entropy

MDPI

Article

Inequalities for Jensen–Sharma–Mittal and Jeffreys–Sharma–Mittal Type f–Divergences

Paweł A. Kluza

Department of Applied Mathematics and Computer Science, University of Life Sciences in Lublin, 28 Głęboka Street, 20-612 Lublin, Poland; pawel.kluza@up.lublin.pl; Tel.: +48-81-531-96-19

Abstract: In this paper, we introduce new divergences called Jensen–Sharma–Mittal and Jeffreys–Sharma–Mittal in relation to convex functions. Some theorems, which give the lower and upper bounds for two new introduced divergences, are provided. The obtained results imply some new inequalities corresponding to known divergences. Some examples, which show that these are the generalizations of Rényi, Tsallis, and Kullback–Leibler types of divergences, are provided in order to show a few applications of new divergences.

Keywords: convex function; Csiszár f-divergence; Sharma–Mittal f-divergence; Jensen–Sharma–Mittal divergence; Jeffreys–Sharma–Mittal divergence

MSC: 94A17; 26D15; 15B51

Citation: Kluza, P.A. Inequalities for Jensen–Sharma–Mittal and Jeffreys–Sharma–Mittal Type f–Divergences. *Entropy* **2021**, *23*, 1688. https://doi.org/10.3390/e23121688

Academic Editors: Boris Ryabko and Takuya Yamano

Received: 21 October 2021
Accepted: 14 December 2021
Published: 16 December 2021

1. Introduction

The Sharma–Mittal entropy was introduced as a new measure of information with two parameters [1]. It has previously been studied in the context of multi-dimensional harmonic oscillator systems [2]. This entropy could also be formulated in the form of exponential families, to which many usual statistical distributions including the Gaussians and discrete multinomials (that is, normalized histograms) belong. In physical applications it plays a major role in the field of thermo-statistics [3].

The Sharma–Mittal entropy is also applied for the analysis of the results of machine learning methods [4,5]. Additionally, the divergence based on considered entropy could be a cost function in the context of so-called the Twin Gaussian Processes [6].

It was originally showed by [7] that the Sharma–Mittal entropy generalized both Tsallis and Rényi entropy in the limiting cases of these two entropies. In [8], authors suggested a physical meaning of Sharma–Mittal entropy, which is the free energy difference between the equilibrium and the off-equilibrium distribution.

Recently, was published a manuscript showing, in opposition to the work [8], that Sharma–Mittal entropy besides the convenient thermodynamic systems does not reduce only to Kullback–Leibler entropy. In [9] Verma and Merigó present the use of Sharma–Mittal entropy under intuitionistic fuzzy environment. Additionally, in [5] Koltcov et al. demonstrate that Sharma–Mittal entropy is a tool for selecting both the number of topics and the values of hyper-parameters, simultaneously controlling for semantic stability, which none of the existing metrics can do.

Another applications of considered entropy are interesting results in the cosmological setup, such as black hole thermodynamics [10]. Namely, it helps us to describe the current accelerated universe by using the vacuum energy in a suitable manner [11]. In addition [12] have established the relation between anomalous diffusion process and Sharma–Mittal entropy.

This paper is based on publications in which we introduced new types of f-divergences [13–16].

In this paper we generalize Sharma–Mittal types divergences in order to obtain new types of divergences and hence the inequalities from which it will be possible to derive new results and generalizations for known divergences in order to estimate the lower and upper bounds which determine the level of the uncertainty measure.

2. Sharma–Mittal Type Divergences

Throughout $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the sets of non-negative and positive numbers, respectively, i.e., $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_{++} = (0, \infty)$.

Let $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ with $p_i, q_i \geq 0$, $i = 1, \ldots, n$. The *relative entropy* (also called *Kullback–Leibler divergence*) is defined by (see [17])

$$H_1(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}. \tag{1}$$

In the above definition, based on continuity arguments, we use a convention that $0 \log(0/q) = 0$ and $p \log(p/0) = +\infty$. Additionally $0 \log(0/0) = 0$.

Let $f\colon \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$, and $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$.

The *Csiszár f-divergence* is defined by (see [15])

$$C_f(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{n} p_i f\left(\frac{q_i}{p_i}\right). \tag{2}$$

with the conventions $0 f\left(\frac{0}{0}\right) = 0$ and $0 f\left(\frac{c}{0}\right) = c \lim_{t \to \infty} \frac{f(t)}{t}$, $c > 0$ (see [18–20]).

The *Tsallis* divergence of order α is defined by (see [17])

$$T_\alpha(\mathbf{p}, \mathbf{q}) = \frac{1}{\alpha - 1} \left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1 \right).$$

The *Rényi* divergence of order α is defined by (see [17,21])

$$H_\alpha(\mathbf{p}, \mathbf{q}) = \frac{1}{\alpha - 1} \log \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}.$$

The *Sharma–Mittal* divergence of order α and degree β is defined by (see [4])

$$SM_{\alpha,\beta}(\mathbf{p}, \mathbf{q}) = \frac{1}{\beta - 1} \left[\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} \right)^{\frac{1-\beta}{1-\alpha}} - 1 \right], \tag{3}$$

for all $\alpha > 0$, $\alpha \neq 1$ and $\beta \neq 1$.

Let $g : I \to \mathbb{R}$ be a convex function on an interval $I \subset \mathbb{R}$. Let $\mathbf{x} = (x_1, \ldots, x_n) \in I^n$ and $p_i \in [0, 1)$ for $i = 1, \ldots n$.

