

Article

A Comparison of Algorithms to Achieve the Maximum Entropy in the Theory of Evidence

Joaquín Abellán *, Aina López-Gay, Maria Isabel A. Benítez and Francisco Javier G. Castellano 

Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain; ainalopez@correo.ugr.es (A.L.-G.); miab@correo.ugr.es (M.I.A.B.); fjc@decsai.ugr.es (F.J.G.C.)

* Correspondence: jabellan@decsai.ugr.es

Abstract

Within the framework of evidence theory, maximum entropy is regarded as a measure of total uncertainty that satisfies a comprehensive set of mathematical properties and behavioral requirements. However, its practical applicability is severely questioned due to the high computational complexity of its calculation, which involves the manipulation of the power set of the frame of discernment. In the literature, attempts have been made to reduce this complexity by restricting the computation to singleton elements, leading to a formulation based on reachable probability intervals. Although this approach relies on a less specific representation of evidential information, it has been shown to provide an equivalent maximum entropy value under certain conditions. In this paper, we present an experimental comparative study of two algorithms for calculating maximum entropy in evidence theory: the classical algorithm, which operates directly on belief functions, and an alternative algorithm based on reachable probability intervals. Through numerical experiments, we demonstrate that the differences between these approaches are less pronounced than previously suggested in the literature. Depending on the type of information representations to which it is applied, the original algorithm based on belief functions can be more efficient than the one using the reachable probability interval approach. This is an interesting result, and a reason for choosing one algorithm over the other depending on the situation.

Keywords: evidence theory; reachable probability intervals; uncertainty measures; maximum of entropy



Academic Editors: Nikolai Leonenko, Brian Dennis, Mark L. Taper and Jose Miguel Ponciano

Received: 30 December 2025

Revised: 30 January 2026

Accepted: 15 February 2026

Published: 21 February 2026

Copyright: © 2026 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

The representation and quantification of uncertainty are central issues in information theory, artificial intelligence, and decision-making under incomplete or imprecise knowledge. Classical probability theory (PT) provides a rigorous and widely accepted framework for uncertainty modeling; however, it requires precise probability assignments that are often unavailable in real-world applications. This limitation has motivated the development of more general models of uncertainty, commonly referred to as theories based on imprecise probabilities [1].

Among these frameworks, evidence theory (ET), also known as Dempster–Shafer theory [2,3], has been extensively employed to manage uncertainty-based information in practical applications such as medical diagnosis [4], statistical classification [5], target identification [6], and face recognition [7]. Evidence theory extends PT by introducing the concept of a *basic probability assignment* (BPA), which generalizes the notion of a probability distribution. Each BPA induces an associated belief function and a plausibility function,

where the belief (respectively, plausibility) of a set represents the minimum (respectively, maximum) degree of support that the available evidence provides for that set.

Quantifying the uncertainty represented by a BPA is a fundamental problem in evidence theory. To this end, numerous uncertainty measures have been proposed, most of which are inspired by Shannon entropy [8], the standard measure of uncertainty in PT. However, extending Shannon entropy to evidence theory is non-trivial, as ET accounts for additional types of uncertainty not present in classical probability models.

As pointed out by Yager [9], two different types of uncertainty arise in evidence theory: *conflict*, which occurs when information supports disjoint sets, and *non-specificity*, which arises when information is assigned to sets with cardinality greater than one. Consequently, any uncertainty measure in ET must be able to jointly capture both conflict and non-specificity in a coherent manner.

Klir and Wierman [10] analyzed the mathematical properties that uncertainty measures in evidence theory should satisfy; this study was later extended by Abellán and Masegosa [11], who also introduced behavioral requirements for such measures. Among all proposals to date, the maximum entropy defined on the closed and convex set of probability distributions (credal set) compatible with a BPA [10] is the only uncertainty measure in evidence theory that satisfies all crucial mathematical properties and behavioral requirements simultaneously.

The Maximum Entropy Principle [12] is applicable in various domains, being principally related to information theory and applications [13–15]. It states that, in the absence of complete information, one should prioritize the most unbiased distribution by maximizing uncertainty, thereby ensuring that no groundless assumptions are introduced into the model. In the fields of evidence theory and uncertainty quantification [16–20], this principle and its associated uncertainty measures are essential for constructing belief distributions that reflect incomplete or ambiguous information.

Despite its strong axiomatic foundations, the practical computation of maximum entropy in evidence theory remains challenging. The algorithm proposed in [21] (also presented in [19]) involves solving constrained optimization problems whose complexity grows exponentially with the size of the frame of discernment. Consequently, numerous alternative uncertainty measures with lower computational complexity have been proposed in recent years. However, these alternatives often fail to satisfy all the required axiomatic properties and behavioral conditions [22–24], contributing to a lack of consensus regarding uncertainty measures in ET. In summary, while maximum entropy exhibits optimal axiomatic behavior, it entails a high computational cost; conversely, alternative measures are computationally efficient but theoretically weaker.

On the other hand, maximum entropy has demonstrated excellent performance in practical applications, particularly in data mining and related fields. For specific classes of belief functions, its computation becomes straightforward and can be executed efficiently, as illustrated in [15,25,26].

An approximation of the maximum entropy of the credal set associated with a BPA was proposed in [27]. This approach computes the maximum entropy over the credal set consistent with belief intervals for singleton elements, where the lower and upper bounds correspond to belief and plausibility values, respectively. Although this measure satisfies all crucial mathematical properties and behavioral requirements, utilizing belief intervals for singletons instead of the full BPA may lead to information loss. Indeed, the credal set associated with a BPA is always contained within the credal set that is compatible with the corresponding belief intervals [27], resulting in an uncertainty measure that may overestimate the level of uncertainty represented by the original BPA. However, this method offers the advantage of bypassing the enumeration of all subsets of the frame of discernment. A key question remains regarding whether this computational advantage persists across all possible scenarios.

The aim of this paper is to provide a comparative study of two algorithms for computing maximum entropy in the theory of evidence: the Meyerowitz et al. algorithm [19,21], which operates directly on belief functions; and the maximum entropy algorithm based on reachable probability intervals [18,27], derived from evidential constraints. The comparison addresses both theoretical aspects and numerical behavior, highlighting the advantages and limitations of each approach under various uncertainty scenarios.

The remainder of this paper is organized as follows. Section 2 introduces the basic concepts of evidence theory, reviews the main uncertainty measures proposed in this framework, and describes the algorithms for computing maximum entropy from a BPA and from belief intervals for singletons. Section 3 presents the comparative study via numerical examples and experiments in which millions of different belief functions are randomly generated. Finally, concluding remarks and directions for future work are provided in Section 4.

2. Background

Let $X = \{x_1, \dots, x_n\}$ 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 (ET), also known as Dempster–Shafer theory [2,3], is based on the concept of a *basic probability assignment* (BPA). A BPA is a mapping $m : \wp(X) \rightarrow [0, 1]$ satisfying $m(\emptyset) = 0$ and $\sum_{A \subseteq X} m(A) = 1$.

