Kullback–Leibler Divergence and Mutual Information of Partitions in Product MV Algebras

Dagmar Markechová [1],* and Beloslav Riečan [2,3]

[1] Department of Mathematics, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, A. Hlinku 1, SK-949 01 Nitra, Slovakia

[2] Department of Mathematics, Faculty of Natural Sciences, Matej Bel University, Tajovského 40, SK-974 01 Banská Bystrica, Slovakia; Beloslav.Riecan@umb.sk

[3] Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, SK-814 73 Bratislava, Slovakia

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

Abstract: The purpose of the paper is to introduce, using the known results concerning the entropy in product MV algebras, the concepts of mutual information and Kullback–Leibler divergence for the case of product MV algebras and examine algebraic properties of the proposed measures. In particular, a convexity of Kullback–Leibler divergence with respect to states in product MV algebras is proved, and chain rules for mutual information and Kullback–Leibler divergence are established. In addition, the data processing inequality for conditionally independent partitions in product MV algebras is proved.

1. Introduction

The notions of entropy and mutual information are fundamental concepts in information theory [1]; they are used as measures of information obtained from a realization of the considered experiments. The standard approach in information theory is based on the Shannon entropy [2]. Consider a finite measurable partition $\mathcal{A}$ of probability space (Ω, S, P) with probabilities $p_1, ..., p_n$ of the corresponding elements of $\mathcal{A}$. We recall that the Shannon entropy of $\mathcal{A}$ is the number $H(\mathcal{A}) = -\sum_{i=1}^{n} F(p_i)$, where the function $F : [0, \infty) \to \Re$ is defined by $F(x) = x \log x$, if $x > 0$, and $F(0) = 0$. Perhaps a crucial point in applications of the Shannon entropy in another scientific field presents the discovery of Kolmogorov and Sinai [3] (see also [4,5]). They showed an existence of non-isomorphic Bernoulli shifts describing independent repetition of random spaces with finite numbers of results. If two dynamical systems are isomorphic, they have the same Kolmogorov-Sinai entropy. So Kolmogorov and Sinai constructed two Bernoulli shifts with different entropies, hence non-isomorphic. It is natural that the mentioned modification of entropy has been used in many mathematical structures. In [6], we have generalized the notion of Kolmogorov–Sinai entropy to the case when the considered probability space is a fuzzy probability space (Ω, M, μ) defined by Piasecki [7]. This structure can serve as an alternative mathematical model of probability theory for the situations where the observed events are described unclearly, vaguely (so called fuzzy events). Other fuzzy generalizations of Shannon's and Kolmogorov–Sinai's entropy can be found e.g., in [8–17]. It is known that there are many possibilities for defining operations with fuzzy sets; an overview can be found in [18]. It should be noted that while the model presented in [6] was based on the Zadeh connectives [19], in our recently published paper [14], the Lukasiewicz connectives were used to define the fuzzy set operations. In [20], the mutual information of fuzzy partitions of a given fuzzy probability

space (Ω, M, μ) has been defined. It was shown that the entropy of fuzzy partitions introduced and studied in [6] can be considered as a special case of their mutual information.

In classical information theory the mutual information is a special case of a more general quantity called Kullback–Leibler divergence (K–L divergence for short), which was originally introduced by Kullback and Leibler in 1951 [21] (see also [22]) as the divergence between two probability distributions. It plays an important role, as a mathematical tool, in the stability analysis of master equations [23] and Fokker–Planck equations [24], and in isothermal equilibrium fluctuations and transient nonequilibrium deviations [25] (see also [24,26]). In [27], we have introduced the concept of K–L divergence for the case of fuzzy probability spaces.

A natural generalization of some family of fuzzy sets is the notion of an MV algebra introduced by Chang [28]. An MV algebra is an algebraic structure which models the Lukasiewicz multivalued logic, and the fragment of that calculus which deals with the basic logical connectives "and", "or", and "not", but in a multivalued context. MV algebras play a similar role in the multivalued logic as Boolean algebras in the classical two-valued logic. Recall also that families of fuzzy sets can be embedded to suitable MV algebras. MV algebras have been studied by many authors (see e.g., [29–33]) and, of course, there are also many results about the entropy on this structure (cf. [34,35]). The theory of fuzzy sets is a rapidly and massively developing area of theoretical and applied mathematical research. In addition to MV algebras, generalizations of MV algebras as D-posets (cf. [36–38]), effect algebras (cf. [39]), or A-posets (cf. [40,41]) are currently subject of intensive research. Some results about the entropy on these structures can be found e.g., in [42–44].

A special class of MV algebras is a class of product MV algebras. They have been introduced independently in [45] from the point of view of probability theory, and in [46] from the point of view of mathematical logic. Product MV algebras have been studied e.g., in [47,48]. A suitable theory of entropy of Kolmogorov type for the case of product MV algebras has been constructed in [35,49,50].

The purpose of this contribution is to define, using the results concerning the entropy in product MV algebras, the concepts of mutual information and Kullback–Leibler divergence for the case of product MV algebras and to study properties of the suggested measures. The main results of the contribution are presented in Sections 3 and 4. In Section 3 the notions of mutual information and conditional mutual information in product MV algebras are introduced and basic properties of the suggested measures are proved, inter alia, the data processing inequality for conditionally independent partitions. In Section 4 we define the Kullback–Leibler divergence in product MV algebras and its conditional version and examine the algebraic properties of the proposed measures. Our results are summarized in the final section.

2. Basic Definitions, Notations and Facts

In this section, we recall some definitions and basic facts which will be used in the following ones. An MV algebra [30] is a system $(M, \oplus, \otimes, *, 0, 1)$, where M is a non-empty set, $\oplus, \otimes$ are binary operations on M, $*$ is a unary operation on M and $0, 1$ are fixed elements of M, such that the following conditions are satisfied:

(i)　　$a \oplus b = b \oplus a$;

(ii)　　$a \oplus (b \oplus c) = (a \oplus b) \oplus c$;

(iii)　$a \oplus 0 = a$;

(iv)　$a \oplus 1 = 1$;

(v)　　$(a^*)^* = a$;

(vi)　$0^* = 1$;

(vii)　$a \oplus a^* = 1$;

(viii) $(a^* \oplus b)^* \oplus b = (a \oplus b^*)^* \oplus a$;

(ix)　$a \otimes b = (a^* \oplus b^*)^*$.

An example of MV algebra is the real interval $[0, 1]$ equipped with the operations $x \oplus y = \min(1, x + y)$, $x \otimes y = \max(0, x + y - 1)$. It is interesting that any MV algebra has a similar structure. In fact, by the Mundici theorem [33] any MV algebra can be represented by a lattice-ordered Abelian group (shortly Abelian *l*-group). Recall that an Abelian *l*-group is an algebraic system $(G, +, \leq)$, where $(G, +)$ is an Abelian group, $(G, \leq)$ is a partially ordered set being a lattice and $a \leq b$ implies $a + c \leq b + c$.

Let $(G, +, \leq)$ be an Abelian *l*-group, 0 be a neutral element of $(G, +)$ and $u \in G$, $u > 0$. On the interval $[0, u] = \{h \in G; 0 \leq h \leq u\}$ we define the following operations: $a^* = u - a$, $a \oplus b = (a + b) \wedge u$; $a \otimes b = (a + b - u) \vee 0$. Then the system $MG = ([0, u], \oplus, \otimes, *, 0, u)$ becomes an MV algebra. The Mundici theorem states that to any MV algebra M there exists an Abelian *l*-group G with a strong unit u (i.e., to every $a \in G$ there exists $n \in \mathbb{N}$ with the property $a \leq nu$) such that $M \cong MG$.

In this contribution we shall consider MV algebras with a product. We recall that the definition of product MV algebra is based on Mundici's categorical representation of MV algebra by an Abelian *l*-group, i.e., the sum in the following definition of product MV algebra, and subsequently in the next text, means the sum in the Abelian *l*-group associated to the given MV algebra. Similarly, the element u is a strong unit of this group. More details can be found in [45,46].

Definition 1. *A product MV algebra is a couple* $(M, \cdot)$, *where M is an MV algebra and $\cdot$ is a commutative and associative operation on M satisfying the following conditions:*

(i)   *for any* $a \in M$, $u \cdot a = a$;
(ii) *if* $a, b, c \in M$, $a + b \leq u$, *then* $c \cdot a + c \cdot b \leq u$, *and* $c \cdot (a + b) = c \cdot a + c \cdot b$.

In addition, we shall consider a finitely additive state defined on a product MV algebra.

Definition 2 [30]. *Let* $(M, \cdot)$ *be a product MV algebra. A map* $m : M \to [0, 1]$ *is said to be a state if the following properties are satisfied:*

(i) $m(u) = 1$;
(ii) *if* $a = \sum_{i=1}^{n} a_i$, *then* $m(a) = \sum_{i=1}^{n} m(a_i)$.

In product MV algebras a suitable entropy theory has been provided in [35,49,50]. In the following we present the main idea and some results of this theory which will be used in the contribution.

Definition 3. *By a partition in a product MV algebra* $(M, \cdot)$ *we mean a finite collection* $\mathcal{A} = \{a_1, ..., a_n\} \subset M$ *such that* $\sum_{i=1}^{n} a_i = u$.