The *Jensen's inequality* is as follows (see [22])

$$g\left(\sum_{i=1}^{n} p_i x_i \right) \leq \sum_{i=1}^{n} p_i g(x_i). \tag{4}$$

When the function $k : \mathbb{R}_+ \to \mathbb{R}$ is convex and the function $l : \mathbb{R}_+ \to \mathbb{R}$ is convex and increasing then the composition of the functions $k \circ l : \mathbb{R}_+ \to \mathbb{R}$ is convex. We assume that the probabilities $p_i \geq 0$ and $q_i > 0$ for $i = 1, \ldots, n$.

It is known (see [4]) that if

$$\alpha = \beta \to 1 \quad \text{then} \quad SM_{\alpha,\beta}(\mathbf{p}, \mathbf{q}) \to H_1(\mathbf{p}, \mathbf{q}),$$

$$\beta \to 1 \quad \text{and} \quad \alpha \in \mathbb{R} \quad \text{then} \quad SM_{\alpha,\beta}(\mathbf{p},\mathbf{q}) \to H_\alpha(\mathbf{p},\mathbf{q}),$$
$$\beta = \alpha \quad \text{then} \quad SM_{\alpha,\beta}(\mathbf{p},\mathbf{q}) \to T_\alpha(\mathbf{p},\mathbf{q}).$$

Let $h : \mathbb{R} \to \mathbb{R}$ be the differentiable function. Then the *Sharma–Mittal h-divergence* is defined as follows:

$$SM_{h,\alpha,\beta}(\mathbf{p},\mathbf{q}) = \frac{h\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta - 1}, \tag{5}$$

for all $\alpha > 0$, $\alpha \neq 1$ and $\beta \neq 1$.

If we assume that $h = id$ then (5) becomes Sharma–Mittal divergence.

When for all $t > 0$, $h(t) = \log(te)$ then (5) becomes Rényi divergence of order α.

We substitute for $t = \left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}$ and we have

$$h(t) = h\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}\right) = \log\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}} e\right) = \log\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}} + 1.$$

Hence, from (5)

$$SM_{h,\alpha,\beta}(\mathbf{p},\mathbf{q}) = \frac{\log\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}}{\beta - 1} = \frac{1}{\alpha - 1}\log\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} = H_\alpha(\mathbf{p},\mathbf{q}).$$

Let $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ be a differentiable function with respect to β and

$$\Psi(\alpha,\beta) = h\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}\right).$$

We assume that $h'(1) = 1$ and $\Psi(\alpha,1) = 1$. Then,

$$\lim_{\beta\to 1}\frac{\Psi(\alpha,\beta) - \Psi(\alpha,1)}{\beta - 1} = \Psi'(\alpha,1) = h'\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}\right)_{|\beta=1} =$$

$$\frac{\partial}{\partial\beta}_{|\beta=1} h\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}\right)\frac{\partial}{\partial\beta}_{|\beta=1}\left(\left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right]^{\frac{1-\beta}{1-\alpha}}\right)\frac{\partial}{\partial\beta}_{|\beta=1}\left(\frac{1-\beta}{1-\alpha}\right) =$$

$$h'(1)\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right)^{\frac{1-1}{1-\alpha}}\log\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right)\left(\frac{-1}{1-\alpha}\right) =$$

$$h'(1)\frac{1}{\alpha-1}\log\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right) = \frac{1}{\alpha-1}\log\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right).$$

Hence, the Sharma–Mittal h–divergence tends to Rényi divergence of order α.

Remark 1. *If, additionally, α tends to 1 then based on the proof of the Equation (11) from [16], Sharma–Mittal h-divergence tends to relative entropy (called Kullback–Leibler divergence).*

Now we define a new *generalized* (h,ϕ) *Sharma–Mittal divergence* as follows

$$SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q}) = \frac{h\left(\left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta - 1}, \tag{6}$$

where $\phi : (0,+\infty) \times \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing, non-negative and differentiable function for $\beta > 1$.

We assume that $\mathcal{F} = \{f_\alpha : (0,\infty) \to \mathbb{R} : \alpha \in \mathbb{R}\}$ is a given family of functions such that $\sum_{i=1}^{n} q_i f_{\alpha|_{\alpha=1}}\left(\frac{p_i}{q_i}\right) = 1$ for $\alpha = 1$ and which are increasing, non-negative for $\alpha > 1$ and such that for every $t \in (0,+\infty)$ the function $\alpha \longmapsto f_\alpha(t)$ is differentiable.

According to [16] if we substitute the function $f_\alpha(\frac{p_i}{q_i})$ from the family $\mathcal{F}$ for $\phi(\frac{p_i}{q_i},\alpha)$ then it stands that

$$\lim_{\beta\to 1} SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q}) = R_{h,f_\alpha}(\mathbf{p},\mathbf{q}).$$

We assume that $h(1) = 1$. Then,

$$\lim_{\beta\to 1} SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q}) = \lim_{\beta\to 1} \frac{h\left(\left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta - 1} =$$

$$h'\left(\left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{0}\right) \cdot \left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{0} \left[\log \sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right] \frac{-1}{1-\alpha} =$$

$$h'(1)\frac{1}{\alpha-1}\log \sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right) = \frac{1}{\alpha-1}\log \sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right) =$$

$$\frac{1}{\alpha-1}\log \sum_{i=1}^{n} q_i f_\alpha\left(\frac{p_i}{q_i}\right) = R_{h,f_\alpha}(\mathbf{p},\mathbf{q}). \tag{7}$$

Remark 2. *If in (6)* $\beta \to 1$ *and* $\phi(\frac{p_i}{q_i},\alpha) = f_\alpha(\frac{p_i}{q_i})$ *then the generalized* (h,ϕ)*–Sharma–Mittal divergence tends to generalized* $(h,\mathcal{F})$*–Rényi divergence.*

The function ϕ is the generalization of the function $f_\alpha(\frac{p_i}{q_i})$ which is used for example in Csiszár f-divergence. Condition $\beta \to 1$ means that the limit of generalized Sharma–Mittal divergence is equal to generalized $(h,\mathcal{F})$–Rényi divergence. Hence we have implications for generalized forms of entropies.

