CrossMark

Uncertainty measure in evidence theory with its applications

Xiaodan Wang[1] · Yafei Song[1]

Abstract Uncertainty measure in evidence theory supplies a new criterion to assess the quality and quantity of knowledge conveyed by belief structures. As generalizations of uncertainty measure in the probabilistic framework, several uncertainty measures for belief structures have been developed. Among them, aggregate uncertainty AU and the ambiguity measure AM are well known. However, the inconsistency between evidential and probabilistic frameworks causes limitations to existing measures. They are quite insensitive to the change of belief functions. In this paper, we consider the definition of a novel uncertainty measure for belief structures based on belief intervals. Based on the relation between evidence theory and probability theory, belief structures are transformed to belief intervals on singleton subsets, with the belief function Bel and the plausibility function Pl as its lower and upper bounds, respectively. An uncertainty measure SU for belief structures is then defined based on interval probabilities in the framework of evidence theory, without changing the theoretical frameworks. The center and the span of the interval is used to define the total uncertainty degree of the belief structure. It is proved that SU is identical to Shannon entropy and AM for Bayesian belief structures. Moreover, the proposed uncertainty measure has a wider range determined by the cardinality of discernment frame, which is more practical. Numerical examples, applications and related analyses are provided to verify the rationality of our new measure.

1 Introduction

With the sharply increasing interest in data fusion, especially for military applications, the evidence theory [3, 31], also known as the Dempster-Shafer theory, or the D-S theory for short, has attracted considerable attention for its effectiveness in modeling and fusing uncertain information. Its application has extended to many areas, such as pattern classification [24, 26], clustering analysis [27], and decision making support [2, 37]. Often presented as a generalization of probability theory, the evidence theory outperforms probability theory due to its capability in uncertainty reasoning regardless of the prior knowledge. In evidence theory, the Dempster's rule, which is famous for its commutativity and associativity, is used to combine belief functions. Although Dempster's combination rule is well-modeled theoretically, its lack of robustness is considered as a limitation by researchers in this field. This is because counterintuitive results are obtained in some cases, especially when there is high conflict among bodies of evidence. So alternative rules of combination have been proposed to improve classical Dempster's rule [7, 14, 25, 29].

Another issue involving in evidence theory is how to measure the uncertainty of belief functions, which is significant for processing and assessing evidence. Three main types of uncertainty have been identified by Klir and Yuan

✉ Yafei Song
yafei_song@163.com

Xiaodan Wang
afeu_wang@163.com

[1] Air Force Engineering University, Xi'an 710051, People's Republic of China

[21]: fuzziness, discord, and non-specificity. Generally, a fuzzy set represents fuzziness, a probability distribution represents only discord, and a classical set simply represents non-specificity. Total uncertainty in evidence theory is modeled by assigning a confidence value (being not necessarily 1) to the whole universe of discourse. This is impossible in the probability-theory framework, due to the axiom of additivity imposing that the probability of an event and the probability of its complement must sum to 1. A consequence is that a belief structure defined in the framework of evidence theory can model both non-specificity and discord, whereas only discord (and even a special kind of discord) can be modeled by a probability distribution. Therefore, in evidence theory, a belief structure (or equivalently a body of evidence) conveys two types of uncertainty: discord and non-specificity.

With the development of evidence theory, various kinds of measures for discord, non-specificity, and the total uncertainty including both of them have been proposed. The discord measure [17], the strife [17], and the confusion [13] were defined to measure the discord in evidence theory. Dubois and Prade [9] generalized the Hartley entropy [12] in the classical set theory to measure the non-specificity degree of belief structure. Yager's specificity measure [38] and Korner's specificity definition [23] can also be used to induce non-specificity measures.

The existing total uncertainty measures are represented by the aggregated measure (*AU*) [11, 28] and the ambiguity measure (*AM*) [15]. Following Maeda's work [28], Harmanec and Klir [11] proposed a measure of aggregate uncertainty they named *AU*, which was defined in the framework of evidential theory by aggregating the non-specificity and discord. It has been proved that *AU* satisfies the requirements proposed by Klir et al. [18]. However, it is completive in computing, and insensitive to changes of evidence. Moreover, these two types of uncertainty coexisting in evidence theory cannot be discriminated by *AU*. In order to overcome the insensitivity problem, Klir and Smith proposed a total uncertainty measure *TU* defined as a linear combination of *AU* and non-specificity measure. However, Jousselme et al. [15] claimed that *TU* measure could not solve the problem of computing complexity. The total uncertainty measure *TU* brings a new problem of determining the parameter δ, which is more subjective. So they presented a new measure of aggregate uncertainty for ambiguity measure (*AM*) [15]. This ambiguity measure satisfies all the requirements for general uncertainty measure. It can also overcome some of the shortcomings of *AU*. So it has been widely applied in evidence theory. However, *AM* measure is incapable of discriminating the uncertainty degrees of different belief structures with identical pignistic probability distribution. Particularly, these two cases, i.e., total ignorance and assigning equal basic probability mass to each element in the universe, are potentially different from each other, the uncertainty degrees of them are regarded as identical by the *AM* measure, which is counterintuitive.

We can note that *AU*, *AM*, and *TU* all develop from Shannon entropy [32] in probability theory. They are defined based on the conversion from the basic probability assignment function (BPA) to probability distribution according to some criteria or constraints. These measures are used to quantify uncertainty by calculating the Shannon entropy of the eliciting probability. Since they are not directly defined in the framework of evidence theory, the information loss caused by the conversion will bring limitations to these measures. Hence, they are all insensitive to the change of BBA. These limitations to some extent reflect the inconsistency between the framework of evidence theory and that of the probability theory [39]. Therefore, to avoid these limitations of traditional uncertainty measures, it is desirable to define an uncertainty measure in the framework of evidence theory, without the conversion from BPA to probability. As an attempt to improve existing uncertainty measures, Deng [6] has proposed Deng entropy to simplify the calculation of uncertainty of BPAs by considering total non-specificity and discord simultaneously.

In this paper, we analyze the uncertainty degree of BPA based on belief intervals consisting of lower and upper probabilities. We conclude that belief intervals convey both the discord and the imprecision, which are corresponding to the two parts in the uncertainty of BPA. So it is feasible to define a total uncertainty measure for a BPA base on probability intervals. In order to agree with the well-known Shannon entropy, we define the discord of belief intervals by borrowing Shannon's idea. The non-specificity of probability interval is quantified by its span.

We notice that Yang and Han [39] have defined a distance-based total uncertainty measure for BPA based on belief intervals. Deng et al. [5] have improved this measure to avoid counter-intuitive results caused by it. They defined the uncertainty measure by calculating the distance between the belief intervals and the so-called most uncertain interval. Although it can overcome some limitations in traditional uncertainty measures, the choice of distance measure is still an open problem. Moreover, this uncertainty measure is not consistent with Shannon entropy when the BPA reduces into the probability distribution. Since evidence theory is usually regarded as the extension of probability theory, it is necessary to define a unified uncertainty measure for both evidence theory and probability theory. Our proposed uncertainty measure is competent enough to quantify uncertainty degree in evidence theory and probability theory simultaneously. Since there is no switch between evidence theory and probability theory, it can overcome

these limitations in traditional measures. It also has desired properties including the probability consistency, the set consistency, the ideal value range and the monotonicity. Furthermore, the new uncertainty measure can provide more rational results when compared with the traditional ones, which can be supported by experimental results and related analyses.

The rest of this paper is organized as follows. Section 2 gives a brief recall of the evidence theory, together with introductions of interval-valued belief structure and operations on interval values. In Section 3, we briefly review the existing uncertainty measure for precise belief structure. A new uncertainty measure for interval-valued belief structures and its properties are proposed in Section 4. Numerical experiments are given to illustrate the performance of the proposed measure in Section 5. This paper is concluded in Section 6.

2 Basics of evidence theory

Dempster-Shafer evidence theory was modeled based on a finite set consisting of mutually exclusive elements, called the frame of discernment denoted by Θ [3]. The power set of Θ, denoted by 2^Θ, contains all possible unions of the sets in Θ including Θ itself. Singleton sets in a frame of discernment Θ are called atomic sets because they do not contain nonempty subsets. The following terminologies are central in the Dempster-Shafer theory.

Let $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ be the frame of discernment. A basic probability assignment (BPA) is a function m: $2^\Theta \rightarrow$ [0,1], satisfying the two following conditions:

$$m(\emptyset) = 0 \tag{1}$$

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{2}$$

where $\emptyset$ denotes empty set, and A is any subset of Θ. Such a function is also called a basic belief assignment by Smets [33], and a belief structure (BS) by Yager [36]. The terminology of belief structure will be adopted in this paper. For each subset $A \subseteq \Theta$, the value taken by the BPA at A is called the basic probability mass of A, denoted by $m(A)$.

A subset A of Θ is called the focal element of a belief structure m if $m(A) > 0$. The set of all focal elements is expressed by F = $\{A | A \subseteq \Theta, \ m(A) > 0\}$.

A Bayesian belief structure (BBS) on Θ is a belief structure on Θ whose focal elements are atomic sets (singletons) of Θ. A categorical belief structure is a normalized belief structure defined as: $m(A) = 1, \forall A \subseteq \Theta$ and $m(B) = 0, \forall B \subseteq \Theta, B \neq A$. A vacuous belief structure on Θ is defined as: $m(\Theta) = 1$ and $m(A) = 0, \forall A \neq \Theta$.

Given a belief structure m on Θ, the belief function and plausibility function which are in one-to-one correspondence with m can be defined respectively as:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{3}$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) = 1 - \sum_{B \cap A = \emptyset} m(B) \tag{4}$$

$Bel(A)$ represents all basic probability masses assigned exactly to A and its smaller subsets, and $Pl(A)$ represents all possible basic probability masses that could be assigned to A and its smaller subsets. As such, $Bel(A)$ and $Pl(A)$ can be interpreted as the lower and upper bounds of probability to which A is supported. So we can consider the belief degree of A as an interval number $BI(A) = [Bel(A), Pl(A)]$.

Definition 1 [33] The pignistic transformation maps a belief structure m to so called pignistic probability function. The pignistic transformation of a belief structure m on $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ is given by

$$BetP(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} \frac{m(B)}{1 - m(\emptyset)}, \forall A \subseteq \Theta \tag{5}$$

where $|A|$ is the cardinality of set A.

