

Article

R-Norm Entropy and R-Norm Divergence in Fuzzy Probability Spaces

Dagmar Markechová ^{1,*}, Batool Mosapour ² and Abolfazl Ebrahimzadeh ³
¹ Department of Mathematics, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, A. Hlinku 1, Nitra SK-949 01, Slovakia

² Department of Mathematics, Farhangian University, Kerman +98-7616914111, Iran; mosapourbatool@gmail.com

³ Young Researchers and Elite Club, Zahedan Branch, Islamic Azad University, Zahedan +98-9816883673, Iran; a.ebrahimzade@iauzah.ac.ir

* Correspondence: dmarkechova@ukf.sk; Tel.: +421-376-408-111; Fax: +421-376-408-020

Received: 12 March 2018; Accepted: 9 April 2018; Published: 11 April 2018



Abstract: In the presented article, we define the R -norm entropy and the conditional R -norm entropy of partitions of a given fuzzy probability space and study the properties of the suggested entropy measures. In addition, we introduce the concept of R -norm divergence of fuzzy P -measures and we derive fundamental properties of this quantity. Specifically, it is shown that the Shannon entropy and the conditional Shannon entropy of fuzzy partitions can be derived from the R -norm entropy and conditional R -norm entropy of fuzzy partitions, respectively, as the limiting cases for R going to 1; the Kullback–Leibler divergence of fuzzy P -measures may be inferred from the R -norm divergence of fuzzy P -measures as the limiting case for R going to 1. We also provide numerical examples that illustrate the results.

Keywords: fuzzy measurable space; fuzzy P -measure; fuzzy probability space; fuzzy partition; R -norm entropy; conditional R -norm entropy; R -norm divergence

1. Introduction

The concept of information entropy was introduced by Claude Shannon in 1948 in his article [1]. It is used in information theory [2] to quantify the amount of information or uncertainty inherent in a system. We remind that Shannon's entropy is defined in the context of a probabilistic model. Consider a measurable partition $\mathcal{A} = \{E_1, E_2, \dots, E_n\}$ of a probability space (X, S, P) (that is a finite collection of measurable subsets $E_k \subset X, k = 1, 2, \dots, n$, such that $\cup_{k=1}^n E_k = X$, and $E_i \cap E_j = \emptyset$ whenever $i \neq j$) with probabilities $p_k = P(E_k), k = 1, 2, \dots, n$. Then the Shannon entropy of \mathcal{A} is defined as the number $h(\mathcal{A}) = -\sum_{k=1}^n p_k \cdot \log_b p_k$ with the convention that $0 \cdot \log_b 0 = 0$ (which is justified by the fact that $\lim_{p \rightarrow 0} p \cdot \log_b \frac{1}{p} = 0$). The base of the logarithm can be any positive real number; depending on the selected base of the logarithm, the entropy is expressed in bits ($b = 2$), nats ($b = e$), or dits ($b = 10$).

The extensions of Shannon's entropy have led to several alternatives of entropy measure, of which the Rényi entropy [3] is one of the most important. The classical logical entropy (cf. [4,5]) and the entropy measure called the R -norm entropy (cf. [6,7]) are other alternative entropy measures. In this article, we will deal with the study of the R -norm entropy. If $P = \{p_1, p_2, \dots, p_n\}$ is a probability distribution, then the R -norm entropy is defined, for every real number $R \in (0, 1) \cup (1, \infty)$, by the formula:

$$H_R(P) = \frac{R}{R-1} \left(1 - \left[\sum_{k=1}^n p_k^R \right]^{\frac{1}{R}} \right).$$

Some results regarding the R -norm entropy measure and its generalizations can be found in [8–12]. The above entropy measures have found many important applications, for example, in statistics, pattern recognition, and coding theory.

In classical probability theory, partitions are defined in the context of the Cantor set theory. In solving many real-life problems, however, the partitions defined in terms of fuzzy set theory [13] are more appropriate. Therefore, many proposals have been made to generalize the classical partitions into fuzzy partitions [14–20]. Fuzzy partitions represent a mathematical tool for modeling random experiments that lead to unclear, vague events. Naturally, there are also many results concerning the Shannon entropy of fuzzy partitions; see e.g., [21–31]. We note that in [32] the results regarding the entropy of fuzzy partitions provided in [25] have been employed to introduce the notions of mutual information and Kullback–Leibler divergence for the fuzzy case. The notion of Kullback–Leibler divergence was introduced in [33] as the distance measure between two probability distributions. It plays significant roles in information theory and various disciplines such as statistics, machine learning, physics, neuroscience, computer science, linguistics, etc.

Since its inception in 1965, the theory of fuzzy sets has advanced in many mathematical disciplines and has found important applications in practice. Currently, the subjects of intense study are algebraic systems based on the theory of fuzzy sets, for example, D-posets [34–36], MV-algebras [37–41], and effect algebras [42]. Some results regarding the above entropy measures and divergence on these structures can be found e.g., in [43–58].

The aim of this article is to study the R -norm entropy of fuzzy partitions and the R -norm divergence in fuzzy probability spaces [59]. The organization of the paper is as follows. In Section 2 we provide basic definitions, terminology and some known results used in the paper. The results of the article are presented in Sections 3 and 4. In Section 3, we define the R -norm entropy and conditional R -norm entropy of fuzzy partitions and examine their properties. In Section 4, the concept of the R -norm divergence for the case of fuzzy probability spaces is proposed and the properties of this distance measure are studied. The results presented in Sections 3 and 4 are illustrated with numerical examples. The paper concludes in Section 5 with a brief summary.

2. Preliminaries

We begin by recalling the basic concepts and the known results used in the paper.

It is known that the concept of fuzzy set, introduced by Zadeh in 1965 [13], extends the classical set theory. In classical set theory, the membership of elements in a set is assessed in binary terms according to the condition—an element either belongs or does not belong to the considered set. By contrast, the fuzzy set is characterized by a membership function which assigns to every element a grade of membership ranging between zero and one. The mathematical model of the fuzzy set is as follows. Let X be a non-empty set. By a fuzzy subset of X we mean a mapping $a : X \rightarrow [0, 1]$ (where the considered fuzzy set is identified with its membership function). The value $a(x)$ is interpreted as a grade of membership of the element $x \in X$ to the considered fuzzy set a .

Definition 1. Let X be a non-empty set, and $M \subset [0, 1]^X$ be a family of fuzzy subsets of X . The pair (X, M) is called a fuzzy measurable space, if the following conditions are satisfied: (i) $1_X \in M$, $(1/2)_X \notin M$; (ii) $a \in M \Rightarrow a^\perp \in M$; (iii) if $a_n \in M$, $n = 1, 2, \dots$, then $\bigcup_{n=1}^\infty a_n \in M$. The family $M \subset [0, 1]^X$ with the properties (i)–(iii) is said to be a fuzzy σ -algebra.

Throughout the paper, the symbols $\bigcup_{n=1}^\infty a_n$ and $\bigcap_{n=1}^\infty a_n$ denote the fuzzy union and the fuzzy intersection of a sequence $\{a_n\}_{n=1}^\infty \subset M$, respectively, by Zadeh [13], i.e., $\bigcup_{n=1}^\infty a_n = \sup_n a_n$, and $\bigcap_{n=1}^\infty a_n = \inf_n a_n$. The symbol a^\perp denotes the complement of fuzzy set $a \in M$, i.e., $a^\perp = 1_X - a$. Here, 1_X denotes the constant function with the value 1; analogously, the symbols $(1/2)_X$ and 0_X denote the constant functions with the value $1/2$, and 0 , respectively. Additionally, the relation \leq denotes the usual order relation of fuzzy subsets of X , i.e., $a \leq b$ if and only if $a(x) \leq b(x)$, for every $x \in X$.

The complementation $\perp : a \rightarrow a^\perp$, $a \in M$, satisfies, for every $a, b \in M$, the conditions: (i) $(a^\perp)^\perp = a$, and (ii) $a \leq b$ implies $b^\perp \leq a^\perp$.

Fuzzy subsets $a, b \in M$ with the property $a \cap b = 0_X$ are said to be separated, fuzzy subsets $a, b \in M$ with the property $a \leq b^\perp$ are said to be W-separated fuzzy sets. Any fuzzy subset $a \in M$ with the property $a \geq a^\perp$ is said to be a W-universum, any fuzzy subset $a \in M$ with the property $a \leq a^\perp$ is said to be a W-empty fuzzy set. A fuzzy set from the fuzzy σ -algebra M is interpreted as a fuzzy event. W-separated fuzzy events are considered to be mutually exclusive events. A W-universum is interpreted as a certain event, a W-empty set as an impossible event. It can be shown that a fuzzy subset $a \in M$ is a W-universum if and only if there exists a fuzzy subset $b \in M$ such that $a = b \cup b^\perp$.

Naturally, the notion of a fuzzy measurable space generalizes the concept of a measurable space (X, S) from the classical measure theory; it suffices to put $M = \{\chi_E : E \in S\}$, where χ_E is the characteristic function of the set $E \in S$. With this procedure, the classical model can be inserted into the fuzzy one.