Remark 3. *Additionally, when in (6)* $\alpha \to 1$ *and* $\phi(\frac{p_i}{q_i},\alpha) = \left(\frac{p_i}{q_i}\right)^{\alpha}$ *then the generalized* (h,ϕ)*–Sharma–Mittal divergence tends to Kullback–Leibler divergence, because we have from Remark 2*

$$\lim_{(\alpha,\beta)\to(1,1)} SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q}) = \lim_{\alpha\to 1} \frac{\log \sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right) - \log 1}{\alpha - 1} =$$

$$\frac{\partial}{\partial\alpha}_{|_{\alpha=1}} \log \sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i}\right)^{\alpha} = \frac{1}{\sum_{i=1}^{n} p_i} \sum_{i=1}^{n} q_i \frac{p_i}{q_i} \log \frac{p_i}{q_i} = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} = H_1(\mathbf{p},\mathbf{q}).$$

Remark 4. *In (6), when the parameter* $\beta = \alpha$*, the function* $h = id$ *and* $\phi(\frac{p_i}{q_i},\alpha) = \left(\frac{p_i}{q_i}\right)^{\alpha}$ *then the generalized* (h,ϕ)*–Sharma–Mittal divergence tends to the Tsallis* f*-divergence or order* α*.*

This work is more theoretical than practical. Therefore, the implications are formulated in the mathematical area that is from constructing general model which gives known specific cases.

3. Jensen–Sharma–Mittal and Jeffreys–Sharma–Mittal Divergences

The Jensen–Shannon divergence (Jensen–Shannon entropy) is defined as follows (see [17]):

$$\mathrm{Jen}\,(\mathbf{p},\mathbf{q}) = \frac{1}{2}SM\Big(\mathbf{p},\frac{\mathbf{p}+\mathbf{q}}{2}\Big)+\frac{1}{2}SM\Big(\mathbf{q},\frac{\mathbf{p}+\mathbf{q}}{2}\Big).$$

The Jeffreys divergence (Jeffreys entropy) is defined as follows (see [17]):

$$\mathrm{Jef}\,(\mathbf{p},\mathbf{q}) = SM(\mathbf{p},\mathbf{q})+SM(\mathbf{q},\mathbf{p}).$$

We introduce a new generalized (h,ϕ) Jensen–Sharma–Mittal divergence defined by

$$\mathrm{Jen}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) = \frac{1}{2}SM_{h,\phi,\alpha,\beta}\Big(\mathbf{p},\frac{\mathbf{p}+\mathbf{q}}{2}\Big)+\frac{1}{2}SM_{h,\phi,\alpha,\beta}\Big(\mathbf{q},\frac{\mathbf{p}+\mathbf{q}}{2}\Big) \tag{8}$$

with assumptions as before.

We similarly introduce a new generalized (h,ϕ) Jeffreys–Sharma–Mittal divergence as follows

$$\mathrm{Jef}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) = SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q})+SM_{h,\phi,\alpha,\beta}(\mathbf{q},\mathbf{p}). \tag{9}$$

Taking into account inequality from [17]:

$$0 \le \mathrm{Jen}\,(\mathbf{p},\mathbf{q}) \le \frac{1}{2}\mathrm{Jef}\,(\mathbf{p},\mathbf{q}),$$

describing the relation between the Jensen–Shannon and Jeffreys divergences, we could formulate the following:

$$0 \le \mathrm{Jen}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) \le \frac{1}{2}\mathrm{Jef}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q}). \tag{10}$$

We define the *Jensen–Sharma–Mittal h-divergence* where, in (8), $\phi\Big(\frac{p_i}{q_i},\alpha\Big) = \Big(\frac{p_i}{q_i}\Big)^{\alpha}$. Then, it takes the form:

$$\mathrm{Jen}\,SM^{h}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) = \frac{1}{2}SM_{h,\alpha,\beta}\Big(\mathbf{p},\frac{\mathbf{p}+\mathbf{q}}{2}\Big)+\frac{1}{2}SM_{h,\alpha,\beta}\Big(\mathbf{q},\frac{\mathbf{p}+\mathbf{q}}{2}\Big) \tag{11}$$

In the same way, we define the *Jeffreys–Sharma–Mittal h–divergence*:

$$\mathrm{Jef}\,SM^{h}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) = SM_{h,\alpha,\beta}(\mathbf{p},\mathbf{q})+SM_{h,\alpha,\beta}(\mathbf{q},\mathbf{p}). \tag{12}$$

Additionally, if the function $h(t) = t$ then we define the Jensen–Sharma–Mittal and the Jeffreys–Sharma–Mittal divergences of order α and degree β, respectively.

$$\mathrm{Jen}\,SM_{\alpha,\beta}(\mathbf{p},\mathbf{q}) = \frac{1}{2}SM_{\alpha,\beta}\Big(\mathbf{p},\frac{\mathbf{p}+\mathbf{q}}{2}\Big)+\frac{1}{2}SM_{\alpha,\beta}\Big(\mathbf{q},\frac{\mathbf{p}+\mathbf{q}}{2}\Big), \tag{13}$$

$$\mathrm{Jef}\,SM_{\alpha,\beta}(\mathbf{p},\mathbf{q}) = SM_{\alpha,\beta}(\mathbf{p},\mathbf{q})+SM_{\alpha,\beta}(\mathbf{q},\mathbf{p}). \tag{14}$$