In particular, given $m(\emptyset) = 0$ and $\theta \in \Theta$, we have

$$BetP(\{\theta\}) = \sum_{\theta \in B} \frac{m(B)}{|B|}, \theta = \theta_1, \ldots, \theta_n, B \subseteq \Theta \tag{6}$$

Then we can get $Bel(A) \leq BetP(A) \leq Pl(A)$.

Definition 2 [3] Given two belief structures m_1 and m_2 on Θ, the belief structure that results from the application of Dempster's combination rule, denoted as $m_1 \oplus m_2$, or m_{12} for short, is given by:

$$m_1 \oplus m_2(A) = \begin{cases} \frac{\sum\limits_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum\limits_{B \cap C = \emptyset} m_1(B) m_2(C)}, \forall A \subseteq \Theta, A \neq \emptyset \\ 0, A = \emptyset \end{cases} \tag{7}$$

3 Existing uncertainty measures for belief structures

In evidence theory, there are two types of uncertainty, namely, the discord and the non-specificity. Some definitions of discord and non-specificity have been proposed [39]. The discord measure and non-specificity measure are also aggregated in different forms to measure the uncertainty of belief structures [15].

3.1 Measure of discord in belief structure

Shannon entropy is the classical measure of discord in probability theory. It is defined as follows.

Definition 3 (Shannon entropy [32]) Let $p = \{p_\theta | \theta \in \Theta\}$ be a probability distribution defined on discernment frame Θ. Then, the Shannon entropy is defined as:

$$S(p) = -\sum_{\theta \in \Theta} p_\theta \log_2 p_\theta \tag{8}$$

In evidence theory, measures of discord are defined to describe the randomness (or discord or conflict) in a BPA [15]. Several widely used definitions are listed below. Despite their different names, they are all proposed for measuring the discord part of the uncertainty in a belief structure.

(1) Confusion measure [13]

$$Conf(m) = -\sum_{A \subseteq \Theta} m(A) \log_2(Bel(A)) \tag{9}$$

(2) Dissonance measure [38]

$$Diss(m) = -\sum_{A \subseteq \Theta} m(A) \log_2(Pl(A)) \tag{10}$$

(3) Discord measure [18]

$$Disc(m) = -\sum_{A \subseteq \Theta} m(A) \log_2 \left[1 - \sum_{B \subseteq \Theta} m(B) \frac{|B - A|}{|B|} \right] \tag{11}$$

(4) Strife measure [17]

$$Stri(m) = -\sum_{A \subseteq \Theta} m(A) \log_2 \left[1 - \sum_{B \subseteq \Theta} m(B) \frac{|A - B|}{|A|} \right] \tag{12}$$

We can note that all of above measures are defined based on Shannon entropy. Klir and Parviz have discussed these definitions in [17], where more detailed information on these measures can be found.

3.2 Measure of non-specificity for belief structure

In probability theory, belief is only assigned singleton subsets of discernment frame. So there is merely discord type uncertainty in the probabilistic framework. Compared with probability theory, basic probability masses in evidence theory may be focused on the subsets with more than one element, which will bring the non-specificity type uncertainty [10, 23]. Non-specificity means two or more alternatives

are left unspecified and represents an imprecision degree. In classical set theory, the only measurable kind of uncertainty is the non-specificity, which is related to its cardinality. The Hartley measure is regarded as the standard measure of non-specificity for classical sets.

Definition 4 (Hartley measure [12]) Let Θ be a frame of discernment, and let A be any subset of Θ. Then, the Hartley measure is defined as:

$$H(A) = \log_2(|A|) \tag{13}$$

In evidence theory, non-specificity is only related to the focal element whose cardinality is larger than one. Considering the basic probability mass on the focal elements, researchers have defined some non-specificity measures [9, 23, 38]. The most commonly used non-specificity measure for belief structures is developed from Hartley measure. It is defined as [9]

$$NS(A) = \sum_{A \subseteq \Theta} m(A) \log_2(|A|) \tag{14}$$

When the m is a Bayesian belief structure, i.e., it only has singleton focal elements, $NS(A)$ reaches the minimum value 0. When m is a categorical belief structure, i.e., $m(A) = 1$, $NS(A) = \log_2(|A|)$, which coincides with the Hartley measure for classical sets. Specially, if m is a vacuous belief structure, i.e., $m(\Theta) = 1$, it reaches the maximum value $\log_2(|\Theta|)$. This definition was proved to have uniqueness by Ramer [30] and it satisfies all the requirements of the non-specificity measure.

3.3 Total uncertainty measure in evidence theory

Based on these two kinds of uncertainty measures, Klir and Wierman gave some axiomatic requirements for general uncertainty measure of belief structures [20]. Harmanec and Klir [11] defined an aggregated uncertainty (AU) to quantify the total uncertainty of a belief structure.

$$AU(m) = \max \left[-\sum_{\theta \in \Theta} p_\theta \log_2 p_\theta \right]$$
$$s.t. \begin{cases} p_\theta \in [0, 1], \forall \theta \in \Theta \\ \sum_{\theta \in \Theta} p_\theta = 1 \\ Bel(A) \leq \sum_{\theta \in A} p_\theta \leq 1 - Bel(\bar{A}), \forall A \subseteq \Theta \end{cases} \tag{15}$$

The definition of AU illustrates that it is the maximum Shannon entropy of all probability distributions under the constraints corresponding to the given BPA. Therefore, it is also called as "upper entropy" [1]. It is can be seen as an aggregated total uncertainty (ATU) measure, which captures both non-specificity and discord. It can be proved that AU satisfies all the requirements for uncertainty measure in evidence theory, including probability consistency, set consistency,

value range, monotonicity, sub-additivity and additivity for the joint BBA in Cartesian space [20]. So it is regarded as a well-justified measure of uncertainty for Dempster-Shafer theory. However, this measure is stuck with several shortcomings, i.e., high computing complexity, highly insensitive to changes of evidence, and no distinction between the two types of uncertainty (discord and non-specificity) [9]. In order to overcome the sensitivity problem, Klir and Smith proposed a total uncertainty measure *TU* defined as a linear combination of *AU* and non-specificity measure *N* [19].

Definition 5 (Total Uncertainty [19]) Let *m* be a BPA defined on a discernment frame Ω, then the total uncertainty $TU(m, \delta)$ is a linear combination of the non-specificity measure $N(m)$ and $AU(m)$ as follows:

$$TU(m, \delta) = \delta AU(m) + (1 - \delta)N(m) \tag{16}$$

where $\delta \in [0, 1]$ is a constant, viewed as a discounting factor of *AU*.

It can be proved that *TU* satisfies all the requirements for uncertainty measure. However, Jousselme et al. [15] pointed that *TU* measure could not solve the problem of computing complexity, and bought a new problem with the choice of the linear parameter δ. So they presented an alternative measure to *AM* for quantifying ambiguity of belief structures [15].

Definition 6 (Ambiguity Measure [15]) Let Θ be a frame of discernment with *n* elements, $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, and let *m* be a BPA defined on Θ. An ambiguity measure *AM* can be defined as:

$$AM(m) = -\sum_{\theta \in \Theta} Bet P_m(\{\theta\}) \log_2 (Bet P_m(\{\theta\})) \tag{17}$$

where $Bet P_m$ is the pignistic probability distribution of *m*.

Besides satisfying all the requirements for general uncertainty measures, the ambiguity measure also overcomes some of the shortcomings of the *AU* measure. The *AM* measure has been a popular uncertainty measure for belief structures thanks to its good performance. However, it is not sensitive to the change of belief structures, particularly, it cannot discriminate the uncertainty degrees of different belief structures with identical pignistic probability distribution. This will be demonstrated by the following example.

3.4 Analysis on available total uncertainty measures

The traditional total uncertainty measures have their own draw-backs. *AM* cannot satisfy the sub-additivity (for joint BBA in Cartesian space) which has been pointed out by Klir

and Lewis [22]. Moreover, *AM* is criticized due to its logical non-monotonicity [22]. Some examples will be presented to analyze the performance of existing total uncertainty measures.

Example 1 The discernment frame is denoted as $\Theta = \{\theta_1, \theta_2, \theta_3\}$. Two BPAs are given as:

$m_1(\{\theta_1, \theta_2\}) = 1/3, m_1(\{\theta_1, \theta_3\}) = 1/2, m_1(\{\theta_2, \theta_3\}) = 1/6;$

$m_2(\{\theta_1, \theta_2, \theta_3\}) = 1/3, m_2(\{\theta_1, \theta_3\}) = 1/2, m_2(\{\theta_2, \theta_3\}) = 1/6.$

The belief function and plausibility function of m_1 and m_2 can be obtained as follows:

$Bel_1(\{\theta_1\}) = 0, \quad Pl_1(\{\theta_1\}) = 5/6;$

$Bel_1(\{\theta_2\}) = 0, \quad Pl_1(\{\theta_2\}) = 1/2;$

$Bel_1(\{\theta_3\}) = 0, \quad Pl_1(\{\theta_1\}) = 2/3;$

$Bel_1(\{\theta_1, \theta_2\}) = 1/3, \quad Pl_1(\{\theta_1, \theta_2\}) = 1;$

$Bel_1(\{\theta_1, \theta_3\}) = 1/2, \quad Pl_1(\{\theta_1, \theta_3\}) = 1;$

$Bel_1(\{\theta_2, \theta_3\}) = 1/6, \quad Pl_1(\{\theta_2, \theta_3\}) = 1;$

$Bel_1(\Theta) = 1, \quad Pl_1(\Theta) = 1.$

$Bel_2(\{\theta_1\}) = 0, \quad Pl_2(\{\theta_1\}) = 5/6;$

$Bel_2(\{\theta_2\}) = 0, \quad Pl_2(\{\theta_2\}) = 1/2;$

$Bel_2(\{\theta_3\}) = 0, \quad Pl_2(\{\theta_1\}) = 1;$

$Bel_2(\{\theta_1, \theta_2\}) = 0, \quad Pl_2(\{\theta_1, \theta_2\}) = 1;$

$Bel_2(\{\theta_1, \theta_3\}) = 1/2, \quad Pl_2(\{\theta_1, \theta_3\}) = 1;$

$Bel_2(\{\theta_2, \theta_3\}) = 1/6, \quad Pl_2(\{\theta_2, \theta_3\}) = 1;$

$Bel_2(\Theta) = 1, \quad Pl_2(\Theta) = 1.$

From these belief and plausibility functions, we can get:

$Bel_2(A) \le Bel_1(A), \quad Pl_2(A) \ge Pl_1(A), \quad \forall A \subseteq \Theta.$

The belief intervals corresponding to m_1 and m_2 can be expressed by $BI_1(A) = [Bel_1(A), Pl_1(A)]$ and $BI_2(A) = [Bel_2(A), Pl_2(A)]$, respectively. Then we have $BI_1(A) \subseteq BI_2(A), \forall A \subseteq \Theta$. That is to say, all the belief intervals of m_1 can be contained by those corresponding belief intervals of m_2, which means that m_2 has a higher level of uncertainty. However, their corresponding *AM* and *AU* listed below are counter-intuitive.

