

Article

A Novel Belief Entropy for Measuring Uncertainty in Dempster-Shafer Evidence Theory Framework Based on Plausibility Transformation and Weighted Hartley Entropy

Qian Pan ^{1,*} , Deyun Zhou ¹, Yongchuan Tang ¹ , Xiaoyang Li ¹ and Jichuan Huang ²

¹ School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China; dyzhounpu@nwpu.edu.cn (D.Z.); tangyongchuan@mail.nwpu.edu.cn (Y.T.); lixiaoyang@mail.nwpu.edu.cn (X.L.)

² First Military Representative Office of Air Force Equipment Department, People's Liberation Army Air Force, Chengdu 610013, China; jichuan1980@163.com

* Correspondence: panq@mail.nwpu.edu.cn; Tel.: +86-29-8843-1267

Received: 22 January 2019; Accepted: 7 February 2019; Published: 10 February 2019



Abstract: Dempster-Shafer evidence theory (DST) has shown its great advantages to tackle uncertainty in a wide variety of applications. However, how to quantify the information-based uncertainty of basic probability assignment (BPA) with belief entropy in DST framework is still an open issue. The main work of this study is to define a new belief entropy for measuring uncertainty of BPA. The proposed belief entropy has two components. The first component is based on the summation of the probability mass function (PMF) of single events contained in each BPA, which are obtained using plausibility transformation. The second component is the same as the weighted Hartley entropy. The two components could effectively measure the discord uncertainty and non-specificity uncertainty found in DST framework, respectively. The proposed belief entropy is proved to satisfy the majority of the desired properties for an uncertainty measure in DST framework. In addition, when BPA is probability distribution, the proposed method could degrade to Shannon entropy. The feasibility and superiority of the new belief entropy is verified according to the results of numerical experiments.

Keywords: Dempster-Shafer evidence theory; uncertainty of basic probability assignment; belief entropy; plausibility transformation; weighted Hartley entropy; Shannon entropy

1. Introduction

Dempster-Shafer evidence theory (DST) [1,2], which was initially introduced by Dempster in the context of statistical inference and then extended by Shafer into a general framework, has drawn great and continued attention in recent years [3–6]. The DST could be regarded as an extension of probability theory (PT). In DST, the probabilities are assigned to basic probability assignments (BPAs), which is presented to generalize the BPA in probability distribution in PT. The DST has shown its effectiveness and advantages in wide applications with uncertainty in terms of decision making, such as knowledge reasoning [7–9], sensor fusion [10–13], reliability analysis [14,15], fault diagnosis [16–18], assessment and evaluation [19–21], image recognition [22,23], and others [24–26].

Decision making in the framework of DST is based on the combination results of BPAs. Nonetheless, how to measure the uncertainty of BPA is still an open issue, which has not been completely solved [27]. The uncertainty of BPA mainly contains discord uncertainty and non-specificity uncertainty. Working out the uncertainty of BPA is the groundwork and precondition of applying

DST to applications [28]. Entropy was initially proposed to measure the uncertainty in statistical thermodynamics [29]. Then Claude Shannon extended this concept to solve the problem of information theory, namely Shannon entropy [30]. Although the Shannon entropy is admitted as an efficient way for measuring uncertainty in PT framework, it is unavailable to be used directly in the DST as the BPA described by sets of probabilities rather than single events [31]. For the sake of better standardizing the uncertainty measure in the framework of DST, Klir and Wierman defined a list of five basic required properties that an uncertainty measure should verify in DST [32]. Many attempts have been made to extend the Shannon entropy for measuring the uncertainty of BPA in the framework of DST, including Dubois and Prade's weighted Hartley entropy [33], Höhle's confusion uncertainty measure [34], Yager's dissonance uncertainty measure [35], Klir and Ramer's discord uncertainty measure [36], Klir and Parviz's strife uncertainty measure [37], Jousselme's ambiguity uncertainty measure [38], and Deng entropy [39]. Generally speaking, these approaches could degenerate to Shannon entropy if the probability values are assigned to single events. A belief entropy following Deng entropy is proposed by Pan and Deng to measure uncertainty in DST [40]. The method borrows from the idea of Deng entropy and is based on the probability interval, which is composed of the belief function and plausibility function. Although this Deng entropy-based method contains more information and could effectively measure the uncertainty in numerical cases, it does not satisfy most of the desired properties. Moreover, the expression of discord uncertainty measure in the method just considers the central values of the lower and upper bounds of the interval, which lacks explicit practical significance. Recently, Jiroušek and Shenoy added four new properties to the set of basic requirements. Thereafter, they define a belief entropy, which could verify six desired properties [41]. Their approach uses the probability mass function (PMF) transformed by plausibility transformation and weighted Hartley entropy to measure the discord and non-specificity uncertainty, respectively. However, the PMF used in the discord uncertainty measure may cause information loss when it is converted from BPA [42]. Hence, the discord uncertainty measure used in Jiroušek and Shenoy's belief entropy needs to be improved.

In this study, inspired by Pan and Deng's uncertainty measure [32] and Jiroušek and Shenoy's uncertainty measure [41], a novel belief entropy is proposed to measure the uncertainty in DST framework. The novel belief entropy has two components, the discord uncertainty measure and non-specificity uncertainty measure. The non-specificity uncertainty measure is the same as Dubois and Prade's weighted Hartley entropy, which could efficiently reflect the scale of each BPA. The discord uncertainty measure is based on the sum of PMFs transformed by plausibility transformation of single events, which are contained in each BPA. The sum of PMFs could be seen as the representative of probability interval with a practical significance. The discord uncertainty measure in the proposed method could capture sufficient information. In addition, the proposed method could satisfy six basic required properties.

The rest of this study is organized as follows. In Section 2, the preliminaries of DST, probability transformation of BPA, and Shannon entropy are briefly introduced. In Section 3, we discuss the desired properties of uncertainty measure in DST framework. Section 4 presents the existing belief entropies and the proposed belief entropy. The property analysis of the proposed belief entropy is also conducted in this section. In Section 5, some significant numerical experiments are carried out to illustrate the feasibility and effectiveness of the proposed belief entropy. Finally, in Section 6, the conclusion and future work are summarized.

2. Preliminaries

Some basic concepts are briefly introduced in this section, including Dempster-Shafer evidence theory [1,2], probability transformation of transforming a BPA to a PMF [43,44], and Shannon entropy [30].

2.1. Dempster-Shafer Evidence Theory

Let $\Theta = \{x_1, x_2, \dots, x_n\}$ be a nonempty finite set of mutually exclusive and collectively exhaustive alternatives. The Θ is called the frame of discernment (FOD). The power set of Θ is denoted by 2^Θ , namely

$$2^\Theta = \{\emptyset, \{x_1\}, \{x_2\}, \dots, \{x_n\}, \{x_1, x_2\}, \dots, \{x_1, x_2, \dots, x_i\}, \dots, \Theta\}, \tag{1}$$

A BPA is a mapping m from power set 2^Θ to $[0, 1]$, which satisfies the condition:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{A \in 2^\Theta} m(A) = 1. \tag{2}$$

A is called a focal element such that $m(A) > 0$. The BPA is also known as mass function.

There are two functions associated with each BPA called belief function $Bel(A)$ and plausibility function $Pl(A)$, respectively. The two functions are defined as follows:

$$\begin{aligned} Bel(A) &= \sum_{B \subseteq A} m(B), \\ Pl(A) &= \sum_{A \cap B \neq \emptyset} m(B). \end{aligned} \tag{3}$$

The plausibility function $Pl(A)$ denotes the degree of BPA that potentially supports A , while the belief function $Bel(A)$ denotes the degree of BPA that definitely supports A . Thus, $Bel(A)$ and $Pl(A)$ could be seen as the lower and upper probability of A .

Suppose m_1 and m_2 are two independent BPAs in the same FOD Θ , and they can be combined by using the Dempster-Shafer combination rule as follows:

$$m(A) = m_1 \oplus m_2 = \begin{cases} \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1-k}, & A \neq \emptyset \\ 0, & A = \emptyset \end{cases}, \tag{4}$$

with

$$k = \sum_{B \cap C = \emptyset} m_1(B) m_2(C), \tag{5}$$

where the k is the conflict coefficient to measure the degree of conflict among BPAs. The operator \oplus denotes the Dempster-Shafer combination rule. Please note that the Dempster-Shafer combination rule is unavailable for combining BPAs such that $k > 0$.

2.2. Probability Transformation

There are many ways to transform a BPA m to a PMF. Here, the pignistic transformation and the plausibility transformation are introduced.

Let m be a BPA on FOD Θ . Its associated probabilistic expression of PMF on Θ is defined as follows:

$$BetP(x) = \sum_{A \in 2^\Theta, x \in A} \frac{m(A)}{|A|}, \tag{6}$$

where the $|A|$ is the cardinality of A . The transformation between m and $BetP(x)$ is called the pignistic transformation.

$Pt(x)$ is a probabilistic expression of PMF that is obtained from m by using plausibility transformation as follows:

$$Pt(x) = \frac{Pl(x)}{\sum_{x \in \Theta} Pl(x)}, \tag{7}$$

where the $Pl(x)$ is the plausibility function of specific element x in Θ . The transformation between m and $Pt(x)$ is called the plausibility transformation.

