the focal element as $\Delta_i = \{A_i = (x \in \varphi_i)\}$. The $BPA$ of focal element is:

$$m(A_i) = \int_{x \in \varphi_i} PDF(x)dx = \int_{m_i}^{n_i} PDF(x)dx \tag{1}$$

## 1.2 Random set presentation of probability box

Under the circumstance of inaccurate probability distribution parameters, generally interval value is adopted. For example, it is known that $x$ is normal distribution $N(\mu, \sigma^2)$. Where .., $\sigma \in [\chi, \gamma]$, then probability distribution curves changes to probability box (p-box).

Assume that the cumulative probability distribution function of parameter $x$ is $CDF(x)$. The upper bound and lower bound of the probability envelopes are $\overline{CDF}(x)$ and $\underline{CDF}(x)$ respectively. Then the probability envelope information can be expressed as random set by adopting ADM (Averaging Diseretization Method)or ODM (Outer Discretization Method) [10].

### 1.2.1 ADM method

Discretize the CDF range $[0,1]$ of upper and lower bounds $\overline{CDF}(x)$ and $\underline{CDF}(x)$ of p-box to $n$ subintervals. The length of each interval is $M_j > 0 (j = 1, 2, \cdots, n)$ and $M_0 = 0$. Then the value of $j$th focal element of $x$ is:

$$A_j = \left[ \overline{CDF}^{-1}\left(\sum_{s=0}^{j-1} M_s + \frac{M_j}{2}\right), \underline{CDF}^{-1}\left(\sum_{s=0}^{j-1} M_s + \frac{M_j}{2}\right) \right] \tag{2}$$

BPA is

$$m_j = M_j \tag{3}$$

### 1.2.2 ODM method

Discretize the CDF range $[0,1]$ of upper and lower bounds $\overline{CDF}(x)$ and $\underline{CDF}(x)$ of p-box to $n$ subintervals. The length of each interval is $M_j > 0 (j = 1, 2, \cdots, n)$ and $M_0 = 0$. Then the value of $j$th focal element of $x$ is

$$A_j = \left[ \overline{CDF}^{-1}\left(\sum_{s=0}^{j-1} M_s\right), \underline{CDF}^{-1}\left(\sum_{s=0}^{j-1} M_s\right) \right] \tag{4}$$

Similarly, BPA is

$$m_j = M_j \tag{5}$$

Where, $\overline{CDF}^{-1}(0) = \lim_{x \to 0^+} \overline{CDF}^{-1}(x)$, $\underline{CDF}^{-1}(0) = \lim_{x \to 0+} \underline{CDF}^{-1}(x)$.

## 1.3 Random set presentation of possibility distribution

Mihai et. Al. [12] systematically studied the description process of random set of possibility distribution, and specified the corresponding distance measuring basis.

Assume that $u$ is fuzzy set defined in discrete area $\Theta$. $\alpha_i$ is the object of fuzzy membership function $\mu_u(\theta)$, $\theta \subseteq \Theta$, in addition

$$0 < \alpha_0 < \alpha_0 < \cdots < \alpha_M \le 1 \tag{6}$$

The random set of fuzzy set $u$ is described as $(\mathscr{F}_i, m_i)$, and

$$\begin{cases} \mathscr{F}_i = \{\theta \in \Theta \mid \mu_u(\theta) \geq \alpha_i, 1 \leq i \leq M\} \\ m_i = \dfrac{\alpha_i - \alpha_{i-1}}{\alpha_i} \end{cases} \tag{7}$$

If u is the fuzzy set defined at continuous domain $\Theta$, discrete $\Theta$ to M embedded focal elements. The above discretization presents $\mu_u(\theta)$ as random set description.

### 1.4 Random set presentation of interval distribution

Specific to single interval $A = [a, b]$, it can be processed as independent focal element. In addition $m(A) = 1$, in case of n interval distribution information and unknown interval belief, each interval information can be used as an independent focal element $A_i$ and $m(A_i) = 1/n$. In case of unequal belief of interval information, generally D-S evidence synthesis formula can integrate the belief range information.

