

# Interval-Valued Uncertainty Based on Entropy and Dempster–Shafer Theory

F. Khalaj<sup>1</sup>, E. Pasha<sup>2\*</sup>, R. Tavakkoli-Moghaddam<sup>3,4</sup>, and M. Khalaj<sup>5</sup>

<sup>1</sup>Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>2</sup>Department of Mathematics, Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran, Iran

<sup>3</sup>School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran

<sup>4</sup>LCFC, Arts et Métier Paris Tech, Metz, France

<sup>5</sup>Department of Industrial Engineering, Robat Karim Branch, Islamic Azad University, Tehran, Iran

## ARTICLE INFO

### Article History

Received 02 Aug 2017

Accepted 20 Nov 2017

### Keywords

Epistemic uncertainty

Aleatory uncertainty

Shannon entropy

Dempster–Shafer theory

Upper and lower bounds

### 2000 Mathematics Subject

Classifications

62D05, 94A20

## ABSTRACT

This paper presents a new structure as a simple method at two uncertainties (i.e., aleatory and epistemic) that result from variabilities inherent in nature and a lack of knowledge. Aleatory and epistemic uncertainties use the concept of the entropy and Dempster–Shafer (D–S) theory, respectively. Accordingly, we propose the generalized Shannon entropy in the D–S theory as a measure of uncertainty. This theory has been originated in the work of Dempster on the use of probabilities with upper and lower bounds. We describe the framework of our approach to assess upper and lower uncertainty bounds for each state of a system. In this process, the uncertainty bound is calculated with the generalized Shannon entropy in the D–S theory in different states of these systems. The probabilities of each state are interval values. In the current study, the effect of epistemic uncertainty is considered between events with respect to the non-probabilistic method (e.g., D–S theory) and the aleatory uncertainty is evaluated by using an entropy index over probability distributions through interval-valued bounds. Therefore, identification of total uncertainties shows the efficiency of uncertainty quantification.

© 2018 The Authors. Published by Atlantis Press SARL.

This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. INTRODUCTION

Several mathematical frameworks can be used to evaluate an uncertainty analysis. Probability theory is the most traditional representation of uncertainty, which is familiar to non-mathematicians. Probability is used as a representation of subjective belief, which is common in a quantitative analysis of events in different applications. Safety assessments should deal with rare events, and thus it is difficult to assess the relative frequencies of these events [1]. The Bayesian approach for the uncertainty analysis is to specify a coherent probability measure as the current state of available knowledge and uses Bayes' theorem to adjust the probabilities as new evidence, which is unveiled. Though Bayesian inference can be employed to determine the probability of decisions correctness based on prior information, it has some disadvantages, namely (1) the knowledge required to generate the prior probability distributions may not be available, (2) instabilities may occur when conflicting data are presented and/or the number of unknown propositions is large compared to the known propositions [2], (3) available information should be characterized by a specific distribution or an exact assertion of the truth of a proposition for the decision maker, and (4) Bayesian inference offers a few opportunities to express incomplete information or partial belief [3].

Imprecise probability is a generic term for any mathematical model, which measures chance or uncertainty without crisp numerical probabilities. The evidence or Dempster–Shafer (D–S) theory of belief structures is one type of imprecise probabilities, which offers a more flexible representation of uncertainty over the crisp probabilistic approach. Based on the Dempster's work, Shafer [4] proposed the D–S theory as an alternative to Bayesian inference. This theory is a generalization of the Bayesian theory and can robustly deal with incomplete data. It allows the representation of both imprecision and uncertainty [5, 6]. Rather than computing probabilities of propositions, it computes probabilities that evidence supports the propositions or alternatives, which deal with uncertainty reasoning based on incomplete information. The D–S theory tackles the prior probability issue by keeping a track of an explicit probabilistic measure of a possible lack of information. It is suitable for taking into account the disparity among knowledge types [7]. Because it is able to provide a federative framework [8] and combines cumulative evidence for changing prior opinions in the light of new evidence [9].

\* Corresponding author. Emails: pashaeinollah@yahoo.com; pasha@khu.ac.ir

The D–S theory is well developed and considered as the most appropriate alternative for uncertainty quantification, which deals with lack of knowledge. For example, Jiang *et al.* [10] proposed a new method to deal with a reliability analysis to solve engineering problems by using the D–S theory under uncertainty. Tang *et al.* [11] presented an evidential uncertainty quantification method to determine uncertainties with imprecise information, which is included in revealing the material constants of the metal fatigue crack growth model. Yin *et al.* [12] proposed an efficient algorithm to solve an acoustic problem under uncertainty and applied the Jacobi polynomial for evidence-theory based on an uncertainty analysis. Then, Yin *et al.* [13] analyzed the response of built-up systems in the mid-frequency range by using the D–S theory and a model of a finite element/statistical energy analysis under uncertainty. Jian *et al.* [14] evaluated flutter risk and the structural stability of the system, which deals with the D–S theory and uncertainty quantification.

In the probability theory, Shannon entropy is one of the best-known measures of uncertainty for a purely probabilistic system [15]. Based on the Shannon entropy concept, many studies have formulated to handle uncertainty in different fields of science. For example, Thapliyal and Taneja [16] used the concept of the Shannon entropy to evaluate the uncertainty as a measure of inaccuracy associated with distributions of order statistics. Olvera-Guerrero *et al.* [17] used the non-linear Shannon entropy to measure the uncertainty of a boiling water reactor to assess a system stability and operation.

