

## Research Article

# The Unifying Frameworks of Information Measures

Shiwei Yu and Ting-Zhu Huang 

School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

Correspondence should be addressed to Ting-Zhu Huang; [tingzhu Huang@126.com](mailto:tingzhu Huang@126.com)

Received 23 July 2017; Accepted 6 February 2018; Published 8 March 2018

Academic Editor: Zhen-Lai Han

Copyright © 2018 Shiwei Yu and Ting-Zhu Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Information measures are capable of providing us with fundamental methodologies to analyze uncertainty and unveiling the substantive characteristics of random variables. In this paper, we address the issues of different types of entropies through  $q$ -generalized Kolmogorov-Nagumo averages, which lead to the propositions of the survival Rényi entropy and survival Tsallis entropy. Therefore, we make an inventory of eight types of entropies and then classify them into two categories: the density entropy that is defined on density functions and survival entropy that is defined on survival functions. This study demonstrates that, for each type of the density entropy, there exists a kind of the survival entropy corresponding to it. Furthermore, the similarity measures and normalized similarity measures are, respectively, proposed for each type of entropies. Generally, functionals of different types of information-theoretic metrics are equally diverse, while, simultaneously, they also exhibit some unifying features in all their manifestations. We present the unifying frameworks for entropies, similarity measures, and normalized similarity measures, which helps us deal with the available information measures as a whole and move from one functional to another in harmony with various applications.

## 1. Introduction

Measures of probabilistic uncertainty and information have attracted growing attentions since Hartley introduced the practical measure of information as the logarithm of the amount of uncertainty associated with finite possible symbol sequences, where the distribution of events is considered to be equally probable [1]. Today, entropy plays a basic role in the definitions of information measures with various applications in different areas. It has been recognized as the fundamental important field intersecting with mathematics, communication, physics, computer science, economics, and so forth [2–5].

The generalized information theory arising from the study of complex systems was intended to expand classical information theory based on probability. The additive probability measures, which are inherent in classical information theory, are extended to various types of nonadditive measures and thus result in different types of functionals that generalize Shannon entropy [6–8]. Generally, the formalization of uncertainty functions involves a considerable diversity. However, it also exhibits some unifying features [9].

*1.1. Entropies Defined on Density Functions.* We consider  $(X, Y)$  as the continuous random variables (r.v.) over a state space  $\Omega \times \Omega$  with the joint density function  $f(x, y)$  and marginal density functions  $f_X(x)$  and  $f_Y(y)$ . We also consider the conditional density function  $f_{X|Y}(x | y)$  of  $X$  given  $Y$  defined over  $\Omega$ . Note that  $f(x)$ ,  $f(y)$ , and  $f(x | y)$  are also used to mean  $f_X(x)$ ,  $f_Y(y)$ , and  $f_{X|Y}(x | y)$ , respectively, if their meanings are clear in context.

Let  $f(x)$  be a density function of r.v.  $X$  with  $\int_{\mathcal{R}} f(x) dx = 1$ . The Khinchin axioms [10] are capable of obtaining the Shannon entropy in a unique way. However, this may be too narrow-minded if one wants to describe complex systems. Therefore, a generalized measure of an r.v.  $X$  with respect to Kolmogorov-Nagumo (KN) averages [11] can be deduced as

$$\langle X \rangle_{\psi} = \psi^{-1} \left( \int_{\mathcal{R}} f(x) \psi \left( \log \frac{1}{f(x)} \right) dx \right), \quad (1)$$

where  $\psi$  is a continuous and strictly monotonic KN function [12] and hence has an inverse  $\psi^{-1}$ .

The KN averages can be extended in different manners to propose more generalized information measures. We use the  $q$ -logarithm function [13] given as

$$\ln_q x = \frac{x^{1-q} - 1}{1 - q}, \quad (x > 0, q \in \mathbb{R}) \quad (2)$$

to replace the logarithm function in (1). Note that  $\log x = \lim_{q \rightarrow 1} \ln_q x$  and  $\ln_q x$  satisfies pseudoadditivity; for example,  $\ln_q(xy) = \ln_q x + \ln_q y + (1 - q)\ln_q x \ln_q y$ . Hence we extend KN averages to a generalized measure of information  $X$  with respect to  $q$ -generalized KN averages defined by

$$\langle X \rangle_\psi^q = \psi^{-1} \left( \int_{\mathbb{R}} f(x) \psi \left( \ln_q \frac{1}{f(x)} \right) dx \right). \quad (3)$$

In terms of Rényi's generalization on axioms of KN averages [14], if  $q = 1$  and  $\psi(x) = x$  in (3), it yields Shannon entropy (SE) [15] defined as

$$H(X) = - \int_{\mathbb{R}} f(x) \log f(x) dx. \quad (4)$$

Based on Shannon entropy, the Shannon mutual information (SMI) [15, 16] of r.v.s  $X$  and  $Y$  was given by

$$\begin{aligned} I(X, Y) &= H(X) + H(Y) - H(X, Y) \\ &= H(X) - H(X | Y), \end{aligned} \quad (5)$$

where  $H(X, Y)$  is the joint Shannon entropy of  $(X, Y)$  and  $H(X | Y)$  is the conditional Shannon entropy of  $X$  given  $Y$ .

If  $q = 1$  and  $\psi$  is chosen as  $\psi(x) = e^{(1-\alpha)x}$  in (3), it yields Rényi entropy (RE) [14] defined by

$$R_\alpha(X) = \frac{1}{1 - \alpha} \log \int_{\mathbb{R}} f^\alpha(x) dx, \quad (6)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

Shannon entropy and Rényi entropy are additive. If  $q = \alpha$  and  $\psi(x) = x$  in (3), we get the pseudoadditive entropy or Tsallis entropy (TE) [17] defined by

$$\begin{aligned} T_\alpha(X) &= \int_{\mathbb{R}} f(x) \ln_\alpha \frac{1}{f(x)} dx \\ &= \frac{1}{\alpha - 1} \left( 1 - \int_{\mathbb{R}} f^\alpha(x) dx \right), \end{aligned} \quad (7)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

We obtain  $H(X) = \lim_{\alpha \rightarrow 1} R_\alpha(X)$  and  $H(X) = \lim_{\alpha \rightarrow 1} T_\alpha(X)$ . Therefore, Rényi entropy and Tsallis entropy can be viewed as interpolation formulas of the Shannon entropy ( $\alpha = 1$ ) and Hartley entropy ( $\alpha = 0$ ). A relation between Rényi and Tsallis entropies can be easily deduced as

$$T_\alpha(X) = \frac{1}{\alpha - 1} (1 - \exp((1 - \alpha) R_\alpha(X))). \quad (8)$$

More recently, interest in generalized information measures increases dramatically in different manners. A respectable number of nonclassical entropies, rather than

Shannon entropy, Rényi entropy, and Tsallis entropy, have already been developed in the study of complex systems.

The exponential entropy (EE) of order  $\alpha$  [18] was defined by

$$\mathcal{E}_\alpha(X) = \left( \int_{\mathbb{R}} f^\alpha(x) dx \right)^{1/(1-\alpha)}, \quad (9)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

We obtain  $R_\alpha(X) = \log \mathcal{E}_\alpha(X)$  and  $H(X) = \lim_{\alpha \rightarrow 1} \log \mathcal{E}_\alpha(X)$ .

**1.2. Entropies Defined on Survival Functions.** As narrated in [19], information measures defined on the density function suffer from several drawbacks, since the distribution function is more regular than the density function. Therefore, the cumulative residual entropy, which was defined on the cumulative distribution function or equivalently the survival function, was proposed as an alternative information measure of uncertainty.

Let  $X = (X_1, X_2, \dots, X_m)$  be a nonnegative r.v. in  $\mathbb{R}_+^m$ . We use the notation  $X > x$  to mean that  $X_i > x_i$  for  $x_i \geq 0$ ,  $i = 1, 2, \dots, m$ . The multivariate survival function of a nonnegative r.v.  $X$  is given as

$$\begin{aligned} \bar{F}_X(x) &= P(X > x) \\ &= P(X_1 > x_1, X_2 > x_2, \dots, X_m > x_m), \end{aligned} \quad (10)$$

where  $x = (x_1, \dots, x_m) \in \mathbb{R}_+^m$  with  $\mathbb{R}_+^m = \{x \in \mathbb{R}^m : x_i \geq 0, i = 1, \dots, m\}$ .

If the density function is replaced by the survival function,  $q$  is set as 1, and  $\psi(x) = x$  in (3), it yields the survival Shannon entropy (SSE) [19] defined as

$$\bar{\mathcal{H}}(X) = - \int_{\mathbb{R}_+^m} \bar{F}_X(x) \log \bar{F}_X(x) dx. \quad (11)$$

Since eight different types of entropies and their corresponding similarity measures will be discussed subsequently, it is worth pointing out that some notations and names of the existing information measures will be changed in harmony with the unifying frameworks throughout this paper.