When in (8) and (9) $\beta \to 1$ and we substitute for $\phi\Big(\frac{p_i}{q_i},\alpha\Big) = f_\alpha\Big(\frac{p_i}{q_i}\Big)$ then we obtain, defined in [16], the generalized $(h,\mathcal{F})$ Jensen–Rényi and Jeffreys–Rényi divergences, respectively:

$$\mathrm{Jen}\,SM^{h,f_\alpha}_{\alpha,1}(\mathbf{p},\mathbf{q}) = \mathrm{Jen}\,R_{h,f_\alpha}(\mathbf{p},\mathbf{q}) = \frac{1}{2}R_{h,f_\alpha}\Big(\mathbf{p},\frac{\mathbf{p}+\mathbf{q}}{2}\Big)+\frac{1}{2}R_{h,f_\alpha}\Big(\mathbf{q},\frac{\mathbf{p}+\mathbf{q}}{2}\Big),$$

$$\operatorname{Jef} SM_{\alpha,1}^{h,f_\alpha}(\mathbf{p},\mathbf{q}) = \operatorname{Jef} R_{h,f_\alpha}(\mathbf{p},\mathbf{q}) = R_{h,f_\alpha}(\mathbf{p},\mathbf{q}) + R_{h,f_\alpha}(\mathbf{q},\mathbf{p}). \tag{15}$$

The following theorem is the generalization and refinement of the inequalities for some known divergences and provides lower and upper bounds for the generalized (h,ϕ) Jeffreys–Sharma–Mittal divergence in order to a more accurate estimation of its uncertainty measure.

Theorem 1. *Let* $\mathbf{p} = (p_1,\ldots,p_n)$ *and* $\mathbf{q} = (q_1,\ldots,q_n)$ *be two discrete probability distributions with* $p_i > 0$, $q_i > 0$, $\frac{q_i}{p_i} \in I_0$, $\frac{p_i}{q_i} \in I_0$, $i = 1,\ldots,n$, *where* $I_0 \subset \mathbb{R}$ *is an interval, such that* $1 \in I_0$. *Let* $\phi : I_0 \times \mathbb{R}_+ \to \mathbb{R}_+$ *be an increasing, non-negative and differentiable function for which* $\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right) \geq 1$ *and* $\sum_{i=1}^{n} p_i \phi\left(\frac{q_i}{p_i},\alpha\right) \geq 1$ *where* $1 < \beta \leq \alpha$, $\alpha \in \mathbb{R}_+ \setminus \{1\}$ *and* $h : I_0 \to \mathbb{R}$ *be a convex and increasing function on* I_0.

Then, the following inequalities are valid:

$$\frac{1}{\alpha-1}\log\prod_{i=1}^{n}\left(\phi\left(\frac{p_i}{q_i},\alpha\right)\right)^{q_i}\left(\phi\left(\frac{q_i}{p_i},\alpha\right)\right)^{p_i} \leq \operatorname{Jef} SM_{\alpha,\beta}^{h,\phi}(\mathbf{p},\mathbf{q}) \leq$$

$$\frac{1}{\beta-1}\left(\operatorname{Jef} C_{h\circ\phi}(\mathbf{p},\mathbf{q}) - 2\right). \tag{16}$$

Proof. Taking into account the assumptions, we could formulate the following inequality:

$$\left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}} \leq \sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right). \tag{17}$$

The function h is increasing and convex, therefore, from (4) and (17) we obtain inequalities:

$$h\left(\left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) \leq h\left(\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right) \leq \sum_{i=1}^{n} q_i (h\circ\phi)\left(\frac{p_i}{q_i},\alpha\right). \tag{18}$$

In the same way, we obtain the following inequalities:

$$h\left(\left[\sum_{i=1}^{n} p_i \phi\left(\frac{q_i}{p_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) \leq h\left(\sum_{i=1}^{n} p_i \phi\left(\frac{q_i}{p_i},\alpha\right)\right) \leq \sum_{i=1}^{n} p_i (h\circ\phi)\left(\frac{q_i}{p_i},\alpha\right). \tag{19}$$

From (9) we have

$$\operatorname{Jef} SM_{\alpha,\beta}^{h,\phi}(\mathbf{p},\mathbf{q}) = SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q}) + SM_{h,\phi,\alpha,\beta}(\mathbf{q},\mathbf{p}) =$$

$$\frac{h\left(\left[\sum_{i=1}^{n} q_i \phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta-1} + \frac{h\left(\left[\sum_{i=1}^{n} p_i \phi\left(\frac{q_i}{p_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta-1}$$

Taking into account (2), (18), (19) and the definition of Jeffreys divergence, it stands that:

$$\operatorname{Jef} SM_{\alpha,\beta}^{h,\phi}(\mathbf{p},\mathbf{q}) \leq \frac{\sum_{i=1}^{n} q_i (h\circ\phi)\left(\frac{p_i}{q_i},\alpha\right) - 1}{\beta-1} + \frac{\sum_{i=1}^{n} p_i (h\circ\phi)\left(\frac{q_i}{p_i},\alpha\right) - 1}{\beta-1} =$$

$$\frac{\sum_{i=1}^{n} q_i (h\circ\phi)\left(\frac{p_i}{q_i},\alpha\right) + \sum_{i=1}^{n} p_i (h\circ\phi)\left(\frac{q_i}{p_i},\alpha\right) - 2}{\beta-1} = \frac{\operatorname{Jef} C_{h\circ\phi}(\mathbf{p},\mathbf{q}) - 2}{\beta-1}. \tag{20}$$

The above inequality is the upper bound for generalized (h,ϕ) Jeffreys–Sharma–Mittal divergence.

