

Article

A Variation of the Algorithm to Achieve the Maximum Entropy for Belief Functions

Joaquín Abellán * , Alejandro Pérez-Lara  and Serafín Moral-García

Department of Computer Science and Artificial Intelligence, University of Granada, 18014 Granada, Spain; alejandropi@correo.ugr.es (A.P.-L.); seramoral@decsai.ugr.es (S.M.-G.)

* Correspondence: jabellan@decsai.ugr.es

Abstract: Evidence theory (TE), based on imprecise probabilities, is often more appropriate than the classical theory of probability (PT) to apply in situations with inaccurate or incomplete information. The quantification of the information that a piece of evidence involves is a key issue in TE. Shannon's entropy is an excellent measure in the PT for such purposes, being easy to calculate and fulfilling a wide set of properties that make it axiomatically the best one in PT. In TE, a similar role is played by the maximum of entropy (ME), verifying a similar set of properties. The ME is the unique measure in TE that has such axiomatic behavior. The problem of the ME in TE is its complex computational calculus, which makes its use problematic in some situations. There exists only one algorithm for the calculus of the ME in TE with a high computational cost, and this problem has been the principal drawback found with this measure. In this work, a variation of the original algorithm is presented. It is shown that with this modification, a reduction in the necessary steps to attain the ME can be obtained because, in each step, the power set of possibilities is reduced with respect to the original algorithm, which is the key point of the complexity found. This solution can provide greater applicability of this measure.

Keywords: belief functions; uncertainty measures; maximum of entropy



Citation: Abellán, J.; Pérez-Lara, A.; Moral-García, S. A Variation of the Algorithm to Achieve the Maximum Entropy for Belief Functions. *Entropy* **2023**, *25*, 867. <https://doi.org/10.3390/e25060867>

Academic Editor: Yong Deng

Received: 8 May 2023

Revised: 18 May 2023

Accepted: 26 May 2023

Published: 29 May 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Managing uncertainty is essential for making decisions. Evidence theory (TE), also known as the Dempster–Shafer theory [1,2], is widely employed to handle uncertainty-based information in practical applications such as *medical diagnosis* [3], *statistical classification* [4], *target identification* [5], *face recognition* [6], or *risk management* [7,8]. TE is also commonly utilized to fuse information from different sources [9–11], a crucial issue for decision making.

Evidence theory extends classical Probability Theory (PT). It is based on the *basic probability assignment* concept (b.p.a.), a generalization of the concept of the probability distribution in PT. Each b.p.a. in TE has a belief function and a plausibility function associated with it. The belief (plausibility) value of a set is the minimum (maximum) support of information represented by the b.p.a. on that set.

In TE, it is essential to quantify the uncertainty-based information represented by a b.p.a. For this purpose, many uncertainty measures in TE have been proposed so far. The starting point of most of them is the Shannon entropy [12], a well-established uncertainty measure in PT that satisfies a large set of properties.

As TE generalizes PT, there are more types of uncertainty in TE than in PT. As pointed out by Yager [13], two types of uncertainty appear in TE. The first one, called *conflict*, arises when the information focuses on disjoint sets. The second type, known as *non-specificity*, appears when the information has a cardinality greater than or equal to two. Hence, a total uncertainty measure in TE must capture both the conflict and non-specificity.

Klir and Weirman [14] carried out a study concerning the set of mathematical properties that have to be verified by every uncertainty measure in TE. Such a study was extended

by Abellán and Masegosa [15]. They also analyzed the behavioral requirements that a total uncertainty measure in TE must satisfy. The maximum entropy on the closed and convex set of probability distributions (credal set) compatible with a b.p.a., proposed in [16], is the only total uncertainty measure in TE so far that verifies all the crucial mathematical properties and behavioral requirements.

However, the algorithms proposed so far in [16,17] to compute the maximum entropy of the credal set associated with a b.p.a. are notably complex. For this reason, in recent years, many alternative measures to the maximum entropy have been introduced. Nonetheless, none of these measures verifies all the required mathematical properties and behaviors for uncertainty measures in TE [18–20]. This is the principal reason for the lack of consensus on the use of uncertainty measures in TE: the ME has an optimal axiomatic behavior but a high complexity of calculus, and more recent ones have had worse axiomatic behavior but a low complexity of calculus.

It must be remarked that the ME has been used with excellent results in practical applications, such as in the data mining area. Its use on special types of belief functions can be of different complexity, and its calculus is immediate. We can find examples of this in [21–24].

An approximation to the maximum entropy on the credal set associated with a b.p.a. was proposed in [25]. Such an approximation consisted of the maximum entropy on the credal set consistent with the belief intervals for singletons, where the lower and upper bounds were, respectively, the belief and plausibility values on singletons. Even though this measure satisfied all the crucial mathematical properties and behavioral requirements, when the belief intervals for singletons are employed to represent the uncertainty-based information instead of the corresponding b.p.a., some information could be lost because the credal set consistent with a b.p.a. is always contained in the one compatible with the associated belief intervals for singletons [25]. In consequence, this uncertainty measure indicated more uncertainty than the one represented by a b.p.a.

In this research, we propose a variation of the algorithm for computing the maximum entropy on the credal set compatible with a b.p.a. We demonstrate that our proposed procedure involves less computational time than the algorithms developed so far for the maximum entropy on the credal set corresponding to a b.p.a. With our proposal, fewer steps are necessary to achieve the probability distribution of maximum entropy on the credal set consistent with a b.p.a. This is shown via some numerical examples and with an experimentation over a huge set of b.p.a. functions randomly generated. The reduction in the computational time of our proposed algorithm makes this measure (maximum of entropy) more suitable for use in practical applications.

The remainder of this paper is structured as follows: Section 2 describes evidence theory, the main uncertainty measures proposed so far in evidence theory, and the algorithm developed so far to compute the maximum entropy on the credal set consistent with a basic probability assignment. Our proposed procedure and different examples of its use, compared with the original one, are presented in Section 3. Moreover, in that section, we present an experiment showing the gains in the time of computing obtained with the new improved algorithm compared with the original one. Concluding remarks and ideas for future work are given in Section 4.

