

# Entropy measures for Atanassov intuitionistic fuzzy sets based on divergence

Ignacio Montes<sup>1</sup>  · Nikhil R. Pal<sup>2</sup> · Susana Montes<sup>1</sup>

Published online: 18 June 2018  
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

In the literature, there are two different approaches to define entropy of Atanassov intuitionistic fuzzy sets (AIFS, for short). The first approach, given by Szmidt and Kacprzyk, measures how far is an AIFS from its closest crisp set, while the second approach, given by Burillo and Bustince, measures how far is an AIFS from its closest fuzzy set. On the other hand, divergence measures are functions that measure how different two AIFSs are. Our work generalizes both types of entropies using local measures of divergence. This results in at least two benefits: depending on the application, one may choose from a wide variety of entropy measures and the local nature provides a natural way of parallel computation of entropy, which is important for large data sets. In this context, we provide the necessary and sufficient conditions for defining entropy measures under both frameworks using divergence measures for AIFS. We show that the usual examples of entropy measures can be obtained as particular cases of our more general framework. Also, we investigate the connection between knowledge measures and divergence measures. Finally, we apply our results in a multi-attribute decision-making problem to obtain the weights of the experts.

**Keywords** Atanassov intuitionistic fuzzy sets · Divergence measures · Entropy

## 1 Introduction

Fuzzy sets were introduced by Zadeh (1965) to model situations where the available information is vague or incomplete. A fuzzy set is characterized by a membership function, which indicates the degree to which an element belongs to the set or satisfies the property described by the fuzzy set. The theory of fuzzy sets has been widely studied both from the theoretical and applied points of view (see Dubois and Prade 2000; Kacprzyk and Pedrycz 2015; Zimmermann 2001, among others).

---

Communicated by C. Kahraman.

---

✉ Ignacio Montes  
imontes@uniovi.es

Nikhil R. Pal  
nrpal159@gmail.com

Susana Montes  
montes@uniovi.es

<sup>1</sup> Department of Statistics and O.R., University of Oviedo, Oviedo, Spain

<sup>2</sup> Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta, India

Over the years, several extensions of fuzzy sets have been proposed: interval valued fuzzy sets, type-2 fuzzy sets Zadeh (1975), hesitant fuzzy sets Torra (2010), and so on. Atanassov (1986) proposed the notion of intuitionistic fuzzy set (AIFS, for short). The idea is quite simple: for any element, an AIFS assigns a membership and a non-membership degree. The former represents the degree to which an element belongs to the set or complies with the property described by the set, while the latter represents the degree to which the element does not belong to the set. The membership and non-membership degrees satisfy a mathematical constraint: their sum cannot exceed one. The difference between one and the sum of both degrees is called hesitation index, which represents the lack of knowledge about whether the element belongs or not to the set. In recent past, research work on the theory of AIFSs has exponentially grown, and it has been successfully applied in decision making (Joshi and Kumar 2017; Szmidt and Kacprzyk 2006; Xu 2007), pattern recognition (Hung and Yang 2004; Liang and Shi 2003) and image segmentation (Melo-Pinto et al. 2013), among others.

For developing useful applications, two important lines of research have attracted the attention of the researchers. One of the approaches involves comparison of AIFSs. In

this context, many different measures of comparison have been suggested in the literature, such as distances or dissimilarities. However, it can be argued that these measures could be inadequate in some contexts. For this reason, we have introduced divergence measures in our previous work (Montes et al. 2011, 2015), and we have shown many interesting mathematical properties as well as their usefulness in many applications (Montes et al. 2012, 2016). The other approach involves the study of entropy measures for AIFSs. In this framework, two different trends can be found: (i) the Szmidt and Kacprzyk (2001) approach, which considers entropy as a measure of fuzziness: it measures how distant is an AIFS to be a crisp set. (ii) The Burrillo and Bustince (1996) approach, which interprets entropy as a measure of intuitionism: it measures how different is an AIFS to be a fuzzy set.

Our aim in this paper is to define entropies, both Szmidt and Kacprzyk (SK) type and Burrillo and Bustince (BB) type, using divergences. For this, after introducing some preliminary notations in Sect. 2, in Sects. 3 and 4 we study how measures of divergence can be used to define entropies under both frameworks. In Sect. 5, we investigate the connection between divergences and knowledge measures (Guo 2016). We provide some concluding remarks in Sect. 7. Some preliminary results of this investigation have been reported in Montes et al. (2018).

## 2 Preliminaries

In this section, we introduce the main notions used throughout the paper. First of all, we introduce fuzzy sets and AIFSs. We also explain a graphical interpretation of AIFSs and define some usual operations between these sets. Then, we recall the definition of divergences for fuzzy sets (Montes et al. 2002) and AIFSs (Montes et al. 2015), emphasizing the property of locality (Montes et al. 2016). We conclude this section explaining in detail the primary objective of this paper.

Throughout this paper, we consider a finite universe  $X$  whose cardinality will be denoted by  $n$ , that is,  $|X| = n$ .

### 2.1 Atanassov intuitionistic fuzzy sets

Fuzzy sets were introduced by Zadeh (1965) as an extension of crisp sets to model vague or linguistic information. While a crisp set  $A$  allows only two possibilities  $x \in A$  or  $x \notin A$ , a fuzzy set  $A$  allows, for any  $x \in X$ , a degree of membership of the element  $x$  to the set. This membership degree, formally defined as a function  $\mu_A : X \rightarrow [0, 1]$ , represents the degree to which an element belongs to  $A$ , or the degree to which it satisfies the property described by the set  $A$ . In this way,

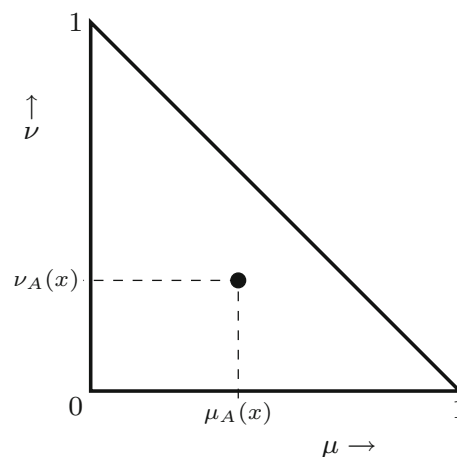


Fig. 1 Graphical representation of AIFSs

$1 - \mu_A(x)$  represents the degree to which  $x$  does not belong to  $A$ .

Atanassov (1986) suggested that the non-membership degree could be different from  $1 - \mu_A(x)$  due to lack of knowledge. To account for this, he proposed an extension of fuzzy sets allowing two degrees: the membership and non-membership degrees, which correspond to the degree to which an element belongs and does not belong to the set, respectively. Formally, an *intuitionistic fuzzy set*, AIFS for short, is defined by  $A = \langle\langle x, \mu_A(x), \nu_A(x) \rangle \mid x \in X \rangle$ , where  $\mu_A(x)$  and  $\nu_A(x)$  denote the membership and non-membership degrees of  $x$  to  $A$ , respectively. An AIFS associates, for every  $x \in X$ , a *hesitation index*, denoted by  $\pi_A$ , and defined by  $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$ . It measures the lack of knowledge about whether  $x$  belongs to or not belongs to  $A$ .

Any fuzzy set  $A$  can be expressed as an AIFS, just by taking  $\nu_A = 1 - \mu_A$ . In particular, for fuzzy sets it holds that  $\pi_A = 0$ . Also, a crisp set is a particular case of an AIFS where for any  $x \in X$ , either  $\mu_A(x) = 1$  and  $\nu_A(x) = 0$ , if  $x \in A$ , or  $\mu_A(x) = 0$  and  $\nu_A(x) = 1$ , if  $x \notin A$ . From now on, we denote by  $\text{AIFS}(X)$  the set of all AIFSs on  $X$ , and by  $\text{FS}(X)$  the set of all fuzzy sets on  $X$ .

Every element  $x \in X$  of any AIFS can be graphically depicted by a pair  $(\mu_A(x), \nu_A(x))$ , as Fig. 1 shows. The segment that goes from  $(1, 0)$  to  $(0, 1)$  corresponds to the pairs  $(\mu_A(x), \nu_A(x))$  where  $\nu_A(x) = 1 - \mu_A(x)$ . If for every  $x \in X$ , the element  $(\mu_A(x), \nu_A(x))$  belongs to this segment,  $A$  is a fuzzy set. Moreover, the further  $(\mu_A(x), \nu_A(x))$  is from the segment  $((1, 0), (0, 1))$ , the greater is the hesitation index  $\pi_A(x)$ . Also, the case of total ignorance, that is when  $\pi_A(x) = 1$ , corresponds to the pair  $(0, 0)$ .

To conclude this subsection, let us recall some basic operations between AIFSs. Given  $A, B \in \text{AIFS}(X)$ , we consider the following operations:

- The *union* of  $A$  and  $B$ , denoted by  $A \cup B$ , is an AIFS whose membership and non-membership degrees are given by:

$$\begin{aligned} \mu_{A \cup B}(x) &= \max\{\mu_A(x), \mu_B(x)\}, \\ \nu_{A \cup B}(x) &= \min\{\mu_A(x), \mu_B(x)\}. \end{aligned}$$

- The *intersection* of  $A$  and  $B$ , denoted by  $A \cap B$ , is an AIFS whose membership and non-membership degrees are given by:

$$\begin{aligned} \mu_{A \cap B}(x) &= \min\{\mu_A(x), \mu_B(x)\}, \\ \nu_{A \cap B}(x) &= \max\{\mu_A(x), \mu_B(x)\}. \end{aligned}$$

- $A$  is *included* in  $B$ , denoted by  $A \subseteq B$ , if  $\mu_A \leq \mu_B$  and  $\nu_A \geq \nu_B$ .
- The *complement* of  $A$ , denoted by  $A^c$ , is defined by:

$$A^c = \langle (x, \nu_A(x), \mu_A(x)) \mid x \in X \rangle.$$

We note here that more general families of unions and intersections can be defined using a T-norm for min and T-conorm (S-norm) for max.

### 2.2 Divergences for AIFSs

One very popular topic of research within AIFSs theory is that of measuring how different two AIFSs are. To compare this kind of sets, although there are many different approaches, for instance similarities or distances, in Montes et al. (2015) we have introduced a new family of measures called *AIF-divergences*. We have also argued that, from our point of view, AIF-divergences are more appropriate than other measures of comparison existing in the literature.

**Definition 1** (Montes et al. 2015) A function  $D$  defined from  $\text{AIFS}(X) \times \text{AIFS}(X)$  to  $\mathbb{R}^+$  is an AIF-divergence if it satisfies the following properties:

- (Div.1)  $D(A, B) = D(B, A)$  for any  $A, B \in \text{AIFS}(X)$ .
- (Div.2)  $D(A, A) = 0$  for any  $A \in \text{AIFS}(X)$ .
- (Div.3)  $D(A \cap C, B \cap C) \leq D(A, B)$  for any  $A, B, C \in \text{AIFS}(X)$ .
- (Div.4)  $D(A \cup C, B \cup C) \leq D(A, B)$  for any  $A, B, C \in \text{AIFS}(X)$ .

Hence, an AIF-divergence is symmetric and it takes the value 0 when comparing an AIFS with itself. And the closer two AIFSs are, the smaller is the AIF-divergence between them.

In (Montes et al. 2015, Lemma 3.2), we proved that any AIF-divergence satisfies the following property:

$$A \subseteq B \subseteq C \Rightarrow D(A, C) \geq \max\{D(A, B), D(B, C)\}. \quad (1)$$

This property will be used later. We next consider one particular family of AIF-divergences that satisfies the local property:

$$\begin{aligned} D(A \cup \{x\}, B \cup \{x\}) - D(A, B) \\ = h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_B(x), \nu_B(x)); \end{aligned}$$

where  $h_{\text{IF}}$  satisfies five locality properties in Theorem 1. This means that if we modify the membership and non-membership degrees of only one element of the sets, the change in the AIF-divergence between the sets depends only on what has been changed.

In Montes et al. (2016), we characterized local AIF-divergences, using the following notation<sup>1</sup>:

$$\mathcal{D} = \{(u_1, u_2, v_1, v_2) \in \mathbb{R}^{4+} \mid u_1 + u_2 \leq 1, v_1 + v_2 \leq 1\}. \quad (2)$$

**Theorem 1** Montes et al. (2015) An AIF-divergence  $D$  is local if and only if there exists a function  $h : \mathcal{D} \rightarrow \mathbb{R}^+$  such that

$$D(A, B) = \sum_{x \in X} h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_B(x), \nu_B(x))$$

and  $h$  satisfies the following properties:

- (AIF-loc.1)  $h_{\text{IF}}(u, v, u, v) = 0$  for any  $(u, v, u, v) \in \mathcal{D}$ .
- (AIF-loc.2)  $h_{\text{IF}}(u_1, u_2, v_1, v_2) = h_{\text{IF}}(v_1, v_2, u_1, u_2)$  for any  $(u_1, u_2, v_1, v_2) \in \mathcal{D}$ .
- (AIF-loc.3) If  $(u_1, u_2, v_1, v_2) \in \mathcal{D}$ ,  $\omega \in [0, 1]$  and  $u_1 \leq \omega \leq v_1$ , it holds that

$$h_{\text{IF}}(u_1, u_2, v_1, v_2) \geq h_{\text{IF}}(u_1, u_2, \omega, v_2).$$

Moreover, if  $\max\{u_2, v_2\} + \omega \leq 1$ , it holds that

$$h_{\text{IF}}(u_1, u_2, v_1, v_2) \geq h_{\text{IF}}(\omega, u_2, v_1, v_2).$$

- (AIF-loc.4) If  $(u_1, u_2, v_1, v_2) \in \mathcal{D}$ ,  $\omega \in [0, 1]$  and  $u_2 \leq \omega \leq v_2$ , it holds that

$$h_{\text{IF}}(u_1, u_2, v_1, v_2) \geq h_{\text{IF}}(u_1, u_2, v_1, \omega).$$

Moreover, if  $\max\{u_1, v_1\} + \omega \leq 1$ , it holds that

$$h_{\text{IF}}(u_1, u_2, v_1, v_2) \geq h_{\text{IF}}(u_1, \omega, v_1, v_2).$$

<sup>1</sup> The notation we are using here is slightly different from that in Montes et al. (2015), where the set  $\mathcal{D}$  was defined by  $\mathcal{D} = \{(u, v) \in \mathbb{R}^{2+} \mid u+v \leq 1\}$ , and then we considered  $\mathcal{D}^2$ . In this paper, we have considered an alternative expression for  $\mathcal{D}$  in Eq. (2) for the sake of mathematical convenience. However, both approaches are equivalent.