By using the convexity of the function h with $h(1)=1$ the following inequality is valid for $\beta>1$:

$$\frac{h\left(\left[\sum_{i=1}^{n} q_i\phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-1}{\beta-1} \geq \frac{\partial}{\partial\beta}_{|_{\beta=1}} h\left(\left[\sum_{i=1}^{n} q_i\phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right). \tag{21}$$

From (7) the above derivative function is equal to: $\frac{1}{\alpha-1}\log\sum_{i=1}^{n} q_i\phi\left(\frac{p_i}{q_i},\alpha\right)$.

The function $f(t)=\log t$ is concave and increasing. Then, it stands that:

$$\frac{1}{\alpha-1}\log\sum_{i=1}^{n} q_i\phi\left(\frac{p_i}{q_i},\alpha\right) \geq \frac{1}{\alpha-1}\sum_{i=1}^{n} q_i\log\left(\phi\left(\frac{p_i}{q_i},\alpha\right)\right). \tag{22}$$

Hence, from (21) and (22) we have the inequality:

$$\frac{h\left(\left[\sum_{i=1}^{n} q_i\phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-1}{\beta-1} \geq \frac{1}{\alpha-1}\sum_{i=1}^{n} q_i\log\left(\phi\left(\frac{p_i}{q_i},\alpha\right)\right). \tag{23}$$

Similarly, we obtain the second inequality:

$$\frac{h\left(\left[\sum_{i=1}^{n} p_i\phi\left(\frac{q_i}{p_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-1}{\beta-1} \geq \frac{1}{\alpha-1}\sum_{i=1}^{n} p_i\log\left(\phi\left(\frac{q_i}{p_i},\alpha\right)\right). \tag{24}$$

We have from (6), (23) and (24) that:

$$SM_{h,\phi,\alpha,\beta}(\mathbf{p},\mathbf{q}) \geq \frac{1}{\alpha-1}\sum_{i=1}^{n} q_i\log\left(\phi\left(\frac{p_i}{q_i},\alpha\right)\right),$$

$$SM_{h,\phi,\alpha,\beta}(\mathbf{q},\mathbf{p}) \geq \frac{1}{\alpha-1}\sum_{i=1}^{n} p_i\log\left(\phi\left(\frac{q_i}{p_i},\alpha\right)\right).$$

Then, by using the definition (9) we have:

$$\mathrm{Jef}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) \geq \frac{1}{\alpha-1}\left(\log\prod_{i=1}^{n}\left(\phi\left(\frac{p_i}{q_i},\alpha\right)\right)^{q_i}+\log\prod_{i=1}^{n}\left(\phi\left(\frac{q_i}{p_i},\alpha\right)\right)^{p_i}\right) =$$

$$\frac{1}{\alpha-1}\log\prod_{i=1}^{n}\left(\phi\left(\frac{p_i}{q_i},\alpha\right)\right)^{q_i}\left(\phi\left(\frac{q_i}{p_i},\alpha\right)\right)^{p_i}. \tag{25}$$

This result is the lower bound of the generalized (h,ϕ) Jeffreys–Sharma–Mittal divergence. Combining (20) and (25) we obtain the expected inequalities (16). □

Corollary 1. *When we substitute for* $\phi\left(\frac{p_i}{q_i},\alpha\right)=\left(\frac{p_i}{q_i}\right)^{\alpha}$ *then from (16) we obtain the inequalities for Jeffreys–Sharma–Mittal h-divergence:*

$$\frac{1}{\alpha-1}\log\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{\alpha(q_i-p_i)} \leq \mathrm{Jef}\,SM^{h}_{\alpha,\beta}(\mathbf{p},\mathbf{q}) \leq$$

$$\frac{\sum_{i=1}^{n} q_i h\left(\left(\frac{p_i}{q_i}\right)^{\alpha}\right) + \sum_{i=1}^{n} p_i h\left(\left(\frac{q_i}{p_i}\right)^{\alpha}\right) - 2}{\beta - 1}.$$

We now formulate the theorem thanks to which the estimation of the generalized (h, ϕ) Jensen–Sharma–Mittal divergence will be possible.

Theorem 2. *Let* $\mathbf{p} = (p_1, \ldots, p_n)$ *and* $\mathbf{q} = (q_1, \ldots, q_n)$ *be two discrete probability distributions with* $p_i > 0$, $q_i > 0$, $\frac{q_i}{p_i} \in I_0$, $\frac{p_i}{q_i} \in I_0$, $i = 1, \ldots, n$, *where* $I_0 \subset \mathbb{R}$ *is an interval such that* $1 \in I_0$. *Let* $\phi : I_0 \times \mathbb{R}_+ \to \mathbb{R}_+$ *be an increasing, non-negative and differentiable function for which* $\sum_{i=1}^{n} \frac{p_i+q_i}{2} \phi\left(\frac{2p_i}{p_i+q_i}, \alpha\right) \geq 1$ *and* $\sum_{i=1}^{n} \frac{p_i+q_i}{2} \phi\left(\frac{2q_i}{p_i+q_i}, \alpha\right) \geq 1$ *where* $1 < \beta \leq \alpha$, $\alpha \in \mathbb{R}_+ \setminus \{1\}$ *and* $h : I_0 \to \mathbb{R}$ *be a convex and increasing function on* I_0.