Definition 2 ([59]). Let (X, M) be a fuzzy measurable space. A map $\mu : M \rightarrow [0, 1]$ is said to be a fuzzy P-measure, the following conditions being satisfied: (i) $\mu(a \cup a^\perp) = 1$, for every $a \in M$; (ii) if $\{a_n\}_{n=1}^\infty$ is a sequence of pairwise W-separated fuzzy sets from M , then $\mu(\bigcup_{n=1}^\infty a_n) = \sum_{n=1}^\infty \mu(a_n)$. The triplet (X, M, μ) is called a fuzzy probability space.

The fuzzy P-measure $\mu : M \rightarrow [0, 1]$ has the properties that correspond to properties of a classical probability measure; the proofs can be found in [59].

- (P1) $\mu(a^\perp) = 1 - \mu(a)$, for every $a \in M$.
- (P2) μ is non-decreasing, i.e., if $a, b \in M$ with $a \leq b$, then $\mu(a) \leq \mu(b)$.
- (P3) $\mu(a \cup b) + \mu(a \cap b) = \mu(a) + \mu(b)$, for every $a, b \in M$.
- (P4) Let $b \in M$. Then $\mu(a \cap b) = \mu(a)$ for all $a \in M$ if and only if $\mu(b) = 1$.
- (P5) If $a, b \in M$ such that $a \leq b^\perp$, then $\mu(a \cap b) = 0$.
- (P6) If $a, b \in M$ such that $a \leq b$, then $\mu(b) = \mu(a) + \mu(a^\perp \cap b)$.

Definition 3 ([14]). A fuzzy partition of a fuzzy probability space (X, M, μ) is a collection $A = \{a_1, a_2, \dots, a_n\}$ of W-separated fuzzy sets from M with the property $\mu(\bigcup_{i=1}^n a_i) = 1$.

In the system of all fuzzy partitions of (X, M, μ) , we define the refinement partial order in the following way. If A and B are two fuzzy partitions of (X, M, μ) , then we say that B is a refinement of A (and write $A \prec B$), if for every $b \in B$ there exists $a \in A$ such that $b \leq a$. Furthermore, for every two fuzzy partitions $A = \{a_1, a_2, \dots, a_n\}$, and $B = \{b_1, b_2, \dots, b_m\}$ of (X, M, μ) , we put $A \vee B = \{a_i \cap b_j; i = 1, 2, \dots, n, j = 1, 2, \dots, m\}$. One can easily verify that the family $A \vee B$ is a family of pairwise W-separated fuzzy sets from M ; moreover, by the property (P4), we have $\mu(\bigcup_{i=1}^n \bigcup_{j=1}^m (a_i \cap b_j)) = \mu((\bigcup_{i=1}^n a_i) \cap (\bigcup_{j=1}^m b_j)) = \mu(\bigcup_{i=1}^n a_i) = 1$. Thus, $A \vee B$ is a fuzzy partition of (X, M, μ) . It represents a combined experiment consisting of a realization of the experiments A and B . Evidently, it holds $A \prec A \vee B$, and $B \prec A \vee B$, i.e., the fuzzy partition $A \vee B$ is a common refinement of fuzzy partitions A and B . If A_1, A_2, \dots, A_n are fuzzy partitions of (X, M, μ) , then we put $\bigvee_{i=1}^n A_i = A_1 \vee A_2 \vee \dots \vee A_n$.

Definition 4. Two fuzzy partitions $A = \{a_1, a_2, \dots, a_n\}$, and $B = \{b_1, b_2, \dots, b_m\}$ of a fuzzy probability space (X, M, μ) are said to be statistically independent, if $\mu(a_i \cap b_j) = \mu(a_i) \cdot \mu(b_j)$, for $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$.

Example 1. Let us consider a classical probability space (X, S, P) , and put $M = \{\chi_E : E \in S\}$. It can be verified that the map $\mu : M \rightarrow [0, 1]$ defined by $\mu(\chi_E) = P(E)$, for every $\chi_E \in M$, is a fuzzy P-measure and the triplet (X, M, μ) is a fuzzy probability space. A classical measurable partition $\mathcal{A} = \{E_1, E_2, \dots, E_n\}$ of a

probability space (X, S, P) can be eventually regarded as a fuzzy partition of (X, M, μ) , considering χ_{E_i} instead of E_i .

The Shannon entropy of fuzzy partition of a fuzzy probability space (X, M, μ) has been introduced and examined in [23], see also [25].

Definition 5 ([23]). We define the entropy of a fuzzy partition $A = \{a_1, a_2, \dots, a_n\}$ of (X, M, μ) by Shannon's formula:

$$H^\mu(A) = -\sum_{i=1}^n \mu(a_i) \cdot \log \mu(a_i). \quad (1)$$

If $A = \{a_1, a_2, \dots, a_n\}$, and $B = \{b_1, b_2, \dots, b_m\}$ are two fuzzy partitions of (X, M, μ) , then we define the conditional entropy of A given B by the formula:

$$H^\mu(A/B) = -\sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j) \cdot \log \frac{\mu(a_i \cap b_j)}{\mu(b_j)} \quad (2)$$

with the convention that $0 \cdot \log \frac{0}{x} = 0$ if $x \geq 0$.

The symbol \log denotes the base 2 logarithm, so the Shannon entropy of fuzzy partition is expressed in bits. The entropy and the conditional entropy of fuzzy partitions have properties that correspond to properties of Shannon's entropy of classical measurable partitions: for every fuzzy partitions A, B, C of a fuzzy probability space (X, M, μ) , it holds:

- (S1) $A \prec B$ implies $H^\mu(A) \leq H^\mu(B)$;
- (S2) $H^\mu(A \vee B/C) = H^\mu(A/C) + H^\mu(B/C \vee A)$;
- (S3) $A \prec B$ implies $H^\mu(A/C) \leq H^\mu(B/C)$;
- (S4) $A \prec B$ implies $H^\mu(C/A) \geq H^\mu(C/B)$;
- (S5) $H^\mu(A/B) \leq H^\mu(A)$ with the equality if and only if A, B are statistically independent;
- (S6) $H^\mu(B \vee C/A) \leq H^\mu(B/A) + H^\mu(C/A)$;
- (S7) $H^\mu(A \vee B) = H^\mu(A) + H^\mu(B/A)$.

The proofs can be found in [23,25]. We remark that in [15–22,26–31], other conceptions of fuzzy partitions and their entropy measures have been introduced. Whereas our approach is based on Zadeh's connectives, in the referenced papers Zadeh's connectives have been replaced by other fuzzy set operations.

We note that in [32], the concept of Kullback–Leibler divergence in the fuzzy probability space was introduced. Let μ, ν be two fuzzy P-measures on a fuzzy measurable space (X, M) , and $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, \nu)$. Then the Kullback–Leibler divergence of fuzzy P-measures μ, ν with respect to A is defined as the number:

$$D_A(\mu \parallel \nu) = \sum_{i=1}^n \mu(a_i) \cdot \log \frac{\mu(a_i)}{\nu(a_i)} \quad (3)$$

with the convention that $x \log \frac{x}{0} = \infty$ if $x > 0$, and $0 \cdot \log \frac{0}{x} = 0$ if $x \geq 0$.

In the following sections, we will use the following known Minkowski inequality: for non-negative real numbers $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$, it holds:

$$\left[\sum_{i=1}^n x_i^R \right]^{\frac{1}{R}} + \left[\sum_{i=1}^n y_i^R \right]^{\frac{1}{R}} \geq \left[\sum_{i=1}^n (x_i + y_i)^R \right]^{\frac{1}{R}}, \text{ for } R > 1,$$

and

$$\left[\sum_{i=1}^n x_i^R \right]^{\frac{1}{R}} + \left[\sum_{i=1}^n y_i^R \right]^{\frac{1}{R}} \leq \left[\sum_{i=1}^n (x_i + y_i)^R \right]^{\frac{1}{R}}, \text{ for } 0 < R < 1.$$

Furthermore, we will use the Jensen inequality which states that for a real convex function φ , real numbers x_1, x_2, \dots, x_m in its domain and non-negative real numbers $\alpha_1, \alpha_2, \dots, \alpha_m$ with $\sum_{j=1}^m \alpha_j = 1$, it holds:

$$\varphi \left(\sum_{j=1}^m \alpha_j x_j \right) \leq \sum_{j=1}^m \alpha_j \varphi(x_j),$$

and the inequality is reversed if φ is a real concave function. The equality holds if and only if $x_1 = x_2 = \dots = x_m$ or φ is linear.

In addition, we will use L'Hôpital's rule: for functions f and g that are differentiable on an open interval U except possibly at a point $a \in U$, if $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$, $g'(x) \neq 0$, for every x in U with $x \neq a$, and $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$ exists, then:

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.$$

3. The R-Norm Entropy of Fuzzy Partitions

In this part we define the R -norm entropy of a fuzzy partition and its conditional version and study the properties of these entropy measures. It is shown that as the limiting cases of the R -norm entropy and the conditional R -norm entropy of fuzzy partitions for R going to 1, we obtain the Shannon entropy $H^\mu(A)$, and the conditional Shannon entropy $H^\mu(A/B)$, respectively, expressed in nats.