Let m be a state on a product MV algebra $(M, \cdot)$. In the set of all partitions of $(M, \cdot)$ the relation $\prec$ is defined in the following way: Let $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ be two partitions of $(M, \cdot)$. We say that $\mathcal{B}$ is a refinement of $\mathcal{A}$ (with respect to m), and write $\mathcal{A} \prec \mathcal{B}$, if there exists a partition $I(1), I(2), ..., I(n)$ of the set $\{1, 2, ..., k\}$ such that $m(a_i) = \sum_{j \in I(i)} m(b_j)$, for every $i = 1, 2, ..., n$. Given two partitions $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ of $(M, \cdot)$, their join $\mathcal{A} \vee \mathcal{B}$ is defined as the system $\mathcal{A} \vee \mathcal{B} = \{a_i \cdot b_j; i = 1, ..., n, j = 1, ..., k\}$, if $\mathcal{A} \neq \mathcal{B}$, and $\mathcal{A} \vee \mathcal{A} = \mathcal{A}$. Since $\sum_{i=1}^{n} \sum_{j=1}^{k} a_i \cdot b_j = \sum_{i=1}^{n} a_i \cdot \sum_{j=1}^{k} b_j = u \cdot (\sum_{j=1}^{k} b_j) = \sum_{j=1}^{k} b_j = u$, the system $\mathcal{A} \vee \mathcal{B}$ is a partition of $(M, \cdot)$, too. If $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ are partitions in a product MV algebra $(M, \cdot)$, then we put $\vee_{i=1}^{n} \mathcal{A}_i = \mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n$.

Let $\mathcal{A} = \{a_1, ..., a_n\}$ be a partition in a product MV algebra $(M, \cdot)$ and m be a state on $(M, \cdot)$. Then the entropy of $\mathcal{A}$ with respect to m is defined by Shannon's formula:

$$H_m(\mathcal{A}) = -\sum_{i=1}^{n} F(m(a_i)), \tag{1}$$

where:

$$F : [0, \infty) \to \Re, \ F(x) = \begin{cases} x \log x, & \text{if } x > 0; \\ 0, & \text{if } x = 0. \end{cases}$$

If $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ are two partitions of $(M, \cdot)$, then the conditional entropy of $\mathcal{A}$ given $\mathcal{B}$ is defined by:

$$H_m(\mathcal{A}/\mathcal{B}) = -\sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log \frac{m(a_i \cdot b_j)}{m(b_j)}.$$

In accordance with the classical theory the log is to the base 2 and the entropy is expressed in bits. Note that we use the convention (based on continuity arguments) that $x \log \frac{x}{0} = \infty$ if $x > 0$, and $0 \log \frac{0}{x} = 0$ if $x \geq 0$.

Example 1. *Consider any product MV algebra $(M, \cdot)$ and a state m defined on M. Then the set $\mathcal{E} = \{u\}$ is a partition of $(M, \cdot)$ such that $\mathcal{E} \prec \mathcal{A}$ for any partition $\mathcal{A}$ of $(M, \cdot)$. Its entropy is $H_m(\mathcal{E}) = 0$. Let $a \in M$ such that $m(a) = p$, where $p \in (0, 1)$. Evidently, $m(u - a) = 1 - p$, and the set $\mathcal{A} = \{a, u - a\}$ is a partition of $(M, \cdot)$. The entropy $H_m(\mathcal{A}) = -p \log p - (1 - p) \log(1 - p)$. In particular, if $p = \frac{1}{2}$, then $H_m(\mathcal{A}) = \log 2 = 1$ bit.*

The entropy and the conditional entropy of partitions in a product MV algebra satisfy all properties analogous to properties of Shannon's entropy of measurable partitions in the classical case; the proofs can be found in [35,49,50]. We present those that will be further exploited. Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be any partitions of a product MV algebra $(M, \cdot)$. Then the following properties hold: (E1) $H_m(\mathcal{A}) \geq 0$; (E2) $\mathcal{B} \prec \mathcal{C}$ implies $H_m(\mathcal{A}/\mathcal{C}) \leq H_m(\mathcal{A}/\mathcal{B})$; (E3) $H_m(\mathcal{A} \vee \mathcal{B} /\mathcal{C}) = H_m(\mathcal{A} /\mathcal{C}) + H_m(\mathcal{B} /\mathcal{C} \vee \mathcal{A})$; (E4) $H_m(\mathcal{A} \vee \mathcal{B}) = H_m(\mathcal{A}) + H_m(\mathcal{B}/\mathcal{A})$; (E5) $H_m(\mathcal{A} \vee \mathcal{B} /\mathcal{C}) \leq H_m(\mathcal{A} /\mathcal{C}) + H_m(\mathcal{B} /\mathcal{C})$.

3. Mutual Information of Partitions in Product MV Algebras

In this section the results concerning the entropy in product MV algebras are used in developing information theory for the case of product MV algebras. We define the notions of mutual information and conditional mutual information of partitions in a product MV algebra and prove basic properties of the proposed measures.

Definition 4. *Let $\mathcal{A}$, $\mathcal{B}$ be partitions in a given product MV algebra $(M, \cdot)$. Then we define the mutual information of $\mathcal{A}$ and $\mathcal{B}$ by the formula:*

$$I_m(\mathcal{A}, \mathcal{B}) = H_m(\mathcal{A}) - H_m(\mathcal{A}/\mathcal{B}). \tag{2}$$

Remark 1. *As a simple consequence of (E4) we get:*

$$I_m(\mathcal{A}, \mathcal{B}) = H_m(\mathcal{A}) + H_m(\mathcal{B}) - H_m(\mathcal{A} \vee \mathcal{B}). \tag{3}$$

Subsequently we see that $I_m(\mathcal{A}, \mathcal{A}) = H_m(\mathcal{A})$, i.e., the entropy of partitions in product MV algebras can be considered as a special case of their mutual information. Moreover, we see that $I_m(\mathcal{A}, \mathcal{B}) = I_m(\mathcal{B}, \mathcal{A})$, and hence we can also write:

$$I_m(\mathcal{A}, \mathcal{B}) = H_m(\mathcal{B}) - H_m(\mathcal{B}/\mathcal{A}). \tag{4}$$

Example 2. *Consider the measurable space (Ω, S), where Ω is the unit interval $[0, 1]$, and S is the σ-algebra of all Borel subsets of $[0, 1]$. Let $\mathbf{F}$ be the family of all S-measurable functions $f : \Omega \to [0, 1]$ (i.e., $[\alpha, \beta] \subset [0, 1] \Rightarrow f^{-1}([\alpha, \beta]) \in S$). $\mathbf{F}$ is the so called full tribe of fuzzy sets [30] (see also [14,29]); it is*

closed also under the natural product of fuzzy sets and represents a special case of product MV algebras. On the product MV algebra F we define a state m by the formula $m(f) = \int_0^1 f(x)dx$, for every $f \in F$. Evidently, the sets $\mathcal{A} = \{x, 1 - x\}$ and $\mathcal{B} = \{x^2, 1 - x^2\}$ are two partitions of F with the m-state values $\frac{1}{2}$, $\frac{1}{2}$ and $\frac{1}{3}$, $\frac{2}{3}$ of the corresponding elements of $\mathcal{A}$ and $\mathcal{B}$, respectively. By simple calculations we obtain the entropy $H_m(\mathcal{A}) = \log 2 = 1$ bit, and the entropy $H_m(\mathcal{B}) = -\frac{1}{3} \cdot \log \frac{1}{3} - \frac{2}{3} \cdot \log \frac{2}{3} = 0.9183$ bit. The join of $\mathcal{A}$ and $\mathcal{B}$ is the system $\mathcal{A} \vee \mathcal{B} = \{x^3, x^2(1 - x), x(1 - x^2), (1 - x)(1 - x^2)\}$ with the m-state values $\frac{1}{4}$, $\frac{1}{12}$, $\frac{1}{4}$, $\frac{5}{12}$ of the corresponding elements. The entropy of $\mathcal{A} \vee \mathcal{B}$ is the number:

$$H_m(\mathcal{A} \vee \mathcal{B}) = -\frac{1}{4} \cdot \log \frac{1}{4} - \frac{1}{12} \cdot \log \frac{1}{12} - \frac{1}{4} \cdot \log \frac{1}{4} - \frac{5}{12} \cdot \log \frac{5}{12} = 1.8250 \text{ bit.}$$

Since:

$$H_m(\mathcal{A}/\mathcal{B}) = -m\left(x^3\right) \cdot \log \frac{m(x^3)}{m(x^2)} - m\left(x(1 - x^2)\right) \cdot \log \frac{m(x(1 - x^2))}{m(1 - x^2)}$$
$$-m\left((1 - x)\,x^2\right) \cdot \log \frac{m((1 - x)\,x^2)}{m(x^2)} - m\left((1 - x)\,(1 - x^2)\right) \cdot \log \frac{m((1 - x)\,(1 - x^2))}{m(1 - x^2)}$$
$$= -\frac{1}{4} \cdot \log \frac{\frac{1}{4}}{\frac{1}{3}} - \frac{1}{4} \cdot \log \frac{\frac{1}{4}}{\frac{2}{3}} - \frac{1}{12} \cdot \log \frac{\frac{1}{12}}{\frac{1}{3}} - \frac{5}{12} \cdot \log \frac{\frac{5}{12}}{\frac{2}{3}} = 0.9067 \text{ bit,}$$

the mutual information of $\mathcal{A}$ and $\mathcal{B}$ is the number:

$$I_m(\mathcal{A}, \mathcal{B}) = H_m(\mathcal{A}) - H_m(\mathcal{A}/\mathcal{B}) = 1 - 0.9067 = 0.0933 \text{ bit.}$$

We can also see that Equation (3) is fulfilled:

$$H_m(\mathcal{A}) + H_m(\mathcal{B}) - H_m(\mathcal{A} \vee \mathcal{B}) = 1 + 0.9183 - 1.8250 = 0.0933 \text{ bit.}$$

In the following we will use the assertions of Propositions 1 and 2.