2.3. Shannon Entropy

Let Ω be a FOD with possible values $\{w_1, w_2, \dots, w_n\}$. The Shannon entropy is explicitly defined as:

$$H_s = \sum_{w_i \in \Omega} p(w_i) \log_2 \left[\frac{1}{p(w_i)} \right], \quad (8)$$

where the $p(w_i)$ is the probability of alternative w_i , which satisfies $\sum_{i=1}^n p(w_i) = 1$. If some $p(w_i) = 0$, we follow the convention that $p(w_i) \log_2 \left[\frac{1}{p(w_i)} \right] = 0$ as $\lim_{x \rightarrow 0^+} x \log_2(x) = 0$. Please note that we will simply use \log for \log_2 in the rest of this paper.

3. Desired Properties of Uncertainty Measures in The DS Theory

In the research of Klir and Wierman [32], Klir and Lewis [45], and Klir [46], five basic required properties are defined for uncertainty measure in DST framework, namely probabilistic consistency, set consistency, range, sub-additivity, and additivity. These requirements are detailed as follows.

- *Probability consistency.* Let m be a BPA on FOD X . If m is a Bayesian BPA, then $H(m) = \sum_{x \in X} m(x) \log \left[\frac{1}{m(x)} \right]$.
- *Additivity.* Let m_X and m_Y be distinct BPAs for FOD X and FOD Y , respectively. The combined BPA $m_X \oplus m_Y$ using Dempster-Shafer combination rules must satisfy the following equality:

$$H(m_X \oplus m_Y) = H(m_X) + H(m_Y), \quad (9)$$

where the $m_X \oplus m_Y$ is a BPA for $\{X, Y\}$. For all $A \times B \in 2^{\{X, Y\}}$, where $A \in 2^X$ and $B \in 2^Y$, we have:

$$(m_X \oplus m_Y)(A \times B) = m_X(A)m_Y(B) \quad (10)$$

- *Sub-additivity.* Let m be a BPA on the space $X \times Y$, with marginal BPAs $m^{\downarrow X}$ and $m^{\downarrow Y}$ on FOD X and FOD Y , respectively. The uncertainty measure must satisfy the following inequality:

$$H(m) \leq H(m^{\downarrow X}) + H(m^{\downarrow Y}) \quad (11)$$

- *Set consistency.* Let m be a BPA on FOD X . If there exists a focal element $A \in X$ and $m(A) = 1$, then an uncertainty measure must degrade to Hartley measure:

$$H(m) = \log |A|. \quad (12)$$

- *Range.* Let m be a BPA on FOD X . The range of an uncertainty measure $H(m)$ must be $[0, \log |X|]$.

These properties illuminated in DST framework start from the verification by Shannon entropy in PT. In DST, there exist more situations of uncertainty than in PT framework [47]. Therewith, by analyzing shortcomings of these properties, Jiroušek and Shenoy add four other desired properties for measuring uncertainty in DST framework, including consistency with DST semantics, non-negativity, maximum entropy, monotonicity [41].

The uncertainty measure for BPA in DST must agree on the DST semantics [48]. Many uncertainty measures are based on the PMFs which are transformed from BPA [49–51]. However, only the plausibility transformation is compatible with the Dempster-Shafer combination rule [41,44]. Therefore, the property of consistency with DST semantics is presented to require the uncertainty measure to satisfy the tenets in DST framework.

- *Consistency with DST semantics.* Let m_1 and m_2 be two BPAs in the same FOD. If an uncertainty measure is based on a probability transformation of BPA, which transforms a BPA m to a PMF P_m , then the PMFs of m_1 and m_2 must satisfy the following condition:

$$P_{m_1 \oplus m_2} = P_{m_1} \otimes P_{m_2}, \quad (13)$$

where \otimes denotes the Bayesian combination rule [41], i.e., pointwise multiplication followed by normalization. Notice that this property is not presupposing the use of probability transformation in the uncertainty measure.

The property of additivity is easy to satisfy by most definitions of uncertainty measure [41]. The property of consistency with DST semantics is regarded as reinforcement of the additivity property, which makes sure that any uncertainty measure in DST framework follows the Dempster-Shafer combination rule.

Since the number of uncertainty type in DST framework is larger than that in PT framework. One can find that uncertainty measures in DST framework prefer a wider range than that in PT framework, namely $[0, \log |X|]$. Thus, in Jiroušek and Shenoy's opinion, the properties of non-negativity, maximum entropy, and monotonicity are pivotal to uncertainty measure in DST framework.

- *Non-negativity.* Let m be a BPA on FOD X . The uncertainty measure $H(m)$ must satisfy the following inequality:

$$H(m) \geq 0, \quad (14)$$

where the equality holds up if and only if m is Bayesian and $m\{(x)\} = 1$ with $x \in X$.

- *Maximum entropy.* Let m be a BPA on FOD X . The vacuous BPA m_v should have the most uncertainty, then the uncertainty measure must satisfy the following inequality:

$$H(m_v) \geq H(m), \quad (15)$$

where the equality holds up if and only if $m = m_v$.

- *Monotonicity.* Let v_X and v_Y be the vacuous BPAs of FOD X and FOD Y , respectively. If $|X| < |Y|$, then $H(v_X) < H(v_Y)$.

The property of set consistency entails that the uncertainty of a vacuous BPA m_v for FOD X is $\log |X|$. The probability consistency entails that the uncertainty of a Bayesian BPA m_e , which has the equally likely probabilities for X , is $\log |X|$ too. However, these two requirements are contradictory as the property of maximum entropy consider $H(m_v) > H(m_e)$. About this contradiction, there is a debatable open issue. Some researchers suggest the uncertainty of these two kinds of BPA should be equal and be the maximum possible uncertainty as we cannot get information to help us make a determinate decision [52,53]. Some other researchers deem the uncertainty of a vacuous BPA to be greater than a Bayesian uniform BPA, which is demonstrated by Ellsberg paradox phenomenon [54–56]. To provide a comprehensive understanding for our definition of uncertainty measure, all the above-mentioned properties are taken into account.

4. The Belief Entropy for Uncertainty Measure in DST Framework

4.1. The Existing Definitions of Belief Entropy of BPAs

The majority of the uncertainty measures have the Shannon entropy as the start point, which plays an important role to address the uncertainty in PT framework. Nevertheless, the Shannon entropy has inherent limitations to handle the uncertainty in DST as there are more types of uncertainty [27,57]. This is reasonable because the BPA includes more information than probabilistic distribution [4]. In the earlier literatures, the definitions of belief entropy only focus on one aspect of discord uncertainty or

non-specificity uncertainty in the BPAs. Then, Yager makes a contribution to distinction between the discord uncertainty and non-specificity uncertainty [35]. Thereafter, the discord and non-specificity are taken into consideration in most of the definitions of belief entropy. Some representative belief entropies and their definitions are listed as follows:

Höhel. One of the earliest uncertainty measures in DST is presented by Höhel as shown [34]:

$$H_o(m) = \sum_{A \in 2^X} m(A) \log \left[\frac{1}{\text{Bel}(A)} \right], \quad (16)$$

where the $\text{Bel}(A)$ is the belief function of proposition A . $H_o(m)$ only considers the discord uncertainty measure.

Nguyen defines the belief entropy of BPA m using the original BPAs [58]:

$$H_n(m) = \sum_{A \in 2^X} m(A) \log \left[\frac{1}{m(A)} \right]. \quad (17)$$

As the definition of $H_o(m)$, $H_n(m)$ only captures the discord part of uncertainty.

Dubois and Prade define the belief entropy using the cardinality of BPAs [33]:

$$H_d(m) = \sum_{A \in 2^X} m(A) \log |A|. \quad (18)$$

$H_d(m)$ considers only the non-specificity portion of the uncertainty. Dubois and Prade's definition could be regarded as the weighted Hartley entropy $H_h(m)$, where $H_h(m) = \log |A|$.

Pal et al. define a belief entropy as [59]:

$$H_p(m) = \sum_{A \in 2^X} m(A) \log \left[\frac{1}{m(A)} \right] + \sum_{A \in 2^X} m(A) \log (|A|). \quad (19)$$

In $H_p(m)$, the first component is the measure of discord uncertainty, and the second component is the measure of non-specificity uncertainty.

Jousselme et al. define a belief entropy based on the pignistic transformation [38]:

$$H_j(m) = \sum_{x \in X} \text{BetP}(x) \log \left[\frac{1}{\text{BetP}(x)} \right], \quad (20)$$

where the $\text{BetP}(x)$ is the PMF of pignistic transformation. The $H_j(m)$ using the Shannon entropy of $\text{BetP}(x)$

Deng defines a belief entropy, namely Deng entropy, as follows [39]:

$$H_{deng}(m) = \sum_{A \in 2^X} m(A) \log \left[\frac{1}{m(A)} \right] + \sum_{A \in 2^X} m(A) \log [2^{|A|} - 1]. \quad (21)$$

The $H_{deng}(m)$ is very similar to the definition of $H_p(m)$, while $H_{deng}(m)$ employs the $2^{|A|} - 1$ instead of $|A|$ to measure the non-specificity uncertainty of the BPA.