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

A given BPA m on X has an associated *belief function* Bel_m and a *plausibility function* Pl_m . These functions are defined as follows:

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

It should be noted that for each $A \subseteq X$, $Bel_m(A) \leq Pl_m(A)$. The interval $[Bel_m(A), Pl_m(A)]$ is referred to as the *belief interval* of A . Furthermore,

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

where \bar{A} denotes the complement of A . Consequently, Bel_m and Pl_m are considered dual or conjugate functions. Either function is sufficient to represent uncertainty-based information in ET; for this purpose, Bel_m is more commonly utilized.

For a given BPA m on X , the set of compatible probability distributions (which corresponds to a closed and convex set, also known as a *credal set*) is defined as follows:

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

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

2.2. Uncertainty Measures in Evidence Theory

Shannon entropy [8] is the standard uncertainty measure in probability theory. Given a probability distribution p defined on a finite set $X = \{x_1, \dots, x_n\}$, Shannon entropy is defined as follows:

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

The type of uncertainty quantified by S is usually referred to as *conflict*, which is the only form of uncertainty present in classical probability theory. Shannon entropy satisfies a well-known set of desirable axiomatic properties [8,10].

In possibility theory, uncertainty is commonly quantified by the Hartley measure [28], defined as follows:

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

This measure captures *non-specificity*, the sole type of uncertainty considered in possibility theory.

As noted by Yager [9], conflict and non-specificity coexist in ET. Conflict arises when information supports mutually exclusive subsets, whereas non-specificity appears when information is assigned to sets with cardinality greater than one. A generalization of the Hartley measure to evidence theory was introduced by Dubois and Prade [29]:

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

GH attains its minimum value (zero) when m is a probability distribution, and its maximum value ($\log_2 n$) when $m(X) = 1$. It constitutes an appropriate measure of non-specificity in ET and can be naturally extended to more general uncertainty frameworks [18].

Numerous attempts have been made to generalize Shannon entropy to evidence theory; however, most proposals fail to satisfy the essential mathematical and behavioral requirements for this framework. A total uncertainty measure capable of jointly capturing conflict and non-specificity was proposed by Harmanec and Klir [19]. This measure, denoted by $S^*(\mathcal{P}_m)$, is defined as the maximum Shannon entropy over the credal set \mathcal{P}_m associated with a BPA m :

$$S^*(\mathcal{P}_m) = \max_{p \in \mathcal{P}_m} S(p). \tag{7}$$

To date, this is the only measure that satisfies all necessary mathematical properties and behavioral requirements for uncertainty measures in evidence theory [19,26].

Despite its strong foundations, computing $S^*(\mathcal{P}_m)$ is computationally demanding. Algorithms proposed in the literature [19,21,30,31] involve solving nonlinear optimization problems with exponential complexity. Consequently, several alternative measures with lower computational costs have been proposed.

One well-known alternative is Deng entropy [22,32–34], defined as follows:

$$E_d(m) = - \sum_{A \subseteq X} m(A) \log_2 \left(\frac{m(A)}{2^{|A|} - 1} \right). \tag{8}$$

In this formulation, the expression can be decomposed into two components: one capturing non-specificity and the other quantifying conflict. However, Deng entropy violates several essential mathematical properties and exhibits problematic behavior in various scenarios [23]. Similarly, Pan [35] introduced an uncertainty measure based on the plausibility transformation:

$$H_{PQ}(m) = - \sum_{A \subseteq X} m(A) \log_2 Pm(A) + GH(m), \tag{9}$$

where $Pm(A) = \sum_{x \in A} Pt(x)$, and $Pt(x)$ is the *plausibility transformation* value. As shown in [27], H_{PQ} also fails to satisfy all required mathematical properties.

Let us now consider the set of belief intervals for singletons elements associated with a BPA m :

$$\mathcal{I}_m = \{ [Bel_m(\{x_i\}), Pl_m(\{x_i\})] \mid i = 1, \dots, n \}. \tag{10}$$

Zhao et al. [36] proposed a measure (H_{inter}) combining Deng entropy with these intervals. Nevertheless, it does not satisfy all crucial mathematical properties required in ET [27].

Let $\mathcal{P}(\mathcal{I}_m)$ denote the credal set consistent with the belief intervals for singleton elements:

$$\mathcal{P}(\mathcal{I}_m) = \{ p \in \mathcal{P}(X) \mid Bel_m(\{x_i\}) \leq p(\{x_i\}) \leq Pl_m(\{x_i\}), i = 1, \dots, n \}. \tag{11}$$

In [27], a new uncertainty measure was proposed as the maximum entropy over $\mathcal{P}(\mathcal{I}_m)$:

$$S^*(\mathcal{P}(\mathcal{I}_m)) = \max_{p \in \mathcal{P}(\mathcal{I}_m)} S(p). \tag{12}$$

This measure satisfies all essential requirements [27]. However, since $\mathcal{P}_m \subseteq \mathcal{P}(\mathcal{I}_m)$, representing uncertainty through singleton belief intervals may result in information loss. Consequently, $S^*(\mathcal{P}(\mathcal{I}_m))$ may indicate a higher level of uncertainty than the original BPA. Its primary advantage lies in the significant reduction of computational complexity, albeit at the cost of potential information loss.

2.3. Maximum Entropy from a Belief Function

The maximum entropy associated with a belief function is obtained by solving a constrained optimization problem over the credal set induced by a basic probability assignment. An exact procedure for this task was proposed by Meyerowitz et al. [21] and subsequently by Harmanec and Klir [19]. The algorithm iteratively constructs the probability distribution that maximizes Shannon entropy while satisfying the evidential constraints. Algorithm 1 calculates the maximum entropy of a BPA m with associated $Bel(\cdot)$ function. The procedure is described as follows:

Algorithm 1 Algorithm to attain the maximum entropy from a Belief function

```

X ← current frame of discernment
Bel ← associated belief function
while X ≠ ∅ and Bel(X) ≠ 0 do
  Select a non-empty subset A ⊆ X maximizing  $\frac{Bel(A)}{|A|}$ 
  If multiple subsets satisfy the condition, select the one with maximum cardinality
  for x ∈ A do
    Assign probability  $\frac{Bel(A)}{|A|}$  to x
  end for
  for B ⊆ X \ A do
    Bel'(B) ← Bel(B ∪ A) − Bel(A)
  end for
  X ← X \ A
  Bel ← Bel'
end while
if X ≠ ∅ then
  for x ∈ X do
    Assign probability 0 to x
  end for
end if

```

The resulting probability distribution maximizes Shannon entropy under the constraints induced by the original belief function. Although exact, this algorithm requires evaluating all subsets of the frame of discernment at each iteration, resulting in exponential computational complexity.

2.4. Maximum Entropy from Reachable Probability Intervals

An alternative approach to computing maximum entropy is based on the reachable probability intervals [1] derived from the belief and plausibility values of singleton elements. Let

$$\mathcal{I} = \{[l_i, u_i]\}_{i=1}^n$$

denote the set of such intervals, where $l_i = Bel_m(\{x_i\})$ and $u_i = Pl_m(\{x_i\})$.

This algorithm builds upon the work in [18] to obtain the maximum entropy on the credal set associated with a reachable set of probability intervals. We introduce the following notation:

- $Min(p, In_x)$: the minimum value of the probability distribution p among the components whose indices belong to the set In_x .
- $Sig(p, In_x)$: the second smallest value of the probability distribution p among the components in In_x . If no such value exists, $Sig(p, In_x) = -1$.
- $N_{min}(p, In_x)$: the number of indices in In_x that attain the minimum value of the probability distribution p .
- $Min(a, b, c)$: the minimum value among the real numbers $\{a, b, c\}$.