(AIF-loc.5) If  $(u_1, u_2, v_1, v_2) \in \mathcal{D}$  and  $\omega \in [0, 1]$ , then if  $\max\{u_2, v_2\} + \omega \leq 1$  it holds that:

$$h_{\text{IF}}(\omega, u_2, \omega, v_2) \leq h_{\text{IF}}(u_1, u_2, v_1, v_2);$$

and if  $\max\{u_1, v_1\} + \omega \leq 1$ , it holds that:

$$h_{\text{IF}}(u_1, \omega, v_1, \omega) \leq h_{\text{IF}}(u_1, u_2, v_1, v_2).$$

Let us prove a useful property of the function  $h_{\text{IF}}$  associated with a local AIF-divergence.

**Proposition 1** Let  $D$  be a local AIF-divergence with associated function  $h_{\text{IF}}$ . Then,  $h_{\text{IF}}(u, v, 1, 0)$  is decreasing on  $u$  and increasing on  $v$ , and  $h_{\text{IF}}(u, v, 0, 1)$  is increasing on  $u$  and decreasing on  $v$ , whenever  $u + v \leq 1$ .

**Proof** Let us prove that  $h_{\text{IF}}(u, v, 1, 0)$  is decreasing in  $u$ . For this, take  $u_1 \leq u_2$  such that  $u_2 + v \leq 1$ . Taking property (AIF-loc.3) into account, it holds that:

$$h_{\text{IF}}(u_1, v, 1, 0) \geq h_{\text{IF}}(u_2, v, 1, 0).$$

On the other hand, let us see that  $h_{\text{IF}}(u, v, 1, 0)$  is increasing in  $v$ . For this, take  $v_1 \leq v_2$ . Define the AIFSs  $A, B, M$  on  $\{x\}$  by:

$$A = \langle (x, u, v_1) \rangle, \quad B = \langle (x, u, v_2) \rangle, \quad M = \langle (x, 1, 0) \rangle.$$

From property (Div.4),  $D(A, M) = D(A \cup B, A \cup M) \leq D(B, M)$ , which means that:

$$h_{\text{IF}}(u, v_2, 1, 0) \geq h_{\text{IF}}(u, v_1, 1, 0).$$

Let us now study the function  $h_{\text{IF}}(u, v, 0, 1)$ . First of all, we show that that it is increasing in the first component. Take  $u_1 \leq u_2$ , and define the AIFSs  $A, B, N$  on  $\{x\}$  by:

$$A = \langle (x, u_2, v) \rangle, \quad B = \langle (x, u_1, v) \rangle, \quad N = \langle (x, 0, 1) \rangle.$$

Using property (Div.3), we obtain that

$$D(B, N) = D(A \cap B, N \cap B) \leq D(A, N),$$

which means that:

$$h_{\text{IF}}(u_2, v, 0, 1) \geq h_{\text{IF}}(u_1, v, 0, 1).$$

Next we show that  $h_{\text{IF}}(u, v, 0, 1)$  is decreasing in  $v$ . Taking  $v_1 \leq v_2$  and using (AIF-loc.4) it holds that:

$$h_{\text{IF}}(u, v_1, 0, 1) \geq h_{\text{IF}}(u, v_2, 0, 1).$$

Divergences for fuzzy sets have already been introduced in Montes et al. (2002). A function  $D : \text{FS}(X) \times \text{FS}(X) \rightarrow \mathbb{R}$  is a *divergence for fuzzy sets* if it satisfies conditions (Div.1) to (Div.4) when we restrict them to  $\text{FS}(X)$ . The property of locality has also been defined for divergences between fuzzy sets, and it was characterized in the following way:

**Theorem 2** (Montes et al. 2002, Prop. 3.4) A function  $D : \text{FS}(X) \times \text{FS}(X) \rightarrow \mathbb{R}^+$  is a local divergence for fuzzy sets if and only if there exists a function  $h_{\text{FS}} : [0, 1] \times [0, 1] \rightarrow \mathbb{R}^+$  such that:

$$D(A, B) = \sum_{x \in X} h_{\text{FS}}(\mu_A(x), \mu_B(x))$$

and it satisfies the following properties:

(loc.1)  $h_{\text{FS}}(u, v) = h_{\text{FS}}(v, u)$  for any  $(u, v) \in [0, 1]^2$ .

(loc.2)  $h_{\text{FS}}(u, u) = 0$  for any  $u \in [0, 1]$ .

(loc.3)  $h_{\text{FS}}(u, z) \geq \max\{h_{\text{FS}}(u, v), h_{\text{FS}}(v, z)\}$  for any  $u, v, z \in [0, 1]$  such that  $u < v < z$ .

### 2.3 Divergence-based entropies of AIFS

For fuzzy sets, the notion of entropy or fuzziness was introduced by Luca and Termini (1972). Since then, many researchers continued working on this topic, such as (Bhandari and Pal 1993; Kosko 1986; Liu 1992; Pal and Bejdek 1994; Trillas and Riera 1978; Yager 1982), among others. In particular, Montes et al. (1998) used local divergences for fuzzy sets as a measure of entropy or fuzziness.

Our objective here is to define entropies for AIFSs by using local AIF-divergences. As we shall explain later, there are two different types of entropies: the one defined by Szmidt and Kacprzyk (2001), that measures how different is an AIFS from its nearest crisp set, and the one defined by Burrillo and Bustince (1996) that measures how different is an AIFS from its closest fuzzy set. As Pal et al. (2013) have explained, the two types of entropies are different and can be interpreted as complementary.

From now on, we consider a local AIF-divergence  $D$  with associated function  $h_{\text{IF}}$ , and we investigate the additional properties that must be imposed on  $h_{\text{IF}}$  to define entropies, with respect to both interpretations, the interpretation of Szmidt and Kacprzyk (SK) and that of Burrillo and Bustince (BB).

For this we make two assumptions: (i) the local AIF-divergence can be expressed by:

$$D(A, B) = \frac{1}{n} \sum_{x \in X} h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_B(x), \nu_B(x))$$

for any  $A, B \subseteq \text{AIFS}(X)$ ; and (ii)  $h_{\text{IF}}$  takes values in  $[0, 1]$ , meaning that for any pair of AIFSs the maximal difference

□

is 1. Here  $D$  is upper bounded and the previous assumptions can be simply understood as a rescaling of the divergence. Therefore, these assumptions are only made for mathematical convenience.

### 3 SK-Entropies

Here we deal with entropies that measure how far is an AIFS to be a crisp set, which will be called SK-entropies. Following the definition of entropy given by Szmidt and Kacprzyk (2001), we introduce the notion of closest crisp set to an AIFS and then use local AIF-divergences to define SK-entropies.

#### 3.1 Szmidt and Kacprzyk's entropy

We first introduce the definition of entropy for AIFSs given by Szmidt and Kacprzyk (2001): it measures how far is an AIFS from its closest crisp set.

**Definition 2** (Szmidt and Kacprzyk (2001)) A function  $E : \text{AIFS}(X) \rightarrow [0, 1]$  is an *entropy* if it satisfies the following axioms:

- ( $I_{SK1}$ )  $E(A) = 0$  if and only if  $A$  is a crisp set.
- ( $I_{SK2}$ )  $E(A) = 1$  if and only if  $\mu_A(x) = \nu_A(x)$  for every  $x \in X$ .
- ( $I_{SK3}$ )  $E(A) = E(A^c)$ .
- ( $I_{SK4}$ )  $E(A) \leq E(B)$  if  $\mu_A(x) \leq \mu_B(x) < \nu_B(x) \leq \nu_A(x)$  or  $\nu_A(x) \leq \nu_B(x) \leq \mu_B(x) \leq \mu_A(x)$  for every  $x \in X$ .

Let us discuss the previous conditions. The condition ( $I_{SK1}$ ) implies that the entropy is zero (nonexistent) if, and only if, the set is crisp. ( $I_{SK3}$ ) says that the entropy is closed with respect to the complement. According to Property ( $I_{SK2}$ ), the entropy takes the maximum value if, and only if, the membership and non-membership degrees coincide. However, in some applications this property can be argued to be rather soft: given two AIFSs  $A$  and  $B$  satisfying  $\mu_A = \nu_A$  and  $\mu_B = \nu_B$ , both sets have the same entropy, regardless of the exact values of the membership degrees  $\mu_A$  and  $\mu_B$  (or  $\pi_A$  and  $\pi_B$ ). Hence, property ( $I_{SK2}$ ) does not take into account the hesitation associated with the AIFSs. For example, consider the AIFSs  $A$  and  $B$  defined by:

$$\begin{aligned} A &= \langle (x, 0.1, 0.1) \mid \forall x \in X \rangle, \\ B &= \langle (x, 0.45, 0.45) \mid \forall x \in X \rangle. \end{aligned} \tag{3}$$

From ( $I_{SK2}$ ),  $E$  satisfies  $E(A) = E(B)$ , but the lack of information associated with  $A$  seems to be greater than that with  $B$ , because  $\pi_A(x) = 0.8 > 0.1 = \pi_B(x)$  for any  $x \in X$ . These AIFSs are graphically depicted in Fig. 2.

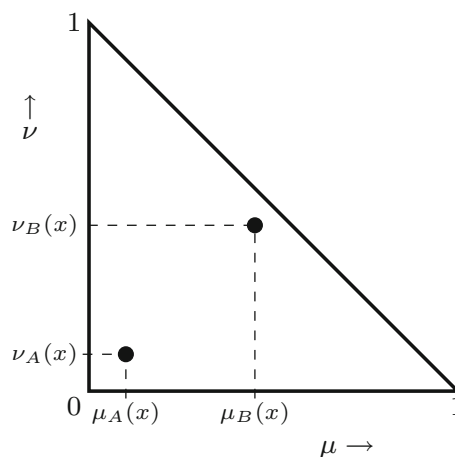


Fig. 2 Graphical representation of the AIFSs in Eq. (3)

Property ( $I_{SK2}$ ) can be slightly modified in order to overcome this drawback as follows:

$$(I_{SK2'}) \ E(A) = 1 \text{ if and only if } \mu_A(x) = \nu_A(x) = 0 \text{ for every } x \in X.$$

This is more plausible as it implies that  $E(A) = 1$  when we have no knowledge about membership and non-membership for every element. This modified property can be equivalently expressed in terms of the hesitation index, since  $E(A) = 1$  if and only if  $\pi_A = 1$ . For the AIFSs defined in Eq. (3), if we consider the property ( $I_{SK2'}$ ), an entropy must satisfy  $E(A) \geq E(B)$ , but the equality is not required.

Finally, condition ( $I_{SK4}$ ) says that the closer is the set to a crisp set, the lower is its entropy.

Taking into account Definition 2 as well as the previous discussion, we consider the following definition of an SK-entropy.

**Definition 3** A mapping  $E : \text{AIFS}(X) \rightarrow [0, 1]$  is a *SK-entropy* if it satisfies the properties ( $I_{SK1}$ ), ( $I_{SK2'}$ ), ( $I_{SK3}$ ) and ( $I_{SK4}$ ).

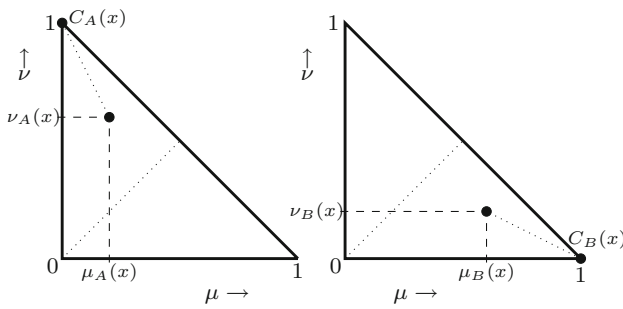
#### 3.2 SK-entropies based on local AIF-divergences

Throughout this subsection, we aim to investigate how SK-entropies can be built using local AIF-divergences. For this, we first introduce the notion of the closest crisp set to an AIFS.

**Definition 4** Given  $A \in \text{AIFS}(X)$ , we define the *closest crisp set to A*, denoted by  $C_A$ , by:

$$x \in C_A \text{ if } \mu_A(x) \geq \nu_A(x), \text{ and } x \notin C_A \text{ otherwise.}$$

This notion had already been considered for fuzzy sets in Montes et al. (1998).



**Fig. 3** Example of the closest crisp set to an AIFS

Since any crisp set is an AIFS with zero hesitation index, we can express the closet crisp set to  $A$  as:

$$\mu_{C_A}(x) = \begin{cases} 1 & \text{if } \mu_A(x) \geq \nu_A(x), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\nu_{C_A}(x) = \begin{cases} 0 & \text{if } \mu_A(x) \geq \nu_A(x), \\ 1 & \text{otherwise.} \end{cases}$$

In Fig. 3, we depict an example of the closest crisp set to an AIFS. In the left panel of the picture, we show an example where  $\mu_A(x) < \nu_A(x)$ , so  $x \notin C_A$ , or equivalently,  $\mu_{C_A}(x) = 0, \nu_{C_A}(x) = 1$ . The opposite happens in the right panel of the picture, where  $\mu_B(x) > \nu_B(x)$ , so  $x \in C_B$ , or equivalently,  $\mu_{C_B}(x) = 1, \nu_{C_B}(x) = 0$ . As we can see in the picture, as long as  $(\mu_A(x), \nu_A(x))$  is above the dotted line which represents the pairs  $(t, t)$ ,  $x \notin C_A$ , while as long as  $(\mu_B(x), \nu_B(x))$  is on or below the dotted line,  $x \in C_B$ .

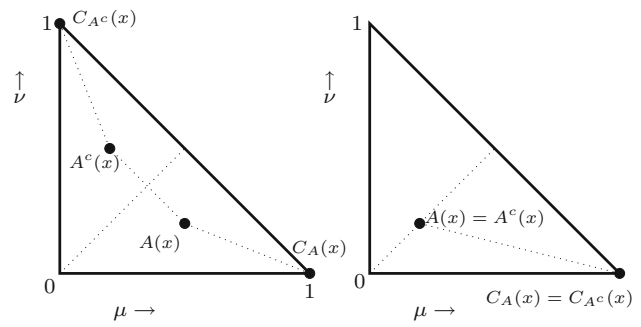
Next proposition shows two simple but useful properties of the closest crisp set to an AIFS.

**Proposition 2** Consider  $A \in \text{AIFS}(X)$ , and let  $C_A$  be its closest crisp set. The following statements hold:

1.  $A$  is a crisp set if and only if  $A = C_A$ .
2. For any  $x \in X$ , if  $\mu_A(x) \neq \nu_A(x)$ , then  $C_A^c(x) = C_{A^c}(x)$ , where  $C_{A^c}$  denotes the closets crisp set to  $A^c$ .

**Proof** Let us prove the first item. Obviously, if  $A = C_A$ ,  $A$  is a crisp set. On the other hand, if  $A$  is a crisp set, for any  $x \in X$  either  $x \in A$ , which implies  $\mu_A(x) = 1$ , or  $x \notin A$ , which implies  $\nu_A(x) = 1$ . In the former case,  $\mu_A(x) > \nu_A(x)$ , and therefore  $x \in C_A$ , while in the second case  $\mu_A(x) < \nu_A(x)$ , which implies  $x \notin C_A$ . We conclude that  $A = C_A$ .