2. Background

Let $X = \{x_1, \dots, x_t\}$ be a finite set of possible alternatives, also known as the *frame of discernment*. Let $\wp(X)$ denote the power set of X .

2.1. Theory of Evidence

Evidence theory (TE), also known as the Dempster–Shafer theory [1,2], is based on the concept of a *basic probability assignment* (b.p.a.). A b.p.a. is a mapping $m : \wp(X) \rightarrow [0, 1]$, satisfying $m(\emptyset) = 0$ and $\sum_{A \subseteq X} m(A) = 1$.

If $A \subseteq X$ verifies that $m(A) > 0$, A is said to be a *focal element* of m .

A given b.p.a. m on X has a *belief function* Bel_m and a *plausibility function* Pl_m associated with it. Such functions are defined as follows:

$$Bel_m(A) = \sum_{B|B \subseteq A} m(B), \quad Pl_m(A) = \sum_{B|B \cap A \neq \emptyset} m(B), \quad \forall A \subseteq X. \quad (1)$$

We may note that for each $A \subseteq X$, $Bel_m(A) \leq Pl_m(A)$. The interval $[Bel_m(A), Pl_m(A)]$ is called the *belief interval* of $A \quad \forall A \subseteq X$. In addition,

$$Pl_m(A) = 1 - Bel_m(\bar{A}), \quad \forall A \subseteq X, \quad (2)$$

where \bar{A} denotes the complement of A . Thereby, Bel_m and Pl_m are called dual or conjugate. One of them is sufficient for representing the uncertainty-based information in TE. For this purpose, Bel_m is often utilized.

For a given b.p.a. m on X , the set of probability distributions compatible with it (a closed and convex set of probability distributions, also called a credal set) is given by:

$$\mathcal{P}_m = \{p \in \mathcal{P}(X) \mid Bel_m(A) \leq p(A) \quad \forall A \subseteq X\}, \quad (3)$$

where $\mathcal{P}(X)$ is the set of all probability distributions on X .

2.2. Uncertainty Measures in Evidence Theory

The Shannon entropy [12] is a well-established uncertainty measure in probability theory. For a probability distribution p on X , the Shannon entropy is defined as follows:

$$S(p) = \sum_{i=1}^t -p(\{x_i\}) \log_2(p(\{x_i\})). \quad (4)$$

The type of uncertainty measured by S is called *conflict*. It is the only type of conflict present in probability theory. It satisfies a set of desirable properties [12,14].

In classical possibility theory, the Hartley measure [26] is a well-established uncertainty measure. It is defined in the following way:

$$H(A) = \log_2(|A|), \quad \forall A \in \wp(X). \quad (5)$$

The type of uncertainty measured by H , the only one existing in possibility theory, is known as *non-specificity*.

As pointed out by Yager [13], conflict and non-specificity coexist in TE; conflict appears when the information is focused on disjoint sets, while non-specificity arises when the information resides in sets with a cardinality greater than one.

A generalization of the Hartley measure to TE was introduced by Dubois and Prade in [27]. It is given by:

$$GH(m) = \sum_{A \subseteq X} m(A) \log_2(|A|). \quad (6)$$

GH reaches its minimum value, 0, when m is a probability distribution, and the maximum value of GH , $\log_2(t)$, is attained when $m(X) = 1$. GH is an appropriate non-specificity measure in TE that satisfies desirable properties. Moreover, it is easily extensible to more general theories than TE [28].

Several attempts to generalize the Shannon entropy to TE have been proposed, but none of them has satisfied all the essential requirements for this type of measure in TE: probability consistency, set consistency, range, additivity, subadditivity, and monotonicity [15].

Next, a total uncertainty measure in TE that quantified both the conflict and non-specificity was proposed. Such a measure, developed by Harmanec and Klir [16], is the maximum entropy on the credal set consistent with the b.p.a. m , \mathcal{P}_m , determined via Equation (3). It is denoted by $S^*(\mathcal{P}_m)$. This measure is suitable for quantifying uncertainty

in TE because it is the only one so far that satisfies all necessary mathematical properties and behavioral requirements for uncertainty measures in TE.

Nevertheless, the algorithms proposed so far in [16,17] for computing $S^*(\mathcal{P}_m)$ (also noted in the literature as $S^*(Bel)$ or $S^*(Bel_m)$, where all the expressions have the same meaning: the maximum of entropy over all the probability distributions associated with a b.p.a. m), are very complex. For this reason, in recent years, many alternative measures to $S^*(\mathcal{P}_m)$ have been proposed.

For instance, the Deng entropy was presented in [18,29–31]. It was defined in the following way:

$$E_d(m) = - \sum_{A \subseteq X} m(A) \log_2 \left(\frac{m(A)}{2^{|A|} - 1} \right) = \sum_{A \subseteq X} m(A) \log_2 (2^{|A|} - 1) - \sum_{A \subseteq X} m(A) \log_2 (m(A)). \tag{7}$$

In Equation (7), the first term captures the non-specificity, while the second one quantifies the conflict part. The idea of this measure is that there must be more uncertainty as the number of alternatives increases. However, the Deng entropy violates most of the required mathematical properties for the uncertainty measures in TE, and its behavior in many cases is problematic [19].

The basis for some recent uncertainty measures in TE is the plausibility transformation [32,33], defined in the following way:

$$Pt(x_i) = \frac{Pl_m(\{x_i\})}{\sum_{j=1}^t Pl_m(\{x_j\})}, \quad \forall i = 1, \dots, t. \tag{8}$$

Jirousek and Shenoy [34] introduced a new uncertainty measure consisting of the sum of the GH and the Shannon entropy of the plausibility transformation:

$$H_{JS}(m) = GH(m) - \sum_{i=1}^t Pt(x_i) \log_2 (Pt(x_i)). \tag{9}$$

The first term of Equation (9) captures the conflict, whereas the second one corresponds to the non-specificity.