The following procedure (Algorithm 2) yields the probability distribution \hat{p} that attains the maximum entropy on $\mathcal{P}(\mathcal{I}_m)$:

Algorithm 2 Algorithm to attain the maximum entropy from reachable probability intervals

```

index_set ← {1, 2, ..., n}
for i = 1 to n do
     $\hat{p}(\{x_i\}) \leftarrow Bel_m(\{x_i\})$ 
end for
mass ←  $1 - \sum_{i=1}^n Bel_m(\{x_i\})$ 
while mass > 0 do
    for i ∈ index_set do
        if  $\hat{p}(\{x_i\}) = Pl_m(\{x_i\})$  then
            index_set ← index_set \ {i}
        end if
    end for
    min ←  $Min(\hat{p}, index\_set)$ 
    second ←  $Sig(\hat{p}, index\_set)$ 
    m ←  $N_{min}(\hat{p}, index\_set)$ 
    for i ∈ index_set do
        if  $\hat{p}(\{x_i\}) = Min(\hat{p}, index\_set)$  then
            if second = -1 then
                 $\hat{p}(\{x_i\}) \leftarrow \hat{p}(\{x_i\}) + Min(Pl(\{x_i\}) - \hat{p}(\{x_i\}), \frac{mass}{m}, 1)$ 
                mass ← mass -  $Min(Pl(\{x_i\}) - \hat{p}(\{x_i\}), \frac{mass}{m}, 1)$ 
            else
                 $\hat{p}(\{x_i\}) \leftarrow \hat{p}(\{x_i\}) + Min(Pl(\{x_i\}) - \hat{p}(\{x_i\}), second - \hat{p}(\{x_i\}), \frac{mass}{m})$ 
                mass ← mass -  $Min(Pl(\{x_i\}) - \hat{p}(\{x_i\}), second - \hat{p}(\{x_i\}), \frac{mass}{m})$ 
            end if
        end if
    end for
end while

```

Unlike the belief-function-based algorithm, this approach exhibits polynomial computational complexity relative to the number of singleton elements.

2.5. Discussion

The two algorithms analyzed in this work address the same conceptual objective—computing maximum entropy within the framework of evidence theory—but rely on fundamentally different representations of uncertainty. Meyerowitz et al.’s algorithm operates directly on belief functions and computes the exact maximum entropy associated with the credal set induced by a BPA. In contrast, the interval-based algorithm computes maximum entropy over a larger credal set defined solely by the belief and plausibility values of singleton elements.

From a theoretical perspective, the Meyerowitz et al. algorithm preserves the full informational content of the belief function, providing an exact characterization of total uncertainty. However, this precision entails high computational complexity that grows exponentially with the size of the frame of discernment. Consequently, its applicability is primarily limited to problems of moderate dimensionality.

The interval-based algorithm significantly reduces computational overhead by restricting the optimization problem to singleton constraints. Although this approach may result in information loss—since the credal set induced by singleton intervals contains only the one associated with the original belief function—the experimental results demonstrate that the resulting maximum entropy values are often remarkably close or even identical to those obtained by the exact algorithm. This suggests that, in many practical scenarios, the loss of specificity introduced by interval representations has a negligible impact on the total uncertainty measure.

Contrary to claims in previous studies, our analysis indicates that the discrepancy between the two approaches is not as pronounced as might be expected. The interval-based method provides a reliable approximation of maximum entropy while offering substantial computational advantages. These findings underscore the importance of balancing representational accuracy with computational feasibility when selecting algorithms for evidence-based models.

3. Comparison of Algorithms

Our principal aim in this work is to compare the performance of the maximum entropy calculation algorithms.

We will introduce several examples and apply the respective algorithms to them in order to compare them. The objective of this comparison is to see how both algorithms work in a specific situation when both are applied within the TE. To do this, we will first study them by applying Meyerowitz et al.’s algorithm, specific to the theory of evidence, with the algorithm adapted to reachable intervals, taking only the singletons. We remark that the calculus of the values of $Bel()$, $Pl()$, l_i and u_i is based on the expressions in the Background section.

The first two examples are free of conflict, while the following two have conflict, allowing us to see how this situation could affect the effectiveness of the algorithms. To simplify, *Algorithm-1* is the algorithm of Meyerowitz et al., and *Algorithm-2* is the algorithm based on the reachable probability intervals of the singletons.

Example 1. Let us start with the set $X = \{a, b, c, d, e\}$, whose mass function m is given by

$$m(\{a, b\}) = 0.4$$

$$m(\{c, d\}) = 0.35$$

$$m(\{e\}) = 0.25$$

We see that subsets $\{a, b\}$, $\{c, d\}$, and $\{e\}$ of X are mutually disjointed, so there is no conflict. With this, we will proceed to apply the algorithms presented above.

1. **Algorithm-1.**

We first construct Table 1 with the values of the belief function for each subset $A \subset X$ and the value of $\frac{Bel(A)}{|A|}$.

Table 1. Values of $Bel(A)$ and $\frac{Bel(A)}{|A|}$.

A	Bel(A)	$\frac{Bel(A)}{ A }$
{a}	0	0
{b}	0	0
{c}	0	0
{d}	0	0
{e}	0.25	0.25
{a,b}	0.40	0.20
{a,c}	0	0
{a,d}	0	0
{a,e}	0.25	0.125
{b,c}	0	0
{b,d}	0	0
{b,e}	0.25	0.125
{c,d}	0.35	0.175
{c,e}	0.25	0.125
{d,e}	0.25	0.125
{a,b,c}	0.40	0.13
{a,b,d}	0.40	0.13
{a,b,e}	0.65	0.216
{a,c,d}	0.35	0.116
{a,c,e}	0.25	0.083
{a,d,e}	0.25	0.083
{b,c,d}	0.35	0.116
{b,c,e}	0.25	0.083
{b,d,e}	0.25	0.083
{c,d,e}	0.60	0.2
{a,b,c,d}	0.75	0.1875
{a,b,c,e}	0.65	0.1625
{a,b,d,e}	0.65	0.1625
{a,c,d,e}	0.60	0.15
{b,c,d,e}	0.6	0.15
{a,b,c,d,e}	1	0.20

- *First iteration.*

We observe that the maximum value of $\frac{Bel(A)}{|A|}$ is for $A = \{e\}$. With this, we have $p_e = 0.25$, and we can make modifications to the belief functions so that we obtain

$$Bel(\{a,b\}) \leftarrow Bel(\{a,b,e\}) - Bel(\{e\}) = 0.65 - 0.25 = 0.4,$$

$$Bel(\{a,b,c,d\}) \leftarrow Bel(\{a,b,c,d,e\}) - Bel(\{e\}) = 1 - 0.25 = 0.75.$$

Now, we take $X \leftarrow \{a,b,c,d,e\} - \{e\} = \{a,b,c,d\}$, $X \neq \emptyset$ and $Bel(X) \neq \emptyset$, performing a new iteration.

- *Second iteration.*

We start from $X = \{a,b,c,d\}$, so we have the values seen in Table 2:

Table 2. Values of $Bel(A)$ and $\frac{Bel(A)}{|A|}$ in the second iteration.