Definition 6. Let $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of a fuzzy probability space (X, M, μ) . The R -norm entropy of A with respect to μ is defined, for a positive real number R not equal to 1, by the formula:

$$H_R^\mu(A) = \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right). \quad (4)$$

Remark 1. For simplicity, we write $\mu(a_i)^R$ instead of $(\mu(a_i))^R$. In the following, we will write $\log \sum_{i=1}^n \mu(a_i)^R$ instead of $\log \left(\sum_{i=1}^n \mu(a_i)^R \right)$.

Theorem 1. For arbitrary fuzzy partition A of a fuzzy probability space (X, M, μ) , the R -norm entropy $H_R^\mu(A)$ is non-negative.

Proof. Assume that $A = \{a_1, a_2, \dots, a_n\}$. We will consider two cases: the case of $R > 1$, and the case of $0 < R < 1$. If $R > 1$, then $\mu(a_i)^R \leq \mu(a_i)$, for $i = 1, 2, \dots, n$, hence $\sum_{i=1}^n \mu(a_i)^R \leq \sum_{i=1}^n \mu(a_i) = 1$. This implies that $\left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \leq 1$. Since $\frac{R}{R-1} > 0$ for $R > 1$, it follows that $H_R^\mu(A) = \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \geq 0$. On the other hand, for $0 < R < 1$, it holds that $\mu(a_i)^R \geq \mu(a_i)$, for $i = 1, 2, \dots, n$, hence $\sum_{i=1}^n \mu(a_i)^R \geq \sum_{i=1}^n \mu(a_i) = 1$. It follows that $\left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \geq 1$. Since $\frac{R}{R-1} < 0$ for $0 < R < 1$, we obtain $H_R^\mu(A) = \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \geq 0$. \square

Example 2. Let $X = [0, 1]$, and $a : X \rightarrow [0, 1]$ be defined by $a(x) = x$, $x \in X$. Consider a fuzzy measurable space (X, M) , where $M = \{1_X, 0_X, a, a^\perp, a \cup a^\perp, a \cap a^\perp\}$. Then it can be easily verified that the mappings $\mu : M \rightarrow [0, 1]$, and $\nu : M \rightarrow [0, 1]$ defined by the equalities

$\mu(1_X) = \mu(a \cup a^\perp) = 1$, $\mu(0_X) = \mu(a \cap a^\perp) = 0$, $\mu(a) = \mu(a^\perp) = \frac{1}{2}$, $v(1_X) = v(a \cup a^\perp) = 1$, $v(0_X) = v(a \cap a^\perp) = 0$, $v(a) = \frac{1}{3}$, $v(a^\perp) = \frac{2}{3}$, are fuzzy P-measures and the systems (X, M, μ) , (X, M, v) are fuzzy probability spaces. The sets $A = \{a, a^\perp\}$, $B = \{a \cup a^\perp\}$, $C = \{1_X\}$ are fuzzy partitions of (X, M, μ) , and (X, M, v) such that $C \prec B \prec A$. We can calculate their R-norm entropy. Evidently, $H_R^\mu(B) = H_R^\mu(C) = H_R^v(B) = H_R^v(C) = 0$; in accordance with the natural requirement, experiments resulting in a certain event have zero R-norm entropy. Furthermore, we have:

$$H_R^\mu(A) = \frac{R}{R-1} \left(1 - \left[\left(\frac{1}{2} \right)^R + \left(\frac{1}{2} \right)^R \right]^{\frac{1}{R}} \right) = \frac{R}{R-1} \left(1 - 2^{\frac{1-R}{R}} \right),$$

$$H_R^v(A) = \frac{R}{R-1} \left(1 - \left[\left(\frac{1}{3} \right)^R + \left(\frac{2}{3} \right)^R \right]^{\frac{1}{R}} \right).$$

If we put $R = \frac{1}{2}$, then $H_R^\mu(A) = 1$, $H_R^v(A) \doteq 0.9428$; for $R = \frac{1}{3}$, we have $H_R^\mu(A) = 1.5$, $H_R^v(A) \doteq 1.4236$; for $R = 2$, we have $H_R^\mu(A) \doteq 0.58579$, and $H_R^v(A) \doteq 0.50928$.

Definition 7. Let $A = \{a_1, a_2, \dots, a_n\}$, and $B = \{b_1, b_2, \dots, b_m\}$ be two fuzzy partitions of a fuzzy probability space (X, M, μ) . Then the conditional R-norm entropy of A given B with respect to μ is defined, for a positive real number R not equal to 1, by the formula:

$$H_R^\mu(A/B) = \frac{R}{R-1} \left(\left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} - \left[\sum_{j=1}^m \sum_{i=1}^n \mu(a_i \cap b_j)^R \right]^{\frac{1}{R}} \right). \quad (5)$$

Remark 2. Let A be a fuzzy partition of a given fuzzy probability space (X, M, μ) . Evidently, if we put $B = \{b\}$, where $b \in M$ is a W-universum, then $H_R^\mu(A/B) = H_R^\mu(A)$.

The following theorem shows the consistency of the conditional R-norm entropy $H_R^\mu(A/B)$, in the case of the limit of R going to 1, with the conditional Shannon entropy $H^\mu(A/B)$ defined by the formula (2), up to a positive multiplicative constant.

Theorem 2. Let $A = \{a_1, a_2, \dots, a_n\}$, and $B = \{b_1, b_2, \dots, b_m\}$ be two fuzzy partitions of a given fuzzy probability space (X, M, μ) . Then $\lim_{R \rightarrow 1} H_R^\mu(A/B) = c \cdot H^\mu(A/B)$, where $c = \frac{1}{\log e}$, and $H^\mu(A/B) = -\sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j) \cdot \log \frac{\mu(a_i \cap b_j)}{\mu(b_j)}$.

Proof. In the proof, we use L'Hôpital's rule $\lim_{R \rightarrow a} \frac{f(R)}{g(R)} = \lim_{R \rightarrow a} \frac{f'(R)}{g'(R)}$, where in this case $a = 1$. For every $R \in (0, 1) \cup (1, \infty)$, we can write:

$$H_R^\mu(A/B) = \frac{1}{1 - \frac{1}{R}} \left(\left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} - \left[\sum_{j=1}^m \sum_{i=1}^n \mu(a_i \cap b_j)^R \right]^{\frac{1}{R}} \right) = \frac{f(R)}{g(R)},$$

where f, g are continuous functions defined for $R \in (0, \infty)$ in the following way:

$$f(R) = \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} - \left[\sum_{j=1}^m \sum_{i=1}^n \mu(a_i \cap b_j)^R \right]^{\frac{1}{R}}, g(R) = 1 - \frac{1}{R}.$$

By continuity of the function g , we get $\lim_{R \rightarrow 1} g(R) = g(1) = 0$. Furthermore, by continuity of the function f , and by the property (P4) of fuzzy P-measure μ , we get $\lim_{R \rightarrow 1} f(R) = f(1) = \sum_{j=1}^m \mu(b_j) -$

$\sum_{j=1}^m \sum_{i=1}^n \mu(a_i \cap b_j) = 1 - \sum_{j=1}^m \mu((\cup_{i=1}^n a_i) \cap b_j) = 1 - \sum_{j=1}^m \mu(b_j) = 1 - 1 = 0$. Using L'Hôpital's rule, this implies:

$$\lim_{R \rightarrow 1} H_R^\mu(A/B) = \frac{\lim_{R \rightarrow 1} f'(R)}{\lim_{R \rightarrow 1} g'(R)}$$

under the assumption that the right-hand side exists. To find the derivative of the function $f(R)$ we use the identity $b^a = e^{a \ln b}$. Let us calculate:

$$\begin{aligned} \frac{d}{dR} f(R) &= e^{\frac{1}{R} \ln \sum_{j=1}^m \mu(b_j)^R} \cdot \left(-\frac{1}{R^2} \cdot \ln \sum_{j=1}^m \mu(b_j)^R + \frac{1}{R} \cdot \frac{1}{\sum_{j=1}^m \mu(b_j)^R} \sum_{j=1}^m \mu(b_j)^R \cdot \ln \mu(b_j) \right) \\ &- e^{\frac{1}{R} \ln \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R} \left(-\frac{1}{R^2} \cdot \ln \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R + \frac{1}{R} \cdot \frac{1}{\sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R} \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R \cdot \ln \mu(a_i \cap b_j) \right). \end{aligned}$$

Since $\lim_{R \rightarrow 1} g'(R) = \lim_{R \rightarrow 1} \frac{1}{R^2} = 1$, we obtain:

$$\begin{aligned} \lim_{R \rightarrow 1} H_R^\mu(A/B) &= \lim_{R \rightarrow 1} f'(R) = \sum_{j=1}^m \mu(b_j) \cdot \ln \mu(b_j) - \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j) \cdot \ln \mu(a_i \cap b_j) \\ &= \sum_{j=1}^m \sum_{i=1}^n \mu(a_i \cap b_j) \cdot \ln \mu(b_j) - \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j) \cdot \ln \mu(a_i \cap b_j) \\ &= - \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j) \cdot \left(\frac{\log \mu(a_i \cap b_j)}{\log e} - \frac{\log \mu(b_j)}{\log e} \right) = -\frac{1}{\log e} \sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j) \cdot \log \frac{\mu(a_i \cap b_j)}{\mu(b_j)}. \quad \square \end{aligned}$$

Theorem 3. Let $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of a fuzzy probability space (X, M, μ) . Then $\lim_{R \rightarrow 1} H_R^\mu(A) = c \cdot H^\mu(A)$, where $H^\mu(A) = -\sum_{i=1}^n \mu(a_i) \cdot \log \mu(a_i)$, and $c = \frac{1}{\log e}$.