Let us now turn to the second item. Since  $\mu_A(x) \neq \nu_A(x)$ , we only have two possibilities, either  $\mu_A(x) > \nu_A(x)$  or  $\mu_A(x) < \nu_A(x)$ . Assume we are in the former case. By definition of  $C_A$ , it holds that  $\mu_{C_A}(x) = 1, \nu_{C_A}(x) = 0$ , which



**Fig. 4** Graphical representation of the second item in Proposition 2. In this figure,  $A(x)$  and  $A^c(x)$  represent the pairs  $(\mu_A(x), \nu_A(x))$  and  $(\nu_A(x), \mu_A(x))$ , respectively

implies that  $x \in C_A$ , and consequently  $x \notin C_A^c$ . Also, since  $\mu_A(x) > \nu_A(x)$ , it holds that:

$$\mu_{A^c}(x) = \nu_A(x) < \mu_A(x) = \nu_{A^c}(x),$$

so  $x \notin C_{A^c}$ .

On the other hand, if  $\nu_A(x) > \mu_A(x)$ , following a similar reasoning we obtain that  $x \in C_A^c$  and  $x \in C_{A^c}$ , so we conclude that  $C_A^c(x) = C_{A^c}(x)$ .  $\square$

From the second item, we deduce that if  $\mu_A(x) \neq \nu_A(x)$  then  $\mu_{C_{A^c}}(x) = \nu_{C_A}(x)$  and  $\nu_{C_{A^c}}(x) = \mu_{C_A}(x)$ .

The second item of Proposition 2 is graphically explained in Fig. 4. As the left-side panel shows, when  $\mu_A(x) > \nu_A(x)$ ,  $C_A(x)$  and  $C_{A^c}(x)$  are just at the opposite corner, so  $C_A^c(x) = C_{A^c}(x)$ . On the other hand, note that the second item requires  $\mu_A(x) \neq \nu_A(x)$ . The reason is that if  $\mu_A(x) = \nu_A(x)$  holds, as in the right-side panel, the closest crisp set to  $A$  and  $A^c$  for  $x$  coincide and is equal to  $\mu_{C_A}(x) = 1, \nu_{C_A}(x) = 0$ . That is why condition  $\mu_A(x) \neq \nu_A(x)$  is required in the second item of the Proposition 2.

So far we have investigated the properties of the closest crisp set to an AIFS. Now, we use this notion to define an SK-entropy in terms of local AIF-divergences. Recall that the aim of an SK-entropy is to measure how different is an AIFS from a crisp set. Therefore, it seems reasonable to measure the entropy of an AIFS as the AIF-divergence between the AIFS and its closest crisp set. Note that, when comparing  $A \in \text{AIFS}(X)$  with  $C_A$  by means of a local AIF-divergence  $D$  induced by the function  $h_{\text{IF}}$ , the domain of  $h_{\text{IF}}$  is no longer  $\mathcal{D}$ , but

$$\begin{aligned} \mathcal{D}_1 = & \{(x, y, 1, 0) \in \mathcal{D} \mid x \geq y\} \\ & \cup \{(x, y, 0, 1) \in \mathcal{D} \mid x < y\}. \end{aligned} \tag{4}$$

For this, only the conditions imposed on  $h_{\text{IF}}$  in the next theorem need to be satisfied in the domain  $\mathcal{D}_1$ .

**Theorem 3** Consider a local AIF-divergence  $D$  induced by a function  $h_{IF}$ , and define  $E : AIFS(X) \rightarrow [0, 1]$  by:

$$E(A) = D(A, C_A) = \frac{1}{n} \sum_{x \in X} h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x))$$

for any  $A \subseteq X$ . Then,  $E$  is an SK-entropy if and only if the function  $h_{IF}$  satisfies the following additional properties on  $\mathcal{D}_1$ :

- (AIF-loc.1')  $h_{IF}(u_1, u_2, v_1, v_2) = 0$  for  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_1$  if and only if  $u_1 = v_1, u_2 = v_2$ .
- (AIF-loc.5)  $h_{IF}(u_1, u_2, v_1, v_2) = h_{IF}(u_2, u_1, v_2, v_1)$  for any  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_1$  such that  $u_1 \neq u_2$ .
- (AIF-loc.6)  $h_{IF}(u_1, u_2, v_1, v_2) = 1$  for  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_1$  if and only if  $u_1 = u_2 = 0$  and  $v_1 = 1, v_2 = 0$ .

**Proof** Let us first prove that if  $h_{IF}$  satisfies these additional properties, then  $E$  is an SK-entropy.

(ISK1) :  $E(A) = 0$  if and only if for any  $x \in X$ :

$$h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) = 0.$$

According to (AIF-loc.1'), this is equivalent to  $\mu_A(x) = \mu_{C_A}(x)$  and  $\nu_A(x) = \nu_{C_A}(x)$ , which from Proposition 2 happens if and only if  $A$  is a crisp set.

(ISK2') :  $E(A) = 1$  if and only if

$$h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) = 1$$

for any  $x \in X$ . From (AIF-loc.6), this happens if and only if  $(u_1, u_2, v_1, v_2) = (0, 0, 1, 0)$ , which is equivalent to  $\mu_A(x) = 0$  and  $\nu_A(x) = 0$ , that is, if and only if  $\pi_A(x) = 1$  for any  $x \in X$ .

(ISK3) : In order to check that  $E(A) = E(A^c)$ , it is enough to check whether the following equality holds for any  $x \in X$ :

$$h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) = h_{IF}(\mu_{A^c}(x), \nu_{A^c}(x), \mu_{C_{A^c}}(x), \nu_{C_{A^c}}(x)).$$

On the one hand, if  $\mu_A(x) \neq \nu_A(x)$ , from Proposition 2,  $C_A^c(x) = C_{A^c}(x)$ , and therefore:

$$\begin{aligned} h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) &= h_{IF}(\nu_A(x), \mu_A(x), \nu_{C_A}(x), \mu_{C_A}(x)) \\ &= h_{IF}(\mu_{A^c}(x), \nu_{A^c}(x), \mu_{C_{A^c}}(x), \nu_{C_{A^c}}(x)), \end{aligned}$$

where the first equality follows from property (AIF-loc.5). On the other hand, if  $\mu_A(x) = \nu_A(x)$ , it trivially holds that

$$h_{IF}(\mu_A(x), \nu_A(x), 1, 0) = h_{IF}(\nu_A(x), \mu_A(x), 1, 0).$$

(ISK4) : Assume that  $\mu_A(x) \leq \mu_B(x) < \nu_B(x) \leq \nu_A(x)$ , which implies that  $\mu_{C_A}(x) = \mu_{C_B}(x) = 0$  and  $\nu_{C_A}(x) = \nu_{C_B}(x) = 1$ . Define the following AIFSs on  $\{x\}$  by:

$$A^* = \langle (x, \mu_A(x), \nu_A(x)) \rangle, \quad B^* = \langle (x, \mu_B(x), \nu_B(x)) \rangle, \\ N = \langle (x, 0, 1) \rangle.$$

It holds that  $N \subseteq A^* \subseteq B^*$ , and therefore from Eq. (1),  $D(A^*, N) \leq D(B^*, N)$ , which implies that

$$\begin{aligned} h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) &= h_{IF}(\mu_A(x), \nu_A(x), 1, 0) \\ &\leq h_{IF}(\mu_B(x), \nu_B(x), 1, 0) \\ &= h_{IF}(\mu_B(x), \nu_B(x), \mu_{C_B}(x), \nu_{C_B}(x)). \end{aligned}$$

On the other hand, assume  $\mu_B(x) \geq \nu_B(x)$  and  $\nu_A(x) \leq \nu_B(x) \leq \mu_B(x) \leq \mu_A(x)$ , which implies that  $\mu_{C_A}(x) = \mu_{C_B}(x) = 1$  and  $\nu_{C_A}(x) = \nu_{C_B}(x) = 0$ .

Now define the following AIFSs on  $\{x\}$ :

$$A^* = \langle (x, \mu_A(x), \nu_A(x)) \rangle, \quad B^* = \langle (x, \mu_B(x), \nu_B(x)) \rangle, \\ M = \langle (x, 1, 0) \rangle.$$

It holds that  $B^* \subseteq A^* \subseteq M$ , which implies, from Eq. (1), that  $D(B^*, M) \geq D(A^*, M)$ , and therefore

$$\begin{aligned} h_{IF}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) &= h_{IF}(\mu_A(x), \nu_A(x), 0, 1) \\ &\leq h_{IF}(\mu_B(x), \nu_B(x), 0, 1) \\ &= h_{IF}(\mu_B(x), \nu_B(x), \mu_{C_B}(x), \nu_{C_B}(x)). \end{aligned}$$

We conclude that if  $h_{IF}$  satisfies the additional conditions,  $E$  is an SK-entropy.

Now we assume that  $E$  is an SK-entropy. We need to prove that  $h_{IF}$  satisfies the additional conditions.

(AIF-loc.1') : Take  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle (x, u_1, u_2) \rangle$ . Note that:

$$E(A) = h_{IF}(u_1, u_2, v_1, v_2).$$

Hence, from (ISK1), it holds that

$$E(A) = h_{IF}(u_1, u_2, v_1, v_2) = 0$$

if and only if  $A$  is a crisp set, which from Proposition 2 is equivalent to  $A = C_A$ . For this equality, it must hold that either  $u_1 = 1$  and  $u_2 = 0$ , which in turn implies that  $v_1 = 1$  and  $v_2 = 0$ , or  $u_1 = 0$  and  $u_2 = 1$ , which implies that  $v_1 = 0$  and  $v_2 = 1$ . In both cases, we conclude that  $h_{IF}(u_1, u_2, v_1, v_2) = 0$  if and only if  $u_1 = v_1$  and  $u_2 = v_2$ .

(AIF – loc.5) : Take  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle (x, u_1, u_2) \rangle$ . It holds that:

$$E(A) = h_{\text{IF}}(u_1, u_2, v_1, v_2).$$

Let us assume that  $u_1 \neq u_2$ . This means that  $A^c = \langle (x, u_2, u_1) \rangle$  and, from the second item in Proposition 2,  $C_A^c = C_{A^c}$ . Hence:

$$E(A^c) = h_{\text{IF}}(u_2, u_1, v_2, v_1).$$

Finally, from (ISK3),  $E(A) = E(A^c)$ , and we deduce that  $h_{\text{IF}}(u_1, u_2, v_1, v_2) = h_{\text{IF}}(u_2, u_1, v_2, v_1)$ .

(AIF – loc.6) : Take  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle (x, u_1, u_2) \rangle$ . Then, it holds that:

$$E(A) = h_{\text{IF}}(u_1, u_2, v_1, v_2).$$

Therefore,  $h_{\text{IF}}(u_1, u_2, v_1, v_2) = 1$  if and only if  $E(A) = 1$ , which from (ISK2'), is equivalent to  $u_1 = u_2 = 0$  and  $v_1 = 1, v_2 = 0$ .

We conclude that if  $E$  is an SK-entropy,  $h_{\text{IF}}$  must satisfy the additional conditions.  $\square$

In Montes et al. (2016), we have shown how an AIF-divergence can be built from a divergence for fuzzy sets. In particular, given a divergence  $D$  for fuzzy sets and a component-wise increasing function  $f : [0, 1] \times [0, 1] \rightarrow [0, 1]$  satisfying  $f(0, 0) = 0$ , the function

$$D_{\text{AIF}}(A, B) = f(D(\mu_A, \mu_B), D(v_A, v_B))$$

for any  $A, B \in \text{AIFS}(X)$  is an AIF-divergence, where  $\mu_A, v_A, \mu_B, v_B$  are considered fuzzy sets (Montes et al. 2015, Prop. 4.7). Furthermore, if  $D$  is local,  $D_{\text{AIF}}$  is local if and only if  $f(x, y) = \alpha x + \beta y$  for some  $\alpha, \beta > 0$  (Montes et al. 2016, Prop. 5.2).

Following a similar reasoning, we can define an SK-entropy using a local divergence for fuzzy sets, just imposing some additional conditions on  $f$  and on the fuzzy divergence. For this, we consider the following domain where the function  $h_{\text{FS}}$  will be defined:

$$\mathcal{D}_2 = \left\{ (x, y) \mid y = 1 \right\} \cup \left\{ (x, y) \mid x \leq \frac{1}{2}, y = 0 \right\}.$$

**Proposition 3** Consider a local divergence  $D$  for fuzzy sets induced by the function  $h_{\text{FS}}$  and let  $f : [0, 1] \times [0, 1] \rightarrow [0, 1]$  be a function satisfying

- (f1)  $f(u, v) = 0$  if and only if  $u = v = 0$ .
- (f2)  $f$  is component-wise increasing.
- (f3)  $f(u, v) = 1$  if and only if  $u = 1$  and  $v = 0$ .

$$(f4) \quad f(u, v) = f(v, u).$$

Then, the function  $E$  defined by:

$$E(A) = \frac{1}{n} \sum_{x \in X} f(h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)), h_{\text{FS}}(v_A(x), v_{C_A}(x))) \quad (5)$$

is an SK-entropy if and only if  $h_{\text{FS}}$  satisfies the following additional conditions in  $\mathcal{D}_2$ :

- (loc.1') For  $(u, v) \in \mathcal{D}_2$ ,  $h_{\text{FS}}(u, v) = 0$  if and only if  $u = v$ ;
- (loc.4) For  $(u, v) \in \mathcal{D}_2$ ,  $h_{\text{FS}}(u, v) = 1$  if and only if  $u = 0, v = 1$ .

**Proof** First we assume that  $h_{\text{FS}}$  satisfies the additional conditions and prove that  $E$  is an SK-entropy.

(ISK1) :  $E(A) = 0$  if and only if

$$f(h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)), h_{\text{FS}}(v_A(x), v_{C_A}(x))) = 0$$

for any  $x \in X$ . From (f1),  $f(u, v) = 0$  if and only if  $u = v = 0$ , which is equivalent to

$$h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)) = h_{\text{FS}}(v_A(x), v_{C_A}(x)) = 0,$$

but from (loc.1') this happens if and only if  $\mu_A(x) = \mu_{C_A}(x)$  and  $v_A(x) = v_{C_A}(x)$ , which from Proposition 2 is equivalent to  $A = C_A$ . So  $A$  is a crisp set.

(ISK2') :  $E(A) = 1$  if and only if

$$f(h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)), h_{\text{FS}}(v_A(x), v_{C_A}(x))) = 1$$

for any  $x \in X$ . From (f3), this happens if and only if  $h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)) = 1$  and  $h_{\text{FS}}(v_A(x), v_{C_A}(x)) = 0$ . Also, from (loc.4) it holds that  $h_{\text{FS}}(u, v) = 1$  if and only if  $u = 0, v = 1$ . However, this happens if and only if  $\mu_A(x) = 0$  and  $\mu_{C_A}(x) = 1$ , which is equivalent to  $v_A(x) = 0$ .