Pan and Deng develop Deng entropy $H_{deng}(m)$ with the definition [40]:

$$H_{pd}(m) = \sum_{A \in 2^X} \frac{1}{2} [\text{Bel}(A) + \text{Pl}(A)] \log \left\{ \frac{1}{\frac{1}{2} [\text{Bel}(A) + \text{Pl}(A)]} \right\} + \sum_{A \in 2^X} m(A) \log [2^{|A|} - 1], \quad (22)$$

where the $\text{Bel}(A)$ and $\text{Pl}(A)$ are the belief function and plausibility function, respectively. $H_{pd}(m)$ uses the central value of the probability interval $[\text{Bel}(A), \text{Pl}(A)]$ to measure the discord uncertainty of BPA.

It is obvious that all these uncertainty measures are the extension of the Shannon entropy in DST. Apart from the aforementioned methods of belief entropy, there are, of course, some other entropy-based uncertainty measures for BPAs in DST framework. One can find an expatiatory and detailed introduction to these methods in the literature [41,47].

Jiroušek and Shenoy define a concept for measuring uncertainty, as follows [41]:

$$H_{JS}(m) = \sum_{x \in X} Pt(x) \log \left[\frac{1}{Pt(x)} \right] + \sum_{A \in 2^X} m(A) \log (|A|). \quad (23)$$

The $H_{JS}(m)$ consists of two components. The first part is Shannon entropy of a PMF based on the plausibility transformation, which is associated with discord uncertainty. The second part is the entropy of Dubois and Prade for measuring non-specificity in BPAs. The $H_{JS}(m)$ satisfies the six desired properties, including consistency with DST semantics, non-negativity, maximum entropy, monotonicity, probability consistency, and additivity. Moreover, the properties of range and set consistency are expanded.

4.2. The Proposed Belief Entropy

Although the $H_{JS}(m)$ can better meet the requirement of the basic properties for uncertainty measure, it has an intrinsic defect. The first part in $H_{JS}(m)$ using Shannon entropy captures only the probability of plausibility transformation, which may lead to information loss. As argued in $H_{pd}(m)$, the probability interval $[Bel(A), Pl(A)]$ can provide more information according to the BPAs in each proposition. However, the $H_{pd}(m)$ considers only the numerical average of the probability interval, which lacks the piratical physical significance. In this study, by combining the merit of $H_{JS}(m)$ and $H_{pd}(m)$, a new definition of belief entropy-based uncertainty measure in DST framework is proposed as follows:

$$H_{PQ}(m) = \sum_{A \in 2^X} m(A) \log \left[\frac{1}{Pm(A)} \right] + \sum_{A \in 2^X} m(A) \log (|A|), \quad (24)$$

where the $Pm(A) = \sum_{x \in A} Pt(x)$ is the summation of plausibility transformation-based PMFs of x contained in A .

Similar to most of the belief entropies, the first component $\sum_{A \in 2^X} m(A) \log [Pm^{-1}(A)]$ in $H_{PQ}(m)$ is designed to measure the discord uncertainty of BPA. The information contained in not only BPAs but also the plausibility function based on $Pt(x)$ is taken into consideration. Since the $Pt(x)$ reflects the support degree of different propositions to element x , it could provide more information than $m(A)$. Furthermore, the $Pm(A) = \sum_{x \in A} Pt(x)$ satisfies the $Bel(A) \leq Pm(A) \leq Pl(A)$, which could be seen as a representative of the probability interval. At length, the second component $\sum_{A \in 2^X} m(A) \log (|m(A)|)$ in H_{PQ} is the same as the $H_d(m)$ to measure the non-specificity uncertainty of BPA. Therefore, we believe that the new proposed belief entropy can be more effective to measure the uncertainty of BPAs in DST framework. The property analysis of $H_{PQ}(m)$ is explored as follows.

(1) Consistency with DST semantics. The first part in $H_{PQ}(m)$ uses $Pt(x)$ based on the plausibility transformation, which is compatible with the definition of the property. The second part is not a Shannon entropy based on probability transformation. Thus, $H_{PQ}(m)$ satisfies the consistency with DST semantics property.

(2) Non-negativity. As $Pm(A) \in [0, 1]$, $m(A) \in [0, 1]$ and $0 < |m(A)|$, thus, $H_{PQ}(m) \geq 0$. If and only if the m is a Bayesian BPA and $m(x) = 1$, $H_{PQ}(m) = 0$. Thus, $H_{PQ}(m)$ satisfies the non-negativity property.

(3) Maximum entropy. Let m_e and m_v be a uniform Bayesian BPA and a vacuous BPA in the same FOD X , respectively. We could obtain $H_{PQ}(m_e) = H_{PQ}(m_v) = \log (|X|)$, therefore $H_{PQ}(m)$ dissatisfies the maximum entropy property.

(4) Monotonicity. Since $H_{PQ}(m_v) = \log (|X|)$, $H_{PQ}(m_v)$ is monotonic in $|X|$. Therefore $H_{PQ}(m)$ satisfies the monotonicity property.

(5) Probability consistency. If m is a Bayesian BPA, then $Pm(x) = Pt(x) = m(x)$ and $H_d(m) = 0$. Hence, $H_{PQ}(m) = \sum_{x \in X} m(x) \log \left[\frac{1}{m(x)} \right]$. Therewith, we know that the $H_{PQ}(m)$ satisfies the probability consistency property.

(6) Set consistency. If a focal element has the whole support degree such that $m(A) = 1$, $H_{PQ}(m) = \log(|X|)$. Hence, the $H_{PQ}(m)$ satisfies the set consistency property.

(7) Range. As $Pm(A)$ includes the support from the other propositions, thus, $m(A) \leq Pm(A)$. Therefore, $\sum_{A \in 2^X} m(A) \log [Pm^{-1}(A)] \leq \sum_{A \in 2^X} m(A) \log [m^{-1}(A)] = H_n(m)$. The range of $H_n(m)$ and $H_d(m)$ both are $[0, \log(|X|)]$. Thus, the range of $H_{PQ}(m)$ is $[0, 2\log(|X|)]$, which means the $H_{PQ}(m)$ dissatisfies the range property.

(8) Additivity. Let m_X and m_Y be two BPAs of FOD X and FOD Y , respectively, $A \subseteq 2^X$, and $B \subseteq 2^Y$. Let $C = A \times B$ be the corresponding joint focal element on $X \times Y$, $x \in X$, and $y \in Y$. Let m be a joint BPA defined on $X \times Y$ which is obtained by using Equation (10). Thus, $m(C) = (m_X \oplus m_Y)(A \times B) = m_X(A)m_Y(B)$. Then the new belief entropy for m is:

$$H_{PQ}(m) = H_{PQ}(m_X \oplus m_Y) = \sum_{C \in 2^{X \times Y}} m(C) \log \left[\frac{|C|}{Pm(C)} \right],$$

where

$$m(C) = m_X(A)m_Y(B),$$

$$Pm(C) = Pm(A \times B) = \sum_{(x,y) \in \{A \times B\}} Pt(x,y) = \sum_{(x,y) \in \{A \times B\}} \frac{Pl(x,y)}{\sum_{(x,y) \in \{X \times Y\}} Pl(x,y)}.$$

As proved in [33], we have $Pl(A \times B) = Pl(A)Pl(B)$. Thus, we know

$$\sum_{(x,y) \in \{A \times B\}} \frac{Pl(x,y)}{\sum_{(x,y) \in \{X \times Y\}} Pl(x,y)} = \sum_{x \in A} \sum_{y \in B} \frac{Pl(x)Pl(y)}{\sum_{x \in X} \sum_{y \in Y} Pl(x)Pl(y)} = \sum_{x \in A} \frac{Pl(x)}{\sum_{x \in X} Pl(x)} \sum_{y \in B} \frac{Pl(y)}{\sum_{y \in Y} Pl(y)}$$

and

$$Pm(C) = \sum_{x \in A} \frac{Pl(x)}{\sum_{x \in X} Pl(x)} \sum_{y \in B} \frac{Pl(y)}{\sum_{y \in Y} Pl(y)} = Pm(A)Pm(B).$$

Consequently,

$$\begin{aligned} H_{PQ}(m_X \oplus m_Y) &= \sum_{C \in 2^{X \times Y}} m(C) \log \left[\frac{|A||B|}{Pm(A)Pm(B)} \right] \\ &= \sum_{A \in 2^X} \sum_{B \in 2^Y} m_X(A)m_Y(B) \log \left[\frac{|A|}{Pm(A)} \right] + \sum_{A \in 2^X} \sum_{B \in 2^Y} m_X(A)m_Y(B) \log \left[\frac{|B|}{Pm(B)} \right] \\ &= \sum_{A \in 2^X} m_X(A) \sum_{B \in 2^Y} m_Y(B) \log \left[\frac{|A|}{Pm(A)} \right] + \sum_{A \in 2^X} m_X(A) \sum_{B \in 2^Y} m_Y(B) \log \left[\frac{|B|}{Pm(B)} \right] \\ &= \sum_{A \in 2^X} m_X(A) \log \left[\frac{|A|}{Pm(A)} \right] + \sum_{B \in 2^Y} m_Y(B) \log \left[\frac{|B|}{Pm(B)} \right] \\ &= H_{PQ}(m_X) + H_{PQ}(m_Y). \end{aligned}$$

Hence, the $H_{PQ}(m)$ satisfies the additivity property.

