

Article

A New Belief Entropy in Dempster–Shafer Theory Based on Basic Probability Assignment and the Frame of Discernment

Jiapeng Li ^{1,*}  and Qian Pan ^{2,*}¹ School of Automation, Northwestern Polytechnical University, Xi'an 710072, China² School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China

* Correspondence: ljpeng@mail.nwpu.edu.cn (J.L.); panq@mail.nwpu.edu.cn (Q.P.)

Received: 30 April 2020; Accepted: 10 June 2020; Published: 20 June 2020



Abstract: Dempster–Shafer theory has been widely used in many applications, especially in the measurement of information uncertainty. However, under the D-S theory, how to use the belief entropy to measure the uncertainty is still an open issue. In this paper, we list some significant properties. The main contribution of this paper is to propose a new entropy, for which some properties are discussed. Our new model has two components. The first is Nguyen entropy. The second component is the product of the cardinality of the frame of discernment (FOD) and Dubois entropy. In addition, under certain conditions, the new belief entropy can be transformed into Shannon entropy. Compared with the others, the new entropy considers the impact of FOD. Through some numerical examples and simulation, the proposed belief entropy is proven to be able to measure uncertainty accurately.

Keywords: Dempster–Shafer theory; basic probability assignment; frame of discernment; uncertainty measure; belief entropy; Shannon entropy

1. Introduction

How to measure uncertainty is a meaningful question to be solved. Our work will also discuss this issue. First of all, we need to know what uncertainty is.

The uncertainty problem is still very extensive. It means not certainly known, questionable, and problematic. Uncertainty can mainly divided into three types: vagueness, which is boundary uncertainty; nonspecificity, which is size (cardinality) uncertainty; and discord, which expresses conflict. Correspondingly, there are some theories to solve these problems: fuzzy set theory [1], probability theory [2], evidence theory [3,4], and rough sets [5]. Besides, some extended theories are also presented for the uncertainty measure, e.g., generalized evidence theory [6], complex numbers [7], fuzzy numbers [8–10], Z numbers [11,12], D numbers theory [13–16], and so on [17–22]. In this paper, we use evidence theory to study these open issues.

In 1960, Dempster [3] proposed upper and lower probabilities to solve the multivalued mapping problem. In 1971, Shafer [4] completed the theory proposed by Dempster and formed evidence theory, also called D-S theory. After years of exploration, D-S theory is a very effective tool for modeling and processing information uncertainty. In 1948, Shannon [23] used the concepts in thermodynamics to define information entropy. Under probability theory, Shannon entropy was very good at measuring the degree of information uncertainty. However, D-S theory is easier than probability theory for getting prior data, and the former has the advantage of fused data. Thus, we introduce D-S theory to replace probability theory in uncertainty.

D-S theory uses basic probability assignment (BPA), which is under the frame of discernment (FOD), to represent the degree of support for a focal element. Different FODs may have different BPAs.

Besides, the core of D-S theory is Dempster's combination rule. Dempster's combination rule provides a way to fuse different BPAs. The proposition of evidence theory provides mathematical support for the establishment of uncertain models.

On the other hand, it is well known in information theory that Hartley [24] and Shannon's measures are both effective ways to deal with information uncertainty. Meanwhile, D-S theory as an extension of probability theory contains much ignorance information. Thus, Höhle [25] was the earliest to combine D-S theory and Shannon entropy, namely Höhle entropy. Subsequently, Nguyen [26], Dubois and Prade [27], Klir [28], Jiroušek and Shenoy [29], Nikhil R. Pal [30,31], Deng [32], Pan and Deng [33], and Wang [34] defined their uncertainty models. Some of them have been successfully applied in real situations [35,36]. However, these models are not effective in some places [37,38].

According to their entropy, most of them are only focused on the BPA of every focal element and the cardinality of an element or use the belief function and plausibility function to measure uncertainty. Therefore, no one focused on FOD. Obviously, the scale of FOD can impact the degree of uncertainty. We combined different models and proposed a new uncertainty measurement, namely B&F entropy, because the uncertainty is determined by both BPA and FOD. In this method, it can well reflect the impact of FOD on uncertainty. In the end, we will give a few examples to compare the new model and others. Besides, we design a simulation to illustrate the feasibility and effectiveness of the proposed model.

The outline of the remainder of the paper is as follows. In Section 2, we briefly review the Hartley and Shannon measures and D-S theory. Some essential properties are briefly introduced in Section 3. Section 4 presents and existing uncertainty measures. In Section 5, we discuss some properties and define a new entropy. In Section 6, some significant numerical examples and simulations are carried out to illustrate the feasibility and effectiveness of the proposed belief entropy. In Section 7, we summarize our findings and conclude with some open questions.

2. Preliminaries

The focus of this paper is based on D-S theory and information entropy. We divide this section into two parts. In the D-S theory section, some basic concepts will briefly be introduced. In the information entropy section, we will introduce two typical representatives, the Hartley measure and Shannon entropy.

2.1. D-S Theory

Dempster–Shafer theory, also called evidence reasoning or evidence theory, originated from Dempster [3] and was developed by his student Shafer [4]. Through a series of improvements and reinforcements, a method of uncertainty reasoning using “evidence” and “combination” was formed. In a way, D-S theory is a promotion of Bayesian reasoning. Dempster–Shafer theory is often applied to pattern recognition [39–44], fault diagnosis [45–48], uncertainty modeling [20,49], clustering [50], decision making [51,52], risk analysis [53–56], and other hot fields [57,58].

The idea of D-S theory is based on the frame of discernment X , which $X = \{x_1, x_2 \cdots x_n\}$, and the set of all subsets in X is called the power set 2^X . In the power set 2^X , it contains $2^{|X|}$ elements. $|X|$ means the cardinality of X , which is the number of elements in X . Under this frame, Dempster and Shafer defined some basic concepts as follows.

2.1.1. Basic Belief Assignment

Based on the above power set 2^X , function $m : 2^X \rightarrow [0, 1]$ satisfies:

$$\begin{cases} \sum_{a \in 2^X} m(a) = 1 \\ m(\emptyset) = 0 \end{cases} \quad (1)$$

The function $m(a)$ is also called a basic probability assignment or mass function. If $m(a) > 0$, then a is a focal element. $m(a)$ means the value of trust that the object belongs to a . The larger $m(a)$ is, the higher the trust value is.

Some definitions of BPA are as follows. The vacuous BPA means entirely unknown for the true result. In contrast, from the Bayesian BPA, we can know to which category the target should belong.

2.1.2. Belief Function

The belief function is the sum of the basic probability assignments for all subsets of a and is given by:

$$Bel(a) = \sum_{b \subseteq a} m(b), \forall a \in 2^X \quad (2)$$

It is the lower limit of support for a .

2.1.3. Plausibility Function

The plausibility function is the sum of the basic probability assignments for all subsets that intersect with a and is given by:

$$Pl(a) = \sum_{b \cap a \neq \emptyset} m(b), \forall a \in 2^X \quad (3)$$

It is the upper limit of support for a .

The value, between the belief function and plausibility function, is the degree of uncertainty for evidence.

2.1.4. Dempster's Combination Rule

Dempster's combination rule is the most commonly used method in evidence fusion. This rule takes into account the degree of conflict between the evidence and defined conflict coefficient k to measure the degree of conflict among different evidence.

Suppose m_1 and m_2 are independent BPAs from the different evidence resources, respectively. The fusion result of m_1 and m_2 under Dempster's combination is as follows:

$$\begin{cases} m(a) = 0, a = \emptyset \\ m(a) = \frac{\sum_{b \cap c = a} m_1(b)m_2(c)}{1 - k}, a \neq \emptyset \end{cases} \quad (4)$$

where k is a conflict coefficient, defined by:

$$k = \sum_{b \cap c = \emptyset} m_1(b)m_2(c) \quad (5)$$

Notice that Dempster's combination rule is invalid, if two bodies of evidence completely conflict ($k = 1$). Furthermore, if $k > 1$, Dempster's combination rule cannot be applied to the two BPAs' fusion.