(ISK3) : In order to prove  $E(A) = E(A^c)$ , we see that:

$$\begin{aligned} & f(h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)), h_{\text{FS}}(v_A(x), v_{C_A}(x))) \\ &= f(h_{\text{FS}}(\mu_{A^c}(x), \mu_{C_{A^c}}(x)), h_{\text{FS}}(v_{A^c}(x), v_{C_{A^c}}(x))) \end{aligned}$$

for any  $x \in X$ . First of all, since  $\mu_{A^c}(x) = v_A(x)$  and  $v_{A^c}(x) = \mu_A(x)$ , we only need to prove that:

$$\begin{aligned} & f(h_{\text{FS}}(\mu_A(x), \mu_{C_A}(x)), h_{\text{FS}}(v_A(x), v_{C_A}(x))) \\ &= f(h_{\text{FS}}(v_A(x), \mu_{C_{A^c}}(x)), h_{\text{FS}}(\mu_A(x), v_{C_{A^c}}(x))). \quad (6) \end{aligned}$$

If  $\mu_A(x) = v_A(x)$ , this means that  $\mu_{C_A}(x) = \mu_{C_{A^c}}(x) = 1$  and  $v_{C_A}(x) = v_{C_{A^c}}(x) = 0$ , which implies that the equality in Eq. (6) holds.



Assume now that  $\mu_A(x) \neq \nu_A(x)$ . From Proposition 2,  $C_A^c(x) = C_{A^c}(x)$ , which means that  $\mu_{C_A}(x) = \nu_{C_{A^c}}(x)$  and  $\nu_{C_A}(x) = \mu_{C_{A^c}}(x)$ . Also, note that  $\mu_{A^c}(x) = \nu_{A^c}(x)$  and  $\nu_{A^c}(x) = \mu_{A^c}(x)$ . Using these facts, as well as property (f4), it holds that:

$$\begin{aligned} & f(h_{FS}(\mu_A(x), \mu_{C_A}(x)), h_{FS}(\nu_A(x), \nu_{C_A}(x))) \\ &= f(h_{FS}(\nu_{A^c}(x), \nu_{C_{A^c}}(x)), h_{FS}(\mu_{A^c}(x), \mu_{C_{A^c}}(x))) \\ &= f(h_{FS}(\mu_{A^c}(x), \mu_{C_{A^c}}(x)), h_{FS}(\nu_{A^c}(x), \nu_{C_{A^c}}(x))) \\ &= f(h_{FS}(\nu_A(x), \mu_{C_{A^c}}(x)), h_{FS}(\mu_A(x), \nu_{C_{A^c}}(x))). \end{aligned}$$

(ISK4) : Assume that  $\mu_A(x) \leq \mu_B(x) < \nu_B(x) \leq \nu_A(x)$ . In this case, it holds that  $\mu_{C_A}(x) = \mu_{C_B}(x) = 0$  and  $\nu_{C_A}(x) = \nu_{C_B}(x) = 1$ . By the property (loc.3) of  $h_{FS}$ , it follows that:

$$\begin{aligned} h_{FS}(\mu_A(x), 0) &\leq h_{FS}(\mu_B(x), 0) \text{ and} \\ h_{FS}(\nu_A(x), 0) &\leq h_{FS}(\nu_B(x), 0), \end{aligned}$$

and by (f2), it follows that

$$\begin{aligned} & f(h_{FS}(\mu_A(x), \mu_{C_A}(x)), h_{FS}(\nu_A(x), \nu_{C_A}(x))) \\ &= f(h_{FS}(\mu_A(x), 0), h_{FS}(\nu_A(x), 1)) \\ &\leq f(h_{FS}(\mu_B(x), 0), h_{FS}(\nu_B(x), 1)) \\ &= f(h_{FS}(\mu_B(x), \mu_{C_B}(x)), h_{FS}(\nu_B(x), \nu_{C_B}(x))). \end{aligned}$$

Assume now that  $\nu_A(x) \leq \nu_B(x) \leq \mu_B(x) \leq \mu_A(x)$ , which implies that  $\mu_{C_A}(x) = \mu_{C_B}(x) = 1$  and  $\nu_{C_A}(x) = \nu_{C_B}(x) = 0$ . Using the property (loc.3) of  $h_{FS}$ , it holds that:

$$\begin{aligned} h_{FS}(\mu_A(x), 1) &\leq h_{FS}(\mu_B(x), 1) \text{ and} \\ h_{FS}(\nu_A(x), 0) &\leq h_{FS}(\nu_B(x), 0). \end{aligned}$$

Also, (f2) implies that:

$$\begin{aligned} & f(h_{FS}(\mu_A(x), \mu_{C_A}(x)), h_{FS}(\nu_A(x), \nu_{C_A}(x))) \\ &= f(h_{FS}(\mu_A(x), 1), h_{FS}(\nu_A(x), 0)) \\ &\leq f(h_{FS}(\mu_B(x), 1), h_{FS}(\nu_B(x), 0)) \\ &= f(h_{FS}(\mu_B(x), \mu_{C_B}(x)), h_{FS}(\nu_B(x), \nu_{C_B}(x))). \end{aligned}$$

On the other hand, let us now assume that  $E$  is an SK-entropy and we prove that  $f$  and  $h_{FS}$  must satisfy the additional conditions.

(loc.1') : Consider  $(u, v) \in \mathcal{D}_2$  such that  $v = 1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle(x, u, 0)\rangle$ . Then,  $C_A(x) = 1$ , hence:

$$E(A) = f(h_{FS}(u, 1), h_{FS}(0, 0)).$$

Also, from (ISK1),  $E(A) = 0$  if and only if  $A$  is a crisp set, which by Proposition 2 is equivalent to  $A = C_A$ . This holds if and only if  $u = 1$ . Finally,  $E(A) = 0$  is equivalent to:

$$f(h_{FS}(u, 1), h_{FS}(0, 0)) = 0,$$

but from (f1) this happens if and only if  $h_{FS}(u, 1) = h_{FS}(0, 0) = 0$ . We therefore conclude that  $u = 1$  if and only if  $h_{FS}(u, 1) = 0$ .

Now consider  $(u, v) \in \mathcal{D}_2$  such that  $v = 0$  and define the AIFS  $A$  on  $\{x\}$  such that  $A = \langle(x, 1 - u, u)\rangle$ . Note that since  $(u, 0) \in \mathcal{D}_2$ , this implies that  $u \leq \frac{1}{2}$ , or equivalently,  $1 - u \geq \frac{1}{2}$ . Then,  $C_A(x) = 1$ , hence:

$$E(A) = f(h_{FS}(1 - u, 1), h_{FS}(u, 0)).$$

Now, from (ISK1),  $E(A) = 0$  if and only if  $A$  is a crisp set, which by Proposition 2 is equivalent to  $A = C_A$ . This happens if and only if  $u = 0$ . Finally,  $E(A) = 0$  is equivalent to

$$f(h_{FS}(1 - u, 1), h_{FS}(u, 0)) = 0,$$

but from (f1) this happens if and only if  $h_{FS}(1 - u, 1) = h_{FS}(u, 0) = 0$ . We conclude that  $u = 0$  if and only if  $h_{FS}(u, 0) = 0$ .

(loc.4) : First of all, take  $(u, v) = (0, 1) \in \mathcal{D}_2$  and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle(x, 0, 0)\rangle$ . From (ISK2'):

$$E(A) = f(h_{FS}(0, 1), h_{FS}(0, 0)) = 1.$$

However, using (f3), this happens if and only if  $h_{FS}(0, 1) = 1$  and  $h_{FS}(0, 0) = 0$ . We conclude that  $h_{FS}(0, 1) = 1$ .

On the other hand, let us see that if  $h_{FS}(u, v) = 1$ , it must hold that  $u = 0, v = 1$ . First of all, assume that  $(u, v) \in \mathcal{D}_2$  such that  $v = 0$  and  $h_{FS}(u, v) = 1$ . This means that  $u < 1 - u$ . Let us define the AIFS  $A$  on  $\{x\}$  by  $A = \langle(x, u, 1 - u)\rangle$ . Using the definition of  $E$  and (f2), it holds that:

$$\begin{aligned} E(A) &= f(h_{FS}(u, 0), h_{FS}(1 - u, 1)) \\ &= f(1, h_{FS}(1 - u, 1)) \geq f(1, 0) = 1, \end{aligned}$$

where the last equality follows from (f3). We conclude that  $E(A) = 1$ . However, from (ISK2') this is equivalent to  $u = 1 - u = 0$ , a contradiction.

Therefore, take  $(u, v) \in \mathcal{D}_2$  such that  $v = 1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle(x, u, 0)\rangle$ . Then, it holds that:

$$E(A) = f(h_{FS}(u, 1), h_{FS}(0, 0)) = f(h_{FS}(u, 1), 0),$$

where the last equality follows from (loc.1). Now, from (f3),  $f(h_{FS}(u, 1), 0) = 1$  if and only if  $h_{FS}(u, 1) = 1$ . However, from (ISK2'),  $E(A) = 1$  if and only if  $u = 0$ . Therefore, we

conclude that  $u = 0$  and  $h_{FS}(u, 1) = 1$  are equivalent when  $v = 1$ .  $\square$

**Remark 1** From (loc.4),  $h_{FS}(u, v) = 1$  if and only if  $u = 0, v = 1$ . This means that  $h_{FS}(v_A(x), v_{C_A}(x))$  cannot take the value 1 because  $v_A(x) = 0$  and  $v_{C_A}(x) = 1$  cannot happen at the same time. That is why the function  $f$  in Proposition 3 is defined in the domain  $[0, 1] \times [0, 1)$ , not including the value 1 in the second component.

### 3.3 Examples of SK-entropies based on AIF-divergences

In the literature, several different measures of comparison have been introduced. In an earlier work (Montes et al. 2016, Section III-C), we showed some examples of local AIF-divergences, like the Hamming distance (Szmidt and Kacprzyk 2000) and Hausdorff distance (Grzegorzewski 2004, denoted by  $l_{AIF}$  and  $d_H$ , respectively, and the two measures proposed by Hong and Kim (1999),<sup>2</sup> denoted by  $D_C$  and  $D_L$ . In this subsection, we consider these four local AIF-divergences and we investigate whether they satisfy the conditions of Theorem 3 and can therefore be used to define SK-entropy measures. Recall that these AIF-divergences are defined by:

$$l_{AIF}(A, B) = \frac{1}{2n} \sum_{x \in X} |\mu_A(x) - \mu_B(x)| + |v_A(x) - v_B(x)| + |\pi_A(x) - \pi_B(x)|. \quad (7)$$

$$d_H(A, B) = \frac{1}{n} \sum_{x \in X} \max\{|\mu_A(x) - \mu_B(x)|, |v_A(x) - v_B(x)|\}. \quad (8)$$

$$D_C(A, B) = \frac{1}{n} \sum_{x \in X} |\mu_A(x) - \mu_B(x)| + |v_A(x) - v_B(x)|. \quad (9)$$

$$D_L(A, B) = \frac{1}{2n} \sum_{x \in X} |\mu_A(x) - v_A(x) - \mu_B(x) + v_B(x)| + |\mu_A(x) - \mu_B(x)| + |v_A(x) - v_B(x)|. \quad (10)$$

In the next example, we show that both Hamming and Hausdorff distances satisfy the conditions of Theorem 3 and, surprisingly, they both induce the same SK-entropy. On the contrary, we show that the Hong and Kim divergences do not induce an SK-entropy.

<sup>2</sup> The original definitions of  $D_C$  and  $D_L$  are slightly different from those of Eqs. (9) and (10). The difference is that in Hong and Kim (1999),  $D_C$  was divided by 2 and  $D_L$  by 4, instead of 2. In this paper, we consider the definitions of Eqs. (9) and (10) just to make  $D_C$  and  $D_L$  to satisfy the normalization property mentioned in Sect. 2.3.

**Example 1** Consider first the Hamming and Hausdorff distances defined in Eqs. (7) and (8). We shall show that their associated functions, that will be denoted by  $h_l$  and  $h_d$ , satisfy the conditions of Theorem 3. First of all, note that the functions  $h_l$  and  $h_d$  inducing the Hamming and Hausdorff distances are given by:

$$h_l(u_1, u_2, v_1, v_2) = \frac{1}{2} \left( |u_1 - v_1| + |u_2 - v_2| + |u_1 + u_2 - v_1 - v_2| \right).$$

$$h_d(u_1, u_2, v_1, v_2) = \max\{|u_1 - v_1|, |u_2 - v_2|\}.$$

Also:

$$h_l(u, v, 1, 0) = \frac{1}{2} \left( (1 - u) + v + (1 - u - v) \right) = 1 - u.$$

$$h_l(u, v, 0, 1) = \frac{1}{2} \left( u + (1 - v) + (1 - u - v) \right) = 1 - v.$$

$$h_d(u, v, 1, 0) = \max\{1 - u, v\} = 1 - u.$$

$$h_d(u, v, 0, 1) = \max\{u, 1 - v\} = 1 - v.$$

Therefore, we can see that  $h_l$  and  $h_d$  coincide in  $\mathcal{D}_1$ . Now, let us prove that  $h_l$ , and consequently also  $h_d$ , satisfy the required properties:

(AIF-loc.1')  $h_l(u, v, 1, 0) = 0$  if and only if  $1 - u = 0$ , or equivalently, if and only if  $u = 1$ . But also,  $u = 1$  is equivalent to  $v = 0$ , because  $u + v \leq 1$ . Similarly,  $h_l(u, v, 0, 1) = 0$  if and only if  $1 - v = 0$ , or equivalently, if and only if  $v = 1$ . But this is equivalent to  $u = 0$  because  $u + v \leq 1$ .

(AIF-loc.5) In (Montes et al. 2016, Section IV), we have proven that  $h_l$  is symmetric.

(AIF-loc.6) Finally,  $h_l(u, v, 1, 0) = 1$  if and only if  $u = 0$ , but this happens if and only if  $v = 0$ . On the other hand,  $h_l(u, v, 0, 1) = 1$  is not possible, because this would mean that  $v = 0$ , but in that case we would compare  $(u, v)$  with  $(1, 0)$ , not with  $(0, 1)$ .

We conclude that both  $h_l$  and  $h_d$  satisfy the additional conditions of Theorem 3, so each of  $l_{AIF}$  and  $d_H$  induces an SK-entropy. Furthermore, since  $h_l = h_d$  in the domain  $\mathcal{D}_1$  that we are considering in the definition of the entropy, we conclude that they induce the same SK-entropy.