Then, the following inequalities are valid:

$$\frac{1}{2(\alpha-1)} \log \prod_{i=1}^{n} \left(\phi\left(\frac{2p_i}{p_i+q_i}, \alpha\right) \phi\left(\frac{2q_i}{p_i+q_i}, \alpha\right)\right)^{\frac{p_i+q_i}{2}} \leq \mathrm{Jen}\, SM_{\alpha,\beta}^{h,\phi}(\mathbf{p}, \mathbf{q}) \leq$$

$$\frac{\sum_{i=1}^{n} (p_i + q_i) \left[(h \circ \phi)\left(\frac{2p_i}{p_i+q_i}, \alpha\right) + (h \circ \phi)\left(\frac{2q_i}{p_i+q_i}, \alpha\right)\right] - 4}{4(\beta - 1)} \tag{26}$$

Proof. Let's consider the function

$$\frac{h\left(\left[\sum_{i=1}^{n} \frac{p_i+q_i}{2} \phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta - 1}. \tag{27}$$

Using the assumptions that the function h is differentiable, convex and $h(1) = 1$ we could formulate the following inequality:

$$\frac{h\left(\left[\sum_{i=1}^{n} \frac{p_i+q_i}{2} \phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) - 1}{\beta - 1} \geq \frac{\partial}{\partial \beta}_{|_{\beta=1}} h\left(\left[\sum_{i=1}^{n} \frac{p_i + q_i}{2} \phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right). \tag{28}$$

Then, (28) is equal to:

$$\left[\log \sum_{i=1}^{n} \frac{p_i + q_i}{2} \phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right)\right] \left(\frac{-1}{1-\alpha}\right) = \frac{1}{\alpha - 1} \log \sum_{i=1}^{n} \frac{p_i + q_i}{2} \phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right).$$

Taking into account concavity of the function log, we have that:

$$\frac{1}{\alpha - 1} \log \sum_{i=1}^{n} \frac{p_i + q_i}{2} \phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right) \geq \frac{1}{\alpha - 1} \sum_{i=1}^{n} \frac{p_i + q_i}{2} \log\left(\phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right)\right).$$

Then, we obtain that (27) is greater than

$$\frac{1}{\alpha - 1} \log \prod_{i=1}^{n} \left(\phi\left(\frac{p_i}{\frac{p_i+q_i}{2}}, \alpha\right)\right)^{\frac{p_i+q_i}{2}}. \tag{29}$$

We do the same with the function

$$\frac{h\left(\left[\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{q_i}{\frac{p_i+q_i}{2}},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-1}{\beta-1}. \tag{30}$$

Hence, we have that (30) is greater than

$$\frac{1}{\alpha-1}\log\prod_{i=1}^{n}\left(\phi\left(\frac{q_i}{\frac{p_i+q_i}{2}},\alpha\right)\right)^{\frac{p_i+q_i}{2}}. \tag{31}$$

Then, combining (27), (29)–(31), and using the definition (8) the following inequality occurs

$$\text{Jen}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q})\geq\frac{1}{2(\alpha-1)}\log\prod_{i=1}^{n}\left(\phi\left(\frac{2p_i}{p_i+q_i},\alpha\right)\phi\left(\frac{2q_i}{p_i+q_i},\alpha\right)\right)^{\frac{p_i+q_i}{2}} \tag{32}$$

and it is the lower bound of the generalized (h,ϕ) Jensen–Sharma–Mittal divergence.

When we consider the function

$$h\left(\left[\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{p_i}{\frac{p_i+q_i}{2}},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) \tag{33}$$

with $1<\beta\leq\alpha$ then for the convex and increasing function h we have from (4) that (33) is smaller than

$$h\left(\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{p_i}{\frac{p_i+q_i}{2}},\alpha\right)\right)\leq\sum_{i=1}^{n}\frac{p_i+q_i}{2}(h\circ\phi)\left(\frac{p_i}{\frac{p_i+q_i}{2}},\alpha\right). \tag{34}$$

In a similar way we conclude the following inequality for the function

$$h\left(\left[\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{q_i}{\frac{p_i+q_i}{2}},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right) \tag{35}$$

and we have

$$h\left(\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{q_i}{\frac{p_i+q_i}{2}},\alpha\right)\right)\leq\sum_{i=1}^{n}\frac{p_i+q_i}{2}(h\circ\phi)\left(\frac{q_i}{\frac{p_i+q_i}{2}},\alpha\right). \tag{36}$$

Then combining (33)–(36) and the definition (8) with the proper transformations we obtain the inequality

$$\text{Jen}\,SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q})\leq\frac{\sum_{i=1}^{n}(p_i+q_i)\left[(h\circ\phi)\left(\frac{2p_i}{p_i+q_i},\alpha\right)+(h\circ\phi)\left(\frac{2q_i}{p_i+q_i},\alpha\right)\right]-4}{4(\beta-1)} \tag{37}$$

which is the upper bound of the generalized (h,ϕ) Jensen–Sharma–Mittal divergence.

When we take into account (32) and (37), then we obtain (26). □

Corollary 2. *When we substitute for* $\phi\left(\frac{2p_i}{p_i+q_i},\alpha\right)=\left(\frac{2p_i}{p_i+q_i}\right)^{\alpha}$ *and for* $\phi\left(\frac{2q_i}{p_i+q_i},\alpha\right)=\left(\frac{2q_i}{p_i+q_i}\right)^{\alpha}$ *then from (26) we obtain the inequalities for Jensen–Sharma–Mittal h–divergence:*

$$\frac{1}{2(\alpha-1)}\log\prod_{i=1}^{n}\left[\frac{4p_iq_i}{(p_i+q_i)^2}\right]^{\frac{\alpha(p_i+q_i)}{2}}\leq \operatorname{Jen}SM^{h}_{\alpha,\beta}(\mathbf{p},\mathbf{q})\leq$$

$$\frac{\sum_{i=1}^{n}(p_i+q_i)\left[h\left(\left(\frac{2p_i}{p_i+q_i}\right)^{\alpha}\right)+h\left(\left(\frac{2q_i}{p_i+q_i}\right)^{\alpha}\right)\right]-4}{4(\beta-1)}.$$