## 2 Hybrid uncertainty analysis based on variable step size discrete random set

### 2.1 Step-size assign strategy of discrete random set

The paper proposed a variable step size discretization method of probability distribution. The principle of step-size principle is sure that there are more distribution points at truncation. Without loss of generality, take the normal cumulative distribution (CDF) with symmetry characteristic as an example. The detailed procedure is:

1) The upper and lower bound of normal cumulative distributed curves are $\overline{CDF}$ and $\underline{CDF}$. Take the value domain 1/2 of CDF as the demarcation point. Set discrete interval number as $n$;

2) Set the lower part of demarcation point $[0, 1/2]$ as an example. The interval $[\underline{CDF}_{(0)}, \underline{CDF}_{(1/2)}]$ and $[\overline{CDF}_{(0)}, \overline{CDF}_{(1/2)}]$ are divided to n/2 intervals. The intersection point with coordinate axis is $\underline{\alpha}_i$ and $0 = \underline{\alpha}_1 < \underline{\alpha}_2 <, \cdots, < \underline{\alpha}_{n/2} = 1/2$;

3) Focal element $A_i = [\overline{CDF}^{-1}(\underline{\alpha}_i), \underline{CDF}^{-1}(\underline{\alpha}_{i+1})]$. The basic probability assignment is $M_i = \underline{\alpha}_{i+1} - \underline{\alpha}_i$ without loss of generality. Set focal element interval in <u>arithmetic</u> <u>sequence</u> progression, and $\underline{\alpha}_1 = d$, then

$$\underline{M}_i = \frac{i}{n(n-1)}, \quad (1 \leq i \leq n/2) \tag{8}$$

### 2.2 Hybrid uncertainty analysis framework

In risk assessment of engineering system, various incomplete and imprecise knowledges are dealt by different theoretical frameworks. The random set has the capability to processing random and incomplete information. The risk model under random set theory framework is

$$R = \varphi(\boldsymbol{X}, \boldsymbol{U}) \tag{9}$$

where as in the formula, $\boldsymbol{X}$ is the deterministic variables; $\boldsymbol{U}$ is the uncertainty variables. As being affected by the incomplete and inaccuracy, $\boldsymbol{U}$ contains random and epistemic uncertainty variables, like as probability, possibility, intervals and other forms. Due to the deterministic variable $\boldsymbol{X}$ is known, the formula (9) can be written as:

$$R = \varphi'(\boldsymbol{U}) \tag{10}$$

In this paper, the uncertainty variable $U_i$ is represented as $(\boldsymbol{\zeta}, \boldsymbol{M})$ by random sets:

$$\begin{aligned} U_i &= (\boldsymbol{\zeta}, \boldsymbol{M}) \\ &= [(\zeta_{i1}, M_{i1}), (\zeta_{i2}, M_{i2}), \cdots, (\zeta_{iN}, M_{iN})] \end{aligned} \tag{11}$$

Let the $M_i(\zeta_i)$ is the basic probability assignment of ith parameter of $U_i$. Then the focal elements of the random relation $(\zeta_R, M_R)$ of $R = \varphi'(U)$ are all Cartesian products $\zeta_R = \zeta_1 \times \zeta_2 \times \cdots \times \zeta_n$.

1) The hybrid uncertainty information acquired includes probability distribution, possibility distribution, p-box and interval distribution etc.

2) Select reasonable variable step size discretization strategy, and carry out random set description for the acquired hybrid uncertainty information.

3) Carry out synthetic fusion strategy to combine inconsistent conflict information using D-S evidence theory.

4) Random extension principle is adopted for hybrid uncertainty propagation. The propagation model is $R = \varphi'(U)$, the input random set represented as $U_i = [(\zeta_{i1}, M_{i1}), (\zeta_{i2}, M_{i2}), \cdots, (\zeta_{iN}, M_{iN})]$.