(9) Sub-additivity. An example of binary-valued variables is given to check whether the $H_{PQ}(m)$ satisfies the sub-additivity as follows with masses

$$m(z_{11}) = m(z_{12}) = 0.1, m(z_{21}) = m(z_{22}) = 0.3, m(X \times Y) = 0.2,$$

where $z_{ij} = (x_i, y_j)$. The marginal BPAs of for X and Y are $m^{\downarrow X}$ and $m^{\downarrow Y}$, respectively, shown as following ones.

$$m^{\downarrow X}(x_1) = 0.2, m^{\downarrow X}(x_2) = 0.6, m^{\downarrow X}(X) = 0.2$$

$$m^{\downarrow Y}(y_1) = 0.4, m^{\downarrow Y}(y_2) = 0.4, m^{\downarrow Y}(Y) = 0.2$$

Thus,

$$Pl(x_1) = 0.4, Pl(x_2) = 0.8, Pl(y_1) = 0.5, Pl(y_2) = 0.5,$$

$$Pl(Z_{11}) = 0.3, Pl(Z_{12}) = 0.3, Pl(Z_{21}) = 0.5, Pl(Z_{22}) = 0.5$$

$$Pt(x_1) = 0.333, Pt(x_2) = 0.667, Pt(y_1) = 0.5, Pt(y_2) = 0.5,$$

$$Pt(Z_{11}) = 0.188, Pt(Z_{12}) = 0.188, Pt(Z_{21}) = 0.312, Pt(Z_{22}) = 0.312$$

$$H_{PQ}(m) = 1.8899, H_{PQ}(m^{\downarrow X}) + H_{PQ}(m^{\downarrow Y}) = 0.8678 + 1 = 1.8687.$$

Obviously, $H_{PQ}(m) > H_{PQ}(m^{\downarrow X}) + H_{PQ}(m^{\downarrow Y})$, thus the $H_{PQ}(m)$ dissatisfies the sub-additivity property.

In summary, the new belief entropy $H_{PQ}(m)$ for uncertainty measure in DST framework satisfies the properties of consistency with DST semantics, non-negativity, set consistency, probability consistency, additivity, monotonicity, and does not satisfy the properties of sub-additivity, maximum entropy, range. An overview of the properties of existing belief entropies for uncertainty measure are listed in Table 1.

Table 1. An overview of the properties of existing belief entropies and the proposed method.

Definition	Cons.w DST	Non-neg	Max. ent	Monoton	Prob. cons	Add	Subadd	Range	Set. cons
Höhle	yes	no	no	no	yes	yes	no	yes	no
Smets	yes	no	no	no	no	yes	no	yes	no
Yager	yes	no	no	no	yes	yes	no	yes	no
Nguyen	yes	no	no	no	yes	yes	no	yes	no
Dubois-Prade	yes	no	yes	yes	no	yes	yes	yes	yes
Klir-Ramner	yes	yes	no	yes	yes	yes	no	no	yes
Klir-Parviz	yes	yes	no	yes	yes	yes	no	no	yes
Pal et al.	yes	yes	no	yes	yes	yes	no	no	yes
George-Pal	yes	no	no	no	no	no	no	no	yes
Maeda-Ichihashi	no	yes	yes	yes	yes	yes	yes	no	yes
Harmanec-Klir	no	yes	no	yes	yes	yes	yes	no	no
Abellán-Moral	no	yes	yes	yes	yes	yes	yes	no	yes
Jousselme et al.	no	yes	no	yes	yes	yes	no	yes	yes
Pouly et al.	no	yes	no	yes	yes	yes	no	no	yes
Jiroušek-Shenoy	yes	yes	yes	yes	yes	yes	no	no	no
Deng	yes	yes	no	yes	yes	no	no	no	no
Pan-Deng	yes	yes	no	yes	yes	no	no	no	no
Proposed method	yes	yes	no	yes	yes	yes	no	no	yes

Additionally, based on combining the advantages of the definition of Jiroušek-Shenoy and Pan-Deng, the new belief entropy involves more information, which can better meet the requirements. The properties of maximum entropy and range that the new belief entropy dissatisfies need further discussion. For maximum entropy properties, we think that the uncertainty of a vacuous BPA and an equally likely Bayesian BPA should be equivalent. There is a classical example.

Assume a bet on a race conducted by four cars, A, B, C , and D . Two experts give their opinion. Expert-1 suggests that the ability of the four drivers and the performance of the four cars are almost the same. Expert-2 has no idea about the traits of each car and driver. The opinion of the Expert-1 could be regarded as a uniform probability distribution with $m(A) = m(B) = m(C) = m(D) = \frac{1}{4}$. while the Expert-2 produces a vacuous BPA with $m(A, B, C, D) = 1$. Based on only one piece of these two pieces of evidence, we have no information to support us to make a certain bet. Besides, it is very convincing that the range property is not suitable for uncertain measure. The range $[0, \log(|X|)]$

can only reflect one aspect of uncertainty, which lacks consideration for multiple uncertainties of a BPA in DST framework. As a consequence, the properties of maximum entropy and range should be extended.

5. Numerical Experiment

In this section, several numerical experiments are verified to demonstrate the reasonability and effectiveness of our proposed new belief entropy.

5.1. Example 1

Let $\Theta = \{x\}$ be the FOD. Given a BPA with $m(x) = 1$, we can obtain the $Pt(x)$ and $Pm(x)$ with:

$$Pt(x) = 1, Pm(x) = 1.$$

Then, the associated Shannon entropy $H_s(m)$ and the proposed belief entropy $H_{PQ}(m)$ are calculated as follows:

$$H_s(m) = 1 \times \log(1) = 0, H_{PQ}(m) = 1 \times \log\left(\frac{1}{1}\right) = 0.$$

Obviously, the above example shows that the Shannon entropy and the proposed belief entropy are equal when the FOD has only one single element, where exists no uncertainty.

5.2. Example 2

Let $\Theta = \{x_1, x_2, x_3, x_4, x_5\}$ be the FOD. A uniform BPA of FOD is given as $m(x_1) = m(x_2) = m(x_3) = m(x_4) = m(x_5) = \frac{1}{5}$. Then,

$$\begin{aligned} Pt(x_1) &= Pt(x_2) = Pt(x_3) = Pt(x_4) = Pt(x_5) = \frac{1}{5}, \\ Pm(x_1) &= Pm(x_2) = Pm(x_3) = Pm(x_4) = Pm(x_5) = \frac{1}{5}, \\ H_s(m) &= \frac{1}{5} \times \log 5 + \frac{1}{5} \times \log 5 = 2.3219, \\ H_{PQ}(m) &= \frac{1}{5} \times \log 5 + \dots \\ &\dots + \frac{1}{5} \times \log 1 = 2.3219. \end{aligned}$$

As shown above, the proposed belief entropy is the same as the Shannon entropy when the BPA is the probability distribution. Sections 5.1 and 5.2 verify that the proposed belief entropy will degenerate into the Shannon entropy when the belief is assigned to singleton elements.

5.3. Example 3

Let $\Theta = \{x_1, x_2, x_3, x_4, x_5\}$ be the FOD. A vacuous BPA of FOD is given as $m(x_1, x_2, x_3, x_4, x_5) = 1$. Then,

$$\begin{aligned} Pt(x_1) &= Pt(x_2) = Pt(x_3) = Pt(x_4) = Pt(x_5) = \frac{1}{5}, \\ Pm(x_1) &= Pm(x_2) = Pm(x_3) = Pm(x_4) = Pm(x_5) = \frac{1}{5}, \\ H_{PQ}(m) &= 1 \times \log 1 + 1 \times \log 5 = 2.3219. \end{aligned}$$

Compared to Section 5.2, we know that the uncertainty of this example is the same as the Section 5.2. This is reasonable. As discussed in Section 4.2, neither the uniform BPA nor the vacuous

BPA in the same FOD could provide more information for a determinate single element. Thus, their uncertainty should be equal.

5.4. Example 4

Two experiments in [40] are recalled in this example. Let $\Theta = \{x_1, x_2, x_3, x_4\}$ be the FOD. Two BPAs are given as m_1 and m_2 . The detailed BPAs are:

$$m_1(x_1) = \frac{1}{4}, m_1(x_2) = \frac{1}{3}, m_1(x_3) = \frac{1}{6}, m_1(x_1, x_2, x_3) = \frac{1}{6}, m_1(x_4) = \frac{1}{12},$$

$$m_2(x_1) = \frac{1}{4}, m_2(x_2) = \frac{1}{3}, m_2(x_3) = \frac{1}{6}, m_2(x_1, x_2) = \frac{1}{6}, m_2(x_4) = \frac{1}{12}.$$