Proposition 1. *If $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ are two partitions of $(M, \cdot)$, then we have:*

(i) $m(a_i) = \sum\limits_{j=1}^{k} m(a_i \cdot b_j)$, *for $i = 1, 2, ..., n$;*

(ii) $m(b_j) = \sum\limits_{i=1}^{n} m(a_i \cdot b_j)$, *for $j = 1, 2, ..., k$.*

Proof. By the assumption $\sum\limits_{j=1}^{k} b_j = u$, therefore, according to Definitions 1 and 2, we get:

$$m(a_i) = m(u \cdot a_i) = m\left(\left(\sum_{j=1}^{k} b_j\right) \cdot a_i\right) = m\left(\sum_{j=1}^{k} (b_j \cdot a_i)\right) = \sum_{j=1}^{k} m\,(a_i \cdot b_j), \text{ for } i = 1, 2, ..., n.$$

The equality (ii) could be obtained in the same way. □

From the following proposition it follows that, for every partitions $\mathcal{A}$, $\mathcal{B}$ of $(M, \cdot)$, the set $\mathcal{A} \vee \mathcal{B}$ is a common refinement of $\mathcal{A}$ and $\mathcal{B}$.

Proposition 2. $\mathcal{A} \prec \mathcal{A} \vee \mathcal{B}$, *for every partitions $\mathcal{A}$, $\mathcal{B}$ of $(M, \cdot)$.*

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$. Since the set $\mathcal{A} \vee \mathcal{B}$ is indexed by $\{(i,j); i = 1, ..., n, j = 1, 2, ..., k\}$, we put $I(i) = \{(i,1), ..., (i,k)\}$, $i = 1, 2, ..., n$. In view of Proposition 1, we have:

$$m(a_i) = \sum_{j=1}^{k} m(a_i \cdot b_j) = \sum_{(l,j) \in I(i)} m(a_l \cdot b_j), \text{ for } i = 1, 2, ..., n.$$

However, this indicates that $\mathcal{A} \prec \mathcal{A} \vee \mathcal{B}$. $\square$

Theorem 1. *For any partitions $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ in a product MV algebra $(M, \cdot)$, we have:*

$$I_m(\mathcal{A} \vee \mathcal{B}, \mathcal{C}) \geq I_m(\mathcal{A}, \mathcal{C}).$$

Proof. By Equation (2) and the properties (E3) and (E4), we get:

$$I_m(\mathcal{A} \vee \mathcal{B}, \mathcal{C}) = H_m(\mathcal{A} \vee \mathcal{B}) - H_m(\mathcal{A} \vee \mathcal{B} / \mathcal{C})$$
$$= H_m(\mathcal{A}) + H_m(\mathcal{B}/\mathcal{A}) - H_m(\mathcal{A}/\mathcal{C}) - H_m(\mathcal{B}/\mathcal{C} \vee \mathcal{A})$$
$$= I_m(\mathcal{A}, \mathcal{C}) + H_m(\mathcal{B}/\mathcal{A}) - H_m(\mathcal{B}/\mathcal{C} \vee \mathcal{A}).$$

According to Proposition 2 $\mathcal{A} \prec \mathcal{C} \vee \mathcal{A}$, and therefore by (E2) $H_m(\mathcal{B}/\mathcal{A}) \geq H_m(\mathcal{B}/\mathcal{C} \vee \mathcal{A})$. It follows the inequality:

$$I_m(\mathcal{A} \vee \mathcal{B}, \mathcal{C}) \geq I_m(\mathcal{A}, \mathcal{C}). \square$$

Proposition 3. *If $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ are two partitions of $(M, \cdot)$, then:*

$$I_m(\mathcal{A}, \mathcal{B}) = \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log \frac{m(a_i \cdot b_j)}{m(a_i) \cdot m(b_j)}. \tag{5}$$

Proof. Since by Proposition 1 it holds:

$$m(a_i) = \sum_{j=1}^{k} m(a_i \cdot b_j), \text{ for } i = 1, 2, ..., n,$$

we get:

$$I_m(\mathcal{A}, \mathcal{B}) = -\sum_{i=1}^{n} m(a_i) \cdot \log m(a_i) + \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log \frac{m(a_i \cdot b_j)}{m(b_j)}$$
$$= -\sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log m(a_i) + \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log \frac{m(a_i \cdot b_j)}{m(b_j)}$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \left[\log \frac{m(a_i \cdot b_j)}{m(b_j)} - \log m(a_i) \right] = \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log \frac{m(a_i \cdot b_j)}{m(a_i) \cdot m(b_j)}. \square$$

Definition 5. *Two partitions $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ of $(M, \cdot)$ are called statistically independent, if $m(a_i \cdot b_j) = m(a_i) \cdot m(b_j)$, for $i = 1, 2, ..., n, j = 1, 2, ..., k$.*

Theorem 2. *Let $\mathcal{A}$, $\mathcal{B}$ be partitions in a product MV algebra $(M, \cdot)$. Then $I_m(\mathcal{A}, \mathcal{B}) \geq 0$ with the equality if and only if the partitions $\mathcal{A}$, $\mathcal{B}$ are statistically independent.*

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$. Then using the inequality $\log x \leq x - 1$, which is valid for all real numbers $x > 0$, with the equality if and only if $x = 1$, we get:

$$m(a_i \cdot b_j) \cdot \log \frac{m(a_i) \cdot m(b_j)}{m(a_i \cdot b_j)} \leq m(a_i \cdot b_j) \cdot \left[\frac{m(a_i) \cdot m(b_j)}{m(a_i \cdot b_j)} - 1 \right] = m(a_i) \cdot m(b_j) - m(a_i \cdot b_j).$$

The equality holds if and only if $\frac{m(a_i) \cdot m(b_j)}{m(a_i \cdot b_j)} = 1$, i.e., when $m(a_i \cdot b_j) = m(a_i) \cdot m(b_j)$. Therefore using Equation (5) and Proposition 1 we have:

$$-I_m(\mathcal{A}, \mathcal{B}) = \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) \cdot \log \frac{m(a_i) \cdot m(b_j)}{m(a_i \cdot b_j)} \leq \sum_{i=1}^{n} \sum_{j=1}^{k} \left[m(a_i) \cdot m(b_j) - m(a_i \cdot b_j) \right]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i) \cdot m(b_j) - \sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i \cdot b_j) = \sum_{i=1}^{n} m(a_i) \cdot \sum_{j=1}^{k} m(b_j) - \sum_{i=1}^{n} m(a_i)$$

$$= m(\sum_{i=1}^{n} a_i) \cdot m(\sum_{j=1}^{k} b_j) - m(\sum_{i=1}^{n} a_i) = m(u) \cdot m(u) - m(u) = 1 \cdot 1 - 1 = 0.$$

It follows that $I_m(\mathcal{A}, \mathcal{B}) \geq 0$ with the equality if and only if $m(a_i \cdot b_j) = m(a_i) \cdot m(b_j)$, for $i = 1, 2, ..., n, j = 1, 2, ..., k$, i.e., when the partitions $\mathcal{A}$, $\mathcal{B}$ are statistically independent. $\square$

From Theorem 2 it follows subadditivity and additivity of entropy in a product MV algebra, as shown by the following theorem.

Theorem 3 (*Subadditivity and additivity of entropy*). *For arbitrary partitions $\mathcal{A}$, $\mathcal{B}$ in a product MV algebra $(M, \cdot)$, it holds $H_m(\mathcal{A} \vee \mathcal{B}) \leq H_m(\mathcal{A}) + H_m(\mathcal{B})$ with the equality if and only if the partitions $\mathcal{A}$, $\mathcal{B}$ are statistically independent.*

Proof. It follows by Equation (3) and Theorem 2. $\square$

Theorem 4. *For arbitrary partitions $\mathcal{A}$, $\mathcal{B}$ in a product MV algebra $(M, \cdot)$, it holds $H_m(\mathcal{A}/\mathcal{B}) \leq H_m(\mathcal{A})$ with the equality if and only if the partitions $\mathcal{A}$, $\mathcal{B}$ are statistically independent.*

Proof. The assertion is a simple consequence of Equation (2) and Theorem 2. $\square$

Definition 6. *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be partitions in a given product MV algebra $(M, \cdot)$. Then the conditional mutual information of $\mathcal{A}$ and $\mathcal{B}$ given $\mathcal{C}$ is defined by the formula*

$$I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}) = H_m(\mathcal{A} / \mathcal{C}) - H_m(\mathcal{A} / \mathcal{B} \vee \mathcal{C}). \tag{6}$$

Remark 2. *Notice that the conditional mutual information is nonnegative, because by the property (E2) $H_m(\mathcal{A} / \mathcal{C}) \geq H_m(\mathcal{A} / \mathcal{B} \vee \mathcal{C})$.*

Theorem 5. *For any partitions $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ in a product MV algebra $(M, \cdot)$, we have:*

$$I_m(\mathcal{A}, \mathcal{B} \vee \mathcal{C}) = I_m(\mathcal{A}, \mathcal{C}) + I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}) = I_m(\mathcal{A}, \mathcal{B}) + I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}).$$

Proof. Let us calculate:

$$I_m(\mathcal{A}, \mathcal{C}) + I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}) = H_m(\mathcal{A}) - H_m(\mathcal{A} / \mathcal{C}) + H_m(\mathcal{A} / \mathcal{C}) - H_m(\mathcal{A} / \mathcal{B} \vee \mathcal{C})$$
$$= H_m(\mathcal{A}) - H_m(\mathcal{A} / \mathcal{B} \vee \mathcal{C}) = I_m(\mathcal{A}, \mathcal{B} \vee \mathcal{C}).$$

In a similar way we obtain also the second equality. □

Theorem 6 (*Chain rules*). *Let* $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ *and* $\mathcal{C}$ *be partitions in a product MV algebra* $(M, \cdot)$. *Then, for* $n = 2, 3, ...,$ *the following equalities hold:*

(i) $H_m(\mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n) = H_m(\mathcal{A}_1) + \sum_{i=2}^n H_m(\mathcal{A}_i / \vee_{k=1}^{i-1} \mathcal{A}_k);$

