axioms

MDPI

Article

Different Types of Entropy Measures for Type-2 Fuzzy Sets

Luis Magdalena *, **Carmen Torres-Blanc**, **Susana Cubillo** and **Jesus Martinez-Mateo**

Departamento de Matemática Aplicada a las Tecnologías de la Información y las Comunicaciones, Universidad Politécnica de Madrid, Campus de Montegancedo, 28660 Boadilla del Monte, Madrid, Spain; mariadelcarmen.torres@upm.es (C.T.-B.); susana.cubillo@upm.es (S.C.); jesus.martinez.mateo@upm.es (J.M.-M.)
* Correspondence: luis.magdalena@upm.es

Abstract: In this work, we consider De Luca and Termini's notion of non-probabilistic entropy, and we extend some entropy-like measures of the degree of fuzziness to type-2 fuzzy sets. With this aim, we first study different entropy measures proposed in the frameworks of fuzzy, intuitionistic, and interval-valued fuzzy sets. Then, we propose three possible novel axiomatizations for entropy in type-2 fuzzy sets. The proposed types of entropy measures evaluate how much a type-2 fuzzy set is non-crisp, non-fuzzy, and non-interval-valued fuzzy. This can also be interpreted as how far a type-2 fuzzy set is from a crisp, fuzzy, or interval-valued fuzzy set. The present contribution is also novel, since we considered the interpretation of type-2 fuzzy sets that is closest to Zadeh's original conception.

Keywords: uncertainty; fuzziness; non-probabilistic fuzzy entropy; type-2 fuzzy set

MSC: 03E72 (Fuzzy set theory), 28E10 (Fuzzy measure theory)

Citation: Magdalena, L.; Torres-Blanc, C.; Cubillo, S.; Martinez-Mateo, J. Different Types of Entropy Measures for Type-2 Fuzzy Sets. *Axioms* **2024**, *13*, 556. https://doi.org/10.3390/axioms13080556

Academic Editor: Boldizsár Tüű-Szabó

Received: 22 June 2024
Revised: 27 July 2024
Accepted: 6 August 2024
Published: 14 August 2024

1. Introduction

The notion of entropy for a fuzzy set was initially studied by Zadeh in 1968 [1] in terms of a probability measure. Zadeh first introduced the concept of a fuzzy event from a fuzzy set and defined the probability of a fuzzy event as the expectation of its membership function. He then defined the entropy of a fuzzy set from probabilities as the usual definition of entropy in information theory. Subsequently, in 1972, De Luca and Termini introduced the novel concept of the non-probabilistic entropy of a fuzzy set [2]. Their aim was to find a measure of uncertainty, fuzziness, or entropy in fuzzy sets that did not depend on probabilities. To this end, they began by studying the axiomatic properties that any measure of entropy should satisfy. Consequently, they put forth the proposition that an entropy-like measure of the degree of fuzziness of a fuzzy set, denoted by $d(f)$ (where d is a function from the set of fuzzy sets to the interval $[0,1]$, and f the membership degree function of a fuzzy set), should satisfy the following properties:

1.1 $d(f)$ must be 0 if and only if f takes values 0 or 1.
1.2 $d(f)$ must assume the maximum value if and only if f always assumes the value $\frac{1}{2}$.
1.3 $d(f)$ must be greater than or equal to $d(f')$ where f' is any sharpened version of f, that is, any fuzzy set such that $f'(x) \geq f(x)$ if $f(x) \geq \frac{1}{2}$, and $f'(x) \leq f(x)$ if $f(x) < \frac{1}{2}$.

The property "sharpened version of" induces a partial order $\leq_1$ in the set of membership functions such that $f_1 \leq_1 f_2$ when f_1 is a sharpened version of f_2. Therefore, $d(f)$ is monotonically non-decreasing with respect to this order. The entropy then refers to the distance measured from the fuzzy set f to its most-sharpened version f' defined as follows:

$$f'(x) = \begin{cases} 1 & \text{if } f(x) \geq \frac{1}{2}, \\ 0 & \text{if } f(x) < \frac{1}{2}. \end{cases}$$

This most-sharpened version transforms the fuzzy set into a crisp set. In fact, this is the closest crisp set to f, as illustrated in Figure 1.

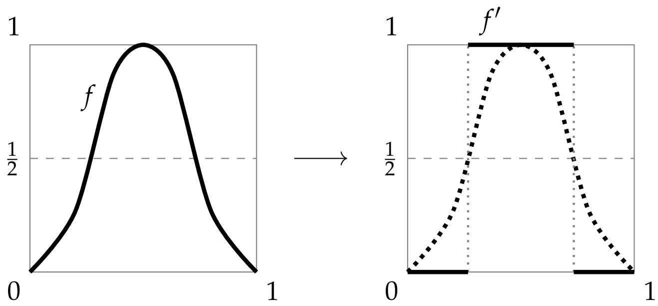

Figure 1. A fuzzy set f (**left**) and its most-sharpened version f' (**right**).

This approach was also employed in [3], where Kaufmann defined an entropy measure as the distance between a fuzzy set and its nearest crisp set. However, other approaches to measuring the entropy of a fuzzy set have also been proposed. For instance, Yager [4,5] defined an entropy measure considering the distance between a fuzzy set and its complementary set. Other authors, such as Kosko [6] and Madrid and Ojeda [7], defined entropy measures using the notion of subsethood degree in fuzzy sets.

Later, two major contributions focused on the entropy of interval-valued fuzzy sets [8,9]. Both works agree on the introduction of an additional axiom that had not been considered by De Luca and Termini in their seminal work: the entropy of an interval-valued fuzzy set must be equal to the entropy of its complementary set. However, both works follow different approaches. Burillo and Bustince [8] quantify the dissimilarity between an interval-valued fuzzy set and a type-1 fuzzy set, whereas Zeng and Li [9] quantify the dissimilarity between an interval-valued fuzzy set and a crisp set. In other words, the former work by Burillo considers the entropy measure to be the distance from the membership degree (that is, an interval) to the closest $[a, a]$ interval, while the latter work by Zeng considers the entropy measure to be the distance to the intervals $[0, 0]$ and $[1, 1]$.

In this contribution, when comparing both proposals for the entropy in interval-valued fuzzy sets, we realized that these rely on different premises, that is, on a different set of axiomatic properties, which are mutually incompatible. Therefore, this implies that there exists at least two different types of entropy in these sets. Other approaches to the entropy of interval-valued fuzzy sets were proposed by Wu and Mendel [10], Liu and Cai [11], and Cornelis and Kerre [12].

The entropy of type-2 fuzzy sets has also been studied, see [13–16], but none of these works consider type-2 fuzzy sets as proposed by Mizumoto and Tanaka [17,18], more in line with Zaded's initial conception [19], that is, where the membership degrees are labels of the variable truth. For instance, in [13–15], the authors consider the membership degrees of a type-2 fuzzy set simply as fuzzy sets. This is functionally equivalent to considering the type-2 fuzzy set as a type-1 fuzzy set defined in a two-dimensional domain, the second dimension being the interval $[0, 1]$, that is, the space of membership degrees. With this approach, they have a different conception of what is the complementary set of a type-2 fuzzy set (as argued in [17]), and also of what is a type-2 fuzzy set with crisp membership degrees. These are central questions in the analysis of entropy; for instance, once introduced by the above-mentioned new axiom, the entropy of a fuzzy set must be equal to the entropy of its complementary set. Consequently, considering membership degrees as labels of the variable truth, or simply as type-1 fuzzy sets, will lead to a different interpretation of entropy.

In this work, we study the entropy of type-2 fuzzy sets considering sets as proposed by Mizumoto and Tanaka. To this end, we first analyze what has been done so far in the case of fuzzy sets and interval-valued fuzzy sets, then we describe the properties that it is reasonable to require of an entropy (including those related to the complementary set), and, finally, we propose three novel types of entropy measures for type-2 fuzzy sets. With each of the proposed types of entropy, we measure how much a type-2 fuzzy set is non-crisp, non-fuzzy, or non-interval-valued.

The remainder of this paper is organized as follows. In Section 2, we analyze entropy in interval-valued fuzzy sets. In a previous contribution [20], we already studied the works we considered most relevant and compared several entropy measures proposed for interval-valued fuzzy sets. Here, we revisit this work as it is the basis of the present study. In Section 3, we introduce type-2 fuzzy sets and use the ideas reviewed in the previous section to propose different types of entropy measures for these sets. Finally, we present our conclusions in Section 4.

2. Entropy in Interval-Valued Fuzzy Sets

Interval-valued fuzzy sets (IVFSs) are characterized by a membership function that assigns to each element of the universe a membership degree in the form of a closed subinterval of the unit interval $[0,1]$, instead of a single value in that interval, as is the case with fuzzy sets. That is,

$$\mu_A(x) = [L_A(x), U_A(x)] \quad \forall x \in X,$$

where $0 \leq L_A(x) \leq U_A(x) \leq 1$.

In the following, we consider different definitions of entropy for IVFSs. However, it is important to note that some of the proposals considered were presented in the framework of intuitionistic fuzzy sets (IFSs), and not for IVFSs. Nevertheless, considering their formal equivalence [21], we analyze all of them from the IVFS point of view. Whether defined for IVFSs or IFSs, the nomenclature can be adapted, if necessary, by applying the following correspondence between an intuitionistic fuzzy set B and its counterpart in an IVFS, the set A, as described in [8]:

$$\mu_A(x) = [L_A(x), U_A(x)] = [\eta_B(x), 1 - \nu_B(x)] \quad \forall x \in X,$$

where the membership degree of the intuitionistic fuzzy set B is $\sigma_B(x) = (\eta_B(x), \nu_B(x))$.

Burillo and Bustince [8] first adapted the axiomatization of a non-probabilistic entropy, as previously defined by De Luca and Termini for fuzzy sets, to both IVFS and IFS frameworks. In their approach, entropy is a tool to measure how far an interval-valued fuzzy set is from a fuzzy set. Thus, a fuzzy set (in which the membership value of each element of the universe is an interval reduced to a point) will have an entropy measure of 0, while the set whose membership values are all the interval $[0,1]$ will have the maximum value, that is, 1 if the function is normalized. Let IVFS(X) denote the set of all interval-valued fuzzy sets in a finite universe X. Then, according to Burillo and Bustince, a function $I\colon \text{IVFS}(X) \to [0,1]$ is an entropy measure on IVFS(X) if it satisfies the following axioms:

2.1 $I(A) = 0 \Leftrightarrow A$ is a fuzzy set, that is, $\mu_A(x) \in [0,1]$ or, equivalently, $L_A(x) = U_A(x)$, $\forall x \in X$.
2.2 $I(A) = 1 \Leftrightarrow \mu_A(x) = [0,1]$, that is, $L_A(x) = 0$ and $U_A(x) = 1$, $\forall x \in X$.
2.3 $I(A) = I(A^c)$, A^c being the complementary set of A, where $\mu_{A^c}(x) = [1 - U_A(x), 1 - L_A(x)]$, $\forall x \in X$.
2.4 If $A \preceq B$ (that is, $[L_A(x), U_A(x)] \subseteq [L_B(x), U_B(x)]$, $\forall x \in X$), then $I(A) \leq I(B)$.

Given that $\preceq$ is an order relation in IVFSs, Axiom 2.4 requires that any entropy measure that satisfies it must be monotonically non-decreasing with respect to that order.

An example of an entropy measure that satisfies these axiomatic properties is as follows:

$$I(A) = \frac{1}{N} \sum_{i=1}^{N} (U_A(x_i) - L_A(x_i)). \tag{1}$$

In a later work, Zeng and Li [9] modified this axiomatization in order to introduce different entropy measures in IVFSs. The proposed entropy is now a measure of how non-crisp an interval-valued fuzzy set is, that is, it measures how far the set is from being crisp. Then, according to Zeng and Li, a function $E\colon \text{IVFS}(X) \to [0,1]$ is an entropy measure on IVFS(X) if it satisfies the following axioms:

3.1 $E(A) = 0 \Leftrightarrow A$ is a crisp set, that is, $\mu_A(x) = [0,0]$ or $\mu_A(x) = [1,1]$, $\forall x \in X$.
3.2 $E(A) = 1 \Leftrightarrow L_A(x) + U_A(x) = 1, \forall x \in X$.
3.3 $E(A) = E(A^c)$, A^c being the complementary set of A.
3.4 If $A \preccurlyeq B$ (that is, if $L_B(x) + U_B(x) \leq 1$, then $L_A(x) \leq L_B(x)$ and $U_A(x) \leq U_B(x)$; and if $L_B(x) + U_B(x) \geq 1$, then $L_A(x) \geq L_B(x)$ and $U_A(x) \geq U_B(x)$), then $E(A) \leq E(B)$.

Again, considering that $\preccurlyeq$ is an order relation in IVFSs, Axiom 3.4 requires that any entropy measure satisfying it must be monotonically non-decreasing with respect to the order $\preccurlyeq$.