A	Bel(A)	$\frac{Bel(A)}{ A }$
{a,b}	0.40	0.20
{c,d}	0.35	0.175
{a,b,c}	0.40	0.13
{a,b,d}	0.40	0.13
{a,c,d}	0.35	0.116
{b,c,d}	0.35	0.116
{a,b,c,d}	0.75	0.1875

In this case, the maximum of $\frac{Bel(A)}{|A|}$ is reached at $\frac{Bel(\{a,b\})}{|\{a,b\}|}$, so we assign $p_a = p_b = 0.20$. Moving to the next step of the algorithm, $X \leftarrow \{a, b, c, d\} - \{a, b\} = \{c, d\} \neq \emptyset$, and

$$Bel(\{c, d\}) \leftarrow Bel(\{a, b, c, d\}) - Bel(\{a, b\}) = 0.75 - 0.40 = 0.35 \neq \emptyset,$$

so a new iteration begins.

- Third iteration.

Given that $X = \{c, d\}$, the only possible non-zero value is that shown in Table 3:

Table 3. Values of $Bel(A)$ and $\frac{Bel(A)}{|A|}$ in the third iteration.

A	Bel(A)	$\frac{Bel(A)}{ A }$
{c, d}	0.35	0.175

Thus, the maximum of $\frac{Bel(A)}{|A|}$ is $A = \{c, d\}$. Thus, $p_c = p_d = 0.175$ and we obtain the new $X = \{c, d\} - \{c, d\} = \emptyset$, with $Bel(\emptyset) = 0$, so we can proceed to calculate the maximum entropy.

$$S^*(Bel) = - \sum_{i \in \{a,b,c,d,e\}} p_i \log_2(p_i) = -(2 \cdot 0.2 \cdot \log_2(0.2) + 2 \cdot 0.175 \cdot \log_2(0.175) + 0.25 \cdot \log_2(0.25)) = \boxed{2.308872}$$

2. Algorithm-2

To apply this algorithm, we need to calculate the probability intervals of the singletons. Calculating the belief and plausibility functions associated with each element of $\{a, b, c, d, e\}$, we obtain the values shown in Table 4:

Table 4. Belief intervals.

x_i	Bel(x_i)	Pl(x_i)	[Bel(x_i), Pl(x_i)]
{a}	0	0.40	[0, 0.40]
{b}	0	0.40	[0, 0.40]
{c}	0	0.35	[0, 0.35]
{d}	0	0.35	[0, 0.35]
{e}	0.25	0.25	[0.25, 0.25]

Therefore, we start with the following set of probability intervals

$$L = \{[0, 0.4], [0, 0.4], [0, 0.35], [0, 0.35], [0.25, 0.25]\}$$

where:

$$l = [0, 0, 0, 0, 0.25], u = [0.40, 0.40, 0.35, 0.35, 0.25]$$

and \hat{p} is the probability vector for which we will calculate the maximum entropy.

- First iteration.

To begin, we initialize with $S = \{1, 2, 3, 4, 5\}$, where 1 corresponds to singleton a, 2 to singleton b, etc. We construct the vector \hat{p} with the lower bound values l_i for each $i \in \{1, 2, 3, 4, 5\}$:

$$\hat{p} = (0, 0, 0, 0, 0.25)$$

and we see that $Sum(\hat{p}) = 0.25 < 1$. We check if $l_i = u_i$ for some $i = 1, \dots, 5$, and we see that it holds for $i = 5$, so we remove it from S , obtaining $S = \{1, 2, 3, 4\}$. We calculate the following values:

- * $s = \text{Sum}(l_i) = 0.25,$
- * $r = \text{Min}(l, S) = 1,$
- * $m = \text{Nmin}(l, S) = 4,$
- * $\text{Sig}(l, S) = -1,$

hence, using the algorithm, we carry out the assignment

$$l_i \leftarrow l_i + \min \left(u_i - l_i, \frac{1-s}{m}, 1 \right) = l_i + 0.1875,$$

updating \hat{p}_i for each $i = 1, \dots, 5,$ so we have

$$\hat{p} = [0.1875, 0.1875, 0.1875, 0.1875, 0.25].$$

- Second iteration.

We start from

$$\hat{p} = (0.1875, 0.1875, 0.1875, 0.1875, 0.25),$$

with $\text{Sum}(\hat{p}) = 1,$ so the algorithm ends and we proceed to calculate the maximum entropy associated with the given distribution:

$$S^*(\text{Bel}) = -(4 \cdot 0.1875) \log_2(0.1875) + 0.25 \cdot \log_2(0.25) = \boxed{2.311278}$$

Example 2. We consider the set $X = \{a, b, c, d, e, f\}$ with a given mass function m defined by

$$\begin{aligned} m(\{a\}) &= 0.1 \\ m(\{b, c\}) &= 0.4 \\ m(\{d, f\}) &= 0.35 \\ m(\{e\}) &= 0.15 \end{aligned}$$

1. **Algorithm-1**

We have the values for the belief function and $\frac{\text{Bel}(A)}{|A|}$ for each subset $A \subseteq X,$ as shown in Table 5:

Table 5. Values of $\text{Bel}(A)$ and $\frac{\text{Bel}(A)}{|A|}$.

A	Bel(A)	$\frac{\text{Bel}(A)}{ A }$
{a}	0.10	0.10
{e}	0.15	0.15
{a, b}	0.10	0.05
{a, c}	0.10	0.05
{a, d}	0.10	0.05
{a, e}	0.25	0.125
{a, f}	0.10	0.05
{b, c}	0.40	0.20
{b, e}	0.15	0.075
{c, e}	0.15	0.075
{d, e}	0.15	0.075
{d, f}	0.35	0.175
{e, f}	0.15	0.075
{a, b, c}	0.50	0.16
{a, b, e}	0.25	0.083
{a, c, e}	0.25	0.083
{a, d, f}	0.45	0.15
{a, e, f}	0.25	0.083
{b, c, d}	0.40	0.13
{b, c, e}	0.55	0.183
{b, c, f}	0.40	0.13

Table 5. Cont.

A	Bel(A)	$\frac{Bel(A)}{ A }$
{b, d, f}	0.35	0.11 $\bar{6}$
{b, e, f}	0.15	0.05
{c, d, f}	0.35	0.11 $\bar{6}$
{c, e, f}	0.15	0.15
{d, e, f}	0.50	0.1 $\bar{6}$
{a, b, c, d}	0.50	0.125
{a, b, c, e}	0.65	0.1625
{a, b, c, f}	0.50	0.125
{a, b, d, f}	0.45	0.1125
{a, b, e, f}	0.25	0.0625
{a, c, d, f}	0.45	0.1125
{a, c, e, f}	0.25	0.0625
{a, d, e, f}	0.60	0.15
{b, c, d, e}	0.55	0.1375
{b, c, d, f}	0.75	0.1875
{b, c, e, f}	0.55	0.1375
{b, d, e, f}	0.50	0.125
{c, d, e, f}	0.50	0.125
{a, b, c, d, e}	0.65	0.13
{a, b, c, d, f}	0.85	0.17
{a, b, c, e, f}	0.65	0.13
{a, b, d, e, f}	0.60	0.12
{a, c, d, e, f}	0.60	0.12
{b, c, d, e, f}	0.9	0.18
{a, b, c, d, e, f}	1	0.1 $\bar{6}$

- First iteration.