To consider the conditional survival entropy, we denote  $F_{Y|X}(x | y)$  as the conditional distribution function of  $X$  given  $Y = y$  and also  $\bar{F}_{X|Y}(x | y)$  as the respective conditional survival function.

The cross survival Shannon entropy (CSSE) of r.v.s  $(X, Y)$  was given by [19]

$$\bar{\mathcal{F}}(X, Y) = \bar{\mathcal{H}}(X) - \bar{\mathcal{H}}(X | Y), \quad (12)$$

where  $\bar{\mathcal{H}}(X | Y)$  is the conditional survival Shannon entropy of r.v.s  $X$  given  $Y$  defined as [19]

$$\begin{aligned} \bar{\mathcal{H}}(X | Y) \\ = -E_Y \left( \int_{\mathbb{R}_+^m} \bar{F}_{X|Y}(x | y) \log \bar{F}_{X|Y}(x | y) dx \right) \end{aligned} \quad (13)$$

and  $E_Y(\cdot)$  here is the expectation with respect to an r.v.  $Y$ . The nonnegativity of CSSE was proven in [19] and thus CSSE was used as a similarity measure in image registration [20]. The generalized versions of SSE in dynamic systems were discussed in [21, 22].

If the density function in (9) is replaced by the survival function, this yields the survival exponential entropy (SEE) [23] of an r.v.  $X \in R_+^m$  with order  $\alpha$  given by

$$\overline{\mathcal{E}}_\alpha(X) = \left( \int_{R_+^m} \overline{F}_X^\alpha(x) dx \right)^{1/(1-\alpha)}, \quad (14)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

As an ongoing research program, generalized information measure offers us a steadily growing inventory of distinct entropy theories. Diversity and unity are two significant features of these theories. The growing diversity of information measures makes it increasingly more realistic to find a certain information measure suitable for a given condition. The unity allows us to view all available information measures as a whole and to move from one measure to another as needed. To that end, motivated by the researching approaches on Shannon entropy, Shannon mutual information [2], SSE [19], and SEE [23], we attempt to study information-theoretic metrics in their manifestations. On one hand, we propose several new types of entropies and their similarity measures; on the other hand, for each type of the existing entropies, except for Shannon entropy, we give the definitions of similarity measures (see Tables 1 and 2). Finally, we deduce the unifying frameworks for information measures emerging from the study of complex systems based on probability.

The remainder of this paper is organized as follows. Section 2 will propose the similarity measures defined on the density function. In Section 3, the survival Rényi entropy and survival Tsallis entropy are presented. In Section 4, we address the similarity measures defined on the survival function. The unifying frameworks of information measures and examples are provided in Section 5. Finally, we conclude this paper in Section 6.

## 2. Similarity Measures Defined on the Density Function

Shannon mutual information measures the information of an r.v.  $X$  conveying about another r.v.  $Y$ . It has been widely used in image registration [24, 25] and pattern recognition [26, 27]. Generally, as SMI is defined on Shannon entropy, each type of entropies would lead to corresponding similarity measures. In application, an idea similarity measure should be nonnegative. To that end, we take the way as [15, 19, 23] to define the similarity measures by linear expectation operator rather than KN average operator weighted by the escort distribution [28]. This section will present the similarity measures defined on the density function corresponding to Rényi entropy, Tsallis entropy, and the exponential entropy, respectively.

### 2.1. Rényi Mutual Information

**Lemma 1.** *Let  $X$  and  $Y$  be r.v.s and let  $\varphi(x)$  be a real convex function. Then*

$$E_Y \left( \int_R \varphi(f(x|y)) dx \right) \geq \int_R \varphi(f(x)) dx. \quad (15)$$

*If  $\varphi(x)$  is strictly convex, the equality holds if and only if  $X$  and  $Y$  are independent. If  $\varphi(x)$  is concave, the inequality is reversed.*

*Proof.* For a real convex function  $\varphi(x)$ , using Jensen's inequality [29], we obtain

$$E_Y(\varphi(f(x|y))) \geq \varphi(E_Y(f(x|y))). \quad (16)$$

The equality holds if  $X$  and  $Y$  are independent. Since  $f(x) = \int_R f(x|y)f(y)dy$ , it is immediate that

$$\begin{aligned} & E_Y \left( \int_R \varphi(f(x|y)) dx \right) \\ &= \int_R \left( \int_R \varphi(f(x|y)) dx \right) f(y) dy \\ &= \int_R \left( \int_R \varphi(f(x|y)) f(y) dy \right) dx \\ &= \int_R E_Y(\varphi(f(x|y))) dx \\ &\geq \int_R \varphi(E_Y(f(x|y))) dx \\ &= \int_R \varphi \left( \int_R f(x|y) f(y) dy \right) dx \\ &= \int_R \varphi(f(x)) dx. \end{aligned} \quad (17)$$

Now consider the "only if" part of the lemma. If  $X$  and  $Y$  are independent, then  $f(x|y) = f(x)$ , and thus the equality holds in (15). On the other hand, if the equality holds in (15), then the equality holds in Jensen's inequality (16), which leads to  $E_Y(\varphi(f(x|y))) = \varphi(E_Y(f(x|y)))$ . Then  $E_Y(f(x|y)) = f(x|y)$  almost surely, since  $\varphi$  is strictly convex. We obtain  $E_Y(f(x|y)) = f(x)$  by (18); hence  $f(x|y) = f(x)$ , which leads to the independence of  $X$  and  $Y$ .

Lemma 1 plays an important role to prove the nonnegativity for the similarity measure to be introduced, which is defined on the density function.  $\square$

**Definition 2.** Let  $X$  and  $Y$  be r.v.s; the conditional Rényi entropy of  $X$  given  $Y$  with order  $\alpha$  is defined by

$$R_\alpha(X|Y) = \frac{1}{1-\alpha} E_Y \left( \log \int_R f^\alpha(x|y) dx \right), \quad (18)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

Motivated by the definitions of the joint Shannon entropy  $H(X, Y) = H(X) + H(Y|X)$  and joint survival Shannon entropy  $\overline{\mathcal{H}}(X, Y) = \overline{\mathcal{H}}(X) + \overline{\mathcal{H}}(Y|X)$ , the joint Rényi entropy can be similarly introduced.

TABLE I: Entropies, conditional entropies, and joint entropies defined on density functions and survival functions.

Name	Density entropy	Name	Survival entropy
Shannon entropy	$H(X) = -\int_R f(x) \log f(x) dx$ $H(X Y) = -E_y \int_R f(x y) \log f(x y) dx$ $H(X, Y) = H(X) + H(Y X) $	Survival Shannon entropy	$\overline{H}(X) = -\int_{R_+^m} \overline{F}_X(x) \log \overline{F}_X(x) dx$ $\overline{H}(X Y) = -E_y \int_{R_+^m} \overline{F}_{X Y}(x y) \log \overline{F}_{X Y}(x y) dx$ $\overline{H}(X, Y) = \overline{H}(X) + \overline{H}(Y X) $
Rényi entropy	$R_\alpha(X) = \frac{1}{1-\alpha} \log \int_R f^\alpha(x) dx$ $R_\alpha(X Y) = \frac{1}{1-\alpha} E_y \log \int_R f^\alpha(x y) dx$ $R_\alpha(X, Y) = R_\alpha(X) + R_\alpha(Y X) $	Survival Rényi entropy	$\overline{R}_\alpha(X) = \frac{1}{1-\alpha} \log \int_{R_+^m} \overline{F}_X^\alpha(x) dx$ $\overline{R}_\alpha(X Y) = \frac{1}{1-\alpha} E_y \log \int_{R_+^m} \overline{F}_{X Y}^\alpha(x y) dx$ $\overline{R}_\alpha(X, Y) = \overline{R}_\alpha(X) + \overline{R}_\alpha(Y X) $
Tsallis entropy	$T_\alpha(X) = \frac{1}{\alpha-1} (1 - \int_R f^\alpha(x) dx)$ $T_\alpha(X Y) = \frac{1}{\alpha-1} E_y (1 - \int_R f^\alpha(x y) dx)$ $T_\alpha(X, Y) = T_\alpha(X) + T_\alpha(Y X) $	Survival Tsallis entropy	$\overline{T}_\alpha(X) = \frac{1}{\alpha-1} (\int_{R_+^m} \overline{F}_X(x) dx - \int_{R_+^m} \overline{F}_X^\alpha(x) dx)$ $\overline{T}_\alpha(X Y) = \frac{1}{\alpha-1} E_y (E_X(X Y=y) - \int_{R_+^m} \overline{F}_{X Y}^\alpha(x y) dx)$ $\overline{T}_\alpha(X, Y) = \overline{T}_\alpha(X) + \overline{T}_\alpha(Y X) $
Exponential entropy	$\mathcal{E}_\alpha(X Y) = E_y (\int_R f^\alpha(x y) dx)^{1/(1-\alpha)}$ $\mathcal{E}_\alpha(X, Y) = \mathcal{E}_\alpha(X) + \mathcal{E}_\alpha(Y X) $	Survival exponential entropy	$\overline{\mathcal{E}}_\alpha(X) = (\int_{R_+^m} \overline{F}_X(x) dx)^{1/(1-\alpha)}$ $\overline{\mathcal{E}}_\alpha(X Y) = E_y (\int_{R_+^m} \overline{F}_{X Y}^\alpha(x y) dx)^{1/(1-\alpha)}$ $\overline{\mathcal{E}}_\alpha(X, Y) = \overline{\mathcal{E}}_\alpha(X) + \overline{\mathcal{E}}_\alpha(Y X) $

TABLE 2: The similarity measures and normalized similarity measures defined on density functions and survival functions.