Gu [18] introduced a new concept of entropy and multi-scale Shannon entropy and then analyzed its application to detect predictive power for the Dow Jones Industrial Average Index (DJIAI). In the fields of component lifetimes of a system, Gupta *et al.* [19] considered the Shannon entropy as a measure of the uncertainty associated with a residual lifetime distribution and acquired the properties of the residual and the entropy function of order statistics. Tahmasebi *et al.* [20] proposed a new extension of cumulative residual entropy using the Shannon entropy as a measure of the uncertainty associated with the lifetime of a system. Toomaj and Doostparast [21] introduced a new concept of a stochastic order for comparing mixed systems and used cumulative residual entropy, which measures the residual uncertainty of a random variable. They applied these concepts to evaluate component lifetimes of competing systems using the Shannon entropy, in order to handle uncertainty. The other applications are fault diagnosis [22], risk [23], etc.

Generally, frameworks, which are considered the uncertainty of the physical events or the behavior of a system, can be categorized into two types of uncertainty. Type one is uncertainty due to stochastic and irreducible variabilities, which is inherent in nature (i.e., aleatory uncertainty), and the other one is resulting from unknown physical phenomena due to a lack of knowledge (i.e., epistemic uncertainty). In Section 2, a brief introduction is provided two uncertainties. For adequate data existing in probability theory, numerous uncertain models (e.g., various types of entropies) are used as the most appropriate evaluation for aleatory uncertainty quantification. Non-probabilistic methods based on interval specifications or alternative mathematical frameworks (e.g., D–S theory) are proposed for possible better representations of epistemic uncertainty.

Obviously, in the analysis process of the system performance, uncertainty appears at different steps of analysis and the interaction between these sources of uncertainty cannot be modeled easily. Thus, in order to evaluate and reduce uncertainty, different models and modern mathematical frameworks have been proposed to quantify both aleatory and epistemic uncertainties in systems (e.g., see [24, 25]). However, the existing methods to evaluate the effect of mixed uncertainty using a piece of information and different mathematical frameworks are efficient, and modern frameworks for mathematical representation of uncertainty are still under development.

Therefore, in this paper, the concepts of D–S theory and Shannon entropy are used to handle uncertainty for systems' failure in each state considering the uncertainty bounds. On the other hand, the objective is to introduce a new technique to calculate uncertainty bounds based on the D–S theory and Shannon entropy. This paper is organized as follows. In Section 2, we introduce the type of uncertainty. Sections 3 and 4 introduce some concepts of the D–S theory and Shannon entropy as a measure of uncertainty. Section 5 describes a heuristic method for the entropy of interval-valued probabilities. It is assumed that there is a system with  $n$  different states. Conclusions are summarized in Section 6.

## 2. TYPES OF UNCERTAINTY

In the field of uncertainty quantification, uncertainty in the governing equations may assume two uncertainties (i.e., aleatory and epistemic). Aleatory uncertainty can be characterized by known probability distributions whilst epistemic uncertainty arises from a lack of knowledge of probabilistic information. It arises from the inherent variation associated with the system under consideration and is irreducible. The sources of aleatory uncertainty are typically represented by using a probabilistic framework. It is referred to inherent uncertainty due to the attribute of intrinsic randomness and probabilistic variables, which is associated with the physical system.

On the other hand, it is a result of a naturally occurring process, variability in the underlying variables or statistical variability, which is stochastic and more collected data in a given model cannot reduce this uncertainty. Thus, stochastic uncertainty in the sequence of possible events is entirely aleatory by nature. Epistemic uncertainty represents any lack of knowledge or information in any phase or activity of the modeling process [26, 27] and it is reducible through the introduction of additional information. Source of this uncertainty is of data completeness and quality. Hence, it can be reduced by more knowledge that is available. Therefore, epistemic uncertainty is reducible as results of an inaccurate scientific understanding of the natural phenomenon being modeled or data quality, completeness or incomplete knowledge of the underlying process.

Frequently, strong statistical information (e.g., probability distribution functions or high-order statistical moments) is not available. Experimental data needed to construct this information are often expensive. Thus no data or only a small collection of data points may be obtainable consequently. In these cases, expert opinion is used to handle epistemic uncertainty in conjunction with the available data to produce weak inferential estimates of parametric characteristics in a form of lower and upper bounds.

Other sources of epistemic uncertainty include limited understanding or misrepresentation of a modeled process, known commonly as a model specification uncertainty. The inclusion of enough additional information about either the model parameters or structure can lead to a reduction in the predicted uncertainty of a model output. Consequently, we can consider epistemic uncertainty as providing (conservative) bounds on an underlying aleatory uncertainty, in which reduction and convergence to the true aleatory uncertainty (or a constant value in some cases) can be obtained given sufficient additional information [28]. To learn more, the reader may find some information about evidence theory [25, 29], possibility theory [30], interval analysis [31], discrete probability distributions [32], fuzzy arithmetic [33] and probability bounds [34, 35]. These theories are different from each other in terms of characterizing the input parameter uncertainty and kind of propagation from a parameter level to model an output level. In this paper, we propose a framework for quantification of epistemic uncertainty. We calculate the upper and lower bounds for uncertainty based on entropy and D–S theory.