We observe that the maximum of $\frac{Bel(A)}{|A|}$ is reached for $A = \{b, c\}$. Thus, for the elements of this set, we have $p_b = 0.2$ and $p_c = 0.20$. With this, we proceed to update the values of the belief function as follows:

$$Bel(\{d, f\}) \leftarrow Bel(\{b, c, d, f\}) - Bel(\{b, c\}) = 0.50 - 0.40 = 0.10,$$

$$Bel(\{a, d, e, f\}) \leftarrow Bel(\{a, b, c, d, e, f\}) - Bel(\{b, c\}) = 1 - 0.40 = 0.60.$$

We update $X = \{a, d, e, f\} \neq \emptyset$, verifying that $Bel(X) \neq 0$, and apply the algorithm again.

- Second iteration.

We start from $X = \{a, d, e, f\}$ with the values in Table 6:

Table 6. Values of $Bel(A)$ and $\frac{Bel(A)}{|A|}$ in the second iteration.

A	Bel(A)	$\frac{Bel(A)}{ A }$
{a}	0.10	0.10
{e}	0.15	0.15
{a, d}	0.10	0.05
{a, e}	0.25	0.125
{a, f}	0.10	0.05
{d, e}	0.15	0.075
{d, f}	0.35	0.175
{e, f}	0.15	0.075
{a, d, e}	0.25	0.08 $\bar{3}$
{a, d, f}	0.45	0.15
{a, e, f}	0.25	0.08 $\bar{3}$
{d, e, f}	0.50	0.1 $\bar{6}$
{a, d, e, f}	0.60	0.15

In this case, the maximum of $\frac{Bel(A)}{|A|}$ is 0.175, a value that corresponds to the set $\{d, f\}$. Thus, $p_d = 0.175 = p_f$ and we can update the belief function:

$$Bel(\{a, e\}) \leftarrow Bel(\{a, d, e, f\}) - Bel(\{d, f\}) = 0.6 - 0.35 = 0.25.$$

So, our new set is $X = \{a, e\} \neq \emptyset$ whose associated belief function is $Bel(X) = 0.25 \neq 0$; therefore, we begin another iteration.

- Third iteration.

We have the values shown in Table 7:

Table 7. Values of $Bel(A)$ and $\frac{Bel(A)}{|A|}$ in the third iteration.

A	Bel(A)	$\frac{Bel(A)}{ A }$
{a}	0.10	0.10
{e}	0.15	0.15
{a, e}	0.25	0.125

From these, we can see that $\frac{Bel(A)}{|A|}$ is maximized if $A = \{e\}$, so we assign $p_e = 0.15$. With this, we update the value of the function Bel:

$$Bel(\{a\}) \leftarrow Bel(\{a, e\}) - Bel(\{e\}) = 0.25 - 0.15 = 0.1.$$

We then have now that $X = \{a\} \neq \emptyset$ with $Bel(X) \neq 0$, so the algorithm is applied again.

- Fourth iteration.

Since we start from the set $X = \{a\}$, we maximize $\frac{Bel(A)}{|A|}$ on this same set, so we assign $p_a = 0.1$. By updating both the set X and the value of its belief function, we obtain $X = \emptyset$ and $Bel(X) = Bel(\emptyset) = 0$, at which point the algorithm terminates, and we can proceed to obtain the value of the maximum entropy:

$$S^*(Bel) = -(0.1 \cdot \log_2(0.1) + 0.15 \cdot \log_2(0.15) + 2 \cdot 0.175 \cdot \log_2(0.175) + 2 \cdot 0.2 \cdot \log_2(0.2)) = \boxed{2.55151}.$$

2. Algorithm-2

First, we will transform the data given by the mass function into reachable intervals. The values of the belief function and the plausibility function associated with each element of $\{a, b, c, d, e, f\}$ are as shown in Table 8:

Table 8. Belief intervals.

x_i	Bel(x_i)	Pl(x_i)	[Bel(x_i), Pl(x_i)]
{a}	0.1	0.10	[0.10, 0.10]
{b}	0	0.40	[0, 0.40]
{c}	0	0.40	[0, 0.40]
{d}	0	0.35	[0, 0.35]
{e}	0.15	0.15	[0.15, 0.15]
{f}	0	0.35	[0, 0.35]

Thus, we initialize the table to correctly apply the algorithm:

$$l = [0.10, 0, 0, 0, 0.15, 0], \quad u = [0.10, 0.40, 0.40, 0.35, 0.15, 0.35].$$

- First iteration.

We assign $S = \{1, 2, 3, 4, 5, 6\}$, and vector \hat{p} is given by

$$\hat{p} = [0.1, 0, 0, 0, 0.15, 0],$$

which satisfies $Sum(\hat{p}) = 0.25 < 1$. We look for indices i such that $l_i = u_i$, with $i = 1, \dots, 6$. We have $l_1 = u_1$ and $l_5 = u_5$, so S becomes $S = \{2, 3, 4, 6\}$. Therefore,

- * $s = Sum(l_i) = 0.25$,
- * $r = Min(l, S) = 2$,
- * $m = Nmin(l, S) = 4$,
- * $f = Sig(l, S) = -1$.

Hence, we apply the assignment $l_i \leftarrow l_i + \min(u_i - l_i, \frac{1-s}{m}, 1)$, with $m = Nmin(l, S) = 4$, $s = 0.25$. Thus, $\frac{1-s}{m} = 0.1875$, and our vector becomes $\hat{p} = [0.10, 0.1875, 0.1875, 0.1875, 0.15, 0.1875]$, before applying the algorithm again.

- **Second iteration.**

We start with $\hat{p} = [0.10, 0.1875, 0.1875, 0.1875, 0.15, 0.1875]$, where $Sum(\hat{p}) = 1$. The algorithm finishes, and we calculate the maximum associated entropy as follows:

$$S^*(Bel) = -(0.10 \cdot \log_2(0.10) + 4 \cdot 0.1875 \cdot \log_2(0.1875) + 0.15 \cdot \log_2(0.15)) = \boxed{2.554145}.$$

Unlike the two previous examples, we will now study two examples in which conflict appears.

Example 3. Given a set $X = \{a, b, c, d\}$, we define the mass function m as

$$\begin{aligned} m(\{a\}) &= 0.15 \\ m(\{b, c\}) &= 0.40 \\ m(\{c, d\}) &= 0.30 \\ m(\{a, d\}) &= 0.15 \end{aligned}$$

1. **Algorithm-1**

The values of the non-zero belief function associated with the $A \subseteq X$ are shown in Table 9:

Table 9. Values of $Bel(A)$ and $\frac{Bel(A)}{|A|}$.