Name	Density similarity measure	Name	Survival similarity measure
Shannon mutual information	$I(X, Y) = H(X) - H(X Y)$	Cross survival Shannon entropy	$\overline{\mathcal{J}}(X, Y) = \overline{\mathcal{H}}(X) - \overline{\mathcal{H}}(X Y)$
	$NI(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)}$		$\overline{\mathcal{N}\mathcal{J}}(X, Y) = \frac{\overline{\mathcal{H}}(X) + \overline{\mathcal{H}}(Y)}{\overline{\mathcal{H}}(X, Y)}$
Rényi mutual information	$IR_\alpha(X, Y) = R_\alpha(X) - R_\alpha(X Y)$	Cross survival Rényi entropy	$\overline{\mathcal{J}}\overline{\mathcal{R}}_\alpha(X, Y) = \overline{\mathcal{R}}_\alpha(X) - \overline{\mathcal{R}}_\alpha(X Y)$
	$NIR_\alpha(X, Y) = \frac{R_\alpha(X) + R_\alpha(Y)}{R_\alpha(X, Y)}$		$\overline{\mathcal{N}\mathcal{J}}\overline{\mathcal{R}}_\alpha(X, Y) = \frac{\overline{\mathcal{R}}_\alpha(X) + \overline{\mathcal{R}}_\alpha(Y)}{\overline{\mathcal{R}}_\alpha(X, Y)}$
Tsallis mutual information	$IT_\alpha(X, Y) = T_\alpha(X) - T_\alpha(X Y)$	Cross survival Tsallis entropy	$\overline{\mathcal{J}}\overline{\mathcal{T}}_\alpha(X, Y) = \overline{\mathcal{T}}_\alpha(X) - \overline{\mathcal{T}}_\alpha(X Y)$
	$NIT_\alpha(X, Y) = \frac{T_\alpha(X) + T_\alpha(Y)}{T_\alpha(X, Y)}$		$\overline{\mathcal{N}\mathcal{J}}\overline{\mathcal{T}}_\alpha(X, Y) = \frac{\overline{\mathcal{T}}_\alpha(X) + \overline{\mathcal{T}}_\alpha(Y)}{\overline{\mathcal{T}}_\alpha(X, Y)}$
Exponential mutual information	$I\mathcal{E}_\alpha(X, Y) = \mathcal{E}_\alpha(X) - \mathcal{E}_\alpha(X Y)$	Cross survival exponential entropy	$\overline{\mathcal{J}}\overline{\mathcal{E}}_\alpha(X, Y) = \overline{\mathcal{E}}_\alpha(X) - \overline{\mathcal{E}}_\alpha(X Y)$
	$NI\mathcal{E}_\alpha(X, Y) = \frac{\mathcal{E}_\alpha(X) + \mathcal{E}_\alpha(Y)}{\mathcal{E}_\alpha(X, Y)}$		$\overline{\mathcal{N}\mathcal{J}}\overline{\mathcal{E}}_\alpha(X, Y) = \frac{\overline{\mathcal{E}}_\alpha(X) + \overline{\mathcal{E}}_\alpha(Y)}{\overline{\mathcal{E}}_\alpha(X, Y)}$

**Definition 3.** The joint Rényi entropy of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $R_\alpha(X, Y) = R_\alpha(X) + R_\alpha(Y | X)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Theorem 4.** For r.v.s  $X$  and  $Y$ , we obtain  $R_\alpha(X) \geq R_\alpha(X | Y)$  and  $R_\alpha(X, Y) \leq R_\alpha(X) + R_\alpha(Y)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* Since  $1/(1 - \alpha) > 0$  and  $t^\alpha$  is concave of  $t > 0$  for all  $0 < \alpha < 1$ , using Lemma 1 and Jensen's inequality, we have

$$\begin{aligned} R_\alpha(X) &= \frac{1}{1 - \alpha} \log \int_{\mathcal{R}} f^\alpha(x) dx \\ &\geq \frac{1}{1 - \alpha} \log E_Y \left( \int_{\mathcal{R}} f^\alpha(x | y) dx \right) \\ &\geq \frac{1}{1 - \alpha} E_Y \left( \log \int_{\mathcal{R}} f^\alpha(x | y) dx \right) \\ &= R_\alpha(X | Y). \end{aligned} \quad (19)$$

Since  $1/(1 - \alpha) < 0$  and  $t^\alpha$  is convex of  $t > 0$  for all  $\alpha > 1$ , using Lemma 1 and Jensen's inequality, we obtain

$$\begin{aligned} R_\alpha(X) &\geq \frac{1}{1 - \alpha} \log E_Y \left( \int_{\mathcal{R}} f^\alpha(x | y) dx \right) \\ &\geq \frac{1}{1 - \alpha} E_Y \left( \log \int_{\mathcal{R}} f^\alpha(x | y) dx \right) \\ &= R_\alpha(X | Y). \end{aligned} \quad (20)$$

Similarly, we have  $R_\alpha(Y) \geq R_\alpha(Y | X)$ , and thus  $R_\alpha(X, Y) = R_\alpha(X) + R_\alpha(Y | X) \leq R_\alpha(X) + R_\alpha(Y)$ .  $\square$

**Definition 5.** The Rényi mutual information (RMI) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as

$$\begin{aligned} IR_\alpha(X, Y) &= R_\alpha(X) + R_\alpha(Y) - R_\alpha(X, Y) \\ &= R_\alpha(X) - R_\alpha(X | Y), \end{aligned} \quad (21)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

It is worth pointing out that the definition of RMI parallels with the definitions of SMI (5) and CSSE (13). The nonnegativity of RMI is ensured by Theorem 4. Considering Theorem 4 that parallels with (5), we can give another form of the definition for RMI as

$$\begin{aligned} IR_\alpha^*(X, Y) &= R_\alpha(X) + R_\alpha(Y) - R_\alpha(X, Y) \\ &= R_\alpha(Y) - R_\alpha(Y | X). \end{aligned} \quad (22)$$

There are no essential differences between these two forms of definitions for RMI. We only consider the similar definitions as (24) for similarity measures throughout this paper.

Using L'Hôpital's rules, it is easy to obtain  $I(X, Y) = \lim_{\alpha \rightarrow 1} IR_\alpha(X, Y)$  and  $H(X, Y) = \lim_{\alpha \rightarrow 1} R_\alpha(X, Y)$ .

The normalized Shannon mutual information (NSMI) [16] of r.v.s  $X$  and  $Y$  was given as

$$NI(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)}. \quad (23)$$

NSMI often acts as a robust similarity measure in image registration [16, 30], attribute abstraction [31], and clustering [32]. Note that  $1 \leq NI(X, Y) \leq 2$ . In a similar way, different forms of the normalized mutual information will be deduced in this work.

**Definition 6.** The normalized Rényi mutual information (NRMI) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined by

$$NIR_\alpha(X, Y) = \frac{R_\alpha(X) + R_\alpha(Y)}{R_\alpha(X, Y)}, \quad (24)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

We immediately obtain  $NI(X, Y) = \lim_{\alpha \rightarrow 1} NIR_\alpha(X, Y)$  by L'Hôpital's rules.