Several examples of entropy measures satisfying these other axioms can be found in the literature, but, in this contribution, we consider only three of them. First, the one proposed by Zeng and Li [9]:

$$E_Z(A) = 1 - \frac{1}{N}\sum_{i=1}^{N}|L_A(x_i) + U_A(x_i) - 1|. \quad (2)$$

Second, the entropy measure proposed by Szmidt and Kacprzyk [22]. It should be noted that this proposal predates Zeng's work, and although it was defined in the framework of IFSs, we have redefined it here for IVFSs.

$$E_S(A) = \frac{1}{N}\sum_{i=1}^{N}\frac{1 - \max(1 - U_A(x_i), L_A(x_i))}{1 - \min(1 - U_A(x_i), L_A(x_i))}. \quad (3)$$

Third, the following other entropy measure proposed by Vlachos and Sergiadis [23]:

$$E_V(A) = \frac{1}{N}\sum_{i=1}^{N}\frac{\min(L_A(x_i), 1 - U_A(x_i)) + \min(U_A(x_i), 1 - L_A(x_i))}{\max(L_A(x_i), 1 - U_A(x_i)) + \max(U_A(x_i), 1 - L_A(x_i))}. \quad (4)$$

Before analyzing the four entropy measures presented above, we consider it interesting to compare both sets of axioms and analyze the purpose of each one. First, comparing Axiom 2.1 and Axiom 3.1, we quickly realize that any fuzzy set has an entropy of 0 for Burillo, while only a crisp set has an entropy of 0 for Zeng. This was to be expected since, as explained above, Burillo measures how far an IVFS is from a fuzzy set (that is, how non-fuzzy it is) while Zeng measures how far an IVFS is from a crisp set (that is, how non-crisp it is). Furthermore, it is noteworthy that this also induces a completely different ordering in both entropy measures, as is shown in the following.

The orders considered in IVFSs are those defined on intervals. Let $A = [L_A, U_A]$ and $B = [L_B, U_B]$ be two intervals; we consider the following partial orders defined on intervals:

- Contained order ($A \subseteq B$). It is said that A is contained in B if and only if $L_B \leq L_A \leq U_A \leq U_B$.
- Lattice order ($A \leq B$). It is said that A is lower than or equal to B if and only if $L_A \leq L_B$ and $U_A \leq U_B$.

Let A and B now be two IVFSs, such that $A(x)$ and $B(x)$ are intervals. We can define two relations (partial orders) in IVFSs from contained and lattice orders, as follows:

- Embedded. It is said that A is embedded in B if and only if $A(x) \subseteq B(x), \forall x \in X$.
- Included. It is said that A is included in B if and only if $A(x) \leq B(x), \forall x \in X$.

It should be noted that both relations correspond to different situations. Inclusion means that the information given by one interval-valued fuzzy set (included in another) is less true than that given by the other. Embedding occurs when one interval-valued fuzzy set (the embedded one) has more precise information about the actual membership function than another.

Now, let us compare two IVFSs A and B with a fuzzy set C, where, for each $x \in X$, $\mu_A(x) = [L_A(x), U_A(x)]$, $\mu_B(x) = [L_B(x), U_B(x)]$, and $\mu_C(x) = c$, with $c \in [L_A(x), U_A(x)]$ and $c \in [L_B(x), U_B(x)]$. Then, we have $B \preceq A$ when $[L_B(x), U_B(x)] \subseteq [L_A(x), U_A(x)]$, $\forall x \in X$, that is, B is closer to C than A (see the top of Figure 2). Thus, to get closer to a fuzzy set, the interval contracts. According to the given definitions, this order is an embedding of IVFSs. This makes sense because embedding occurs when an interval-valued fuzzy set has more precise information, and, in such a case, the entropy should also be lower. For a deeper analysis of embedding in IVFSs, see [24].

On the other hand, let us now compare two IVFSs A and B with a crisp set C, that is, either $\mu_C(x) = 0$ or $\mu_C(x) = 1$. Now, we have $B \preccurlyeq A$ when $L_B(x) \leq L_A(x), U_B(x) \leq U_A(x)$ if $L_A(x) + U_A(x) \leq 1$ (or $L_A(x) \leq L_B(x), U_A(x) \leq U_B(x)$ if $L_A(x) + U_A(x) \geq 1$), that is, B is closer to 0 (or 1, respectively) than A (see center and bottom of Figure 2). Thus, to get closer to a crisp set, the whole interval shifts to the left (or right). This order is then based on the lattice order for intervals, but it is not an inclusion in IVFSs. That is because, when A is closer to 0 than B, A should be included in B, but when A is closer to 1 than B, B should be included in A. In both cases, the entropy of A will be lower than that of B, and, consequently, inclusion does not directly correlate with entropy.

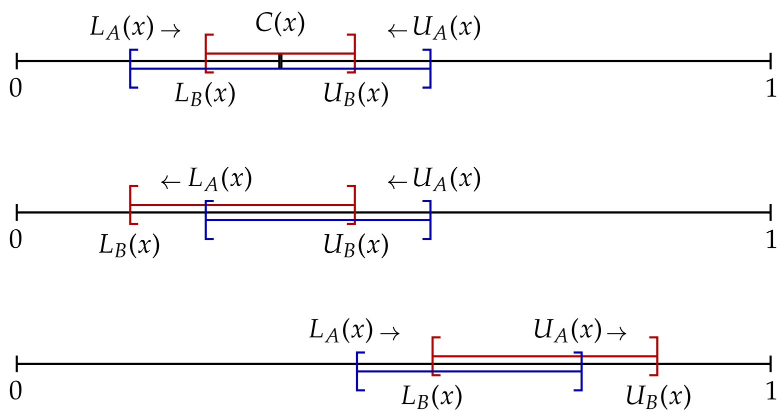

Figure 2. Orders in intervals (Intervals A in blue and B in red). The top shows the interval order for the relationship $B \preceq A$, while the center and bottom show the interval order for relationship $B \preccurlyeq A$.

Consequently, when analyzing the orders used in Axioms 2.4 and 3.4 to compare intervals, we realize that, for Burillo, two intervals are comparable only if one is inside the other ($\subseteq$), while for Zeng, two intervals are comparable only if one is shifted to the left (or right) of the other ($\leq$). That is, assuming $L_A(x) < L_B(x)$, we obtain the following conditions:

1. If $U_A(x) < U_B(x)$, then A and B are not comparable for $\preceq$.
2. If $U_B(x) < U_A(x)$, then A and B are not comparable for $\preccurlyeq$.

Therefore, for two intervals to be comparable in both approaches simultaneously, we need $L_A(x) = L_B(x)$ or $U_A(x) = U_B(x)$.

Let us now compare the four entropy measures presented above. The entropy measures proposed by Burillo, Zeng, Szmidt, and Vlachos, originally defined in terms of the lower and upper bound of an interval, are here rewritten, now considering the interval in terms of its centroid and half-width, denoted by $C_A(x)$ and $\delta_A(x)$, respectively. Both centroid and half-width in terms of the lower and upper bounds of the interval are as follows:

$$C_A(x) = \frac{L_A(x) + U_A(x)}{2}, \tag{5}$$

$$\delta_A(x) = \frac{U_A(x) - L_A(x)}{2}. \tag{6}$$

The function C_A, which applies to each $x \in X$ the value $C_A(x)$, is referred to as the centroid function of A.

Therefore, an interval $[L_A(x), U_A(x)]$ can also be expressed in terms of centroid and half-width as $[C_A(x) - \delta_A(x), C_A(x) + \delta_A(x)]$. Furthermore, considering that the maximum value for $\delta_A(x)$ is bounded by the distance to the closest endpoint of the interval, it is clear that

$$\delta_A(x) \leq \min(C_A(x), 1 - C_A(x)).$$

New expressions for the four entropy measures can now easily be derived. Burillo's entropy can then be rewritten as follows:

$$I(A) = \frac{1}{N}\sum_{i=1}^{N}(U_A(x_i) - L_A(x_i)) = \frac{1}{N}\sum_{i=1}^{N}2\delta_A(x_i), \tag{7}$$

Zeng's entropy as

$$\begin{aligned} E_Z(A) &= 1 - \frac{1}{N}\sum_{i=1}^{N}|L_A(x_i) + U_A(x_i) - 1| \\ &= 1 - \frac{1}{N}\sum_{i=1}^{N}|2C_A(x_i) - 1| \\ &= 1 - \frac{2}{N}\sum_{i=1}^{N}\left|C_A(x_i) - \frac{1}{2}\right|, \end{aligned} \tag{8}$$

Szmidt's entropy as

$$\begin{aligned} E_S(A) &= \frac{1}{N}\sum_{i=1}^{N}\frac{1 - \max(1 - U_A(x_i), L_A(x_i))}{1 - \min(1 - U_A(x_i), L_A(x_i))} \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{\min(C_A(x_i), (1 - C_A(x_i))) + \delta_A(x_i)}{\max(C_A(x_i), (1 - C_A(x_i))) + \delta_A(x_i)}, \end{aligned} \tag{9}$$

and, finally, Vlachos' entropy as

$$E_V(A) = \frac{1}{N}\sum_{i=1}^{N}\frac{\min(L_A(x_i), 1 - U_A(x_i)) + \min(U_A(x_i), 1 - L_A(x_i))}{\max(L_A(x_i), 1 - U_A(x_i)) + \max(U_A(x_i), 1 - L_A(x_i))} \tag{10}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{\min(C_A(x_i), 1 - C_A(x_i))}{\max(C_A(x_i), 1 - C_A(x_i))}. \tag{11}$$

Comparing Different Entropy Measures

With this reinterpretation of the four entropy measures, we can now easily analyze and compare $I(A)$, $E_Z(A)$, $E_S(A)$, and $E_V(A)$ considering the following propositions already introduced in [20].

Proposition 1. *$E_S(A) \leq E_Z(A)$ $\forall A$, with $E_S(A) = E_Z(A)$ only when $\delta_A(x_i) = \min(C_A(x_i), 1 - C_A(x_i))$ $\forall i$, that is, when, for every interval, either $L_A(x_i) = 0$ or $U_A(x_i) = 1$.*

Proposition 2. *$E_V(A) \leq E_S(A)$ $\forall A$, with $E_V(A) = E_S(A)$ only when $\delta_A(x_i) = 0$ $\forall i$, that is, when working with a type-1 fuzzy set.*

Proposition 3. *$I(A) \leq E_S(A) \leq E_Z(A)$ $\forall A$, with $I(A) = E_S(A) = E_Z(A)$ only when $\delta_A(x_i) = \min(C_A(x_i), 1 - C_A(x_i))$ $\forall i$, that is, when, for every interval, either $L_A(x_i) = 0$ or $U_A(x_i) = 1$.*

In summary, it holds that

$$E_V(A) \leq E_S(A) \leq E_Z(A) \quad \forall A, \tag{12}$$

and

$$I(A) \leq E_S(A) \quad \forall A. \tag{13}$$

However, we cannot define an overall total order since, in some cases, we have $I(A) < E_V(A)$, while in others we have $E_V(A) < I(A)$ (as shown in Example 1).

It should also be noted that, while $I(A)$ considers only the width of the interval (the value 0 corresponds to intervals with $\delta_A(x_i) = 0$, that is, type-1 fuzzy sets), $E_Z(A)$ and $E_V(A)$ only refer to the centroid of the interval (the value 0 corresponds to intervals with $C_A(x_i) \in \{0,1\}$, that is, crisp sets). Instead, $E_S(A)$ considers both the centroid and width (the value 0 corresponds to intervals with $C_A(x_i) \in \{0,1\}$ and, consequently, $\delta_A(x_i) = 0$, that is, again, crisp sets).

These ideas, together with the inequalities given by Equations (12) and (13), are illustrated in Figures 3 and 4, and in Example 1.

First, Figure 3 shows two IVFSs (A and B) that have the same centroid function, that is, both produce the same closest type-1 fuzzy set (red dashed set in A). In fact, this type-1 fuzzy set is B. As a result $E_V(A) = E_V(B)$, $E_Z(A) = E_Z(B)$. However, at the same time, A being the IVFS with maximum fuzziness for this centroid function, and B being the IVFS with minimum fuzziness for this centroid function, we have that

$$I(A) = E_S(A) = E_Z(A),$$

and

$$I(B) = 0, \quad E_S(B) = E_V(B).$$

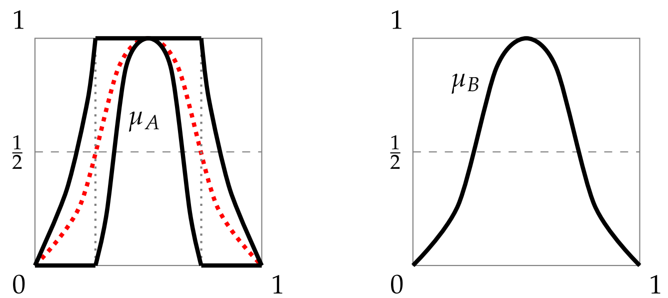

Figure 3. Two IVFSs with the same centroid function: A with maximum width of intervals (centroid function represented in red) and B with minimum width of intervals.