2.2. Origin of Information Entropy

Different authors have measured information uncertainty in a variety of ways, and Hartley and Shannon laid the foundation for it. The information entropy and their extended models have been applied to many fields [59]. Next, we will briefly introduce the Hartley measure and Shannon entropy.

2.2.1. Hartley Measure

Suppose X is an FOD and a is a subset of X . Then, the Hartley measure [24] is defined as:

$$H(a) = \log_2 |a| \quad (6)$$

where $|a|$ means the cardinality of a .

Obviously, the measurement is proportional to the cardinality of a . When a is a singleton of X , $H(a) = 0$, this means there is no conflict. Unfortunately, the measurement method of Hartley does not show the effect of the probability distribution on the degree of uncertainty.

2.2.2. Shannon Entropy

In 1948, Shannon [23] proposed information entropy, namely Shannon entropy. His model uses the concept of entropy from thermodynamics.

$$H(x) = \sum_{x \in X} P(x) \log_2 \frac{1}{P(x)} \quad (7)$$

where $P(x)$ is the probability of x and $P(x)$ satisfies $\sum_{x \in X} P(x) = 1$.

As he said in his thesis, the role of information is used to eliminate the uncertainty. Shannon entropy is an excellent way to measure and eliminate uncertainty. It played a crucial role in solving the probability problem. We can conclude from his definition that it is based on the probability distribution. With the emergence of D-S theory, the information entropy was given a new meaning. The format of our new model is also derived from the Shannon entropy.

3. Properties of the Uncertainty Measure in D-S Theory

According to Klir and Wierman [60] and Klir and Folger [61], we introduce some important properties of entropy for D-S theory, including non-negativity, maximum, monotonicity, probability consistency, additivity, sub-additivity, and range. These properties for a measure that captures both discord and non-specificity are defined as follows.

3.1. Non-Negativity

Suppose m is a BPA on FOD X ; the entropy $H(m)$ must be:

$$H(m) > 0 \quad (8)$$

where this is equality if and only if $m(x) = 1$ and $x \in X$.

Only when entropy satisfies the non-negativity property, it provides a standard for measurement uncertainty.

3.2. Maximum Entropy

It makes sense that the vacuous BPA m_v for uncertainty is larger than other normal BPAs m_n . Thus, the maximum entropy property is defined as:

$$H(m_v) > H(m_n) \quad (9)$$

3.3. Monotonicity

As the number of focal elements in FOD increases, so should the degree of uncertainty. The monotonicity property is defined as:

$$H(m_X) > H(m_Y) \quad (10)$$

where m_X and m_Y are the vacuous BPAs for FOD X and FOD Y . Meanwhile, $|X| > |Y|$.

3.4. Probability Consistency

Let m_B be a Bayesian BPA, and then, the entropy should be the same as Shannon entropy. Therefore, the probability consistency property follows as:

$$H(m_B) = H_S(P_X) = \sum_{x \in X} P_X(x) \log_2 \frac{1}{P_X(x)} \quad (11)$$

where H_S is the Shannon entropy and P_X is the BPA of X corresponding to m_b .

3.5. Additivity

Let m_X and m_Y be independent BPAs on FOD X and FOD Y , respectively. \oplus means Dempster's combination rule. Thus, the additivity property is defined as:

$$H(m_X \oplus m_Y) = H(m_X) + H(m_Y) \quad (12)$$

where $m_X \oplus m_Y$ is a BPA for FOD $\{X, Y\}$. Note that $m(a \times b) = m_X(a)m_Y(b)$, where m is the m_X and m_Y combined by Dempster's combination rule.

3.6. Sub-Additivity

Let m be a BPA on FOD $\{X \times Y\}$. Let $m^{\downarrow X}$ and $m^{\downarrow Y}$ be the marginal BPAs of FOD X and FOD Y . Then, define:

$$H(m) \leq H(m^{\downarrow X}) + H(m^{\downarrow Y}) \quad (13)$$

3.7. Range

As Klir and Wierman defined, the range of $H(m)$ is $[0, \log_2 |X|]$.

4. The Development of Entropy Based on D-S Theory

In this section, some belief entropies of BPAs in D-S theory proposed by others are reviewed. We also discuss whether or not these models satisfy the properties we list.

Yager [62] defined the belief entropy using the conflict coefficient between two focal elements, simplified as follows:

$$H_Y(m) = - \sum_{a \in 2^X} m(a) \log_2 Pl(a) \quad (14)$$

where $Pl(a)$ is the plausibility function associated with a under m . The entropy of Yager only measures the degree of conflict between evidence. $H_Y(m)$ only satisfies the additivity property.

Dubois [27] used a new information measurement method to get the new formula of entropy.

$$H_D(m) = \sum_{a \in 2^X} m(a) \log_2 |a| \quad (15)$$

From the definition of Dubois, this entropy only answers the question of the non-specific part of the uncertainty. If m is a Bayesian BPA, then $H_D(m) = 0$. It is noticeable that $H_D(m)$ is clearly a weighted Hartley [24] measure. $H_D(m)$ satisfies the maximum entropy and monotonicity properties.

Nguyen [26] defined a new entropy according to Shannon entropy.

$$H_N(m) = \sum_{a \in 2^X} m(a) \log_2 \frac{1}{m(a)} \quad (16)$$

From the definition format, it only uses the BPA to capture the part of the conflict. This is inaccurate for uncertain measurements. It only satisfies the probabilistic consistency property and the additivity property.

Lamata and Moral [63] used the entropy theory proposed by Yager and Dubois.

$$\begin{aligned} H_{L\&M}(m) &= H_Y(m) + H_D(m) \\ &= \sum_{a \in 2^X} m(a) \log_2 \frac{|a|}{Pl(a)} \end{aligned} \quad (17)$$

They both have two components: one measures the innate contradiction, while the other measures the imprecision of the information. This definition does not satisfy the maximum entropy and sub-additivity properties.

Jiroušek and Shenoy [29] entropy is a combination of the Shannon and Dubois definitions.

$$\begin{aligned} H_{J\&S}(m) &= H_S(Pl_P_m) + H_D(m) \\ &= \sum_{x \in X} Pl_P_m \log_2 \frac{1}{Pl_P_m} + \sum_{a \in 2^X} m(a) \log_2 |a| \end{aligned} \quad (18)$$

where Pl_P_m is the normalized result of plausibility function Pl_m . The first part is the measurement of conflict based on Shannon entropy, and the second part is to measure the non-specificity portion of uncertainty. The entropy of $H_{J\&S}(m)$ satisfies non-negativity, maximum entropy, monotonicity, probability consistency, and additivity.

Klir and Ramer [28] defined:

$$H_{K\&R}(m) = - \sum_{a \in 2^X} m(a) \log_2 \sum_{b \in 2^X} m(b) \frac{|b - a|}{|b|} \quad (19)$$

Due to the Yager entropy not concluded the broader view of conflict (it only considered the conflict situation of $B \cap A = \emptyset$), Klir and Ramer proposed a new method to solve this problem. It is easy to see that this entropy can measure the conflict of evidential claims within each body of evidence in bits. However, under certain conditions, it is difficult for $H_{K\&R}(m)$ to express the aspect of uncertainty. It just does not satisfy the maximum entropy property.

Nikhil R. Pal [30,31] focused on nonspecificity and randomness under a total uncertainty environment.

$$\begin{aligned} H_P(m) &= \sum_{a \in 2^X} m(a) \log_2 \frac{|a|}{m(a)} \\ &= \sum_{a \in 2^X} m(a) \log_2 \frac{1}{m(a)} + \sum_{a \in 2^X} m(a) \log_2 |a| \end{aligned} \quad (20)$$

They summed up the methods proposed by Lamata and Moral and Klir and Ramer. It was pointed out that there would be mistakes against common sense in certain situations. The first part is, in some sense, analogous to Yager's entropy, and the second part measures the conflict of the body of evidence. It does not satisfy the maximum entropy property.