Remark 5. *It could be seen that the lower bounds for both Jeffreys (25) and Jensen (32) Sharma–Mittal* (h,ϕ) *divergences are independent of the function h.*

Remark 6. *Taking into account the inequality (10) we obtain the alternative upper bound for the Jensen–Sharma–Mittal and the lower bound for the Jeffreys–Sharma–Mittal generalized* (h,ϕ) *divergences, respectively.*

$$\operatorname{Jen}SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q})\leq\frac{h\left(\left[\sum_{i=1}^{n}q_i\phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)+h\left(\left[\sum_{i=1}^{n}q_i\phi\left(\frac{p_i}{q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-2}{2(\beta-1)},$$

$$\operatorname{Jef}SM^{h,\phi}_{\alpha,\beta}(\mathbf{p},\mathbf{q})\geq\frac{h\left(\left[\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{2p_i}{p_i+q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-1}{\beta-1}+$$

$$\frac{h\left(\left[\sum_{i=1}^{n}\frac{p_i+q_i}{2}\phi\left(\frac{2q_i}{p_i+q_i},\alpha\right)\right]^{\frac{1-\beta}{1-\alpha}}\right)-1}{\beta-1}.$$

4. Applications

In this section we show how our theory works.

4.1. Bounds for Sharma–Mittal Divergences

For the functions $h(t)=t$, $\phi(t,\alpha)=t^{\alpha}$ and based on Theorems 1 and 3 we obtain the lower and upper bounds for Jeffreys–Sharma–Mittal and Jensen–Sharma–Mittal divergences, respectively, as follows

$$\frac{1}{\alpha-1}\log\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{\alpha(q_i-p_i)}\leq\operatorname{Jef}SM_{\alpha,\beta}(\mathbf{p},\mathbf{q})\leq\frac{\sum_{i=1}^{n}(p_iq_i)^{\alpha}\left(q_i^{1-2\alpha}+p_i^{1-2\alpha}\right)-2}{\beta-1}\tag{38}$$

$$\frac{1}{2(\alpha-1)}\log\prod_{i=1}^{n}\left(\frac{2\sqrt{p_iq_i}}{p_i+q_i}\right)^{\alpha(p_i+q_i)}\leq\operatorname{Jen}SM_{\alpha,\beta}(\mathbf{p},\mathbf{q})\leq\frac{\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)^{1-\alpha}\left(p_i^{\alpha}+q_i^{\alpha}\right)-2}{2(\beta-1)}.\tag{39}$$

Remark 7. *The above lower bounds (38) and (39) are the same for Rényi types divergences because they are independent of the parameter* β *which in that case approaches 1.*

Remark 8. *Substituting different values for the parameters* α, β, *such that* $1<\beta\leq\alpha$ *and taking into account the assumptions from the Theorems 1 and 3 about the functions h and* ϕ *we could formulate new types of divergences and related inequalities which are based on the generalized* (h,ϕ) *Sharma–Mittal divergence.*

4.2. Bounds for Tsallis Divergences

When we make the same assumptions as for Sharma–Mittal divergences with additional that $\beta = \alpha$ we obtain the bounds for Tsallis type divergences as follows

$$\frac{1}{\alpha-1}\log\prod_{i=1}^{n}\left(\frac{p_i}{q_i}\right)^{\alpha(q_i-p_i)} \leq \operatorname{Jef} T_\alpha(\mathbf{p},\mathbf{q}) \leq \frac{\sum_{i=1}^{n}(p_iq_i)^\alpha\left(q_i^{1-2\alpha}+p_i^{1-2\alpha}\right)-2}{\alpha-1}$$

$$\frac{1}{2(\alpha-1)}\log\prod_{i=1}^{n}\left(\frac{2\sqrt{p_iq_i}}{p_i+q_i}\right)^{\alpha(p_i+q_i)} \leq \operatorname{Jen} T_\alpha(\mathbf{p},\mathbf{q}) \leq \frac{\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)^{1-\alpha}\left(p_i^\alpha+q_i^\alpha\right)-2}{2(\alpha-1)}.$$

4.3. Bounds for Kullback–Leibler Divergences

When we have the same situation as in case of Tsallis divergence that is $h(t) = t$, $\phi(t,\alpha) = t^\alpha$, $\alpha = \beta$ and additionally both α and β approach 1 then we obtain new upper bounds for Jeffreys and Jensen–Shannon divergences, respectively.

$$\operatorname{Jef} S(\mathbf{p},\mathbf{q}) \leq \sum_{i=1}^{n}(p_i+q_i)\log p_iq_i,$$

$$\operatorname{Jen} S(\mathbf{p},\mathbf{q}) \leq \sum_{i=1}^{n}[p_i\log p_i + q_i\log q_i - (p_i+q_i)\log(p_i+q_i)].$$

The last inequality is equivalent to $\operatorname{Jen} S(\mathbf{p},\mathbf{q}) \geq 2log2$.

5. Summary

In this paper, new types of entropy have been defined, which are generalizations of others known and used so far in information theory.

The manuscript deals more with issues in the field of pure mathematics, therefore the standard axioms of entropy used in thermodynamics could, in this case, be extended by other assumptions and properties.

These divergences have been introduced for new physical interpretations which could be generated.