The corresponding $H_{PQ}(m_1)$ and $H_{PQ}(m_2)$ are calculated as follows:

$$Pl_{m_1}(x_1) = \frac{5}{12}, Pl_{m_1}(x_2) = \frac{1}{2}, Pl_{m_1}(x_3) = \frac{1}{3}, Pl_{m_1}(x_4) = \frac{1}{12},$$

$$Pl_{m_2}(x_1) = \frac{5}{12}, Pl_{m_2}(x_2) = \frac{1}{2}, Pl_{m_2}(x_3) = \frac{1}{6}, Pl_{m_2}(x_4) = \frac{1}{12},$$

$$Pt_{m_1}(x_1) = \frac{5}{16}, Pt_{m_1}(x_2) = \frac{6}{16}, Pt_{m_1}(x_3) = \frac{4}{16}, Pt_{m_1}(x_4) = \frac{1}{16},$$

$$Pt_{m_2}(x_1) = \frac{5}{14}, Pt_{m_2}(x_2) = \frac{6}{14}, Pt_{m_2}(x_3) = \frac{2}{14}, Pt_{m_2}(x_4) = \frac{1}{14},$$

$$H_{PQ}(m_1) = \frac{1}{4} \times \log\left(\frac{1}{5/16}\right) + \frac{1}{3} \times \log\left(\frac{1}{6/16}\right) + \frac{1}{6} \times \log\left(\frac{1}{4/16}\right) + \dots$$

$$\dots + \frac{1}{6} \times \log\left(\frac{3}{15/16}\right) + \frac{1}{12} \times \log\left(\frac{1}{1/16}\right) = 1.8375,$$

$$H_{PQ}(m_2) = \frac{1}{4} \times \log\left(\frac{1}{5/14}\right) + \frac{1}{3} \times \log\left(\frac{1}{6/14}\right) + \frac{1}{6} \times \log\left(\frac{1}{2/14}\right) + \dots$$

$$\dots + \frac{1}{6} \times \log\left(\frac{2}{13/14}\right) + \frac{1}{12} \times \log\left(\frac{1}{1/14}\right) = 1.7485.$$

It can be seen from the results of $H_{PQ}(m_1)$ and $H_{PQ}(m_2)$, the belief entropy of m_1 is larger than the m_2 . This is logical because the $m_1(x_1, x_2, x_3) = \frac{1}{6}$ has one more single element than $m_2(x_1, x_2) = \frac{1}{6}$, which implies that the $m_1(x_1, x_2, x_3)$ contains more information. Thus, the m_1 should be more uncertain.

5.5. Example 5

Consider a target recognition problem in [60]. Target detection results provided by two independent sensors. Let A, B, C , and D be the potential target types. The results are represented by BPAs shown as follows.

$$m_1(A, B) = 0.4, m_1(C, D) = 0.6,$$

$$m_2(A, C) = 0.4, m_2(B, C) = 0.6.$$

Then the corresponding uncertainty measure with $H_{den}(m)$, $H_{pd}(m)$ and $H_{PQ}(m)$ are calculated as:

$$\begin{aligned}
& Bel_{m_1}(A, B) = 0.4, Pl_{m_1}(A, B) = 0.4, Bel_{m_1}(C, D) = 0.6, Pl_{m_1}(C, D) = 0.6, \\
& Bel_{m_2}(A, C) = 0.4, Pl_{m_2}(A, C) = 1.0, Bel_{m_2}(B, C) = 0.6, Pl_{m_2}(B, C) = 1.0, \\
& H_d(m_1) = 0.4 \log \frac{2^{|2|} - 1}{0.4} + 0.6 \log \frac{2^{|2|} - 1}{0.6} = 2.5559, \\
& H_d(m_2) = 0.4 \log \frac{2^{|2|} - 1}{0.4} + 0.6 \log \frac{2^{|2|} - 1}{0.6} = 2.5559, \\
& H_{pd}(m_1) = \frac{0.4 + 0.4}{2} \log \frac{2^{|2|} - 1}{(0.4 + 0.4)/2} + \frac{0.6 + 0.6}{2} \log \frac{2^{|2|} - 1}{(0.6 + 0.6)/2} = 2.5559, \\
& H_{pd}(m_2) = \frac{0.4 + 1.0}{2} \log \frac{2^{|2|} - 1}{(0.4 + 1.0)/2} + \frac{0.6 + 1.0}{2} \log \frac{2^{|2|} - 1}{(0.6 + 1.0)/2} = 2.9952, \\
& Pl_{m_1}(A) = 0.4, Pl_{m_1}(B) = 0.4, Pl_{m_1}(C) = 0.6, Pl_{m_1}(D) = 0.6, \\
& Pl_{m_2}(A) = 0.4, Pl_{m_2}(B) = 0.6, Pl_{m_2}(C) = 1.0, \\
& Pt_{m_1}(A) = 0.2, Pt_{m_1}(B) = 0.2, Pt_{m_1}(C) = 0.3, Pt_{m_1}(D) = 0.3, \\
& Pt_{m_2}(A) = 0.2, Pt_{m_2}(B) = 0.3, Pt_{m_2}(C) = 0.5, \\
& H_{PQ}(m_1) = 0.4 \times \log \frac{2}{0.4} + 0.6 \times \log \frac{2}{0.6} = 1.9710, \\
& H_{PQ}(m_2) = 0.4 \times \log \frac{2}{0.7} + 0.6 \times \log \frac{2}{0.8} = 1.3390.
\end{aligned}$$

Though the two BPAs have the same value, the BPA m_1 has four potential targets, namely A, B, C, D , while the BPA m_2 has just three potential targets, namely A, B, C . As verified in [60], it is intuitively expected that m_1 has a larger uncertainty than m_2 . According to the above calculation results, the $H_{den}(m)$ illustrates that the two BPAs have the same uncertainty, and the $H_{pd}(m)$ suggests that the m_2 has a larger uncertainty. Therefore, both $H_{den}(m)$ and $H_{pd}(m)$ are unable to reflect the prospective difference. The proposed belief entropy can effectively quantify this divergence by considering not only the information contained in each focal element but also the mutual support degree among different focal elements. Therefore, it is safe to say that the capability of the proposed belief entropy $H_{PQ}(m)$ is unavailable in the $H_{den}(m)$ and $H_{pd}(m)$.

5.6. Example 6

Let $\Theta = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ be the FOD. Two BPAs are given as follows.

$$\begin{aligned}
m_1(x_1, x_2) &= \frac{1}{3}, m_1(x_3, x_4) = \frac{1}{3}, m_1(x_5, x_6) = \frac{1}{3}, \\
m_2(x_1, x_2, x_3) &= \frac{1}{2}, m_2(x_4, x_5, x_6) = \frac{1}{2}.
\end{aligned}$$

According to the Jiroušek-Shenoy entropy $H_{JS}(m)$ in Equation (23) and the proposed belief entropy $H_{PQ}(m)$ in Equation (24), both kinds of entropy consists of the discord uncertainty measure and the non-specificity uncertainty measure. Then the $H_{JS}(m)$ and $H_{PQ}(m)$ are calculated as follows.

$$\begin{aligned}
Pl_{m_1}(x_1) &= \frac{1}{3}, Pl_{m_1}(x_2) = \frac{1}{3}, Pl_{m_1}(x_3) = \frac{1}{3}, Pl_{m_1}(x_4) = \frac{1}{3}, Pl_{m_1}(x_5) = \frac{1}{3}, Pl_{m_1}(x_6) = \frac{1}{3}, \\
Pl_{m_2}(x_1) &= \frac{1}{2}, Pl_{m_2}(x_2) = \frac{1}{2}, Pl_{m_2}(x_3) = \frac{1}{2}, Pl_{m_2}(x_4) = \frac{1}{2}, Pl_{m_2}(x_5) = \frac{1}{2}, Pl_{m_2}(x_6) = \frac{1}{2}, \\
Pt_{m_1}(x_1) &= \frac{1}{6}, Pt_{m_1}(x_2) = \frac{1}{6}, Pt_{m_1}(x_3) = \frac{1}{6}, Pt_{m_1}(x_4) = \frac{1}{6}, Pt_{m_1}(x_5) = \frac{1}{6}, Pt_{m_1}(x_6) = \frac{1}{6}, \\
Pt_{m_2}(x_1) &= \frac{1}{6}, Pt_{m_2}(x_2) = \frac{1}{6}, Pt_{m_2}(x_3) = \frac{1}{6}, Pt_{m_2}(x_4) = \frac{1}{6}, Pt_{m_2}(x_5) = \frac{1}{6}, Pt_{m_2}(x_6) = \frac{1}{6}, \\
H_{JS}(m_1) &= H_{JS}^{dis}(m_1) + H_{JS}^{nos-spe}(m_1) = 6 \times \frac{1}{6} \times \log\left(\frac{1}{1/6}\right) + 3 \times \frac{1}{3} \times \log(2) = 3.5850, \\
H_{JS}(m_2) &= H_{JS}^{dis}(m_2) + H_{JS}^{nos-spe}(m_2) = 6 \times \frac{1}{6} \times \log\left(\frac{1}{1/6}\right) + 2 \times \frac{1}{2} \times \log(3) = 4.1699, \\
H_{PQ}(m_1) &= H_{PQ}^{dis}(m_1) + H_{PQ}^{nos-spe}(m_1) = 3 \times \frac{1}{3} \times \log\left(\frac{1}{1/3}\right) + 3 \times \frac{1}{3} \times \log(2) = 2.5850, \\
H_{PQ}(m_2) &= H_{PQ}^{dis}(m_2) + H_{PQ}^{nos-spe}(m_2) = 2 \times \frac{1}{2} \times \log\left(\frac{1}{1/2}\right) + 2 \times \frac{1}{2} \times \log(3) = 2.5850.
\end{aligned}$$

The results calculated by the $H_{JS}^{dis}(m_1)$ and $H_{JS}^{dis}(m_2)$ are the same, which are equal to $\log(6)$. This outcome is counterintuitive. The BPAs in m_1 are completely different from that in m_2 , thus the $H_{JS}^{dis}(m)$ and $H_{JS}^{nos-spe}(m)$ of m_1 are expected to be distinguished from those ones of m_2 . However, only the $H_{JS}^{nos-spe}(m_1)$ and $H_{JS}^{nos-spe}(m_2)$ are different. The reason for this situation is that the discord uncertainty measure $H_{JS}^{dis}(m)$ in $H_{JS}(m)$ overly concerns the conflict involved in single elements and ignores the information contained in the original BPAs. The discord uncertainty measure $H_{PQ}^{dis}(m)$ in $H_{PQ}(m)$ combines the original BPAs with the probability distribution of single elements included in the BPA can better resolve the limitations. In short, this example indicates the effectiveness for measuring the discord uncertainty of the proposed belief entropy.