## 2.2. Tsallis Mutual Information

**Definition 7.** The conditional Tsallis entropy of r.v.  $X$  given  $Y$  with order  $\alpha$  is defined as

$$T_\alpha(X | Y) = \frac{1}{\alpha - 1} E_Y \left( 1 - \int_{\mathcal{R}} f^\alpha(x | y) dx \right), \quad (25)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Definition 8.** The joint Tsallis entropy of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $T_\alpha(X, Y) = T_\alpha(X) + T_\alpha(Y | X)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Theorem 9.** For two r.v.s  $X$  and  $Y$ , we have  $T_\alpha(X) \geq T_\alpha(X | Y)$  and  $T_\alpha(X, Y) \leq T_\alpha(X) + T_\alpha(Y)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* Since  $1/(\alpha - 1) < 0$  and  $t^\alpha$  is concave of  $t$  for all  $0 < \alpha < 1$ , using Lemma 1, we obtain

$$\begin{aligned} T_\alpha(X) - E_Y(T_\alpha(X | Y)) &= \frac{1 - \int_{\mathcal{R}} f^\alpha(x) dx}{\alpha - 1} - E_Y \left( \frac{1 - \int_{\mathcal{R}} f^\alpha(x | y) dx}{\alpha - 1} \right) \\ &= \frac{1}{\alpha - 1} \left( E_Y \left( \int_{\mathcal{R}} f^\alpha(x | y) dx \right) - \int_{\mathcal{R}} f^\alpha(x) dx \right) \\ &\geq \frac{1}{\alpha - 1} \left( \int_{\mathcal{R}} f^\alpha(x) dx - \int_{\mathcal{R}} f^\alpha(x) dx \right) = 0. \end{aligned} \quad (26)$$

The inequality holds for all  $\alpha > 1$ , since  $1/(\alpha - 1) > 0$  and  $t^\alpha$  is convex of  $t$ .

It is trivial to verify that  $T_\alpha(X, Y) \leq T_\alpha(X) + T_\alpha(Y)$ .  $\square$

**Definition 10.** The Tsallis mutual information (TMI) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $IT_\alpha(X, Y) = T_\alpha(X) - T_\alpha(X | Y)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

Using Theorem 9, we have  $IT_\alpha(X, Y) \geq 0$ .

**Definition 11.** The normalized Tsallis mutual information (NTMI) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined by

$$NIT_\alpha(X, Y) = \frac{T_\alpha(X) + T_\alpha(Y)}{T_\alpha(X, Y)}, \quad (27)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

It is easy to verify that

$$\begin{aligned} H(X) &= \lim_{\alpha \rightarrow 1} T_\alpha(X), \\ H(X, Y) &= \lim_{\alpha \rightarrow 1} T_\alpha(X, Y), \\ I(X, Y) &= \lim_{\alpha \rightarrow 1} IT_\alpha(X, Y), \\ NI(X, Y) &= \lim_{\alpha \rightarrow 1} NIT_\alpha(X, Y), \end{aligned} \quad (28)$$

by L'Hôpital's rules, and  $T_\alpha(X, X) = T_\alpha(X)$ ,  $NIT_\alpha(X, Y) \geq 1$ .

### 2.3. Exponential Mutual Information

**Definition 12.** The conditional exponential entropy of r.v.  $X$  given  $Y$  with order  $\alpha$  is defined by

$$\mathcal{E}_\alpha(X | Y) = E_Y \left( \int_R f^\alpha(x | y) dx \right)^{1/(1-\alpha)}, \quad (29)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Definition 13.** The joint exponential entropy of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\mathcal{E}_\alpha(X, Y) = \mathcal{E}_\alpha(X) + \mathcal{E}_\alpha(Y | X)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Theorem 14.** For two r.v.s  $X$  and  $Y$ ,

$$\begin{aligned} & \left( E_Y \left( \int_R f^\alpha(x | y) dx \right) \right)^{1/(1-\alpha)} \\ & \leq \left( \int_R f^\alpha(x) dx \right)^{1/(1-\alpha)} \end{aligned} \quad (30)$$

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

*Proof.* Since  $t^\alpha$  is concave of  $t > 0$  for all  $0 < \alpha < 1$ , using Lemma 1, we obtain  $E_Y(\int_R f^\alpha(x | y) dx) \leq \int_R f^\alpha(x) dx$ . The equality is true, since  $1/(1-\alpha) > 1$  for all  $0 < \alpha < 1$ .

Similarly, since  $t^\alpha$  is convex of  $t > 0$  for all  $\alpha > 1$ , using Lemma 1, we obtain  $E_Y(\int_R f^\alpha(x | y) dx) \geq \int_R f^\alpha(x) dx$ . We complete the proof by considering that  $t^{1/(1-\alpha)}$  is decreasing in  $t > 0$  for all  $\alpha > 1$ .  $\square$

**Theorem 15.** For two r.v.s  $X$  and  $Y$ , we obtain  $\mathcal{E}_\alpha(X) \geq \mathcal{E}_\alpha(X | Y)$  and  $\mathcal{E}_\alpha(X, Y) \leq \mathcal{E}_\alpha(X) + \mathcal{E}_\alpha(Y)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* Since  $t^{1/(1-\alpha)}$  is convex of  $t > 0$  for all  $0 < \alpha < 1$ , using (30) and Jensen's inequality, we obtain

$$\begin{aligned} \mathcal{E}_\alpha(X) &= \left( \int_R f^\alpha(x) dx \right)^{1/(1-\alpha)} \\ &\geq \left( E_Y \left( \int_R f^\alpha(x | y) dx \right) \right)^{1/(1-\alpha)} \\ &\geq E_Y \left( \int_R f^\alpha(x | y) dx \right)^{1/(1-\alpha)} = \mathcal{E}_\alpha(X | Y). \end{aligned} \quad (31)$$

Since  $t^{1/(1-\alpha)}$  is concave and is decreasing in  $t > 0$  for all  $\alpha > 1$ , similarly we obtain

$$\begin{aligned} \mathcal{E}_\alpha(X) &\geq \left( E_Y \left( \int_R f^\alpha(x | y) dx \right) \right)^{1/(1-\alpha)} \\ &\geq E_Y \left( \int_R f^\alpha(x | y) dx \right)^{1/(1-\alpha)} \\ &= \mathcal{E}_\alpha(X | Y). \end{aligned} \quad (32)$$

It is trivial to verify that  $\mathcal{E}_\alpha(X, Y) \leq \mathcal{E}_\alpha(X) + \mathcal{E}_\alpha(Y)$ .  $\square$

**Definition 16.** The exponential mutual information (EMI) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $I\mathcal{E}_\alpha(X, Y) = \mathcal{E}_\alpha(X) - \mathcal{E}_\alpha(X | Y)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

Using Theorem 15, we have  $I\mathcal{E}_\alpha(X, Y) \geq 0$ .

**Definition 17.** The normalized exponential mutual information (NEMI) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined by

$$NI\mathcal{E}_\alpha(X, Y) = \frac{\mathcal{E}_\alpha(X) + \mathcal{E}_\alpha(Y)}{\mathcal{E}_\alpha(X, Y)}, \quad (33)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

## 3. Entropies Defined on the Survival Function

The existing survival Shannon entropy and the survival exponential entropy extended the corresponding functionals from the density function to the survival function. In this section, we will propose the survival Rényi entropy and the survival Tsallis entropy defined on the survival function which, respectively, parallel with the classical Rényi entropy and Tsallis entropy defined on the density function.

**3.1. Survival Rényi Entropy.** If the density function is replaced by the survival function,  $q$  is set as 1, and  $\psi$  is chosen as  $\psi(x) = e^{(1-\alpha)x}$  in (3), it yields the survival Rényi entropy.

**Definition 18.** The survival Rényi entropy (SRE) of an r.v.  $X$  with order  $\alpha$  is defined as

$$\overline{\mathcal{R}}_\alpha(X) = \frac{1}{1-\alpha} \log \int_{R_+^m} \overline{F}_X^\alpha(x) dx, \quad (34)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

It is immediate to see that  $\overline{\mathcal{E}}_\alpha(X) = \exp(\overline{\mathcal{R}}_\alpha(X))$ .

**Definition 19.** The conditional survival entropy of r.v.  $X$  given  $Y$  with order  $\alpha$  is defined as

$$\overline{\mathcal{R}}_\alpha(X | Y) = \frac{1}{1-\alpha} E_Y \left( \log \int_{R_+^m} \overline{F}_{X|Y}^\alpha(x | y) dx \right), \quad (35)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Definition 20.** The joint survival Rényi entropy of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\overline{\mathcal{R}}_\alpha(X, Y) = \overline{\mathcal{R}}_\alpha(X) + \overline{\mathcal{R}}_\alpha(Y | X)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Definition 21.** The joint survival exponential entropy of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\overline{\mathcal{E}}_\alpha(X, Y) = \overline{\mathcal{E}}_\alpha(X) + \overline{\mathcal{E}}_\alpha(Y | X)$ , where  $\alpha > 0$ ,  $\alpha \neq 1$ , and  $\overline{\mathcal{E}}_\alpha(Y | X) = \exp(\overline{\mathcal{R}}_\alpha(Y | X))$  is the conditional survival exponential entropy of r.v.  $Y$  given  $X$ .