A	$Bel(A)$	$\frac{Bel(A)}{ A }$
$\{a\}$	0.15	0.15
$\{b, c\}$	0.40	0.20
$\{c, d\}$	0.30	0.15
$\{a, d\}$	0.30	0.15
$\{a, b, c\}$	0.55	0.183
$\{a, c, d\}$	0.60	0.20
$\{b, c, d\}$	0.70	0.23
$\{a, b, c, d\}$	1	0.25

- **First iteration.**

We identify that set A , which means the $\frac{Bel(A)}{|A|}$ maximal is $\{a, b, c, d\}$. Since the chosen set coincides with X , we assign probabilities to all its elements:

$$p_a = p_b = p_c = p_d = 0.25.$$

Thus, for each $B \subseteq X - A$, we have $Bel(B) = 0$, and the new set is set \emptyset . Therefore, we can proceed to calculate the value of the maximum entropy associated with this distribution as follows:

$$S^*(Bel) = -4 \cdot 0.25 \cdot \log_2(0.25) = \boxed{2}.$$

2. **Algorithm-2**

For each element of $\{a, b, c, d, e\}$, we have its associated values, expressed in Table 10:

Table 10. Belief intervals.

x_i	$Bel(x_i)$	$Pl(x_i)$	$[Bel(x_i), Pl(x_i)]$
$\{a\}$	0.15	0.30	[0.15, 0.30]
$\{b\}$	0	0.40	[0, 0.40]
$\{c\}$	0	0.70	[0, 0.70]
$\{d\}$	0	0.45	[0, 0.45]

We can initialize the following:

$$l = [0.15, 0, 0, 0], \quad u = [0.30, 0.30, 0.70, 0.45].$$

- *First iteration.*

Let us assign $S = \{1, 2, 3, 4\}$, so for each $i = 1, 2, 3, 4$ we have the vector:

$$\hat{p} = [0.15, 0, 0, 0].$$

It is true that $Sum(l) = 0.15 < 1$, so we proceed to check if $l_i = u_i$ holds for some i . It does not hold for any $i = 1, 2, 3, 4$, so $S = \{1, 2, 3, 4\}$. We calculate the following:

- * $s = Sum(\hat{p}) = 0.15$,
- * $r = Min(l, S) = 2$,
- * $m = Nmin(l, S) = 3$,
- * $f = Sig(l, S) = 1$, so we move to step 8.

We now assign $l_i + \min(u_i - l_i, l_f - l_r, \frac{1-s}{m}) \rightarrow l_i$, where $\min(u_i - l_i, l_f - l_r, \frac{1-s}{m}) = 0.15$

- *Second iteration.*

The new vector \hat{p} is given by:

$$\hat{p} = [0.15, 0.15, 0.15, 0.15],$$

verifying $Sum(\hat{p}) = 0.60 < 1$. Furthermore, no $i \in \{1, 2, 3, 4\}$ satisfies $l_i = u_i$, so S is not modified. Thus,

- * $s = Sum(\hat{p}) = 0.60$,
- * $r = Min(l, S) = 1$,
- * $m = Nmin(l, S) = 4$,
- * $f = Sig(l, S) = -1$.

Hence, we assign $l_i \leftarrow l_i + \min(u_i - l_i, \frac{1-s}{m}, 1) = l_i + \frac{1-0.60}{4} = l_i + 0.10$.

- *Third iteration.*

The vector \hat{p} updated with the values from the previous iteration is as follows:

$$\hat{p} = [0.25, 0.25, 0.25, 0.25],$$

verifying $\text{Sum}(\hat{p}) = 1$, and thus terminating the algorithm. Thus, we have arrived at the fact that the probability vector of maximum entropy is $\hat{p} = [0.25, 0.25, 0.25, 0.25]$, and we proceed to obtain the value of the maximum entropy:

$$S^*(\text{Bel}) = -4 \cdot 0.25 \cdot \log_2(0.25) = \boxed{2}.$$

Example 4. We define the following mass function m on the set $X = \{a, b, c, d, e\}$:

$$\begin{aligned} m(\{a\}) &= 0.10 \\ m(\{b, c, d\}) &= 0.35 \\ m(\{d, e\}) &= 0.20 \\ m(\{a, c, e\}) &= 0.25 \\ m(\{b, e\}) &= 0.10 \end{aligned}$$

1. **Algorithm-1**

The values of $\text{Bel}(A)$ for each $A \subseteq X$ are shown in Table 11:

Table 11. Values of $\text{Bel}(A)$ and $\frac{\text{Bel}(A)}{|A|}$.

A	Bel(A)	$\frac{\text{Bel}(A)}{ A }$
{a}	0.10	0.10
{a, b}	0.10	0.05
{a, c}	0.10	0.05
{a, d}	0.10	0.05
{a, e}	0.10	0.05
{b, e}	0.10	0.05
{d, e}	0.20	0.10
{a, b, c}	0.10	0.033
{a, b, d}	0.10	0.033
{a, b, e}	0.20	0.066
{a, c, d}	0.10	0.033
{a, c, e}	0.35	0.116
{a, d, e}	0.30	0.10
{b, c, d}	0.35	0.116
{b, c, e}	0.10	0.033
{b, d, e}	0.30	0.10
{c, d, e}	0.20	0.066
{a, b, c, d}	0.45	0.1125
{a, b, c, e}	0.45	0.1125
{a, b, d, e}	0.40	0.10
{a, c, d, e}	0.55	0.1375
{b, c, d, e}	0.65	0.15625
{a, b, c, d, e}	1	0.20

- *First iteration.*

As in the previous example, we observe that the set that maximizes $\frac{\text{Bel}(A)}{|A|}$ is $A = \{a, b, c, d, e\}$, and thus assign probabilities to all elements in the following form:

$$p_a = p_b = p_c = p_d = p_e = 0.20.$$

We then proceed to take the set $X = \emptyset$, with $\text{Bel}(X) = 0$, so we move directly to calculating the maximum entropy:

$$S^*(Bel) = -5 \cdot 0.20 \cdot \log_2(0.20) = \boxed{2.3219}.$$

2. **Algorithm-2**

Table 12 shows the reachable interval associated with each element of $\{a, b, c, d, e\}$:

Table 12. Belief intervals.

x_i	$Bel(x_i)$	$Pl(x_i)$	$[Bel(x_i), Pl(x_i)]$
$\{a\}$	0.10	0.35	[0.10, 0.35]
$\{b\}$	0	0.45	[0, 0.45]
$\{c\}$	0	0.60	[0, 0.60]
$\{d\}$	0	0.55	[0, 0.55]
$\{d\}$	0	0.55	[0, 0.55]

Therefore, we have:

$$l = [0.10, 0, 0, 0, 0], u = [0.35, 0.45, 0.60, 0.55, 0.55].$$

- *First iteration.*

We assign $S = \{1, 2, 3, 4, 5\}$, $l_i \rightarrow \hat{p}_i$, so that:

$$\hat{p} = [0.10, 0, 0, 0, 0],$$

$Sum(\hat{p}) = 0.10 < 1$. Furthermore, no $i \in \{1, 2, 3, 4, 5\}$ satisfies $l_i = u_i$, so S remains the same. We now obtain the following values:

- * $s = Sum(\hat{p}) = 0.10$,
- * $r = Min(l, S) = 2$,
- * $m = Nmin(l, S) = 4$,
- * $f = Sig(l, S) = 1$.

Since $f = 1$, we perform the assignment presented in step 8, $l_i \leftarrow l_i + \min(u_i - l_i, l_f - l_r, \frac{1-s}{m}) = l_i + \min(0.45, 0.10, 0.225) = l_i + 0.10$ and proceed to the next iteration.