Proof. The claim is a direct consequence of the previous theorem; it suffices to put $B = \{1_X\}$. \square

In the following, the properties of the R -norm entropy of fuzzy partitions are discussed.

Theorem 4. For arbitrary fuzzy partitions A, B and C of a fuzzy probability space (X, M, μ) , it holds that:

$$H_R^\mu(A \vee B/C) = H_R^\mu(A/C) + H_R^\mu(B/A \vee C). \quad (6)$$

Proof. Suppose that $A = \{a_1, a_2, \dots, a_n\}$, $B = \{b_1, b_2, \dots, b_m\}$, $C = \{c_1, c_2, \dots, c_r\}$. Let us calculate:

$$\begin{aligned} H_R^\mu(A \vee B/C) &= \frac{R}{R-1} \left(\left[\sum_{k=1}^r \mu(c_k)^R \right]^{\frac{1}{R}} - \left[\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r \mu(a_i \cap b_j \cap c_k)^R \right]^{\frac{1}{R}} \right) \\ &= \frac{R}{R-1} \left(\left[\sum_{k=1}^r \mu(c_k)^R \right]^{\frac{1}{R}} - \left[\sum_{i=1}^n \sum_{k=1}^r \mu(a_i \cap c_k)^R \right]^{\frac{1}{R}} \right) \\ &+ \frac{R}{R-1} \left(\left[\sum_{i=1}^n \sum_{k=1}^r \mu(a_i \cap c_k)^R \right]^{\frac{1}{R}} - \left[\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r \mu(a_i \cap b_j \cap c_k)^R \right]^{\frac{1}{R}} \right) = H_R^\mu(A/C) + H_R^\mu(B/A \vee C). \quad \square \end{aligned}$$

Theorem 5. For arbitrary fuzzy partitions A, B of a fuzzy probability space (X, M, μ) , it holds that:

$$H_R^\mu(A \vee B) = H_R^\mu(A) + H_R^\mu(B/A). \quad (7)$$

Proof. The claim is a direct consequence of the previous theorem; it suffices to put $C = \{1_X\}$. \square

In the following theorem, using the notion of conditional R -norm entropy of fuzzy partitions, chain rules for the R -norm entropy of fuzzy partitions are established.

Theorem 6. Let A_1, A_2, \dots, A_n and C be fuzzy partitions of a fuzzy probability space (X, M, μ) . Then, for $n = 2, 3, \dots$, the following equalities hold:

- (i) $H_R^\mu(A_1 \vee A_2 \vee \dots \vee A_n) = H_R^\mu(A_1) + \sum_{i=2}^n H_R^\mu(A_i / \bigvee_{k=1}^{i-1} A_k)$;
- (ii) $H_R^\mu(\bigvee_{i=1}^n A_i / C) = H_R^\mu(A_1 / C) + \sum_{i=2}^n H_R^\mu(A_i / (\bigvee_{k=1}^{i-1} A_k \vee C))$.

Proof. The proof can be done using mathematical induction and Theorems 4 and 5. \square

In the following we prove that the R -norm entropy $H_R^\mu(A)$ is a concave function on the class of all fuzzy P -measures on a given fuzzy measurable space (X, M) .

Proposition 1. Let μ, ν be two fuzzy P -measures on a given fuzzy measurable space (X, M) . Then, for every real number $\lambda \in [0, 1]$, the map $\lambda\mu + (1 - \lambda)\nu : X \rightarrow [0, 1]$ is a fuzzy P -measure on (X, M) .

Proof. It is straightforward. \square

Theorem 7. Let A be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, \nu)$. Then, for every real number $\lambda \in [0, 1]$, this inequality holds:

$$\lambda H_R^\mu(A) + (1 - \lambda) H_R^\nu(A) \leq H_R^{\lambda\mu + (1-\lambda)\nu}(A).$$

Proof. Let $A = \{a_1, a_2, \dots, a_n\}$, and $\lambda \in [0, 1]$. Putting $x_i = \lambda\mu(a_i)$, and $y_i = (1 - \lambda)\nu(a_i)$, for $i = 1, 2, \dots, n$, in the Minkowski inequality, we obtain for $R > 1$:

$$\lambda \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} + (1 - \lambda) \left[\sum_{i=1}^n \nu(a_i)^R \right]^{\frac{1}{R}} \geq \left[\sum_{i=1}^n (\lambda\mu(a_i) + (1 - \lambda)\nu(a_i))^R \right]^{\frac{1}{R}},$$

and for $0 < R < 1$:

$$\lambda \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} + (1 - \lambda) \left[\sum_{i=1}^n \nu(a_i)^R \right]^{\frac{1}{R}} \leq \left[\sum_{i=1}^n (\lambda\mu(a_i) + (1 - \lambda)\nu(a_i))^R \right]^{\frac{1}{R}}.$$

This means that the function $\mu \mapsto \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}}$ is convex in μ for $R > 1$, and concave in μ for $0 < R < 1$. Therefore, the function $\mu \mapsto 1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}}$ is concave in μ for $R > 1$, and convex in μ for $0 < R < 1$. Evidently, $\frac{R}{R-1} > 0$ for $R > 1$, and $\frac{R}{R-1} < 0$ for $0 < R < 1$. According to definition of the R -norm entropy $H_R^\mu(A)$, we obtain that for every $R \in (0, 1) \cup (1, \infty)$, the R -norm entropy $\mu \mapsto H_R^\mu(A)$ is a concave function on the family of all fuzzy P -measures on a given fuzzy measurable space (X, M) . Thus, for every $\lambda \in [0, 1]$, it holds that:

$$\lambda H_R^\mu(A) + (1 - \lambda) H_R^\nu(A) \leq H_R^{\lambda\mu + (1-\lambda)\nu}(A). \quad \square$$

Proposition 2. Let A, B be fuzzy partitions of a fuzzy probability space (X, M, μ) such that $A \prec B$. Then there exists a partition $\{I_1, I_2, \dots, I_n\}$ of the set $\{1, 2, \dots, m\}$ such that $\mu(a_i) = \sum_{j \in I_i} \mu(b_j)$, for $i = 1, 2, \dots, n$.

Proof. By the assumption, for every $b \in B$ there exists $a \in A$ such that $b \leq a$. Let us denote by I_i the subset of the set $\{1, 2, \dots, m\}$ such that for every $j \in I_i$, it holds that $b_j \leq a_i$, $i = 1, 2, \dots, n$. Then the set $\{I_1, I_2, \dots, I_n\}$ is a partition of the set $\{1, 2, \dots, m\}$ and $\bigcup_{j \in I_i} b_j \leq a_i$, for $i = 1, 2, \dots, n$. By monotonicity of fuzzy P -measure μ , $\mu(a_i) \geq \mu(\bigcup_{j \in I_i} b_j) = \sum_{j \in I_i} \mu(b_j)$, for $i = 1, 2, \dots, n$. Summing over $i = 1, 2, \dots, n$, we get $\sum_{i=1}^n \mu(a_i) \geq \sum_{i=1}^n \sum_{j \in I_i} \mu(b_j) = \sum_{j=1}^m \mu(b_j) = 1$. Since $\sum_{i=1}^n \mu(a_i) = 1$, it follows that $\mu(a_i) = \sum_{j \in I_i} \mu(b_j)$, for $i = 1, 2, \dots, n$. \square

Theorem 8. Let A, B, C be fuzzy partitions of a fuzzy probability space (X, M, μ) such that $A \prec B$. Then:

- (i) $H_R^\mu(A) \leq H_R^\mu(B)$;
- (ii) $H_R^\mu(A/C) \leq H_R^\mu(B/C)$.

Proof. (i) Assume that $A = \{a_1, a_2, \dots, a_n\}$, $B = \{b_1, b_2, \dots, b_m\}$, $A \prec B$. According to Proposition 2 there exists a partition $\{I_1, I_2, \dots, I_n\}$ of the set $\{1, 2, \dots, m\}$ such that $\mu(a_i) = \sum_{j \in I_i} \mu(b_j)$, for $i = 1, 2, \dots, n$. For the case of $R > 1$, we obtain:

$$\mu(a_i)^R = \left(\sum_{j \in I_i} \mu(b_j) \right)^R \geq \sum_{j \in I_i} \mu(b_j)^R, \text{ for } i = 1, 2, \dots, n,$$

and consequently:

$$\sum_{i=1}^n \mu(a_i)^R \geq \sum_{j=1}^m \mu(b_j)^R.$$

Hence

$$\left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \geq \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}}.$$

Since $\frac{R}{R-1} > 0$ for $R > 1$, we conclude that:

$$H_R^\mu(A) = \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \leq \frac{R}{R-1} \left(1 - \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} \right) = H_R^\mu(B).$$

For the case of $0 < R < 1$, we get:

$$\mu(a_i)^R = \left(\sum_{j \in I_i} \mu(b_j) \right)^R \leq \sum_{j \in I_i} \mu(b_j)^R, \text{ for } i = 1, 2, \dots, n,$$

and consequently:

$$\sum_{i=1}^n \mu(a_i)^R \leq \sum_{j=1}^m \mu(b_j)^R.$$

Therefore:

$$\left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \leq \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}}.$$