Jousselme [64] entropy is based on pignistic transformation $BetP_m(a)$ [65].

$$H_J(m) = \sum_{a \in 2^X} BetP_m(a) \log_2 \frac{1}{BetP_m(a)} \quad (21)$$

He finally proved that as the evidence changes, the entropy becomes more sensitive.

Deng [32] defined an entropy:

$$H_{Deng}(m) = \sum_{a \in 2^X} m(a) \log_2 \left[\frac{2^{|a|} - 1}{m(a)} \right] \quad (22)$$

As proven by Joaquín Abellán [66], the Deng entropy does not satisfy the monotonicity, additivity, and subadditivity properties.

Pan and Deng [33] developed Deng entropy and defined it as follows:

$$H_{Bel}(m) = - \sum_{a \in 2^X} \frac{Bel(a) + Pl(a)}{2} \log_2 \frac{Bel(a) + Pl(a)}{2(2^{|a|} - 1)} \quad (23)$$

where $Bel(a)$ and $Pl(a)$ are the belief function and plausibility function, respectively. $H_{Bel}(m)$ uses the interval probability to measure the discord and non-specificity uncertainty of BPA. It does not satisfy the maximum entropy, additivity, sub-additivity, and range properties.

W [34] is another modified model based on Deng entropy:

$$H_W(m) = - \sum_{a \in X} m(a) \log_2 \left(\frac{m(a)}{2^{|a|} - 1} (1 + \epsilon)^{f(|X|)} \right) \quad (24)$$

where ϵ is a constant and $\epsilon \geq 0$, $f(|X|)$ is the function about the cardinality of X . ϵ is a change number; it can take different values to represent different entropies. However, as the parameter ϵ changes, it has little effect on the value of W entropy [34].

5. A New Belief Entropy Based on Evidence Theory

As introduced at the start of the first chapter of Shafer's book [4], D-S theory is a theory of evidence. That means using the mathematical form to express the degree of support for evidence.

Based on the entropy proposed by previous scholars, for the measurement method of information uncertainty, there remain several aspects of the frame of discernment about which relatively little is known. In D-S theory, if we have the same cardinality of BPA, but different FODs, the results of uncertainty should be changed. However, most of them we listed above only focused on the value of BPA or the cardinality of every BPA, and the effect of FOD was totally ignored. Thus, these definitions cannot measure the degree of uncertainty under different FOD. To improve these deficiencies, we suggest that the FOD is also important for the measurement of uncertainty. Therefore, we introduce the scale of FOD to our new entropy. The new belief entropy based on D-S theory, namely B&F entropy, is defined as follows:

$$H_{B\&F}(m) = \sum_{a \in 2^X} m(a) \log_2 \frac{|a|^{|X|}}{m(a)} \quad (25)$$

where $|a|$ denotes the cardinality of the focal element a and $|X|$ equals the number of elements in FOD.

Like some of the definitions we mentioned, the new definition can be represented by a combination of other entropies. Thus, the new entropy also can be expressed as:

$$\begin{aligned} H_{B\&F}(m) &= \sum_{a \in 2^X} m(a) \log_2 \frac{1}{m(a)} + \sum_{a \in 2^X} m(a) \log_2 |a|^{|X|} \\ &= H_N(m) + |X| H_D(m) \end{aligned} \quad (26)$$

where $H_N(m)$ is Nguyen's entropy and $H_D(m)$ is Dubois' entropy. Obviously, the new entropy is a combination of $H_N(m)$ and $|X|$ times $H_D(m)$. Similar to most of the belief entropies, the first component $\sum_{a \in 2^X} m(a) \log_2 m^{-1}(a)$ in the new belief entropy is designed to measure the discord uncertainty of BPA. At length, the second component $\sum_{a \in 2^X} m(a) \log_2 |a|^{|X|}$ is the measure of non-specificity of the mass function among various focal elements [27,32,61]. In addition, it can capture the information about the size of cardinality. When m is a Bayesian BPA or the cardinality of FOD equals one, the new entropy degenerates to Pal's definition.

The most important information about FOD is the quantity of the focal element, namely $|X|$. If $|X|$ is modified, the accuracy of uncertain measurement will be affected. Here, we use an example to show that $|X|$ is the best way to represent the information of FOD.

As shown in Figure 1, it is obvious that $\log_2|X|$ and $2^{|X|}$ cannot reflect the effect of FOD on entropy very well. When the cardinality of FOD is greater than 10, $\log_2|X|$ is almost constant, but $2^{|X|}$ is very large. Thus, $|X|$ can well contain the information of the FOD size.

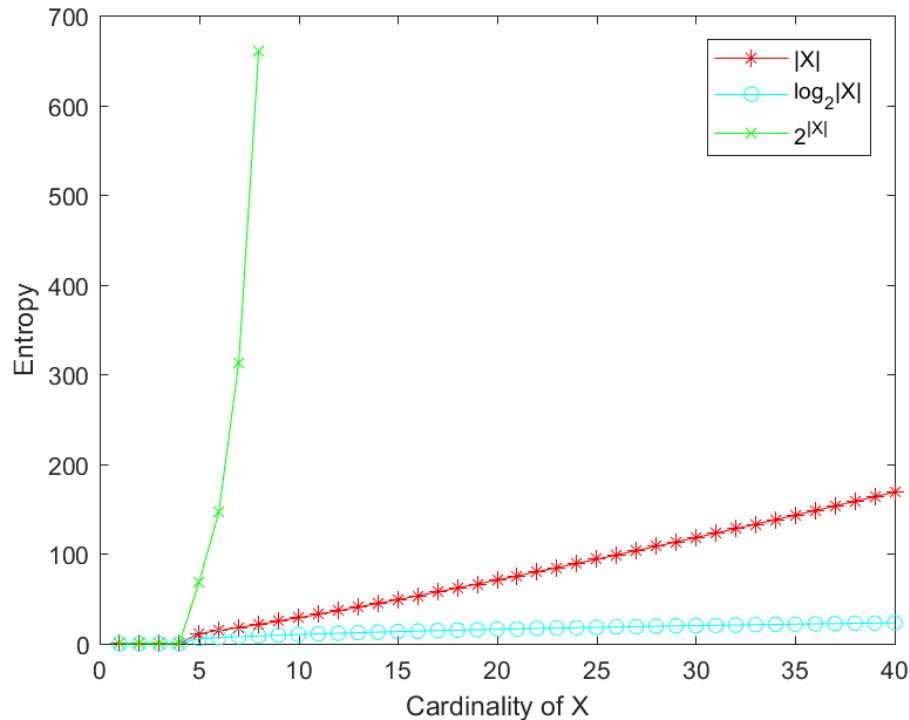


Figure 1. Comparison of different frame of discernment (FOD) information.

The new entropy connects the degree of information uncertainty and the FOD, meanwhile improving the information uncertainty measurement method.

According to Section 3, the basic properties of the new belief entropy are proven as follows:

(P1) Non-negativity:

Let $|a|$ be a cardinality of the focal element and $|X|$ be a cardinality of FOD. It is obvious that $|a|^{|X|} > 1$; thus, $H_{B\&F}(m) \geq 0$, if and only if m is the Bayesian BPA and $m = 1$. Therefore, the new definition satisfies the non-negativity property.

(P2) Maximum entropy:

Let m_b be a Bayesian BPA and m_v be a vacuous BPA, then $H_{B\&F}(m_b) = \log_2|X|$, $H_{B\&F}(m_v) = |X|\log_2|X|$. Although, according to our calculations, $H_{B\&F}(m_v) > H_{B\&F}(m_b)$, it does not mean $H_{B\&F}(m_v)$ is the maximum value. Later, we will further explain the max value through simulation. In this part, we just give some simple explanations.

From the definition of Nguyen we introduced, this entropy does not satisfy the maximum entropy, as it consists of Nguyen's entropy and Dubois entropy. Thus, the maximum entropy is not satisfied with the new belief entropy.

(P3) Monotonicity:

We suppose that m_v denotes the vacuous BPA, then $H_{B\&F}(m_v) = |X|\log_2|X|$. Obviously, $H_{B\&F}(m_v)$ increases with $|X|$. Therefore, $H_{B\&F}(m)$ satisfies the monotonicity property.