(ii) $H_m(\vee_{i=1}^n \mathcal{A}_i / \mathcal{C}) = H_m(\mathcal{A}_1 / \mathcal{C}) + \sum_{i=2}^n H_m(\mathcal{A}_i / (\vee_{k=1}^{i-1} \mathcal{A}_k) \vee \mathcal{C});$

(iii) $I_m(\vee_{i=1}^n \mathcal{A}_i, \mathcal{C}) = I_m(\mathcal{A}_1, \mathcal{C}) + \sum_{i=2}^n I_m(\mathcal{A}_i, \mathcal{C} / \vee_{k=1}^{i-1} \mathcal{A}_k).$

Proof. (i) By the property (E4) we have:

$$H_m(\mathcal{A}_1 \vee \mathcal{A}_2) = H_m(\mathcal{A}_1) + H_m(\mathcal{A}_2 / \mathcal{A}_1).$$

By (E3) and (E4) we get:

$$\begin{aligned} H_m(\mathcal{A}_1 \vee \mathcal{A}_2 \vee \mathcal{A}_3) &= H_m(\mathcal{A}_1) + H_m(\mathcal{A}_2 \vee \mathcal{A}_3 / \mathcal{A}_1) \\ &= H_m(\mathcal{A}_1) + H_m(\mathcal{A}_2 / \mathcal{A}_1) + H_m(\mathcal{A}_3 / \mathcal{A}_2 \vee \mathcal{A}_1) \\ &= H_m(\mathcal{A}_1) + \sum_{i=2}^3 H_m(\mathcal{A}_i / \vee_{k=1}^{i-1} \mathcal{A}_k). \end{aligned}$$

Now let us suppose that the result is true for a given $n \in \mathbb{N}$. Then:

$$\begin{aligned} H_m(\mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n \vee \mathcal{A}_{n+1}) &= H_m(\mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n) + H_m(\mathcal{A}_{n+1} / \mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n) \\ &= H_m(\mathcal{A}_1) + \sum_{i=2}^n H_m(\mathcal{A}_i / \vee_{k=1}^{i-1} \mathcal{A}_k) + H_m(\mathcal{A}_{n+1} / \mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n) \\ &= H_m(\mathcal{A}_1) + \sum_{i=2}^{n+1} H_m(\mathcal{A}_i / \vee_{k=1}^{i-1} \mathcal{A}_k). \end{aligned}$$

(ii) For $n = 2$, using (E3) we obtain:

$$H_m(\mathcal{A}_1 \vee \mathcal{A}_2 / \mathcal{C}) = H_m(\mathcal{A}_1 / \mathcal{C}) + H_m(\mathcal{A}_2 / \mathcal{A}_1 \vee \mathcal{C}).$$

Suppose that the result is true for a given $n \in \mathbb{N}$. Then:

$$\begin{aligned} H_m(\mathcal{A}_1 \vee \mathcal{A}_2 \vee ... \vee \mathcal{A}_n \vee \mathcal{A}_{n+1} / \mathcal{C}) &= H_m(\vee_{i=1}^n \mathcal{A}_i / \mathcal{C}) + H_m(\mathcal{A}_{n+1} / \mathcal{A}_1 \vee ... \vee \mathcal{A}_n \vee \mathcal{C}) \\ &= H_m(\mathcal{A}_1 / \mathcal{C}) + \sum_{i=2}^n H_m(\mathcal{A}_i / (\vee_{k=1}^{i-1} \mathcal{A}_k) \vee \mathcal{C}) + H_m(\mathcal{A}_{n+1} / (\vee_{k=1}^n \mathcal{A}_k) \vee \mathcal{C}) \\ &= H_m(\mathcal{A}_1 / \mathcal{C}) + \sum_{i=2}^{n+1} H_m(\mathcal{A}_i / (\vee_{k=1}^{i-1} \mathcal{A}_k) \vee \mathcal{C}). \end{aligned}$$

(iii) By Equation (2), the equalities (i) and (ii) of this theorem, and Equation (6), we obtain:

$$\begin{aligned} I_m(\vee_{i=1}^n \mathcal{A}_i, \mathcal{C}) &= H_m(\vee_{i=1}^n \mathcal{A}_i) - H_m(\vee_{i=1}^n \mathcal{A}_i / \mathcal{C}) \\ &= H_m(\mathcal{A}_1) + \sum_{i=2}^n H_m(\mathcal{A}_i / \vee_{k=1}^{i-1} \mathcal{A}_k) - H_m(\mathcal{A}_1 / \mathcal{C}) - \sum_{i=2}^n H_m(\mathcal{A}_i / (\vee_{k=1}^{i-1} \mathcal{A}_k) \vee \mathcal{C}) \\ &= I_m(\mathcal{A}_1, \mathcal{C}) + \sum_{i=2}^n \left(H_m(\mathcal{A}_i / \vee_{k=1}^{i-1} \mathcal{A}_k) - H_m(\mathcal{A}_i / (\vee_{k=1}^{i-1} \mathcal{A}_k) \vee \mathcal{C}) \right) \\ &= I_m(\mathcal{A}_1, \mathcal{C}) + \sum_{i=2}^n I_m(\mathcal{A}_i, \mathcal{C} / \vee_{k=1}^{i-1} \mathcal{A}_k). \quad \square \end{aligned}$$

Definition 7. *Let* $\mathcal{A}$, $\mathcal{B}$ *and* $\mathcal{C}$ *be partitions in a product MV algebra* $(M, \cdot)$. *We say that* $\mathcal{A}$ *is conditionally independent to* $\mathcal{C}$ *given* $\mathcal{B}$ *(and write* $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$ *) if* $I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}) = 0$.

Theorem 7. *For partitions* $\mathcal{A}$, $\mathcal{B}$ *and* $\mathcal{C}$ *in a product MV algebra* $(M, \cdot)$, $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$ *if and only if* $\mathcal{C} \to \mathcal{B} \to \mathcal{A}$.

Proof. Let $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$. Then $0 = I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}) = H_m(\mathcal{A} / \mathcal{B}) - H_m(\mathcal{A} / \mathcal{B} \vee \mathcal{C})$. Therefore by (E4) we get:

$$H_m(\mathcal{A} / \mathcal{B}) = H_m(\mathcal{A} / \mathcal{B} \vee \mathcal{C}) = H_m(\mathcal{A} \vee \mathcal{B} \vee \mathcal{C}) - H_m(\mathcal{B} \vee \mathcal{C}).$$

Let us calculate:

$$I_m(\mathcal{C}, \mathcal{A} / \mathcal{B}) = H_m(\mathcal{C} / \mathcal{B}) - H_m(\mathcal{C} / \mathcal{A} \vee \mathcal{B})$$
$$= H_m(\mathcal{C} \vee \mathcal{B}) - H_m(\mathcal{B}) - H_m(\mathcal{A} \vee \mathcal{B} \vee \mathcal{C}) + H_m(\mathcal{A} \vee \mathcal{B})$$
$$= H_m(\mathcal{A} \vee \mathcal{B}) - H_m(\mathcal{B}) - H_m(\mathcal{A} / \mathcal{B}) = H_m(\mathcal{A} / \mathcal{B}) - H_m(\mathcal{A} / \mathcal{B}) = 0.$$

The results means that $\mathcal{C} \to \mathcal{B} \to \mathcal{A}$. The reverse implication is evident. $\square$

Remark 3. *According to the above theorem, we may say that $\mathcal{A}$ and $\mathcal{C}$ are conditionally independent given $\mathcal{B}$ and write $\mathcal{A} \leftrightarrow \mathcal{B} \leftrightarrow \mathcal{C}$ instead of $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$.*

Theorem 8. *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be partitions in a given product MV algebra $(M, \cdot)$ such that $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$. Then we have*:

(i) $I_m(\mathcal{A} \vee \mathcal{B}, \mathcal{C}) = I_m(\mathcal{B}, \mathcal{C})$;
(ii) $I_m(\mathcal{B}, \mathcal{C}) = I_m(\mathcal{C}, \mathcal{A}) + I_m(\mathcal{C}, \mathcal{B} / \mathcal{A})$;
(iii) $I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}) \leq I_m(\mathcal{A}, \mathcal{B})$;
(iv) $I_m(\mathcal{A}, \mathcal{B}) \geq I_m(\mathcal{A}, \mathcal{C})$ *(data processing inequality)*.