5) If relative error $\Delta E$ cannot satisfy the requirements, perform uncertainty information discrete strategy optimization. Else, turn to 6).

6) Calculate risk uncertainty intervals.

## 3 Case Study: Quality-spring-damper system response

### 3.1 Quality-spring-damper system model

The motivation of ouality-spring-damper system model is $Y \cos \omega t$ and the corresponding function is [13].

$$m\ddot{x} + c\dot{x} + kx = Y \cos \omega t \tag{12}$$

Where, m is the mass of oscillator, c is the damper viscosity, $k$ is the elasticity coefficient, $\omega$ and $D_s$ are frequency and load amplitude respectively. The steady amplification coefficient of the system is:

$$D_s = \frac{k}{\sqrt{(k - m\omega^2)^2 + (c\omega)^2}} \tag{13}$$

Parameters m, c, k and $\omega$ are mutually independent. The specific information is: m is given by symmetric triangular probability distribution, where $m_{min} = 10$, $m_{mod} = 11$ and $m_{max} = 12$. c is given by three independent intervals with equally credibility, namely $c_1 = [5, 10]$, $c_2 = [15, 20]$ and $c_3 = [25, 25]$. k is given by three independent triangular probability distribution which is as shown in Table 1:

Table 1  Uncertainty parameter of *k*.

| N | Min | mod | max |
|---|---|---|---|
| k1 | [100,110] | [160,170] | [210,220] |
| k2 | [90,110] | [150,180] | [210,220] |
| k3 | [80,120] | [130,180] | [200,230] |

In Table 1, min refers to the lower bound of probability distribution; mod refers to probability distribution mode and max refers to the upper bound of probability distribution.

$\omega$ is given by the possibility distribution, as well as $\omega_l = [2, 2.3]$, $\omega_m = [2.5, 2.7]$ and $\omega_r = [3, 3.5]$; where, l refers to the lower possibility distribution; m refers to the maximum possible value of possibility distribution; and r refers to the upper bound of the possibility distribution.

### 3.2 Random set presentation of hybrid uncertainty information

In order to analyze the uncertainty of $D_s$, it is necessary to convert the above four types of uncertainty parameters to random set description, and calculate the focal elements and its BPA.

For probability distribution parameter m, discrete $[m_{min}, m_{max}]$ to n subintervals $A_{m,i} = [a_i, b_i]$ according to the method specified in 2.1. Each subinterval $[a_i, b_i]$ is a random focal element. $p(m)$ is defined as probability density function (PFD) of m. $F_m(m)$ is the cumulative distribution function (CDF) of m. And $M_m(A_{m,i})$ is the BPA of $A_{m,i}$:

$$M_m(A_{m,i}) = \int_{A_{m,i}} p(m)dm = F_m(b_i) - F_m(a_i) \tag{14}$$

Here set the discretization precision n as 10 and 20 respectively thus to obtain random set focal elements of m and BPAs as shown in Fig. 1.
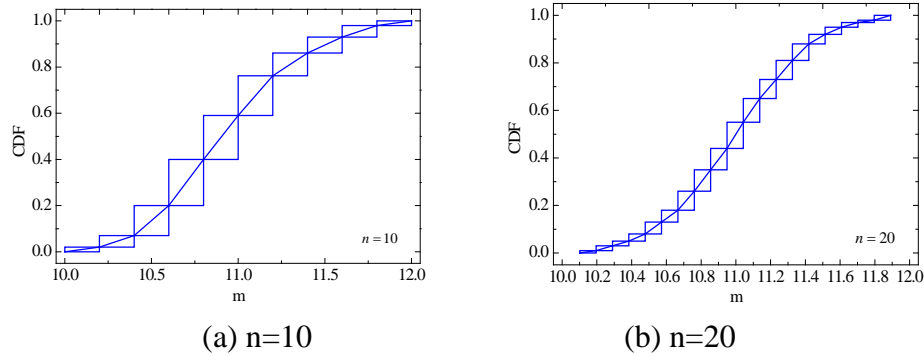


(a) n=10          (b) n=20

Fig. 1  Random set representation of m.