(P4) Probability consistency:

When m_b is a Bayesian BPA, then $|a| = 1$, $H_{B\&F}(m) = \sum_{a \in 2^X} m(a) \log_2 \frac{1}{m(a)}$. From this result, we conclude that the new belief entropy satisfies the probability consistency property.

(P5) Additivity and sub-additivity:

Let $c = a \times b \in 2^{X \times Y}$, where a, b, c is a focal element and X, Y means the FOD. Meanwhile, $a \in 2^X$ and $b \in 2^Y$. According to the definition of the above properties, $m(c) = m^{\downarrow X}(a) \times m^{\downarrow Y}(b)$, where $m^{\downarrow X}$ is the marginal BPA for X and $m^{\downarrow Y}$ is the marginal BPA for Y .

$$\begin{aligned}
 H_{B\&F}(c) &= \sum_{c \in 2^{X \times Y}} m(c) \log_2 \frac{|c|^{|X||Y|}}{m(c)} \\
 &= \sum_{\substack{a \in 2^X \\ b \in 2^Y}} m^{\downarrow X}(a) m^{\downarrow Y}(b) \log_2 (|a| |b|)^{|X||Y|} - \sum_{\substack{a \in 2^X \\ b \in 2^Y}} m^{\downarrow X}(a) m^{\downarrow Y}(b) \log_2 (m^{\downarrow X}(a) m^{\downarrow Y}(b)) \\
 &= \sum_{\substack{a \in 2^X \\ b \in 2^Y}} m^{\downarrow X}(a) m^{\downarrow Y}(b) \log_2 |a|^{|X||Y|} + \sum_{\substack{a \in 2^X \\ b \in 2^Y}} m^{\downarrow X}(a) m^{\downarrow Y}(b) \log_2 |b|^{|X||Y|} \\
 &\quad - \sum_{\substack{a \in 2^X \\ b \in 2^Y}} m^{\downarrow X}(a) m^{\downarrow Y}(b) \log_2 m^{\downarrow X}(a) - \sum_{\substack{a \in 2^X \\ b \in 2^Y}} m^{\downarrow X}(a) m^{\downarrow Y}(b) \log_2 m^{\downarrow Y}(b) \\
 &= \sum_{b \in 2^Y} m^{\downarrow Y}(b) \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 |a|^{|X||Y|} + \sum_{a \in 2^X} m^{\downarrow X}(a) \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 |b|^{|X||Y|} \\
 &\quad - \sum_{b \in 2^Y} m^{\downarrow Y}(b) \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 m^{\downarrow X}(a) - \sum_{a \in 2^X} m^{\downarrow X}(a) \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 m^{\downarrow Y}(b) \\
 &= \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 |a|^{|X||Y|} + \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 |b|^{|X||Y|} \\
 &\quad - \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 m^{\downarrow X}(a) - \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 m^{\downarrow Y}(b) \\
 &= \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 \frac{|a|^{|X||Y|}}{m^{\downarrow X}(a)} + \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 \frac{|b|^{|X||Y|}}{m^{\downarrow Y}(b)} \\
 &\geq \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 \frac{|a|^{|X|}}{m^{\downarrow X}(a)} + \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 \frac{|b|^{|Y|}}{m^{\downarrow Y}(b)} \\
 &= H_{B\&F}(m^{\downarrow X}(a)) + H_{B\&F}(m^{\downarrow Y}(b))
 \end{aligned}$$

We can see from the above proof that the new entropy satisfies the additivity property, if and only if $|X| = |Y| = 1$. Otherwise, the new belief entropy neither satisfies the additivity property nor sub-additivity.

To be more intuitive, we consider the following example:

Let Z be the product of FOD $X = \{x_1, x_2\}$ and FOD $Y = \{y_1, y_2, y_3\}$. We have that BPA on Z is m , and the marginal BPAs on X and Y are $m^{\downarrow X}$ and $m^{\downarrow Y}$. We suppose the case on Z is shown as follows:

$$\begin{aligned}
 m(\{z_{11}\}) &= 0.1, m(\{z_{13}\}) = 0.1 \\
 m(\{z_{12}\}) &= 0.1, m(\{z_{21}\}) = 0.1 \\
 m(\{z_{22}\}) &= 0.2, m(\{z_{23}\}) = 0.2 \\
 m(Z) &= 1 - m(\{z_{11}\}) - m(\{z_{13}\}) - m(\{z_{12}\}) - m(\{z_{21}\}) - m(\{z_{22}\}) - m(\{z_{23}\}) = 0.2
 \end{aligned}$$

where $Z_{i,j} = \{x_i, y_j\}$. Thus, the BPAs on X and Y are:

$$\begin{aligned}
 m^{\downarrow X}(\{x_1\}) &= 0.3, m^{\downarrow X}(\{x_2\}) = 0.5 \\
 m^{\downarrow X}(X) &= 1 - m^{\downarrow X}(\{x_1\}) - m^{\downarrow X}(\{x_2\}) = 0.2 \\
 m^{\downarrow Y}(\{y_1\}) &= 0.2, m^{\downarrow Y}(\{y_2\}) = 0.3, m^{\downarrow Y}(\{y_3\}) = 0.3 \\
 m^{\downarrow Y}(Y) &= 1 - m^{\downarrow Y}(\{y_1\}) - m^{\downarrow Y}(\{y_2\}) - m^{\downarrow Y}(\{y_3\}) = 0.2
 \end{aligned}$$

The calculation results are as follows:

$$\begin{aligned}
 |Z| &= 6, |X| = 2, |Y| = 3 \\
 H_{B\&F}(m) &= \sum_{z \in 2^{X \times Y}} m(z) \log_2 \frac{|z|^{|X||Y|}}{m(z)} \\
 &= 2 \times 0.1 \times \log_2 \frac{1^6}{0.1} + 2 \times 0.1 \times \log_2 \frac{1^6}{0.1} + 2 \times 0.2 \times \log_2 \frac{1^6}{0.2} + 0.2 \times \log_2 \frac{6^6}{0.2} \\
 &= 0.6644 + 0.6644 + 0.9288 + 3.5663 \\
 &= 5.8239 \\
 H_{B\&F}(m^{\downarrow X}) &= \sum_{a \in 2^X} m^{\downarrow X}(a) \log_2 \frac{|a|^{|X|}}{m^{\downarrow X}(a)} \\
 &= 0.3 \times \log_2 \frac{1^2}{0.3} + 0.5 \times \log_2 \frac{1^2}{0.5} + 0.2 \times \log_2 \frac{2^2}{0.2} \\
 &= 0.5211 + 0.5 + 0.8644 \\
 &= 1.8855 \\
 H_{B\&F}(m^{\downarrow Y}) &= \sum_{b \in 2^Y} m^{\downarrow Y}(b) \log_2 \frac{|b|^{|Y|}}{m^{\downarrow Y}(b)} \\
 &= 0.2 \times \log_2 \frac{1^3}{0.2} + 0.3 \times \log_2 \frac{1^3}{0.3} + 0.3 \times \log_2 \frac{1^3}{0.3} \\
 &\quad + 0.2 \times \log_2 \frac{3^3}{0.2} \\
 &= 0.4644 + 0.5211 + 0.5211 + 1.4154 \\
 &= 2.9220 \\
 H_{B\&F}(m^{\downarrow X}) + H_{B\&F}(m^{\downarrow Y}) &= 1.8855 + 2.9220 = 4.8075
 \end{aligned}$$

Obviously, $H_{B\&F}(m) > H_{B\&F}(m^{\downarrow X}) + H_{B\&F}(m^{\downarrow Y})$. Therefore, the additivity and sub-additivity properties are not satisfied with the new entropy.

(P6) Range:

As demonstrated by the maximum entropy property, the value of the new entropy $H_{B\&F}(m_v) = |X| \log_2 |X|$, and $H_{B\&F}(m_v) > \log_2 |X|$. Thus, it does not satisfy the range property.

From the above results we proved, the new belief entropy satisfies the non-negativity, monotonicity, and probability consistency properties, and does not satisfy the maximum entropy, additivity, subadditivity, and range properties.