In [35], an uncertainty measure also based on the plausibility transformation was proposed. It is defined as follows:

$$H_{PQ}(m) = - \sum_{A \subseteq X} m(A) \log_2 (Pm(A)) + GH(m), \tag{10}$$

where for each $A \subseteq X$, $m(A) = \sum_{x \in A} Pt(x)$. The first term quantifies the conflict, while the second one captures the non-specificity.

As shown in [25], the H_{JS} does not satisfy all the required mathematical properties for the uncertainty measures in TE. The same situation occurs with the H_{PQ} .

Let us consider the set of belief intervals for singletons associated with m :

$$\mathcal{I}_m = \{[Bel_m(\{x_i\}), Pl_m(\{x_i\})] \mid i = 1, \dots, t\}. \tag{11}$$

The uncertainty measure proposed in [36] combines the Deng entropy with the belief intervals for singletons. It is defined as:

$$\begin{aligned} H_{inter}(m) = & - \sum_{x \in X} \frac{Bel_m(\{x\}) + Pl_m(\{x\})}{2} \\ & \times \log_2 \left[\frac{Bel_m(\{x\}) + Pl_m(\{x\})}{2} \times \exp(-(Pl_m(\{x\}) - Bel_m(\{x\}))) \right] \\ & - \sum_{A \subseteq X \mid |A| \geq 2} m(A) \times \log_2 \left[\frac{m(A)}{2^{|A|} - 1} \exp(-(Pl_m(A) - Bel_m(A))) \right]. \end{aligned} \tag{12}$$

In Equation (12), the first term captures the conflict, and the second one quantifies the non-specificity. As demonstrated in [25], the measure also does not satisfy all the crucial mathematical properties for the uncertainty measures in TE.

Let $\mathcal{P}(\mathcal{I}_m)$ denote the credal set consistent with the belief intervals for singletons associated with m . It is determined by:

$$\mathcal{P}(\mathcal{I}_m) = \{p \in \mathcal{P}(X) \mid Bel_m(\{x_i\}) \leq p(\{x_i\}) \leq Pl_m(\{x_i\}), \forall i = 1, \dots, t\}, \quad (13)$$

with $\mathcal{P}(X)$ being the set of all probability distributions on X .

In [25], an uncertainty measure that consists of the maximum entropy on the credal set given in Equation (13) was proposed:

$$S^*(\mathcal{P}(\mathcal{I}_m)) = \max_{p \in \mathcal{P}(\mathcal{I}_m)} S(p). \quad (14)$$

$S^*(\mathcal{P}(\mathcal{I}_m))$ verified all the essential mathematical properties and behavioral requirements for uncertainty measures in TE, as demonstrated in [25].

Nevertheless, it always holds that $\mathcal{P}_m \subseteq \mathcal{P}(\mathcal{I}_m)$. Consequently, using \mathcal{I}_m rather than m to represent uncertainty may imply some loss of information. Therefore, $S^*(\mathcal{P}(\mathcal{I}_m))$ might indicate more uncertainty than the one involved in m . The principal advantage of the use of the maximum entropy in this case is the notable reduction in the complexity though this implies a possible loss of information.

2.3. Algorithm to Compute the Maximum Entropy

For the calculation of the maximum entropy, it is necessary to solve a nonlinear optimization problem. To solve this issue, Meyerowitz et al. [17] proposed an algorithm for the calculation of the maximum entropy given a belief function. The algorithm follows these steps:

Input: A belief function Bel on the frame of discernment X .

1. Find a nonempty set $A \in \wp(X)$, such that $\frac{Bel(A)}{|A|}$ is maximal. If there exists more than one set A that attains that maximal, choose the one with maximal cardinality.
2. For $x \in A$, put $p_x = \frac{Bel(A)}{|A|}$.
3. For each $B \in \wp(X \setminus A)$, put $Bel'(B) = Bel(B \cup A) - Bel(A)$.
4. Put $X = X \setminus A$.
5. If $X \neq \emptyset$ and $Bel'(X) > 0$, then go to Step 1.
6. If $Bel(X) = 0$ and $X \neq \emptyset$, then put $p_x = 0, \forall x \in X$.
7. Calculate $S^*(Bel) = -\sum_{x \in X} p_x \log_2 p_x$.

Meyerowitz et al.'s algorithm provides a process for taking the probabilities that maximize the Shannon entropy, starting from a given belief function. This algorithm has a complexity of $2^{|X|}$ because in each iteration it is necessary to check which set maximizes $\frac{Bel(A)}{|A|}, \forall A \subseteq X$.

3. A Computational Improvement of Meyerowitz et al.'s Algorithm

In the following, we present the improvement of Meyerowitz et al.'s algorithm. For this new procedure, we need to define an accumulator variable (acu), which we initialize by taking $acu = 1$; this variable can be considered as the probability to be distributed among the elements of the frame of discernment X . The algorithm follows these steps:

Input: A belief function Bel on the frame of discernment X .

1. Find a nonempty set $A \in \wp(X)$, such that $\frac{Bel(A)}{|A|}$ is maximal. If there exists more than one set A that attains that maximal, choose the one with maximal cardinality.
2. Find a nonempty set $B \in \wp(X)$, such that $\frac{Bel(B)}{|B|}$ is minimal. If there exists more than one set B that attains that minimal, choose the one with minimal cardinality.
3. For $x \in A$, put $p_x = \frac{Bel(A)}{|A|}$.

4. For each $y \in B$, put $p_y = \frac{Pl(B)}{|B|}$.
5. Put $acu = acu - Bel(A) - Pl(B)$.
6. For each $C \in \wp(X \setminus \{A \cup B\})$, put $Bel'(C) = Bel(C \cup A) - Bel(A)$.
7. Put $X = X \setminus \{A \cup B\}$.
8. For each $C \in \wp(X)$, put $Pl'(C) = acu - Bel'(\bar{C})$.
9. If $X \neq \emptyset$ and $X \neq \emptyset$, then go to the first step.
10. If $Bel(X) = 0$ and $X \neq \emptyset$, then put $p_x = 0, \forall x \in X$.
11. Calculate $S^*(Bel) = -\sum_{x \in X} p_x \log_2 p_x$.