Theorem 22 will show the relation between SRE and Shannon entropy.

**Theorem 22.** For an r.v.  $X$ , one has

$$\overline{\mathcal{R}}_\alpha(X) \geq \frac{1}{1-\alpha} (H(X) - \alpha), \quad 0 < \alpha < 1, \quad (36)$$

$$\overline{\mathcal{R}}_\alpha(X) \leq \frac{1}{1-\alpha} (H(X) - \alpha), \quad \alpha > 1.$$

*Proof.* Using the log-sum inequality [2, 19], we obtain

$$\int_{\mathbb{R}_+^m} f(x) \log \frac{f(x)}{\overline{F}_X^\alpha(x)} dx \geq -\log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx \quad (37)$$

and thus  $\int_{\mathbb{R}_+^m} f(x) \log f(x) dx - \int_{\mathbb{R}_+^m} f(x) \log \overline{F}_X^\alpha(x) dx \geq -\log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx$ . We obtain

$$\begin{aligned} H(X) + \alpha \int_{\mathbb{R}_+^m} f(x) \log \overline{F}_X(x) dx \\ \leq \log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx. \end{aligned} \quad (38)$$

Since  $d\overline{F}_X(x)/dx = -f(x)$ , (38) can be written as

$$\begin{aligned} \log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx \geq H(X) - \alpha \int_1^0 \log u du \\ = H(X) - \alpha. \end{aligned} \quad (39)$$

Note that  $1/(1-\alpha) > 0$  for all  $0 < \alpha < 1$  and  $1/(1-\alpha) < 0$  for all  $\alpha > 1$ . We complete the proof by multiplying  $1/(1-\alpha)$  on both sides of (39).  $\square$

**Theorem 23.** Let  $X$  and  $Y$  be two r.v.s; then

$$\overline{\mathcal{R}}_\alpha(X, Y) \geq \frac{1}{1-\alpha} (H(X, Y) - 2\alpha), \quad 0 < \alpha < 1, \quad (40)$$

$$\overline{\mathcal{R}}_\alpha(X, Y) \leq \frac{1}{1-\alpha} (H(X, Y) - 2\alpha), \quad \alpha > 1. \quad (41)$$

*Proof.* If  $0 < \alpha < 1$ , using Theorem 22, one has

$$\begin{aligned} \overline{\mathcal{R}}_\alpha(X, Y) &= \overline{\mathcal{R}}_\alpha(X) + \overline{\mathcal{R}}_\alpha(Y | X) \\ &\geq \frac{1}{1-\alpha} (H(X) + H(Y | X) - 2\alpha) \\ &= \frac{1}{1-\alpha} (H(X, Y) - 2\alpha). \end{aligned} \quad (42)$$

Similarly, the inequality in (41) holds for all  $\alpha > 1$ .

Since  $\overline{\mathcal{R}}_\alpha(X) = \log \overline{\mathcal{E}}_\alpha(X)$ , some properties of SRE can be similarly deduced by the proof approaches as SEE [23] and SSE [19]. We list these properties as propositions and neglect their proofs unless we can provide the improved, different, or more concise versions.  $\square$

**Proposition 24.** If  $E_X(X_i^p) < \infty$ , for some  $p > m/\alpha$ ,  $i = 1, 2, \dots, m$ , then  $|\overline{\mathcal{R}}_\alpha(X)| < \infty$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* For sets  $A_i \subseteq \mathcal{R}$ ,  $i = 1, 2, \dots, m$ , using Hölder's inequality [33], we have

$$P(A_1 \cap A_2 \cap \dots \cap A_m) \leq \prod_{i=1}^m P(A_i)^{1/m}. \quad (43)$$

For all  $\alpha > 0$ , we obtain

$$\begin{aligned} \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx &= \int_{\mathbb{R}_+^m} (P(X > x))^\alpha dx \\ &= \int_{\mathbb{R}_+^m} \prod_{i=1}^m (P(X_i > x_i))^\alpha dx \\ &\leq \int_{\mathbb{R}_+^m} \prod_{i=1}^m (P(X_i > x_i))^{\alpha/m} dx \\ &= \prod_{i=1}^m \int_0^\infty (P(X_i > x_i))^{\alpha/m} dx_i \\ &= \prod_{i=1}^m \left( \int_0^1 F_{X_i}^{\alpha/m}(x_i) dx_i + \int_1^\infty F_{X_i}^{\alpha/m}(x_i) dx_i \right) \\ &\leq \prod_{i=1}^m \left( 1 + \int_1^\infty \left( \frac{1}{x_i^p} E_X(X_i^p) \right)^{\alpha/m} dx_i \right). \end{aligned} \quad (44)$$

The inequality in (44) follows from Markov's inequality [19, 34], where the integral on the right side exists if  $p\alpha/m > 1$ , that is, if  $p < m/\alpha$ . Hence,  $\int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx$  exists for all  $\alpha > 0$ .

(i) If  $\int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx \geq 1$ , note that  $\log t < t$  for all  $t \geq 1$ ; then

$$\begin{aligned} |\overline{\mathcal{R}}_\alpha(X)| &= \left| \frac{1}{1-\alpha} \log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx \right| \\ &< \frac{1}{|1-\alpha|} \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx < \infty. \end{aligned} \quad (45)$$

(ii) If  $0 < \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx < 1$ , on one hand, for all  $\alpha > 1$ , since  $\int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx$  exists for all  $\alpha > 0$ , we see  $0 < \overline{\mathcal{R}}_\alpha(X) < \infty$ . On the other hand, for all  $0 < \alpha < 1$ , we obtain  $1 - \alpha > 0$  and  $\log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx < 0$ . We complete the proof, since  $-\infty < \overline{\mathcal{R}}_\alpha(X) < 0$  for all  $0 < \alpha < 1$ .  $\square$

**Proposition 25.** If the components  $X_i$ ,  $i = 1, 2, \dots, m$ , of r.v.  $X = (X_1, X_2, \dots, X_m)$  are independent, then  $\overline{\mathcal{R}}_\alpha(X) = \sum_{i=1}^m \overline{\mathcal{R}}_\alpha(X_i)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proposition 25 is the immediate result of Theorem 15 in [23].*

*The Shannon entropy of a sum of independent variables is larger than that of either; for example,  $H(X + Y) \geq \max\{H(X), H(Y)\}$ . SRE has this similar property.*

**Proposition 26.** Let  $X$  and  $Y$  be independent r.v.s; then  $\overline{\mathcal{R}}_\alpha(X + Y) \geq \max\{\overline{\mathcal{R}}_\alpha(X), \overline{\mathcal{R}}_\alpha(Y)\}$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .



*Proof.* Since  $X$  and  $Y$  are independent, one has  $P(X + Y > u) = \int f_Y(v)P(X > u - v)dv$ . Since  $1 - \alpha > 0$  and  $t^\alpha$  is concave of  $t > 0$  for all  $0 < \alpha < 1$ , using Jensen's inequality, one has

$$P^\alpha(X + Y > u) \geq \int f_Y(v)P^\alpha(X > u - v)dv. \quad (46)$$

Integrating both sides of  $u$  from 0 to  $\infty$ ,

$$\begin{aligned} & \int P^\alpha(X + Y > u) du \\ & \geq \int \int f_Y(u)P^\alpha(X > u - v)dv du \\ & = \int f_Y(v)dv \int_v^\infty P^\alpha(X > u - v) du \\ & = \int P^\alpha(X > x) dx. \end{aligned} \quad (47)$$

Then

$$\log \int_{R_+^m} P^\alpha(X + Y > u) du \geq \log \int_{R_+^m} P^\alpha(X > x) dx. \quad (48)$$

Multiplying by  $1/(1 - \alpha)$  on both sides of the above leads to  $\overline{\mathcal{R}}_\alpha(X + Y) \geq \overline{\mathcal{R}}_\alpha(X)$ , and exchanging  $X$  and  $Y$  leads to  $\overline{\mathcal{R}}_\alpha(X + Y) \geq \overline{\mathcal{R}}_\alpha(Y)$ .

We complete the proof by using Jensen's inequality and considering that  $1 - \alpha < 0$  and  $t^\alpha$  is convex of  $t > 0$  for all  $\alpha > 1$  in a similar way.  $\square$

**Proposition 27.** Let  $X = (X_1, X_2, \dots, X_m)$  and let  $Y = (Y_1, Y_2, \dots, Y_m)$  be r.v.s with  $Y_i = k_i X_i$ ,  $k_i \neq 0$ ,  $i = 1, 2, \dots, m$ , for a constant vector  $k = (k_1, k_2, \dots, k_m) \in R^m$ ; then  $\overline{\mathcal{R}}_\alpha(Y) = \sum_{i=1}^m \log|k_i| + \overline{\mathcal{R}}_\alpha(X)$  for all  $0 < \alpha < 1$ .