6. Numerical Example and Simulation

In the first part of this section, some examples are given to illustrate the effectiveness of the new belief entropy. The influence of different BPAs on B&F entropy is shown in the second section.

6.1. Numerical Example

6.1.1. Example 1

Let FOD $X = \{a\}$, and we get a BPA from the sensor as $m(\{a\}) = 1$. Shannon entropy and the new definition proposed by the authors' calculation results are as follows:

$$H_{B\&F}(m) = H_S(m) = 1 \times \log_2 1 = 0$$

6.1.2. Example 2

Suppose there are three FODs $X_1 = \{x_1, x_2\}$, $X_2 = \{x_1, x_2, x_3, x_4\}$, $X_3 = \{x_1, x_2, x_3, x_4, x_5\}$. Every Bayesian BPA m_1, m_2, m_3 of these FODs is equal. Their BPAs are as follows:

$$m_1(\{x_1\}) = m_1(\{x_2\}) = 0.5$$

$$m_2(\{x_1\}) = m_2(\{x_2\}) = m_2(\{x_3\}) = m_2(\{x_4\}) = 0.25$$

$$m_3(\{x_1\}) = m_3(\{x_2\}) = m_3(\{x_3\}) = m_3(\{x_4\}) = m_3(\{x_5\}) = 0.2$$

The new belief entropy is calculated as follows:

$$H_{B\&F}(m_1) = 2 \times 0.5 \times \log_2 \frac{1^2}{0.5} = 1$$

$$H_{B\&F}(m_2) = 4 \times 0.25 \times \log_2 \frac{1^4}{0.25} = 2$$

$$H_{B\&F}(m_3) = 5 \times 0.2 \times \log_2 \frac{1^5}{0.2} = 2.3219$$

It is obvious that uncertainty increases as the number of focal elements increases. This is reasonable.

6.1.3. Example 3

Using the FOD raised by Example 2 and the vacuous BPAs $m_1(\{x_1, x_2\}) = 1$, $m_2(\{x_1, x_2, x_3, x_4\}) = 1$, $m_3(\{x_1, x_2, x_3, x_4, x_5\}) = 1$, the new entropy results are calculated as follows:

$$H_{B\&F}(m_1) = 1 \times \log_2 \frac{2^2}{1} = 2$$

$$H_{B\&F}(m_2) = 1 \times \log_2 \frac{4^4}{1} = 8$$

$$H_{B\&F}(m_3) = 1 \times \log_2 \frac{5^5}{1} = 11.6096$$

Comparing Example 2 and Example 3, it is easy to get that the results of the vacuous BPA are bigger than the results of the Bayesian BPA.

6.1.4. Example 4

In this example, we compare the difference between Pal entropy and B&F entropy. Let FOD $X_1 = \{x_1, x_2, x_3, x_4, x_5\}$ and $X_2 = \{x_1, x_2, x_3, x_4\}$. Meanwhile, suppose the following two situations exist:

$$C_1 : m_1(\{x_1, x_2, x_3\}) = 0.3, m_1(\{x_4, x_5\}) = 0.7$$

$$C_2 : m_2(\{x_1, x_2, x_3\}) = 0.3, m_2(\{x_3, x_4\}) = 0.7$$

Thus, the Pal entropy and B&F entropy results calculated and compared are the following:

$$\begin{aligned}
C_1 : H_{B\&F}(m_1) &= 0.3 \times \log_2 \frac{3^5}{0.3} + 0.7 \times \log_2 \frac{2^5}{0.7} \\
&= 2.8985 + 3.8602 = 6.7587 \\
C_1 : H_P(m_1) &= 0.3 \times \log_2 \frac{3}{0.3} + 0.7 \times \log_2 \frac{2}{0.7} \\
&= 0.9966 + 1.0602 = 2.0568 \\
C_2 : H_{B\&F}(m_2) &= 0.3 \times \log_2 \frac{3^4}{0.3} + 0.7 \times \log_2 \frac{2^4}{0.7} \\
&= 2.4230 + 3.1602 = 5.5832 \\
C_2 : H_P(m_2) &= 0.3 \times \log_2 \frac{3}{0.3} + 0.7 \times \log_2 \frac{2}{0.7} \\
&= 0.9966 + 1.0602 = 2.0568
\end{aligned}$$

We can draw the following conclusions:

$$H_{B\&F}(m_1) > H_{B\&F}(m_2), H_P(m_1) = H_P(m_2)$$

By comparison, we can conclude that the result of $B\&F$ is more reasonable. Because of C_2 has fewer focal elements and they have the same element x_3 in two BPAs, therefore, the uncertainty of C_1 should be bigger than the uncertainty of C_2 .

From an overall view, as long as the focal elements for every BPA are equal, the results of Pal entropy keep constant, even if the number of focal elements on FOD is different. This is unreasonable. However, for the new belief entropy, it reflects the impact of the number of FODs on information uncertainty. Obviously, the degree of information uncertainty is proportional to FOD. Thus, the new definition proposed in this paper is more reasonable for the above Section 6.1.4.

6.1.5. Example 5

In this example, we suppose a FOD that has ten focal elements, $X = \{x_1, x_2 \dots x_{10}\}$ and four mass functions, $m(\{3, 4, 5\}) = 0.05$, $m(\{7\}) = 0.05$, $m(\{B_i\}) = 0.8$, $m(X) = 0.1$, where B_i is the subset of 2^X and i is equal to the cardinality of B . We chose ten subsets of 2^X to assignment B and used Dubois entropy, Deng entropy, Pan–Deng entropy, and the new belief entropy for comparison. In Section 4, we already listed these definitions of entropy. When B_i changes, their values can be calculated by MATLAB. The calculation results of these definitions are shown in the following Table 1.

Table 1. The value of different definitions when B_i changes.

Cases	Dubois Entropy	Deng Entropy	Pan–Deng Entropy	New Entropy
$B_1 = \{1\}$	0.4114	2.6623	16.1443	5.1363
$B_2 = \{1, 2\}$	1.2114	3.9303	17.4916	13.1363
$B_3 = \{1, 2, 3\}$	1.6794	4.9082	19.8608	17.8160
$B_4 = \{1, 2 \dots 4\}$	2.0114	5.7878	20.8229	21.1363
$B_5 = \{1, 2 \dots 5\}$	2.2690	6.6256	21.8314	23.7118
$B_6 = \{1, 2 \dots 6\}$	2.4794	7.4441	22.7521	25.8160
$B_7 = \{1, 2 \dots 7\}$	2.6573	8.2532	24.1331	27.5952
$B_8 = \{1, 2 \dots 8\}$	2.8114	9.0578	25.0685	29.1363
$B_9 = \{1, 2 \dots 9\}$	2.9474	9.8600	26.0212	30.4957
$B_{10} = \{1, 2 \dots 10\}$	3.0690	10.6612	27.1947	31.7118

Table 1 and Figure 2 show that the new belief entropy is larger than Deng entropy and Dubois entropy. On the other hand, the growth trend of the new belief entropy is slower than Deng entropy and Pan–Deng entropy and the same as Dubois entropy. For example, we chose B_1 , B_2 and B_9 , B_{10} to illustrate the impact of each additional element in B_i on uncertainty, under different cardinality of B_i . From the Table 1, we can get:

$$H_{Dubois}(m_2) - H_{Dubois}(m_1) = 0.8 > H_{Dubois}(m_{10}) - H_{Dubois}(m_9) = 0.1216$$

$$H_{Deng}(m_2) - H_{Deng}(m_1) = 1.2680 \approx H_{Deng}(m_{10}) - H_{Deng}(m_9) = 0.8012$$

$$H_{P\&D}(m_2) - H_{P\&D}(m_1) = 1.3473 \approx H_{P\&D}(m_{10}) - H_{P\&D}(m_9) = 1.1735$$

$$H_{B\&F}(m_2) - H_{B\&F}(m_1) = 8 > H_{B\&F}(m_{10}) - H_{B\&F}(m_9) = 1.2161$$