Figure 4 illustrates how $E_V(A(x_i)) \leq E_S(A(x_i)) \leq E_Z(A(x_i))$, and, consequently, how $E_V(A) \leq E_S(A) \leq E_Z(A), \forall A$.

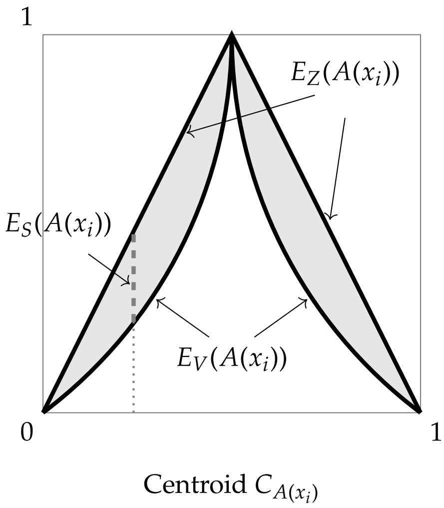

Figure 4. Ordering among entropy measures E_Z (upper curve), E_S (taking values in the shaded area), and E_V (lower curve) with respect to centroid.

Finally, Example 1 considers different situations showcasing the previous ideas.

Example 1. *Let A, B, and C be interval-valued fuzzy sets defined as follows:*

$$A = \{\langle x_1, [0,0]\rangle, \langle x_2, [0.4, 0.4]\rangle, \langle x_3, [1,1]\rangle, \langle x_4, [0.6, 0.6]\rangle, \langle x_5, [0,0]\rangle\},$$
$$B = \{\langle x_1, [0, 0.4]\rangle, \langle x_2, [0.2, 0.8]\rangle, \langle x_3, [0.5, 1]\rangle, \langle x_4, [0.3, 0.7]\rangle, \langle x_5, [0, 0.5]\rangle\},$$
$$C = \{\langle x_1, [0, 0.4]\rangle, \langle x_2, [0, 0.8]\rangle, \langle x_3, [0.5, 1]\rangle, \langle x_4, [0.2, 1]\rangle, \langle x_5, [0, 0.5]\rangle\}.$$

For the set A, the entropy measures are

$$I(A) = 0,$$
$$E_Z(A) = 1 - \frac{1}{5}(1 + 0.2 + 1 + 0.2 + 1) = 0.32,$$
$$E_S(A) = \frac{1}{5}(0 + \frac{2}{3} + 0 + \frac{2}{3} + 0) = \frac{4}{15} = 0.27,$$
$$E_V(A) = \frac{1}{5}(0 + \frac{2}{3} + 0 + \frac{2}{3} + 0) = \frac{4}{15} = 0.27.$$

For the set B, they are

$$I(B) = \frac{1}{5}(0.4 + 0.6 + 0.5 + 0.4 + 0.5) = 0.48,$$
$$E_Z(B) = 1 - \frac{1}{5}(0.6 + 0 + 0.5 + 0 + 0.5) = 0.68,$$
$$E_S(B) = \frac{1}{5}(0.2 + 1 + 0.5 + 1 + 0.5) = 0.64,$$
$$E_V(B) = \frac{1}{5}(0.25 + 1 + \frac{1}{3} + 1 + \frac{1}{3}) = \frac{7}{12} = 0.58.$$

Furthermore, finally, for the set C, they are

$$\begin{aligned} I(C) &= \frac{1}{5}(0.4+0.8+0.5+0.8+0.5) = 0.6, \\ E_Z(C) &= 1 - \frac{1}{5}(0.6+0.2+0.5+0.2+0.5) = 0.6, \\ E_S(C) &= \frac{1}{5}(0.4+0.8+0.5+0.8+0.5) = 0.6, \\ E_V(C) &= \frac{1}{5}(0.25+\frac{2}{3}+\frac{1}{3}+\frac{2}{3}+\frac{1}{3}) = 0.45. \end{aligned}$$

Thus, we have that

$$\begin{aligned} I(A) < E_V(A) = E_S(A) < E_Z(A), \\ I(B) < E_V(B) < E_S(B) < E_Z(B), \\ E_V(C) < I(C) = E_S(C) = E_Z(C). \end{aligned}$$

As expected, according to Equations (12) and (13), we see in these examples that the inequalities $E_V(Y) \leq E_S(Y) \leq E_Z(Y)$ and $I(Y) \leq E_S(Y)$, respectively, hold, while it is shown that there is no ordering between $I(Y)$ and $E_V(Y)$.

3. Entropy in Type-2 Fuzzy Sets

Type-2 fuzzy sets were introduced by Zadeh in 1975 [19] as an extension of fuzzy sets (type-1 fuzzy sets). A type-2 fuzzy set (T2FS), according to the definition given by Mizumoto and Tanaka [17,18], is a set A in which the membership degree of any element is a fuzzy set on $[0,1]$ (see Figure 5). Therefore, it is determined by a membership function

$$\mu_A\colon X \to [0,1]^{[0,1]} = \text{Map}\,([0,1],[0,1]),$$

that is, $\mu_A(x)$ is a fuzzy set in the interval $[0,1]$ and, also, the membership degree of an element $x \in X$ to the set A. Equivalently, for any $x \in X$, $\mu_A(x)$ (which we denote for convenience by $\mu(x)$, as in Figure 5, μ_x or simply μ_A) is a function from $[0,1]$ to $[0,1]$. In the following, we denote by T2FS(X) the set of all type-2 fuzzy sets in a finite universe X. It should be noted that, although throughout this paper we only consider entropy measures defined in a finite universe, the set of properties and entropy measures proposed here can be easily adapted to any universe.

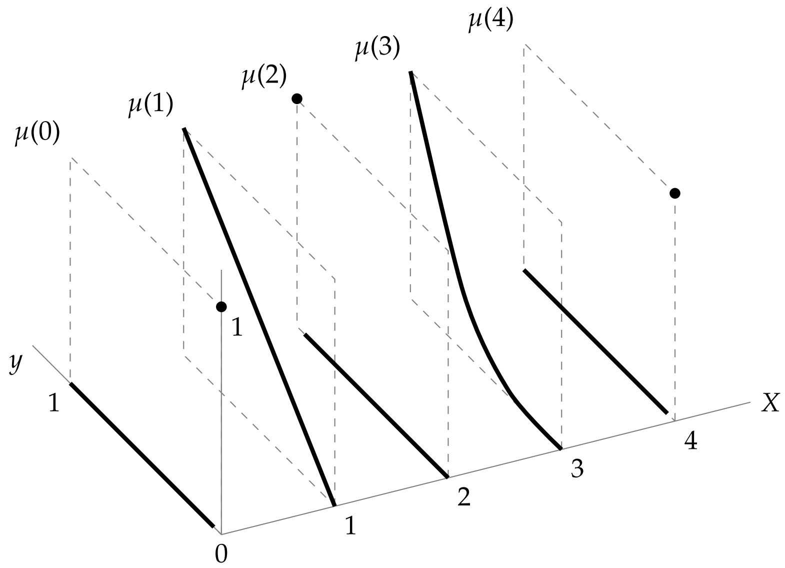

Figure 5. Example of a T2FS with discrete universe $X = \{0, 1, 2, 3, 4\}$.

It should be noted that it is possible to perceive the membership degrees of a T2FS as labels of the variable truth. In practical situations, these labels are usually normal (that is, $\sup\{f(x) : x \in [0,1]\} = 1$) and convex (that is, for any $x \leq y \leq z$, it holds that $f(y) \geq f(x) \wedge f(z)$). In this work, we will only consider this type of label, and we will denote the set of all such labels by

$$\mathcal{L} = \{f \in [0,1]^{[0,1]} : f \text{ normal and convex}\}. \tag{14}$$

The idea shown in Section 2, behind the definition of different entropy measures in interval-valued fuzzy sets, suggests the possibility of doing a similar approach in the case of type-2 fuzzy sets. Therefore, based on this idea, we first study how to introduce two different types of entropy measures by making a comparison between type-2 fuzzy sets and crisp or fuzzy sets. In other words, we want to study whether it is possible to find a measure of how non-crisp (or non-fuzzy) a type-2 fuzzy set is. Furthermore, in the case of type-2 fuzzy sets, we can also propose a third type of measure of how much a set is non-interval-valued.

3.1. Type-C Entropy in Type-2 Fuzzy Sets

In a first approach, we define the entropy of a T2FS as a measure of how much it is non-crisp, and we will refer to this entropy as Type-C, where C stands for crisp. Then, we begin by studying the entropy values at its endpoints (Axioms 3.1 and 3.2 proposed by Zeng, see Section 2). On the one hand, we assume that the entropy of a crisp set must be 0. Let us note that the membership degrees of a crisp T2FS could only be the singleton labels $\bar{0}$ or $\bar{1}$:

- $\bar{0}(0) = 1$ and $\bar{0}(y) = 0 \quad \forall y \in [0,1], y \neq 0$;
- $\bar{1}(1) = 1$ and $\bar{1}(y) = 0 \quad \forall y \in [0,1], y \neq 1$.

Thus, the entropy must be 0 when $\mu(x) = \bar{0}$ or $\mu(x) = \bar{1}$, $\forall x \in X$ (as shown in Figure 6).

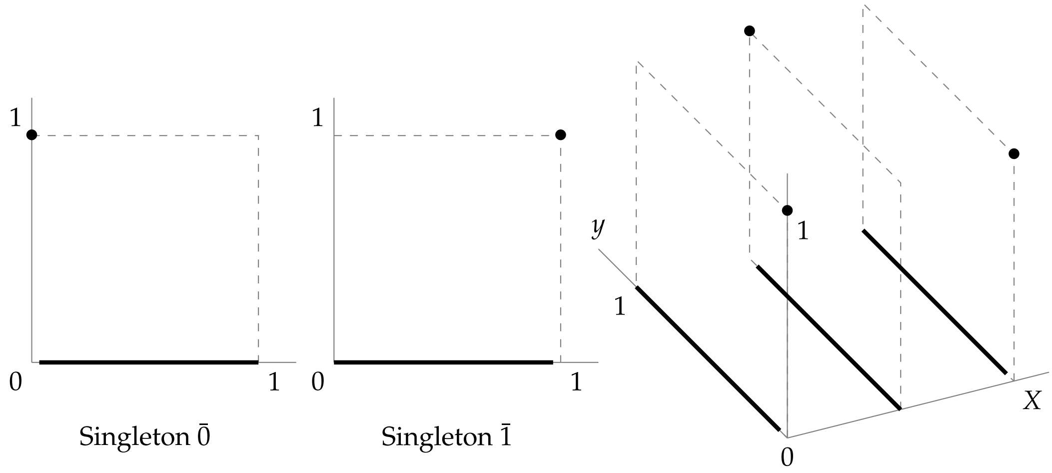

Figure 6. Crisp type-2 fuzzy set.

On the other hand, the entropy must take a maximum value when the set is as non-crisp as possible. That is, its labels should be as far away from the two singletons $\bar{0}$ and $\bar{1}$ as possible. Let $\mu(x)$ be a singleton $\bar{a}$, that is, $\bar{a}(a) = 1$ and $\bar{a}(y) = 0$, $\forall y \in [0,1], y \neq a$. The one furthest from the singletons $\bar{0}$ and $\bar{1}$ is the singleton $\bar{\frac{1}{2}}$. Thus, the entropy must be at least maximal when all the membership degrees are the singleton $\bar{\frac{1}{2}}$.

Furthermore, we have to modify the order $\preccurlyeq$ proposed by Zeng in Axiom 3.4 (see Section 2). This order, proposed for IVFSs, has to be adapted for the T2FS case. We then introduce the order $\sqsubseteq$, defined on the membership degrees of T2FS and characterized in the set $\mathcal{L}$ of functions from $[0,1]$ to $[0,1]$, normal and convex (see [25,26]). To this end, we first introduce some auxiliary functions as in [25,26], give their properties and illustrate their meaning.

Definition 1 ([26])**.** *Let $f \in [0,1]^{[0,1]}$; we define $f^L, f^R \in [0,1]^{[0,1]}$ as follows:*

$$f^L(x) = \sup\{f(y) : y \le x\}, \quad f^R(x) = \sup\{f(y) : y \ge x\}.$$

It should be noted that f^L and f^R are monotonically increasing and decreasing, respectively, but it also holds that $f \le f^L$, $f \le f^R$, $(f^L)^L = f^L$, and $(f^R)^R = f^R$, for all $f \in [0,1]^{[0,1]}$, where $\le$ is the usual order in the set of functions (that is, $f \le g$ if and only if $f(x) \le g(x)$, $\forall x \in [0,1]$). An example of both f^L and f^R is shown in Figure 7.