Let us now show the SK-entropy that they induce:

$$E(A) = 1 - \frac{1}{n} \sum_{x \in X} \max\{\mu_A(x), v_A(x)\}. \quad (11)$$

Now if  $\mu_A(x) \geq v_A(x)$ , then  $\mu_{C_A}(x) = 1, v_{C_A}(x) = 0$  so:

$$h_l(\mu_A(x), v_A(x), \mu_{C_A}(x), v_{C_A}(x)) = h_l(\mu_A(x), v_A(x), 1, 0) = 1 - \mu_A(x).$$

On the other hand, if  $\mu_A(x) < v_A(x), \mu_{C_A}(x) = 0, v_{C_A}(x) = 1$ , and then:

$$\begin{aligned}
 h_I(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) \\
 = h_I(\mu_A(x), \nu_A(x), 0, 1) = 1 - \nu_A(x).
 \end{aligned}$$

Substituting these values, we obtain the following:

$$\begin{aligned}
 E(A) &= \frac{1}{n} \left( \sum_{x|\mu_A(x) \geq \nu_A(x)} 1 - \mu_A(x) \right. \\
 &\quad \left. + \sum_{x|\mu_A(x) < \nu_A(x)} 1 - \nu_A(x) \right) \\
 &= 1 - \frac{1}{n} \left( \sum_{x|\mu_A(x) \geq \nu_A(x)} \mu_A(x) \right. \\
 &\quad \left. + \sum_{x|\mu_A(x) < \nu_A(x)} \nu_A(x) \right) \\
 &= 1 - \frac{1}{n} \sum_{x \in X} \max \{ \mu_A(x), \nu_A(x) \}.
 \end{aligned}$$

Let us now see that the local AIF-divergences of Hong and Kim defined in Eqs. (9) and (10) do not satisfy the conditions of Theorem 3. It can be easily seen that the functions  $h_C$  and  $h_L$  associated with  $D_C$  and  $D_L$ , respectively, are given by:

$$\begin{aligned}
 h_C(u_1, u_2, v_1, v_2) &= |u_1 - v_1| + |u_2 - v_2|. \\
 h_L(u_1, u_2, v_1, v_2) &= \frac{1}{2} (|u_1 - u_2 - v_1 + v_2| \\
 &\quad + |u_1 - v_1| + |u_2 - v_2|).
 \end{aligned}$$

However, these functions do not satisfy one of the conditions of Theorem 3 because for  $\alpha = \mu_A(x) = \nu_A(x) > 0$ , it happens that  $\mu_{C_A}(x) = 1, \nu_{C_A}(x) = 0$  and:

$$h_C(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x)) = h_C(\alpha, \alpha, 1, 0) = 1,$$

but  $\alpha \neq 0$ . Therefore, neither  $h_C$  nor  $h_D$  satisfies (AIF-loc.6). Hence, neither  $D_C$  nor  $D_L$  induces an SK-entropy.

Next example shows that the SK-entropy defined by Guo and Song Guo and Song (2014) by:

$$\begin{aligned}
 E(A) &= \frac{1}{n} \sum_{x \in X} (1 - |\mu_A(x) \\
 &\quad - \nu_A(x)|) \cdot \left( \frac{2 - \mu_A(x) - \nu_A(x)}{2} \right)
 \end{aligned}$$

can also be defined as in Theorem 3 through a local AIF-divergence.

**Example 2** Consider now the function  $D$  given for any  $A, B \in \text{AIFS}(X)$  is:

$$\begin{aligned}
 D(A, B) &= \frac{1}{n} \sum_{x \in X} (|\mu_A(x) - \mu_B(x)| \\
 &\quad + |\nu_A(x) - \nu_B(x)|) \cdot \left( \frac{1 + |\pi_A(x) - \pi_B(x)|}{2} \right).
 \end{aligned}$$

It can be easily proven that this function is a local AIF-divergence with associated function:

$$\begin{aligned}
 h_{\text{IF}}(u_1, u_2, v_1, v_2) \\
 = (|u_1 - v_1| + |u_2 - v_2|) \cdot \left( \frac{1 + |u_1 + u_2 - v_1 - v_2|}{2} \right).
 \end{aligned}$$

The function  $h_{\text{IF}}$  also satisfies the additional conditions of Theorem 3; hence, it induces an SK-entropy. Note that when  $u \geq v$ , we obtain:

$$h_{\text{IF}}(u, v, 1, 0) = (1 - u + v) \cdot \left( \frac{2 - u - v}{2} \right),$$

while for  $u < v$ , we obtain:

$$h_{\text{IF}}(u, v, 0, 1) = (1 - v + u) \cdot \left( \frac{2 - u - v}{2} \right).$$

Therefore, the SK-entropy induced by  $D$  is given by:

$$\begin{aligned}
 E(A) &= \frac{1}{n} \left( \sum_{x|\mu_A(x) \geq \nu_A(x)} (1 - \mu_A(x) \right. \\
 &\quad \left. + \nu_A(x)) \cdot \left( \frac{2 - \mu_A(x) - \nu_A(x)}{2} \right) \right. \\
 &\quad \left. + \sum_{x|\mu_A(x) < \nu_A(x)} (1 + \mu_A(x) \right. \\
 &\quad \left. - \nu_A(x)) \cdot \left( \frac{2 - \mu_A(x) - \nu_A(x)}{2} \right) \right) \\
 &= \frac{1}{n} \sum_{x \in X} (1 - |\mu_A(x) \\
 &\quad - \nu_A(x)|) \cdot \left( \frac{2 - \mu_A(x) - \nu_A(x)}{2} \right).
 \end{aligned}$$

To conclude this section, let us consider two examples of SK-entropies built using the procedure in Proposition 3.

**Example 3** Consider the Hamming distance for fuzzy sets, which is defined for any  $A, B \in FS(X)$  by:

$$h_{FS}(A, B) = \frac{1}{n} \sum_{x \in X} |\mu_A(x) - \mu_B(x)|, \tag{12}$$

where  $\mu_A$  and  $\mu_B$  denote the membership function of two fuzzy sets  $A$  and  $B$ . The Hamming distance for fuzzy sets is

known to be a local divergence for fuzzy sets, and its associated function  $h$  is given by  $h(u, v) = |u - v|$ . This function  $h$  satisfies the additional conditions required in Proposition 3. Here,  $h(u, v) = 0$  if and only if  $u = v = 0$ ; on the other hand,  $h(u, v) = |u - v| = 1$  if and only if either  $u = 1, v = 0$  or  $u = 0, v = 1$ . However,  $(u, v) = (1, 0) \notin \mathcal{D}_2$ . Hence,  $h(u, v) = 1$  for  $(u, v) \in \mathcal{D}_2$  if and only if  $u = 0, v = 1$ . Also, we consider the function  $f$  given by  $f(u, v) = u + v - uv$ , which is usually called the product t-conorm (see Klement et al. (2000) for details in t-norms and t-conorms). As any t-conorm,  $f$  satisfies conditions (f1), (f2) and (f4). Also, since  $f$  is the t-conorm associated with a strict t-norm (algebraic product), it also satisfies (f3).<sup>3</sup>

Therefore, applying Proposition 3 we can define an SK-entropy by using Eq. (5). Note that, if  $\mu_A(x) \geq \nu_A(x)$ ,  $\mu_{C_A}(x) = 1, \nu_{C_A}(x) = 0$ , and it holds that:

$$\begin{aligned} f(h(\mu_A(x), \mu_{C_A}(x)), h(\nu_A(x), \nu_{C_A}(x))) &= f(h(\mu_A(x), 1), h(\nu_A(x), 0)) \\ &= f(1 - \mu_A(x), \nu_A(x)) \\ &= 1 - \mu_A(x) + \nu_A(x) - (1 - \mu_A(x))\nu_A(x) \\ &= (1 - \mu_A(x))(1 - \nu_A(x)) + \nu_A(x). \end{aligned}$$

Also, if  $\mu_A(x) < \nu_A(x), \mu_{C_A}(x) = 0, \nu_{C_A}(x) = 1$ , and then:

$$\begin{aligned} f(h(\mu_A(x), \mu_{C_A}(x)), h(\nu_A(x), \nu_{C_A}(x))) &= f(h(\mu_A(x), 0), h(\nu_A(x), 1)) \\ &= f(\mu_A(x), 1 - \nu_A(x)) \\ &= 1 - \nu_A(x) + \mu_A(x) - (1 - \nu_A(x))\mu_A(x) \\ &= (1 - \mu_A(x))(1 - \nu_A(x)) + \mu_A(x). \end{aligned}$$

Therefore, the SK-entropy defined using Eq. (5) is given by:

$$\begin{aligned} E(A) &= \frac{1}{n} \left( \sum_{x|\mu_A(x) \geq \nu_A(x)} (1 - \mu_A(x))(1 - \nu_A(x)) + \nu_A(x) \right. \\ &\quad \left. + \sum_{x|\mu_A(x) < \nu_A(x)} (1 - \mu_A(x))(1 - \nu_A(x)) + \mu_A(x) \right) \\ &= \frac{1}{n} \sum_{x \in X} (1 - \mu_A(x))(1 - \nu_A(x)) + \min\{\mu_A(x), \nu_A(x)\}. \end{aligned}$$

**Example 4** Consider again the Hamming distance for fuzzy sets defined in Eq. (12) and take the function  $f$  given by

<sup>3</sup> Recall that any t-conorm  $f$  satisfies  $f(u, v) = 1$  if and only if either  $u = 1$  or  $v = 1$ . However, remember that in the statement of Proposition 3 we are restricting the domain of  $f$  to the set  $[0, 1] \times [0, 1)$ , hence  $f(u, v) = 1$  if and only if  $u = 1$ .

$f(u, v) = \max\{u, v\}$ . This function is also a t-conorm satisfying (f1) to (f4), so using the Hamming distance for fuzzy sets and the maximum t-conorm, we can apply Proposition 3 to define an SK-entropy. Let us note that for  $\mu_A(x) \geq \nu_A(x), \mu_{C_A}(x) = 1, \nu_{C_A}(x) = 0$ , so:

$$\begin{aligned} f(h(\mu_A(x), \mu_{C_A}(x)), h(\nu_A(x), \nu_{C_A}(x))) &= f(h(\mu_A(x), 1), h(\nu_A(x), 0)) \\ &= \max\{1 - \mu_A(x), \nu_A(x)\} = 1 - \mu_A(x), \end{aligned}$$

and if  $\mu_A(x) < \nu_A(x), \mu_{C_A}(x) = 0, \nu_{C_A}(x) = 1$ , therefore:

$$\begin{aligned} f(h(\mu_A(x), \mu_{C_A}(x)), h(\nu_A(x), \nu_{C_A}(x))) &= f(h(\mu_A(x), 0), h(\nu_A(x), 1)) \\ &= \max\{\mu_A(x), 1 - \nu_A(x)\} = 1 - \nu_A(x). \end{aligned}$$

However, taking Example 1 into account, the SK-entropy defined using the Hamming distance for fuzzy sets and the maximum t-conorm coincides with the SK-entropy defined from the Hamming and Hausdorff distances for AIFSs, given in Eq. (11).

### 4 BB-entropies

We now investigate the other type of entropies, which measures how different is an AIFS from being a fuzzy set. For this we consider the definition of entropy given by Burrillo and Bustince (1996) and we investigate whether we can define an entropy using local AIF-divergences.

#### 4.1 Burrillo and Bustince’s entropy

To the best of our knowledge, the first proposal of entropy for AIFSs was given by Burrillo and Bustince (1996).

**Definition 5** (Burrillo and Bustince (1996)) A mapping  $I : \text{AIFS}(X) \rightarrow [0, 1]$  is called *entropy* if it satisfies the following properties:

- ( $I_{BB1}$ )  $I(A) = 0$  if and only if  $A \in \text{FS}(X)$ .
- ( $I_{BB2}$ )  $I(A) = 1$  if and only if  $\mu_A = \nu_A = 0$ .
- ( $I_{BB3}$ )  $I(A) = I(A^c)$ .
- ( $I_{BB4}$ )  $I(A) \geq I(B)$  if  $\mu_A \leq \mu_B$  and  $\nu_A \leq \nu_B$ .

This type of entropy measures how intuitionistic is and AIFS, or in other words, how different is an AIFS from a fuzzy set. The first property ( $I_{BB1}$ ) says that the entropy is zero if, and only if, the hesitation index is zero, or equivalently, if and only if the AIFS is a fuzzy set. ( $I_{BB2}$ ) says that the entropy is maximal if and only if the hesitation index is 1, which means that there is a total lack of information. The third condition

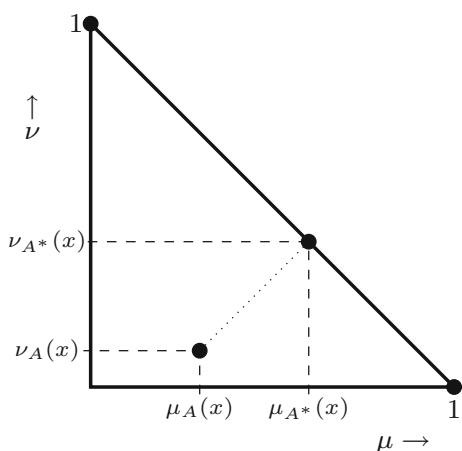


Fig. 5 Closest fuzzy set to an AIFS

says that the entropy is closed under complementaries, while ( $I_{BB4}$ ) means that the greater the hesitation index, the greater the entropy.

In what follows, a function  $I$  satisfying properties ( $I_{BB1}$ ) to ( $I_{BB4}$ ) will be called *BB-entropy*.

### 4.2 BB-entropies based on local AIF-divergences

Our aim is now to define BB-entropies using local AIF-divergences in a similar manner as we did in Sect. 3.2. For this, we define the closest fuzzy set to an AIFS.

**Definition 6** Given  $A \in \text{AIFS}(X)$ , we define the *closest fuzzy set* to  $A$ , denoted by  $A^*$ , by  $\mu_{A^*}(x) = \mu_A(x) + \frac{\pi_A(x)}{2}$ .

It can be seen that  $1 - \mu_{A^*}(x) = \nu_A(x) + \frac{\pi_A(x)}{2}$ . The interpretation of  $A^*$  can be easily seen in Fig. 5. In this figure, the fuzzy sets are those elements that belong to the line from  $(1,0)$  to  $(0,1)$ . In order to define the closest fuzzy set to an AIFS, we find the point having the shortest distance from the point  $(\mu_A(x), \nu_A(x))$  on the  $(0, 1) - (1, 0)$  line. This results in an equal distribution of the hesitation index into the membership and non-membership values.

Now, we define the BB-entropy of an AIFS as the AIF-divergence between the AIFS and its closest fuzzy set. Therefore, the domain of the function  $h_{\text{IF}}$  associated with the local AIF-divergence  $D$  will be

$$\mathcal{D}_3 = \left\{ \left( u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2} \right) \mid u+v \leq 1 \right\}.$$

**Theorem 4** Consider a local AIF-divergence  $D$  with associated function  $h_{\text{IF}}$  and define the function  $I$  by:

$$\begin{aligned} I(A) &= D(A, A^*) \\ &= \frac{1}{n} \sum_{x \in X} h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_{A^*}(x), 1 - \mu_{A^*}(x)). \end{aligned} \tag{13}$$

Then,  $I$  is a BB-entropy if and only if  $h_{\text{IF}}$  satisfies the following additional properties:

(AIF-loc.1'') For  $(u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2}) \in \mathcal{D}_3$ , it holds that  $h_{\text{IF}}(u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2}) = 0$  if and only if  $u + v = 1$ .