Under the random set description of interval parameter c, focal elements include $A_{c,1} = [5,10]$, $A_{c,2} = [5,10]$ and $A_{c,3} = [5,10]$. The interval parameters are mutually independent. The basic probability assignment is equal, namely $m(A_{c,1}) = m(A_{c,1}) = m(A_{c,1}) = 1/3$.

### 3.3 Results and discussion

Hereby, random focal elements and BPAs of four uncertainty parameters are obtained. Uniform and variable step size discretization method is adopted respectively. Uncertainty envelope curves of $D_s$ under different discrete precisions and discretization step size strategies with random mapping and are as shown in Fig. 2. The expectation $\mu$ of $D_s$ and uncertainty measurement d are as shown in Table 2. In addition, in order to compare and analyze the impact of different discrete strategies on output $D_s$.
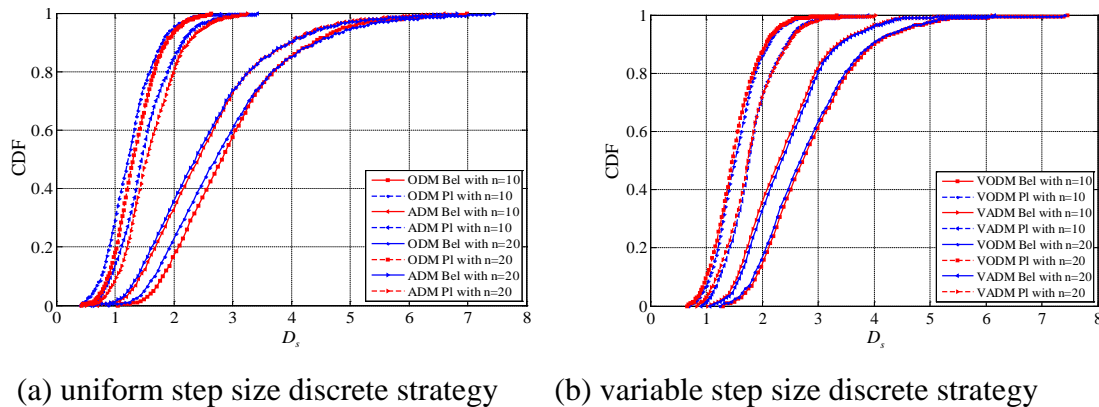


(a) uniform step size discrete strategy     (b) variable step size discrete strategy

Fig. 2  Response of $D_s$ under different discrete strategy.

Table 2 $\mu$ and $d$ under different discrete strategies

| | $\mu$ | $d$ | | $\mu$ | $d$ | $\Delta E_{inf}^{A-B}$ | $\Delta E_{inf}^{A-B}$ |
|---|---|---|---|---|---|---|---|
| $ODM_{10}$ | [1.2171,2.7169] | 2.538 | $ODM_{20}$ | [1.2218,2.7085] | 2.374 | -0.39 | 0.31 |
| $VODM_{10}$ | [1.3057,2.6934] | 2.515 | $VODM_{20}$ | [1.3074,2.6821] | 2.297 | -0.13 | 0.42 |
| $ADM_{10}$ | [1.4502,2.3996] | 2.137 | $ADM_{20}$ | [1.4557,2.3962] | 2.029 | -0.38 | 0.14 |
| $VADM_{10}$ | [1.5422,2.3732] | 2.028 | $VADM_{20}$ | [1.5441,2.3668] | 1.912 | -0.12 | 0.27 |

1) Comparing with Fig 2, it can be obtained combining with the data of Table 2 that, when the discrete step size allocation strategy unchanged, the value fine degree of random focal elements can be improved through improving the discrete precision, but it has small influence on uncertainty measurement. While the discrete precision improving by twice under VODM, the focal element number will increase 17.3 times, the relative error change is only 0.42%.