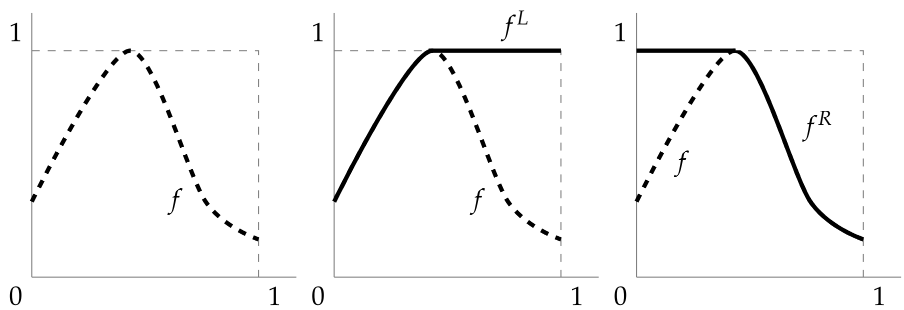

Figure 7. Example of f^L and f^R (solid lines).

The following characterization of the partial order $\sqsubseteq$ on $\mathcal{L}$ was given in [25] (see Figure 8).

Theorem 1 ([25])**.** *Let $f, g \in \mathcal{L}$. Then, $f \sqsubseteq g$ if and only if $g^L \le f^L$ and $f^R \le g^R$.*

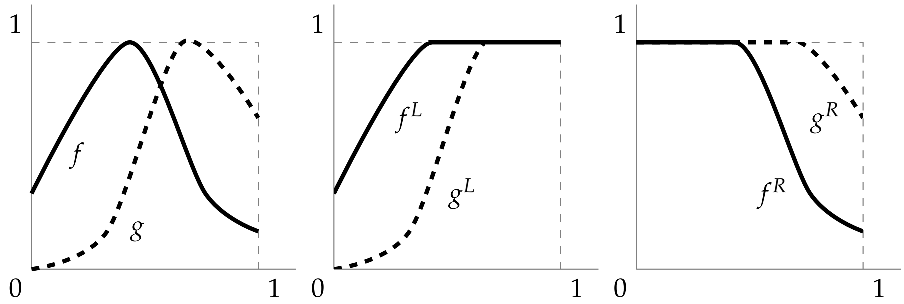

Figure 8. Example where $f \sqsubseteq g$ (f solid line and g dashed line).

Let us suppose now that the membership degree μ_A in a T2FS is a characteristic function of an interval, that is, $\mu_A = \overline{[L_A, U_A]}$ such that $\mu_A(y) = 1$ when $y \in [L_A, U_A]$, and $\mu_A(y) = 0$ otherwise. Then, it is verified that (see Figure 9).

$$U_A = \int_{y=0}^{y=1} (\mu_A)^R(y)\, dy, \tag{15}$$

$$L_A = 1 - \int_{y=0}^{y=1} (\mu_A)^L(y)\, dy. \tag{16}$$

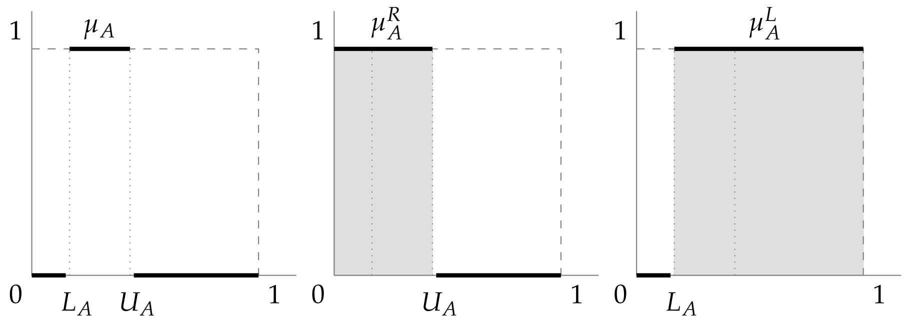

Figure 9. Reinterpreting U_A and L_A in terms of the area (shaded region) under the curves μ_A^R and μ_A^L, that is, the integral in Equations (16) and (15), μ_A being the characteristic function of an interval.

We then propose the following extension of the membership degree in IVFSs (that is, endpoints in an interval) to a membership degree in T2FSs. In other words, while in IVFSs we use the interval endpoints $[L_A(x_i), U_A(x_i)]$ to identify the membership degree, we propose to modify these endpoints to identify the membership degree in T2FSs as follows (see Figure 10):

$$\left[1 - \int_{y=0}^{y=1} (\mu_A(x_i))^L(y)\, dy, \int_{y=0}^{y=1} (\mu_A(x_i))^R(y)\, dy\right].$$

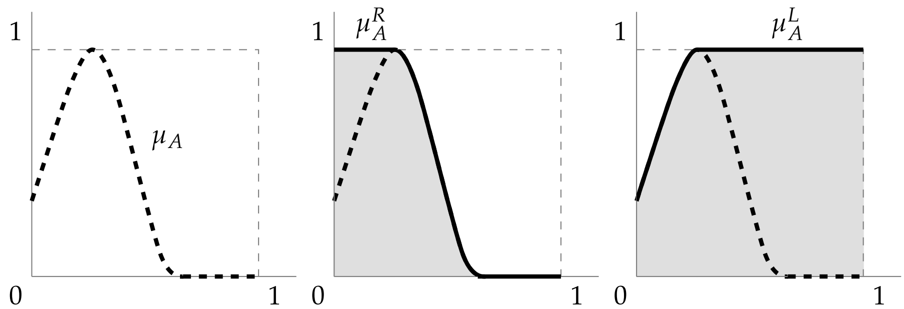

Figure 10. Areas (shaded regions) under the curves μ_A^R and μ_A^L (solid lines).

Similarly, the interval $[L_B(x_i), U_B(x_i)]$ is now as follows:

$$\left[1 - \int_{y=0}^{y=1} (\mu_B(x_i))^L(y)\, dy, \int_{y=0}^{y=1} (\mu_B(x_i))^R(y)\, dy\right].$$

Furthermore, the inequality $L_B(x_i) + U_B(x_i) \leq 1$ in IVFSs becomes, in $\mathcal{L}$,

$$1 - \int_{y=0}^{y=1} (\mu_B(x_i))^L(y)\, dy + \int_{y=0}^{y=1} (\mu_B(x_i))^R(y)\, dy \leq 1 \Leftrightarrow$$
$$\Leftrightarrow \int_{y=0}^{y=1} (\mu_B(x_i))^L(y)\, dy \geq \int_{y=0}^{y=1} (\mu_B(x_i))^R(y)\, dy.$$

Thus, the relationship $L_A(x_i) \leq L_B(x_i)$; $U_A(x_i) \leq U_B(x_i)$ is now rewritten as $\mu_A(x_i) \sqsubseteq \mu_B(x_i)$. Finally, we can define the following order relation on T2FSs with membership degrees in $\mathcal{L}$, that is, an extension of the order $\preccurlyeq$ defined in Section 2.

Definition 2. *Let A and B be type-2 fuzzy sets with membership degrees $\mu_A(x)$ and $\mu_B(x)$, respectively, which are normal and convex functions for each $x \in X$. Then, $A \sqsubseteq_{LR} B$ if and only if $\forall x \in X$ the following conditions are met:*

- $\mu_A(x) \sqsubseteq \mu_B(x)$ *when* $\int_{y=0}^{y=1} (\mu_B(x))^L(y)\, dy > \int_{y=0}^{y=1} (\mu_B(x))^R(y)\, dy$;
- $\mu_A(x) = \mu_B(x)$ *when* $\int_{y=0}^{y=1} (\mu_B(x))^L(y)\, dy = \int_{y=0}^{y=1} (\mu_B(x))^R(y)\, dy$;
- $\mu_B(x) \sqsubseteq \mu_A(x)$ *when* $\int_{y=0}^{y=1} (\mu_B(x))^L(y)\, dy < \int_{y=0}^{y=1} (\mu_B(x))^R(y)\, dy$.

It should be noted that, if $\int_{y=0}^{y=1} (\mu_B(x))^L(y)\, dy > \int_{y=0}^{y=1} (\mu_B(x))^R(y)\, dy$ and $\mu_A(x) \sqsubseteq \mu_B(x)$, then

$$\int_{y=0}^{y=1} (\mu_A(x))^L(y)\, dy > \int_{y=0}^{y=1} (\mu_A(x))^R(y)\, dy,$$

since

$$\int_{y=0}^{y=1} (\mu_A(x))^L(y)\, dy \geq \int_{y=0}^{y=1} (\mu_B(x))^L(y)\, dy > \int_{y=0}^{y=1} (\mu_B(x))^R(y)\, dy \geq \int_{y=0}^{y=1} (\mu_A(x))^R(y)\, dy.$$

Similarly, if $\int_{y=0}^{y=1} (\mu_B(x))^R(y)\, dy > \int_{y=0}^{y=1} (\mu_B(x))^L(y)\, dy$ and $\mu_B(x) \sqsubseteq \mu_A(x)$, then

$$\int_{y=0}^{y=1} (\mu_A(x))^R(y)\, dy > \int_{y=0}^{y=1} (\mu_A(x))^L(y)\, dy.$$

Furthermore, it is easy to prove that the relation $\sqsubseteq_{LR}$ is antisymmetric and transitive.

Examples of both the partial order $\sqsubseteq$ in $\mathcal{L}$ and the order relation $\sqsubseteq_{LR}$ in T2FSs with membership degrees in $\mathcal{L}$ are illustrated in Figures 8 and 11, respectively.

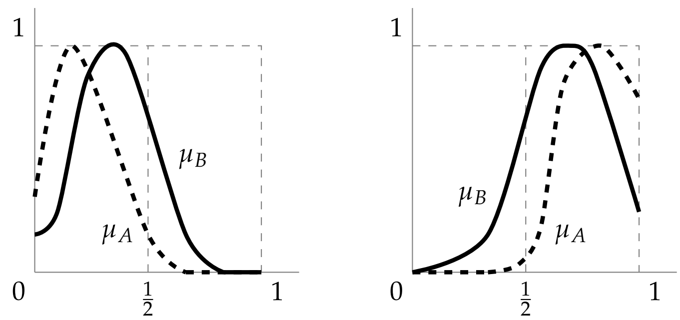

Figure 11. Examples where $A \sqsubseteq_{LR} B$ (A dashed line and B solid line).

Next, based on what has been exposed above, we propose the following axiomatization.

Definition 3. *Let A and B be type-2 fuzzy sets with membership degrees $\mu_A(x)$ and $\mu_B(x)$, respectively, which are normal and convex functions for each $x \in X$. We say that a function $E_C\colon \mathrm{T2FS(X)} \to [0,1]$ is a Type-C entropy measure on* T2FS(X) *if it satisfies the following axioms:*

C.1 $E_C(A) = 0 \Leftrightarrow \mu_A(x) = \bar{0}$ *or* $\mu_A(x) = \bar{1}$, $\forall x \in X$.

C.2 $E_C(A) = 1 \Leftrightarrow \int_{y=0}^{y=1} (\mu_A(x))^L(y)\,dy = \int_{y=0}^{y=1} (\mu_A(x))^R(y)\,dy$, $\forall x \in X$.

C.3 $E_C(A) = E_C(A^c)$, *A^c being the complementary set of A, where* $(\mu_{A^c}(x))(y) = (\mu_A(x))(1-y)$, $\forall x \in X, \forall y \in [0,1]$.

C.4 *If* $A \sqsubseteq_{LR} B$, *then* $E_C(A) \le E_C(B)$.

It is at this point where considering membership degrees as labels of the variable truth, or simply as fuzzy sets, makes the difference. This question is considered in the following remarks.

Remark 1. *Axiom C.1 considers that a type-2 fuzzy set with crisp membership degrees is one that has $\mu_A(x) = \bar{0}$ or $\mu_A(x) = \bar{1}$, $\forall x \in X$. This is consistent with the use of labels of the variable truth. On the other hand, when simply using fuzzy sets as membership degrees, a type-2 fuzzy set with crisp membership degrees is the one having $(\mu_A(x))(y) = 0$ or $(\mu_A(x))(y) = 1$, $\forall x \in X, y \in [0,1]$. This is functionally equivalent to considering the type-2 fuzzy set as a type-1 fuzzy set defined in a two-dimensional domain (the second dimension being the interval $[0,1]$). Obviously, both interpretations differ.*

Remark 2. *Going now to Axiom C.3, considering membership degrees as labels of the variable truth will lead to a definition of the complementary set of a type-2 fuzzy set as $(\mu_{A^c}(x))(y) = (\mu_A(x))(1-y)$, $\forall x \in X, \forall y \in [0,1]$. On the other hand, considering membership degrees simply as fuzzy sets will lead to a complementary set (implicitly) defined as $(\mu_{A^c}(x))(y) = 1 - (\mu_A(x))(y)$, $\forall x \in X, \forall y \in [0,1]$. The point is that this is not the complementary, this is the negation of A (see [17]). Furthermore, that is the reason why some authors replace the axiom stating that "the entropy of a set and its complementary set should be equal", with a new one stating that "the entropy of a set and its negated set should be equal". However, this is not consistent with the definition of entropy for fuzzy sets and interval-valued fuzzy sets.*

Some examples of entropy measures that satisfy Definition 3 are given below.