**Proposition 28.** Let  $X_n$  be a sequence of  $m$ -dimensional non-negative r.v.s converging in law to r.v.  $X$ . If all  $X_n$  are bounded in  $L^p$  for some  $p > m/\alpha$ , then  $\overline{\mathcal{R}}_\alpha(X) = \lim_{n \rightarrow \infty} \overline{\mathcal{R}}_\alpha(X_n)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

**Proposition 29.** If  $X_i$ ,  $i = 1, 2, \dots, m$ , of the r.v.  $X = (X_1, X_2, \dots, X_m)$  are independent, then  $\overline{\mathcal{R}}_\alpha(X) = \sum_{i=1}^m \overline{\mathcal{R}}_\alpha(X_i)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

**3.2. Survival Tsallis Entropy.** If the density function is replaced by the survival function,  $q = \alpha$ , and  $\psi(x) = x$  in (3), or, equivalently, the density function is replaced by the survival function and logarithm function is replaced by  $q$ -logarithm function in (7), it yields the survival Tsallis entropy.

**Definition 30.** The survival Tsallis entropy (STE) of an r.v.  $X$  with order  $\alpha$  is defined as

$$\begin{aligned} \overline{\mathcal{T}}_\alpha(X) &= \int_{R_+^m} \overline{F}_X(x) \ln_\alpha \frac{1}{\overline{F}_X(x)} dx \\ &= \frac{1}{\alpha - 1} \left( \int_{R_+^m} \overline{F}_X(x) dx - \int_{R_+^m} \overline{F}_X^\alpha(x) dx \right), \end{aligned} \quad (49)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

Note that  $\overline{F}_X(x) = \int_x^\infty f_X(t)dt$ . Using integration by parts formula, we obtain

$$\begin{aligned} \int_{R_+^m} \overline{F}_X(x) dx &= x\overline{F}_X(x) \Big|_0^\infty - \int_{R_+^m} x d\overline{F}_X(x) \\ &= \int_{R_+^m} x f_X(x) dx = E_X(X). \end{aligned} \quad (50)$$

Hence, the survival Tsallis entropy can also be written as  $\overline{\mathcal{T}}_\alpha(X) = (E_X(X) - \int_{R_+^m} \overline{F}_X^\alpha(x) dx)/(\alpha - 1)$ . It is easy to see that  $\overline{\mathcal{H}}(X) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{T}}_\alpha(X)$  and

$$\begin{aligned} \overline{\mathcal{T}}_\alpha(X) &= \frac{1}{\alpha - 1} \left( E_X(X) - \exp\left((1 - \alpha)\overline{\mathcal{R}}_\alpha(X)\right) \right) \\ &= \frac{1}{\alpha - 1} \left( E_X(X) - (1 - \alpha) \log \overline{\mathcal{E}}_\alpha(X) \right). \end{aligned} \quad (51)$$

**Definition 31.** The conditional survival Tsallis entropy of r.v.  $X$  given  $Y$  with order  $\alpha$  is defined as

$$\begin{aligned} \overline{\mathcal{T}}_\alpha(X | Y) &= \frac{1}{\alpha - 1} \\ &\cdot E_Y \left( E_X(X | Y = y) - \int_{R_+^m} \overline{F}_{X|Y}^\alpha(x | y) dx \right), \end{aligned} \quad (52)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

Note that  $E_Y(E_X(X | Y = y)) = E_Y(\int_{R_+^m} \overline{F}_{X|Y}(x | y) dx) = E_X(X)$ . Then  $\overline{\mathcal{T}}_\alpha(X | Y) = (1/(\alpha - 1))(E_X(X) - E_Y \int_{R_+^m} \overline{F}_{X|Y}^\alpha(x | y) dx)$ .

**Definition 32.** The joint survival Tsallis entropy of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\overline{\mathcal{T}}_\alpha(X, Y) = \overline{\mathcal{T}}_\alpha(X) + \overline{\mathcal{T}}_\alpha(Y | X)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

It is easy to see that  $\overline{\mathcal{H}}(X) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{T}}_\alpha(X)$  and  $\overline{\mathcal{H}}(X, Y) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{T}}_\alpha(X, Y)$  using L'Hôpital's rules.

**Theorem 33.** Let  $X$  be an r.v.; then

$$\begin{aligned} \overline{\mathcal{T}}_\alpha(X) &\geq \frac{1}{\alpha - 1} \left( E_X(X) - e^{H(X) - \alpha} \right), \quad 0 < \alpha < 1, \\ \overline{\mathcal{T}}_\alpha(X) &\leq \frac{1}{\alpha - 1} \left( E_X(X) - e^{H(X) - \alpha} \right), \quad \alpha > 1. \end{aligned} \quad (53)$$

*Proof.* It is easy to verify using (51) and Theorem 22.

Since there is a relation among  $\overline{\mathcal{T}}_\alpha(X)$ ,  $\overline{\mathcal{R}}_\alpha(X)$ , and  $\overline{\mathcal{E}}_\alpha(X)$  by (51), some properties of STE can be deduced by the theorems of SEE and SRE. We only list these properties as propositions and provide necessary explanations for their proofs.  $\square$

**Proposition 34.** If  $E(X_i^p) < \infty$  for some  $p > m/\alpha$ ,  $i = 1, 2, \dots, m$ , then  $\overline{\mathcal{T}}_\alpha(X) < \infty$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* It is easy to verify that  $\overline{\mathcal{F}}_\alpha(X) \geq 0$ . By the proof of Proposition 24, we see that  $\int_{\mathbb{R}_+^m} \overline{F}_X(x) dx$  and  $\int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx$  exist. Thus,  $\overline{\mathcal{F}}_\alpha(X) < \infty$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .  $\square$

**Proposition 35.** Let  $X = (X_1, X_2, \dots, X_m)$  and let  $Y = (Y_1, Y_2, \dots, Y_m)$  be r.v.s with  $Y_i = k_i X_i$ ,  $k_i \neq 0$ ,  $i = 1, 2, \dots, m$ , and let  $k = (k_1, k_2, \dots, k_m) \in \mathbb{R}^m$  be a constant vector; then  $\overline{\mathcal{F}}_\alpha(Y) = \prod_{i=1}^m |k_i| \overline{\mathcal{F}}_\alpha(X)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

This proposition can be proven using (30) in [23].

**Proposition 36.** Let  $X_n$  be a sequence of  $m$ -dimensional nonnegative r.v.s converging to  $X$ :  $0 < \alpha < 1$ . If all  $X_n$  are bounded in  $L^p$  for some  $p > m/\alpha$ , then  $\overline{\mathcal{F}}_\alpha(X) = \lim_{n \rightarrow \infty} \overline{\mathcal{F}}_\alpha(X_n)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

It is immediate using (38) in [23].

**Proposition 37.** Let  $X$  and  $Y$  be nonnegative and independent r.v.s; then  $\overline{\mathcal{F}}_\alpha(X+Y) \geq \max\{\overline{\mathcal{F}}_\alpha(X), \overline{\mathcal{F}}_\alpha(Y)\}$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* It can be proven in a similar way as Proposition 26 by considering  $\alpha - 1 < 0$  and  $t - t^\alpha$  is convex of  $t > 0$  for all  $0 < \alpha < 1$  and  $\alpha - 1 > 0$  and  $t - t^\alpha$  is concave of  $t > 0$  for all  $\alpha > 1$ .  $\square$

## 4. Similarity Measures Defined on the Survival Function

Paralleling with the similarity measures and the normalized similarity measures defined on the density function in Section 2, this section will focus on the corresponding similarity measures and the normalized similarity measures defined on the survival function. Traditionally, the kernel point is the proofs of the nonnegativity for the similarity measures to be introduced.

### 4.1. Cross Survival Rényi Entropy and Cross Survival Exponential Entropy

**Lemma 38.** Let  $X$  and  $Y$  be r.v.s and let  $\varphi(x)$  be a real convex function; then

$$E_Y \left( \int_{\mathbb{R}_+^m} \varphi(\overline{F}_{X|Y}(x|y)) dx \right) \geq \int_{\mathbb{R}_+^m} \varphi(\overline{F}_X(x)) dx. \quad (54)$$

If, moreover,  $\varphi(x)$  is strictly convex, then equality holds in (54) if and only if  $X$  and  $Y$  are independent. The inequality is reversed if  $\varphi(x)$  is concave.

Lemma 38 was proven in [23]. It is a cornerstone to prove the nonnegativity of each form of the similarity measure to be introduced according to the survival function.