5.7. Example 7

Let $\Theta = (1, 2, \dots, 14, 15)$ be a FOD with 15 elements. The mass functions of Θ is denoted as:

$$m(3, 4, 5) = 0.05, \quad m(7) = 0.05, \quad m(A) = 0.8, \quad m(\Theta) = 0.1$$

The proposition A is a variable subset of 2^Θ with the number of single elements changing from 1 to 14. To verify the merit and effectiveness of the proposed belief entropy, eight uncertainty measures listed in Table 1 are selected for comparison, including Dubois and Prade's weighted Hartley entropy [33], Höhle's confusion uncertainty measure [34], Yager's dissonance uncertainty measure [35], Klir and Ramer's discord uncertainty measure [36], Klir and Parviz's strife uncertainty measure [37], George and Pal's conflict uncertainty measure [61], Pan and Deng's uncertainty measure [40], Jiroušek and Shenoy's uncertainty measure [41]. The experimental results are shown in Table 2. The Höhle's confusion uncertainty measure (M_2), Yager's dissonance uncertainty measure (M_3), Klir and Ramer's discord uncertainty measure (M_4), Klir and Parviz's strife uncertainty measure (M_5), and George and Pal's conflict uncertainty measure (M_6) are plotted in Figure 1. The Dubois and Prade's weighted Hartley entropy (M_1), Pan and Deng's uncertainty measure (M_7), Jiroušek and Shenoy's uncertainty measure (M_8), and proposed belief entropy (M_9) are plotted in Figure 2.

Table 2. The value of different uncertainty measures.

Cases	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9
$A=\{1\}$	0.4699	0.6897	0.3953	6.4419	3.3804	0.3317	16.1443	3.8322	1.9757
$A=\{1, 2\}$	1.2699	0.6897	0.3953	5.6419	3.2956	0.3210	17.4916	4.4789	2.3362
$A=\{1, 2, 3\}$	1.7379	0.6897	0.1997	4.2823	2.9709	0.2943	19.8608	4.8870	2.5232
$A=\{1, 2, 3, 4\}$	2.0699	0.6897	0.1997	3.6863	2.8132	0.2677	20.8229	5.2250	2.7085
$A=\{1, 2, 3, 4, 5\}$	2.3275	0.6198	0.1997	3.2946	2.7121	0.2410	21.8314	5.5200	2.8749
$A=\{1, 2, \dots, 6\}$	2.5379	0.6198	0.1997	3.2184	2.7322	0.2383	22.7521	5.8059	3.0516
$A=\{1, 2, \dots, 7\}$	2.7158	0.5538	0.0074	2.4562	2.5198	0.2220	24.1131	6.0425	3.0647
$A=\{1, 2, \dots, 8\}$	2.8699	0.5538	0.0074	2.4230	2.5336	0.2170	25.0685	6.2772	3.2042
$A=\{1, 2, \dots, 9\}$	3.0059	0.5538	0.0074	2.3898	2.5431	0.2108	26.0212	6.4921	3.3300
$A=\{1, 2, \dots, 10\}$	3.1275	0.5538	0.0074	2.3568	2.5494	0.2037	27.1947	6.6903	3.4445
$A=\{1, 2, \dots, 11\}$	3.2375	0.5538	0.0074	2.3241	2.5536	0.1959	27.9232	6.8743	3.5497
$A=\{1, 2, \dots, 12\}$	3.3379	0.5538	0.0074	2.2920	2.5562	0.1877	29.1370	7.0461	3.6469
$A=\{1, 2, \dots, 13\}$	3.4303	0.5538	0.0074	2.2605	2.5577	0.1791	30.1231	7.2071	3.7374
$A=\{1, 2, \dots, 14\}$	3.5158	0.5538	0.0074	2.2296	2.5582	0.1701	31.0732	7.3587	3.8219

M_1 is the Dubois and Prade’s weighted Hartley entropy; M_2 is the Höhle’s confusion uncertainty measure; M_3 is the Yager’s dissonance uncertainty measure; M_4 is the Klir and Ramer’s discord uncertainty measure; M_5 is the Klir and Parviz’s strife uncertainty measure; M_6 is the George and Pal’s conflict uncertainty measure; M_7 is the Pan and Deng’s uncertainty measure; M_8 is the Jiroušek and Shenoy’s uncertainty measure; M_9 is the proposed belief entropy.

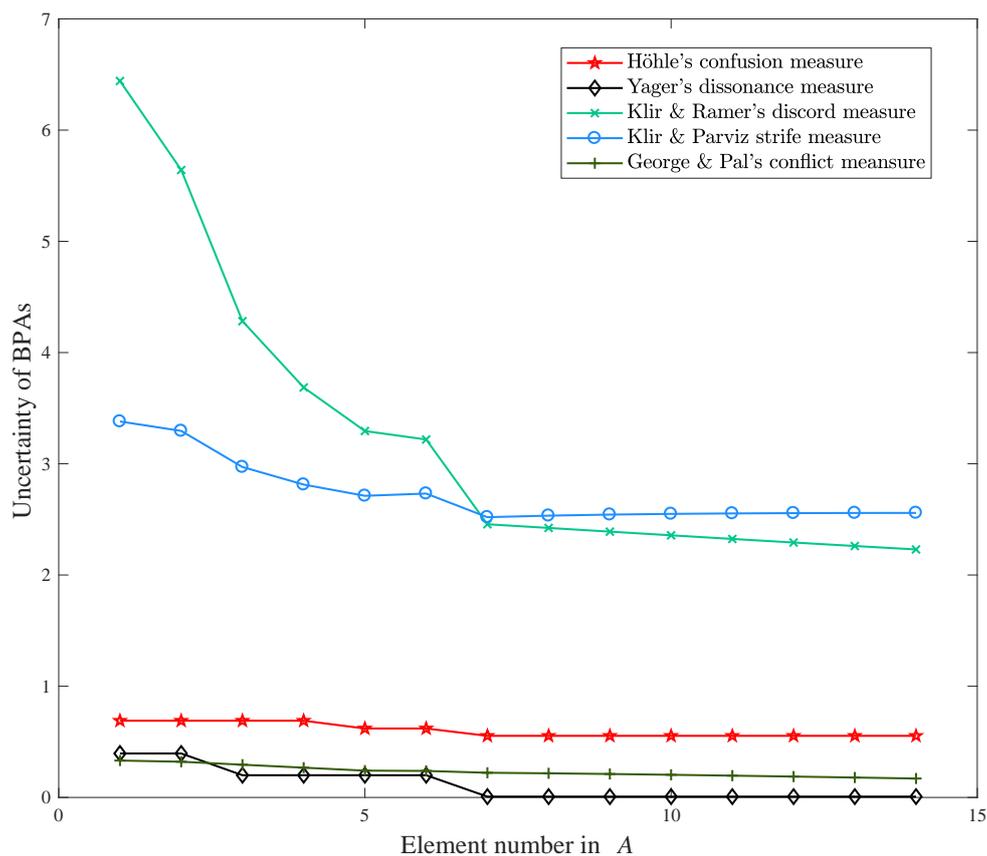


Figure 1. Results comparison of M_2 , M_3 , M_4 , M_5 , and M_6 in DST (Dempster-Shafer evidence theory).