$AM(m_1) = 1.5546, \quad AM(m_2) = 1.5100;$

$AU(m_1) = AU(m_2) = \log_2 3 = 1.5850$

We can see that *AM* decreases the total uncertainty of m_2 in this example although there is a clear increment of uncertainty of m_2 compared with that of m_1. The *AU* measure does not bring any decrease on the uncertainty of m_2, but it generates two equal values, which is also counter-intuitive, since the uncertainty degree of m_2 should be higher than that of m_1.

Taking a closer examination of *AU*, we can note that *AU* tries to find a probability to maximize the Shannon entropy.

In this example, the Shannon entropy gets its maximum value $\log_2 3 = 1.5850$ when the probability masses assigned to all singletons are equal, i.e., $p(\{\theta_1\}) = p(\{\theta_2\}) = p(\{\theta_3\}) = 1/3$. Moreover, the uniform probability distribution $p(\theta_1) = p(\theta_2) = p(\theta_3) = 1/3$ satisfies all the constraints (shown in Eq. 15) established by Bel_1, Pl_1 and Bel_2, Pl_2. Therefore, both AU_1 and AU_2 take the maximum Shannon entropy value $\log_2 3 = 1.5850$.

Example 2 Let $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ denote the discernment frame. Three belief structures on Θ are given as:

$$m_1(\Theta) = 1;$$

$$m_2(\{\theta_1, \theta_2\}) = 0.5, \quad m_2(\{\theta_3, \theta_4\}) = 0.5;$$

$$m_3(\{\theta_1\}) = m_3(\{\theta_2\}) = m_3(\{\theta_3\}) = m_3(\{\theta_4\}) = 0.25.$$

It is easy to get that $Bet P_m(\{\theta_1\}) = Bet P_m(\{\theta_2\}) = Bet P_m(\{\theta_3\}) = Bet P_m(\{\theta_4\}) = 0.25$ for all of these belief structures, i.e., they have the same pignistic probability distribution. Therefore, their uncertainty degrees calculated by AM are identical. We can also get $AU(m_1) = AU(m_2) = AU(m_3) = 2$. Intuitively, among these four belief structures, m_1 represents totally ignorance, so it is the most uncertain. As for m_3, the basic probability masses assigned to four singleton focal elements are equivalent. m_3 indicates that the support degree of each element is certain, while other two belief structures give interval probabilities for each element. Hence, the uncertain degree of m_3 should be the smallest. However, the AM measure cannot discriminate the differences hidden in the uncertainty of these three belief structures.

As we can see above, both AM and AU are defined in the framework of probability theory. They are calculated by transforming a BPA to certain probability distribution. Such a switch between different frameworks might cause problems in representing the uncertainty in belief functions. Since the belief functions theory is not a successful generalization of the probability theory, there exists inconsistency between the two frameworks. Information loss may be caused in such switch, which brings drawbacks to AM and AU. In our work, we will design the uncertainty measure of belief structures directly in the framework of evidence theory as introduced in the next section.

4 Uncertainty measure based on interval probabilities

4.1 Belief intervals and interval probabilities

In evidence theory, the belief interval $[Bel(A), Pl(A)]$ can be obtained based on the basic probability assigned to each focal element. The belief intervals on all singletons can be regarded as interval probability distribution on the discernment frame. To facilitate the following exposition, we would like give some background knowledge about interval probability distribution [8, 34, 35].

Definition 7 Let $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ be the frame of discernment, $[a_i, b_i]$ $(i = 1, 2, \cdots, n)$ be n intervals with $0 \leq a_i \leq b_i \leq 1$. $P(\theta_i) = [a_i, b_i]$ $(i = 1, 2, \cdots, n)$ constitute an interval probability distribution on Θ such that:

(1) $\sum\limits_{i=1}^{n} a_i \leq 1$ *and* $\sum\limits_{i=1}^{n} b_i \geq 1$;

(2) $\sum\limits_{i=1}^{n} b_i - (b_k - a_k) \geq 1$ and $\sum\limits_{i=1}^{n} a_i + (b_k - a_k) \leq 1 \forall k \in \{1, \cdots, n\}$;

(3) $p(H) = 0, \ \forall H \notin \Theta$.

Theorem 1 *For a belief structure m on the discernment frame $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, all belief intervals $[Bel(\theta_i), Pl(\theta_i)]$ $(i = 1, 2, \cdots, n)$ constitute an interval probability distribution on Θ.*

Proof Let F_j $(j = 1, 2, \cdots, N)$ be the focal elements of m, F be the set of focal elements. Non-singleton focal elements in F is expressed by $B_1, B_2, \cdots, B_q$, $q \leq N$. We can get $0 \leq Bel(\theta_i) \leq Pl(\theta_i) \leq 1$ by definitions of belief function and probability function.

(1) For a singleton $\theta_i \in \Theta$, $Bel(\theta_i)$ and $m(\theta_i)$ are identical. Then we have:

$$\sum_{i=1}^{n} Bel(\theta_i) = \sum_{i=1}^{n} m(\theta_i) \leq \sum_{j=1}^{N} m(F_j) = 1.$$

According to $Pl(\theta_i) = \sum\limits_{F_j \cap \theta_i \neq \emptyset} m(F_j) = m(\theta_i) + \sum\limits_{B_k \cap \theta_i \neq \emptyset} m(B_k)$, we can get:

$$\sum_{i=1}^{n} Pl(\theta_i) = \sum_{i=1}^{n} \left(m(\theta_i) + \sum_{B_k \cap \theta_i \neq \emptyset} m(B_k) \right)$$
$$\geq \sum_{i=1}^{n} m(\theta_i) + \sum_{k=1}^{q} m(B_k) = \sum_{j=1}^{N} m(F_j) = 1.$$

(2) Given an arbitrary non-singleton focal elements B_j satisfying $B_j \cap \theta_k \neq \emptyset$, $\forall \theta_k \in \Omega$, we have $\exists \theta_i \in \Omega$, $\theta_i \neq \theta_k$, $B_j \cap \theta_i \neq \emptyset$. So the following inequality holds:

$$\sum_{i \neq k} \sum_{B_j \cap \theta_i \neq \emptyset} m(B_j) \geq \sum_{j=1}^{q} m(B_j).$$

For $\forall k \in \{1, 2, \cdots, n\}$, we have

$$\sum_{i \neq k} Pl(\theta_i) = \sum_{i \neq k} \left(\sum_{F_j \cap \theta_i \neq \emptyset} m(F_j) \right)$$

$$= \sum_{i \neq k} \left(m(\theta_i) + \sum_{B_j \cap \theta_i \neq \emptyset} m(B_j) \right)$$

$$= \sum_{i \neq k} m(\theta_i) + \sum_{i \neq k} \left(\sum_{B_j \cap \theta_i \neq \emptyset} m(B_j) \right)$$

$$\geq \sum_{i \neq k} m(\theta_i) + \sum_{j=1}^{q} m(B_j)$$

$$= \sum m(F_j) - m(\theta_k)$$

$$= 1 - Bel(\theta_k)$$

That is to say $\sum_{i=1}^{n} Pl(\theta_i) - Pl(\theta_k) \geq 1 - Bel(\theta_k)$ holds for any $k \in \{1, 2, \cdots, n\}$. So we get $\sum_{i=1}^{n} Pl(\theta_i) - (Pl(\theta_k) - Bel(\theta_k)) \geq 1, \forall k \in \{1, 2, \cdots, n\}$.

As such, for $\forall k \in \{1, 2, \cdots, n\}$, we still have

$$\sum_{i \neq k} Bel(\theta_i) = \sum_{i \neq k} m(\theta_i) = \sum m(F_j) - \sum m(B_j) - m(\theta_k)$$

$$\leq \sum m(F_j) - \sum_{B_j \cap \theta_k \neq \emptyset} m(B_j) - m(\theta_k)$$

$$= 1 - \sum_{F_j \cap \theta_k \neq \emptyset} m(F_j) = 1 - Pl(\theta_k)$$

which means $\sum_{i=1}^{n} Bel(\theta_i) - Bel(\theta_k) \leq 1 - Pl(\theta_k)$ holds for any $k \in \{1, 2, \cdots, n\}$. Thus, $\sum_{i=1}^{n} Bel(\theta_i) + (Pl(\theta_k) - Bel(\theta_k)) \leq 1, \forall k \in \{1, 2, \cdots, n\}$.

(3) $p(H) = 0, \forall H \notin \Theta$ is straightforward.

So all belief intervals $[Bel(\theta_i), Pl(\theta_i)]$ $(i = 1, 2, \cdots, n)$ constitute an interval probability distribution over Θ. $\square$

On the basis of interval probabilities, we can measure the uncertainty of belief structures. In an interval probability distribution $[Bel(\theta_i), Pl(\theta_i)]$ over $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, its uncertainty degree can be represented by the interval as analyzed below.

Case I: The most uncertain case is $Bel(\theta_i) = 0$ and $Pl(\theta_i) = 1$ for each $\theta_i \in \Theta$, i.e., the probability intervals of all elements are all equal to [0,1]. The equal probability intervals indicate the greatest discord degree, and the interval [0,1] represents the largest non-specificity.

Case II: The clearest case is $Bel(\theta_k) = Pl(\theta_k) = 1$ and $Bel(\theta_i) = Pl(\theta_i) = 0$ for $i \neq k$. That is to say,

θ_k is assured to occur, while others are assured to never occur.

Case III: When $Bel(\theta_i) = a_i$, $Pl(\theta_i) = b_i$, $\forall a_i, b_i \in [0, 1]$, the probability interval of θ_i is $[a_i, b_i]$. For θ_i, the degree of non-specificity can be represented by its imprecision $b_i - a_i$. The degree of discord is quantified by the relation between all probability intervals.