Note that, for steps 1 and 2, the same set of subsets is used. We note this characteristic because it is important: we only need to calculate and use once the power set of X . We observe that the cardinality of the resulting frame of discernment is reduced more than in the original algorithm. We must remember that the use of the powerset of the frame of discernment is the principal drawback found in the original algorithm. Hence, if the size of such a set is reduced, it is obvious that we can gain time for the calculus of that measure.

3.1. Justification

The idea behind this algorithm is based on the use of the property $Pl(A) = 1 - Bel(\bar{A}) \quad \forall A \in \wp(X)$, which relates the belief and plausibility functions. To achieve this, we take the original algorithm as a starting point. Let $A_i (i \in \{1, \dots, n\})$ be ordered disjoint sets, obtained after n iterations of Meyerowitz et al.'s algorithm. These sets are those that verify that $\frac{Bel(A)}{|A|}$ is maximal at each iteration of the algorithm (for simplicity, we assume that $p_x \neq 0 : x \in A_n$. In the case where this is not satisfied, we consider the sets A_i with $1 \leq i \leq n - 1$).

Within the algorithm, we defined the function $Bel'(B) = Bel(B \cup A_1) - Bel(A_1) \quad \forall B \subseteq X \setminus A$. Now, our goal is to find a way to relate successive iterations of the function Bel to the function Pl . For this purpose, the following expression was proposed, where Bel^n is the belief function at the n -th iteration of the original algorithm:

$$Bel^n(A_{n+1}) = Bel(\cup_{i=1}^{n+1} A_i) - Bel(\cup_{i=1}^n A_i). \tag{15}$$

To prove the correctness of this equation, we give the following property.

Proposition 1. *Let Bel and Bel^n be the belief functions in the first and last iteration of Meyerowitz et al.'s algorithm, respectively. Then, the following is verified.*

$$Bel^n(A_{n+1}) = Bel(\cup_{i=1}^{n+1} A_i) - Bel(\cup_{i=1}^n A_i). \tag{16}$$

Proof. We use induction to prove the equality:

- For $n = 1$:

$$Bel'(A_2) = Bel(A_2 \cup A_1) - Bel(A_1). \tag{17}$$

- For $n = 2$:

$$\begin{aligned} Bel^2(A_3) &= Bel'(A_3 \cup A_2) - Bel'(A_2), \\ Bel^2(A_3) &= Bel(A_3 \cup A_2 \cup A_1) - Bel(A_1) - Bel(A_2 \cup A_1) + Bel(A_1), \\ Bel^2(A_3) &= Bel(A_3 \cup A_2 \cup A_1) - Bel(A_2 \cup A_1). \end{aligned} \tag{18}$$

- Now, we assume true for n , then

$$\begin{aligned}
 Bel^{n+1}(A_{n+2}) &= Bel^n(A_{n+2} \cup A_{n+1}) - Bel^n(A_{n+1}) \\
 &= Bel^n(A_{n+2} \cup A_{n+1}) - Bel(\cup_{i=1}^{n+1} A_i) + Bel(\cup_{i=1}^n A_i) \\
 &= Bel^{n-1}(A_{n+2} \cup A_{n+1} \cup A_n) - Bel^{n-1}(A_n) - Bel(\cup_{i=1}^{n+1} A_i) + Bel(\cup_{i=1}^n A_i) \\
 &= Bel^{n-1}(A_{n+2} \cup A_{n+1} \cup A_n) - Bel(\cup_{i=1}^n A_i) + Bel(\cup_{i=1}^{n-1} A_i) \\
 &\quad - Bel(\cup_{i=1}^{n+1} A_i) + Bel(\cup_{i=1}^n A_i) \\
 &= Bel^{n-1}(A_{n+2} \cup A_{n+1} \cup A_n) + Bel(\cup_{i=1}^{n-1} A_i) - Bel(\cup_{i=1}^{n+1} A_i) \\
 &\dots \\
 &= Bel'(\cup_{i=2}^{n+2} A_i) + Bel(A_1) - Bel(\cup_{i=1}^{n+1} A_i) \\
 &= Bel(\cup_{i=1}^{n+2} A_i) - Bel(A_1) + Bel(A_1) - Bel(\cup_{i=1}^{n+1} A_i) \\
 &= Bel(\cup_{i=1}^{n+2} A_i) - Bel(\cup_{i=1}^{n+1} A_i).
 \end{aligned}
 \tag{19}$$

Hence, it is verified for $n + 1$. \square

Meyerowitz et al.’s algorithm ends after a finite number of steps; in our case, we will say n . Therefore, in the last iteration, we would obtain that $Bel^{n-1}(A_n)$ was the last set for which the probability was calculated. So, if we apply Equation (17), we obtain:

$$\begin{aligned}
 Bel^{n-1}(A_n) &= Bel(\cup_{i=1}^n A_i) - Bel(\cup_{i=1}^{n-1} A_i) \\
 &= 1 - Bel(\cup_{i=1}^{n-1} A_i) \\
 &= Pl(A_n).
 \end{aligned}
 \tag{20}$$

Having established this new relationship, we must take into account that, to implement it in the algorithm, it is necessary to take the minimum of the plausibility functions. This is because the relationship is between the original plausibility function and the belief function at the n -th iteration, which is the smallest of all.

The next step is to know how the successive plausibility functions are calculated over the iterations of the algorithm and how this affects the calculation of Bel' . For this purpose, the following property is stated.

Proposition 2. *The calculation of Bel' and Pl' follows the following equation:*

$$Bel'(C) = Bel(C \cup A_1) - Bel(A_1), \quad \forall C \subseteq X \setminus \{A_1 \cup A_n\}.
 \tag{21}$$

Proof. We begin by studying what happens to the function Bel' .