### 3. D–S THEORY OF EVIDENCE

The D–S theory is a mathematical theory of evidence first proposed by Dempster in early 1967 and then extended by Shafer in 1976. It makes inferences based on uncertain knowledge from different information sources. This theory allows strengthening or erosion of beliefs by combining additional sources of evidence, even in the presence of partly contradictory evidence. There are three critical functions in the D–S theory, namely the basic probability assignment function (*BPA* or  $m$ ), the belief function ( $Bel$ ), and the plausibility function ( $Pl$ ). For a finite set of mutually exclusive and exhaustive propositions (i.e.,  $\Omega$ ), a power set  $2^\Omega$  is the set of all the subsets of  $\Omega$  including itself and a null set  $\emptyset$ . The basic probability assignment is a critical variable of evidence theory and does not refer to probability in the classical sense. For any subset  $A$  of  $2^\Omega$ , the *BPA*, represented as  $m(A)$ , defines a mapping of  $2^\Omega$  to the interval between 0 and 1. Formally, this description of  $m$  can be represented by:

$$m: 2^\Omega \rightarrow [0, 1] \quad (1)$$

$$m(\emptyset) = 0 \quad (2)$$

$$\sum_{A \in 2^\Omega} m(A) = 1 \quad (3)$$

The upper and lower bounds of an interval can be defined from the *BPA*. This probability interval contains the precise probability of a set of interest in the classical sense and is bounded by two measures, namely  $Bel$  and  $Pl$ . The lower bound  $Bel$  for a set  $A$  is defined as the sum of all the *BPAs* of the proper subsets ( $B$ ) of the set of interest ( $A$ ) ( $B \subseteq A$ ). The upper bound  $Pl$  is the sum of all the *BPAs* of the sets ( $B$ ) that intersect the set of interest ( $A$ ) ( $B \cap A \neq \emptyset$ ). According to the definition above, the  $Bel$  and  $Pl$  for a set  $A$  can be described by Eqs. (4) and (5), respectively.

$$Bel(A) = \sum_{B|B \subseteq A} m(B) \quad (4)$$

$$Pl(A) = \sum_{B|A \cap B \neq \emptyset} m(B) \quad (5)$$

It is readily provable that  $Bel$  and  $Pl$  have the following relation.

$$Bel(\sim A) = Pl(A) \quad (6)$$

**Example 1.** Assuming  $\Omega = \{a, b, c\}$ ,  $2^\Omega = \{\emptyset, a, b, c, a \cup b, a \cup c, b \cup c, a \cup b \cup c\}$ , then;  $Bel(a) = m(a)$  and  $Pl(a) = m(a) + m(a \cup b) + m(a \cup c) + m(a \cup b \cup c)$ . The uncertainty of  $A$  can be represented as;  $u(A) = Pl(A) - Bel(A)$ , where  $u(A)$  is a measure of the uncertainty level of  $A$ .  $[Bel(A), Pl(A)]$  is called the confidence interval and contains more important information regarding  $A$ . *BPA* of evidence can be combined according to Dempster's rule. Considering that the subset  $A$  has two information sources  $S_1$  and  $S_2$  and let  $m_1$  and  $m_2$  be *BPAs* of sources  $S_1$  and  $S_2$ , respectively. *BPA* of the subset  $A$  can be obtained by:

$$m(A) = m(B) \oplus m(C) = \frac{1}{1-k} \sum_{B \cap C = A} m(B)m(C) \quad (7)$$

where  $k$  is the degree of conflict in two sources of evidence. It is *BPA* that the combination assigns to the null subset and represents contradictory evidence. The denominator in Dempster's rule  $1 - k$  is a normalization factor. The effect of the normalization factor  $1 - k$  in Eq. (7)

is the elimination of conflicting pieces of information between the two sources which is combined. The combined  $k$  does not exist and the two evidence sources  $S_1$  and  $S_2$  are said to be in full contradiction [36]. This study uses the D–S theory as the framework for representing the uncertainty of failure identification for states of a system.

**Example 2.** Assume that a system's failure states include  $a$ ,  $b$ , and  $c$ . The fault hypothesis set is  $\Omega = \{a, b, c\}$ . Assume that two evidences  $A$  and  $B$  are obtained. BPAs of faults supported by such evidences are  $m(A)_a = 0.7$ ,  $m(A)_{a,b} = 0.2$  and  $m(A)_{b,c} = 0.2$ ; therefore,  $m(A)_\Omega = 0.1$ , because  $m(A)_a + m(A)_{a,b} + m(A)_\Omega = 1$ , for evidence  $B$ ,  $m(B)_b = 0.4$ ,  $m(B)_{b,c} = 0.2$  and therefore  $m(B)_\Omega = 0.4$ . These two evidences do not support any other subsets of  $2^\Omega$ . The given  $m(A)_a$  value of 0.7 and  $m(A)_{a,b}$  value of 0.20 indicate that evidence  $A$  provides a belief value of 70% supporting hypothesis  $a$  and a belief value of 20% supporting hypothesis  $b$  or  $c$  and supporting hypothesis  $a$  or  $b$  or  $c$  with a belief value of 10%. The belief and plausibility functions for evidence  $A$  can be derived by:

$$\begin{aligned}
 \text{bel}(A)_a &= m(A)_a = 0.7 \\
 \text{bel}(A)_b &= m(A)_b = 0.0 \\
 \text{bel}(A)_c &= m(A)_c = 0.0 \\
 \text{bel}(A)_{a,b} &= m(A)_a + m(A)_b + m(A)_{a,b} = 0.7 + 0.0 + 0.2 = 0.9 \\
 \text{bel}(A)_{a,c} &= m(A)_a + m(A)_c + m(A)_{a,c} = 0.7 \\
 \text{bel}(A)_{b,c} &= m(A)_b + m(A)_c + m(A)_{b,c} = 0.0 \\
 \text{bel}(A)_{a,b,c} &= m(A)_a + \dots + m(A)_{b,c} = 1.0 \\
 \\ 
 \text{pl}(A)_a &= m(A)_a + m(A)_{a,b} + m(A)_{a,c} + m(A)_\Omega = 1.0 \\
 \text{pl}(A)_b &= m(A)_b + m(A)_{a,b} + m(A)_{b,c} + m(A)_\Omega = 0.3 \\
 \text{pl}(A)_c &= m(A)_c + m(A)_{a,c} + m(A)_{b,c} + m(A)_\Omega = 0.1 \\
 \text{pl}(A)_{a,b} &= m(A)_a + m(A)_b + m(A)_{a,b} + m(A)_{a,c} + m(A)_{b,c} + m(A)_\Omega = 1.0 \\
 \text{pl}(A)_{a,c} &= m(A)_a + m(A)_c + m(A)_{a,b} + m(A)_{a,c} + m(A)_{b,c} + m(A)_\Omega = 1.0 \\
 \text{pl}(A)_{b,c} &= m(A)_b + m(A)_c + m(A)_{a,b} + m(A)_{a,c} + m(A)_{b,c} + m(A)_\Omega = 0.3 \\
 \text{pl}(A)_{a,b,c} &= m(A)_a + \dots + m(A)_\Omega = 1.0
 \end{aligned} \tag{8}$$

Therefore, we have:

$$\begin{aligned}
 m(A)_a &= 0.7 & m(B)_b &= 0.4 \\
 m(A)_{a,b} &= 0.2 & m(B)_{b,c} &= 0.2 \\
 m(A)_\Omega &= 0.1 & m(B)_\Omega &= 0.4
 \end{aligned}$$

By applying the D–S theory of combination on sources of information  $A$  and  $B$ , the following data are generated based on these evidence sources as summarized in Table 1. The results are as follows.

$$\begin{aligned}
 m(C) &= m(A) \oplus m(B) \\
 k &= 0.28 + 0.14 = 0.42, \text{ degree of conflict} \\
 1 - k &= 0.58 \text{ normalization factor} \\
 \\ 
 m(C)_a &= 0.28/0.58 = 0.48 \\
 m(C)_b &= 0.08 + 0.04 + 0.04/0.58 = 0.27 \\
 m(C)_c &= 0.0; m(A)_{a,b} = \frac{0.08}{0.58} = 0.13; \\
 \\ 
 m(C)_{a,c} &= 0.0 \\
 m(C)_{b,c} &= \frac{0.02}{0.58} = 0.03; \\
 \\ 
 m(C)_\Omega &= 0.04/0.58 = 0.06
 \end{aligned}$$

Similarly, belief and plausibility functions and belief interval can be determined by using the corresponding equation described earlier, as shown in Table 2. By combining  $m(A)$  with  $m(B)$  using Eqs. (7) and (8), a new BPA,  $m(C)$  characterized for all the contain subsets.

**Table 1** | Combination of information  $A$  and  $B$ .

	$m(A)_a = 0.7$	$m(A)_b = 0.0$	$m(A)_c = 0.0$	$m(A)_{a,b} = 0.2$	$m(A)_{a,c} = 0.0$	$m(A)_{b,c} = 0.0$	$m(A)_{a,b,c} = 0.1$
$m(B)_a = 0.0$	$\{a\}$ 0.0	$\{\phi\}$ 0.0	$\{\phi\}$ 0.0	$\{a\}$ 0.0	$\{a\}$ 0.0	$\{\phi\}$ 0.0	$\{a\}$ 0.0
$m(A)_b = 0.4$	$\{\phi\}$ 0.28	$\{b\}$ 0.0	$\{\phi\}$ 0.0	$\{b\}$ 0.0	$\{\phi\}$ 0.0	$\{b\}$ 0.0	$\{b\}$ 0.0
$m(B)_c = 0.0$	$\{\phi\}$ 0.0	$\{\phi\}$ 0.0	$\{c\}$ 0.0	$\{\phi\}$ 0.0	$\{c\}$ 0.0	$\{c\}$ 0.0	$\{c\}$ 0.0
$m(B)_{a,b} = 0.0$	$\{a\}$ 0.0	$\{b\}$ 0.0	$\{\phi\}$ 0.0	$\{a, b\}$ 0.0	$\{a\}$ 0.0	$\{b\}$ 0.0	$\{a, b\}$ 0.0
$m(B)_{a,c} = 0.0$	$\{a\}$ 0.0	$\{\phi\}$ 0.0	$\{c\}$ 0.0	$\{a\}$ 0.0	$\{a, c\}$ 0.0	$\{c\}$ 0.0	$\{a, c\}$ 0.0
$m(B)_{b,c} = 0.2$	$\{\phi\}$ 0.14	$\{b\}$ 0.0	$\{c\}$ 0.0	$\{b\}$ 0.04	$\{c\}$ 0.0	$\{b, c\}$ 0.0	$\{b, c\}$ 0.02
$m(B)_{a,b,c} = 0.4$	$\{a\}$ 0.28	$\{b\}$ 0.0	$\{c\}$ 0.0	$\{a, b\}$ 0.08	$\{a, c\}$ 0.0	$\{b, c\}$ 0.0	$\Omega = 0.04$