Proposition 4. *Let $A \in$ T2FS(X) be any type-2 fuzzy set on X, with a finite universe $|X| = N$, and membership degrees $\mu_A(x)$ being normal and convex functions for each $x \in X$. The following functions are Type-C entropy measures on* T2FS(X)*:*

$$E_C^Z(A) = 1 - \frac{1}{N}\sum_{i=1}^{N}\left|F_A^R(x_i) - F_A^L(x_i)\right|, \tag{17}$$

$$E_C^S(A) = \frac{1}{N}\sum_{i=1}^{N}\frac{\min(F_A^R(x_i), F_A^L(x_i))}{\max(F_A^R(x_i), F_A^L(x_i))}, \tag{18}$$

$$E_C^V(A) = \frac{1}{N}\sum_{i=1}^{N}\frac{\min(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \min(F_A^R(x_i), F_A^L(x_i))}{\max(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \max(F_A^R(x_i), F_A^L(x_i))}, \tag{19}$$

where

$$F_A^L(x_i) = \int_{y=0}^{y=1} (\mu_A(x_i))^L(y)\,dy, \tag{20}$$

$$F_A^R(x_i) = \int_{y=0}^{y=1} (\mu_A(x_i))^R(y)\,dy, \tag{21}$$

The proof of this proposition is given in the Appendix A.

It should be noted that we use the superscripts Z, S, and V for the three proposed Type-C entropy measures, because they are based on the original measures proposed for IVFSs by Zeng, Szmidt, and Vlachos, respectively. In fact, these entropy measures for T2FSs are extensions of the entropy measures E_Z, E_S, and E_V given for IVFSs. We recall the extension of interval endpoints in IVFSs to T2FSs:

$$[L_A(x_i), U_A(x_i)] \to \left[1 - F_A^L(x_i), F_A^R(x_i)\right]. \tag{22}$$

With this transformation, we obtain the following extensions:

$$\begin{aligned} E_Z(A) &= 1 - \frac{1}{N}\sum_{i=1}^{N} |U_A(x_i) - (1 - L_A(x_i))| \to \\ E_C^Z(A) &= 1 - \frac{1}{N}\sum_{i=1}^{N} \left|F_A^R(x_i) - F_A^L(x_i)\right|, \end{aligned} \tag{23}$$

$$\begin{aligned} E_S(A) &= \frac{1}{N}\sum_{i=1}^{N} \frac{\min(U_A(x_i), 1 - L_A(x_i))}{\max(U_A(x_i), 1 - L_A(x_i))} \to \\ E_C^S(A) &= \frac{1}{N}\sum_{i=1}^{N} \frac{\min(F_A^R(x_i), F_A^L(x_i))}{\max(F_A^R(x_i), F_A^L(x_i))}, \end{aligned} \tag{24}$$

and

$$\begin{aligned} E_V(A) &= \frac{1}{N}\sum_{i=1}^{N} \frac{\min(L_A(x_i), 1 - U_A(x_i)) + \min(U_A(x_i), 1 - L_A(x_i))}{\max(L_A(x_i), 1 - U_A(x_i)) + \max(U_A(x_i), 1 - L_A(x_i))} \to \\ E_C^V(A) &= \frac{1}{N}\sum_{i=1}^{N} \frac{\min(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \min(F_A^R(x_i), F_A^L(x_i))}{\max(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \max(F_A^R(x_i), F_A^L(x_i))}. \end{aligned} \tag{25}$$

Remark 3. *It should be noted that $E_C^S(A) = E_C^V(A)$ when $F_A^L(x_i) + F_A^R(x_i) = 1$ $\forall i = 1, \ldots, n$ (that is, $\mu_A(x_i)$ is a singleton and, thus, its integral is 0), and $E_C^S(A) = E_C^Z(A)$ when $F_A^L(x_i) = 1$ or $F_A^R(x_i) = 1$ (that is, $\forall i = 1, \ldots, n$, $\mu_A(x_i)(0) = 1$ or $\mu_A(x_i)(1) = 1$).*

Let us now compare these three entropy measures. First, without loss of generality, we suppose that

$$F_A^L(x_i) = \int_{y=0}^{y=1} (\mu_A(x_i))^L(y)\, dy < \int_{y=0}^{y=1} (\mu_A(x_i))^R(y)\, dy = F_A^R(x_i).$$

It should be noted that, by symmetry, the result does not change if we assume the opposite, that is, $F_A^R(x_i) < F_A^L(x_i)$. It should also be noted that $\int_{y=0}^{y=1} (\mu_A(x_i))(y)\, dy = F_A^R(x_i) + F_A^L(x_i) - 1$, that is,

$$\int_{y=0}^{y=1} (\mu_A(x_i))(y)\, dy = \int_{y=0}^{y=1} (\mu_A(x_i))^R(y)\, dy + \int_{y=0}^{y=1} (\mu_A(x_i))^L(y)\, dy - 1.$$

Then, it holds that

$$\frac{\min(F_A^R(x_i), F_A^L(x_i))}{\max(F_A^R(x_i), F_A^L(x_i))} = \frac{F_A^L(x_i)}{F_A^R(x_i)},$$

and, also,

$$\frac{\min(1-F^L_A(x_i), 1-F^R_A(x_i)) + \min(F^R_A(x_i), F^L_A(x_i))}{\max(1-F^L_A(x_i), 1-F^R_A(x_i)) + \max(F^R_A(x_i), F^L_A(x_i))} =$$
$$= \frac{1-F^R_A(x_i)+F^L_A(x_i)}{1-F^L_A(x_i)+F^R_A(x_i)} = \frac{2F^L_A(x_i) - \int_{y=0}^{y=1}(\mu_A(x_i))(y)\,dy}{2F^R_A(x_i) - \int_{y=0}^{y=1}(\mu_A(x_i))(y)\,dy}.$$

Since $\int_{y=0}^{y=1}(\mu_A(x_i))(y)\,dy \leq F^L_A(x_i) < F^R_A(x_i)$, it follows that

$$\frac{2F^L_A(x_i) - \int_{y=0}^{y=1}(\mu_A(x_i))(y)\,dy}{2F^R_A(x_i) - \int_{y=0}^{y=1}(\mu_A(x_i))(y)\,dy} \leq \frac{2F^L_A(x_i)}{2F^R_A(x_i)} = \frac{F^L_A(x_i)}{F^R_A(x_i)}.$$

Therefore,

$$\frac{\min(1-F^L_A(x_i), 1-F^R_A(x_i)) + \min(F^R_A(x_i), F^L_A(x_i))}{\max(1-F^L_A(x_i), 1-F^R_A(x_i)) + \max(F^R_A(x_i), F^L_A(x_i))} \leq \frac{\min(F^R_A(x_i), F^L_A(x_i))}{\max(F^R_A(x_i), F^L_A(x_i))},$$

that is, $E^V_C(A) \leq E^S_C(A)$.

On the other hand, let us rewrite $E^Z_C(A)$ as follows:

$$\begin{aligned} E^Z_C(A) &= 1 - \frac{1}{N}\sum_{i=1}^{N}\left|F^R_A(x_i) - F^L_A(x_i)\right| \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(1 - \left|F^R_A(x_i) - F^L_A(x_i)\right|\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(1 - F^R_A(x_i) + F^L_A(x_i)\right), \end{aligned}$$

where the last equality holds, since $F^L_A(x_i) < F^R_A(x_i)$.

It is straightforward to verify that

$$\frac{F^L_A(x_i)}{F^R_A(x_i)} \leq 1 - F^R_A(x_i) + F^L_A(x_i).$$

Indeed, this implies that

$$\begin{aligned} F^L_A(x_i) &\leq F^R_A(x_i) - F^R_A(x_i)F^R_A(x_i) + F^L_A(x_i)F^R_A(x_i) \\ F^L_A(x_i) - F^L_A(x_i)F^R_A(x_i) &\leq F^R_A(x_i) - F^R_A(x_i)F^R_A(x_i) \\ F^L_A(x_i)(1-F^R_A(x_i)) &\leq F^R_A(x_i)(1-F^R_A(x_i)), \end{aligned}$$

which is satisfied since by the assumptions $F^L_A(x_i) < F^R_A(x_i)$ and $0 \leq F^L_A(x_i), F^R_A(x_i) \leq 1$. Therefore, the inequality $E^S_C(A) \leq E^Z_C(A)$ also holds.

In summary, it holds that

$$E^V_C(A) \leq E^S_C(A) \leq E^Z_C(A) \quad \forall A \in \text{T2FS(X)}. \tag{26}$$

It should be noted that this result is similar to the case of interval-valued fuzzy sets, in which $E_V(A)$, $E_S(A)$ and $E_Z(A)$, $\forall A \in \text{IVFS}(X)$ satisfy the same relation (see Section 2).

Example 2. *Let* $X = \{a, b, c\}$ *and* $\chi \colon X \to [0,1]^{[0,1]}$*, such that*

$$\chi(a) = \chi_a = \frac{\bar{1}}{4},$$

$$\chi(b) = \chi_b, \quad \text{where } \chi_b(y) = \begin{cases} 8y-4 & \text{if } \frac{1}{2} \leq y \leq \frac{5}{8} \\ 1 & \text{if } \frac{5}{8} \leq y \leq \frac{7}{8} \\ 8-8y & \text{if } \frac{7}{8} \leq y \leq 1 \\ 0 & \text{elsewhere}, \end{cases}$$

$$\chi(c) = \chi_c, \quad \text{where } \chi_c(y) = \begin{cases} 2y & \text{if } y \leq \frac{1}{2} \\ 2-2y & \text{if } y \geq \frac{1}{2}. \end{cases}$$

These membership degrees thus defined are illustrated in Figure 12.

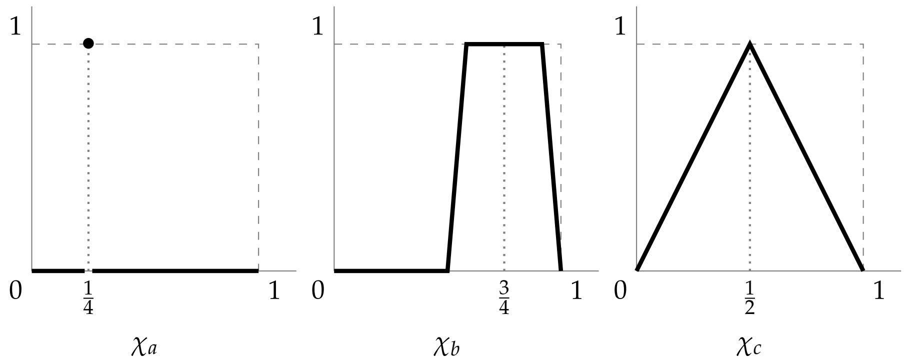

Figure 12. Membership degrees in the type-2 fuzzy set of Example 2.

Then,

$$F_\chi^L(a) = \int_{y=0}^{y=1} (\chi(a))^L(y)\,dy = \frac{3}{4},\ F_\chi^L(b) = \frac{7}{16},\ F_\chi^L(c) = \frac{3}{4},$$
$$F_\chi^R(a) = \int_{y=0}^{y=1} (\chi(a))^R(y)\,dy = \frac{1}{4},\ F_\chi^R(b) = \frac{15}{16},\ F_\chi^R(c) = \frac{3}{4},$$

and, thus, $E_C^Z(\chi) \approx 0.666$, $E_C^V(\chi) \approx 0.555$, *and* $E_C^S(\chi) = 0.6$.