For the calculation of the function Bel' , we do not take into account the set A_n ; this was excluded and not taken into account in the following iterations.

We have to study how the function Pl' is calculated. For this purpose, we use Equation (15).

Considering the sets A_i with $2 \leq i \leq n - 1$, we have:

$$Bel^{n-2}(A_{n-1}) = Bel'(\cup_{i=2}^{n-1} A_i) - Bel'(\cup_{i=2}^{n-2} A_i).
 \tag{22}$$

In the improvement of the algorithm, we defined $Bel'(\cup_{i=2}^{n-1} A_i) = acu$. Hence, making use of (15), we have:

$$Bel^{n-2}(A_{n-1}) = acu - Bel'(\cup_{i=2}^{n-2} A_i).
 \tag{23}$$

In addition, since the set $A_i : i \in \{1, \dots, n\}$ is disjoint, and we defined the new frame of discernment as $X \setminus \{A_1 \cup A_n\}$, we have that the set $\cup_{i=2}^{n-2} A_i = \overline{A_{n-1}}$. Therefore, we obtain:

$$Pl'(A_{n-1}) = acu - Bel'(\cup_{i=2}^{n-2} A_i). \tag{24}$$

□

3.2. Example 1

Given the frame of discernment $X = \{a, b, c, d\}$, if we take the belief function Bel defined by the following basic probability assignment m (see [14]), then:

$$\begin{aligned} m(\{a\}) &= 0.26 \\ m(\{b\}) &= 0.26 \\ m(\{c\}) &= 0.26 \\ m(\{a, b\}) &= 0.07 \\ m(\{a, c\}) &= 0.01 \\ m(\{a, d\}) &= 0.01 \\ m(\{b, c\}) &= 0.01 \\ m(\{b, d\}) &= 0.01 \\ m(\{c, d\}) &= 0.01 \\ m(\{a, b, c, d\}) &= 0.1 \end{aligned}$$

- **Meyerowitz et al.'s algorithm**

- **First iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{a, b\}$. So, we have $p_a = p_b = 0.295$. Now, we update the function Bel , and as $X = \{c, d\}$, and there are sets whose function Bel is nonzero, we need a second iteration.
- **Second iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{c\}$. Thereby, it holds that $p_c = 0.28$. Now, we update the function Bel , and since $X = \{d\}$, and there are sets whose function Bel is nonzero, we need a third iteration.
- **Third iteration:** For this last iteration, we have that $X = \{d\}$, and we have that $Bel(\{d\}) = p_d = 0.13$. With this, we arrive at $X = \emptyset$; so, we can now calculate $S^*(Bel)$.

Now, we proceed to the calculation of the maximum entropy:

$$\begin{aligned} S^*(Bel) &= - \sum_{i \in \{a, b, c, d\}} p_i \log_2 p_i \\ &= - 2 \times 0.295 \log_2(0.295) - 0.28 \log_2(0.28) - 0.13 \log_2(0.13) \\ &= 1.93598. \end{aligned} \tag{25}$$

- **Improvement of Meyerowitz et al.'s algorithm**

- **First iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{a, b\}$. Therefore, we have $p_a = p_b = 0.295$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{d\}$, and we have $p_d = 0.13$. We update the value of $acu = 0.28$, the function Bel , and the function Pl , and since $X = \{c\}$, and there are sets whose function Bel is nonzero, we need a second iteration.
- **Second iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{c\}$. So, it is satisfied that $p_c = 0.28$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is the same that we obtained for $A, B = \{c\}$, and we have $p_c = 0.28$. With this we arrive at $X = \emptyset$. Now, we can calculate $S^*(Bel)$.

Now, we proceed to the calculation of the maximum entropy:

$$\begin{aligned}
 S^*(Bel) &= - \sum_{i \in \{a,b,c,d\}} p_i \log_2 p_i \\
 &= - 2 \times 0.295 \log_2(0.295) - 0.28 \log_2(0.28) - 0.13 \log_2(0.13) \\
 &= 1.93598.
 \end{aligned} \tag{26}$$

In this example, we reduced one step, i.e., we needed the calculus of the power set one time less. In this case, we can observe that the probability distribution where the maximum of entropy was attained, was very close to the initial mass values of the singletons. Hence, the original algorithm needed little effort to find that maximum. Thus, the improvement with the new algorithm was not large.

3.3. Example 2

Given the frame of discernment $X = \{a, b, c, d, e, f\}$, we take the belief function Bel defined by the following basic probability assignment m :

$$\begin{aligned}
 m(\{a\}) &= 0.25, \\
 m(\{a, b\}) &= 0.18, \\
 m(\{a, b, c\}) &= 0.16, \\
 m(\{a, b, c, d\}) &= 0.15, \\
 m(\{a, b, c, d, e\}) &= 0.14, \\
 m(\{a, b, c, d, e, f\}) &= 0.12.
 \end{aligned}$$

- **Meyerowitz et al.'s algorithm**

- **First iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{a\}$. So, we have $p_a = 0.25$. Now, we update the function Bel , and as $X = \{b, c, d, e, f\}$, and there are sets whose function Bel is nonzero, we need a second iteration.
- **Second iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{b\}$. Thereby, it holds that $p_b = 0.18$. Now, we update the function Bel , and since $X = \{c, d, e, f\}$, and there are sets whose function Bel is nonzero, we need a third iteration.
- **Third iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{c\}$. In this way, we have $p_c = 0.16$. Now, we update the function Bel , and as $X = \{d, e, f\}$, and there are sets whose function Bel is non-zero, we need a fourth iteration.
- **Fourth iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{d\}$. Hence, it is satisfied that $p_d = 0.15$. Now, we update the function Bel , and since $X = \{e, f\}$, and there are sets whose function Bel is nonzero, we need a fifth iteration.
- **Fifth iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{e\}$. Consequently, it holds that $p_e = 0.14$. Now, we update the function Bel , and as $X = \{f\}$, and there are sets whose function Bel is nonzero, we need a sixth iteration.
- **Sixth iteration:** For this last iteration, we have $X = \{f\}$, and we have $Bel(\{f\}) = p_f = 0.12$. With this, we arrive at $X = \emptyset$; so, we can now calculate $S^*(Bel)$.