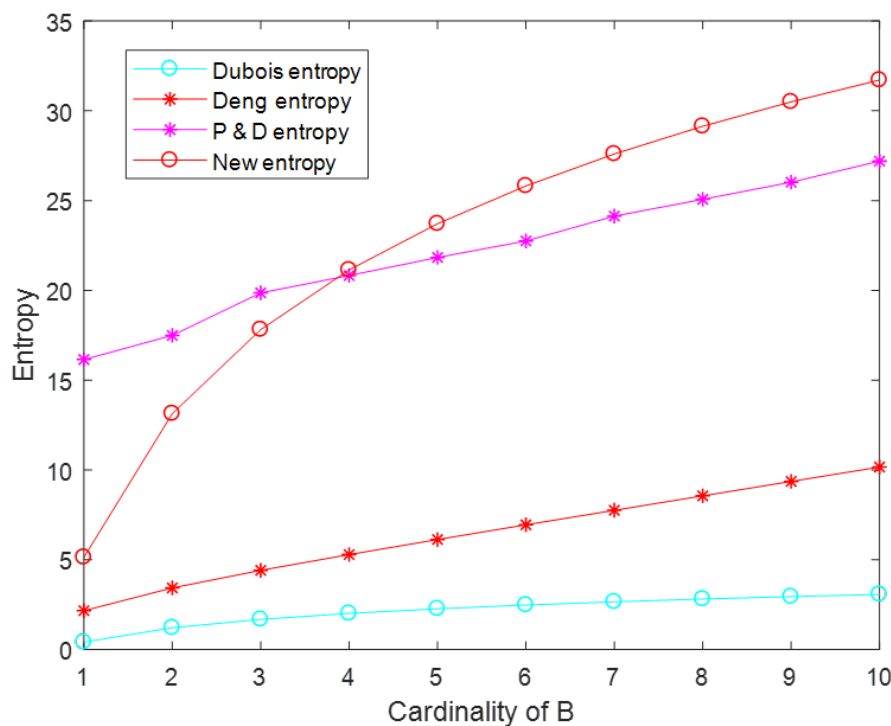


Figure 2. Comparison between the new belief entropy and other entropies.

Where the P & D entropy in Figure 2 is the $H_{Bel}(m)$ we listed in Section 4.

Although the four entropy values in Figure 2 increased, their slopes were different. Deng entropy and Pan–Deng entropy increased linearly, while the slopes of Dubois entropy and the new entropy decreased with the increase of the cardinality of B . We believe that the growth trend of the latter was more reasonable. This was because the scale of B was an important indicator to measure the change of information uncertainty, which should change with the size of cardinality. With the same cardinality of B_i , our new belief entropy was larger than the Dubois entropy. It could well reflect the degree of uncertainty. Therefore, through comprehensive analysis, we considered that the new belief entropy was more accurate.

Yager entropy, Pal entropy, Klir and Rammer entropy, and Jiroušek and Shenoy entropy are plotted in Figure 3.

From Figure 3, it can be seen that these definitions kept a small value. The degrees of uncertainty measured by Klir and Rammer and Yager decreased visibly with the increasing of the elements in B .

This was understandable. The uncertainty measures proposed by Pal and Jiroušek and Shenoy were nearly linear with the cardinality of B . They had the same growth trend as Deng entropy.

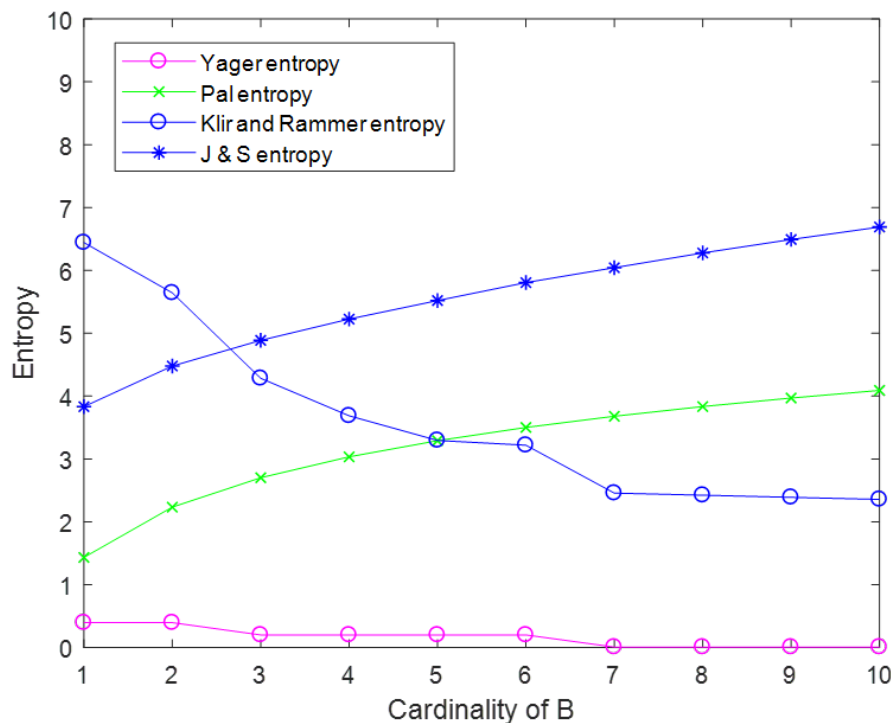


Figure 3. Results' comparison of other entropies.

Where the J & S entropy in Figure 3 is the $H_{J\&S}(m)$ we listed in Section 4.

6.1.6. Example 6

In recent years, much research has been modified based on Deng entropy theory [33,34,37]. In this example, we chose to compare W entropy and our new model.

Although W entropy takes into account the scale of FOD, the effect of the scale of FOD on W entropy is very limited [34]. As Equation (26) shows, the value of our new model would change exponentially with the scale of FOD. As they showed in their examples, when ϵ increased from zero to 10, the change trend of W entropy was almost the same as Deng entropy. However, as we demonstrated in Section 6.1.5, the growth trend of B&F entropy was different from Deng entropy. Therefore, we could see the effectiveness and superiority of the proposed entropy.

6.1.7. Example Summary

Based on the examples proposed above, we list some typical cases that may affect the new belief entropy and compare it with other entropies. From Section 6.1.2 and Section 6.1.3, we could see that the new entropy was more sensitive to the vacuous BPA. Section 6.1.4 shows the limitations of the general entropy, and the new entropy could solve the problem caused by the different number of FODs. Section 6.1.5 reflected the change of the new entropy and other entropies as the number of elements increased. In Section 6.1.6, we made a simple comparison between W entropy and B&F entropy.

6.2. Simulation

Here, we use MATLAB to complete the test. This test could more intuitively feel how the new belief entropy changed with the different BPAs.

We supposed an FOD $X = \{x_1, x_2\}$. This FOD had three BPAs, $m(\{x_1\}) = p_1$, $m(\{x_2\}) = p_2$, and $m(\{x_1, x_2\}) = 1 - p_1 - p_2$. p_1 and p_2 can take any value from zero to one. However, according to D-S theory in Section 2, we limited the value of these BPAs, where $m(\{x_1\}) + m(\{x_2\}) = p_1 + p_2 \leq 1$. Obviously, $m(\{x_1, x_2\})$ exists only when $m(\{x_1\}) + m(\{x_2\}) < 1$. The simulation results are as Figures 4 and 5 show, where the x-axis is $m(\{x_1\})$, the y-axis is $m(\{x_2\})$, and the z-axis means the value of the new entropy.