In the example, it holds that $E_C^V(\chi) \leq E_C^S(\chi) \leq E_C^Z(\chi)$*, as expected since it was proved above for any type-2 fuzzy set; see Equation (26).*

3.2. Type-F Entropy in Type-2 Fuzzy Sets

Following a different approach, we now redefine the entropy of a T2FS as a measure of how much the set is non-fuzzy, and we will refer to this entropy as Type-F, where F stands for fuzzy. Again, as in Section 3.1, we begin by studying the entropy at its endpoints. On the one hand, the entropy of a set must be 0 if it is indeed a fuzzy set, that is, the label associated with any element of the universe is a singleton. Then, $\forall x \in X$, we have that $\mu_A(x) = \bar{a}$ with $a \in [0,1]$ (as shown in Figure 13), that is,

$$(\mu_A(x))(y) = \bar{a}(y) = \begin{cases} 1 & \text{if } y = a \\ 0 & \text{if } y \neq a. \end{cases}$$

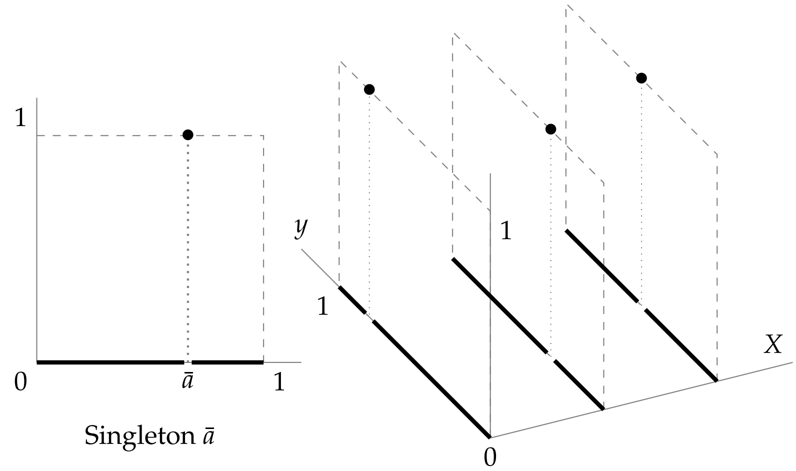

Figure 13. Type-2 fuzzy set being a fuzzy set.

On the other hand, the entropy must take a maximum value when the membership degree of any element is as far from being a singleton as possible, that is, when the label $\mathbf{1_{(0,1)}}$ is such that

$$\mathbf{1_{(0,1)}}(\mathbf{y}) = \begin{cases} a \in [0,1] & \text{if } y = 0 \\ 1 & \text{if } 0 < y < 1 \\ b \in [0,1] & \text{if } y = 1. \end{cases}$$

It should be noted that labels $\mathbf{1_{(0,1)}}$ span a surface of area of 1, while singletons have an area of 0. Therefore, it seems intuitively reasonable that the larger the area of the membership degrees in a T2FS, the larger its entropy should be. Taking this into account, and given that, if one curve is below another, it implies that its area is smaller, we consider the following order relation on T2FSs with membership degrees in $\mathcal{L}$.

Definition 4. *Let A and B be type-2 fuzzy sets with membership degrees $\mu_A(x)$ and $\mu_B(x)$, respectively, which are normal and convex functions for each $x \in X$. Then, $A \preccurlyeq_F B$ if $\mu_A(x) \leq \mu_B(x)$, $\forall x \in X$, where $\leq$ is the usual order in the set of functions (that is, $\mu_A(x) \leq \mu_B(x)$ if and only if $(\mu_A(x))(y) \leq (\mu_B(x))(y)$, $\forall y \in [0,1]$).*

This order is similar to the one considered by Burillo in [8] (only adapted to T2FSs). Therefore, the axiomatization we propose in the following could also be considered as an extension of the one proposed in [8].

Definition 5. *Let A and B be type-2 fuzzy sets with membership degrees $\mu_A(x)$ and $\mu_B(x)$, respectively, which are normal and convex functions for each $x \in X$. We say that a function $E_F\colon \mathrm{T2FS(X)} \to [0,1]$ is a Type-F entropy measure on $\mathrm{T2FS}(X)$ if it satisfies the following axioms:*

F.1 $E_F(A) = 0 \Leftrightarrow \mu_A(x) = \bar{a}$ *for some* $a \in [0,1]$, $\forall x \in X$.
F.2 $E_F(A) = 1 \Leftrightarrow \int_{y=0}^{y=1} (\mu_A(x))(y)\, dy = 1$, $\forall x \in X$.
F.3 $E_F(A) = E_F(A^c)$.
F.4 *If* $A \preccurlyeq_F B$, *then* $E_F(A) \leq E_F(B)$.

An example for this entropy measure is as follows.

Proposition 5. *Let $A \in \mathrm{T2FS}(X)$ be any type-2 fuzzy set on X, with a finite universe $|X| = N$, and membership degrees $\mu_A(x)$ being normal and convex functions for each $x \in X$. The following function is a Type-F entropy measure on $\mathrm{T2FS}(X)$:*

$$E_F(A) = \frac{1}{N} \sum_{i=1}^{N} \int_{y=0}^{y=1} (\mu_A(x_i))(y)\, dy. \quad (27)$$

We omit the proof of this proposition since it is almost straightforward and analogous to the proof of Proposition 4 (see Appendix A).

Let us remember how we modified the endpoints of an interval to extend it to the type-2 case in Section 3.1, Equation (22):

$$[L_A(x_i), U_A(x_i)] \to \left[1 - F_A^L(x_i), F_A^R(x_i)\right].$$

It is then obvious that this entropy measure is an adapted version of $I(A)$ proposed by Burillo for IVFSs, Equation (1), since it also considers the width of the interval. It should be noted that $E_F(A)$ also satisfies

$$E_F(A) = \frac{1}{N}\sum_{i=1}^{N}\left(F_A^R(x_i) + F_A^L(x_i) - 1\right). \tag{28}$$

Example 3. *Using, again, the set of Example 2, we obtain*

$$\int_{y=0}^{y=1} \chi_a(y)\,dy = 0, \qquad \int_{y=0}^{y=1} \chi_b(y)\,dy = \frac{3}{8}, \qquad \int_{y=0}^{y=1} \chi_c(y)\,dy = \frac{1}{2},$$

and, thus,

$$E_F(\chi) = \frac{1}{3}\left(0 + \frac{3}{8} + \frac{1}{2}\right) = \frac{1}{3}\cdot\frac{7}{8} = \frac{7}{24}.$$

3.3. Type-IV Entropy in Type-2 Fuzzy Sets

Finally, we define the entropy of a T2FS as a measure of how much it is a non-interval-valued fuzzy set, and we refer to this entropy as Type-IV, where IV stands for interval-valued. Again, we start by studying the entropy at its endpoints. On the one hand, it seems reasonable that the entropy must be 0 if all of the membership degrees are closed subintervals of $[0,1]$. However, to properly define this measure, it is necessary to consider for each membership value (that is, the linguistic label) which is the closest interval-valued fuzzy set. To do this, we set the values that are below $\frac{1}{2}$ to be 0, and the other values that are greater than or equal to $\frac{1}{2}$ to be 1 (see Figure 14).

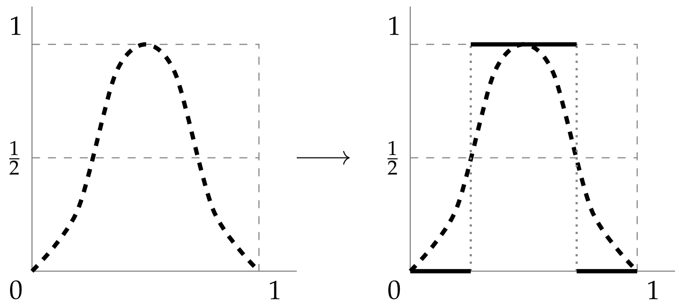

Figure 14. Membership degrees of a type-2 fuzzy set (**left**) and its nearest interval-valued fuzzy set (**right**).

Accordingly, the distance between a membership degree (or label) and its nearest IVFS must be maximum when almost all values are $\frac{1}{2}$ (that is, up to a maximum of three values different from $\frac{1}{2}$ are allowed, since the labels are convex and normal functions). Some of these maximum entropy sets are shown in Figure 15. Let F be the set of labels with almost all values equal to $\frac{1}{2}$. Specifically, labels in F may have only a maximum of three values that are not equal to $\frac{1}{2}$; then,

$$F = \left\{f \in \mathcal{L} : f(y) = \frac{1}{2}, \forall y \in (0,a)\cup(a,1)\right\}. \tag{29}$$

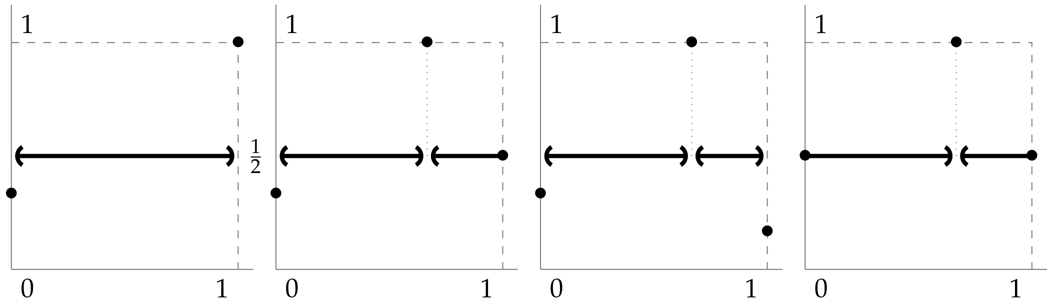

Figure 15. Labels with maximum distance to their nearest interval-valued fuzzy set.

Furthermore, let F' be the set of labels with 0 distance to its nearest interval-valued fuzzy set (see Figure 16):

$$F' = \left\{ f \in \mathcal{L} : f = \overline{(a,b)} \right\}, \tag{30}$$

where

$$\overline{(a,b)}(y) = \begin{cases} 0 & \text{if } y \in [0,a), \\ c \in [0,1] & \text{if } y = a, \\ 1 & \text{if } y \in (a,b), \\ d \in [0,1] & \text{if } y = b, \\ 0 & \text{if } y \in (b,1]. \end{cases}$$

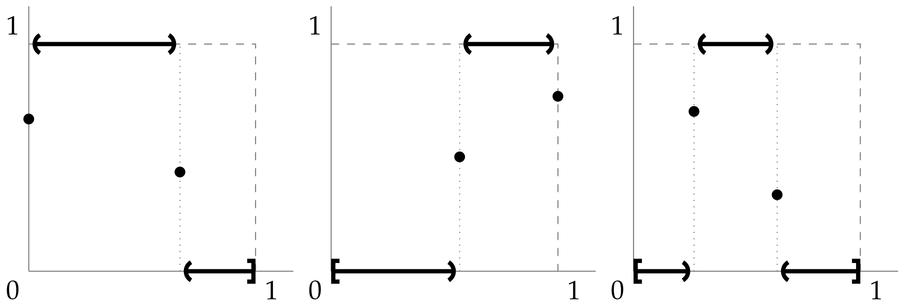

Figure 16. Labels with 0 distance to their nearest interval-valued fuzzy set.

Next, we define the following order relation on T2FSs with membership degrees in $\mathcal{L}$ (see Figure 17).

Definition 6. *Let A and B be type-2 fuzzy sets with membership degrees $\mu_A(x)$ and $\mu_B(x)$, respectively, which are normal and convex functions for each $x \in X$. Then, $A \preccurlyeq_{IV} B$ if it is verified that $\forall x \in X, \forall y \in [0,1]$, the following conditions are met:*

- *If $(\mu_B(x))(y) < \frac{1}{2}$, then $(\mu_A(x))(y) \leq (\mu_B(x))(y)$;*
- *If $(\mu_B(x))(y) = \frac{1}{2}$, then $(\mu_A(x))(y) = (\mu_B(x))(y)$;*
- *If $(\mu_B(x))(y) > \frac{1}{2}$, then $(\mu_A(x))(y) \geq (\mu_B(x))(y)$.*

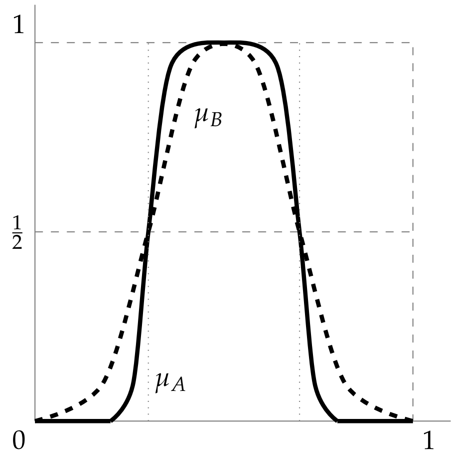

Figure 17. Example of $A \preccurlyeq_{IV} B$ (A solid line and B dashed line).

Based on these considerations, we propose the following axiomatic properties for this entropy measure.

Definition 7. *Let A and B be type-2 fuzzy sets with membership degrees $\mu_A(x)$ and $\mu_B(x)$, respectively, which are normal and convex functions for each $x \in X$. We say that a function $E_I\colon \mathrm{T2FS}(X) \to [0,1]$ is a Type-IV entropy measure on $\mathrm{T2FS}(X)$ if it satisfies the following axioms:*

I.1 $E_I(A) = 0 \Leftrightarrow \mu_A(x) \in F', \forall x \in X$.
I.2 $E_I(A) = 1 \Leftrightarrow \mu_A(x) \in F, \forall x \in X$.
I.3 $E_I(A) = E_I(A^c)$.
I.4 *If $A \preccurlyeq_{IV} B$, then $E_I(A) \leq E_I(B)$.*

An example of this entropy measure is as follows (see Figure 18).