Since $\frac{R}{R-1} < 0$ for $0 < R < 1$, we have:

$$H_R^\mu(A) = \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \leq \frac{R}{R-1} \left(1 - \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} \right) = H_R^\mu(B).$$

(ii) By the assumption, for every $b \in B$ there exists $a \in A$ such that $b \leq a$. Hence, for arbitrary element $b \cap c$ of fuzzy partition $B \vee C$ there exists $a \cap c \in A \vee C$ such that $b \cap c \leq a \cap c$. This means that $A \vee C \prec B \vee C$. Therefore, we get:

$$H_R^\mu(A/C) = H_R^\mu(A \vee C) - H_R^\mu(C) \leq H_R^\mu(B \vee C) - H_R^\mu(C) = H_R^\mu(B/C). \quad \square$$

Theorem 9. Let A, B be statistically independent fuzzy partitions of a fuzzy probability space (X, M, μ) . Then:

$$H_R^\mu(A/B) = H_R^\mu(A) - \frac{R-1}{R} H_R^\mu(A) \cdot H_R^\mu(B).$$

Proof. Let $A = \{a_1, a_2, \dots, a_n\}$, and $B = \{b_1, b_2, \dots, b_m\}$. By the assumption, $\mu(a_i \cap b_j) = \mu(a_i) \cdot \mu(b_j)$, for $i = 1, 2, \dots, n, j = 1, 2, \dots, m$. Therefore we can write:

$$\begin{aligned} H_R^\mu(A/B) &= \frac{R}{R-1} \left(\left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} - \left[\sum_{j=1}^m \sum_{i=1}^n \mu(a_i \cap b_j)^R \right]^{\frac{1}{R}} \right) \\ &= \frac{R}{R-1} \left(\left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} - \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \\ &= \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} - 1 + \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} + \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} - \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \\ &= \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) - \frac{R-1}{R} \left(\frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) \cdot \frac{R}{R-1} \left(1 - \left[\sum_{j=1}^m \mu(b_j)^R \right]^{\frac{1}{R}} \right) \right) \\ &= H_R^\mu(A) - \frac{R-1}{R} H_R^\mu(A) \cdot H_R^\mu(B). \quad \square \end{aligned}$$

In view of Theorems 5 and 9, the R -norm entropy does not have the property of additivity, but it satisfies the property that is called pseudo-additivity, as stated in the following theorem.

Theorem 10. (Pseudo-additivity). Let A, B be statistically independent fuzzy partitions of a fuzzy probability space (X, M, μ) . Then:

$$H_R^\mu(A \vee B) = H_R^\mu(A) + H_R^\mu(B) - \frac{R-1}{R} H_R^\mu(A) \cdot H_R^\mu(B).$$

Proof. The result follows by combining Theorems 5 and 9. \square

4. The R -Norm Divergence of Fuzzy P-Measures

In this part, the concept of the R -norm divergence of fuzzy P-measures is defined. In order to avoid expressions like $\frac{0}{0}$, we will use in this section the following simplification: for any fuzzy partition $A = \{a_1, a_2, \dots, a_n\}$ of a fuzzy probability space (X, M, μ) , we assume that $\mu(a_i) > 0$, for $i = 1, 2, \dots, n$. Note that this is without loss of generality, because $\sum_{i=1}^n \mu(a_i) = \sum_{i: \mu(a_i) > 0} \mu(a_i)$. We will prove basic properties of this quantity. The results are illustrated with numerical examples.

Definition 8. Let μ, v be two fuzzy P-measures on a fuzzy measurable space (X, M) , and $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. The R -norm divergence of fuzzy P-measures μ, v with respect to A is defined, for a positive real number R not equal to 1, as the number:

$$D_R^A(\mu \parallel v) = \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right). \quad (8)$$

Remark 3. It is easy to see that, for any fuzzy partition A of a fuzzy probability space (X, M, μ) , we have $D_R^A(\mu \parallel \mu) = 0$.

The following theorem states that the R -norm divergence $D_R^A(\mu \parallel v)$ is consistent, in the case of the limit of R going to 1, with the Kullback–Leibler divergence $D_A(\mu \parallel v)$ defined by formula (3), up to a positive multiplicative constant.

Theorem 11. Let $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. Then $\lim_{R \rightarrow 1} D_R^A(\mu \parallel v) = c \cdot D_A(\mu \parallel v)$, where $D_A(\mu \parallel v) = \sum_{i=1}^n \mu(a_i) \log \frac{\mu(a_i)}{v(a_i)}$, and $c = \frac{1}{\log e}$.

Proof. For every $R \in (0, 1) \cup (1, \infty)$, we can write:

$$D_R^A(\mu \parallel v) = \frac{1}{1 - \frac{1}{R}} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right) = \frac{f(R)}{g(R)},$$

where f, g are continuous functions defined for $R \in (0, \infty)$ by the formulas:

$$f(R) = \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1, g(R) = 1 - \frac{1}{R}.$$

By continuity of the functions f, g , we get $\lim_{R \rightarrow 1} f(R) = f(1) = \sum_{i=1}^n \mu(a_i) v(a_i)^0 - 1 = \sum_{i=1}^n \mu(a_i) - 1 = 1 - 1 = 0$, and $\lim_{R \rightarrow 1} g(R) = g(1) = 0$. Using L'Hôpital's rule this implies that:

$$\lim_{R \rightarrow 1} D_R^A(\mu \parallel v) = \frac{\lim_{R \rightarrow 1} f'(R)}{\lim_{R \rightarrow 1} g'(R)}$$

under the assumption that the right-hand side exists. Let us calculate the derivative of the function $f(R)$:

$$\begin{aligned} \frac{d}{dR} f(R) &= e^{\frac{1}{R} \ln \sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R}} \cdot \left(-\frac{1}{R^2} \cdot \ln \sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right. \\ &\quad \left. + \frac{1}{R} \cdot \frac{1}{\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R}} \sum_{i=1}^n \left(-\mu(a_i)^R v(a_i)^{1-R} \cdot \ln v(a_i) + v(a_i)^{1-R} \mu(a_i)^R \cdot \ln \mu(a_i) \right) \right). \end{aligned}$$

Since $\lim_{R \rightarrow 1} g'(R) = \lim_{R \rightarrow 1} \frac{1}{R^2} = 1$, we get:

$$\begin{aligned} \lim_{R \rightarrow 1} D_R^A(\mu \parallel v) &= \lim_{R \rightarrow 1} f'(R) = \sum_{i=1}^n (\mu(a_i) \cdot \ln \mu(a_i) - \mu(a_i) \cdot \ln v(a_i)) \\ &= \sum_{i=1}^n \left(\mu(a_i) \cdot \frac{\log \mu(a_i)}{\log e} - \mu(a_i) \cdot \frac{\log v(a_i)}{\log e} \right) = \frac{1}{\log e} \sum_{i=1}^n \mu(a_i) \cdot \log \frac{\mu(a_i)}{v(a_i)} = \frac{1}{\log e} D_A(\mu \parallel v). \quad \square \end{aligned}$$

Remark 4. Evidently, if the Kullback–Leibler divergence $D_A(\mu \parallel v)$ is expressed in terms of a natural logarithm, then it is the limiting case of the R -norm divergence $D_R^A(\mu \parallel v)$ for R going to 1.

Let $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. In [32], it has been shown that for the Kullback–Leibler divergence $D_A(\mu \parallel v)$ it holds the Gibbs inequality $D_A(\mu \parallel v) \geq 0$ with the equality if and only if $\mu(a_i) = v(a_i)$, for $i = 1, 2, \dots, n$. This result allows us to interpret the Kullback–Leibler divergence $D_A(\mu \parallel v)$ as a distance measure between two fuzzy P-measures (over the same fuzzy partition). In the following theorem, we present an analogy of this result for the case of the R -norm divergence.

Theorem 12. Let $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. Then $D_R^A(\mu \parallel v) \geq 0$ with the equality if and only if $\mu(a_i) = v(a_i)$, for $i = 1, 2, \dots, n$.

Proof. We shall consider two cases: the case of $R > 1$, and the case of $0 < R < 1$.