(AIF-loc.7) The function  $h^*$  defined by

$$h^*(u, v) = h_{\text{IF}} \left( u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2} \right),$$

for  $u + v \leq 1$ , is symmetric and decreasing in both  $u$  and  $v$ .

(AIF-loc.8)  $h_{\text{IF}}(u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2}) = 1$  for  $u + v \leq 1$  if and only if  $u = v = 0$ .

**Proof ( $I_{BB1}$ )** : By definition,  $I(A) = 0$  if and only if  $h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_{A^*}(x), 1 - \mu_{A^*}(x)) = 0$  for any  $x \in X$ , but from (AIF-loc.1'') this is equivalent to  $\mu_A(x) + \nu_A(x) = 1$  for any  $x \in X$ , or in other words,  $A$  is a fuzzy set.

( $I_{BB2}$ ) : By definition,  $I(A) = 1$  if and only if

$$h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_{A^*}(x), 1 - \mu_{A^*}(x)) = 1$$

for any  $x \in X$ , and by (AIF-loc.8) this is equivalent to  $\mu_A(x) = \nu_A(x) = 0$  for any  $x \in X$ .

( $I_{BB3}$ ) : In order to prove that  $I(A) = I(A^c)$ , it is enough to realize that, by (AIF-loc.7), it holds that:

$$\begin{aligned} h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_{A^*}(x), 1 - \mu_{A^*}(x)) \\ = h_{\text{IF}}(\nu_A(x), \mu_A(x), \mu_{A^{c*}}(x), 1 - \mu_{A^{c*}}(x)), \end{aligned}$$

and that

$$\begin{aligned} \mu_{A^{c*}}(x) &= \mu_{A^c}(x) + \frac{\pi_{A^c}(x)}{2} \\ &= \nu_A(x) + \frac{\pi_A(x)}{2} = \nu_{A^*}(x). \\ \nu_{A^{c*}}(x) &= \nu_{A^c}(x) + \frac{\pi_{A^c}(x)}{2} \\ &= \mu_A(x) + \frac{\pi_A(x)}{2} = \mu_{A^*}(x). \end{aligned}$$

( $I_{BB4}$ ) : Assume that  $\mu_A(x) \leq \mu_B(x)$  and  $\nu_A(x) \leq \nu_B(x)$ . Then:

$$\begin{aligned}
 &h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_{A^*}(x), 1 - \mu_{A^*}(x)) \\
 &= h_{\text{IF}}\left(\mu_A(x), \nu_A(x), \frac{1 + \mu_A(x) - \nu_A(x)}{2}, \frac{1 + \nu_A(x) - \mu_A(x)}{2}\right) \\
 &\geq h_{\text{IF}}\left(\mu_B(x), \nu_A(x), \frac{1 + \mu_B(x) - \nu_A(x)}{2}, \frac{1 + \nu_A(x) - \mu_B(x)}{2}\right) \\
 &\geq h_{\text{IF}}\left(\mu_B(x), \nu_B(x), \frac{1 + \mu_B(x) - \nu_B(x)}{2}, \frac{1 + \nu_B(x) - \mu_B(x)}{2}\right) \\
 &= h_{\text{IF}}(\mu_B(x), \nu_B(x), \mu_{B^*}(x), 1 - \mu_{B^*}(x)),
 \end{aligned}$$

where the inequalities follows from (AIF-loc.7).

Now assume that  $I$  is a BB-entropy and let us prove that  $h_{\text{IF}}$  satisfies the additional conditions.

(AIF-loc.1''): Take  $u, v$  such that  $u + v \leq 1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle(x, u, v)\rangle$ . Then, it holds that:

$$I(A) = h_{\text{IF}}\left(u, v, \frac{1 + u - v}{2}, \frac{1 + v - u}{2}\right),$$

but from (IBB1),  $I(A) = 1$  if and only if  $u = v = 0$ .

(AIF-loc.7): Let us prove that  $h^*$  is symmetric. Take  $u, v$  such that  $u + v \leq 1$ , and define  $A = \langle(x, u, v)\rangle$ . From (IBB3), it holds that:

$$\begin{aligned}
 h^*(u, v) &= h_{\text{IF}}\left(u, v, \frac{1 + u - v}{2}, \frac{1 + v - u}{2}\right) \\
 &= I(A) = I(A^c) \\
 &= h_{\text{IF}}\left(v, u, \frac{1 + v - u}{2}, \frac{1 + u - v}{2}\right) = h^*(v, u).
 \end{aligned}$$

Let us now see that  $h^*$  is decreasing in the first component. Take  $u_1, u_2, v$  such that  $u_1 \leq u_2$  and  $u_2 + v \leq 1$  and define the AIFSs  $A, B$  on  $\{x\}$  by  $A = \langle(x, u_1, v)\rangle$  and  $B = \langle(x, u_2, v)\rangle$ . Then, from (IBB4) it holds that:

$$\begin{aligned}
 h^*(u_1, v) &= h_{\text{IF}}\left(u_1, v, \frac{1 + u_1 - v}{2}, \frac{1 + v - u_1}{2}\right) \\
 &= I(A) \geq I(B) \\
 &= h_{\text{IF}}\left(u_1, v, \frac{1 + u_2 - v}{2}, \frac{1 + v - u_2}{2}\right) = h^*(u_2, v).
 \end{aligned}$$

With a similar reasoning we can proof that  $h^*$  is also decreasing in the second component.

(AIF-loc.8): Take  $u, v$  such that  $u + v \leq 1$ , and define the AIFS  $A$  on  $\{x\}$  by  $A = \langle(x, u, v)\rangle$ . Then:

$$I(A) = \left(u, v, \frac{1 + u - v}{2}, \frac{1 + v - u}{2}\right).$$

But from (IBB2),  $I(A) = 1$  if and only if  $u = v = 0$ . □

There is an alternative way of defining a BB-entropy using AIF-divergences. This second approach is based on the comparison of the AIFS  $A$  with the fuzzy sets  $A^+$ , with

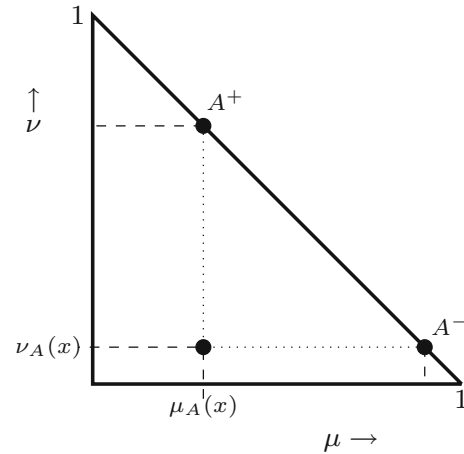


Fig. 6 Graphical representation of the sets  $A^+$  and  $A^-$

membership function  $\mu_{A^+}(x) = \mu_A(x)$ , and  $A^-$ , with membership function  $\mu_{A^-}(x) = \nu_A(x)$ . We first compute the AIF-divergence between  $A$  and  $A^+$  and between  $A$  and  $A^-$ , and then aggregate them. The fuzzy sets  $A^+$  and  $A^-$ , as well as their associated AIFS  $A$ , are graphically shown in Fig. 6.

For the next result, the domain  $\mathcal{D}_4$  of the function  $h_{\text{IF}}$  is given by:

$$\begin{aligned}
 \mathcal{D}_4 &= \{(u, v, u, 1 - u) \mid u + v \leq 1\} \\
 &\cup \{(u, v, 1 - v, v) \mid u + v \leq 1\}.
 \end{aligned}$$

**Proposition 4** Consider a local AIF-divergence  $D$  with associated function  $h_{\text{IF}}$  satisfying the following additional properties on  $\mathcal{D}_4$ :

- (AIF-loc.1''')  $h_{\text{IF}}(u_1, u_2, v_1, v_2) = 0$  for  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_4$  if and only if  $u_1 + u_2 = 1$ .
- (AIF-loc.5)  $h_{\text{IF}}(u_1, u_2, v_1, v_2) = h_{\text{IF}}(u_2, u_1, v_2, v_1)$  for any  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_4$ .
- (AIF-loc.9)  $h_{\text{IF}}(u, v, u, 1 - u)$  and  $h_{\text{IF}}(u, v, 1 - v, v)$  are decreasing in both  $u$  and  $v$ , for  $u + v \leq 1$ .
- (AIF-loc.10)  $h_{\text{IF}}(u_1, u_2, v_1, v_2) = 1$  for  $(u_1, u_2, v_1, v_2) \in \mathcal{D}_4$  if and only if  $u_1 = u_2 = 0$ .

Consider also a function  $f : [0, 1] \times [0, 1] \rightarrow [0, 1]$  such that

- (f1)  $f(u, v) = 0$  if and only if  $u = v = 0$ .
- (f2)  $f$  is component-wise increasing.
- (f5)  $f(u, v) = 1$  if and only if  $u = v = 1$ .
- (f6)  $f(u, v) = f(v, u)$ .

The function  $I : \text{AIFS}(X) \rightarrow [0, 1]$  defined by:

$$I(A) = \frac{1}{n} \sum_{x \in X} f(h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)), h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x))) \tag{14}$$

is a BB-entropy.

**Proof** (*I<sub>BB1</sub>*) :  $I(A) = 0$  if and only if

$$f(h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)), h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x))) = 0$$

for any  $x \in X$ . From (f1),  $I(A) = 0$  is equivalent to

$$h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)) = 0$$

and

$$h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x)) = 0,$$

which by (AIF-loc.1''') is equivalent to  $\mu_A(x) + \nu_A(x) = 1$  for any  $x \in X$ , or equivalent, if  $A$  is a fuzzy set.

(*I<sub>BB2</sub>*) :  $I(A) = 1$  if and only if

$$f(h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)), h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x))) = 1,$$

for any  $x \in X$ . Also, from (f5),  $I(A) = 1$  is equivalent to

$$h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)) = 1$$

and

$$h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x)) = 1,$$

which by (AIF-loc.10) holds if and only if  $\mu_A(x) = \nu_A(x) = 0$  for any  $x \in X$ .

(*I<sub>BB3</sub>*) : In order to check the equality  $I(A) = I(A^c)$ , we only need to prove that

$$f(h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \mu_A(x), \nu_A(x)), h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \mu_A(x))) = f(h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), \mu_{A^c}(x), 1 - \mu_{A^c}(x)), h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), 1 - \nu_{A^c}(x), \nu_{A^c}(x)))$$

for any  $x \in X$ . For this, note that from (AIF-loc.5), it holds that:

$$\begin{aligned} &h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), \mu_{A^c}(x), 1 - \mu_{A^c}(x)) \\ &= h_{\text{IF}}(\nu_A(x), \mu_A(x), \nu_A(x), 1 - \nu_A(x)) \\ &= h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x)). \\ &h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), 1 - \nu_{A^c}(x), \nu_{A^c}(x)) \\ &= h_{\text{IF}}(\nu_A(x), \mu_A(x), 1 - \mu_A(x), \mu_A(x)) \\ &= h_{\text{IF}}(\mu_A(x), \nu_A(x), \nu_A(x), 1 - \mu_A(x)). \end{aligned}$$

Therefore, taking (f6) into account, it holds that:

$$\begin{aligned} &f(h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)), h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x))) \\ &= f(h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), 1 - \nu_{A^c}(x), \nu_{A^c}(x)), h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), \mu_{A^c}(x), 1 - \mu_{A^c}(x))) \\ &= f(h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), \mu_{A^c}(x), 1 - \mu_{A^c}(x)), h_{\text{IF}}(\mu_{A^c}(x), \nu_{A^c}(x), 1 - \nu_{A^c}(x), \nu_{A^c}(x))). \end{aligned}$$

(*I<sub>BB4</sub>*) : Take  $A, B \in \text{AIFS}(X)$  such that  $\mu_A \leq \mu_B$  and  $\nu_A \leq \nu_B$ . By property (AIF-loc.9), it holds that:

$$\begin{aligned} &h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)) \\ &\geq h_{\text{IF}}(\mu_B(x), \nu_A(x), \mu_B(x), 1 - \mu_B(x)) \\ &\geq h_{\text{IF}}(\mu_B(x), \nu_B(x), \mu_B(x), 1 - \mu_B(x)). \\ &h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x)) \\ &\geq h_{\text{IF}}(\mu_B(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x)) \\ &\geq h_{\text{IF}}(\mu_B(x), \nu_B(x), 1 - \nu_B(x), \nu_B(x)). \end{aligned}$$

Therefore, using (f2) we conclude that:

$$\begin{aligned} &f(h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)), h_{\text{IF}}(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x))) \geq \\ &f(h_{\text{IF}}(\mu_B(x), \nu_B(x), \mu_B(x), 1 - \mu_B(x)), h_{\text{IF}}(\mu_B(x), \nu_B(x), 1 - \nu_B(x), \nu_B(x))). \end{aligned}$$

So,  $I(A) \geq I(B)$ . □

At a first glance, one may think that the converse implication in the previous proposition also holds. That is, given a function  $f$  satisfying (f1), (f2), (f5) and (f6), and  $I$  as defined in Eq. (14), then  $I$  is a BB-entropy, if and only if,  $h_{\text{IF}}$  satisfies properties (AIF-loc.1'''), (AIF-loc.5), (AIF-loc.9) and (AIF-loc.10). However, as the next example shows, the equivalence cannot be guaranteed.

**Example 5** Consider the function  $D$  given by:

$$D(A, B) = \frac{1}{n} \sum_{x \in X} |\mu_A(x) - \mu_B(x)| + |v_A(x) - v_B(x)|^2.$$

This function is a local AIF-divergence whose associated function  $h_{IF}$  is given by:

$$h_{IF}(u_1, u_2, v_1, v_2) = |u_1 - v_1| + |u_2 - v_2|^2.$$

Applying Eq. (14) to this function  $h_{IF}$ ,  $I(A)$  is given by:

$$\begin{aligned} I(A) &= \frac{1}{n} \sum_{x \in X} f(|\mu_A(x) - \mu_A(x)| + |1 - \mu_A(x) - v_A(x)|^2, \\ &\quad |\mu_A(x) - 1 + v_A(x)| + |v_A(x) - v_A(x)|^2) \\ &= \frac{1}{n} \sum_{x \in X} f(|1 - \mu_A(x) - v_A(x)|^2, |1 - \mu_A(x) - v_A(x)|) \\ &= \frac{1}{n} \sum_{x \in X} f(\pi_A(x)^2, \pi_A(x)). \end{aligned} \tag{15}$$

Consider the function  $f : [0, 1] \times [0, 1] \rightarrow [0, 1]$  given by  $f(x, y) = \frac{x+y}{2}$ , which satisfies (f1), (f2), (f5) and (f6). Substituting in Eq. (15), we obtain the following:

$$\begin{aligned} I(A) &= \frac{1}{n} \sum_{x \in X} \frac{\pi_A(x) + \pi_A(x)^2}{2} \\ &= \frac{1}{2n} \sum_{x \in X} \pi_A(x)(1 + \pi_A(x)). \end{aligned}$$

Now we show that this function is a BB-entropy:

(**IBB1**) :  $I(A) = 0$  if and only if  $\pi_A(x)(1 + \pi_A(x)) = 0$  for any  $x \in X$ , but this is equivalent to  $\pi(A) = 0$  for any  $x \in X$ , so  $A$  is a fuzzy set.