Generalized Sharma–Mittal and consequently Jensen–Sharma–Mittal and Jeffrey–Sharma–Mittal divergences have been defined for obtaining better estimates for known entropies, which will allow to more accurately determination of the dispersion measure of different distributions.

The derived inequalities have both upper and lower limits for the considered f-divergences. As a consequence, we obtain specific estimates for some new order measures. Hence they provide much wider interpretation possibilities in comparing probability distributions in the sense of mutual distances in different spaces.

In the era of advancing quantum mechanics, scientists are striving to build a quantum computer with very high computing power. The obtained results, despite their mathematical and analytical complexity, will very quickly generate specific numerical intervals which are an estimation of new introduced entropies. Therefore, such results as in this paper will be very useful in developing information theory issues.

This work is from the area of pure mathematics, therefore it is more theoretical than practical and makes it possible to find the existing known entropies by means of new defined generalizations. These generalizations can be used for interpreting various physical phenomena. The aim of this manuscript was to provide some new theoretical solutions for physicists who, with their knowledge and experience, will be able to look for new applications.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author wishes to thank anonymous referees for their helpful suggestions improving the readability of the paper.

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

References

1. Sharma, B.D.; Mittal, D.P. New nonadditive measures of inaccuracy. *J. Math. Sci.* **1975**, *10*, 122–133.
2. Üzengi Aktürk, O.; Aktürk, E.; Tomak, M. Can Sobolev inequality be written for Sharma–Mittal entropy? *Int. J. Theor. Phys.* **2008**, *47*, 3310–3320. [CrossRef]
3. Naudts, J. *Generalised Thermostatistics*; Springer: Berlin, Germany, 2011.
4. Elhoseiny, M.; Elgammal, A. Generalized Twin Gaussian Processes using Sharma–Mittal Divergence. *arXiv* **2015**, arXiv:1409.7480v5.
5. Koltcov, S.; Ignatenko, V.; Koltsova, O. Estimating Topic Modeling Performance with Sharma–Mittal Entropy. *Entropy* **2019**, *21*, 660, doi:10.3390/e21070660. [CrossRef] [PubMed]
6. Bo, L.; Sminchisescu, C. Twin gaussian processes for structured prediction. *Int. J. Comput. Vis.* **2010**, *87*, 28–52. [CrossRef]
7. Masi, M. A step beyond Tsallis and Rényi entropies. *Phys. Lett. A* **2005**, *338*, 217–224. [CrossRef]
8. Aktürk, E.; Bagci, G.; Sever, R. Is Sharma–Mittal entropy really a step beyond Tsallis and Rényi entropies? *arXiv* **2007**, arxiv:cond-mat/0703277.
9. Verma, R.; Merigó, J.M. On Sharma–Mittal's Entropy under Intuitionistic Fuzzy Environment. *Cybern. Syst.* **2021**, *52*, 498–521. [CrossRef]
10. Ghaffari, S.; Ziaie, A.H.; Moradpour, H.; Asghariyan, F.; Feleppa, F.; Tavayef, M. Black hole thermodynamics in Sharma–Mittal generalized entropy formalism. *Gen. Relativ. Gravit.* **2019**, *51*, 93, doi:10.1007/s10714-019-2578-2. [CrossRef]
11. Demirel, E.C.G. Dark energy model in higher–dimensional FRW universe with respect to generalized entropy of Sharma and Mittal of flat FRW space–time. *Can. J. Phys.* **2019**, *97*, 1185–1186. doi:10.1139/cjp-2018-0784. [CrossRef]
12. Frank, T.; Daffertshofer, A. Exact time–dependent solutions of the Rényi Fokker–Planck equation and the Fokker–Planck equations related to the entropies proposed by Sharma and Mittal. *Phys. A Stat. Mech. Appl.* **2000**, *285*, 351–366. [CrossRef]
13. Kluza, P.A.; Niezgoda, M. Inequalities for relative operator entropies. *Electron. J. Linear Algebra* **2014**, *27*, 851–864. [CrossRef]
14. Kluza, P.A.; Niezgoda, M. Generalizations of Crooks and Lin's results on Jeffreys-Csiszár and Jensen-Csiszár f-divergences. *Phys. A Stat. Mech. Appl.* **2016**, *463*, 383–393. doi:10.1016/j.physa.2016.07.062. [CrossRef]
15. Kluza, P.A.; Niezgoda, M. On Csiszár and Tsallis type f-divergences induced by superqudratic and convex functions. *Math. Inequal. Appl.* **2018**, *21*, 455–467.
16. Kluza, P.A. On Jensen–Rényi and Jeffreys-Rényi type f-divergences induced by convex functions. *Phys. A Stat. Mech. Appl.* **2020**, *548*, 1–10. doi:10.1016/j.physa.2019.122527. [CrossRef]
17. Crooks, G.E. On Measures of Entropy and Information. Available online: http://threeplusone.com/info (accessed on 22 September 2018).
18. Csiszár, I. Information–type measures of differences of probability distributions and indirect observations. *Stud. Sci. Math. Hung.* **1967**, *2*, 299–318.
19. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memory-Less Systems*; Academic Press: New York, NY, USA, 1981.
20. Dragomir, S.S. Inequalities for the Csiszár f-Divergence in Information Theory. RGMIA Monographs, Victoria University, 2000. Available online: http://rgmia.org/monographs/csiszar.htm (accessed on 1 February 2001).
21. Baez, J.C. Rényi entropy and free energy. *arXiv* **2011**, arXiv:1102.2098.
22. Dragomir, S.S; Pecaric, J.E.; Persson, L.E. Properties of some functionals related to Jensen's inequality. *Acta Math. Hungar.* **1996**, *70*, 129–143. [CrossRef]