Proof. (i) By the assumption we have $I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}) = 0$. Hence using the chain rule for the mutual information (Theorem 6 (iii)), we obtain:

$$I_m(\mathcal{A} \vee \mathcal{B}, \mathcal{C}) = I_m(\mathcal{B} \vee \mathcal{A}, \mathcal{C}) = I_m(\mathcal{B}, \mathcal{C}) + I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}) = I_m(\mathcal{B}, \mathcal{C}).$$

(ii) By the equality (i) of this theorem and Theorem 5, we can write:

$$I_m(\mathcal{B}, \mathcal{C}) = I_m(\mathcal{A} \vee \mathcal{B}, \mathcal{C}) = I_m(\mathcal{C}, \mathcal{B} \vee \mathcal{A}) = I_m(\mathcal{C}, \mathcal{A}) + I_m(\mathcal{C}, \mathcal{B} / \mathcal{A}).$$

(iii) From (ii) it follows the inequality $I_m(\mathcal{B}, \mathcal{C}) \geq I_m(\mathcal{C}, \mathcal{B} / \mathcal{A})$. Interchanging $\mathcal{A}$ and $\mathcal{C}$ (we can do it based on Theorem 7) we obtain:

$$I_m(\mathcal{A}, \mathcal{B}) \geq I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}).$$

(iv) By the assumption we have $I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}) = 0$. Therefore by Theorem 5 we get:

$$I_m(\mathcal{A}, \mathcal{B} \vee \mathcal{C}) = I_m(\mathcal{A}, \mathcal{B}) + I_m(\mathcal{A}, \mathcal{C} / \mathcal{B}) = I_m(\mathcal{A}, \mathcal{B}).$$

Thus by the same theorem we can write:

$$I_m(\mathcal{A}, \mathcal{B}) = I_m(\mathcal{A}, \mathcal{B} \vee \mathcal{C}) = I_m(\mathcal{A}, \mathcal{C}) + I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}).$$

Since $I_m(\mathcal{A}, \mathcal{B} / \mathcal{C}) \geq 0$, it holds $I_m(\mathcal{A}, \mathcal{B}) \geq I_m(\mathcal{A}, \mathcal{C})$. $\square$

In the following, a concavity of entropy $H_m(\mathcal{A})$ and concavity of mutual information $I_m(\mathcal{A}, \mathcal{B})$ as functions of m are studied. We recall, for the convenience of the reader, the definitions of convex and concave function:

A real-valued function f is said to be convex over an interval $[a, b]$ if for every $x_1, x_2 \in [a, b]$ and for any real number $\alpha \in [0, 1]$:

$$f\left(\alpha x_1 + (1 - \alpha)x_2\right) \leq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

A real-valued function f is said to be concave over an interval $[a, b]$ if for every $x_1, x_2 \in [a, b]$ and for any real number $\alpha \in [0, 1]$:

$$f\left(\alpha x_1 + (1 - \alpha)x_2\right) \geq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

In the following, we will use the symbol F to denote the family of all states on a given product MV algebra $(M, \cdot)$. It is easy to prove the following proposition:

Proposition 4. *If $m_1, m_2 \in F$, then, for every real number $\alpha \in [0, 1]$, $\alpha m_1 + (1 - \alpha)m_2 \in F$.*

Theorem 9 (*Concavity of entropy*). *Let $\mathcal{A}$ be a partition in a given product MV algebra $(M, \cdot)$. Then, for every $m_1, m_2 \in F$, and every real number $\alpha \in [0, 1]$, the following inequality holds:*

$$\alpha\, H_{m_1}(\mathcal{A}) + (1 - \alpha)H_{m_2}(\mathcal{A}) \leq H_{\alpha\, m_1 + (1 - \alpha)m_2}(\mathcal{A}).$$

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$. Since the function F is convex, we get:

$$\alpha\, H_{m_1}(\mathcal{A}) + (1 - \alpha)\, H_{m_2}(\mathcal{A}) = -\alpha \sum_{i=1}^{n} F(m_1(a_i)) - (1 - \alpha) \sum_{i=1}^{n} F(m_2(a_i))$$

$$= -\sum_{i=1}^{n} \left(\alpha\, F(m_1(a_i)) + (1 - \alpha)\, F(m_2(a_i))\right) \leq -\sum_{i=1}^{n} F(\alpha\, m_1(a_i) + (1 - \alpha)m_2(a_i))$$

$$= -\sum_{i=1}^{n} F((\alpha\, m_1 + (1 - \alpha)m_2)(a_i)) = H_{\alpha\, m_1 + (1 - \alpha)m_2}(\mathcal{A}),$$

which proves that the entropy $m \mapsto H_m(\mathcal{A})$ is a concave function on the family F. $\square$

In the proof of concavity of mutual information $I_m(\mathcal{A}, \mathcal{B})$ we will need the assertion of Proposition 5. First, we introduce the following notation. Let m be a state on a product MV algebra $(M, \cdot)$, $a, b \in M$. Then we denote:

$$\dot{m}(a/b) = \begin{cases} \frac{m(a \cdot b)}{m(b)}, & \text{if } m(b) > 0; \\ 0, & \text{if } m(b) = 0. \end{cases}$$

Proposition 5. *If $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$ are two partitions of $(M, \cdot)$, then*

$$H_m(\mathcal{B}\,/\mathcal{A}) = -\sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i) \cdot F(\dot{m}(b_j/a_i)). \tag{7}$$

Proof. Let us calculate:

$$-\sum_{i=1}^{n} \sum_{j=1}^{k} m(a_i) \cdot F(\dot{m}(b_j/a_i)) = -\sum_{i:\, m(a_i)>0} \sum_{j=1}^{k} m(a_i) \cdot F\left(\frac{m(b_j \cdot a_i)}{m(a_i)}\right)$$

$$= -\sum_{i:\, m(a_i)>0} \sum_{j=1}^{k} m(a_i) \cdot \frac{m(b_j \cdot a_i)}{m(a_i)} \cdot \log \frac{m(b_j \cdot a_i)}{m(a_i)} = -\sum_{i:\, m(a_i)>0} \sum_{j=1}^{k} m(b_j \cdot a_i) \cdot \log \frac{m(b_j \cdot a_i)}{m(a_i)}$$

$$= -\sum_{i=1}^{n} \sum_{j=1}^{k} m(b_j \cdot a_i) \cdot \log \frac{m(b_j \cdot a_i)}{m(a_i)} = H_m(\mathcal{B}\,/\mathcal{A}).$$

In the last step, we used the implication $m(a_i) = 0 \Rightarrow m(b_j \cdot a_i) = 0$ which follows from the equality $m(a_i) = \sum_{j=1}^{k} m(a_i \cdot b_j)$ shown in Proposition 1. $\square$

Remark 4. *By Proposition 5 there exists $c_{ij} = -F(\dot{m}(b_j/a_i)) \geq 0$ such that*

$$H_m(\mathcal{B}\,/\mathcal{A}) = \sum_{i=1}^{n}\sum_{j=1}^{k} c_{ij} \cdot m(a_i).$$

Definition 8. *Let $\mathcal{A} = \{a_1, ..., a_n\}$, $\mathcal{B} = \{b_1, ..., b_k\}$ be two partitions of $(M, \cdot)$. Put*

$$K = \left\{ m \in F;\ H_m(\mathcal{B}\,/\mathcal{A}) = \sum_{i=1}^{n}\sum_{j=1}^{k} c_{ij} \cdot m(a_i) \right\}.$$

Theorem 10 *(Concavity of mutual information). The mutual information $m \mapsto I_m(\mathcal{A}, \mathcal{B})$ is a concave function on the family $\mathcal{K}$.*

Proof. By Equation (4) we can write:

$$I_m(\mathcal{A}, \mathcal{B}) = H_m(\mathcal{B}) - H_m(\mathcal{B}\,/\mathcal{A}).$$

In view of Theorem 9 and Remark 4, the function $m \mapsto I_m(\mathcal{A}, \mathcal{B})$ is the sum of two concave functions on the family $\mathcal{K}$: $m \mapsto H_m(\mathcal{B})$, and $m \mapsto -H_m(\mathcal{B}\,/\mathcal{A})$. Since the sum of two concave functions is itself concave, we have the statement. $\square$

4. Kullback–Leibler Divergence in Product MV Algebras

In this section we introduce the concept of Kullback–Leibler divergence in product MV algebras. We prove basic properties of this measure; in particular, Gibb's inequality. Finally, using the notion of conditional Kullback–Leibler divergence we establish a chain rule for Kullback–Leibler divergence with respect to additive states defined on a given product MV algebra. In the proofs we use the following known log-sum inequality: for non-negative real numbers $x_1, x_2, ..., x_n, y_1, y_2, ..., y_n$, it holds:

$$\sum_{i=1}^{n} x_i \cdot \log \frac{x_i}{y_i} \geq \left(\sum_{i=1}^{n} x_i \right) \cdot \log \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} \tag{8}$$

with the equality if and only if $\frac{x_i}{y_i}$ is constant. Recall that we use the convention that $x \log \frac{x}{0} = \infty$ if $x > 0$, and $0 \log \frac{0}{x} = 0$ if $x \geq 0$.

Definition 9. *Let m_1, m_2 be states defined on a given product MV algebra $(M, \cdot)$, and $\mathcal{A} = \{a_1, ..., a_n\}$ be a partition of $(M, \cdot)$. Then we define the Kullback–Leibler divergence $D_{\mathcal{A}}(m_1 \parallel m_2)$ by:*

$$D_{\mathcal{A}}(m_1 \parallel m_2) = \sum_{i=1}^{n} m_1(a_i) \cdot \log \frac{m_1(a_i)}{m_2(a_i)}.$$

Remark 5. *It is obvious that $D_{\mathcal{A}}(m \parallel m) = 0$. The Kullback–Leibler divergence is not a metric in a true sense since it is not symmetric, i.e., the equality $D_{\mathcal{A}}(m_1 \parallel m_2) = D_{\mathcal{A}}(m_2 \parallel m_1)$ is not necessarily true (as shown in the following example), and does not satisfy the triangle inequality.*

Example 3. *Consider any product MV algebra $(M, \cdot)$ and two states m_1, m_2 defined on M. Let $a \in M$ such that $m_1(a) = p$, and $m_2(a) = q$, where $p, q \in (0, 1)$. Evidently, $m_1(u - a) = 1 - p$, $m_2(u - a) = 1 - q$, and the set $\mathcal{A} = \{a, u - a\}$ is a partition of $(M, \cdot)$. Let us calculate:*

$$D_{\mathcal{A}}(m_1 \| m_2) = p \cdot \log \frac{p}{q} + (1 - p) \cdot \log \frac{1 - p}{1 - q}, \text{ and } D_{\mathcal{A}}(m_2 \| m_1) = q \cdot \log \frac{q}{p} + (1 - q) \cdot \log \frac{1 - q}{1 - p}.$$

If $p = q$, then $D_{\mathcal{A}}(m_1 \| m_2) = D_{\mathcal{A}}(m_2 \| m_1) = 0$. If $p = \frac{1}{2}$, $q = \frac{1}{4}$, then we have:

$$D_{\mathcal{A}}(m_1 \| m_2) = \frac{1}{2} \cdot \log \frac{\frac{1}{2}}{\frac{1}{4}} + \frac{1}{2} \cdot \log \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{1}{2} \cdot \log 2 + \frac{1}{2} \cdot \log \frac{2}{3} = 0.\,207519 \ bit,$$

and:

$$D_{\mathcal{A}}(m_2 \| m_1) = \frac{1}{4} \cdot \log \frac{\frac{1}{4}}{\frac{1}{2}} + \frac{3}{4} \cdot \log \frac{\frac{3}{4}}{\frac{1}{2}} = \frac{1}{4} \cdot \log 2 + \frac{3}{4} \cdot \log \frac{3}{2} = 0.188722 \ bit.$$

The result means that $D_{\mathcal{A}}(m_1 \| m_2) \neq D_{\mathcal{A}}(m_2 \| m_1)$, in general.