**Table 2** | Interval-valued uncertainty.

<i>subsets</i>	<i>m</i> (C)	<i>bel</i> (C)	<i>pl</i> (C)	<i>belief interval</i>	<i>U</i> (C) = <i>pl</i> (C) – <i>bel</i> (C)
$\emptyset$	0.0	0.0	1.0	[0.0, 1.0]	0.0
{ <i>a</i> }	0.48	0.48	0.67	[0.48, 0.67]	0.19
{ <i>b</i> }	0.27	0.27	0.49	[0.27, 0.49]	0.22
{ <i>c</i> }	0.0	0.0	0.09	[0.0, 0.09]	0.09
{ <i>a, b</i> }	0.13	0.88	0.97	[0.88, 0.97]	0.09
{ <i>a, c</i> }	0.0	0.48	0.7	[0.48, 0.7]	0.22
{ <i>b, c</i> }	0.03	0.3	0.49	[0.3, 0.49]	0.19
$\Omega$	0.06	1.0	1.0	[1.0, 1.0]	0.0

Table 2 represents BPA (or mass function), belief, plausibility, belief interval and uncertainty value of the subsets. Belief intervals allow reasoning about the certainty of our beliefs or the degree of certainty. Accordingly, small differences between belief and plausibility show less certainty about our belief. Conversely, large differences show the more uncertainty about our belief.

The probability in each subset lies somewhere between belief and plausibility values. A belief function represents the evidence or information, which is supported by each subset directly. Thus, the probability cannot be less than this value. A plausibility function represents an upper bound on the degree of support that can be assigned to each subset if more specific information becomes available. It is the maximum share of the evidence, which can possibly have for all subsets. Therefore, a plausibility function is the maximum possible value of probability.

In this example, comparing the impact of the uncertainty about each subset with other uncertainties shows that, in states  $c$  ( $\{c\}$ ) and  $a$  or  $b$  ( $\{a, b\}$ ), it is the lowest uncertainty bound (expect of  $\Omega$  and  $\emptyset$ ) between all uncertainty bounds. It is equal to 0.09, while the belief degree on states  $a$  or  $b$  ( $\{a, b\}$ ) is the highest one (i.e., 0.88). On the other hand, with the minimum uncertainty, the most likely fault or failure occur in states  $a$  or  $b$ . Therefore, based on evidence from two samples, state  $a$  or  $b$  has the highest probable value of event on the system behavior. Although, a determination of a fault state in this example is obvious and simple; however, Eq. (9) may generally be considered as a comparison index to express relations between subsets using belief intervals [37], which will have a similar result based on the above reasons.

$$P(x_i > x_j) = \frac{\max\{0, pl(x_i) - bel(x_j)\} - \max\{0, bel(x_i) - pl(x_j)\}}{[pl(x_i) - bel(x_i)] + [pl(x_j) - bel(x_j)]} \quad (9)$$

## 4. SHANNON ENTROPY

Shannon [15] presented entropy as a measurement of the uncertainty level of information. The entropy is a function of the probability distribution function. Let  $x$  be a nominal attribute on a finite set  $X = \{x_1, \dots, x_n\}$  with probability distribution function  $p(x)$ . The entropy  $H(X)$  of  $X$  is defined by:

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (10)$$

The Shannon entropy can be taken as a measure of the uncertainty about the realization of a random variable. Entropy has the following properties: (1) by convention,  $0 \log 0 = 0$  and (2) in experiment,  $X = \{x_1, \dots, x_n\}$ , if  $p(x_i) = 1$  and the rest is equal to 0, then  $H(x) = 0$ , implying that there is not any uncertainty in experiment  $A$  and a decisive conclusion can be made. On the contrary, if one knows nothing about experimental results beforehand, then  $p(x_i) = 1/n$  ( $i = 1, 2, 3, \dots, n$ ),  $H(X)$  reaches its maximum value  $H^{\max}(x) = \log_2(n)$ . Obviously, the set has a maximal extent of uncertainty in this case. The Shannon entropy is used to measure the uncertainty level of evidence in this study [15].