Proposition 6. *Let $A \in \mathrm{T2FS}(X)$ be any type-2 fuzzy set on X, with finite universe $|X| = N$, and membership degrees $\mu_A(x)$ being normal and convex functions for each $x \in X$. The following function is a Type-IV entropy measure on $\mathrm{T2FS}(X)$:*

$$E_I(A) = \frac{2}{N} \sum_{i=1}^{N} \int_{y=0}^{y=1} \min((\mu_A(x))(y), 1 - (\mu_A(x))(y))\, dy. \tag{31}$$

Again, we omit the proof of this proposition, and, instead, we refer to the analogous proof of Proposition 4.

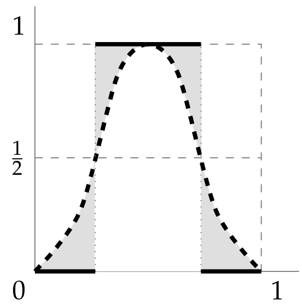

Figure 18. The shaded area represents the distance between a label (or membership degree in a T2FS, dashed line) and its nearest interval fuzzy set.

It should be noted that $2\int_{y=0}^{y=1} \min((\mu_A(x))(y), 1 - (\mu_A(x))(y))\, dy$ in $E_I(A)$ coincides with the expression in Example III. 1 proposed by De Miguel in [14] to obtain the entropy of a membership degree. However, it is a mere coincidence since their meanings are completely different. In [14], the entropy of a T2FS is defined as for a fuzzy set, that is, for each element x of the universe, the authors calculate the entropy of the membership degree of x, that is, the entropy of f_x. Thus, they assign to each element x a measure of how much the function f_x is not a crisp set. On the other hand, in the entropy that we consider here, we propose only one entropy value for the T2FS, and, with this aim, we assign to each element x a measure of how much its label f_x is not an IVFS. It should be noted, for instance, that, while we interpret the image on the right side of Figure 14 as an IVFS, De Miguel would interpret it as a crisp set, and we find the same interpretation in [13,15].

Example 4. *Using, once again, the set of Example 2, we have*

$$\int_{y=0}^{y=1} min(1 - \chi_a(y), \chi_a(y))\, dy = 0,$$
$$\int_{y=0}^{y=1} min(1 - \chi_b(y), \chi_b(y))\, dy = \frac{1}{16},$$
$$\int_{y=0}^{y=1} min(1 - \chi_c(y), \chi_c(y))\, dy = \frac{1}{4}.$$

Thus,

$$E_I(\chi) = \frac{2}{3}\left(0 + \frac{1}{16} + \frac{1}{4}\right) = \frac{5}{24}.$$

3.4. Comparing the Different Measures

In the following two tables, we show eight very simple type-2 fuzzy sets corresponding to labels that are somehow close to the label "true" (named A to H, Table 1) and three type-2 fuzzy sets corresponding to labels that are close to the label "unknown" (named I to K, Table 2). For each one, we compute the five entropy measures presented in Section 3. In all cases, the universe X consists of a single element, that is $X = \{x\}$, and the respective membership degrees represent different labels of the variable truth. The labels in Table 1 range from "absolutely true" (A, the singleton $\bar{1}$) to "true", considering values close to the singleton $\bar{1}$ (B and C), around the singleton $\frac{3}{4}$ (D to G), and, finally, the label "true" (H). In Table 2, we get different potential expressions of the concept "unknown".

In the tables, we can see how the entropy measures of Type C produce higher values as we move further away from "absolutely true" and get closer to "unknown" (exactly the same would happen if considering labels around "false" or closer to "unknown"). There, the maximum value for the Type-C entropy measures is obtained (Table 2). We can interpret this situation as considering an entropy that measures the uncertainty of the given information. We can assume that "absolutely true" and "absolutely false" represent the maximum certainty (value 0), while "unknown" represents the maximum uncertainty (value 1).

At the same time, it is easy to see that the entropy measures of Type F and Type I do not behave in a similar way. The entropy measure can be 0 even for the complete uncertainty represented by the label "true with degree $\frac{1}{2}$". These measures somehow analyze the dispersion of the set that represents the label. This property is usually referred to as imprecision, see [27], representing the area below the membership function. In that sense, a quite precise expression of complete uncertainty, like "true with degree $\frac{1}{2}$", will obtain an entropy measure of 0. Any singleton (no matter whether it represents true, false, or unknown) is considered precise and is assigned the value 0.

Therefore, we have two families of entropy measures that can be described as evaluating uncertainty (Type-C) and imprecision (Type-F and Type-I). Is there any difference inside each of the two families?

Entropy measures of Type C characterize uncertainty, but in the case of E_C^S, imprecision is also taken into account. The higher the imprecision, the higher the value of E_C^S, but always ranging between the value of E_C^V (obtained for minimum imprecision) and E_C^Z

(obtained for maximum imprecision). In fact, all singletons in Table 1 (a singleton has an imprecision of 0, as said before) produce $E_C^V = E_C^S$ (see A, D, and I in Tables 1 and 2). On the other hand, for maximum imprecision (any set A such that $\forall x \in X$, $(\mu_A(x))(0) = 1$ or $(\mu_A(x))(1) = 1$), we get $E_C^S = E_C^Z$ (see A, B, C and H in Table 1).

Now considering entropy measures of Type F and I, while both of them consider imprecision, Type F measures imprecision as defined in the literature (area below the membership function), but Type I measures imprecision considering the area comprised between the membership function and either 0 or 1 (the closer value; see the shaded area in Figure 18).

Table 1. Comparing measures of entropy for labels close to "true".

	$E_C^V(\chi)$	$E_C^S(\chi)$	$E_C^Z(\chi)$	$E_F(\chi)$	$E_I(\chi)$
$A = \{\langle x, \textit{true with degree } 1\rangle\}$ $= \{\langle x, \mu_A(x) = \bar{1}\rangle\}$ (singleton)	0	0	0	0	0
$B = \{\langle x, \textit{true with degree approximately } 1\rangle\}$ $= \{\langle x, \mu_B(x) = [\frac{3}{4}, 1, 1, 1]\rangle\}$ (triangle)	$\frac{1}{15}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$C = \{\langle x, \textit{true with degree greater than } \frac{3}{4}\rangle\}$ $= \{\langle x, \mu_C(x) = \overline{[\frac{3}{4}, 1]}\rangle\}$ (interval)	$\frac{1}{7}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	0
$D = \{\langle x, \textit{true with degree } \frac{3}{4}\rangle\}$ $= \{\langle x, \mu_D(x) = \bar{\frac{3}{4}}\rangle\}$ (singleton)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	0	0
$E = \{\langle x, \textit{true with degree approximately } \frac{3}{4}\rangle\}$ $= \{\langle x, \mu_E(x) = [\frac{5}{8}, \frac{3}{4}, \frac{3}{4}, \frac{7}{8}]\rangle\}$ (triangle)	$\frac{1}{3}$	$\frac{5}{13}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$
$F = \{\langle x, \textit{true with degree } [\frac{5}{8}, \frac{7}{8}]\rangle\}$ $= \{\langle x, \mu_F(x) = \overline{[\frac{5}{8}, \frac{7}{8}]}\rangle\}$ (interval)	$\frac{1}{3}$	$\frac{3}{7}$	$\frac{1}{2}$	$\frac{1}{4}$	0
$G = \{\langle x, \textit{true with degree approx. } [\frac{5}{8}, \frac{7}{8}]\rangle\}$ $= \{\langle x, \mu_G(x) = [\frac{1}{2}, \frac{5}{8}, \frac{7}{8}, 1]\rangle\}$ (trapezoid)	$\frac{1}{3}$	$\frac{7}{15}$	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{1}{8}$
$H = \{\langle x, \textit{true}\rangle\}$ (triangle) $= \{\langle x, (\mu_H(x))(y) = y\rangle\}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Table 2. Comparing measures of entropy for labels close to "unknown".

	$E_C^V(\chi)$	$E_C^S(\chi)$	$E_C^Z(\chi)$	$E_F(\chi)$	$E_I(\chi)$
$I = \{\langle x, \textit{true with degree } \frac{1}{2}\rangle\}$ $= \{\langle x, \mu_I(x) = \bar{\frac{1}{2}}\rangle\}$ (singleton)	1	1	1	0	0
$J = \{\langle x, \textit{not very true and not very false}\rangle\}$ $= \{\langle x, \mu_J(x) = [0, \frac{1}{2}, \frac{1}{2}, 1]\rangle\}$ (triangle)	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$
$K = \{\langle x, \textit{lack of information (unknown)}\rangle\}$ $= \{\langle x, \mu_K(x) = \overline{[0, 1]}\rangle\}$ (interval)	1	1	1	1	0

4. Conclusions

The present paper extended to type-2 fuzzy sets (T2FSs) some entropy measures previously defined in the frameworks of fuzzy, intuitionistic, and interval-valued fuzzy sets (IVFSs). The analysis of those previous ideas offered insights for different approaches to study the entropy in T2FSs. Here, we considered three of those potential approaches, materialized in the novel proposal of three types of entropy measures. The main difference between these approaches relies on the reference set, that is, the set with minimum entropy (the set for which the entropy measure should take value 0). Three references were considered: crisp sets, fuzzy sets, and IVFSs. Thus, the first approach to entropy measures how much a T2FS is not a crisp set (how far from the closest crisp set it is). The second approach measures how far a T2FS is from being a fuzzy set. Finally, the third approach measures how far the T2FS is from being an IVFS (how far is the closest IVFS). These three approaches are described as entropy measures of Type-C, Type-F, and Type-IV, referring to crisp, fuzzy, and interval-valued fuzzy sets, respectively.

Moreover, all three proposals interpret T2FSs according to the structure proposed by Mizumoto and Tanaka [17], more in line with Zadeh's original conception of a T2FS. In that sense, the membership degrees of the T2FS are considered as labels of the variable truth and not simply as two-dimensional type-1 fuzzy sets. Based on the proposed axiomatizations, several examples of entropy measures on T2FSs have been defined, extending to T2FSs the most commonly used entropy measures in fuzzy sets, and IVFSs.

Finally, for better understanding, we have illustrated each idea or key concept with drawings and examples throughout the paper, including an overall comparison through a final example.

Author Contributions: The authors L.M., C.T.-B., S.C. and J.M.-M. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been partially supported by the Government of Spain (grants PID2021-122905NB-C22 and PID2020-112502RB-C41), Comunidad de Madrid (Convenio Plurianual con la Universidad Politécnica de Madrid en la línea de actuación Programa de Excelencia para el Profesorado Universitario).

Data Availability Statement: Data are contained within the article.

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

Appendix A

Proof of Proposition 4. Let us verify that the functions E_C^Z, E_C^Ss and E_C^V satisfy the four Type-C entropy axioms for any $A \in \text{T2FS}(X)$.

Axiom C.1: Let $\mu_A(x_i) = \bar{0}$ or $\mu_A(x_i) = \bar{1}$, $\forall x_i \in X$. If $\mu_A(x_i) = \bar{0}$, then

$$F_A^L(x_i) = \int_{y=0}^{y=1} \bar{0}^L(y)\,dy = 1, \quad F_A^R(x_i) = \int_{y=0}^{y=1} \bar{0}^R(y)\,dy = 0.$$

Otherwise, if $\mu_A(x_i) = \bar{1}$, then

$$F_A^L(x_i) = \int_{y=0}^{y=1} \bar{1}^L(y)\,dy = 0, \quad F_A^R(x_i) = \int_{y=0}^{y=1} \bar{1}^R(y)\,dy = 1.$$

Thus, $E_C^Z(A) = 0$, $E_C^S(A) = 0$ and $E_C^V(A) = 0$.

In contrast, let $E_C^Z(A) = 0$. Then, $\frac{1}{N}\sum_{i=1}^N \left|F_A^R(x_i) - F_A^L(x_i)\right| = 1$. Since $0 \le F_A^R(x_i) \le 1$ and $0 \le F_A^L(x_i) \le 1$, it follows that $0 \le \left|F_A^R(x_i) - F_A^L(x_i)\right| \le 1$, and, hence, it must hold that $\left|F_A^R(x_i) - F_A^L(x_i)\right| = 1$, $\forall x_i \in X$. Then, we have that $F_A^R(x_i) = 1$ and $F_A^L(x_i) = 0$, or $F_A^R(x_i) = 0$ and $F_A^L(x_i) = 1$. Therefore, $\mu_A(x_i) = \bar{1}$ or $\mu_A(x_i) = \bar{0}$, $\forall x_i \in X$.

Let $E_C^S(A) = 0$, that is,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\min(F_A^R(x_i), F_A^L(x_i))}{\max(F_A^R(x_i), F_A^L(x_i))} = 0.$$

Then, $F_A^R(x_i) = 0$ or $F_A^L(x_i) = 0$, $\forall x_i \in X$. Therefore, necessarily, $\mu_A(x_i) = \bar{0}$ or $\mu_A(x_i) = \bar{1}$, $\forall x_i \in X$.