**Theorem 39.** Let  $X$  and  $Y$  be r.v.s; then  $\overline{\mathcal{R}}_\alpha(X) \geq \overline{\mathcal{R}}_\alpha(X|Y)$  and  $\overline{\mathcal{E}}_\alpha(X) \geq \overline{\mathcal{E}}_\alpha(X|Y)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* Since  $t^\alpha$  is concave of  $t > 0$  and  $1/(1-\alpha) > 0$  for all  $0 < \alpha < 1$ , note that  $\log t$  is strictly concave of  $t$ . Using Lemma 38 and Jensen's inequality, we obtain

$$\begin{aligned} \overline{\mathcal{R}}_\alpha(X) &= \frac{1}{1-\alpha} \log \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx \\ &\geq \frac{1}{1-\alpha} \log E_Y \left( \int_{\mathbb{R}_+^m} \overline{F}_{X|Y}^\alpha(x|y) dx \right) \\ &\geq \frac{1}{1-\alpha} E_Y \left( \log \int_{\mathbb{R}_+^m} \overline{F}_{X|Y}^\alpha(x|y) dx \right) \\ &= E_Y \left( \frac{1}{1-\alpha} \log \int_{\mathbb{R}_+^m} \overline{F}_{X|Y}^\alpha(x|y) dx \right) \\ &= \overline{\mathcal{R}}_\alpha(X|Y). \end{aligned} \quad (55)$$

Similarly, considering that  $t^\alpha$  is convex of  $t > 0$  and  $1/(1-\alpha) < 0$  for all  $\alpha > 1$ , the conclusion is the same.

It is trivial to verify that  $\overline{\mathcal{E}}_\alpha(X) \geq \overline{\mathcal{E}}_\alpha(X|Y)$ .

For r.v.s  $X$  and  $Y$ , using Theorem 39, we have

$$\begin{aligned} &\left( \int_{\mathbb{R}_+^m} \overline{F}_X^\alpha(x) dx \right)^{1/(1-\alpha)} \\ &\geq \left( E_Y \int_{\mathbb{R}_+^m} \overline{F}_{X|Y}^\alpha(x|y) dx \right)^{1/(1-\alpha)} \end{aligned} \quad (56)$$

for all  $0 < \alpha$  and  $\alpha \neq 1$ .  $\square$

**Definition 40.** The cross survival Rényi entropy (CSRE) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\overline{\mathcal{F}}\overline{\mathcal{R}}_\alpha(X, Y) = \overline{\mathcal{R}}_\alpha(X) - \overline{\mathcal{R}}_\alpha(X|Y)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Definition 41.** The cross survival exponential entropy (CSEE) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\overline{\mathcal{F}}\overline{\mathcal{E}}_\alpha(X, Y) = \overline{\mathcal{E}}_\alpha(X) - \overline{\mathcal{E}}_\alpha(X|Y)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

Using Theorem 39 and (56), we obtain  $\overline{\mathcal{F}}\overline{\mathcal{R}}_\alpha(X, Y) \geq 0$  and  $\overline{\mathcal{F}}\overline{\mathcal{E}}_\alpha(X, Y) \geq 0$ .

**Definition 42.** The normalized cross survival Rényi entropy (NCSRE) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as

$$\overline{\mathcal{N}}\overline{\mathcal{F}}\overline{\mathcal{R}}_\alpha(X, Y) = \frac{\overline{\mathcal{R}}_\alpha(X) + \overline{\mathcal{R}}_\alpha(Y)}{\overline{\mathcal{R}}_\alpha(X, Y)}, \quad (57)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

**Definition 43.** The normalized cross survival exponential entropy (NCSEE) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as

$$\overline{\mathcal{N}}\overline{\mathcal{F}}\overline{\mathcal{E}}_\alpha(X, Y) = \frac{\overline{\mathcal{E}}_\alpha(X) + \overline{\mathcal{E}}_\alpha(Y)}{\overline{\mathcal{E}}_\alpha(X, Y)}, \quad (58)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

*Definition 44.* The normalized cross survival Shannon entropy (NCSSE) of r.v.s  $X$  and  $Y$  is defined as

$$\overline{\mathcal{N}\mathcal{F}}(X, Y) = \frac{\overline{\mathcal{H}}(X) + \overline{\mathcal{H}}(Y)}{\overline{\mathcal{H}}(X, Y)}. \quad (59)$$

**Proposition 45.** Let  $X$  and  $Y$  be two r.v.s; then  $\overline{\mathcal{H}}(X, Y) \leq \overline{\mathcal{H}}(X) + \overline{\mathcal{H}}(Y)$ ,  $\overline{\mathcal{R}}_\alpha(X, Y) \leq \overline{\mathcal{R}}_\alpha(X) + \overline{\mathcal{R}}_\alpha(Y)$ , and  $\overline{\mathcal{E}}_\alpha(X, Y) \leq \overline{\mathcal{E}}_\alpha(X) + \overline{\mathcal{E}}_\alpha(Y)$ , for  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* Since  $-t \log t$  is convex of  $t$  for  $t > 0$ , using Lemma 38, it is immediate that  $\overline{\mathcal{H}}(X, Y) \leq \overline{\mathcal{H}}(X) + \overline{\mathcal{H}}(Y)$ . The rest can be similarly proven by considering the range of  $\alpha$  and using Lemma 38.

Using L'Hôpital's rules, it is easy to see that  $\overline{\mathcal{F}}(X, Y) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{F}\mathcal{R}}_\alpha(X, Y)$  and  $\overline{\mathcal{N}\mathcal{F}}(X, Y) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{N}\mathcal{F}\mathcal{R}}_\alpha(X, Y)$ .  $\square$

#### 4.2. Cross Survival Tsallis Entropy

**Theorem 46.** For two r.v.s  $X$  and  $Y$ , one has  $\overline{\mathcal{T}}_\alpha(X) \geq \overline{\mathcal{T}}_\alpha(X | Y)$  and  $\overline{\mathcal{T}}_\alpha(X, Y) \leq \overline{\mathcal{T}}_\alpha(X) + \overline{\mathcal{T}}_\alpha(Y)$  for all  $\alpha > 0$  and  $\alpha \neq 1$ .

*Proof.* Since  $\overline{F}_X(x) = E_Y(\overline{F}_{X|Y}(x | y))$ , using Lemma 38, considering  $1/(\alpha - 1) < 0$  and  $t^\alpha$  is concave of  $t > 0$  for all  $0 < \alpha < 1$ , we obtain

$$\begin{aligned} & \overline{\mathcal{T}}_\alpha(X) - \overline{\mathcal{T}}_\alpha(X | Y) \\ &= \frac{1}{\alpha - 1} \left( E_Y \left( \int_{R_+^m} \overline{F}_{X|Y}^\alpha(x | y) dx \right) \right. \\ & \quad \left. - \int_{R_+^m} \overline{F}_X^\alpha(x) dx \right) \geq \frac{1}{\alpha - 1} \left( \int_{R_+^m} \overline{F}_X^\alpha(x) dx \right. \\ & \quad \left. - \int_{R_+^m} \overline{F}_X^\alpha(x) dx \right) = 0. \end{aligned} \quad (60)$$

For  $\alpha > 1$ , similarly the conclusion is the same.

It is immediate to deduce the rest of conclusion by Definition 32.  $\square$

*Definition 47.* The cross survival Tsallis entropy (CSTE) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as  $\overline{\mathcal{F}\mathcal{T}}_\alpha(X, Y) = \overline{\mathcal{T}}_\alpha(X) - \overline{\mathcal{T}}_\alpha(X | Y)$ , where  $\alpha > 0$  and  $\alpha \neq 1$ .

Using Theorem 46, we obtain  $\overline{\mathcal{F}\mathcal{T}}_\alpha(X, Y) \geq 0$  immediately.

*Definition 48.* The normalized cross survival Tsallis entropy (NCSTE) of r.v.s  $X$  and  $Y$  with order  $\alpha$  is defined as

$$\overline{\mathcal{N}\mathcal{F}\mathcal{T}}_\alpha(X, Y) = \frac{\overline{\mathcal{T}}_\alpha(X) + \overline{\mathcal{T}}_\alpha(Y)}{\overline{\mathcal{T}}_\alpha(X, Y)}, \quad (61)$$

where  $\alpha > 0$  and  $\alpha \neq 1$ .

We obtain  $\overline{\mathcal{F}}(X, Y) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{F}\mathcal{T}}_\alpha(X, Y)$  and  $\overline{\mathcal{N}\mathcal{F}}(X, Y) = \lim_{\alpha \rightarrow 1} \overline{\mathcal{N}\mathcal{F}\mathcal{T}}_\alpha(X, Y)$  by L'Hôpital's rules.