(**IBB2**) :  $I(A) = 1$  if and only if  $\frac{1}{2}\pi_A(x)(1 + \pi_A(x)) = 1$  for any  $x \in X$ , but this is equivalent to  $\pi_A(x) = 1$  for any  $x \in X$ , so  $\mu_A = v_A = 0$ .

(**IBB3**) : Trivially,  $I(A) = I(A^c)$  holds.

(**IBB4**) : Take  $\mu_A \leq \mu_B$  and  $v_A \leq v_B$ . This implies that  $\pi_A \geq \pi_B$ , and therefore  $\pi_A(1 + \pi_A) \geq \pi_B(1 + \pi_B)$ , which implies that  $I(A) \geq I(B)$ .

So  $I$  is a BB-entropy. However, the function  $h_{IF}$  does not satisfy property (AIF-loc.5): take  $(0.6, 0.1, 0.6, 0.4)$ ,  $(0.1, 0.6, 0.4, 0.6) \in \mathcal{D}_4$ . It holds that:

$$\begin{aligned} h_{IF}(0.6, 0.1, 0.6, 0.4) &= |0.6 - 0.6| + |0.1 - 0.4|^2 = 0.09. \\ h_{IF}(0.1, 0.6, 0.4, 0.6) &= |0.1 - 0.4| + |0.6 - 0.6|^2 = 0.3. \end{aligned}$$

Since both values do not coincide,  $h_{IF}$  does not satisfy property (AIF-loc.5).

We conclude that the sufficient conditions given in Proposition 4 are not necessary.

### 4.3 Examples of BB-entropies based on AIF-divergences

We consider again the four local AIF-divergences defined in Sect. 3.3, the Hamming and Hausdorff distances  $l_{AIF}$ ,  $d_H$  and the AIF-divergences defined by Hong and Kim  $D_C$ ,  $D_L$ . In order to make them to satisfy the normalization properties mentioned in Sect. 2.3, in this section we consider also  $d_H^*$ , which is defined by  $d_H^* = 2d_H$ . We first apply Theorem 4 to these four local AIF-divergences.

**Proposition 5** Consider the AIF-divergences  $l_{AIF}$ ,  $d_H^*$ ,  $D_C$  and  $D_L$ . They satisfy the conditions in Theorem 4, so each of them induces a BB-entropy. Also, the BB-entropies they induce coincide and it is given by:

$$I(A) = \frac{1}{n} \sum_{x \in X} \pi_A(x). \tag{16}$$

**Proof** First of all, let us see that  $h_l$ ,  $h_{d^*}$ ,  $h_C$  and  $h_L$  coincide in the domain  $\mathcal{D}_3$ :

$$\begin{aligned} h_l \left( u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2} \right) &= \frac{1}{2} \left( \left| \frac{1-u-v}{2} \right| + \left| \frac{1-u-v}{2} \right| + |1-u-v| \right) \\ &= 1-u-v. \\ h_{d^*} \left( u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2} \right) &= 2 \max \left\{ \frac{1-u-v}{2}, \frac{1-u-v}{2} \right\} = 1-u-v. \\ h_C \left( u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2} \right) &= \left| \frac{1-u-v}{2} \right| + \left| \frac{1-u-v}{2} \right| = 1-u-v. \\ h_L \left( u, v, \frac{1+u-v}{2}, \frac{1+v-u}{2} \right) &= \frac{1}{2} \left( \left| u - \frac{1+u-v}{2} - v + \frac{1-u+v}{2} \right| \right. \\ &\quad \left. + \left| \frac{1-u-v}{2} \right| + \left| \frac{1-u-v}{2} \right| \right) = 1-u-v. \end{aligned}$$

Thus, we see that all  $h_l$ ,  $h_{d^*}$ ,  $h_C$ ,  $h_L$  coincide in  $\mathcal{D}_3$ . Now, let us see that they satisfy the conditions on Theorem 4.



(AIF - loc.1'') : It holds that:

$$h_l\left(u, v, \frac{1+u-v}{2}, \frac{1-u+v}{2}\right) = 1-u-v = 0$$

if and only if  $u + v = 1$ .

(AIF - loc.7) :  $h_l^*(u, v) = 1 - u - v$ , so obviously  $h_l^*$  is symmetric and decreasing in both  $u, v$ .

(AIF - loc.8) : It holds that:

$$h_l\left(u, v, \frac{1+u-v}{2}, \frac{1-u+v}{2}\right) = 1-u-v = 1$$

if and only if  $u = v = 0$ .

Therefore,  $h_l, h_{d^*}, h_C, h_L$ , all satisfy the conditions of Theorem 4, so each of Hamming and Hausdorff distances,  $D_C$ , and  $D_L$ , induces a BB-entropy measure. Furthermore, since  $h_l, h_{d^*}, h_C, h_L$  coincide in  $\mathcal{D}_3$ , all of them induce the same BB-entropy. Using Eq. (13), we obtain the following formula:

$$\begin{aligned} I(A) &= \frac{1}{n} \sum_{x \in X} h(\mu_A(x), \nu_A(x), \mu_{A^*}(x), 1 - \mu_{A^*}(x)) \\ &= \frac{1}{n} \sum_{x \in X} \pi_A(x). \end{aligned}$$

□

The BB-entropy obtained in the previous proposition has already been proposed in Burrillo and Bustince (1996) and used in other papers, such as in Szmidt et al. (2014).

Let us now apply the procedure described in Proposition 5 to the four local AIF-divergences. As the next result shows, the four measures induce again the same BB-entropy.

**Proposition 6** Consider the four local AIF-divergences  $l_{AIF}, d_H, D_C, D_L$  and a function  $f : [0, 1] \times [0, 1] \rightarrow [0, 1]$  satisfying properties (f1), (f5) and (f6). Then, these four local AIF-divergences satisfy the conditions of Proposition 5, so each of them induces a BB-entropy. Indeed, they induce the same BB-entropy, given by:

$$I(A) = \frac{1}{n} \sum_{x \in X} f(\pi_A(x), \pi_A(x)).$$

**Proof** First of all, let us see that  $h_l, h_{d^*}, h_C, h_L$  coincide in  $\mathcal{D}_4$  and that they take the value  $1 - u - v$ .

$$\begin{aligned} h_l(u, v, u, 1-u) &= \frac{1}{2} (|u-u| + |1-u-v| + |1-u-v|) \\ &= 1-u-v. \end{aligned}$$

$$h_{d^*}(u, v, u, 1-u) = \max\{|u-u|, |1-u-v|\}$$

$$= 1-u-v.$$

$$h_C(u, v, u, 1-u) = |u-u| + |1-u-v| = 1-u-v.$$

$$h_L(u, v, u, 1-u)$$

$$= \frac{1}{2} (|u-u-(1-u)+v| + |u-u| + |1-u-v|)$$

$$= 1-u-v.$$

Similarly, it can be seen that:

$$h_l(u, v, 1-v, v) = h_{d^*}(u, v, 1-v, v)$$

$$= h_C(u, v, 1-v, v)$$

$$= h_L(u, v, 1-v, v) = 1-u-v.$$

Next we show that they satisfy the conditions on Proposition 5:

(AIF-loc.1''') :  $h_l(u, v, u, 1-u) = h_l(u, v, 1-v, v) = 1-u-v = 0$  if and only if  $u + v = 1$ .

(AIF-loc.5) :  $h_l$  is trivially symmetric in the domain  $\mathcal{D}_4$  since both  $h_l(u, v, u, 1-u)$  and  $h_l(u, v, 1-v, v)$  coincide and take the value  $1-u-v$ .

(AIF-loc.9) : Again, since

$$h_l(u, v, u, 1-u) = h_l(u, v, 1-v, v) = 1-u-v,$$

so it is decreasing in both  $u$  and  $v$ .

(AIF-loc.10) :  $h_l(u, v, u, 1-u) = h_l(u, v, 1-v, v) = 1-u-v = 1$  if and only if  $u = v = 0$ .

We can see that  $h_l$ , and in a similar manner also  $h_{d^*}, h_C, h_L$ , satisfy the properties of Proposition 5, and therefore they define a BB-entropy which is given by:

$$\begin{aligned} I(A) &= \frac{1}{n} \sum_{x \in X} f(h_l(\mu_A(x), \nu_A(x), \mu_A(x), 1 - \mu_A(x)), \\ &\quad h_l(\mu_A(x), \nu_A(x), 1 - \nu_A(x), \nu_A(x))) \\ &= \frac{1}{n} \sum_{x \in X} f(\pi_A(x), \pi_A(x)). \end{aligned}$$

□

In particular, now consider the functions:

$$f_1(u, v) = \sqrt{u \cdot v}, \quad f_2(u, v) = 1 - (1 - \sqrt{u \cdot v})^k,$$

$$f_3(u, v) = \frac{1}{e} \sqrt{u \cdot v} \cdot e^{\sqrt{u \cdot v}},$$

where  $k$  is an integer. It is easy to see that all three satisfy (f1), (f2), (f5) and (f6). So we obtain the following BB-entropies:

$$I_1(A) = \frac{1}{n} \sum_{x \in X} \pi_A(x).$$

$$I_2(A) = \frac{1}{n} \sum_{x \in X} (1 - (1 - \pi_A(x))^k).$$

$$I_3(A) = \frac{1}{en} \sum_{x \in X} \pi_A(x) e^{\pi_A(x)}.$$

The former has already been obtained in Proposition 6, while the second and the third were already presented as examples of BB-entropies in Burrillo and Bustince (1996).

## 5 Knowledge measures

In a recent paper Guo (2016), the notion of *knowledge measure* was introduced by Guo. The aim of this measure is to quantify the amount of knowledge conveyed by an AIFS.

**Definition 7** (Guo (2016)) A mapping  $K : \text{AIFS}(X) \rightarrow [0, 1]$  is called a knowledge measure if  $K$  has the following properties:

- (KP1)  $K(A) = 1$  if and only if  $A$  is crisp.
- (KP2)  $K(A) = 0$  if and only if  $\pi_A = 1$ .
- (KP3)  $K(A^c) = K(A)$ .
- (KP4)  $K(A) \geq K(B)$  if  $\mu_A(x) \leq \mu_B(x) < \nu_B(x) \leq \nu_A(x)$  or  $\nu_A(x) \leq \nu_B(x) \leq \mu_B(x) \leq \mu_A(x)$  for any  $x \in X$ .

We notice that a knowledge measure is nothing but the complement of an SK-entropy:  $K$  is a knowledge measure if and only if  $E = 1 - K$  is a SK-entropy. In this way, we can simply adapt our results from Sect. 3 to build knowledge measures using local AIF-divergences. For instance, we can easily write Theorem 3 in terms of knowledge measures:

**Corollary 1** Consider a local AIF-divergence  $D$  with associated function  $h_{\text{IF}}$ , and define the function  $K$  by:

$$K(A) = 1 - D(A, C_A) \\ = 1 - \frac{1}{n} \sum_{x \in X} h_{\text{IF}}(\mu_A(x), \nu_A(x), \mu_{C_A}(x), \nu_{C_A}(x))$$

for any  $A \in \text{AIFS}(X)$ . Then,  $K$  is a knowledge measure if and only if  $h_{\text{IF}}$  satisfies the conditions (AIF-loc.1'), (AIF-loc.5) and (AIF-loc.6).

The proof is analogous to that of Theorem 3 and therefore, omitted. Also, we can adapt other results from this section, as well as the examples given in Sect. 3.3, which give rise to the following knowledge measures:

$$K_1(A) = \frac{1}{n} \sum_{x \in X} \max\{\mu_A(x), \nu_A(x)\}. \\ K_2(A) = 1 - \frac{1}{n} \sum_{x \in X} (1 - |\mu_A(x) - \nu_A(x)|) \left( \frac{2 - \mu_A(x) - \nu_A(x)}{2} \right).$$

$$K_3(A) = 1 - \frac{1}{n} \sum_{x \in X} \left( (1 - \mu_A(x))(1 - \nu_A(x)) + \min\{\mu_A(x), \nu_A(x)\} \right).$$

$K_1$  is a knowledge measure that can be built using the Hamming and Hausdorff distances.  $K_2$  is a knowledge measure already mentioned in Guo (2016) and Guo and Song (2014), while  $K_3$  is another knowledge measure that can be built using the Hamming distance for fuzzy sets as we did in Example 3.

## 6 Application to multi-attribute group decision making

In this section, we present an application of our results in a multi-attribute decision-making (MADM, for short) problem. In detail, we continue with the approach given in (Nguyen 2015, Section 6), where knowledge measures were used to obtain the weights of the experts.

We first introduce the main notations. In MADM problems,  $X = \{x_1, \dots, x_n\}$  denotes a set of alternatives and  $A = \{a_1, \dots, a_m\}$  are the attributes with a weight vector  $w = (w_1, \dots, w_m)$ . The alternatives are evaluated by a set of experts  $\{e_1, \dots, e_l\}$ . Their evaluations are given as IFSs in a matrix form:  $S^{(k)}$ , which is an  $n \times m$  matrix such that  $s_{i,j}^{(k)} = \langle x_i, \mu_{i,j}^{(k)}, \nu_{i,j}^{(k)} \rangle$  denotes the AIFS that represents the evaluation of the expert  $e_k$  of the alternative  $x_i$  on the attribute  $a_j$ .

In Nguyen (2015), knowledge measures are used to determine weight vector for the experts, following these steps:

- Step 1 For each expert  $e_k$ , we compute the individual overall evaluation values of alternative  $x_i$  by using the following intuitionistic fuzzy weighted averaging operator Xu and Cai (2010):

$$z_i^{(k)} = \left\langle x_i, 1 - \prod_{j=1}^m (1 - \mu_{i,j}^{(k)})^{w_j}, \prod_{j=1}^m (\nu_{i,j}^{(k)})^{w_j} \right\rangle,$$

- Step 2 For each expert  $e_k$ , we compute the knowledge of its overall evaluation  $z_i$ , denoted by  $K(z^{(k)})$ .
- Step 3 We define the weights of the experts by normalizing the values  $K(z^{(i)})$ :

$$\lambda_k = \frac{K(z^{(k)})}{\sum_{i=1}^l K(z^{(i)})}.$$

From our comments in Sect. 5, we deduce that this approach is made in terms of SK-entropies, which means that it determines the weights of experts by measuring the lack of information of any expert about whether the alternative

is adequate or not. Then, the smaller is the lack of information, the greater is the weight. However, as we have explained before, we could also use BB-entropies instead of SK-entropies to measure the indecision of the experts.