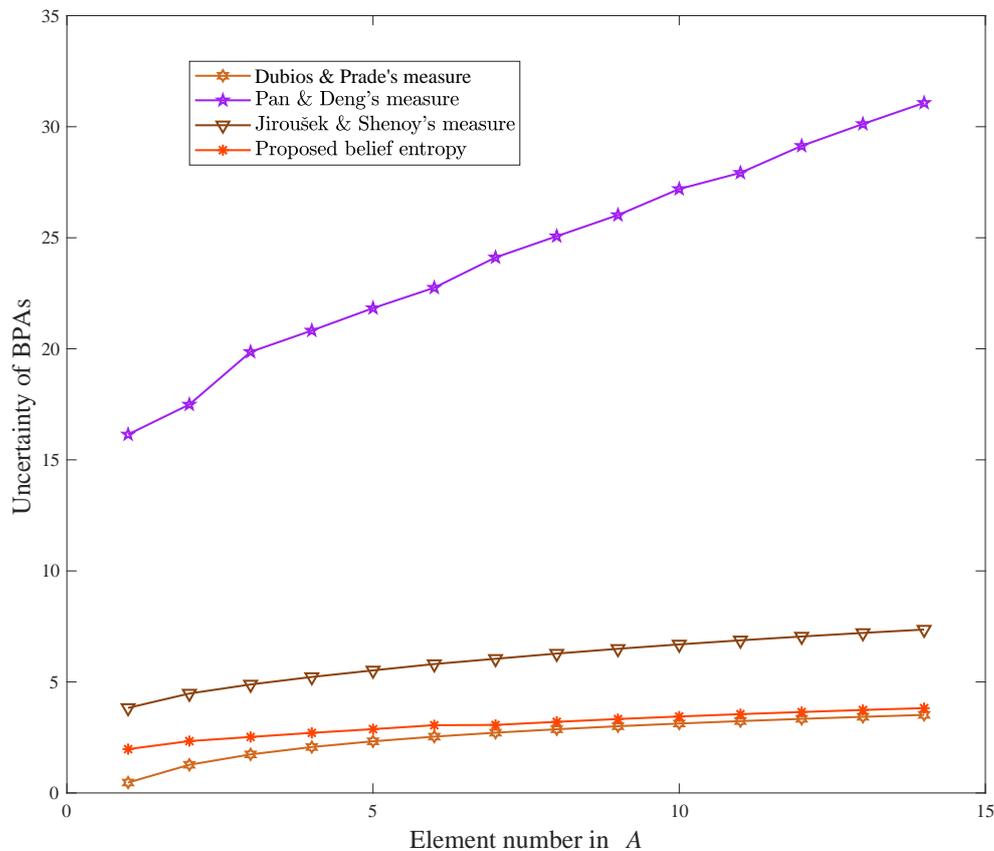


Figure 2. Results comparison of M_1 , M_7 , M_8 , and M_9 in DST.

As shown in Figure 1, it is obvious that the uncertain degree measured by the George and Pal's conflict measure is almost unchanged when the element number increase in proposition A . Similarly, the Höhle's confusion uncertainty measure and Yager's dissonance uncertainty measure have the same situation to reflect the variation on uncertain degree in this case. Thus, these three uncertainty measures cannot detect the change in proposition A . Although the uncertainty degrees obtained by the Klir and Ramer's discord uncertainty measure and Klir and Parviz's strife uncertainty measure change with the growth of element number in A , the variation trends of both methods are contrary to expectation that the uncertainty degree increases with the augment of the element number in A . These methods only measure the discord uncertainty of the BPAs, but ignore the non-specificity uncertainty of the BPAs. Besides, from the Table 2, we can find that the Yager's dissonance uncertainty measure has the minimum uncertainty degree. This is because this method uses the plausibility function to measure the discord uncertainty. The plausibility function contains all the support degree to the single events from other propositions, which could lead to information redundant and uncertainty reducing incorrectly. To sum up, the uncertainty degree obtained by George and Pal's method, Höhle's method, Yager's method, Klir and Ramer's method, and Klir and Parviz's method are unreasonable and counterintuitive, which means these methods cannot measure the uncertainty in this case aright.

From Figure 2, it can be seen that the uncertainty degrees measured by Dubois and Prade's weighted Hartley entropy, Pan and Deng's uncertainty measure, Jiroušek and Shenoy's uncertainty measure, and the proposed belief entropy are increasing visibly with the rising of the element number in A . These methods consider not only the discord uncertainty but also the non-specificity uncertainty. Furthermore, Pan and Deng's uncertainty measure is the largest among all the methods in Table 2. This is understandable. The non-specificity uncertainty measure in Pan and Deng's method is exponential, while the others are linear. As the number of elements in A increases, the uncertainty

degree of Pan and Deng's method increases faster than the other methods. Non-specificity uncertainty measure using exponential form may cause the possible uncertainty degree from the discord part to be significantly smaller than the ones from the non-specificity part. Additionally, Jiroušek and Shenoy's uncertainty measure is larger than the proposed belief entropy. Compared to the Jiroušek and Shenoy's uncertainty measure, which uses the probability distribution of single element obtained by plausibility transformation to measure the discord uncertainty, the proposed belief entropy measure that one by using the information of each mass function and the single element each BPA contains. The redundant information is removed, and the possible values of discord uncertainty is decreased notably in the proposed method. More importantly, except for the proposed method, the other three uncertainty measures have shortcomings. The Dubois and Prade's weighted Hartley entropy does not consider the discord uncertainty of BPAs. The Pan and Deng's uncertainty measure cannot measure accurately two similar BPAs in Section 5.5. The discord uncertainty measure of Jiroušek and Shenoy's uncertainty measure is irrational in Section 5.6. Thus, the proposed belief entropy is the only effective approach for uncertainty measure among these given methods in this case. Therefore, the proposed belief entropy, which considers the information contained in BPAs and single elements, is reasonable and effective for uncertainty measure in Dempster-Shafer framework.

6. Conclusions

How to measure the uncertainty of BPA in the framework of DST is an open issue. In this study, the main contribution is that a new belief entropy is proposed to quantify the uncertainty of BPA. The proposed belief entropy is comprised of the discord uncertainty measurement and the non-specificity uncertainty measurement. In particular, in the discord uncertainty measure component, the idea of probability interval and conversion BPA to probability using the plausibility transformation are combined. The new method takes advantage of the information of not only the BPAs, but also the total support degree of the single events contained in the BPAs. By addressing appropriate information in a BPA, which means less information loss and less information redundancy, the proposed belief entropy could measure the uncertainty of BPA efficiently. In addition, the proposed belief entropy could satisfy six desired properties of consistency with DST semantics, non-negativity, set consistency, probability consistency, additivity, and monotonicity. The results of numerical experiments demonstrate that the proposed belief entropy can be more effective and accurate when compared to the existing uncertainty measures in the framework of DST. Future work of this study will be focused on extending the proposed method to open-world assumptions and applying it to solve problems in real applications.

Author Contributions: Conceptualization, Q.P. and D.Z.; Data curation, Q.P. and J.H.; Formal analysis, Q.P.; Methodology, Q.P.; Software, Q.P. and Y.T.; Validation, X.L.; Writing—original draft, Q.P.; Writing—review & editing, Q.P.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 61603299 and 61602385), Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (Grant CX201705).

Acknowledgments: The authors greatly appreciate the reviews' valuable suggestions and the editor's encouragement. The authors greatly appreciate Boya Wei, the research assistant in Shanxi Normal University, for her advice on manuscript writing.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*; Springer: Berlin, Germany, 2008; pp. 57–72.
2. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976; Volume 42.
3. Wang, J.; Qiao, K.; Zhang, Z. An improvement for combination rule in evidence theory. *Futur. Gener. Comput. Syst.* **2019**, *91*, 1–9. [[CrossRef](#)]