- *Second iteration*

We now start from the vector:

$$\hat{p} = [0.10, 0.10, 0.10, 0.10, 0.10],$$

such that $Sum(\hat{p}) = 0.50 < 1$. Again, no $l_i = u_i$ for $i = 1, 2, \dots, 5$, then

- * $s = Sum(\hat{p}) = 0.50$;
- * $r = Min(l, S) = 1$;
- * $m = Nmin(l, S) = 5$;
- * $f = Sig(l, S) = -1$, so we go to step 7.

We assign $l_i \leftarrow l_i + \min(u_i - l_i, \frac{1-s}{m}, 1) = l_i + 0.10$, and the algorithm is applied again.

- *Third iteration.*

Since we have $\hat{p} = [0.20, 0.20, 0.20, 0.20, 0.20]$, we can verify that $Sum(\hat{p}) = 1$, thus terminating the algorithm. Therefore, the probabilities with maximum entropy for the given intervals are $\hat{p}_a = \hat{p}_b = \hat{p}_c = \hat{p}_d = \hat{p}_e = 0.20$, and we calculate the associated maximum entropy:

$$S^*(Bel) = -5 \cdot 0.20 \cdot \log_2(0.20) = \boxed{2.3219}.$$

Having seen these examples, we can highlight some differences between the algorithms. To do this, we will use Table 13:

Table 13. Differences between the two algorithms.

Example	Algorithm	Conflict	Cardinality	Iterations	Maximum of the Entropy
1	Algorithm-1	No	5	3	2.309982
	Algorithm-2		5	2	2.311278
2	Algorithm-1	No	6	4	2.55151
	Algorithm-2		6	2	2.554145
3	Algorithm-1	Yes	4	1	2
	Algorithm-2		4	3	2
4	Algorithm-1	Yes	5	1	2.3219
	Algorithm-2		5	3	2.3219

From this first numerical comparison, we can make the following comments.

In the case of the examples where conflict appears, the number of iterations used by Algorithm-1 is lower than the number of iterations used by Algorithm-2. Furthermore, the maximum entropy coincides for both theories. Therefore, we might think that, for cases with conflict, Algorithm-1 is more efficient with respect to the number of iterations needed. However, in the case where there is no conflict, we can think that the opposite occurs.

In the following Subsection, we show what happens when we carry out an extensive experimentation with a very large number of different BPAs.

Experimentation and Computational Analysis

To evaluate the computational efficiency of both algorithms, we conducted a series of experiments generating Basic Probability Assignments (BPAs) on frames of discernment with sizes $n = 5$, $n = 6$, and $n = 10$. Both algorithms were implemented in the C++ programming language and executed on a system equipped with an Intel Core i5 1.8GHz CPU and 8GB of RAM.

To assess performance across different evidential structures, we randomly generated one million BPAs with conflict (C) and one million without conflict (NC). For conflicting cases (C), all possible subsets could serve as focal elements; values were generated in the range $[0, 100]$ and subsequently normalized. For the non-conflicting cases (NC), we considered combinations of disjoint sets. For $n = 5$ and $n = 6$, all possible disjoint combinations were explored, while for $n = 10$, focal sets were restricted to cardinalities between 2 and 8 to maximize mass distribution. The results are summarized in Table 14.

Table 14. Processing time (seconds) for calculating maximum entropy across 1 million randomly generated BPAs for $n \in \{5, 6, 10\}$.

Size	With Conflict (C)		No Conflict (NC)	
	Algorithm-1 (s)	Algorithm-2 (s)	Algorithm-1 (s)	Algorithm-2 (s)
$n = 5$	14.19	16.45	18.39	18.26
$n = 6$	83.32	96.95	85.90	79.54
$n = 10$	23539.77	30012.53	24650.91	23660.00

The experimental results align with the theoretical expectations derived from the numerical examples. In scenarios involving conflict, where BPAs typically possess a larger number of focal elements, Algorithm-1 consistently outperforms Algorithm-2. This is noteworthy because a higher count of focal elements often presents a computational challenge. The relative underperformance of Algorithm-2 in these cases is attributed

to the high overhead of calls to auxiliary functions. For $n = 6$, Algorithm-1 shows an improvement of approximately 16.5% over Algorithm-2, a disparity that becomes even more pronounced at $n = 10$, reaching a performance gain of 27.5%.

In contrast, in conflict-free scenarios, the performance hierarchy is partially reversed. Although the differences are negligible for $n = 5$, Algorithm-2 shows an improvement of approximately 8% for $n = 6$. In these cases, the number of focal sets is restricted by the disjointness constraint, which benefits the iterative structure of Algorithm-2.

For $n = 10$, where the number of possible focal sets increases exponentially ($2^{10} - 1 = 1023$), the behavioral patterns persist. In conflicting scenarios, Algorithm-1's superiority suggests that its efficiency scales better with the number of elements in the universal set. In contrast, for non-conflict scenarios, both algorithms exhibit similar performance, with Algorithm-2 maintaining a slight advantage (below 3%). This convergence suggests that as n increases, the computational burden of calculating the *Bel* and *Pl* functions becomes the dominant factor in terms of overall complexity.

Theoretical analysis confirms these observations. Algorithm-1 solves a constrained maximization problem over the full credal set, leading to an exponential time and space complexity of $O(2^n)$. Although Algorithm-2 employs a constructive strategy that appears polynomial ($O(n^2)$), its execution time remains inherently linked to the number of focal sets $|\wp(X)|$. When this number is large, the frequency of calls to auxiliary functions results in an overall exponential complexity comparable to that of Algorithm-1.

In conclusion, our experimental study suggests a pragmatic approach to algorithm selection: Algorithm-2 is preferable for belief structures where conflict is absent or minimal. However, for general applications where conflict is likely—a more common occurrence in real-world data—the classical Algorithm-1 remains the more robust and efficient choice.

4. Conclusions and Future Work

The computational cost associated with the algorithm of Meyerowitz et al. has traditionally been the primary drawback to using maximum entropy as a measure to quantify uncertainty and information within Evidence Theory (ET).

This paper critically analyzes the assertion that is frequently found in the recent literature [27]:

"The exact computation of maximum entropy in Evidence Theory, as performed by the algorithm of Meyerowitz et al. (Algorithm-1), is characterized by exponential computational complexity with respect to the size of the frame of discernment. In contrast, the interval-based formulation (Algorithm-2) reduces the problem to a polynomial-time optimization at the cost of a controlled loss of information. Since Algorithm-2 typically yields maximum entropy values close to those obtained with Algorithm-1, it constitutes a preferable alternative in practical applications, particularly for large frames of discernment."

Our analysis reveals that when evaluating these algorithms, it is essential to distinguish between scenarios where conflict is absent and those where it is present. Theoretically, Algorithm-1 is expected to perform worse in the presence of conflict, as conflict typically increases the number of focal elements. Under such conditions, it has been generally assumed that Algorithm-2 would be more efficient. However, our experimental results demonstrate that this expected behavior does not consistently occur. This discrepancy can be attributed to the extensive number of function calls required by Algorithm-2, which effectively offsets its theoretical computational advantages.

As future work, we intend to investigate alternative procedures for computing maximum entropy in ET that offer superior computational efficiency, even if this entails obtaining approximate rather than exact values, as is currently the case with Algorithm-2.

Overall, the findings of this study challenge the common assumption that transforming a belief function into an equivalent representation based on reachable probability intervals necessarily facilitates the computation of maximum entropy.