Now, we proceed to the calculation of the maximum entropy:

$$\begin{aligned}
 S^*(Bel) &= - \sum_{i \in \{a,b,c,d,e,f\}} p_i \log_2 p_i \\
 &= -0.25 \log_2(0.25) - 0.18 \log_2(0.18) - 0.16 \log_2(0.16) \\
 &\quad - 0.15 \log_2(0.15) - 0.14 \log_2(0.14) - 0.12 \log_2(0.12) \\
 &= 2.5430.
 \end{aligned} \tag{27}$$

- **Improvement of Meyerowitz et al.'s algorithm**

- **First iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{a\}$. Therefore, we have $p_a = 0.25$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{f\}$, and we have that $p_f = 0.12$. We update the value of $acu = 0.63$, the function Bel , and the function Pl , and since $X = \{b, c, d, e\}$, and there are sets whose function Bel is nonzero, we need a second iteration.
- **Second iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{b\}$. So, it is satisfied that $p_b = 0.18$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{e\}$, and we have that $p_e = 0.14$. We update the value of $acu = 0.31$, the function Bel , and the function Pl , and as $X = \{c, d\}$, and there are sets whose function Bel is nonzero, we need a third iteration.
- **Third iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{c\}$. Thus, we have $p_c = 0.16$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{d\}$, and we have that $p_d = 0.15$. With this, we arrive at $X = \emptyset$. Now, we can calculate $S^*(Bel)$.

Now, we proceed to the calculation of the maximum entropy:

$$\begin{aligned}
 S^*(Bel) &= - \sum_{i \in \{a,b,c,d,e,f\}} p_i \log_2 p_i \\
 &= -0.25 \log_2(0.25) - 0.18 \log_2(0.18) - 0.16 \log_2(0.16) \\
 &\quad - 0.15 \log_2(0.15) - 0.14 \log_2(0.14) - 0.12 \log_2(0.12) \\
 &= 2.5430.
 \end{aligned} \tag{28}$$

In this example, we reduced the number of steps to the middle one. This example used a b.p.a. with no conflict because all the focal sets shared an element.

3.4. Example 3

Given the frame of discernment $X = \{a, b, c, d, e, f, g, h\}$, we take the belief function Bel defined by the following basic probability assignment m :

$$\begin{aligned}
 m(\{a\}) &= 0.23, \\
 m(\{g\}) &= 0.05, \\
 m(\{a, b\}) &= 0.21, \\
 m(\{g, h\}) &= 0.02, \\
 m(\{a, b, c\}) &= 0.18, \\
 m(\{a, b, c, d\}) &= 0.13, \\
 m(\{a, b, c, d, e\}) &= 0.1, \\
 m(\{a, b, c, d, e, f\}) &= 0.08.
 \end{aligned}$$

- **Meyerowitz et al.'s algorithm**

- **First iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{a\}$. So, we have $p_a = 0.23$. Now, we update the function Bel , and as $X = \{b, c, d, e, f, g, h\}$, and there are sets whose function Bel is nonzero, we need a second iteration.
- **Second iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{b\}$. Thereby, it holds that $p_b = 0.21$. Now, we update the function Bel , and since $X = \{c, d, e, f, g, h\}$, and there are sets whose function Bel is nonzero, we need a third iteration.
- **Third iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{c\}$. In this way, we have $p_c = 0.18$. Now, we update the function Bel , and as $X = \{d, e, f, g, h\}$, and there are sets whose function Bel is nonzero, we need a fourth iteration.
- **Fourth iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{d\}$. Hence, it is satisfied that $p_d = 0.13$. Now, we update the function Bel , and since $X = \{e, f, g, h\}$, and there are sets whose function Bel is nonzero, we need a fifth iteration.
- **Fifth iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{e\}$. Consequently, it holds that $p_e = 0.1$. Now, we update the function Bel , and as $X = \{f, g, h\}$, and there are sets whose function Bel is nonzero, we need a sixth iteration.
- **Sixth iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{f\}$. In this way, we have $p_f = 0.08$. Now, we update the function Bel , and as $X = \{g, h\}$, and there are sets whose function Bel is nonzero, we need a seventh iteration.
- **Seventh iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{g\}$. Hence, it is satisfied that $p_g = 0.05$. Now, we update the function Bel , and as $X = \{h\}$, and there are sets whose function Bel is nonzero, we need an eighth iteration.
- **Eighth iteration:**
For this last iteration, we have $X = \{h\}$, and we have $Bel(\{h\}) = p_h = 0.02$. With this, we arrive at $X = \emptyset$; so, we can now calculate $S^*(Bel)$.
Now, we proceed to the calculation of the maximum entropy:

$$\begin{aligned}
 S^*(Bel) &= - \sum_{i \in X} p_i \log_2 p_i \\
 &= - 0.23 \log_2(0.23) - 0.21 \log_2(0.21) - 0.18 \log_2(0.18) \\
 &\quad - 0.13 \log_2(0.13) - 0.1 \log_2(0.1) - 0.08 \log_2(0.08) \\
 &\quad - 0.05 \log_2(0.05) - 0.02 \log_2(0.02) \\
 &= 2.74112.
 \end{aligned}
 \tag{29}$$

• **Improvement of Meyerowitz et al.'s algorithm**

- **First iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{a\}$. Therefore, we have $p_a = 0.23$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{h\}$, and we have $p_h = 0.02$. We update the value of $acu = 0.75$, the function Bel , and the function Pl , and since $X = \{b, c, d, e, f, g\}$, and there are sets whose function Bel is nonzero, we need a second iteration.
- **Second iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is attained for $A = \{b\}$. So, it is satisfied that $p_b = 0.21$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{g\}$, and we have that $p_g = 0.05$. We update the value of $acu = 0.49$, the function Bel , and the function Pl , and as

$X = \{c, d, e, f\}$, and there are sets whose function Bel is nonzero, we need a third iteration.