## 5. ENTROPY AND INTERVAL-VALUED PROBABILITY

Uncertainty is usually related to lack of knowledge about future events. Thus, a definition of risk can be considered as the amount of lacking information. Therefore, in this approach, we adopt the idea that uncertainty reflects how much we do not know about the future. Entropy is the basic notion in the information theory field. Informally, we can define entropy as the measure of uncertainty of a system that at a given moment can be in one of  $n$  states, in which a set of all the possible states is defined, and the probability of some system states is known. We can define entropy formally as follows.

Let us assume that there are  $n$  different states, where a system can be in  $S_1, S_2, \dots, S_n$ . In the following equation, the probability of system  $S$  in state  $s_i$  is  $P_i$ .

$$P(S \sim s_i) = P_i, \quad i = 1, 2, \dots, n \quad (11)$$

If values of all  $P_i$  in Eq. (11) are known, then the entropy of a system can be calculated by:

$$H = - \sum_{i=1}^n P_i \log P_i \quad (12)$$

Let us consider how the entropy can be calculated if we are dealing with the interval-valued probability values. Obviously, the entropy itself will be interval-valued. In this section, we introduce a generalized definition of entropy, suitable for interval-valued probabilities. As before, we assume that system can be in  $n$  states; however, the probability that a system is in the  $i$ -th state is the interval and is equal to  $[P_i^{\min}, P_i^{\max}]$ . In this case, when probabilities are single-valued rather than interval-valued, it is required that the sum of probabilities is set to 1 as shown below.

$$\sum_i P_i = 1 \quad (13)$$

If probabilities are interval-valued, then this constraint can be rewritten by:

$$\sum_i P_i^{\min} \leq 1 \leq \sum_i P_i^{\max} \quad (14)$$

It is easy to show that Eq. (13) is a special case of Eq. (14), when  $P_i^{\min} = P_i^{\max} = P_i$  (where  $P_i$  is the probability for each  $i$ ). Moreover, if we define  $P_i^{\min} = Bel_i$  and  $P_i^{\max} = Pl_i$ , it can be shown that Eq. (16) holds.

$$Bel(x_i) \leq P(x_i) \leq Pl(x_i) \quad i = 1, 2, \dots, 2^n \quad (15)$$

$$\sum_i P(x_i) = 1$$

$$\begin{aligned} \forall i : & Bel(x_i) \leq Pl(x_i) \\ \Rightarrow & Bel(x_i) \log Bel(x_i) \\ \geq & -Pl(x_i) \log Pl(x_i) \end{aligned} \quad (16)$$

It should be noted that the states with lower probability values are more informative. Thus, we can expect that in order to calculate the upper boundary of entropy  $H^{\max}$ , we should use the lower probability bounds  $Bel_i$ . We can find the upper boundary of entropy for a system by:

$$H^{\max}(X) = - \sum_{i=1}^n Bel(x_i) \log Bel(x_i) \quad (17)$$

and the lower boundary of entropy:

$$H^{\min}(X) = - \sum_{i=1}^n Pl(x_i) \log Pl(x_i) \quad (18)$$

**Example 3.** Assume that three information sources  $A, B$  and  $C$  are under consideration with different states (e.g.,  $\Omega = \{a, b, c\}$ ). Suppose that the system's failure in state  $a$  based on three evidence sources is determined. Additionally, the belief and plausibility functions by using Eqs. (4) and (5) are respectively derived by:

$$\begin{cases} bel(A)_a = 0.7 \\ pl(A)_a = 1.0 \end{cases} \quad \begin{cases} bel(B)_a = 0.6 \\ pl(B)_a = 1.0 \end{cases} \quad \begin{cases} bel(C)_a = 0.0 \\ pl(C)_a = 0.6 \end{cases}$$

The uncertainty interval in the case at hand is as follows:

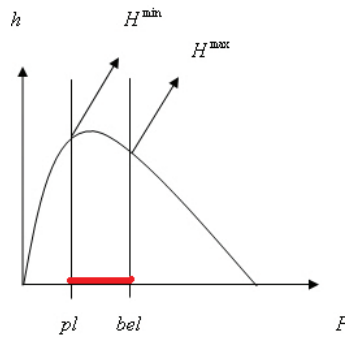
$$H^{\min} = - (0.1 \times \log 0.1 + 1.0 \times \log 1.0 + 0.6 \times \log 0.6) = 0.13$$

$$H^{\max} = - (0.7 \times \log 0.7 + 0.6 \times \log 0.6 + 0.0 \times \log 0.0) = 0.24$$

Therefore, it can be noticed that based on evidence from three samples, the uncertainty interval for the system's failure in state  $a$  is  $[0.13, 0.24]$ .