4. Jiao, Z.; Gong, H.; Wang, Y. A DS evidence theory-based relay protection system hidden failures detection method in smart grid. *IEEE Trans. Smart Grid* **2018**, *9*, 2118–2126. [[CrossRef](#)]
5. Jiang, W.; Zhan, J. A modified combination rule in generalized evidence theory. *Appl. Intell.* **2017**, *46*, 630–640. [[CrossRef](#)]
6. de Oliveira Silva, L.G.; de Almeida-Filho, A.T. A multicriteria approach for analysis of conflicts in evidence theory. *Inf. Sci.* **2016**, *346*, 275–285. [[CrossRef](#)]
7. Liu, Z.G.; Pan, Q.; Dezert, J.; Martin, A. Combination of classifiers with optimal weight based on evidential reasoning. *IEEE Trans. Fuzzy Syst.* **2018**, *26*, 1217–1230. [[CrossRef](#)]
8. Mi, J.; Li, Y.F.; Peng, W.; Huang, H.Z. Reliability analysis of complex multi-state system with common cause failure based on evidential networks. *Reliab. Eng. Syst. Saf.* **2018**, *174*, 71–81. [[CrossRef](#)]
9. Zhao, F.J.; Zhou, Z.J.; Hu, C.H.; Chang, L.L.; Zhou, Z.G.; Li, G.L. A new evidential reasoning-based method for online safety assessment of complex systems. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *48*, 954–966. [[CrossRef](#)]
10. Xiao, F. Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy. *Inf. Fusion* **2019**, *46*, 23–32. [[CrossRef](#)]
11. Tang, Y.; Zhou, D.; Chan, F.T.S. An Extension to Deng's Entropy in the Open World Assumption with an Application in Sensor Data Fusion. *Sensors* **2018**, *18*, 1902. [[CrossRef](#)]
12. Zhu, J.; Wang, X.; Song, Y. Evaluating the Reliability Coefficient of a Sensor Based on the Training Data within the Framework of Evidence Theory. *IEEE Access* **2018**, *6*, 30592–30601. [[CrossRef](#)]
13. Xiao, F.; Qin, B. A Weighted Combination Method for Conflicting Evidence in Multi-Sensor Data Fusion. *Sensors* **2018**, *18*, 1487. [[CrossRef](#)] [[PubMed](#)]
14. Kang, B.; Chhipi-Shrestha, G.; Deng, Y.; Hewage, K.; Sadiq, R. Stable strategies analysis based on the utility of Z-number in the evolutionary games. *Appl. Math. Comput.* **2018**, *324*, 202–217. [[CrossRef](#)]
15. Zheng, X.; Deng, Y. Dependence assessment in human reliability analysis based on evidence credibility decay model and IOWA operator. *Ann. Nuclear Energy* **2018**, *112*, 673–684. [[CrossRef](#)]
16. Lin, Y.; Li, Y.; Yin, X.; Dou, Z. Multisensor Fault Diagnosis Modeling Based on the Evidence Theory. *IEEE Trans. Reliab.* **2018**, *67*, 513–521. [[CrossRef](#)]
17. Song, L.; Wang, H.; Chen, P. Step-by-step Fuzzy Diagnosis Method for Equipment Based on Symptom Extraction and Trivalent Logic Fuzzy Diagnosis Theory. *IEEE Trans. Fuzzy Syst.* **2018**, *26*, 3467–3478. [[CrossRef](#)]
18. Gong, Y.; Su, X.; Qian, H.; Yang, N. Research on fault diagnosis methods for the reactor coolant system of nuclear power plant based on DS evidence theory. *Ann. Nucl. Energy* **2018**, *112*, 395–399. [[CrossRef](#)]
19. Zheng, H.; Deng, Y. Evaluation method based on fuzzy relations between Dempster–Shafer belief structure. *Int. J. Intell. Syst.* **2018**, *33*, 1343–1363.
20. Song, Y.; Wang, X.; Zhu, J.; Lei, L. Sensor dynamic reliability evaluation based on evidence theory and intuitionistic fuzzy sets. *Appl. Intell.* **2018**, *48*, 3950–3962. [[CrossRef](#)]
21. Ruan, Z.; Li, C.; Wu, A.; Wang, Y. A New Risk Assessment Model for Underground Mine Water Inrush Based on AHP and D–S Evidence Theory. *Mine Water Environ.* **2019**, 1–9, doi:10.1007/s10230-018-00575-0. [[CrossRef](#)]
22. Ma, X.; Liu, S.; Hu, S.; Geng, P.; Liu, M.; Zhao, J. SAR image edge detection via sparse representation. *Soft Comput.* **2018**, *22*, 2507–2515. [[CrossRef](#)]
23. Moghaddam, H.A.; Ghodratinama, S. Toward semantic content-based image retrieval using Dempster–Shafer theory in multi-label classification framework. *Int. J. Multimed. Inf. Retr.* **2017**, *6*, 317–326. [[CrossRef](#)]
24. Torous, J.; Nicholas, J.; Larsen, M.E.; Firth, J.; Christensen, H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid. Ment. Health* **2018**, *21*, 116–119. [[CrossRef](#)] [[PubMed](#)]
25. Liu, T.; Deng, Y.; Chan, F. Evidential supplier selection based on DEMATEL and game theory. *Int. J. Fuzzy Syst.* **2018**, *20*, 1321–1333. [[CrossRef](#)]
26. Orient, G.; Babuska, V.; Lo, D.; Mersch, J.; Wapman, W. A Case Study for Integrating Comp/Sim Credibility and Convolved UQ and Evidence Theory Results to Support Risk Informed Decision Making. In *Model Validation and Uncertainty Quantification*; Springer: Berlin, Germany, 2019; Volume 3, pp. 203–208.
27. Li, Y.; Xiao, F. Bayesian Update with Information Quality under the Framework of Evidence Theory. *Entropy* **2019**, *21*, 5. [[CrossRef](#)]

28. Dietrich, C.F. *Uncertainty, Calibration and Probability: the Statistics of Scientific and Industrial Measurement*; Routledge: London, NY, USA, 2017.
29. Rényi, A. *On Measures of Entropy and Information*; Technical Report; Hungarian Academy of Sciences: Budapest, Hungary, 1961.
30. Shannon, C. A mathematical theory of communication. *ACM SIGMOBILE Mob. Comput. Commun. Rev.* **2001**, *5*, 3–55. [[CrossRef](#)]
31. Zhou, M.; Liu, X.B.; Yang, J.B.; Chen, Y.W.; Wu, J. Evidential reasoning approach with multiple kinds of attributes and entropy-based weight assignment. *Knowl. Syst.* **2019**, *163*, 358–375. [[CrossRef](#)]
32. Klir, G.J.; Wierman, M.J. *Uncertainty-Based Information: Elements of Generalized Information Theory*; Springer: Berlin, Germany, 2013; Volume 15.
33. Dubois, D.; Prade, H. Properties of measures of information in evidence and possibility theories. *Fuzzy Sets Syst.* **1987**, *24*, 161–182. [[CrossRef](#)]
34. Hohle, U. Entropy with respect to plausibility measures. In Proceedings of the 12th International Symposium on Multiple-Valued Logic, Paris, France, 25–27 May 1982.
35. Yager, R.R. Entropy and specificity in a mathematical theory of evidence. *Int. J. Gen. Syst.* **1983**, *9*, 249–260. [[CrossRef](#)]
36. Klir, G.J.; Ramer, A. Uncertainty in the Dempster-Shafer theory: a critical re-examination. *Int. J. Gen. Syst.* **1990**, *18*, 155–166. [[CrossRef](#)]
37. Klir, G.J.; Parviz, B. A note on the measure of discord. In *Uncertainty in Artificial Intelligence*; Elsevier: Amsterdam, The Netherlands, 1992; pp. 138–141.
38. Jousselme, A.L.; Liu, C.; Grenier, D.; Bossé, É. Measuring ambiguity in the evidence theory. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2006**, *36*, 890–903. [[CrossRef](#)]
39. Deng, Y. Deng entropy. *Chaos, Solitons & Fractals* **2016**, *91*, 549–553.
40. Pan, L.; Deng, Y. A New Belief Entropy to Measure Uncertainty of Basic Probability Assignments Based on Belief Function and Plausibility Function. *Entropy* **2018**, *20*, 842. [[CrossRef](#)]
41. Jiroušek, R.; Shenoy, P.P. A new definition of entropy of belief functions in the Dempster–Shafer theory. *Int. J. Approx. Reason.* **2018**, *92*, 49–65. [[CrossRef](#)]
42. Yang, Y.; Han, D. A new distance-based total uncertainty measure in the theory of belief functions. *Knowl. Syst.* **2016**, *94*, 114–123. [[CrossRef](#)]
43. Smets, P. Decision making in the TBM: the necessity of the pignistic transformation. *Int. J. Approx. Reason.* **2005**, *38*, 133–148. [[CrossRef](#)]
44. Cobb, B.R.; Shenoy, P.P. On the plausibility transformation method for translating belief function models to probability models. *Int. J. Approx. Reason.* **2006**, *41*, 314–330. [[CrossRef](#)]
45. Klir, G.J.; Lewis, H.W. Remarks on “Measuring ambiguity in the evidence theory”. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2008**, *38*, 995–999. [[CrossRef](#)]
46. Klir, G.J. *Uncertainty and Information: Foundations of Generalized Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
47. Abellán, J. Analyzing properties of Deng entropy in the theory of evidence. *Chaos Solitons Fractals* **2017**, *95*, 195–199. [[CrossRef](#)]
48. Smets, P. Decision making in a context where uncertainty is represented by belief functions. In *Belief Functions in Business Decisions*; Springer: Berlin, Germany, 2002; pp. 17–61.
49. Daniel, M. On transformations of belief functions to probabilities. *Int. J. Intell. Syst.* **2006**, *21*, 261–282. [[CrossRef](#)]
50. Cuzzolin, F. On the relative belief transform. *Int. J. Approx. Reason.* **2012**, *53*, 786–804. [[CrossRef](#)]
51. Shahpari, A.; Seyedin, S. A study on properties of Dempster-Shafer theory to probability theory transformations. *Iran. J. Electr. Electron. Eng.* **2015**, *11*, 87.
52. Jaynes, E.T. Where do we stand on maximum entropy? *Maximum Entropy Formalism* **1979**, *15*, 15–118.
53. Klir, G.J. Principles of uncertainty: What are they? Why do we need them? *Fuzzy Sets Syst.* **1995**, *74*, 15–31.
54. Ellsberg, D. Risk, ambiguity, and the Savage axioms. *Q. J. Econ.* **1961**, *75*, 643–669. [[CrossRef](#)]
55. Dubois, D.; Prade, H. Properties of measures of information in evidence and possibility theories. *Fuzzy Sets Syst.* **1999**, *100*, 35–49.
56. Abellan, J.; Moral, S. Completing a total uncertainty measure in the Dempster-Shafer theory. *Int. J. Gen. Syst.* **1999**, *28*, 299–314.

57. Li, Y.; Deng, Y. Generalized Ordered Propositions Fusion Based on Belief Entropy. *Int. J. Comput. Commun. Control* **2018**, *13*. [[CrossRef](#)]
58. Nguyen, H.T. On entropy of random sets and possibility distributions. *Anal. Fuzzy Inf.* **1987**, *1*, 145–156.
59. Pal, N.R.; Bezdek, J.C.; Hemasinha, R. Uncertainty measures for evidential reasoning II: A new measure of total uncertainty. *Int. J. Approx. Reason.* **1993**, *8*, 1–16. [[CrossRef](#)]
60. Zhou, D.; Tang, Y.; Jiang, W. An improved belief entropy and its application in decision-making. *Complexity* **2017**, *2017*. [[CrossRef](#)]
61. George, T.; Pal, N.R. Quantification of conflict in Dempster-Shafer framework: A new approach. *Int. J. Gen. Syst.* **1996**, *24*, 407–423. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).