Theorem 11. *Let m_1, m_2 be states defined on a product MV algebra $(M, \cdot)$, and $\mathcal{A} = \{a_1, ..., a_n\}$ be a partition of $(M, \cdot)$. Then $D_{\mathcal{A}}(m_1 \| m_2) \geq 0$ (Gibb's inequality) with the equality if and only if $m_1(a_i) = m_2(a_i)$, for $i = 1, 2, ..., n$.*

Proof. If we put $x_i = m_1(a_i)$ and $y_i = m_2(a_i)$, for $i = 1, 2, ..., n$, then $x_1, x_2, ..., x_n, y_1, y_2, ..., y_n$ are non-negative real numbers such that $\sum_{i=1}^{n} x_i = 1$ and $\sum_{i=1}^{n} y_i = 1$. Indeed, $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} m_1(a_i) = m_1(\sum_{i=1}^{n} a_i) = m_1(u) = 1$; analogously we obtain $\sum_{i=1}^{n} y_i = 1$. Thus, using the log-sum inequality we can write:

$$D_{\mathcal{A}}(m_1 \| m_2) = \sum_{i=1}^{n} m_1(a_i) \cdot \log \frac{m_1(a_i)}{m_2(a_i)} = \sum_{i=1}^{n} x_i \cdot \log \frac{x_i}{y_i}$$

$$\geq \left(\sum_{i=1}^{n} x_i \right) \cdot \log \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} = 1 \cdot \log \frac{1}{1} = 0$$

with the equality if and only if $\frac{m_1(a_i)}{m_2(a_i)} = \alpha$ for $i = 1, 2, ..., n$, where α is constant. Taking the sum for all $i = 1, 2, ..., n$, we obtain $\sum_{i=1}^{n} m_1(a_i) = \alpha \sum_{i=1}^{n} m_2(a_i)$, which implies that $\alpha = 1$. This means that $D_{\mathcal{A}}(m_1 \| m_2) = 0$ if and only if $m_1(a_i) = m_2(a_i)$, for $i = 1, 2, ..., n$. □

Theorem 12. *Let $\mathcal{A}$ be a partition of $(M, \cdot)$ and ν be a state on $(M, \cdot)$ uniform over $\mathcal{A}$. Then, for the entropy of $\mathcal{A}$ with respect to any state m from F, we have*:

$$H_m(\mathcal{A}) = \log card \mathcal{A} - D_{\mathcal{A}}(m \| \nu).$$

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$. Then $\nu(a_i) = \frac{1}{n}$, for $i = 1, 2, ..., n$. Let us calculate:

$$D_{\mathcal{A}}(m \| \nu) = \sum_{i=1}^{n} m(a_i) \cdot \log \frac{m(a_i)}{\nu(a_i)} = \sum_{i=1}^{n} m(a_i) \cdot \log \frac{m(a_i)}{\frac{1}{n}}$$

$$= \sum_{i=1}^{n} m(a_i) \cdot \left(\log m(a_i) - \log n^{-1} \right) = \sum_{i=1}^{n} m(a_i) \cdot \log m(a_i) + \log n = \log card \mathcal{A} - H_m(\mathcal{A}). \ \square$$

As a consequence we obtain the following property of entropy of partitions in product MV algebras.

Corollary 1. *For any partition $\mathcal{A}$ of $(M, \cdot)$, it holds $H_m(\mathcal{A}) \leq \log card \mathcal{A}$, with the equality if and only if m is uniform over the partition $\mathcal{A}$.*

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$ and consider a state ν on $(M, \cdot)$ uniform over $\mathcal{A}$, i.e., it holds $\nu(a_i) = \frac{1}{n}$, for $i = 1, 2, ..., n$. Then, by Theorem 12 we get:

$$D_{\mathcal{A}}(m\|\nu) = \log card\mathcal{A} - H_m(\mathcal{A}).$$

Since by Theorem 11 $D_{\mathcal{A}}(m \| \nu) \geq 0$, it holds the inequality:

$$H_m(\mathcal{A}) \leq \log card\mathcal{A}.$$

Further, by Theorem 11 $D_{\mathcal{A}}(m \| \nu) = 0$ if and only if $m(a_i) = \nu(a_i)$, for $i = 1, 2, ..., n$. This means that the equality $H_m(\mathcal{A}) = \log card\mathcal{A}$ holds if and only if $m(a_i) = \frac{1}{n}$, for $i = 1, 2, ..., n$. $\square$

Theorem 13 *(Convexity of K–L divergence). Let $\mathcal{A}$ be a partition in a product MV algebra $(M, \cdot)$. The K–L divergence $D_{\mathcal{A}}(m_1 \| m_2)$ is convex in the pair (m_1, m_2), i.e., if $(m'_1, m'_2), (m''_1, m''_2)$ are pairs of states from F, then, for any real number $\alpha \in [0, 1]$, the following inequality holds:*

$$D_{\mathcal{A}}(\alpha m'_1 + (1 - \alpha)m''_1 \| \alpha m'_2 + (1 - \alpha)m''_2) \leq \alpha D_{\mathcal{A}}(m'_1\|m'_2) + (1 - \alpha)D_{\mathcal{A}}(m''_1\|m''_2). \tag{9}$$

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$ and fix $i \in \{1, 2, ..., n\}$. Putting $x_1 = \alpha m'_1(a_i)$, $x_2 = (1 - \alpha)m''_1(a_i)$, $y_1 = \alpha m'_2(a_i)$, $y_2 = (1 - \alpha)m''_2(a_i)$ in the log-sum inequality, we obtain:

$$(\alpha m'_1(a_i) + (1 - \alpha)m''_1(a_i)) \cdot \log \frac{\alpha m'_1(a_i) + (1 - \alpha)m''_1(a_i)}{\alpha m'_2(a_i) + (1 - \alpha)m''_2(a_i)}$$

$$\leq \alpha m'_1(a_i) \cdot \log \frac{\alpha m'_1(a_i)}{\alpha m'_2(a_i)} + (1 - \alpha)m''_1(a_i) \cdot \log \frac{(1 - \alpha)m''_1(a_i)}{(1 - \alpha)m''_2(a_i)}.$$

Summing these inequalities over $i = 1, 2, ..., n$, we obtain the inequality (9). $\square$

The result of Theorem 13 is illustrated in the following example.

Example 4. *Consider the product MV algebra F from Example 2 and the real functions F_1, F_2, F_3, F_4 defined by $F_1(x) = x$, $F_2(x) = x^2$, $F_3(x) = x^3$, $F_4(x) = x^4$, for every $x \in \Re$. On the product MV algebra **F** we define the states m_1, m_2, m_3, m_4 by the following formulas:*

$$m_1(f) = \int_0^1 f(x)dF_1(x) = \int_0^1 f(x)dx, f \in F;$$

$$m_2(f) = \int_0^1 f(x)dF_2(x) = \int_0^1 f(x)2xdx, f \in F;$$

$$m_3(f) = \int_0^1 f(x)dF_3(x) = \int_0^1 f(x)3x^2dx, f \in F;$$

$$m_4(f) = \int_0^1 f(x)dF_4(x) = \int_0^1 f(x)4x^3dx, f \in F.$$

In addition, we will consider the partition $\mathcal{A} = \{x, 1 - x\}$ of F. It is easy to calculate that it has the m_1-state values $\frac{1}{2}, \frac{1}{2}$; the m_2-state values $\frac{2}{3}, \frac{1}{3}$; the m_3-state values $\frac{3}{4}, \frac{1}{4}$; and the m_4-state values $\frac{4}{5}, \frac{1}{5}$ of the corresponding elements. In the previous theorem we put $\alpha = 0.2$. We will show that:

$$D_{\mathcal{A}}(0.2m_1 + 0.8m_3 \| 0.2m_2 + 0.8m_4) \leq 0.2D_{\mathcal{A}}(m_1\|m_2) + 0.8D_{\mathcal{A}}(m_3\|m_4). \tag{10}$$

Let us calculate:

$$D_{\mathcal{A}}(m_1 \| m_2) = \frac{1}{2} \cdot \log \frac{\frac{1}{2}}{\frac{2}{3}} + \frac{1}{2} \cdot \log \frac{\frac{1}{2}}{\frac{1}{3}} = 0.085 \ bit;$$

$$D_{\mathcal{A}}(m_3 \| m_4) = \frac{3}{4} \cdot \log \frac{\frac{3}{4}}{\frac{4}{5}} + \frac{1}{4} \cdot \log \frac{\frac{1}{4}}{\frac{1}{5}} = 0.01065 \ bit;$$

$$D_{\mathcal{A}}(0.2m_1 + 0.8m_3 \| 0.2m_2 + 0.8m_4) = 0.7 \cdot \log \frac{0.7}{0.7733} + 0.3 \cdot \log \frac{0.3}{0.2267} = 0.020682 \ bit.$$

Since $0.020682 \leq 0.2 \cdot 0.085 + 0.8 \cdot 0.01065 = 0.02552$, *the inequality (10) holds.*

In the final part, we define the conditional Kullback–Leibler divergence and, using this notion, we establish the chain rule for Kullback–Leibler divergence.