From Eq. (16), it follows that  $H^{\min} < H^{\max}$ . Moreover, as mentioned above, entropy in a system with interval-valued probability is equal to:

$$H = [H^{\min}, H^{\max}] \quad (19)$$



**Figure 1** | Upper and lower bounds of uncertainty.

The obtained definition of this interval-valued probability is a generalization of the “traditional” entropy of a system with single-valued probabilities (or mass function). Now let us examine whether the additivity feature holds for the generalized entropy in the D–S theory, as defined in Eqs. (17) and (18). If additivity holds, it means whether we have two independent systems (say,  $X$  and  $Y$ ), then entropy based on the D–S theory of a system obtained by joining systems  $X$  and  $Y$  is equal to the sum of individual generalized entropies in the D–S theory for  $X$  and  $Y$ . In other words, if additivity holds, then we have:

$$H(X, Y) = H(X) + H(Y) \quad (20)$$

If we define generalized entropy in the D–S theory as Eqs. (17) and (18), then it can be shown that we have two systems (i.e.,  $X$  and  $Y$ ) with states accordingly  $x_1, x_2, \dots, x_n$  and  $y_1, y_2, \dots, y_m$ , with  $P(X \sim x_i) = p_i$  and  $P(Y \sim y_j) = r_j$ . Then, lower and upper bounds of the two systems should be summed separately to obtain the corresponding bounds of the summed entropy.

$$\begin{aligned} H^{\min}(X, Y) &= - \left( \sum_{i=1}^n \sum_{j=1}^m pl(x_i, y_j) \log pl(x_i, y_j) \right) \\ &= - \left( \sum_{i=1}^n \sum_{j=1}^m pl(x_i, y_j) \log pl(x_i) pl(y_j|x_i) \right) \\ &= - \sum_{i=1}^n \sum_{j=1}^m pl(x_i, y_j) \log pl(x_i) - \sum_{i=1}^n \sum_{j=1}^m pl(x_i, y_j) \log pl(y_j|x_i) \\ &= - \sum_{i=1}^n pl(x_i) \log pl(x_i) - \sum_{j=1}^m pl(y_j) \log pl(y_j) \\ &\approx H^{\min}(X) + H^{\min}(Y) \end{aligned} \quad (21)$$

$$\begin{aligned} H^{\max}(X, Y) &= - \left( \sum_{i=1}^n bel(x_i) \log bel(x_i) + \sum_{j=1}^m bel(y_j) \log bel(y_j) \right) \\ &\approx H^{\max}(X) + H^{\max}(Y) \end{aligned} \quad (22)$$

Therefore, we can determine the approximate uncertainty bounds (i.e., lower and upper uncertainty bounds) of two independent systems that should be summed separately to obtain the corresponding bounds of each system. These values from each interval contribute to  $H^{\min}$  and  $H^{\max}$  are shown in Fig. 1.

In addition, the relative entropy is another important concept of entropy. It is known as a measure of distance between two probability distributions. It is evaluated divergence between probabilities or mass function of two systems  $X$  ( $X = x_1, x_2, \dots, x_n$ ) and  $Y$  ( $Y = y_1, y_2, \dots, y_m$ ), where,  $P(X \sim x_i) = p_i$  and  $P(Y \sim y_j) = r_j$ .



Let  $m(x_i)$  and  $m(y_i)$ ,  $(i = 1, 2, \dots, n)$  be two basic mass functions of systems  $X$  and  $Y$ . If these values are considered as the estimation of exact probabilities under uncertainty, then the relative entropy of measurement  $m(x_i)$  with respect to the measure  $m(y_i)$  is defined by:

$$\begin{aligned}
 H(m(X) \parallel m(Y)) &= \sum_{i=1}^n m(x_i) \log \frac{m(x_i)}{m(y_i)} \\
 &= - \sum_{i=1}^n m(x_i) \log \frac{m(y_i)}{m(x_i)} \\
 &= -E \left( \log \frac{m(y_i)}{m(x_i)} \right) \\
 &\geq -\log \left( E, \frac{m(y_i)}{m(x_i)} \right) \\
 &= -\log \left( \sum_{i=1}^n m(x_i), \frac{m(y_i)}{m(x_i)} \right) \\
 &= 0
 \end{aligned} \tag{23}$$

Obviously, if and only if  $(x_i) = m(y_i)$ ,  $H(m(X) \parallel m(Y)) = 0$ . Thus, if  $bel(x_i)$  and  $bel(y_i)$ , where  $(i = 1, 2, \dots, n)$ , are belief functions of  $X$  and  $Y$ , then we have:

$$H^{\max}(m(X) \parallel m(Y)) = \sum_{i=1}^n bel(x_i) \log \frac{bel(x_i)}{bel(y_i)} \geq 0 \tag{24}$$

Furthermore, the lower bound for the relative entropy with respect to plausibility functions  $pl(x_i)$  and  $pl(y_i)$  of  $X$  and  $Y$  is as follows:

$$H^{\min}(m(X) \parallel m(Y)) = \sum_{i=1}^n pl(x_i) \log \frac{pl(x_i)}{pl(y_i)} \geq 0$$

## 6. CONCLUSION

One way to assess the uncertainty between proposed theories under imprecise probabilities is the D–S theory. In probability theory, the best-known measure of uncertainty in a purely probabilistic system is the Shannon's entropy. This paper incorporates information and probabilities for different states of systems considered as an interval-valued probability, and then a heuristic approach is used to generalize entropy in the D–S theory for an uncertainty model. This approach is proposed to approximate upper and lower uncertainties for a presented formal problem. Since it is hard to solve this formal problem directly, the proposed method is a proper method to find an approximate solution. Accordingly, different examples are presented for the simple proposed method. The results show the efficiency of the method to find an approximate solution to quantify uncertainty bounds. Furthermore, the new definition of uncertainty bounds for two independent systems is explained, which is held by the additivity property. Then, quasi-additivity of uncertainty bounds is shown in Fig. 1. Finally, the upper and lower bounds of relative entropy are presented to evaluate the divergence between two mass functions.