Case IV: When $Bel(\theta_i) = Pl(\theta_i) = a_i$, $\forall a_i \in [0, 1]$, $\sum_{i=1}^{n} a_i = 1$, the probability interval reduces to a precise probability distribution. It means that there is no imprecision and whether θ_i occurs or not cannot be clearly determined.

Therefore, the information related to the uncertainty carried by a probability interval includes both the discord part and the non-specificity part, which coincides with the construction of the BPA's uncertainty. Thus, given a BPA, we can fully utilize the information of the probability intervals to measure its total uncertainty. Then, how to use the information of belief intervals? Given the interval probability $[Bel(\theta_i), Pl(\theta_i)]$ $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, we can use the relationship between all central values $(Pl(\theta_i) + Bel(\theta_i))/2$ to assess the discord degree, and use the imprecision degree $Pl(\theta_i) - Bel(\theta_i)$ to measure the non-specificity degree. Then, a new uncertainty measure for belief structures can be constructed based on probability intervals.

4.2 Interval probability-based uncertainty measure

As is mentioned previously, this belief interval is equivalent to an interval probability distribution whose lower and upper bounds are the belief function Bel and the plausibility function Pl, respectively. For an interval probability distribution, its uncertainty contains both discord and non-specificity. The discord part represents the difference between the probability masses assigned to all propositions. In other words, discord is determined by the distribution of all probability masses. So we can use the Shannon entropy to quantify the discord of interval probability distribution. Moreover, interval values can be compared by comparing their central values. Therefore, we can measure the discord of interval probability distribution on the basis of Shannon entropy and central probabilities. We have also pointed that the non-specificity of interval probability can be quantified by its imprecision degree, which is related to the span of interval probability. So in the definition of uncertainty measure for belief structure, both the Shannon entropy and the span of the interval probability should be taken into account. Hence, the interval probabilities-based uncertainty measure for a belief structure can be defined as follow.

Definition 8 Let m be a belief structure on the discernment frame $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$. The interval probabilities-based uncertainty measure of m can be expressed as below:

$$SU(m) = \sum_{i=1}^{n} \left[-\frac{Bel(\theta_i) + Pl(\theta_i)}{2} \log_2 \frac{Bel(\theta_i) + Pl(\theta_i)}{2} \right.$$
$$\left. + \frac{Pl(\theta_i) - Bel(\theta_i)}{2} \right] \tag{18}$$

The newly proposed *SU* has no drawbacks as those of *AU* and *AM* pointed out in Examples 1 and 2. Detailed comparative analysis will be presented in Section 5.

Here, we provide an illustrative example to show the calculation of *SU*.

Example 3 Suppose that the discernment frame is $\Theta = \{\theta_1, \theta_2, \theta_3\}$. A BPA m over Θ is $m(\{\theta_1\}) = 0.2$, $m(\{\theta_1\}) = 0.1$, $m(\{\theta_2, \theta_3\}) = 0.4$, $m(\Theta) = 0.3$.

First calculate the belief function and plausibility function for singletons.

$Bel(\{\theta_1\}) = m(\{\theta_1\}) = 0.2,$

$Pl(\{\theta_1\}) = m(\{\theta_1\}) + m(\Theta) = 0.5,$

$Bel(\{\theta_2\}) = m(\{\theta_2\}) = 0.1,$

$Pl(\{\theta_2\}) = m(\{\theta_2\}) + m(\{\theta_2, \theta_3\}) + m(\Theta) = 0.8,$

$Bel(\{\theta_3\}) = m(\{\theta_3\}) = 0,$

$Pl(\{\theta_3\}) = m(\{\theta_2, \theta_3\}) + m(\Theta) = 0.7.$

The probability intervals are:

$BI(\{\theta_1\}) = [0.2, 0.5],$

$BI(\{\theta_2\}) = [0.1, 0.8],$

$BI(\{\theta_3\}) = [0, 0.7].$

The total uncertainty of m can be calculated as:

$$SU(m)$$
$$= \sum_{i=1}^{3} \left[-\frac{Bel(\theta_i) + Pl(\theta_i)}{2} \log_2 \frac{Bel(\theta_i) + Pl(\theta_i)}{2} \right.$$
$$\left. + \frac{Pl(\theta_i) - Bel(\theta_i)}{2} \right]$$
$$= -0.35 \log_2 0.35 + 0.15 - 0.45 \log_2 0.45 + 0.35$$
$$- 0.35 \log_2 0.35 + 0.35$$
$$= 2.4286$$

We find that the uncertainty degree of m is 2.4286, which is greater than $\log_2 3 = 1.585$. It seems that the proposed uncertainty measure violated the requirement about the range of uncertainty degree [20]. In the flowing subsection, we will discuss the range of the proposed *SU*.

4.3 Some properties of the proposed SU

Theorem 2 *The uncertainty measure SU is probability consistency, i.e., when m reduces to a Bayesian belief structure, SU is identical to Shannon entropy.*

Proof If m reduces to a Bayesian belief structure on $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, $m(\{\theta_i\}) = p_i$ for $i = 1, 2, \cdots, n$ we can get: $Bel(\{\theta_i\}) = Pl(\{\theta_i\}) = p_i$. So the interval probability distribution reduces to a precise probability distribution over $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$.

Then it follows that

$$SU(m)$$
$$= \sum_{i=1}^{n} \left[-\frac{Bel(\theta_i) + Pl(\theta_i)}{2} \log_2 \frac{Bel(\theta_i) + Pl(\theta_i)}{2} \right.$$
$$\left. + \frac{Pl(\theta_i) - Bel(\theta_i)}{2} \right]$$
$$= -\sum_{i=1}^{n} \frac{Bel(\theta_i) + Pl(\theta_i)}{2} \log_2 \frac{Bel(\theta_i) + Pl(\theta_i)}{2}$$
$$= -\sum_{i=1}^{n} p_i \log_2 p_i$$

Hence, *SU* is identical to the Shannon entropy. □

Theorem 3 *The range of measure SU is* $[0, |\Theta|]$, *where* $|\Theta|$ *is the cardinality of* Θ.

Proof For $0 \leq x \leq y \leq 1$, we can define a function $f(x, y)$ as:

$$f(x, y) = -\frac{x + y}{2} \log_2 \left(\frac{x + y}{2} \right) + \frac{y - x}{2}$$

In order to obtain the maximum of $f(x, y)$, we should calculate the derivative:

$$\frac{\partial f}{\partial x} = -\frac{1}{2} \log_2 \left(\frac{x + y}{2} \right) - \frac{1}{2 \ln 2} - \frac{1}{2} = -\frac{1}{2} \log_2 [e(x + y)],$$

$$\frac{\partial f}{\partial y} = -\frac{1}{2} \log_2 \left(\frac{x + y}{2} \right) - \frac{1}{2 \ln 2} + \frac{1}{2} = -\frac{1}{2} \left[\log_2 \frac{e(x + y)}{4} \right].$$

Given $0 \leq x \leq y \leq 1$, $0 \leq x + y \leq 2 < e$, we have:

1). $\frac{\partial f}{\partial x} > 0$ for $x + y < \frac{1}{e}$, $\frac{\partial f}{\partial x} < 0$ for $x + y > \frac{1}{e}$, and $\frac{\partial f}{\partial x} = 0$ for $x + y = \frac{1}{e}$;

2). $\frac{\partial f}{\partial y} > 0$ for $x + y < \frac{4}{e}$, $\frac{\partial f}{\partial y} < 0$ for $x + y > \frac{4}{e}$, and $\frac{\partial f}{\partial y} = 0$ for $x + y = \frac{4}{e}$.

Therefore, $f(x, y)$ is not a monotone function on x or y. The maximum of f can be obtained at the stationary point or the bounds of (x, y). Moreover, $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$ cannot hold simultaneously, i.e. the stationary point doesn't exist. Hence, we can get the maximum of f at three possible positions: (0, 1), (0, 0), and (1,1).

Since $f(1, 1) = 0$, $f(0, 0) = 0$, $f(0, 1) = 1$, we can get $f_{\max}(x, y) = f(0, 1) = 1$.

The graph of the function shown in Fig. 1 also illustrates $f_{\max}(x, y) = f(0, 1) = 1$, which is identical to the theoretical analysis.

Therefore, for a belief structure m on $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$, $SU(m)$ can get its maximum value $SU_{\max}(m) = n = |\Theta|$ if and only if $Bel(\theta_i) = 0$, $Pl(\theta_i) = 1$, for $i = 1, 2, \cdots, n$. This case can be obtained iff m is a vacuous belief structure.

Now, we are looking for the lower bound.

Since $0 \leq Bet(\theta_i) \leq Pl(\theta_i) \leq 1$, we can have

$$-\frac{Bet(\theta_i) + Pl(\theta_i)}{2} \log_2\left(\frac{Bet(\theta_i) + Pl(\theta_i)}{2}\right) = 0,$$

$$\frac{Pl(\theta_i) - Bet(\theta_i)}{2} \geq 0.$$

So $SU(m)$ get its minimum by equaling the two functions to 0 for all $A_i \in \Omega$, which indicates that the focal elements of m are all singleton.

Then we have:

$$m(\theta_i) = Bet(\theta_i) = Pl(\theta_i) = 0 \quad \text{or}$$
$$m(\theta_i) = Bet(\theta_i) = Pl(\theta_i) = 1.$$

Considering the condition that $\sum_{i=1}^{n} m(\theta_i) = 1$, we can get:

$m(\{\theta_k\}) = 1$ and $m(\{\theta_i\}) = 0$, for $i \neq k$, i.e., m is a categorical belief structure.

Therefore, $SU(m)$ can get its minimum when m is a categorical belief structure.

Finally, we can get the range of SU is $[0, |\Omega|]$. $\qquad\square$

This theorem indicates that a vacuous BPA contains the most uncertainty, while a categorical BPA with singleton focal element has the least uncertainty, which coincides with the intuitive analysis. It is worth noting that the range of SU is identical to the axiomatic range of uncertainty proposed by Deng [4], but different with Klir and Wierman's requirement that the range of uncertainty measure is $[0, log_2|\Theta|]$ [20]. Klir has also mentioned that the range of uncertainty for evidence is $[0, N]$, where 0 must be assigned to the unique uncertainty function that describes full
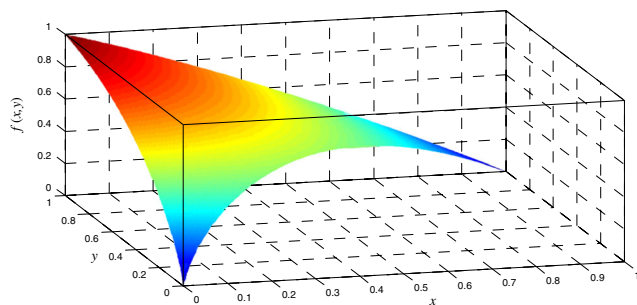

Fig. 1 Graph of $f(x, y)$

certainty and N depends on the cardinality of the universal set involved and on the chosen unit of measurement [16]. It is apparent that the range of our similarity measure satisfies this requirement.