2) It can be seen from Fig. 2 that the maximum relative error of uncertainty outputs is at the left truncation part. It is caused by the error between the discrete random set and the original envelope curves under the uniform step size discretization strategy. In order to reduce truncation error, more discrete points shall be configured to gradually approach to the original envelop curve. However, with the increasing of discretization points, the calculation cost presents exponential rise. Consequently, it is necessary to optimize discrete points step strategy with the same discrete points.

3) It can be seen from Fig 2, the uncertainty measure under ADM is less than that of uncertainty measure obtained through ODM method. In addition, the outputs under ODM completely envelope the outputs under ADM, neither improving discretization accuracy nor changing discrete step allocation strategy, the enveloping nature will remain changed, namely the conservative uncertainty estimations got by ODM method.

## 4 Conclusions

In order to reduce truncation relative error, a variable step size random focal elements allocation strategy was proposed. Through comparative analysis, select reasonable discretization points and step size configuration which can effectively decrease the truncation relative error. In order to obtain more conservative assessment or design results, ODM discretization method can be adopted; in order to obtain more fine focal element description and less uncertainty measurement, ADM method can be adopted and realize through improving the discretization precision and changing the discretization step size strategy.

Finally, the feasibility and effectiveness of the proposed method in the paper was verified by a nonlinear mechanical system displacement. The method also can be applied on uncertainty analysis of reliability assessment about other practical engineering issues.

## References

[1] Baudrit C, Couso I, Dubois D. Joint propagation of probability and possibility in risk analysis: Towards a formal framework. International Journal of Approximate Reasoning 2007; 45(1):82-105.
[2] Aven T. On how to define, understand and describe risk. Reliability Engineering & System Safety 2010; 95(6):623-31.
[3] Aven T. The risk concept—historical and recent development trends. Reliability Engineering & System Safety 2012; 99(99):33–44.
[4] Bae HR, Grandhi RV, Canfield RA. An approximation approach for uncertainty quantification using evidence theory. Reliability Engineering & System Safety 2004; 86(3):215-25.
[5] Moens D, Hanss M. Review: Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: Recent advances. Finite Elements in Analysis & Design 2011; 47(1):4-16.
[6] Rohmer J, Baudrit C. The use of the possibility theory to investigate the epistemic uncertainties within scenario-based earthquake risk assessments. Natural Hazards 2011; 56(3):613-32.

[7] Agarwal H, Renaud JE, Preston EL, Padmanabhan D. Uncertainty quantification using evidence theory in multidisciplinary design optimization. Reliability Engineering & System Safety 2004; 85(1–3):281-94.

[8] Shah HR, Hosder S, Winter T. A Mixed Uncertainty Quantification Approach with Evidence Theory and Stochastic Expansions. 2013.

[9] Guyonnet D, Bourgine B, Dubois D, Fargier H, Co⋏ Me B, Chilès JP. Hybrid Approach for Addressing Uncertainty in Risk Assessments. Journal of Environmental Engineering 2003; 129(1):68-78.

[10] Tonon F. Using random set theory to propagate epistemic uncertainty through a mechanical system. Reliability Engineering [?] System Safety 2004; 85(1):169-81.

[11] Ferson S, Ginzburg LR. Different methods are needed to propagate ignorance and variability. Reliability Engineering & System Safety 1996; 54(2):133-44.

[12] Florea MC, Jousselme AL, Grenier D, Bossé É. Approximation techniques for the transformation of fuzzy sets into random sets. Fuzzy Sets & Systems 2008; 159(3):270-88.

[13] Oberkampf WL, Helton JC, Joslyn CA, Wojtkiewicz SF, Ferson S. Challenge problems: uncertainty in system response given uncertain parameters. Reliability Engineering & System Safety 2004; 85(1–3):11–9.