Consider the case of $R > 1$. The inequality follows from Jensen's inequality for the function φ defined by $\varphi(x) = x^{1-R}$, for every $x \in [0, \infty)$, and putting $\alpha_i = \mu(a_i)$, $x_i = \frac{v(a_i)}{\mu(a_i)}$, for $i = 1, 2, \dots, n$. The assumption that $R > 1$ implies $1 - R < 0$, hence the function φ is convex. Therefore, by Jensen's inequality we obtain:

$$1 = (\sum_{i=1}^n v(a_i))^{1-R} = \left(\sum_{i=1}^n \mu(a_i) \frac{v(a_i)}{\mu(a_i)} \right)^{1-R} \leq \sum_{i=1}^n \mu(a_i) \left(\frac{v(a_i)}{\mu(a_i)} \right)^{1-R} = \sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R}, \quad (9)$$

and consequently:

$$\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} \geq 1.$$

Since $\frac{R}{R-1} > 0$ for $R > 1$, it follows that:

$$D_R^A(\mu \parallel v) = \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right) \geq 0.$$

For $0 < R < 1$, the function φ defined by $\varphi(x) = x^{1-R}$, for every $x \in [0, \infty)$, is concave. Hence, using the Jensen inequality, we obtain:

$$1 = \left(\sum_{i=1}^n v(a_i) \right)^{1-R} = \left(\sum_{i=1}^n \mu(a_i) \frac{v(a_i)}{\mu(a_i)} \right)^{1-R} \geq \sum_{i=1}^n \mu(a_i) \left(\frac{v(a_i)}{\mu(a_i)} \right)^{1-R} = \sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R},$$

and consequently:

$$\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} \leq 1.$$

Since $\frac{R}{R-1} < 0$ for $0 < R < 1$, we conclude that:

$$D_R^A(\mu \parallel v) = \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right) \geq 0.$$

The equality in (9) holds if and only if $\frac{v(a_i)}{\mu(a_i)}$ is constant, for $i = 1, 2, \dots, n$, i.e., if and only if $v(a_i) = c\mu(a_i)$, for $i = 1, 2, \dots, n$. Taking the sum over all $i = 1, 2, \dots, n$, we get $\sum_{i=1}^n v(a_i) = c\sum_{i=1}^n \mu(a_i)$, which implies that $c = 1$. Therefore, $v(a_i) = \mu(a_i)$, for $i = 1, 2, \dots, n$. This means that $D_R^A(\mu \parallel v) = 0$ if and only if $\mu(a_i) = v(a_i)$, for $i = 1, 2, \dots, n$. \square

In the example that follows, it is shown that the equality $D_R^A(\mu \parallel v) = D_R^A(v \parallel \mu)$ is not necessarily true which means that the R -norm divergence $D_R^A(\mu \parallel v)$ is not symmetrical. Therefore, it is not a metric in a true sense.

Example 3. Consider the fuzzy probability spaces (X, M, μ) , (X, M, v) defined in Example 2 and the fuzzy partition $A = \{a, a^\perp\}$ of (X, M, μ) , (X, M, v) . Let us calculate the R -norm divergencies $D_R^A(\mu \parallel v)$, and $D_R^A(v \parallel \mu)$:

$$D_R^A(\mu \parallel v) = \frac{R}{R-1} \left(\left[\left(\frac{1}{2} \right)^R \left(\frac{1}{3} \right)^{1-R} + \left(\frac{1}{2} \right)^R \left(\frac{2}{3} \right)^{1-R} \right]^{\frac{1}{R}} - 1 \right),$$

$$D_R^A(v \parallel \mu) = \frac{R}{R-1} \left(\left[\left(\frac{1}{3} \right)^R \left(\frac{1}{2} \right)^{1-R} + \left(\frac{2}{3} \right)^R \left(\frac{1}{2} \right)^{1-R} \right]^{\frac{1}{R}} - 1 \right).$$

Put $R = 2$. Elementary calculations show that $D_R^A(\mu \parallel v) \doteq 0.1213$, and $D_R^A(v \parallel \mu) \doteq 0.1082$, thus $D_2^A(\mu \parallel v) \neq D_2^A(v \parallel \mu)$. For $R = \frac{1}{3}$, we have $D_R^A(\mu \parallel v) \doteq 0.0188$, and $D_R^A(v \parallel \mu) \doteq 0.01908$, i.e., $D_{1/3}^A(\mu \parallel v) \neq D_{1/3}^A(v \parallel \mu)$. This means that $D_R^A(\mu \parallel v) \neq D_R^A(v \parallel \mu)$, in general.

Theorem 13. Let μ, v be two fuzzy P -measures on a fuzzy measurable space (X, M) , and $A = \{a_1, a_2, \dots, a_n\}$ be a fuzzy partition of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. In addition, let v be uniform over A , i.e., $v(a_i) = \frac{1}{n}$, for $i = 1, 2, \dots, n$. Then, it holds that:

$$H_R^\mu(A) = \frac{R}{R-1} \left(1 - n^{\frac{1-R}{R}} \right) - n^{\frac{1-R}{R}} \cdot D_R^A(\mu \parallel v). \quad (10)$$

Proof. Let us calculate:

$$\begin{aligned} D_R^A(\mu \parallel v) &= \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right) \\ &= \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R n^{\frac{R-1}{R}} \right]^{\frac{1}{R}} - 1 \right) = \frac{R}{R-1} \cdot n^{\frac{R-1}{R}} \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} - \frac{R}{R-1} \\ &= -n^{\frac{R-1}{R}} \cdot \frac{R}{R-1} \left(1 - \left[\sum_{i=1}^n \mu(a_i)^R \right]^{\frac{1}{R}} \right) + n^{\frac{R-1}{R}} \cdot \frac{R}{R-1} - \frac{R}{R-1} \\ &= -n^{\frac{R-1}{R}} \cdot H_R^\mu(A) + \frac{R}{R-1} \left(n^{\frac{R-1}{R}} - 1 \right). \end{aligned}$$

It follows that:

$$H_R^\mu(A) = \frac{R}{R-1} \left(1 - n^{\frac{1-R}{R}} \right) - n^{\frac{1-R}{R}} \cdot D_R^A(\mu \parallel v). \quad \square$$

Example 4. Consider the fuzzy P -measures μ, v from Example 2 and the fuzzy partition $A = \{a, a^\perp\}$ of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. The fuzzy P -measure μ is uniform over A . Put $R = 2$. Based on previous results, we have $D_R^A(v \parallel \mu) \doteq 0.1082$, and $H_R^v(A) \doteq 0.50928$. Let us calculate:

$$\frac{R}{R-1} \left(1 - 2^{\frac{1-R}{R}} \right) - 2^{\frac{1-R}{R}} \cdot D_R^A(v \parallel \mu) \doteq 2 \left(1 - 2^{-\frac{1}{2}} \right) - 2^{-\frac{1}{2}} \cdot 0.1082 \doteq 0.50928.$$

Thus, the Equality (10) holds.

As a direct consequence of Theorems 12 and 13, we obtain the following property of the R -norm entropy of fuzzy partitions:

Corollary 1. For arbitrary fuzzy partition $A = \{a_1, a_2, \dots, a_n\}$ of a fuzzy probability space (X, M, μ) , it holds that:

$$H_R^\mu(A) \leq \frac{R}{R-1} \left(1 - n^{\frac{1-R}{R}} \right)$$

with the equality if and only if the fuzzy P -measure μ is uniform over A .

Theorem 14. Let μ_1, μ_2, v be fuzzy P -measures on a fuzzy measurable space (X, M, μ_1) , and A be a fuzzy partition of fuzzy probability spaces $(X, M, \mu_2), (X, M, v)$. Then, for every real number $\lambda \in [0, 1]$, it holds that:

$$D_R^A(\lambda\mu_1 + (1-\lambda)\mu_2 \parallel v) \leq \lambda D_R^A(\mu_1 \parallel v) + (1-\lambda) D_R^A(\mu_2 \parallel v).$$

Proof. Assume that $A = \{a_1, a_2, \dots, a_n\}$, and $\lambda \in [0, 1]$. Putting $x_i = \lambda\mu_1(a_i)v(a_i)^{\frac{1-R}{R}}$, and $y_i = (1-\lambda)\mu_2(a_i)v(a_i)^{\frac{1-R}{R}}, i = 1, 2, \dots, n$, in the Minkowski inequality, we get for $R > 1$:

$$\begin{aligned}
\left[\sum_{i=1}^n (\lambda \mu_1(a_i) + (1-\lambda) \mu_2(a_i))^R v(a_i)^{1-R} \right]^{\frac{1}{R}} &= \left[\sum_{i=1}^n \left((\lambda \mu_1(a_i) + (1-\lambda) \mu_2(a_i)) v(a_i)^{\frac{1-R}{R}} \right)^R \right]^{\frac{1}{R}} \\
&= \left[\sum_{i=1}^n \left(\lambda \mu_1(a_i) v(a_i)^{\frac{1-R}{R}} + (1-\lambda) \mu_2(a_i) v(a_i)^{\frac{1-R}{R}} \right)^R \right]^{\frac{1}{R}} \\
&\leq \left[\sum_{i=1}^n \left(\lambda \mu_1(a_i) v(a_i)^{\frac{1-R}{R}} \right)^R \right]^{\frac{1}{R}} + \left[\sum_{i=1}^n \left((1-\lambda) \mu_2(a_i) v(a_i)^{\frac{1-R}{R}} \right)^R \right]^{\frac{1}{R}} \\
&= \lambda \left[\sum_{i=1}^n \mu_1(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} + (1-\lambda) \left[\sum_{i=1}^n \mu_2(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}},
\end{aligned}$$

and for $0 < R < 1$:

$$\begin{aligned}
&\left[\sum_{i=1}^n (\lambda \mu_1(a_i) + (1-\lambda) \mu_2(a_i))^R v(a_i)^{1-R} \right]^{\frac{1}{R}} \\
&\geq \left[\sum_{i=1}^n \left(\lambda \mu_1(a_i) v(a_i)^{\frac{1-R}{R}} \right)^R \right]^{\frac{1}{R}} + \left[\sum_{i=1}^n \left((1-\lambda) \mu_2(a_i) v(a_i)^{\frac{1-R}{R}} \right)^R \right]^{\frac{1}{R}} \\
&= \lambda \left[\sum_{i=1}^n \mu_1(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} + (1-\lambda) \left[\sum_{i=1}^n \mu_2(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}}.
\end{aligned}$$