Theorem 4 *If a BPA m is a categorical BPA focusing on A, then the uncertainty of m is $|A|$.*

Proof Suppose that m is a categorical BPA on $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ and $m(A) = 1$. Without any loss of generality, we can let $A = \{\theta_1, \theta_2, \cdots, \theta_k\}$, $k \leq n$.

Then we can get the belief function and plausibility function for each singleton:

$$Bel(\{\theta_1\}) = Bel(\{\theta_2\}) = \cdots Bel(\{\theta_k\} = 0,$$
$$Pl(\{\theta_1\}) = Pl(\{\theta_2\}) = \cdots Pl(\{\theta_k\} = 1,$$
$$Bel(\{\theta_{k+1}\}) = Bel(\{\theta_{k+2}\}) = \cdots Bel(\{\theta_n\} = 0,$$
$$Pl(\{\theta_{k+1}\}) = Pl(\{\theta_{k+2}\}) = \cdots Pl(\{\theta_n\} = 0.$$

So we can calculate the uncertainty of m as:

$$SU(m)$$
$$= \sum_{i=1}^{n}\left[-\frac{Bel(\theta_i) + Pl(\theta_i)}{2} \log_2 \frac{Bel(\theta_i) + Pl(\theta_i)}{2}\right.$$
$$\left. + \frac{Pl(\theta_i) - Bel(\theta_i)}{2}\right]$$
$$= \sum_{i=1}^{k}\left(-\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\right) = k = |A|$$

Hence, the uncertainty of m is identical to the cardinality of A. $\qquad\square$

We notice that $m(A) = 1$ says that A is assure to occur. In such case, the belied structure m coincides with a classical set A. Since the set consistency in [20] is proposed on the basis of Hartley measure, this theorem indicates that the measure SU is not set consistent. However, we have in mind that the uncertainty degree of a classical set is related to its cardinality. So we can give the generalized form of uncertainty measure for classical set A as:

$$U(A) = f(|A|) \tag{19}$$

where f is an increasing function with respect to $|A|$.

In such sense, we can say our proposed uncertainty measure SU is set consistent. Considering the generalized form of uncertainty measure for classical set, we call this set consistency as generalized set consistency.

5 Comparative examples

Firstly, let us revisit Example 1. By Definition 8, we can get that $SU(m_1) = 2.5546$, $SU(m_2) = 2.6929$. This indicated

that m_2 has a higher level of uncertainty, which is consistent to previous intuitive analysis.

Then we reconsider Example 2, where the three belief structures have the same pignistic probability. By Eq. (18), we have $SU(m_1) = 4$, $SU(m_2) = 3$, $SU(m_3) = 2$. We can note that our proposed uncertainty measure SU is competent to discriminate the different uncertainty level for different belief structures, although they have the same pignistic probability distribution. This result coincides with our previous intuitive analysis.

In the sequel, examples presented in [39] are used for reference.

Example 4 Let m be a belief structure in $\Theta = \{\theta_1, \theta_2\}$. A BPA is given as: $m(\{\theta_1\}) = a$, $m(\{\theta_2\}) = b$, $m(\Theta) = 1 - a - b$, $a, b \in [0, 0.5]$. Here, we calculate AU, AM, SU values
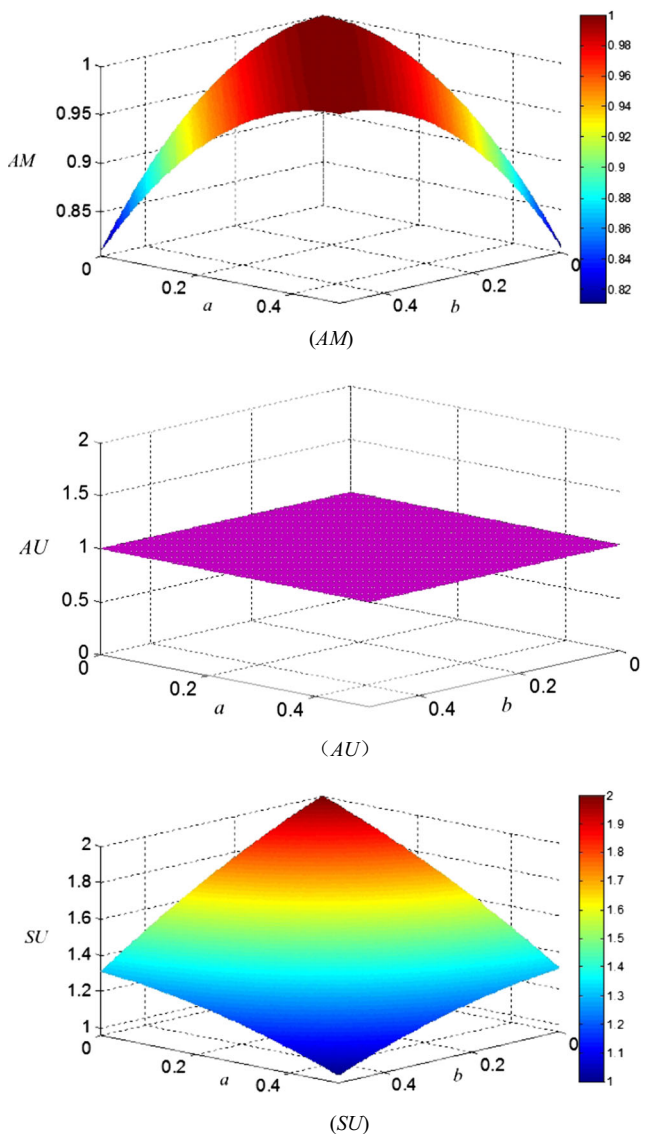
corresponding to different values of a and b. The change of AU, AM and SU values with the change of a and b are illustrated in Fig. 2.

As shown in Fig. 2, AM reaches its maximum value when $a = b$, regardless of the value of a and b. Because when $a = b$, the pignistic probability is uniformly distributed, i.e., $BetP(\{\theta_1\}) = BetP(\{\theta_2\}) = 1/2$. This is counter-intuitive, as analyzed in Example 2. The effect of total ignorance to the uncertainty degree cannot be reflected by AM. For example, $m_1(\{\theta_1\}) = m_1(\{\theta_2\}) = 0.5$ and $m_2(\{\theta_1\}) = m_2(\{\theta_2\}) = 0.25$, $m_2(\Theta) = 0.5$, there exists $AM(m_1) = AM(m_2)$. Obviously, it is irrational to say that m_1 and m_2 have the same degree of total uncertainty.

AU never changes with the change of a and b. The value of AU is always $\log_2 2 = 1$. Since $Bel(\{\theta_1\}) = a$, $Pl(\{\theta_1\}) = a + 1 - a - b = 1 - b$, $Bel(\{\theta_2\}) = b$, $Pl(\{\theta_2\}) = b + 1 + a - b = 1 - a$, the constraints for calculating AU in this example are:

$$a \le p(\{\theta_1\}) \le 1 - b, \quad b \le p(\{\theta_2\}) \le 1 - a.$$

In the calculation of AU, we try to find a probability distribution with the maximum Shannon entropy. Since $a, b \in [0, 0.5]$, the uniformly distributed $p(\{\theta_1\}) = p(\{\theta_2\}) = 0.5$ always satisfies the constraints above. So no matter how a and b change, $p(\{\theta_1\}) = p(\{\theta_2\}) = 0.5$ is always taken when calculating AU. Hence, AU always equals to $\log_2 2 = 1$. However, intuitively, the degree of uncertainty should change with the change of a and b. So, AU cannot well describe the degree of total uncertainty here.

In Fig. 2, we can see that SU brings out rational results. It reaches the maximum when a=b=0, i.e., $m(\Theta) = 1$. When $a + b$ is fixed, the maximum values are reached if $a = b$ holds. For example, suppose that $a + b = 0.3$. The change of SU value with respect to a is shown in Fig. 3. We can see that SU get the maximum value when $a = 0.15$, i.e., $a = b = 0.15$. Since the fixation of $a + b$ indicates that $m(\Theta) = 1 - (a + b)$ is fixed, the non-specificity part is fixed. The case of $a = b$ makes the discord part reach its maximum value. Therefore, the maximum SU can be reached when $a = b$.

Similarly, in the case of $a = b$, the value of SU with the change of a is shown in Fig. 4. We can see that when $a = b$,
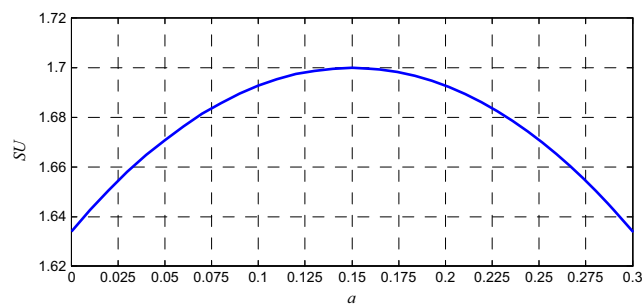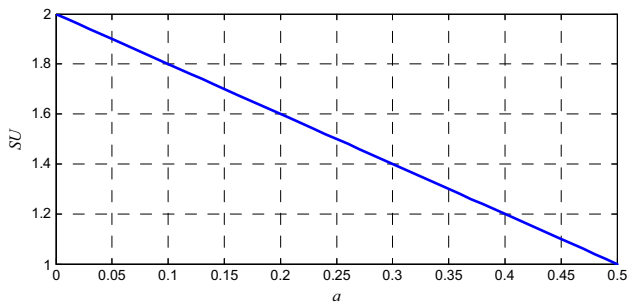

(AM)

(AU)

(SU)

Fig. 2 The change of total uncertainty measures in Example 4


Fig. 3 The change of SU when $a + b = 0.3$

Fig. 4 The change of *SU* when $a = b$

the value of *SU* is decreasing with the increase of a. It is shown that the uncertainty degree gets the maximum value when $a = b = 0$. Since a = b indicates that the discord part uncertainty is equal to 1, the value of SU is determined by the non-specificity part uncertainty, which is related to the basic probability assigned to Θ. So *SU* gets the maximum value in the case of $a = b = 0$. This is consistent to intuitive.