Definition 10. *Let m_1, m_2 be states on a given product MV algebra $(M, \cdot)$ and $\mathcal{A} = \{a_1, ..., a_n\}$, $\mathcal{B} = \{b_1, ..., b_k\}$ be two partitions of $(M, \cdot)$. Then we define the conditional Kullback–Leibler divergence $D_{\mathcal{B}/\mathcal{A}}(m_1 \| m_2)$ by:*

$$D_{\mathcal{B}/\mathcal{A}}(m_1 \| m_2) = \sum_{i=1}^{n} m_1(a_i) \sum_{j=1}^{k} \dot{m}_1(b_j/a_i) \cdot \log \frac{\dot{m}_1(b_j/a_i)}{\dot{m}_2(b_j/a_i)}.$$

Theorem 14 *(Chain rule for K–L divergence). Let m_1, m_2 be states on a given product MV algebra $(M, \cdot)$. If $\mathcal{A}$, $\mathcal{B}$ are two partitions of $(M, \cdot)$, then:*

$$D_{\mathcal{A} \vee \mathcal{B}}(m_1 \| m_2) = D_{\mathcal{A}}(m_1 \| m_2) + D_{\mathcal{B}/\mathcal{A}}(m_1 \| m_2). \tag{11}$$

Proof. Assume that $\mathcal{A} = \{a_1, ..., a_n\}$ and $\mathcal{B} = \{b_1, ..., b_k\}$. We will consider the following two cases: (i) there exists $i_0 \in \{1, ..., n\}$ such that $m_2(a_{i_0}) = 0$; (ii) $m_2(a_i) > 0$ for $i = 1, 2, ..., n$. In the first case, both sides of Equation (11) are equal to ∞, thus the equality holds. Let us now assume that $m_2(a_i) > 0$, for $i = 1, 2, ..., n$. We get:

$$D_{\mathcal{A}}(m_1 \| m_2) + D_{\mathcal{B}/\mathcal{A}}(m_1 \| m_2)$$

$$= \sum_{i=1}^{n} m_1(a_i) \cdot \log \frac{m_1(a_i)}{m_2(a_i)} + \sum_{i=1}^{n} m_1(a_i) \sum_{j=1}^{k} \dot{m}_1(b_j/a_i) \cdot \log \frac{\dot{m}_1(b_j/a_i)}{\dot{m}_2(b_j/a_i)}$$

$$= \sum_{i:m_1(a_i)>0} \sum_{j=1}^{k} m_1(a_i \cdot b_j) \cdot \log \frac{m_1(a_i)}{m_2(a_i)} + \sum_{i:m_1(a_i)>0} \sum_{j=1}^{k} m_1(a_i \cdot b_j) \cdot \log \frac{\dot{m}_1(b_j/a_i)}{\dot{m}_2(b_j/a_i)}$$

$$= \sum_{i:m_1(a_i)>0} \sum_{j=1}^{k} m_1(a_i \cdot b_j) \cdot \left(\log \frac{m_1(a_i)}{m_2(a_i)} + \log \frac{\dot{m}_1(b_j/a_i)}{\dot{m}_2(b_j/a_i)} \right)$$

$$= \sum_{i:m_1(a_i)>0} \sum_{j=1}^{k} m_1(a_i \cdot b_j) \cdot \log \frac{m_1(a_i) \dot{m}_1(b_j/a_i)}{m_2(a_i) \dot{m}_2(b_j/a_i)} = \sum_{i:m_1(a_i)>0} \sum_{j=1}^{k} m_1(a_i \cdot b_j) \cdot \log \frac{m_1(a_i \cdot b_j)}{m_2(a_i \cdot b_j)}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{k} m_1(a_i \cdot b_j) \cdot \log \frac{m_1(a_i \cdot b_j)}{m_2(a_i \cdot b_j)} = D_{\mathcal{A} \vee \mathcal{B}}(m_1 \| m_2).$$

In the last step, analogously as in the proof of Proposition 5, we used the implication $m_1(a_i) = 0 \Rightarrow m_1(a_i \cdot b_j) = 0$ which follows from the equality $m_1(a_i) = \sum_{j=1}^{k} m_1(a_i \cdot b_j)$ shown in Proposition 1. $\square$

In the following example, we illustrate the result of the previous theorem.

Example 5. *Consider the product MV algebra F and the partitions $\mathcal{A} = \{x, 1-x\}$, $\mathcal{B} = \{x^2, 1-x^2\}$ of the product MV algebra F from Example 2. In addition, let m_1, m_2 be the states on F, defined in Example 4. Then the partitions $\mathcal{A}$ and $\mathcal{B}$ have the m_1-state values $\frac{1}{2}$, $\frac{1}{2}$ and $\frac{1}{3}$, $\frac{2}{3}$ of the corresponding elements, respectively, and the m_2-state values $\frac{2}{3}$, $\frac{1}{3}$ and $\frac{1}{2}$, $\frac{1}{2}$ of the corresponding elements, respectively. The join of partitions $\mathcal{A}$ and $\mathcal{B}$ is the system $\mathcal{A} \vee \mathcal{B} = \{x^3, x^2(1-x), x(1-x^2), (1-x)(1-x^2)\}$; it has the m_1-state values $\frac{1}{4}$, $\frac{1}{12}$, $\frac{1}{4}$, $\frac{5}{12}$, and the m_2-state values $\frac{2}{5}$, $\frac{1}{10}$, $\frac{4}{15}$, $\frac{7}{30}$ of the corresponding elements. By simple calculations we obtain:*

$$D_{\mathcal{A}}(m_1\|m_2) = 0.085 \ bit, \ D_{\mathcal{A}\vee\mathcal{B}}(m_1\|m_2) = 0.134 \ bit, \ D_{\mathcal{B}/\mathcal{A}}(m_1\|m_2) = 0.049 \ bit.$$

It is possible to verify that $D_{\mathcal{A}\vee\mathcal{B}}(m_1 \| m_2) = D_{\mathcal{A}}(m_1 \| m_2) + D_{\mathcal{B}/\mathcal{A}}(m_1 \| m_2)$.

5. Discussion

In this paper, we have extended the study of entropy in product MV algebras. The main aim of the paper was to introduce, using known results concerning the entropy in product MV algebras, the concepts of mutual information and Kullback–Leibler divergence for the case of product MV algebras and examine algebraic properties of the proposed measures. Our results have been presented in Sections 3 and 4.

In Section 3 we have introduced the notions of mutual information and conditional mutual information of partitions of product MV algebras and proved some basic properties of the suggested measures. It was shown that the entropy of partitions of product MV algebras can be considered as a special case of their mutual information. Specifically, it was proved that from the properties of mutual information it follows subadditivity and additivity of entropy (Theorem 3). Theorem 6 provides the chain rule for mutual information. In addition, the data processing inequality for conditionally independent partitions in product MV algebras is proved. Moreover, a concavity of mutual information has been studied.

In Section 4 the notion of Kullback–Leibler divergence in product MV algebras was introduced and the basic properties of this measure were shown. In particular, a convexity of Kullback–Leibler divergence with respect to additive states defined on a given product MV algebra is proved. Theorem 11 admits interpretation of Kullback–Leibler divergence as a measure of how different two states on a common product MV algebra (over the same partition) are. The relationship between KL-divergence and entropy is provided in Theorem 12: the more a state $m \in F$ diverges from the state $\nu \in F$ uniform over $\mathcal{A}$ (over the same partition $\mathcal{A}$) the lesser the entropy $H_m(\mathcal{A})$ is and vice versa. Finally, a conditional version of the Kullback–Leibler divergence in product MV algebras has been defined and the chain rule for Kullback–Leibler divergence with respect to additive states defined on a given product MV algebra has been established.

Notice that in [14] (see also [29,30]) the entropy on a full tribe F of fuzzy sets has been studied. The tribe F is closed also under the natural product of fuzzy sets and it represents a special case of product MV algebras. Accordingly, the theory presented in this contribution can also be applied for the mentioned case of tribes of fuzzy sets.

In [51–55] a more general fuzzy theory—intuitionistic fuzzy sets (IF-sets for short) has been developed. While a fuzzy set is a mapping $\mu_A : \Omega \to [0, 1]$ (where the considered fuzzy set is identified with its membership function μ_A), the Atanassov IF-set is a pair $A = (\mu_A, \nu_A)$ of functions $\mu_A, \nu_A : \Omega \to [0, 1]$ with $\mu_A + \nu_A \leq 1$. The function μ_A is interpreted as a membership function of IF-set A, and the function ν_A as a non-membership function of IF-set A. Evidently, any fuzzy set $\mu_A : \Omega \to [0, 1]$ can be considered as an IF-set $A = (\mu_A, 1 - \mu_A)$. Any result holding for IF-sets is applicable also to fuzzy sets. Of course, the opposite implication is not true; the theory of intuitionistic fuzzy sets presents a non-trivial generalization of the fuzzy set theory. So IF-sets present possibilities for modeling a larger class of real situations. Note that some results about the entropy on IF-sets can be found e.g., in [56–59]. These results could be used in developing information theory for the case of IF-sets.