This means that the function $\mu \mapsto \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}}$ is convex in μ for $R > 1$, and concave in μ for $0 < R < 1$. The same applies to the function $\mu \mapsto \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1$. Since $\frac{R}{R-1} > 0$ for $R > 1$, and $\frac{R}{R-1} < 0$ for $0 < R < 1$, we conclude that the function $\mu \mapsto D_R^A(\mu \| v)$ is convex on the family of all fuzzy P-measures on a given fuzzy measurable space (X, M) . Thus, for every real number $\lambda \in [0, 1]$, it holds that:

$$D_R^A(\lambda \mu_1 + (1-\lambda) \mu_2 \| v) \leq \lambda D_R^A(\mu_1 \| v) + (1-\lambda) D_R^A(\mu_2 \| v). \quad \square$$

Theorem 15. Let $A = \{a_1, a_2, \dots, a_n\}$ be any fuzzy partition of a fuzzy probability space (X, M, μ) . Then:

- (i) $0 < R < 1$ implies $D_R^A(\mu \| v) \leq D_A(\mu \| v)$;
- (ii) $R > 1$ implies $D_R^A(\mu \| v) \geq D_A(\mu \| v)$,

where

$$D_A(\mu \| v) = \sum_{i=1}^n \mu(a_i) \log \frac{\mu(a_i)}{v(a_i)}.$$

Proof. In the proof we use the Jensen inequality for the concave function φ defined by $\varphi(x) = \log x$, for $x \in (0, \infty)$, and putting $\alpha_i = \mu(a_i)$, $x_i = \left(\frac{v(a_i)}{\mu(a_i)} \right)^{R-1}$, for $i = 1, 2, \dots, n$. Since the logarithm satisfies the condition $\log x \leq x - 1$, for all real numbers $x > 0$, we get:

$$\log \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} \leq \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1. \quad (11)$$

Suppose that $0 < R < 1$. Then $\frac{R}{R-1} < 0$ and using the inequality (11) and the Jensen inequality, we can write:

$$\begin{aligned}
 D_R^A(\mu \parallel v) &= \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right) \\
 &\leq \frac{R}{R-1} \log \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} = \frac{1}{R-1} \log \sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \\
 &= \frac{1}{R-1} \log \sum_{i=1}^n \mu(a_i) \left(\frac{\mu(a_i)}{v(a_i)} \right)^{R-1} \leq \frac{1}{R-1} \sum_{i=1}^n \mu(a_i) \log \left(\frac{\mu(a_i)}{v(a_i)} \right)^{R-1} \\
 &= \sum_{i=1}^n \mu(a_i) \log \frac{\mu(a_i)}{v(a_i)} = D_A(\mu \parallel v).
 \end{aligned}$$

The case of $R > 1$ can be obtained in the similar way. \square

Example 5. Consider the fuzzy P-measures μ, v from Example 2 and the fuzzy partition $A = \{a, a^\perp\}$ of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. Based on the results from Example 3, we have $D_2^A(\mu \parallel v) \doteq 0.1213$, $D_2^A(v \parallel \mu) \doteq 0.1082$, $D_{1/3}^A(\mu \parallel v) \doteq 0.0188$, $D_{1/3}^A(v \parallel \mu) \doteq 0.01908$. By simple calculation we get that $D_A(\mu \parallel v) \doteq 0.084963$, and $D_A(v \parallel \mu) \doteq 0.081703$. Thus, for $R = \frac{1}{3}$ we have $D_R^A(\mu \parallel v) \leq D_A(\mu \parallel v)$, $D_R^A(v \parallel \mu) \leq D_A(v \parallel \mu)$; and for $R = 2$ we have $D_R^A(\mu \parallel v) \geq D_A(\mu \parallel v)$, and $D_R^A(v \parallel \mu) \geq D_A(v \parallel \mu)$, which is consistent with the statement in the previous theorem.

We conclude our contribution with the formulation of a chain rule for the R -norm divergence in the fuzzy case. First, we define the conditional version of the R -norm divergence of fuzzy P-measures.

Definition 9. Let $A = \{a_1, a_2, \dots, a_n\}$, $B = \{b_1, b_2, \dots, b_m\}$ be two fuzzy partitions of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. Then, we define the conditional divergence of fuzzy P-measures μ, v with respect to B assuming a realization of A , for a positive real number R not equal to 1, as the number:

$$D_R^{B/A}(\mu \parallel v) = \frac{R}{R-1} \left(\left[\sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R v(a_i \cap b_j)^{1-R} \right]^{\frac{1}{R}} - \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} \right).$$

Theorem 16. Let A, B be two fuzzy partitions of fuzzy probability spaces $(X, M, \mu), (X, M, v)$. Then

$$D_R^{A \vee B}(\mu \parallel v) = D_R^A(\mu \parallel v) + D_R^{B/A}(\mu \parallel v).$$

Proof. Assume that $A = \{a_1, a_2, \dots, a_n\}$, $B = \{b_1, b_2, \dots, b_m\}$. Then we have:

$$\begin{aligned}
 D_R^A(\mu \parallel v) + D_R^{B/A}(\mu \parallel v) &= \frac{R}{R-1} \left(\left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} - 1 \right) \\
 &+ \frac{R}{R-1} \left(\left[\sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R v(a_i \cap b_j)^{1-R} \right]^{\frac{1}{R}} - \left[\sum_{i=1}^n \mu(a_i)^R v(a_i)^{1-R} \right]^{\frac{1}{R}} \right) \\
 &= \frac{R}{R-1} \left(\left[\sum_{i=1}^n \sum_{j=1}^m \mu(a_i \cap b_j)^R v(a_i \cap b_j)^{1-R} \right]^{\frac{1}{R}} - 1 \right) = D_R^{A \vee B}(\mu \parallel v). \quad \square
 \end{aligned}$$

5. Conclusions

In this article, we have extended the study of entropy measures and distance measures in the fuzzy case. Our goal was to introduce the concepts of R -norm entropy and R -norm divergence for the case of fuzzy probability spaces and to derive basic properties of these measures. Our results are presented in Sections 3 and 4.

In Section 3, we have defined the R -norm entropy and conditional R -norm entropy of fuzzy partitions of a given fuzzy probability space and have examined the properties of the proposed entropy measures. In particular, it has been shown that the R -norm entropy of fuzzy partitions does not have the property of additivity, but it satisfies the property called pseudo-additivity, as stated in Theorem

10. In Theorem 6, chain rules for the R -norm entropy of fuzzy partitions are provided. Moreover, it was shown that the Shannon entropy and the conditional Shannon entropy of fuzzy partitions can be derived from the R -norm entropy and conditional R -norm entropy of fuzzy partitions, respectively, as the limiting cases for $R \rightarrow 1$.

In Section 4, the concept of R -norm divergence of fuzzy P-measures was introduced and the properties of this quantity have been proven. Specifically, it was shown that the Kullback–Leibler divergence defined and studied in [32] can be derived from the R -norm divergence of fuzzy P-measures, as the limiting case for $R \rightarrow 1$. The result of Theorem 12 allows us to interpret the R -norm divergence as a distance measure between two fuzzy P-measures. Theorem 13 provides a relationship between the R -norm divergence and the R -norm entropy of fuzzy partitions; Theorem 15 provides a relationship between the R -norm divergence and the Kullback–Leibler divergence of fuzzy P-measures. In addition, the concavity of R -norm entropy (Theorem 7) and convexity of R -norm divergence (Theorem 14) have been demonstrated. Finally, using the suggested concept of conditional R -norm divergence of fuzzy P-measures, the chain rule for the R -norm divergence of fuzzy P-measures was established.

In the proofs, the Jensen inequality, L'Hôpital's rule, and the Minkowski inequality were used. The results presented in Sections 3 and 4 are illustrated with numerical examples.

Acknowledgments: The authors are grateful to the editor and the anonymous referees for their valuable comments and suggestions. The authors acknowledge Constantine the Philosopher University in Nitra for covering the costs of publishing in open access.

Author Contributions: All authors contributed significantly to the theoretical work as well as to the creation of illustrative examples. Dagmar Markechová wrote the paper. All authors have read and approved the final manuscript.