Example 5 Suppose that the discernment frame is $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and a BPA over Θ is $m(\Theta) = 1$. We make changes to the BPA step by step. In each step, $m(\Theta)$ has a decrease of $\Delta = 0.05$ and each singleton mass $m(\{\theta_i\})$, $i = 1,2,3$ has an increase of $\Delta/3$. Finally, $m(\Theta)$ reaches 0 and $m(\{\theta_i\}) = 1/3$. We calculate total uncertainty measures including *SU*, *AM*, *AU*, at each step. The changes of uncertainty values are shown in Fig. 5.

We can see that, the total uncertainty measures *AU* and *AM* stay unchanged and they are always at the value $\log_2 3$. Intuitively, when the mass assignments are transferred from the full set Θ to those singletons, the uncertainty level should decrease. Hence, the results of *AU* and *AM* are counter-intuitive. The counter-intuitive results of *AU* and *AM* are caused by their definitions based on some probabilistic transformation from BPAs. In this example, for the probabilistic transformation used in the calculation of *AM* and *AU*, a uniformly distributed probability: $P(\theta_i) = 1/3$,
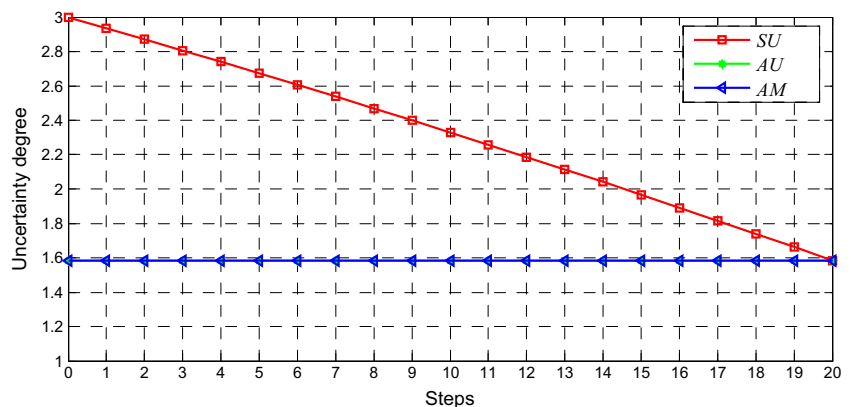
$i = 1, 2, 3$, is applied. Therefore, *AM* and *AU* never change, when the BPA changes from a vacuous BPA to a Bayesian BPA in the final. As we can see in Fig. 4, our new measure *SU* can provide rational results. It becomes smaller and smaller when the BPA changes from a vacuous one to a Bayesian one. Therefore, it is inappropriate to define the total uncertainty measure for belief structures by switching the evidence theory framework to probability framework. It should be better not to directly design the total uncertainty measure in the framework of evidence theory. This is also the motivation of our work in this paper.

Example 6 Suppose that the discernment frame is $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and a BPA over Θ is $m(\Theta) = 1$. We make changes to the BPA step by step. In each step, $m(\Theta)$ has a decrease of $\Delta = 0.05$ and the singleton mass $m(\{\theta_1\})$ has an increase of Δ. In the final step, $m(\Theta) = 0$ and $m(\{\theta_1\}) = 1$. We calculate total uncertainty measures including *SU*, *AM*, *AU*, at each step. The change of uncertainty values are shown in Fig. 6.

In this example, the BPA changes from a vacuous one to a categorical one. As shown in Fig. 5, the total uncertainty values calculated based on *AU*, *AM* and *SU* all become smaller and smaller, which is intuitive. However, we note that *AU* is not sensitive to the BPA's change at the beginning stage. This can be analyzed as following. Since constraint for calculating *AU* is $Bel(\{\theta_1\}) \leq p(\{\theta_1\}) \leq 1$, the uniformly distributed $p(\{\theta_1\}) = p(\{\theta_2\}) = p(\{\theta_3\}) = 1/3$ always satisfies the constraints above when if $Bel(\{\theta_1\}) = m(\{\theta_1\}) \leq 1/3$. So the uncertainty degree *AU* stays unchanged in the first 6 steps. This has already been criticized by Jousselme et al. [15]. Although both *AM* and *SU* change with the change of BPA in each step, we can see that the *SU* measure is more sensitive to the change of BPA.

Example 7 Suppose that the discernment frame is $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8\}$. A BPA over Θ is $m(\Theta) = 1$. We make changes to the BPA step by step. In each step, $m(\Theta)$ has a decrease of $\Delta = 0.05$, and the mass assignment for

Fig. 5 The change of uncertainty degree in Example 5

Fig. 6 The change of uncertainty degree in Example 6
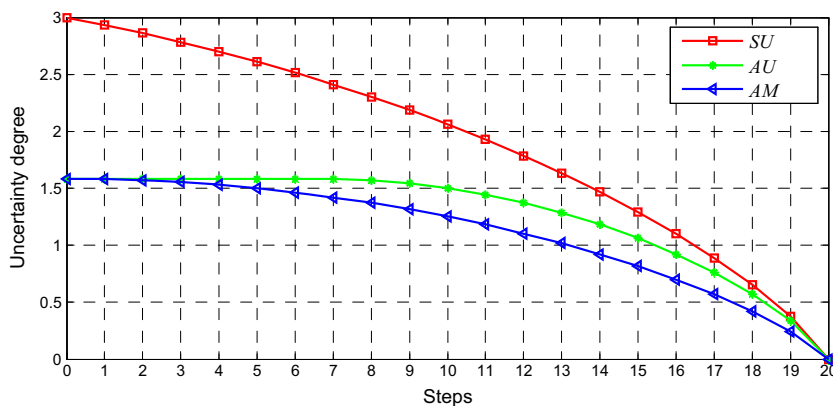
one focal element B with cardinality $s < 8$ has an increase of $\Delta = 0.05$. In the final step, $m(B) = 1$. Here, we can set $s = 2, 4, 6$, respectively. Given an s value, we repeat the whole procedure of BPA change. Under the different s, we investigate the changes of values for uncertainty measures including SU, AM and AU.

For the convenience of comparison, we normalize the uncertainty values in the interval [0,1]. The normalized uncertainty measures corresponding to AU, AM and SU are defined as following, respectively.

$$AU' = AU / \log_2 n \tag{20}$$

$$AM' = AM / \log_2 n \tag{21}$$

$$SU' = SU / n \tag{22}$$

where n is the cardinality of discernment frame.

For different cardinality of B, the changes of values for normalized uncertainty measures including SU', AU', and AM' are shown in Fig. 7.

As shown in Fig. 7, the total uncertainty SU, AU, and AM all decrease when more mass assignments are transferred to B. Intuitively, when the mass assignments are transferred to a focal element with smaller cardinality, the non-specificity decreases faster. So a smaller s will bring a greater decrease on the total uncertainty. Comparatively, AM is not sensitive to the change of the BPA, since it is calculated based on the pignistic probability, which changes slightly with the change of BPAs. In Fig. 7, we can also see that in the earlier steps, AU stays unchanged for different s. Moreover, with the increase of s, the unchanged stage of AU is extended. The reason is analyzed as follows. At the beginning, the BPA is a vacuous one, the uncertainty degree AU can obtain its maximum value in the case of $p(\{\theta_i\}) = 1/8$, $i = 1, 2, \cdots, 8$. When more probability masses are transferred to B, from the constrains for the calculation of AU, we can get that the uniformly distributed probability can be used to calculate AU in the case of $m(B) \leq s/8$. Then it follows that $AU = \log_2 8 = 3$ always holds if $N \leq 2.5s$, N is

the step number. Therefore, AU stays unchanged in the first $2.5s$ steps.

Example 8 Suppose that the discernment frame is $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}\}$. A BPA over Θ is $m(\Theta) = 1\text{-}a$ and $m(A) = a$, $|A| = s \leq 10$. Given an a, and the initial $|A|=1$, we make changes to the BPA step by step. $|A|$ increases by 1 in each step. In the final, the BPA becomes a vacuous one. We set $a = 0.3, 0.5, 0.8$, respectively. Given an a value, we repeat the whole procedure of BPA change. Under the different a, the changes of normalized uncertainty measures including SU', AU' and AM' are shown in Fig. 7.

As shown in Fig. 8, total uncertainty measures including AU, AM, and our SU all increase with the change of BPA at each step. This is intuitively reasonable, since the increase of cardinality of A will make the non-specificity increase. It can be noted that if a is given a larger value, they will increase faster. Because with the increase of $|A|$, relatively more mass assignments are transferred to a larger size focal element. Figure 8 shows that AU stay unchanged in when $|A|$ is larger than a certain value. By the definition of AU, we can get that when $|A| \geq 10a$, AU get its maximum value $\log_2 10$. Compared with our proposed SU, AM is not sensitive to the BPA's change.

From above illustrative examples, we can see that our proposed uncertainty measure SU is more sensitive to the change of BPAs. Due to the complicity of optimization algorithm, the calculation burden of AU is heavier those of SU and AM. Moreover, since our proposed uncertainty measure SU prevents the probabilistic switch of BPA, it is more convenient for direct application.

6 Application of uncertainty measure *SU*

In this section, we apply our new total uncertainty measure in the evaluation of sensors' importance for identification fusion to further show its rationality.