To give a possibility to applied MV algebra results also to families of IF-experiments, one can use the Mundici characterization of MV algebras. In the family of IF-sets it is natural to define the partial ordering relation $\leq$ in the following way: if $A = (\mu_A, \nu_A)$, and $B = (\mu_B, \nu_B)$ are two IF-sets, then $A \leq B$ if and only if $\mu_A \leq \mu_B$, and $\nu_A \geq \nu_B$. Namely, in the fuzzy case $\mu_A \leq \mu_B$ implies $\nu_A = 1 - \mu_A \geq 1 - \mu_B = \nu_B$. Therefore we can consider the Abelian *l*-group $(\Re^2, +, \leq)$ putting $A + B = (\mu_A + \mu_B, 1 - (1 - \nu_A + 1 - \nu_B)) = (\mu_A + \mu_B, \nu_A + \nu_B - 1)$ with the zero element $0 = (0, 1)$. (In fact, $A + 0 = (\mu_A, \nu_A) + (0, 1) = (\mu_A, \nu_A) = A$.) The partial ordering $\leq$ in the *l*-group $(\Re^2, +, \leq)$ is defined by the prescription $A \leq B$ if and only if $\mu_A \leq \mu_B$, and $\nu_A \geq \nu_B$. Then a suitable MV algebra is e.g., the system $M = \{(\mu_A, \nu_A); (0, 1) \leq (\mu_A, \nu_A) \leq (1, 0)\}$. Moreover, this MV algebra is a product MV algebra with the product defined by $A \cdot B = (\mu_A \cdot \mu_B, 1 - (1 - \nu_A) \cdot (1 - \nu_B)) = (\mu_A \cdot \mu_B, \nu_A + \nu_B - \nu_A \cdot \nu_B)$. The presented MV algebra approach gives a possible elegant and practical way for obtaining new results also in the intuitionistic fuzzy case. We note that this approach was used to construct the Kolmogorov-type entropy theory for IF systems in [58], drawing on entropy results for product MV-algebras published in [35,49,50]. In this way it is also possible to develop the theory of information and K–L divergence for IF-sets.

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

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

References

1. Gray, R.M. *Entropy and Information Theory*; Springer: Berlin/Heidelberg, Germany, 2009.
2. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
3. Kolmogorov, A.N. New Metric Invariant of Transitive Dynamical Systems and Automorphisms of Lebesgue Spaces. *Dokl. Russ. Acad. Sci.* **1958**, *119*, 861–864.
4. Sinai, Y.G. *Ergodic Theory with Applications to Dynamical Systems and Statistical Mechanics*; Springer: Berlin/Heidelberg, Germany, 1990.
5. Sinai, Y.G. On the Notion of Entropy of a Dynamical System. *Dokl. Russ. Acad. Sci.* **1959**, *124*, 768–771.
6. Markechová, D. The entropy of fuzzy dynamical systems and generators. *Fuzzy Sets Syst.* **1992**, *48*, 351–363. [CrossRef]
7. Piasecki, K. Probability of fuzzy events defined as denumerable additive measure. *Fuzzy Sets Syst.* **1985**, *17*, 271–284. [CrossRef]
8. Mesiar, R. The Bayes principle and the entropy on fuzzy probability spaces. *Int. J. Gen. Syst.* **1991**, *20*, 67–72. [CrossRef]
9. Mesiar, R.; Rybárik, J. Entropy of Fuzzy Partitions—A General Model. *Fuzzy Sets Syst.* **1998**, *99*, 73–79. [CrossRef]
10. Dumitrescu, D. Entropy of a fuzzy dynamical system. *Fuzzy Sets Syst.* **1995**, *70*, 45–57. [CrossRef]
11. Rahimi, M.; Riazi, A. On local entropy of fuzzy partitions. *Fuzzy Sets Syst.* **2014**, *234*, 97–108. [CrossRef]
12. Rahimi, M.; Assari, A.; Ramezani, F. A Local Approach to Yager Entropy of Dynamical Systems. *Int. J. Fuzzy Syst.* **2015**, *1*, 1–10. [CrossRef]
13. Srivastava, P.; Khare, M.; Srivastava, Y.K. m-Equivalence, entropy and F-dynamical systems. *Fuzzy Sets Syst.* **2001**, *121*, 275–283. [CrossRef]
14. Markechová, D.; Riečan, B. Entropy of Fuzzy Partitions and Entropy of Fuzzy Dynamical Systems. *Entropy* **2016**, *18*, 19. [CrossRef]
15. Riečan, B. An entropy construction inspired by fuzzy sets. *Soft Comput.* **2003**, *7*, 486–488.
16. Riečan, B. On a type of entropy of dynamical systems. *Tatra Mt. Math. Publ.* **1992**, *1*, 135–140.
17. Riečan, B. On some modifications of the entropy of dynamical systems. *Atti Semin. Mat. Fis. dell'Univ. Modena* **1994**, *42*, 157–166.
18. Dubois, D.; Prade, M. A review of fuzzy set aggregation connectives. *Inf. Sci.* **1985**, *36*, 85–121. [CrossRef]

19. Zadeh, L.A. Fuzzy Sets. *Inf. Control* **1965**, *8*, 338–358. [CrossRef]
20. Markechová, D. Entropy and mutual information of experiments in the fuzzy case. *Neural Netw. World* **2013**, *23*, 339–349. [CrossRef]
21. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
22. Kullback, S. *Information Theory and Statistics*; John Wiley & Sons: New York, NY, USA, 1959.
23. Schnakenberg, J. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **1976**, *48*, 571–585. [CrossRef]
24. Risken, H. *The Fokker-Planck Equation, Methods of Solution and Applications*; Springer: New York, NY, USA, 1984.
25. Qian, H. Relative Entropy: Free Energy Associated with Equilibrium Fluctuations and Nonequilibrium Deviations. *arXiv*, 2001; arXiv:math-ph/0007010v2.
26. Ellis, R.S. *Entropy, Large Deviations, and Statistical Mechanics*; Springer: New York, NY, USA, 1985.
27. Markechová, D. Kullback–Leibler Divergence and Mutual Information of Experiments in the Fuzzy Case. *Axioms* **2017**, *6*, 5. [CrossRef]
28. Chang, C.C. Algebraic analysis of many valued logics. *Trans. Am. Math. Soc.* **1958**, *88*, 467–490. [CrossRef]
29. Riečan, B.; Mundici, D. Probability on MV-algebras. In *Handbook of Measure Theory*; Pap, E., Ed.; Elsevier: Amsterdam, The Netherlands, 2002; pp. 869–910.
30. Riečan, B.; Neubrunn, T. *Integral, Measure and Ordering*; Springer: Dordrecht, The Netherlands, 1997.
31. Dvurečenskij, A.; Pulmannová, S. *New Trends in Quantum Structures*; Springer: Dordrecht, The Netherlands, 2000.
32. 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).
33. Mundici, D. Interpretation of AFC*-algebras in Lukasiewicz sentential calculus. *J. Funct. Anal.* **1986**, *56*, 889–894.
34. 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.
35. Riečan, B. Kolmogorov–Sinaj entropy on MV-algebras. *Int. J. Theor. Phys.* **2005**, *44*, 1041–1052. [CrossRef]
36. Kôpka, F.; Chovanec, F. D-posets. *Math. Slovaca* **1994**, *44*, 21–34.
37. Kôpka, F. Quasiproduct on Boolean D-posets. *Int. J. Theor. Phys.* **2008**, *47*, 26–35. [CrossRef]
38. Frič, R. On D-posets of fuzzy sets. *Math. Slovaca* **2014**, *64*, 545–554. [CrossRef]
39. Foulis, D.J.; Bennet, M.K. Effect algebras and unsharp quantum logics. *Found. Phys.* **1994**, *24*, 1331–1352. [CrossRef]
40. Frič, R.; Papčo, M. Probability domains. *Int. J. Theor. Phys.* **2010**, *49*, 3092–3100. [CrossRef]
41. Skřivánek, V.; Frič, R. Generalized random events. *Int. J. Theor. Phys.* **2015**, *54*, 4386–4396. [CrossRef]
42. 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.
43. Giski, Z.E.; 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.
44. Ebrahimi, M.; Mosapour, B. The Concept of Entropy on D-posets. *Cankaya Univ. J. Sci. Eng.* **2013**, *10*, 137–151.
45. Riečan, B. On the product MV-algebras. *Tatra Mt. Math.* **1999**, *16*, 143–149.
46. Montagna, F. An algebraic approach to propositional fuzzy logic. *J. Log. Lang. Inf.* **2000**, *9*, 91–124. [CrossRef]
47. Jakubík, J. On product MV algebras. *Czech. Math J.* **2002**, *52*, 797–810. [CrossRef]
48. Di Nola, A.; Dvurečenskij, A. Product MV-algebras. *Mult. Valued Log.* **2001**, *6*, 193–215.
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. Atanassov, K. *Intuitionistic Fuzzy Sets: Theory and Applications*; Physica Verlag: New York, NY, USA, 1999.
52. Atanassov, K. Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **1986**, *20*, 87–96. [CrossRef]
53. Atanassov, K. More on intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **1989**, *33*, 37–45. [CrossRef]
54. Atanassov, K.; Riečan, B. On two operations over intuitionistic fuzzy sets. *J. Appl. Math. Stat. Inform.* **2006**, *2*, 145–148. [CrossRef]

55. Riečan, B. Probability theory on IF events. In *Algebraic and Proof-Theoretic Aspects of Non-Classical Logics*; Papers in Honor of Daniele Mundici on the Occasion of his 60th Birthday; Lecture Notes in Computer Science; Springer: New York, NY, USA, 2007; pp. 290–308.

56. Farnoosh, R.; Rahimi, M.; Kumar, P. Removing noise in a digital image using a new entropy method based on intuitionistic fuzzy sets. In Proceedings of the International Conference on Fuzzy Systems, Vancouver, BC, Canada, 24–29 July 2016; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2016; pp. 1328–1332.

57. Burillo, P.; Bustince, H. Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets Syst.* **1996**, *78*, 305–316. [CrossRef]

58. Ďurica, M. Entropy on IF-events. *Notes Intuit. Fuzzy Sets* **2007**, *13*, 30–40.

59. Szmidt, E.; Kacprzyk, J. Entropy for intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **2001**, *118*, 467–477. [CrossRef]