Hence, we propose to modify the previous procedure as follows:

- Step 2\* Let us fix a local AIF-divergence and the SK- and BB-entropies it defines, denoted by  $E_{SK}$  and  $E_{BB}$ . For each expert  $e_k$ , we compute the SK- and BB-entropy of its overall evaluation:  $E_{SK}(z^{(k)})$  and  $E_{BB}(z^{(k)})$ .
- Step 3\* We define the weight of each expert by normalizing the entropies:

$$\alpha_k = \frac{1 - E_{SK}(z^{(k)})}{\sum_{i=1}^l (1 - E_{SK}(z^{(i)}))},$$

$$\beta_k = \frac{1 - E_{BB}(z^{(k)})}{\sum_{i=1}^l (1 - E_{BB}(z^{(i)}))}.$$

Once that we have obtained these values, we can proceed as follows:

1. The weights  $\alpha_k$  are computed by measuring the lack of information of any expert about whether the alternative is adequate or not. Thus, the weight decreases as the lack of information decreases.
2. The weights  $\beta_k$  are computed by measuring the determination of the experts, in the sense that more the indeterminacy of an expert, the closer  $s_{i,j}^{(k)}$  is to be a crisp set, and so the greater is the expert's weight.

Following the first interpretation and taking into account our comments in Sect. 5, our framework includes the approach of Nguyen (2015) as a particular case. We next apply this approach in the following example, which first appeared in (Nguyen 2015, Example 4).

**Example 6** Consider an MADM problem that consists of choosing an air-conditioning system between three alternatives  $x_1, x_2, x_3$ . In order to make the decision, five attributes are analyzed: good quality ( $a_1$ ), easiness to operate ( $a_2$ ), being economical ( $a_3$ ), good service ( $a_4$ ) and price ( $a_5$ ), where their weight vector is  $w = (0.2, 0.299, 0.106, 0.156, 0.239)$ . Three experts  $e_1, e_2, e_3$  evaluate the alternatives, and they give the following IFSs:

$$S^{(1)} = \begin{pmatrix} \langle 0.8, 0.1 \rangle & \langle 0.7, 0.1 \rangle & \langle 0.7, 0.2 \rangle & \langle 0.9, 0 \rangle & \langle 0.5, 0.4 \rangle \\ \langle 0.7, 0.1 \rangle & \langle 0.8, 0.2 \rangle & \langle 0.6, 0.4 \rangle & \langle 0.7, 0.1 \rangle & \langle 0.4, 0.6 \rangle \\ \langle 0.8, 0.2 \rangle & \langle 0.9, 0.1 \rangle & \langle 0.7, 0 \rangle & \langle 0.7, 0.2 \rangle & \langle 0.5, 0.5 \rangle \end{pmatrix}$$

$$S^{(2)} = \begin{pmatrix} \langle 0.9, 0.1 \rangle & \langle 0.8, 0.1 \rangle & \langle 0.7, 0 \rangle & \langle 0.9, 0.1 \rangle & \langle 0.7, 0.3 \rangle \\ \langle 0.7, 0.2 \rangle & \langle 0.8, 0.1 \rangle & \langle 0.9, 0.1 \rangle & \langle 0.7, 0.3 \rangle & \langle 0.7, 0.2 \rangle \\ \langle 0.7, 0.1 \rangle & \langle 0.9, 0 \rangle & \langle 0.8, 0 \rangle & \langle 0.8, 0.2 \rangle & \langle 0.3, 0.6 \rangle \end{pmatrix}$$

$$S^{(3)} = \begin{pmatrix} \langle 0.8, 0 \rangle & \langle 0.7, 0.1 \rangle & \langle 0.9, 0 \rangle & \langle 0.8, 0.1 \rangle & \langle 0.6, 0.4 \rangle \\ \langle 0.8, 0.2 \rangle & \langle 0.7, 0.3 \rangle & \langle 0.8, 0.1 \rangle & \langle 0.9, 0.1 \rangle & \langle 0.3, 0.6 \rangle \\ \langle 0.9, 0.1 \rangle & \langle 0.8, 0 \rangle & \langle 0.8, 0.1 \rangle & \langle 0.9, 0 \rangle & \langle 0.4, 0.5 \rangle \end{pmatrix}$$

Using Step 2, the individual overall evaluation of the experts is given by:

$$z_1 = \langle (x_1, 0.737, 0), (x_2, 0.677, 0.219), (x_3, 0.775, 0) \rangle.$$

$$z_2 = \langle (x_1, 0.82, 0), (x_2, 0.701, 0.217), (x_3, 0.7625, 0) \rangle.$$

$$z_3 = \langle (x_1, 0.752, 0), (x_2, 0.727, 0.245), (x_3, 0.797, 0) \rangle.$$

Now, consider the Hamming distance  $l_{AIF}$  and the SK- and BB-entropies they induce, denoted by  $E_{SK}$  and  $E_{BB}$ , which are computed using Eqs. (11) and (16). Then, following Step 2\* we obtain the following values:

	$z_1$	$z_2$	$z_3$
$E_{SK}(z_i)$	0.271	0.239	0.241
$E_{BB}(z_i)$	0.204	0.167	0.160

Thus, following Step 3\*, we obtain the final weight vectors:

$$\alpha = (0.324, 0.339, 0.337), \quad \beta = (0.322, 0.338, 0.340).$$

Let us compare the two weight vectors. On the one hand, the weight vector  $\alpha$  is given in terms of the amount of knowledge of the experts. This means that expert  $e_2$  has a slightly greater weight than  $e_3$ , and the least informative expert is  $e_1$ . On the other hand, the weight vector  $\beta$  is given in terms of the determination of each expert. With respect to this second approach,  $e_3$  is the most determined expert, so she has a slightly greater weight than  $e_2$  and  $e_1$ .

Finally, if we want to take into account the amount of information that experts have, we should use the weights  $\alpha_k$ , while if we want to take into account the determination of the experts, we should use the weights  $\beta_k$ . If we want to take into account both standpoints, we could aggregate both weights by means of any appropriate aggregation operator. In any case, we have to notice that both weights come from the same divergence measure, so they have a common starting point and they are related in all the cases.

As soon as we determine the weights, we can use the usual procedures from MADM to determine which is the most adequate alternative.

## 7 Conclusions

In the framework of AIFSs, two different ways of defining entropies can be found in the literature. Szmidt and Kacprzyk

defined entropy as a measure of how far an AIFS is from a fuzzy set, while Burrillo and Bustince defined entropy as a measure of how far an AIFS is from its closest fuzzy set.

In this work, we have generalized both approaches using local AIF-divergence measures, which are functions that measure how different two AIFSs are, to define entropies of AIFSs. In the framework of Szmidt and Kacprzyk, we have defined the closest crisp set to an AIFS and then we have defined the SK-entropy as the AIF-divergence between an AIFS and its closest crisp set. In the framework of Bustince and Burrillo, we have defined its closest fuzzy set, and then we have defined the BB-entropy as the AIF-divergence between the AIFS and its closest fuzzy set.

In both approaches, we have studied the properties that must be imposed in the AIF-divergence to guarantee that they define either an SK-entropy or a BB-entropy. We have also seen that the usual examples of SK- and BB-entropies can be obtained using local AIF-divergences. Measures of entropy have many applications in areas like image segmentation and multi-attribute decision making. Our generalizations offer a few distinct benefits. First, depending on an application, a user can choose an appropriate measure from a large set of possibilities. Second, the local nature of divergence makes it a trivial task to parallelize the computation of entropy, which is very important for large data sets. Third, such a measure being the divergence between two sets (e.g., an AIFS and its closest crisp set), it is quite easy to understand how and why, for example, an image segmentation algorithm using entropy works. We have also seen that SK-entropies and knowledge measures are equivalent, so we can also apply our results to define knowledge measures using local AIF-divergences.

As a future research, we aim to apply entropies and knowledge measures defined from AIF-divergence measures in image processing, as was done in Bhandari et al. (1992) and Farnoosh et al. (2016), or to pattern recognition, as was done in Deng et al. (2015) and Meng and Chen (2016).

**Acknowledgements** A shorter version of this paper was published at EUSFLAT Conference (Montes et al. 2018). The current version has been updated with additional results, comments and proofs. We acknowledge the financial support by project TIN2014-59543-P.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Atanassov K (1986) Intuitionistic fuzzy sets. *Fuzzy Sets Syst* 20:87–96
- Bhandari D, Pal N (1993) Some new information measures for fuzzy sets. *Inf Sci* 67(3):209–228
- Bhandari D, Pal N, Majumder D (1992) Fuzzy divergence, probability measure of fuzzy events and image thresholding. *Pattern Recogn Lett* 13(12):857–867
- Burrillo P, Bustince H (1996) Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets Syst* 78:305–316
- De Luca A, Termini S (1972) A definition of nonprobabilistic entropy in the setting of fuzzy theory. *J Gen Syst* 5:301–312
- Deng G, Jiang Y, Fu J (2015) Monotonic similarity measures between intuitionistic fuzzy sets and their relationship with entropy and inclusion measure. *Inf Sci* 316:348–369
- Dubois D, Prade H (eds.) (2000) *Fundamentals of Fuzzy Sets*. Springer
- Farnoosh R, Rahimi M, Kumar P (2016) Removing noise in a digital image using a new entropy method based on intuitionistic fuzzy sets. In: 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
- Grzegorzewski P (2004) Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric. *Fuzzy Sets Syst* 148:319–328
- Guo K (2016) Knowledge measure for Atanassov's intuitionistic fuzzy sets. *IEEE Trans Fuzzy Syst* 24(5):1072–1078
- Guo K, Song Q (2014) On the entropy for Atanassov's intuitionistic fuzzy sets: an interpretation from the perspective of amount of knowledge. *Appl Soft Comput* 24:328–340
- Hong D, Kim C (1999) A note on similarity measures between vague sets and between elements. *Inf Sci* 115:83–96
- Hung W, Yang M (2004) Similarity measures of intuitionistic fuzzy sets based on Hausdorff distances. *Pattern Recogn Lett* 25:1603–1611
- Joshi R, Kumar S (2017) Parametric  $(r, s)$ -norm entropy on intuitionistic fuzzy sets with a new approach in multi attribute decision making. *Fuzzy Inf Eng* 9(2):181–203
- Kacprzyk J, Pedrycz, W (eds.) (2015) *Springer Handbook of Computational Intelligence*. Springer
- Klement E, Mesiar R, Pap E (2000) *Triangular norms*. Kluwer Academic Publishers, Dordrecht
- Kosko B (1986) Fuzzy entropy and conditioning. *Inf Sci* 40:166–177
- Liang Z, Shi P (2003) Similarity measures on intuitionistic fuzzy sets. *Pattern Recogn Lett* 24:2687–2693
- Liu X (1992) Entropy, distance measure and similarity measure of fuzzy sets and their relation. *Fuzzy Sets Syst* 53:305–318
- Melo-Pinto P, Couto P, Bustince H, Barrenechea E, Pagola M, Fernandez J (2013) Image segmentation using Atanassov's intuitionistic fuzzy sets. *Expert Syst Appl* 40(1):15–26
- Meng F, Chen X (2016) Entropy and similarity measure of atanassov's intuitionistic fuzzy sets and their application to pattern recognition based on fuzzy measures. *Pattern Analysis and Applications* 19(1):11–20. <https://doi.org/10.1007/s10044-014-0378-6>
- Montes I, Janis V, Montes S (2011) An axiomatic definition of divergence for intuitionistic fuzzy sets. *Proc EUSFLAT Conf* 1:547–553
- Montes I, Janis V, Montes S (2012) Local IF-divergences. *Commun Comput Inf Sci* 298:491–500
- Montes I, Janis V, Pal N, Montes S (2016) Local divergences for atanassov intuitionistic fuzzy sets. *IEEE Trans Fuzzy Syst* 24(2):360–373
- Montes I, Montes S, Pal N (2018) On the use of divergences for defining entropies for Atanassov intuitionistic fuzzy sets. *Adv Intell Syst Comput* 642:554–565
- Montes I, Pal N, Janis V, Montes S (2015) Divergence measures for intuitionistic fuzzy sets. *IEEE Trans Fuzzy Syst* 23(2):444–456

- Montes S, Couso I, Bertoluzza C (1998) Some classes of fuzziness measures from local divergences. *Belg J Oper Res Stat Comput Sci* 38(2–3):37–49
- Montes S, Couso I, Gil P, Bertoluzza C (2002) Divergence measure between fuzzy sets. *Int J Approx Reason* 30(2):91–105
- Nguyen H (2015) A new knowledge-based measure for intuitionistic fuzzy sets and its application in multiple attribute group decision making. *Expert Syst Appl* 42:8766–8774
- Pal N, Bejdek J (1994) Measuring fuzzy uncertainty. *IEEE Trans Fuzzy Syst* 2(2):107–118
- Pal N, Bustince H, Pagola M, Mukherjee U, Goswami D, Beliakov G (2013) Uncertainty with Atanassov intuitionistic fuzzy sets: fuzziness and lack of knowledge. *Inf Sci* 228:61–74
- Szmidt E, Kacprzyk J (2000) Distances between intuitionistic fuzzy sets. *Fuzzy Sets Syst* 114:505–518
- Szmidt E, Kacprzyk J (2001) Entropy for intuitionistic fuzzy sets. *Fuzzy Sets Syst* 118:467–477
- Szmidt E, Kacprzyk J (2006) An application of intuitionistic fuzzy set similarity measures to a multi-criteria decision making problem. In: *Artificial Intelligence and Soft Computing ICAISC. Lecture Notes in Computer Science*, vol. 4029, pp. 314–323. Springer
- Szmidt E, Kacprzyk J, Bujnowski P (2014) How to measure the amount of knowledge conveyed by Atanassov's intuitionistic fuzzy sets. *Inf Sci* 257:276–285
- Torra V (2010) Hesitant fuzzy sets. *Int J Intell Syst* 25:529–539
- Trillas E, Riera T (1978) Entropies in finite fuzzy sets. *Inf Sci* 15:159–168
- Xu Z (2007) Some similarity measures of intuitionistic fuzzy sets and their application to multiple attribute decision making. *Fuzzy Optim Decis Making* 6:109–121
- Xu Z, Cai X (2010) Nonlinear optimization models for multiple attribute group decision making with intuitionistic fuzzy information. *Int J Intell Syst* 25(6):489–513
- Yager R (1982) On measure of fuzziness and fuzzy complements. *Int J Gen Syst* 8:169–180
- Zadeh L (1965) Fuzzy sets. *Inf Control* 8:338–353
- Zadeh L (1975) The concept of a linguistic variable and its application to approximate reasoning. *Inf Sci* 8:199–249
- Zimmermann HJ (2001) *Fuzzy set theory and its applications*. Springer

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.