Fig. 7 The change of total
uncertainty degree in Example 7


$(s=2)$

$(s=4)$

$(s=6)$
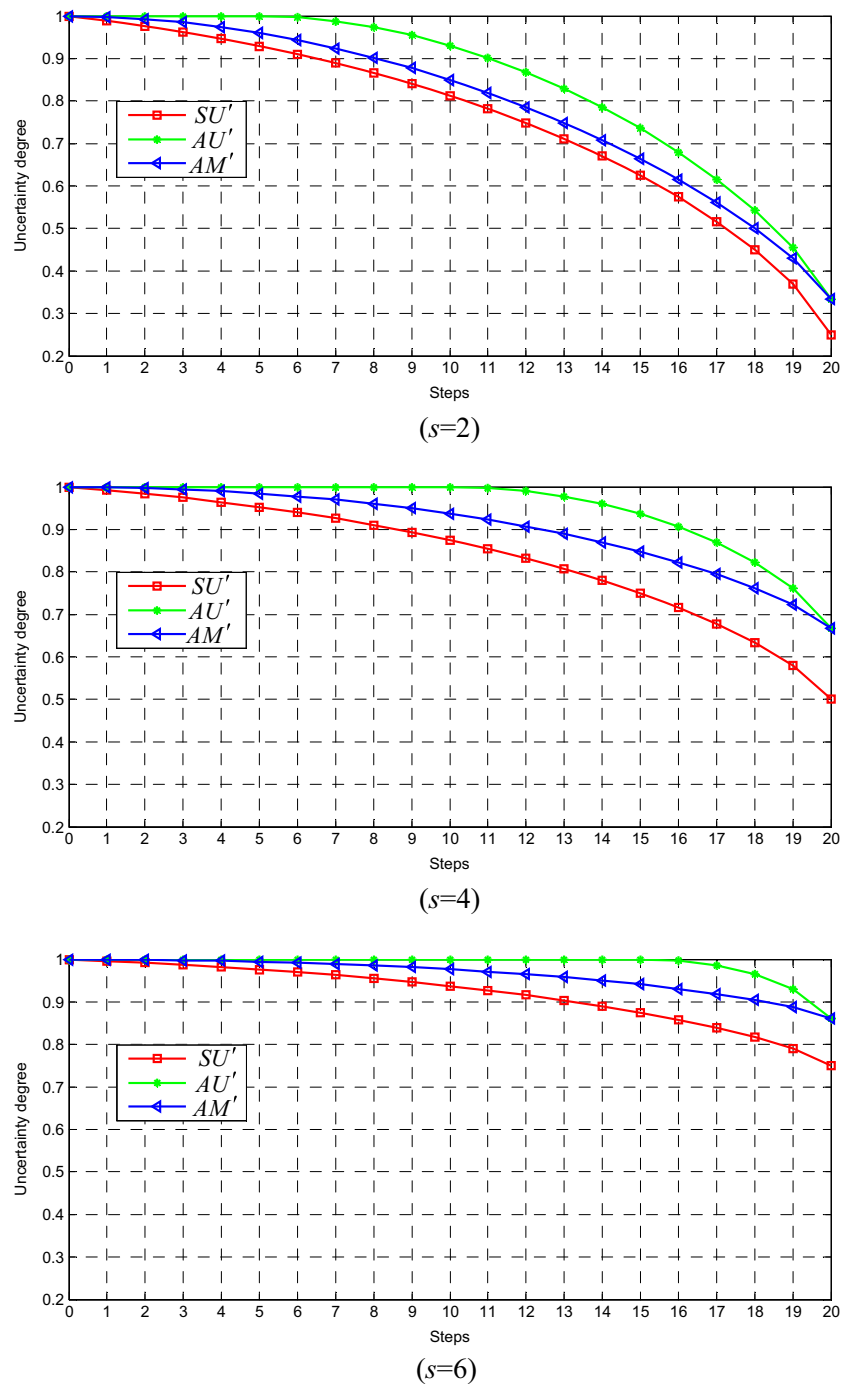
A successful application of evidence theory is identification fusion, which has received considerable attention in both military and civilian areas [24, 26, 27]. In target identification, due to the limitation of sensors and interference from environments, the information derived from different sensors is usually imperfect. Hence, the target cannot be identified by single sensor. It is necessary to fuse information coming from multiple sensors to achieve better results. Due to its flexibility in combination and ease of use by the decision maker, evidence theory has been widely applied in multi-sensor data fusion.

When fusing the information coming from different sensors, it is necessary to access the importance of each sensor. Generally, the importance of sensor is quantified by a weighting factor, which is evaluated based on the similarity (distance) between BPAs provided by different sensors [7, 25]. If the BPA provided by a sensor is similar to others, then it is supported by other sensors in a great degree, and

Fig. 8 The change of uncertainty degree in Example 8

(a=0.3)

(a=0.5)

(a=0.8)
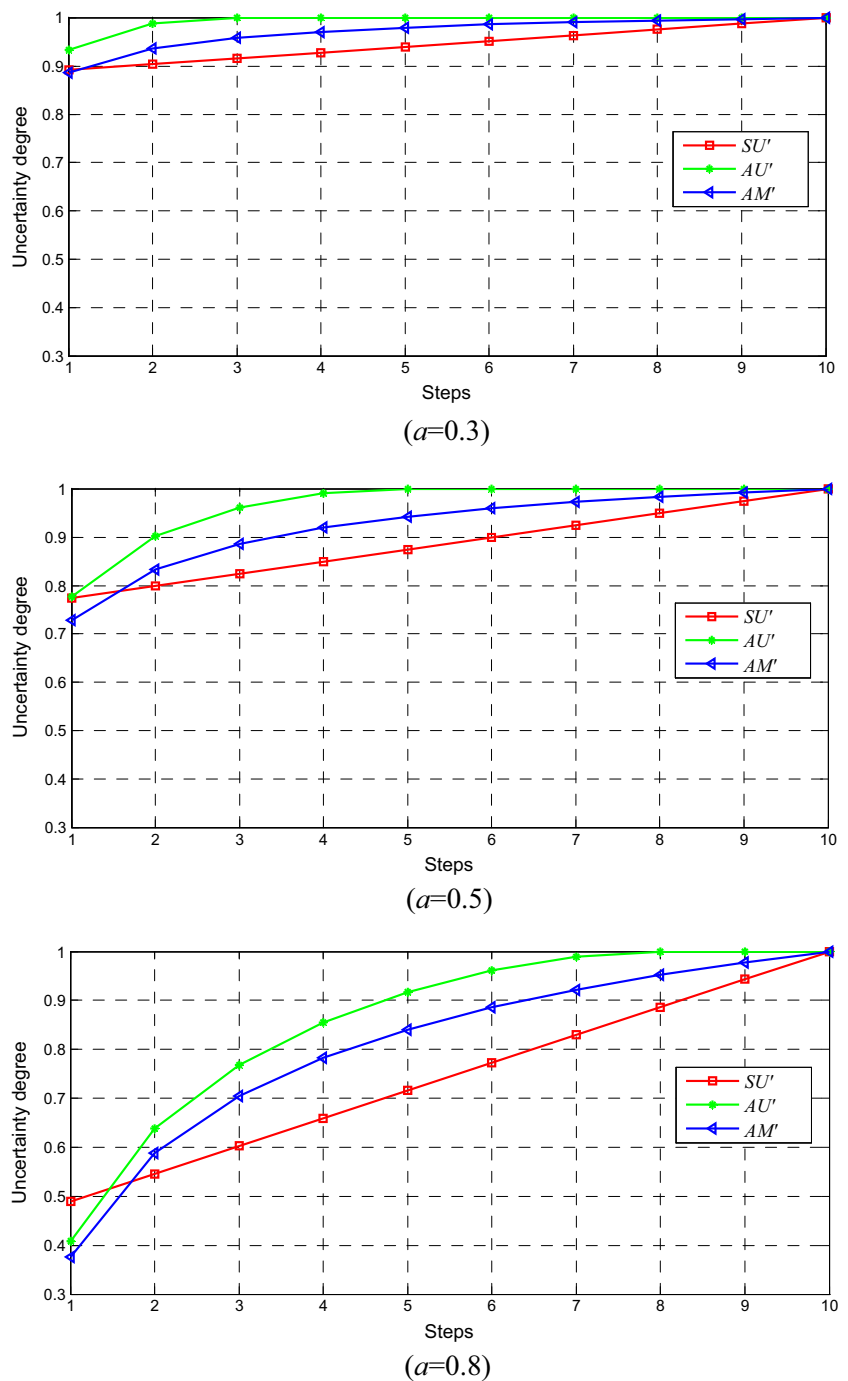
should be assigned a higher weighting factor. In our opinion, the importance of a sensor is also determined by the quality of the information it provided. We can suppose that if the information has a lower uncertainty degree, the corresponding sensor is more important. So the weighting factor of a sensor is simultaneously determined by the similarity degree between BPAs and the uncertainty degree of the BPA corresponding to it. Here, we divide a sensor's importance into two parts, credibility and discriminability, which are determined by similarity (distance) measure and uncertainty

measure, respectively. Intuitively, the sensor with smaller total uncertainty measure should be better (have higher discriminability). Hence, based on the uncertainty measure, we can get the discriminability part importance degree of each sensor. Considering the range of each uncertainty measure, we can use the normalized uncertainty measures to calculate the discriminability.

In a target recognition system based on multi-sensor, p sensors $S_1, S_2, \cdots, S_p$ are employed to identify the target. The BPAs obtained from these sensors are $m_1, m_2, \cdots, m_p$,

which are defined over the discernment frame $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$. Then the relative discriminability of each sensor can be calculated by:

$$D_j = \frac{1 - U'(m_j)}{p - \sum\limits_{k=1}^{p} U'(m_k)}, \quad j = 1, 2, \cdots, p \qquad (23)$$

where U' is the normalized uncertainty measure for belief structure.

Suppose that five sensors $S_1, S_2, \cdots, S_5$ are employed to classify the identification of aerial targets. Three possible types of targets are *Airplane*, *Helicopter*, and *Rocket*. denoted as A, H and R, respectively. So the discernment frame Ω can be written as $\{A, H, R\}$. The sensor readings on the classes are expressed by the BBAs detailed as following:

$m_1(\{A\}) = 0.5, m_1(\{A, H\}) = 0.3, m_1(\{H, R\}) = 0.2;$

$m_2(\{A\}) = 0.8, m_2(\{A, R\}) = 0.1, m_2(\{H, R\}) = 0.1;$
$m_3(\{A\}) = 0.6, m_3(\{H\}) = 0.2, m_3(\{A, R\}) = 0.2;$
$m_4(\{A\}) = 0.6, m_4(\{A, H\}) = 0.2, m_4(\{H, R\}) = 0.2;$
$m_5(\{A\}) = 0.8, m_5(\{H\}) = 0.1, m_5(\{H, R\}) = 0.1.$

The normalized uncertainty values of each BPA calculated based on SU', AU' and AM' are shown in Table 1. We can see that the uncertainty values calculated by our proposed SU are different to each other. So the discriminability of each sensor can be ranked as $D_2 > D_5 > D_4 > D_3 > D_1$. For uncertainty measure AU, it is shown that $AU(m_2) = AU(m_5)$ and $AU(m_3) = AU(m_4)$. Hence, the discriminability degrees of S_2 and S_5 are evaluated as equal, so it is with S_3 and S_4. Table 1 also shows that the uncertainty measure AM is not sensitive to the difference between of m_3 and m_4. Based on AM, the discriminability of sensor S_3 and S_4 cannot be well ranked.