- **Third iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{c\}$. Thus, we have $p_c = 0.18$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{f\}$, and we have that $p_f = 0.08$. We update the value of $acu = 0.23$, the function Bel , and the function Pl , and as $X = \{d, e\}$, and there are sets whose function Bel is nonzero, we need a fourth iteration.

- **Fourth iteration:** The maximum of the function $\frac{Bel(A)}{|A|} \forall A \subseteq X$ is reached for $A = \{d\}$. Thus, we have $p_d = 0.13$. Now, we see the minimum value of $\frac{Pl(B)}{|B|} \forall B \subseteq X$, which in this case is $B = \{e\}$, and we have $p_e = 0.1$. With this we arrive at $X = \emptyset$. Consequently, we can now calculate $S^*(Bel)$.

Now, we proceed to the calculation of the maximum entropy:

$$\begin{aligned}
 S^*(Bel) &= - \sum_{i \in X} p_i \log_2 p_i \\
 &= - 0.23 \log_2(0.23) - 0.21 \log_2(0.21) - 0.18 \log_2(0.18) \\
 &\quad - 0.13 \log_2(0.13) - 0.1 \log_2(0.1) - 0.08 \log_2(0.08) \\
 &\quad - 0.05 \log_2(0.05) - 0.02 \log_2(0.02) \\
 &= 2.74112.
 \end{aligned}
 \tag{30}$$

Again, in this example we needed the middle steps to find the maximum of entropy. In this case, we used a b.p.a. with conflict. Clearly, the greater the $|X|$, the greater the improvement.

3.5. Experiments

We carried out a series of experiments generating b.p.a.s on sets of $n = 4, n = 5$, and $n = 6$ elements (sizes of the frames of discernment), to examine the time of processing of both algorithms: the original and the improved ones. We implemented both algorithms in C++ programming language and ran them in a computer with an Intel Core i5 processor, a CPU of 1.8 GHz, and 8 G of RAM. Each b.p.a. was randomly generated with assigned mass values with the constraint that none of them was above $0.5/n$. This characteristic ensured we could obtain b.p.a.s with a good level of sharing. Obviously, if a b.p.a. has masses only on the singleton sets or only it is focused on one set, both algorithms are immediate and similar, but these situations do not appear with the constraint used. We think that, with this constraint, the differences of both algorithms can be seen in a better way. The results can be seen in Table 1.

Table 1. Time of processing (seconds) of the calculus of both algorithms on 10 million b.p.a.s randomly generated on sets of 4, 5, and 6 elements.

	Original Alg.	Improved Alg.	Percentage of Improvement
$n = 4$	29.10	26.69	8.28%
$n = 5$	120.05	101.20	15.70%
$n = 6$	1077.24	868.66	19.36%

As we can see in Table 1, the reduction in time increases with the number of values. We had about an 8% improvement on sets of four elements, more than a 15% improvement on sets of five elements, and a more than 19% improvement on sets of six elements. As we showed in the previous examples, the level of improvement increased with the number of elements. The increment in the performance was duplicated from $n = 4$ to $n = 6$, with $n = 6$ very close to 20%, which shows the importance of the new algorithm when the size of the frame of discernment increases.

As expected, we observed that the time of processing increased strongly with the size of the frame of discernment, which is a characteristic of algorithms that need to enumerate all the elements of the powerset of a universal set. However, we should not forget that the number of the first two columns correspond to seconds taken for the application of the algorithms on 10 million b.p.a.s on each set of size n . Hence, the average time of the calculus for obtaining the maximum entropy for one b.p.a. can be achieved dividing those values (from columns 1 and 2) by 10 million.

4. Conclusions and Future Work

The computational cost of Meyerowitz et al.'s algorithm has been the principal drawback of using the maximum entropy as a measure to quantify the uncertainty-based information in TE. In this work, a variation of that algorithm was presented. The key point of this new proposal is the double use of each enumeration of the subsets of the power set of the frame of discernment. This new procedure implies an important reduction in the set of elements in the resulting frame of discernment of each step in Meyerowitz et al.'s algorithm. Hence, the number of steps necessary to achieve the maximum entropy of a b.p.a. is reduced. The experiments carried out showed that the improvement in the time can be close to 20% for a set with a cardinality of six elements, showing that this improvement can be greater when we increase the size of the frame of discernment. The outcome presented may have greater applicability in TE. As future work and application of the algorithm presented here, we want to apply it in some real areas, such as the one based on the information obtained by sensors, where TE has been recently widely used.

We believe that it is possible to continue lowering the complexity of the calculation of the maximum entropy via new algorithms. Obtaining these new algorithms with better computational behavior and their comparison with those used in this work are part of our future work in this area of research.

Author Contributions: Formal analysis, A.P.-L. and J.A.; Funding acquisition, J.A.; Investigation, A.P.-L. and J.A.; Methodology, J.A. and A.P.-L.; Project administration, J.A. and A.P.-L.; Supervision, J.A.; Writing—original draft, A.P.-L., S.M.-G. and J.A.; Writing—review and editing, A.P.-L., S.M.-G. and J.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by UGR-FEDER funds under Project A-TIC-344-UGR20; and by the “FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades” under Project P20_00159.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339. [[CrossRef](#)]
2. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976.
3. Beynon, M.; Curry, B.; Morgan, P. The Dempster–Shafer theory of evidence: An alternative approach to multicriteria decision modelling. *Omega* **2000**, *28*, 37–50. [[CrossRef](#)]
4. Dencœur, T. A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*; Yager, R.R., Liu, L., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 737–760.
5. Buede, D.M.; Girardi, P. A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **1997**, *27*, 569–577. [[CrossRef](#)]
6. Ip, H.H.S.; Ng, J.M.C. Human face recognition using Dempster-Shafer theory. In Proceedings of the 1st International Conference on Image Processing, Austin, TX, USA, 13–16 November 1994; Volume 2, pp. 292–295. [[CrossRef](#)]
7. Zheng, H.; Tang, Y. Deng Entropy Weighted Risk Priority Number Model for Failure Mode and Effects Analysis. *Entropy* **2020**, *22*, 280. [[CrossRef](#)]
8. Tang, Y.; Tan, S.; Zhou, D. An Improved Failure Mode and Effects Analysis Method Using Belief Jensen–Shannon Divergence and Entropy Measure in the Evidence Theory. *Arab. J. Sci. Eng.* **2022**, *48*, 7163–7176. [[CrossRef](#)]