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

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: This research has been supported by the Spanish Ministry of Science, Innovation, and Universities under project PID2024.159012NA.I00.

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

References

- Walley, P. *Statistical Reasoning with Imprecise Probabilities*; Chapman & Hall/CRC Monographs on Statistics & Applied Probability; Taylor & Francis: Boca Raton, FL, USA, 1991; Volume 42.
- Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339. <https://doi.org/10.1214/aoms/1177698950>.
- Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press Princeton: Princeton, NJ, USA, 1976.
- Beynon, M.; Curry, B.; Morgan, P. The Dempster–Shafer theory of evidence: An alternative approach to multicriteria decision modelling. *Omega* **2000**, *28*, 37–50. [https://doi.org/https://doi.org/10.1016/S0305-0483\(99\)00033-X](https://doi.org/https://doi.org/10.1016/S0305-0483(99)00033-X).
- 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. https://doi.org/10.1007/978-3-540-44792-4_29.
- Buede, D.M.; Girardi, P. A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion. *IEEE Trans. Syst. Man, Cybern. Part A Syst. Humans* **1997**, *27*, 569–577. <https://doi.org/10.1109/3468.618256>.
- Ip, H.H.S.; Ng, J.M.C. Human face recognition using Dempster-Shafer theory. In *Proceedings of 1st International Conference on Image Processing*; IEEE: New York, NY, USA, 1994; Volume 2, pp. 292–295. <https://doi.org/10.1109/ICIP.1994.413578>.
- Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Yager, R.R. Entropy and specificity in a mathematical theory of evidence. *Int. J. Gen. Syst.* **1983**, *9*, 249–260. <https://doi.org/10.1080/03081078308960825>.
- Klir, G.; Wierman, M. *Uncertainty-Based Information: Elements of Generalized Information Theory*; Studies in Fuzziness and Soft Computing; Physica-Verlag: Heidelberg, Germany, 1999.
- Abellán, J.; Masegosa, A. Requirements for total uncertainty measures in Dempster-Shafer theory of evidence. *Int. J. Gen. Syst.* **2008**, *37*, 733–747. <https://doi.org/10.1080/03081070802082486>.
- Jaynes, E. On the rationale of maximum-entropy methods. *Proc. IEEE* **1982**, *70*, 939–952. <https://doi.org/10.1109/PROC.1982.12425>.
- Provencher Langlois, G.; Buch, J.; Darbon, J. Efficient First-Order Algorithms for Large-Scale, Non-Smooth Maximum Entropy Models with Application to Wildfire Science. *Entropy* **2024**, *26*, 691. <https://doi.org/10.3390/e26080691>.
- Stoyanov, J.M.; Tagliani, A.; Novi Inverardi, P.L. Maximum Entropy Criterion for Moment Indeterminacy of Probability Densities. *Entropy* **2024**, *26*, 121. <https://doi.org/10.3390/e26020121>.
- Abellán, J. Ensembles of decision trees based on imprecise probabilities and uncertainty measures. *Inf. Fusion* **2013**, *14*, 423–430.
- Klir, G.J. *Uncertainty and Information: Foundations of Generalized Information Theory*; John Wiley And Sons, Inc.: Hoboken, NJ, USA, 2005. <https://doi.org/10.1002/0471755575>.
- Klir, G.J.; Smith, R.M. On Measuring Uncertainty and Uncertainty-Based Information: Recent Developments. *Ann. Math. Artif. Intell.* **2001**, *32*, 5–33. <https://doi.org/10.1023/A:1016784627561>.
- Abellán, J.; Moral, S. Maximum of Entropy for Credal Sets. *Int. J. Uncertainty Fuzziness -Knowl.-Based* **2003**, *11*, 587–597. <https://doi.org/10.1142/S021848850300234X>.

19. Harmanec, D.; Klir, G.J. Measuring total uncertainty in Dempster-Shafer Theory: A novel approach. *Int. J. Gen. Syst.* **1994**, *22*, 405–419. <https://doi.org/https://doi.org/10.1080/03081079408935225>.
20. Xiao, F. On the Maximum Entropy Negation of a Complex-Valued Distribution. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 3259–3269. <https://doi.org/10.1109/TFUZZ.2020.3016723>.
21. Meyerowitz, A.; Richman, F.; Walker, E. Calculating maximum-entropy probability densities for belief functions. *Int. J. Uncertainty Fuzziness-Knowl.-Based Syst.* **1994**, *02*, 377–389. <https://doi.org/https://doi.org/10.1142/S0218488594000316>.
22. Deng, Y. Deng entropy. *Chaos Solitons Fractals* **2016**, *91*, 549–553. <https://doi.org/10.1016/j.chaos.2016.07.014>.
23. Abellán, J. Analyzing properties of Deng entropy in the theory of evidence. *Chaos Solitons Fractals* **2017**, *95*, 195–199. <https://doi.org/10.1016/j.chaos.2016.12.024>.
24. 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. <https://doi.org/10.1109/TSMC.2017.2770128>.
25. Abellán, J.; Moral, S. Building classification trees using the total uncertainty criterion. *Int. J. Intell. Syst.* **2003**, *18*, 1215–1225. <https://doi.org/10.1002/int.10143>.
26. Abellán, J.; Masegosa, A.R. Bagging schemes on the presence of class noise in classification. *Expert Syst. Appl.* **2012**, *39*, 6827–6837. <https://doi.org/10.1016/j.eswa.2012.01.013>.
27. Moral-García, S.; Abellán, J. Maximum of Entropy for Belief Intervals Under Evidence Theory. *IEEE Access* **2020**, *8*, 118017–118029. <https://doi.org/10.1109/ACCESS.2020.3003715>.
28. Hartley, R.V.L. Transmission of Information1. *Bell Syst. Tech. J.* **1928**, *7*, 535–563. <https://doi.org/10.1002/j.1538-7305.1928.tb01236.x>.
29. Dubois, D.; Prade, H. A note on measures of specificity for fuzzy sets. *Int. J. Gen. Syst.* **1985**, *10*, 279–283. <https://doi.org/10.1080/03081078508934893>.
30. Maeda, Y.; Nguyen, H.T.; Ichihashi, H. Maximum entropy algorithms for uncertainty measures. *Int. J. Uncertain. Fuzziness-Knowl.-Based Syst.* **1993**, *01*, 69–93. <https://doi.org/https://doi.org/10.1142/S021848859300005X>.
31. Huynh, V.N.; Nakamori, Y. Notes on reducing algorithm complexity for computing an aggregate uncertainty measure. *IEEE Trans. Syst. Man, Cybern. Part A Syst. Hum.* **2010**, *40*, 205–209. <https://doi.org/https://doi.org/10.1109/TSMCA.2009.2030962>.
32. Cui, H.; Liu, Q.; Zhang, J.; Kang, B. An Improved Deng Entropy and Its Application in Pattern Recognition. *IEEE Access* **2019**, *7*, 18284–18292.
33. Kang, B.; Deng, Y. The Maximum Deng Entropy. *IEEE Access* **2019**, *7*, 120758–120765.
34. 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.
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. <https://doi.org/10.3390/e21020163>.
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*, 1122–1137. <https://doi.org/http://dx.doi.org/10.3390/e21111122>.

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.