For clarity, Table 2 presents the relative discriminability degree derived from normalized uncertainty values. We find that uncertainty measure SU can provide the discriminability order of all sensors. The largest discriminability degree is assigned to the sensor S_2 by SU and AM, and the smallest one is assigned to S_1 by SU and AM. However, S_3 and S_4 get the same discriminability degree based on AM. Table 2 shows that AU assigns the largest discriminability to S_2 and S_5, and the second largest one to S_3 and S_4. So it is difficult to determine the discriminability of each sensor based on AU.

Table 1 Normalized uncertainty values for all BPAs

Measures	m_1	m_2	m_3	m_4	m_5
SU '	0.5787	0.3159	0.4523	0.4415	0.3281
AU '	0.9464	0.5817	0.8650	0.8650	0.5817
AM '	0.7799	0.4717	0.7298	0.7298	0.5579

Table 2 Relative discriminability of each sensor

Measures	D_1	D_2	D_3	D_4	D_5
SU '	0.1461	0.2373	0.1900	0.1937	0.2330
AU '	0.0462	0.3605	0.1164	0.1164	0.3605
AM '	0.1271	0.3052	0.1561	0.1561	0.2554

In summary, our proposed uncertainty measure SU can be well applied in discriminability evaluation for identification fusion based on multi-sensor, and it outperforms AM and AU.

The relative discriminability can be used to evaluate the importance of each sensor in information fusion. Considering the method proposed in [7], we can fuse the information provided by five sensors by the modified Murphy's combination rule [29]. The weighted average of all BPAs can be given as:

$$\bar{m}(X) = \sum_{i=1}^{5} D_i \cdot m_i(X), \ \ X \subseteq \Theta \qquad (24)$$

Based on the proposed uncertainty measure SU, we can get the weighting factor of each sensor as:

$w_1 = 0.1461, w_2 = 0.2373,$
$w_3 = 0.1900, w_4 = 0.1937, w_5 = 0.2330.$

So the weighted average BPA can be obtained as:

$\bar{m}(\{A\}) = 0.68, \bar{m}(\{H\}) = 0.06,$
$\bar{m}(\{A, H\}) = 0.08, \bar{m}(\{A, R\}) = 0.06, \bar{m}(\{H, R\}) = 0.12.$

Then we use the classical Dempster's rule to combine the weighted average BPA 4 times, which is the same as Murphy's approach [29]. The final result can be expressed as:

$m(\{A\}) = 0.9965, m(\{H\}) = 0.0030,$
$m(\{A, H\}) = 0.0004, m(\{H, R\}) = 0.0001.$

Hence, the target is recognized as an *Airplane*. Reexamining these BPAs provided by five sensors, we can note that all of them assign the maximum probability mass to the focus element $\{A\}$. This indicates that their fusion result should also pointed to $\{A\}$ as the identification of this target. Fortunately, the combination rule based on our proposed uncertainty measure and modified Murphy' approach can get reasonable result identical to intuitive analysis.

7 Conclusions

In this paper we mainly investigated the uncertainty measure of belief structures. An alternative uncertainty measure is defined based on interval probabilities. We first critically review existing uncertainty measures for belief structures. Then we prove that the interval probability $[Bel(\cdot), Pl(\cdot)]$

is an interval probability distribution. The presentation of our uncertainty measure *SU* then follows. It has been proved that *SU* is probability consistency. The range of *SU* is illustrated by both theoretical proof and numerical examples. Although this wider range is different from that of existing measures, it is determined by the cardinality of the discernment frame, which is of great significance in the application of uncertainty measure. Illustrative examples have reveals that our proposed measure is more sensitive to the change of belief structures than *AM* and *AU*. Furthermore, our new measure can provide more rational results in practical applications such as the sensor evaluation. As for the computational complexity of our proposal, we would like compare it with *AM* and *AU*. Considering the expression of *SU* and *AM*, we can notice that *SU* need some addition and arithmetic average operations, which are also necessary in the calculation of *AM*. But they are much simpler than logarithm operation and nonlinear optimization. So we can conclude that our proposed uncertainty measure will not confront heavy computation burden.

References

1. Abellán J, Masegosa A (2008) Requirements for total uncertainty measures in Dempster–Shafer theory of evidence. Int J Gen Syst 37(6):733–747
2. Awasthi A, Chauhan SS (2011) Using AHP and Dempster-Shafer theory for evaluating sustainable transport solutions. Environ Modell Softw 26(6):787–796
3. Dempster A (1967) Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 38(2):325–339
4. Deng X, Deng Y (2014) On the axiomatic requirement of range to measure uncertainty. Physica A 406:163–168
5. Deng X, Xiao F, Deng Y (2016) An improved distance-based total uncertainty measure in belief function theory. Appl Intell 46(4):898–915
6. Deng Y (2016) Deng entropy. Chaos, Solitons & Fractals 91:549–553
7. Deng Y, Shi W, Zhu Z, Liu Q (2004) Combining belief functions based on distance of evidence. Decis Support Syst 38(3):489–493
8. Denoeux T (1999) Reasoning with imprecise belief structures. Int J Approx Reason 20(1):79–111
9. Dubois D, Prade H (1985) A note on measures of specificity for fuzzy sets. Int J Gen Syst 10(4):279–283
10. Dubois D, Prade H (1999) Properties of measures of information in evidence and possibility theories. Fuzzy Sets Syst 100(1):35–49
11. Harmanec D, Klir GJ (1994) Measuring total uncertainty in Dempster-Shafer theory. Int J Gen Syst 22(4):405–419
12. Hartley RV (1928) Transmission of information. Bell Syst Tech J 7(3):535–563
13. Höhle U (1982) Entropy with respect to plausibility measures. In: Proceedings of the 12th IEEE international symposium on multiple-valued logic, pp 167–169
14. Jiang W, Zhan J (2017) A modified combination rule in generalized evidence theory. Appl Intell 46(3):630–640
15. Jousselme AL, Liu C, Grenier D, Bossé É (2006) Measuring ambiguity in the evidence theory. IEEE Trans Syst Man Cybern Syst Hum 36(5):890–903
16. Klir GJ (2001) Measuring uncertainty and uncertainty-based information for imprecise probabilities. In: International fuzzy systems association (IFSA) world congress and 20th North American fuzzy information processing society (NAFIPS) international conference. IEEE, pp 1927–1934
17. Klir GJ, Parviz B (1992) A note on the measure of discord. In: Proceedings of the Eighth international conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., pp 138–141
18. Klir GJ, Ramer A (1990) Uncertainty in Dempster-Shafer theory: a critical re-examination. Int J Gen Syst 18(2):155–166
19. Klir GJ, Smith RM (1999) Recent developments in generalized information theory. Int J Fuzzy Syst 1(1):1–13
20. Klir GJ, Wierman MJ (1999) Uncertainty-based information: elements of generalized information theory, vol 15. Springer Science & Business Media
21. Klir GJ, Yuan B (1995) Fuzzy sets and fuzzy logic: theory and applications. Prentice-Hall, Upper Saddle River
22. Klir G, Lewis HW (2008) Remarks on "measuring ambiguity in the evidence theory". IEEE Trans Syst Man Cybern Syst Hum 38(4):995–999
23. Körner R, Näther W (1995) On the specificity of evidences. Fuzzy Sets Syst 71(2):183–196
24. Liu Z, Pan Q, Dezert J, Martin A (2016) Adaptive imputation of missing values for incomplete pattern classification. Pattern Recog 52:85–95
25. Liu Z, Pan Q, Dezert J, Mercier G (2011) Combination of sources of evidence with different discounting factors based on a new dissimilarity measure. Decis Support Syst 52:133–141
26. Liu Z, Pan Q, Dezert J, Mercier G (2014) Credal classification rule for uncertain data based on belief functions. Pattern Recog 47(7):2532–2541
27. Liu Z, Pan Q, Dezert J, Mercier G (2015) Credal c-means clustering method based on belief functions. Knowl-Based Syst 74:119–132
28. Maeda Y, Nguyen HT, Ichihashi H (1993) Maximum entropy algorithms for uncertainty measures. Int J Uncertainty Fuzzy and Know-Based Syst 1(1):69–93
29. Murphy CK (2000) Combining belief functions when evidence conflicts. Decis Support Syst 29(1):1–9
30. Ramer A (1987) Uniqueness of information measure in the theory of evidence. Fuzzy Sets Syst 24(2):183–196
31. Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton
32. Shannon CE (2001) A mathematical theory of communication. ACM SIGMOBILE Mobile Comput Commun Rev 5(1):3–55
33. Smets P (2005) Decision making in the TBM: the necessity of the pignistic transformation. Int J Approx Reason 38(2):133–147
34. Wang YM, Yang JB, Xu DL, Chin KS (2006) The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees. Eur J Oper Res 175(1):35–66
35. Wang YM, Yang JB, Xu DL, Chin KS (2007) On the combination and normalization of interval-valued belief structures. Inf Sci 177:1230–1247
36. Yager RR (2001) Dempster-Shafer belief structures with interval valued focal weights. Int J Intell Syst 16:497–512
37. Yager RR, Alajlan N (2015) Dempster–Shafer belief structures for decision making under uncertainty. Knowl-Based Syst 80:58–66
38. Yager RR (1983) Entropy and specificity in a mathematical theory of evidence. Int J Gen Syst 9(4):249–260
39. Yang Y, Han D (2016) A new distance-based total uncertainty measure in the theory of belief functions. Knowl-Based Syst 94:114–123

Xiaodan Wang received her PH.D of computer science from North Western Polytechnical University of China. She is now a professor and PH.D advisor in Air Force Engineering University of China. Her research interest includes: pattern recognition, machine learning, and artificial intelligence.

Yafei Song is a lecturer at the College of Air and Missile Defense, Air Force Engineering University (AFEU). He received his Ph. D. degree from AFEU in 2015. He is working as a postdoctoral researcher in AFEU since April 2017. His research interest covers pattern recognition, intelligent information processing, and evidential reasoning.