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

References

- Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
- Gray, R.M. *Entropy and Information Theory*; Springer: Berlin/Heidelberg, Germany, 2009.
- Rényi, A. On measures of information and entropy. In *Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability*, Berkeley, CA, USA, 20 June–30 July 1960; University of California Press: Berkeley, CA, USA, 1961; pp. 547–561.
- Ellerman, D. An Introduction to Logical Entropy and Its Relation to Shannon Entropy. *Int. J. Semant. Comput.* **2013**, *7*, 121–145. [[CrossRef](#)]
- Ellerman, D. Logical Information Theory: New Foundations for Information Theory. *Log. J. IGPL* **2017**, *25*, 806–835. [[CrossRef](#)]
- Arimoto, S. Information theoretical considerations on estimation problems. *Inf. Control* **1971**, *19*, 181–194. [[CrossRef](#)]
- Boekke, D.E.; Van Der Lubbe, J.C.A. The R -Norm Information Measure. *Inf. Control* **1980**, *45*, 136–155. [[CrossRef](#)]
- Hooda, D.S.; Ram, A. Characterization of a Generalized Measure of R -norm Entropy. *Caribb. J. Math. Comput. Sci.* **2002**, *8*, 18–31.
- Hooda, D.S.; Sharma, D.K. Generalized R -norm Information Measures. *J. Appl. Math. Stat. Inform.* **2008**, *4*, 153–168.
- Hooda, D.S.; Bajaj, R.K. On Generalized R -norm Information Measures of Fuzzy Information. *J. Appl. Math. Stat. Inform.* **2008**, *4*, 199–212.
- Kumar, S.; Choudhary, A. Generalized Parametric R -norm Information Measure. *Trends Appl. Sci. Res.* **2012**, *7*, 350–369.
- Kumar, S.; Choudhary, A.; Kumar, R. Some More Results on a Generalized Parametric R -Norm Information Measure of Type α . *J. Appl. Sci. Eng.* **2014**, *17*, 447–453. [[CrossRef](#)]
- Zadeh, L.A. Fuzzy Sets. *Inf. Control* **1965**, *8*, 338–358. [[CrossRef](#)]
- Piasecki, K. Fuzzy partitions of sets. *BUSEFAL* **1986**, *25*, 52–60.
- De Baets, B.; Mesiar, R. T-Partitions. *Fuzzy Sets Syst.* **1998**, *97*, 211–223. [[CrossRef](#)]

16. Mesiar, R.; Reusch, B.; Thiele, H. Fuzzy equivalence relations and fuzzy partition. *J. Mult. Valued Log. Soft Comput.* **2006**, *12*, 167–181.
17. Jayaram, B.; Mesiar, R. I-fuzzy equivalence relations and I-fuzzy partitions. *Inf. Sci.* **2009**, *179*, 1278–1297. [CrossRef]
18. Montes, S.; Couso, I.; Gil, P. Fuzzy delta-epsilon partitions. *Inf. Sci.* **2003**, *152*, 267–285. [CrossRef]
19. Montes, S.; Couso, I.; Gil, P. One-to-one correspondence between ε -partitions, $(1 - \varepsilon)$ -equivalences and ε -pseudometrics. *Fuzzy Sets Syst.* **2001**, *124*, 87–95. [CrossRef]
20. Dumitrescu, D. Fuzzy partitions with the connectives T_∞ , S_∞ . *Fuzzy Sets Syst.* **1992**, *47*, 193–195. [CrossRef]
21. Dumitrescu, D. Fuzzy measures and entropy of fuzzy partitions. *J. Math. Anal. Appl.* **1993**, *176*, 359–373. [CrossRef]
22. Riečan, B. An entropy construction inspired by fuzzy sets. *Soft Comput.* **2003**, *7*, 486–488.
23. Markechová, D. The entropy of fuzzy dynamical systems and generators. *Fuzzy Sets Syst.* **1992**, *48*, 351–363. [CrossRef]
24. Markechová, D. Entropy of complete fuzzy partitions. *Math. Slovaca* **1993**, *43*, 1–10.
25. Markechová, D. Entropy and mutual information of experiments in the fuzzy case. *Neural Netw. World* **2013**, *23*, 339–349. [CrossRef]
26. Mesiar, R. The Bayes principle and the entropy on fuzzy probability spaces. *Int. J. Gen. Syst.* **1991**, *20*, 67–72. [CrossRef]
27. Mesiar, R.; Rybárik, J. Entropy of Fuzzy Partitions—A General Model. *Fuzzy Sets Syst.* **1998**, *99*, 73–79. [CrossRef]
28. Rahimi, M.; Riaz, A. On local entropy of fuzzy partitions. *Fuzzy Sets Syst.* **2014**, *234*, 97–108. [CrossRef]
29. Srivastava, P.; Khare, M.; Srivastava, Y.K. m-Equivalence, entropy and F-dynamical systems. *Fuzzy Sets Syst.* **2001**, *121*, 275–283. [CrossRef]
30. Khare, M. Fuzzy σ -algebras and conditional entropy. *Fuzzy Sets Syst.* **1999**, *102*, 287–292. [CrossRef]
31. Markechová, D.; Riečan, B. Entropy of Fuzzy Partitions and Entropy of Fuzzy Dynamical Systems. *Entropy* **2016**, *18*, 19. [CrossRef]
32. Markechová, D. Kullback–Leibler Divergence and Mutual Information of Experiments in the Fuzzy Case. *Axioms* **2017**, *6*, 5. [CrossRef]
33. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
34. Kôpka, F.; Chovanec, F. D-posets. *Math. Slovaca* **1994**, *44*, 21–34.
35. Kôpka, F. Quasiproduct on Boolean D-posets. *Int. J. Theor. Phys.* **2008**, *47*, 26–35. [CrossRef]
36. Frič, R. On D-posets of fuzzy sets. *Math. Slovaca* **2014**, *64*, 545–554. [CrossRef]
37. Riečan, B.; Mundici, D. Probability on MV-algebras. In *Handbook of Measure Theory*; Pap, E., Ed.; Elsevier: Amsterdam, The Netherlands, 2002; pp. 869–910.
38. Mundici, D. MV Algebras: A Short Tutorial. 2007. Available online: http://www.matematica.uns.edu.ar/IXCongresoMonteiro/Comunicaciones/Mundici_tutorial.pdf (accessed on 26 May 2007).
39. Jakubík, J. On product MV algebras. *Czechoslov. Math. J.* **2002**, *52*, 797–810. [CrossRef]
40. Di Nola, A.; Dvurečenskij, A. Product MV-algebras. *Mult. Valued Log.* **2001**, *6*, 193–215.
41. Dvurečenskij, A.; Pulmannová, S. *New Trends in Quantum Structures*; Springer: Dordrecht, The Netherlands, 2000.
42. Foulis, D.J.; Bennet, M.K. Effect algebras and unsharp quantum logics. *Found. Phys.* **1994**, *24*, 1331–1352. [CrossRef]
43. Markechová, D.; Riečan, B. Logical Entropy of Fuzzy Dynamical Systems. *Entropy* **2016**, *18*, 157. [CrossRef]
44. Di Nola, A.; Dvurečenskij, A.; Hyčko, M.; Manara, C. Entropy on Effect Algebras with the Riesz Decomposition Property II: MV-Algebras. *Kybernetika* **2005**, *41*, 161–176.
45. Riečan, B. Kolmogorov–Sinaj entropy on MV-algebras. *Int. J. Theor. Phys.* **2005**, *44*, 1041–1052. [CrossRef]
46. Di Nola, A.; Dvurečenskij, A.; Hyčko, M.; Manara, C. Entropy on Effect Algebras with the Riesz Decomposition Property I: Basic Properties. *Kybernetika* **2005**, *41*, 143–160.
47. Eslami Giski, Z.; Ebrahimi, M. Entropy of Countable Partitions on effect Algebras with the Riesz Decomposition Property and Weak Sequential Effect Algebras. *Cankaya Univ. J. Sci. Eng.* **2015**, *12*, 20–39.
48. Ebrahimi, M.; Mosapour, B. The Concept of Entropy on D-posets. *Cankaya Univ. J. Sci. Eng.* **2013**, *10*, 137–151.
49. Petrovičová, J. On the entropy of partitions in product MV-algebras. *Soft Comput.* **2000**, *4*, 41–44. [CrossRef]
50. Petrovičová, J. On the entropy of dynamical systems in product MV-algebras. *Fuzzy Sets Syst.* **2001**, *121*, 347–351. [CrossRef]

51. Markechová, D.; Riečan, B. Kullback–Leibler Divergence and Mutual Information of Partitions in Product MV Algebras. *Entropy* **2017**, *19*, 267. [[CrossRef](#)]
52. Markechová, D.; Mosapour, B.; Ebrahimzadeh, A. Logical Divergence, Logical Entropy, and Logical Mutual Information in Product MV-Algebras. *Entropy* **2018**, *20*, 129. [[CrossRef](#)]
53. Ebrahimzadeh, A.; Eslami Giski, Z.; Markechová, D. Logical Entropy of Dynamical Systems—A General Model. *Mathematics* **2017**, *5*, 4. [[CrossRef](#)]
54. Ebrahimzadeh, A. Logical entropy of quantum dynamical systems. *Open Phys.* **2016**, *14*, 1–5. [[CrossRef](#)]
55. Ebrahimzadeh, A. Quantum conditional logical entropy of dynamical systems. *Ital. J. Pure Appl. Math.* **2016**, *36*, 879–886.
56. Ebrahimzadeh, A.; Jamalzadeh, J. Conditional logical entropy of fuzzy σ -algebras. *J. Intell. Fuzzy Syst.* **2017**, *33*, 1019–1026. [[CrossRef](#)]
57. Eslami Giski, Z.; Ebrahimzadeh, A. An introduction of logical entropy on sequential effect algebra. *Indag. Math.* **2017**, *28*, 928–937. [[CrossRef](#)]
58. Mohammadi, U. The Concept of Logical Entropy on D-posets. *J. Algebra Struct. Appl.* **2016**, *1*, 53–61.
59. Piasecki, K. Probability of fuzzy events defined as denumerable additive measure. *Fuzzy Sets Syst.* **1985**, *17*, 271–284. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).