9. Basir, O.; Yuan, X. Engine fault diagnosis based on multi-sensor information fusion using Dempster-Shafer evidence theory. *Inf. Fusion* **2007**, *8*, 379–386. [[CrossRef](#)]
10. Frittella, S.; Manoorkar, K.; Palmigiano, A.; Tzimoulis, A.; Wijnberg, N. Toward a Dempster-Shafer theory of concepts. *Int. J. Approx. Reason.* **2020**, *125*, 14–25. [[CrossRef](#)]
11. Chen, T.M.; Venkataramanan, V. Dempster-Shafer theory for intrusion detection in ad hoc networks. *IEEE Internet Comput.* **2005**, *9*, 35–41. [[CrossRef](#)]
12. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
13. Yager, R.R. Entropy and specificity in a mathematical theory of evidence. *Int. J. Gen. Syst.* **1983**, *9*, 249–260. [[CrossRef](#)]
14. Klir, G.; Wierman, M. *Uncertainty-Based Information: Elements of Generalized Information Theory*; Studies in Fuzziness and Soft Computing; Physica-Verlag HD : Heidelberg, Germany, 1999.
15. Abellán, J.; Masegosa, A. Requirements for total uncertainty measures in Dempster-Shafer theory of evidence. *Int. J. Gen. Syst.* **2008**, *37*, 733–747. [[CrossRef](#)]
16. Harmanec, D.; Klir, G.J. Measuring total uncertainty in Dempster-Shafer Theory: A novel approach. *Int. J. Gen. Syst.* **1994**, *22*, 405–419. [[CrossRef](#)]
17. Meyerowitz, A.; Richman, F.; Walker, E. Calculating maximum-entropy probability densities for belief functions. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **1994**, *2*, 377–389. [[CrossRef](#)]
18. Deng, Y. Deng entropy. *Chaos Solitons Fractals* **2016**, *91*, 549–553. [[CrossRef](#)]
19. Abellán, J. Analyzing properties of Deng entropy in the theory of evidence. *Chaos Solitons Fractals* **2017**, *95*, 195–199. [[CrossRef](#)]
20. Abellán, J.; Bossé, É. Critique of Recent Uncertainty Measures Developed Under the Evidence Theory and Belief Intervals. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *50*, 1186–1192. [[CrossRef](#)]
21. Abellán, J.; Moral, S. Building classification trees using the total uncertainty criterion. *Int. J. Intell. Syst.* **2003**, *18*, 1215–1225. [[CrossRef](#)]
22. Abellán, J. Ensembles of decision trees based on imprecise probabilities and uncertainty measures. *Inf. Fusion* **2013**, *14*, 423–430. [[CrossRef](#)]
23. Abellán, J.; Mantas, C.J.; Castellano, J.G. AdaptiveCC4.5: Credal C4.5 with a rough class noise estimator. *Expert Syst. Appl.* **2018**, *92*, 363–379. [[CrossRef](#)]
24. Moral-García, S.; Mantas, C.J.; Castellano, J.G.; Benítez, M.D.; Abellán, J. Bagging of credal decision trees for imprecise classification. *Expert Syst. Appl.* **2020**, *141*, 112944. [[CrossRef](#)]
25. Moral-García, S.; Abellán, J. Maximum of Entropy for Belief Intervals Under Evidence Theory. *IEEE Access* **2020**, *8*, 118017–118029. [[CrossRef](#)]
26. Hartley, R.V.L. Transmission of Information1. *Bell Syst. Tech. J.* **1928**, *7*, 535–563. . [[CrossRef](#)]
27. Dubois, D.; Prade, H. A note on measures of specificity for fuzzy sets. *Int. J. Gen. Syst.* **1985**, *10*, 279–283. [[CrossRef](#)]
28. Abellán, J.; Moral, S. A Non-specificity measure for convex sets of probability distributions. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2000**, *8*, 357–367. [[CrossRef](#)]
29. Cui, H.; Liu, Q.; Zhang, J.; Kang, B. An Improved Deng Entropy and Its Application in Pattern Recognition. *IEEE Access* **2019**, *7*, 18284–18292. [[CrossRef](#)]
30. Kang, B.; Deng, Y. The Maximum Deng Entropy. *IEEE Access* **2019**, *7*, 120758–120765. [[CrossRef](#)]
31. Zhu, R.; Chen, J.; Kang, B. Power Law and Dimension of the Maximum Value for Belief Distribution With the Maximum Deng Entropy. *IEEE Access* **2020**, *8*, 47713–47719. [[CrossRef](#)]
32. Voorbraak, F. A computationally efficient approximation of Dempster-Shafer theory. *Int. J. Man-Mach. Stud.* **1989**, *30*, 525–536. . [[CrossRef](#)]
33. 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](#)]
34. Jirousek, 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](#)]
35. Pan, Q.; Zhou, D.; Tang, Y.; Li, X.; Huang, J. A Novel Belief Entropy for Measuring Uncertainty in Dempster-Shafer Evidence Theory Framework Based on Plausibility Transformation and Weighted Hartley Entropy. *Entropy* **2019**, *21*, 163. [[CrossRef](#)] [[PubMed](#)]
36. Zhao, Y.; Ji, D.; Yang, X.; Fei, L.; Zhai, C. An Improved Belief Entropy to Measure Uncertainty of Basic Probability Assignments Based on Deng Entropy and Belief Interval. *Entropy* **2019**, *21*, 1122. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.