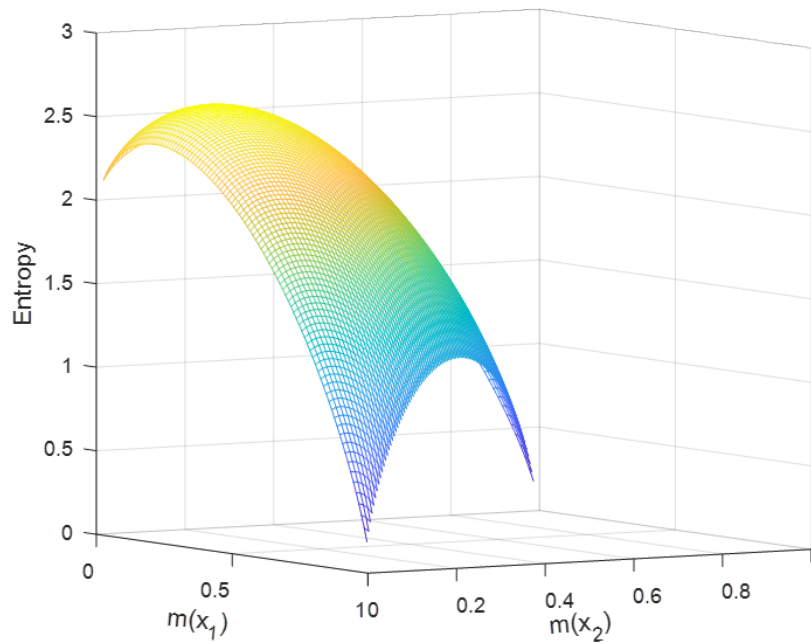


Figure 4. The value of the new belief entropy with changes of BPA.

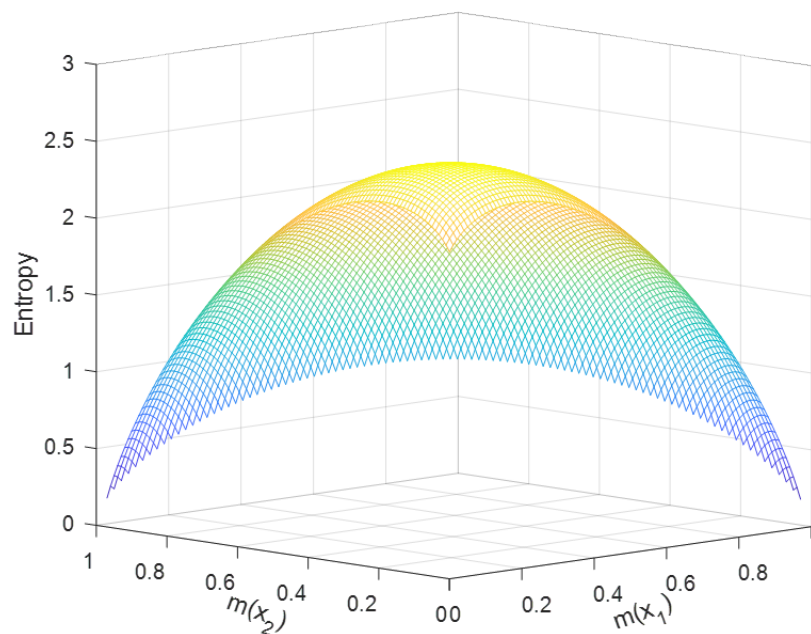


Figure 5. The value of the new belief entropy with changes of BPA.

When $p_1 + p_2 = 1$, the max value of the new entropy $H_{B\&F}(m) = 1$, where $p_1 = p_2 = 0.5$. When considering the BPA $m(\{x_1, x_2\}), m(x_1) = 0.17, m(x_2) = 0.16$ and $m(\{x_1, x_2\}) = 0.67$ could get the max value of the new entropy, $H_{B\&F(max)}(m) = 2.585$. From this max value result, we obtained that the new definition did not satisfy the maximum entropy property.

Analysis: These simulation results suggested that the main trend of the new entropy was changing with different BPAs. It also indicated that the new entropy increased as the vacuous BPA increased, when $p_1 + p_2 < 1$, which was reasonable. Therefore, the new entropy could reflect well the degree of measurement of information uncertainty.

7. Conclusions and Discussion

First of all, we reviewed some earlier definitions proposed by Hartley, Shannon, Yager, Nguyen, Lamata and Moral, Jiroušek and Shenoy, Klir and Ramer, Dubois, Nikhil R. Pal, Joussemlé, Deng, and Pan–Deng. However, none of them reflected the number of FODs' effect on uncertainty.

We discussed an open issue, which was how to measure information uncertainty. Our principle was to include as much known information as possible under D-S theory. Thus, in this paper, we considered the cardinality of FOD and defined a new model to measure uncertainty. Meanwhile, some properties of the new entropy were discussed. The result of the examples and simulation proved that the new entropy could be more effective and accurate when compared with other entropies.

When the target belonged to the set of clusters and the total number of targets could not be determined, our method could get the information uncertainty from the target accurately. Compared with traditional methods, the new entropy was easy to calculate. This meant that in the same time, it could process more data. In future work, we will apply it to solve practical problems and improve it in real applications.

Author Contributions: Conceptualization, Q.P.; Data curation, J.L.; Formal analysis, J.L.; Funding acquisition, J.L.; Investigation, J.L.; Methodology, J.L.; Project administration, J.L.; Resources, Q.P.; Software, J.L.; Supervision, Q.P.; Validation, J.L.; Visualization, J.L.; Writing—original draft, J.L.; Writing—review & editing, Q.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors greatly appreciate the reviews' suggestions and the editor's encouragement.

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

References

1. Zadeh, L.A. Fuzzy sets. *Inf. Control.* **1965**, *8*, 338–353.
2. Feller, W. *An Introduction to Probability Theory and Its Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2008; Volume 2.
3. 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.
4. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976; Volume 42.
5. Pawlak, Z. Rough sets. *Int. J. Comput. Inf. Sci.* **1982**, *11*, 341–356.
6. Deng, Y. Generalized evidence theory. *Appl. Intell.* **2015**, *43*, 530–543.
7. Yager, R.R.; Abbasov, A.M. Pythagorean membership grades, complex numbers, and decision making. *Int. J. Intell. Syst.* **2013**, *28*, 436–452.
8. Song, Y.; Wang, X.; Quan, W.; Huang, W. A new approach to construct similarity measure for intuitionistic fuzzy sets. *Soft Comput.* **2019**, *23*, 1985–1998.
9. Pan, Y.; Zhang, L.; Li, Z.; Ding, L. Improved fuzzy Bayesian network-based risk analysis with interval-valued fuzzy sets and DS evidence theory. In *IEEE Transactions on Fuzzy Systems*; IEEE: Piscataway, NJ, USA, 2019.
10. Düğenci, M. A new distance measure for interval valued intuitionistic fuzzy sets and its application to group decision making problems with incomplete weights information. *Appl. Soft Comput.* **2016**, *41*, 120–134.
11. Liu, Q.; Tian, Y.; Kang, B. Derive knowledge of Z-number from the perspective of Dempster–Shafer evidence theory. *Eng. Appl. Artif. Intell.* **2019**, *85*, 754–764.