## ACKNOWLEDGEMENTS

The authors would like to thank the Editor-in-Chief and autonomous reviewers for their constructive comments and helpful suggestions, which greatly improved the presentation of this paper.

## REFERENCES

1. G. Apostolakis. *Science*. 250 (1990), 1359–1364.
2. M.J. Roemer, G.J. Kacprzyński, R.F. Orsagh, *IEEE Aerospace Conf. Proc.* 6(10–17) (2001), 2979–2988.
3. M. J. Ducey, *For. Ecol. Manage.* 150 (2001), 199–211.
4. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton 1976.
5. M. Beynon, D. Cosker, D. Marshall, *Expert Syst. Appl.* 20 (2001), 357–367.
6. S.L. He'garat-Masclé, Richard D., Otte C. *Integr. Comput. Aid. Eng.* 10 (2003), 9–22.
7. V. Kaftandjian, O. Dupuis, D. Babot, Y.M. Zhu. *Pattern Recognit. Lett.* 24(1–3) (2003), 547–564.
8. S. Fabre, A. Appriou, X. Briottet. *Inform. Fusion.* 2(1) (2001), 49–71.
9. C.R. Parikh, M.J. Pont, N.B. Jones, *Pattern Recognit. Lett.* 22(6–7) (2001), 777–785.
10. C. Jiang, Z. Zhang, X. Han, J. Liu, *Comput. Struct.* 129 (2013), 1–12.



11. H. Tang, D. Li, J. Li, S. Xue, *Int. J. Fatigue*. 99(1) (2017), 163–174.
12. S. Yin, D. Yu, H. Yin, B. Xia, *Comput. Methods Appl. Mech. Eng.* 322 (2017), 419–440.
13. S. Yin, D. Yu, H. Yin, H. Lü, B. Xia, *Mech. Syst. Signal Process.* 93 (2017), 204–224.
14. T. Jian, W. Zhigang, Y. Chao, *Chinese J. Aeronaut.* 28(1) (2015), 164–171.
15. C.E. Shannon, *Bell Syst. Tech J.* 27(379–423) (1948), 623–656.
16. R. Thapliyal, H.C. Taneja, *J. Stat. Theory Appl.* 12(2) (2013), 200–207.
17. O.A. Olvera-Guerrero, P. Guerrero, G. Espinosa-Paredes, *Ann. Nucl. Energy*. 108 (2017), 1–9.
18. R. Gu, *Physica A. Stat. Mech. Appl.* 484 (2017), 215–224.
19. R.C. Gupta, H.C. Taneja, R. Thapliyal, *J. Stat. Theory Appl.* 13(1) (2014), 27–37.
20. S. Tahmasebi, M. Eskandarzadeh, A.A. Jafari, *J. Stat. Theory Appl.* 16(2) (2017), 165–177.
21. A. Toomaj, M. Doostparast, *J. Stat. Theory Appl.* 13(2) (2014), 189–195.
22. Z. Sun, P. Liu, Z. Wang, *Energy Procedia*. 105 (2017), 2354–2359.
23. R. Kiani Mavi, M. Goh, N. Kiani Mavi, *Procedia Soc. Behav. Sci.* 235(24) (2016), 216–225.
24. C.J. Roy, W.L. Oberkampf, *Comput. Methods Appl. Mech. Eng.* 200 (2011), 2131–2144.
25. G. Tang, L.P. Swiler, M.S. Eldred, G. Iaccarino, *Proceedings of the 12th AIAA Non-Deterministic Approaches Conference*, 2010.
26. F.O. Hoffman, J.S. Hammonds, *Risk Anal.* 14 (1994), 707–712.
27. G. Li, Z. Lu, L. Li, B. Ren, *Appl. Math. Model.* 40(9–10) (2016), 5703–5716.
28. J.J. Jakeman, B.M. Michael Eldred, D. Xiu, *J. Comput. Phys.* 229 (2010), 4648–4663.
29. J.C. Helton, J.D. Johnson, W.L. Oberkampf, C.B. Storlie, A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory, Technical Report 5557, Sandia National Laboratories, 2006.
30. D. Dubois, H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum, New York, 1998.
31. L.P. Swiler, T.L. Paez, R.L. Mayes, and M.S. Eldred. *Structural Dynamics, and Materials Conference Palm Springs*, CA 2009.
32. M. M. Granger, M. Henrion, *Uncertainty a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, University Press, Cambridge, 1992.
33. J.H. Purba, D.T. Sony Tjahyani, A.S. Ekariansyah, H. Tjahjono, *Ann. Nucl. Energy*. 85 (2015), 1189–1199.
34. S. Ferson, J. G. Hajagos, *Reliab. Eng. Syst. Safety*. 85(1–3) (2004), 135–152.
35. Q. Zhang, Z. Zeng, E. Zio, R. Kang, *Appl. Soft. Comput.* 56 (2017), 570–579.
36. H. Li, Y. Bao, J. Ou, *Mech. Syst. Signal Process.* 22 (2008), 1427–1440.
37. J. Guo, *J. Intell. Fuzzy Syst.* 30 (2016), 869–881.