Similarly, let $E_C^V(A) = 0$, that is,

$$\sum_{i=1}^{N}\frac{\min(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \min(F_A^R(x_i), F_A^L(x_i))}{\max(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \max(F_A^R(x_i), F_A^L(x_i))} = 0.$$

Then, $\mu_A(x_i) = \bar{0}$ or $\mu_A(x_i) = \bar{1}$, $\forall x_i \in X$. Thus, Axiom *C.1* follows.

Axiom C.2: Let $F_A^R(x_i) = F_A^L(x_i)$, $\forall x_i \in X$. It is straightforward to see that $E_C^Z(A) = E_C^S(A) = E_C^V(A) = 1$.

Conversely, let $E_C^Z(A) = 1$. Then, $\frac{1}{N}\sum_{i=1}^{N}\left|F_A^R(x_i) - F_A^L(x_i)\right| = 0$. Since $\left|F_A^R(x_i) - F_A^L(x_i)\right| \geq 0$, it follows that $\left|F_A^R(x_i) - F_A^L(x_i)\right| = 0$, $\forall x_i \in X$, and, hence, $F_A^R(x_i) = F_A^L(x_i)$, $\forall x_i \in X$.

Next, let $E_C^S(A) = 1$, and we denote

$$\Delta(x_i) = \frac{\min(F_A^R(x_i), F_A^L(x_i))}{\max(F_A^R(x_i), F_A^L(x_i))}.$$

Since $\Delta(x_i) \leq 1$, it necessarily has to be $\Delta(x_i) = 1$, $\forall x_i \in X$. Therefore, $F_A^R(x_i) = F_A^L(x_i)$, $\forall x_i \in X$.

Finally, let $E_C^V(A) = 1$. Then, $\forall x_i \in X$,

$$\frac{\min(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \min(F_A^R(x_i), F_A^L(x_i))}{\max(1 - F_A^L(x_i), 1 - F_A^R(x_i)) + \max(F_A^R(x_i), F_A^L(x_i))} = 1.$$

Assuming that $\min(F_A^R(x_i), F_A^L(x_i)) = F_A^L(x_i)$,

$$\frac{1 - F_A^R(x_i) + F_A^L(x_i)}{1 - F_A^L(x_i)) + F_A^R(x_i)} = 1,$$

and, hence, $F_A^L(x_i) = F_A^R(x_i)$.

Similarly, assuming that $\min(F_A^R(x_i), F_A^L(x_i)) = F_A^R(x_i)$, then $F_A^L(x_i) = F_A^R(x_i)$.

Thus, Axiom *C.2* follows.

Axiom C.3: First, we prove $F_{A^c}^L(x_i) = F_A^R(x_i)$, $\forall x_i \in X$. That is,

$$
\begin{aligned}
F^L_{A^c}(x_i) &= \int_{y=0}^{y=1} (\mu_{A^c}(x_i))^L(y)\, dy \\
&= \int_{y=0}^{y=1} \sup\{(\mu_{A^c}(x_i))(z) : z \leq y\}\, dy \\
&= \int_{y=0}^{y=1} \sup\{(\mu_{A}(x_i))(1-z) : z \leq y\}\, dy \\
&= \int_{y=0}^{y=1} \sup\{(\mu_{A}(x_i))(1-z) : 1-z \geq 1-y\}\, dy \\
&= \int_{y=0}^{y=1} \sup\{(\mu_{A}(x_i))(z') : z' \geq 1-y\}\, dy \\
&= \int_{y=0}^{y=1} (\mu_A(x_i))^R(1-y)\, dy \\
&= \int_{z=1}^{z=0} (\mu_A(x_i))^R(z)\, (-dz) \\
&= \int_{z=0}^{z=1} (\mu_A(x_i))^R(z)\, dz = F^R_A(x_i).
\end{aligned}
$$

Similarly, we can prove that $F^R_{A^c}(x_i) = F^L_A(x_i)$. Thus, Axiom *C.3* follows.

Axiom C.4: Let $A \sqsubseteq_{LR} B$. Assuming that $F^L_B(x_i) > F^R_B(x_i)$, $\forall x_i \in X$. Then, $\mu_A(x_i) \sqsubseteq \mu_B(x_i)$, and, hence, $F^L_A(x_i) \geq F^L_B(x_i) > F^R_B(x_i) \geq F^R_A(x_i)$, that is,

$$F^L_A(x_i) - F^R_A(x_i) \geq F^L_B(x_i) - F^R_B(x_i) \Rightarrow \left|F^R_A(x_i) - F^L_A(x_i)\right| \geq \left|F^R_B(x_i) - F^R_L(x_i)\right|$$

and also,

$$1 - \left|F^R_A(x_i) - F^L_A(x_i)\right| \leq 1 - \left|F^R_B(x_i) - F^R_L(x_i)\right| \quad \forall x_i \in X.$$

Thus, $E^Z_C(A) \leq E^Z_C(B)$. Moreover,

$$\min(F^R_A(x_i), F^L_A(x_i)) = F^R_A(x_i) \leq F^R_B(x_i) = \min(F^R_B(x_i), F^L_B(x_i)),$$

$$\max(F^R_A(x_i), F^L_A(x_i)) = F^L_A(x_i) \geq F^L_B(x_i) = \max(F^R_B(x_i), F^L_B(x_i)).$$

Thus, $E^S_C(A) \leq E^S_C(B)$. Finally,

$$\min(1 - F^L_A(x_i), 1 - F^R_A(x_i)) = 1 - F^L_A(x_i) \leq 1 - F^L_B(x_i) = \min(1 - F^L_B(x_i), 1 - F^R_B(x_i)),$$

$$\max(1 - F^L_A(x_i), 1 - F^R_A(x_i)) = 1 - F^R_A(x_i) \geq 1 - F^R_B(x_i) = \max(1 - F^L_B(x_i), 1 - F^R_B(x_i)).$$

Thus, $E^V_C(A) \leq E^V_C(B)$.

Similarly, assuming that $F^R_B(x_i) > F^L_B(x_i)$, $\forall x_i \in X$. Then, $\mu_B(x_i) \sqsubseteq \mu_A(x_i)$, that is, $F^R_A(x_i) \geq F^R_B(x_i) > F^L_B(x_i) \geq F^L_A(x_i)$, and the inequality $E^Z_C(A) \leq E^Z_C(B)$ also follows.

Therefore, Axiom *C.4* and the proposition follow. □

References

1. Zadeh, L.A. Probability measures of Fuzzy events. *J. Math. Anal. Appl.* **1968**, *23*, 421–427. https://doi.org/10.1016/0022-247X(68)90078-4.
2. De Luca, A.; Termini, S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Inf. Control* **1972**, *20*, 301–312. https://doi.org/10.1016/S0019-9958(72)90199-4.
3. Kaufmann, A. *Introduction to the Theory of Fuzzy Subsets*; Academic Press: Cambridge, MA, USA, 1975.
4. Yager, R.R. On the Measure of Fuzziness and Negation Part I: Membership in the Unit Interval. *Int. J. Gen. Syst.* **1979**, *5*, 221–229. https://doi.org/10.1080/03081077908547452.
5. Yager, R.R. On the measure of fuzziness and negation. II. Lattices. *Inf. Control* **1980**, *44*, 236–260. https://doi.org/10.1016/S0019-9958(80)90156-4.

6. Kosko, B. Fuzzy entropy and conditioning. *Inf. Sci.* **1986**, *40*, 165–174. https://doi.org/10.1016/0020-0255(86)90006-X.
7. Madrid, N.; Ojeda-Aciego, M. Measures of inclusion and entropy based on the φ-index of inclusion. *Fuzzy Sets Syst.* **2021**, *423*, 29–54. https://doi.org/10.1016/j.fss.2021.01.011.
8. Burillo, P.; Bustince, H. Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets Syst.* **1996**, *78*, 305–316. https://doi.org/10.1016/0165-0114(96)84611-2.
9. Zeng, W.; Li, H. Relationship between similarity measure and entropy of interval valued fuzzy sets. *Fuzzy Sets Syst.* **2006**, *157*, 1477–1484. https://doi.org/10.1016/j.fss.2005.11.020.
10. Wu, D.; Mendel, J.M. Uncertainty measures for interval type-2 fuzzy sets. *Inf. Sci.* **2007**, *177*, 5378–5393. https://doi.org/10.1016/j.ins.2007.07.012.
11. Liu, S.; Cai, R. Uncertainty of Interval Type-2 Fuzzy Sets Based on Fuzzy Belief Entropy. *Entropy* **2021**, *23*, 1265. https://doi.org/10.3390/e23101265.
12. Cornelis, C.; Kerre, E.E. Inclusion Measures in Intuitionistic Fuzzy Set Theory. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*; ECSQARU 2003; Lecture Notes in Computer Science; Nielsen, T.; Zhang, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2003; Volume 2711, pp. 345–356.
13. Hwang, C.M.; Yang, M.S.; Hung, W.L.; Stanley Lee, E. Similarity, inclusion and entropy measures between type-2 fuzzy sets based on the Sugeno integral. *Math. Comput. Model.* **2011**, *53*, 1788–1797. https://doi.org/10.1016/j.mcm.2010.12.057.
14. De Miguel, L.; Santos, H.; Sesma-Sara, M.; Bedregal, B.; Jurio, A.; Bustince, H.; Hagras, H. Type-2 fuzzy entropy sets. *IEEE Trans. Fuzzy Syst.* **2017**, *25*, 993–1005. https://doi.org/10.1109/TFUZZ.2016.2593497.
15. Qin, J.; Xu, T.; Zheng, P. Axiomatic framework of entropy measure for type-2 fuzzy sets with new representation method and its application to product ranking through online reviews. *Appl. Soft. Comput.* **2022**, *130*, 109689. https://doi.org/10.1016/j.asoc.2022.109689.
16. Yao, D.; Liu, X.; Zhang, X.; Wang, C. Type-2 fuzzy cross-entropy and entropy measures and Their Applications. *J. Intell. Fuzzy Syst.* **2016**, *30*, 2169–2180. https://doi.org/10.3233/IFS-151986.
17. Mizumoto, M.; Tanaka, K. Some properties of fuzzy sets of type 2. *Inf. Control* **1976**, *31*, 312–340. https://doi.org/10.1016/S0019-9958(76)80011-3.
18. Mizumoto, M.; Tanaka, K. Fuzzy sets of type-2 under algebraic product and algebraic sum. *Fuzzy Sets Syst.* **1981**, *5*, 277–290. https://doi.org/10.1016/0165-0114(81)90056-7.
19. Zadeh, L. The concept of a linguistic variable and its application to approximate reasoning—I. *Inf. Sci.* **1975**, *8*, 199–249. https://doi.org/10.1016/0020-0255(75)90036-5.
20. Magdalena, L.; Cubillo, S.; Torres-Blanc, C. Comparing different approaches to entropy for Interval Valued Fuzzy Sets. In Proceedings of the IEEE International Conference on Fuzzy Systems, Incheon, Republic of Korea, 13–17 August 2023; pp. 1–8.
21. Cornelis, C.; Deschrijver, G.; Kerre, E.E. Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: Construction, classification, application. *Int. J. Approx. Reason.* **2004**, *35*, 55–95. https://doi.org/10.1016/S0888-613X(03)00072-0.
22. Szmidt, E.; Kacprzyk, J. Entropy for intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **2001**, *118*, 467–477. https://doi.org/10.1016/S0165-0114(98)00402-3.
23. Vlachos, I.K.; Sergiadis, G.D. Subsethood, entropy, and cardinality for interval-valued fuzzy sets—An algebraic derivation. *Fuzzy Sets Syst.* **2007**, *158*, 1384–1396. https://doi.org/10.1016/j.fss.2006.12.018.
24. Bouchet, A.; Sesma-Sara, M.; Ochoa, G.; Bustince, H.; Montes, S.; Díaz, I. Measures of embedding for interval-valued fuzzy sets. *Fuzzy Sets Syst.* **2023**, *467*, 108505. https://doi.org/10.1016/j.fss.2023.03.008.
25. Harding, J.; Walker, C.; Walker, E. Lattices of convex normal functions. *Fuzzy Sets Syst.* **2008**, *159*, 1061–1071. https://doi.org/10.1016/j.fss.2007.09.009.
26. Walker, C.L.; Walker, E.A. The algebra of fuzzy truth values. *Fuzzy Sets Syst.* **2005**, *149*, 309–347. https://doi.org/10.1016/j.fss.2003.12.003.
27. Gonzalez, A.; Pons, O.; Vila, M. Dealing with uncertainty and imprecision by means of fuzzy numbers. *Int. J. Approx. Reason.* **1999**, *21*, 233–256. https://doi.org/10.1016/S0888-613X(99)00024-9.