12. Jiang, W.; Cao, Y.; Deng, X. A novel Z-network model based on Bayesian network and Z-number. In *IEEE Transactions on Fuzzy Systems*; IEEE: Piscataway, NJ, USA, 2019.
13. Deng, Y. D numbers: theory and applications. *J. Inf. Comput. Sci.* **2012**, *9*, 2421–2428.
14. Liu, B.; Deng, Y. Risk Evaluation in Failure Mode and Effects Analysis Based on D Numbers Theory. *Int. J. Comput. Commun. Control* **2019**, *14*, 672–691.
15. Deng, X.; Jiang, W. Evaluating green supply chain management practices under fuzzy environment: a novel method based on D number theory. *Int. J. Fuzzy Syst.* **2019**, *21*, 1389–1402.
16. Zhao, J.; Deng, Y. Performer Selection in Human Reliability Analysis: D numbers Approach. *Int. J. Comput. Commun. Control* **2019**, *14*, 437–452.
17. George, T.; Pal, N.R. Quantification of conflict in Dempster–Shafer framework: a new approach. *Int. J. Gen. Syst.* **1996**, *24*, 407–423.
18. Sabahi, F.; Akbarzadeh-T, M.R. A qualified description of extended fuzzy logic. *Inf. Sci.* **2013**, *244*, 60–74.
19. Deng, Y.; Liu, Y.; Zhou, D. An improved genetic algorithm with initial population strategy for symmetric TSP. *Math. Probl. Eng.* **2015**, *2015*, 212794.
20. Yang, Y.; Han, D. A new distance-based total uncertainty measure in the theory of belief functions. *Knowl.-Based Syst.* **2016**, *94*, 114–123.
21. Sabahi, F.; Akbarzadeh-T, M.R. Introducing validity in fuzzy probability for judicial decision-making. *Int. J. Approx. Reason.* **2014**, *55*, 1383–1403.
22. Deng, Y. Fuzzy analytical hierarchy process based on canonical representation on fuzzy numbers. *J. Comput. Anal. Appl.* **2017**, *22*, 201–228.
23. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
24. Hartley, R.V. Transmission of information 1. *Bell Syst. Tech. J.* **1928**, *7*, 535–563.
25. Hohle, U. Entropy with respect to plausibility measures. In *Proceedings of the 12th IEEE International Symposium on Multiple-Valued Logic*, Paris, France, 25–26 May 1982.
26. Nguyen, H.T. On entropy of random sets and possibility distributions. *Anal. Fuzzy Inf.* **1987**, *1*, 145–156.
27. Dubois, D.; Prade, H. Properties of measures of information in evidence and possibility theories. *Fuzzy Sets Syst.* **1987**, *24*, 161–182.
28. Klir, G.J.; Ramer, A. Uncertainty in the Dempster–Shafer theory: a critical re-examination. *Int. J. Gen. Syst.* **1990**, *18*, 155–166.
29. 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.
30. Pal, N.R.; Bezdek, J.C.; Hemasinha, R. Uncertainty measures for evidential reasoning I: A review. *Int. J. Approx. Reason.* **1992**, *7*, 165–183.
31. 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.
32. Deng, Y. Deng entropy. *Chaos Solitons Fractals* **2016**, *91*, 549–553.
33. 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.
34. Wang, D.; Gao, J.; Wei, D. A New Belief Entropy Based on Deng Entropy. *Entropy* **2019**, *21*, 987.
35. Frikha, A.; Moalla, H. Analytic hierarchy process for multi-sensor data fusion based on belief function theory. *Eur. J. Oper. Res.* **2015**, *241*, 133–147.
36. Khodabandeh, M.; Shahri, A.M. Uncertainty evaluation for a Dezert–Smarandache theory-based localization problem. *Int. J. Gen. Syst.* **2014**, *43*, 610–632.
37. Zhou, D.; Tang, Y.; Jiang, W. A modified belief entropy in Dempster–Shafer framework. *PLoS ONE* **2017**, *12*, e0176832.
38. Tang, Y.; Zhou, D.; He, Z.; Xu, S. An improved belief entropy-based uncertainty management approach for sensor data fusion. *Int. J. Distrib. Sens. Networks* **2017**, *13*, 1550147717718497.
39. Denoeux, T. A k-nearest neighbor classification rule based on Dempster–Shafer theory. In *Classic Works of the Dempster–Shafer Theory of Belief Functions*; Springer: Berlin, Germany, 2008; pp. 737–760.
40. Liu, Z.G.; Pan, Q.; Dezert, J. A new belief-based K-nearest neighbor classification method. *Pattern Recognit.* **2013**, *46*, 834–844.
41. Ma, J.; Liu, W.; Miller, P.; Zhou, H. An evidential fusion approach for gender profiling. *Inf. Sci.* **2016**, *333*, 10–20.

42. Liu, Z.G.; Pan, Q.; Dezert, J.; Mercier, G. Credal classification rule for uncertain data based on belief functions. *Pattern Recognit.* **2014**, *47*, 2532–2541.
43. Han, D.; Liu, W.; Dezert, J.; Yang, Y. A novel approach to pre-extracting support vectors based on the theory of belief functions. *Knowl.-Based Syst.* **2016**, *110*, 210–223.
44. Liu, Z.g.; Pan, Q.; Dezert, J.; Martin, A. Adaptive imputation of missing values for incomplete pattern classification. *Pattern Recognit.* **2016**, *52*, 85–95.
45. Jiang, W.; Wei, B.; Xie, C.; Zhou, D. An evidential sensor fusion method in fault diagnosis. *Adv. Mech. Eng.* **2016**, *8*, 1687814016641820.
46. Yuan, K.; Xiao, F.; Fei, L.; Kang, B.; Deng, Y. Modeling sensor reliability in fault diagnosis based on evidence theory. *Sensors* **2016**, *16*, 113.
47. Yuan, K.; Xiao, F.; Fei, L.; Kang, B.; Deng, Y. Conflict management based on belief function entropy in sensor fusion. *SpringerPlus* **2016**, *5*, 638.
48. Jiang, W.; Xie, C.; Zhuang, M.; Shou, Y.; Tang, Y. Sensor data fusion with z-numbers and its application in fault diagnosis. *Sensors* **2016**, *16*, 1509.
49. Yager, R.R.; Liu, L. *Classic Works of the Dempster–Shafer Theory of Belief Functions*; Springer: Berlin, Germany, 2008; Volume 219.
50. Liu, Z.g.; Pan, Q.; Dezert, J.; Mercier, G. Credal c-means clustering method based on belief functions. *Knowl.-Based Syst.* **2015**, *74*, 119–132.
51. Yager, R.R.; Alajlan, N. Decision making with ordinal payoffs under Dempster–Shafer type uncertainty. *Int. J. Intell. Syst.* **2013**, *28*, 1039–1053.
52. Merigó, J.M.; Casanovas, M. Induced aggregation operators in decision making with the Dempster–Shafer belief structure. *Int. J. Intell. Syst.* **2009**, *24*, 934–954.
53. Wang, Y.M.; Elhag, T.M. A comparison of neural network, evidential reasoning and multiple regression analysis in modelling bridge risks. *Expert Syst. Appl.* **2007**, *32*, 336–348.
54. Su, X.; Deng, Y.; Mahadevan, S.; Bao, Q. An improved method for risk evaluation in failure modes and effects analysis of aircraft engine rotor blades. *Eng. Fail. Anal.* **2012**, *26*, 164–174.
55. Fu, C.; Yang, J.B.; Yang, S.L. A group evidential reasoning approach based on expert reliability. *Eur. J. Oper. Res.* **2015**, *246*, 886–893.
56. Zhang, X.; Mahadevan, S.; Deng, X. Reliability analysis with linguistic data: An evidential network approach. *Reliab. Eng. Syst. Saf.* **2017**, *162*, 111–121.
57. Yager, R.R. Arithmetic and other operations on Dempster–Shafer structures. *Int. J. Man-Mach. Stud.* **1986**, *25*, 357–366.
58. Li, Y.; Deng, Y. Intuitionistic evidence sets. *IEEE Access* **2019**, *7*, 106417–106426.
59. Song, Y.; Deng, Y. Divergence measure of belief function and its application in data fusion. *IEEE Access* **2019**, *7*, 107465–107472.
60. Klir, G.J.; Wierman, M.J. *Uncertainty-Based Information: Elements of Generalized Information Theory*; Springer: Berlin, Germany, 2013; Volume 15.
61. Klir, G.; Folger, T. *Fuzzy Sets, Uncertainty, and Information*; Prentice Hall: Englewood Cliffs, NJ, 1988.
62. Yager, R.R. Entropy and specificity in a mathematical theory of evidence. *Int. J. Gen. Syst.* **1983**, *9*, 249–260.
63. LAMATA, M.T.; MORAL, S. Measures of entropy in the theory of evidence. *Int. J. Gen. Syst.* **1988**, *14*, 297–305.
64. Jousselme, A.L.; Liu, C.; Grenier, D.; Bossé, É. Measuring ambiguity in the evidence theory. In *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*; IEEE: Piscataway, NJ, USA, 2006; Volume 36, 890–903.
65. Smets, P. Constructing the Pignistic Probability Function in a Context of Uncertainty. *UAI* **1989**, *89*, 29–40.
66. Abellán, J. Analyzing properties of Deng entropy in the theory of evidence. *Chaos Solitons Fractals* **2017**, *95*, 195–199.