Note that  $\overline{\mathcal{F}}(X, Y) \neq \overline{\mathcal{F}}(Y, X)$ ,  $\overline{\mathcal{F}\mathcal{R}}_\alpha(X, Y) \neq \overline{\mathcal{F}\mathcal{R}}_\alpha(Y, X)$ ,  $\overline{\mathcal{F}\mathcal{E}}_\alpha(X, Y) \neq \overline{\mathcal{F}\mathcal{E}}_\alpha(Y, X)$ , and  $\overline{\mathcal{F}\mathcal{T}}_\alpha(X, Y) \neq \overline{\mathcal{F}\mathcal{T}}_\alpha(Y, X)$ , whereas  $I(X, Y)$  is symmetric; for example,

$I(X, Y) = I(Y, X)$ . We can define the symmetric versions of the similarity measures and the normalized similarity measures, taking the cross survival Rényi entropy as an example, as

$$\begin{aligned} \overline{\mathcal{S}\mathcal{F}\mathcal{R}}_\alpha(X, Y) &= \frac{1}{2} \left( \overline{\mathcal{F}\mathcal{R}}_\alpha(X, Y) + \overline{\mathcal{F}\mathcal{R}}_\alpha(Y, X) \right), \\ \overline{\mathcal{S}\mathcal{N}\mathcal{F}\mathcal{R}}_\alpha(X, Y) &= \frac{1}{2} \left( \overline{\mathcal{N}\mathcal{F}\mathcal{R}}_\alpha(X, Y) + \overline{\mathcal{N}\mathcal{F}\mathcal{R}}_\alpha(Y, X) \right). \end{aligned} \quad (62)$$

The similar way can be used to define the symmetric similarity measures and normalized similarity measures for those defined on the density function.

## 5. Unifying Frameworks and Examples

In this section, based on the generalized denotations on the entropies discussed previously, we will classify the fourteen types of entropies in two categories and then deduce the unifying presentations for entropies, similarity measures, and normalized similarity measures. Examples are also provided to unveil some properties of the information measures.

As enumerated in Table 1, different types of entropies have been discussed in this paper. There are three components in each item: entropy, conditional entropy, and joint entropy. In general, entropies in Table 1 can be classified into two categories: one is defined on the density function and the other is defined on the survival function. For simplicity, we refer to them as the density entropy and survival entropy, respectively. It is demonstrated that, for each type of the density entropy in Column 2, there is a survival entropy in Column 4 corresponding to it.

*5.1. The Unifying Frameworks of Information Measures.* For convenience, we view Shannon entropy, Rényi entropy, Tsallis entropy, and the exponential entropy as the classical density entropy and view their corresponding survival entropies as the classical survival entropy. We can see that the classical density entropy, the classical survival entropy, and their conditional entropy and joint entropy share similar presentations.

Let  $\mathcal{H}_\alpha(X)$  be one type of the generalized density entropy or the generalized survival entropy of r.v.  $X$  with order  $\alpha$ . If  $\alpha = 1$ , then  $\mathcal{H}_1(X) = \mathcal{H}(X)$  means the Shannon entropy or the survival Shannon entropy. In these notations,  $\mathcal{H}_\alpha(X | Y)$  is the conditional entropy with order  $\alpha > 0$ . The corresponding joint entropy of r.v.s  $X$  and  $Y$  with order  $\alpha > 0$  can be introduced as

$$\mathcal{H}_\alpha(X, Y) = \mathcal{H}_\alpha(X) + \mathcal{H}_\alpha(Y | X). \quad (63)$$

For r.v.s  $X$  and  $Y$ , one has  $\mathcal{H}_\alpha(X, Y) \leq \mathcal{H}_\alpha(X) + \mathcal{H}_\alpha(Y)$  for all  $\alpha > 0$ .

Entropies, conditional entropies, and joint entropies are listed in Table 1. The similarity measures and the normalized similarity measures are shown in Table 2 in detail, where the similarity measure is followed by the normalized one in each item. In a similar way, the similarity measure can

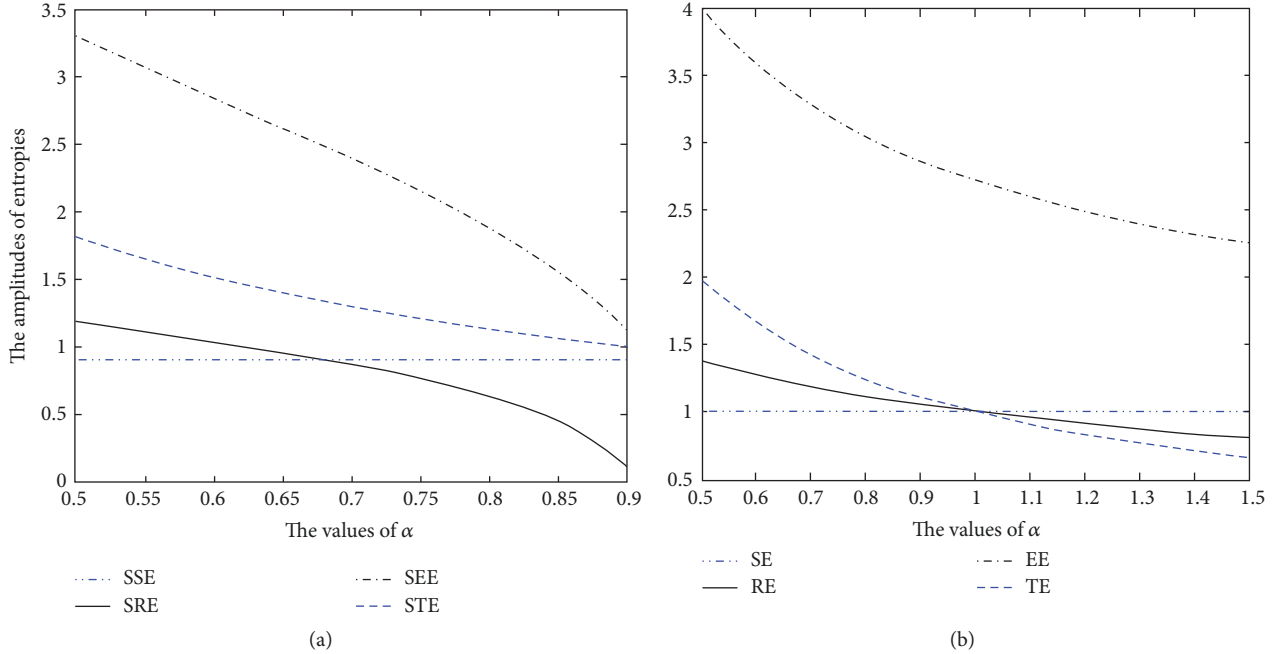


FIGURE 1: Plots of eight different types of entropies. (a) Survival entropies. (b) Density entropies.

be classified into the density similarity measure defined on the density function and the survival similarity measure defined on the survival function and so can the normalized similarity measures. Therefore, the unifying presentations for a similarity and a normalized similarity measure associated with a type of entropies can be deduced as

$$\mathcal{F}_\alpha(X, Y) = \mathcal{H}_\alpha(X) + \mathcal{H}_\alpha(X | Y), \quad (64)$$

$$\mathcal{N}\mathcal{F}_\alpha(X, Y) = \frac{\mathcal{H}_\alpha(X) + \mathcal{H}_\alpha(Y)}{\mathcal{H}_\alpha(X, Y)}. \quad (65)$$

Note that  $\mathcal{F}_1(X, Y) = \mathcal{F}(X, Y)$  and  $\mathcal{N}\mathcal{F}_1(X, Y) = \mathcal{N}\mathcal{F}(X, Y)$ .

We obtain  $\mathcal{F}_\alpha(X, Y) \geq 0$  for all  $\alpha > 0$ . Their symmetric versions can be, respectively, given by

$$\mathcal{S}\mathcal{F}_\alpha(X, Y) = \frac{1}{2} (\mathcal{F}_\alpha(X, Y) + \mathcal{F}_\alpha(Y, X)), \quad (66)$$

$$\mathcal{S}\mathcal{N}\mathcal{F}_\alpha(X, Y) = \frac{1}{2} (\mathcal{N}\mathcal{F}_\alpha(X, Y) + \mathcal{N}\mathcal{F}_\alpha(Y, X)). \quad (67)$$

The unifying frameworks make it possible to view all the available entropies listed in Table 1 as a whole and to move from one to another as necessary. Subsequently, the similarity measures and the normalized similarity measures are simultaneously obtained.

### 5.2. Three Examples.

*Example 1.* Let  $X$  be an r.v. corresponding to the exponential distribution with mean  $1/\lambda$  and density function  $f(x) = \lambda e^{-\lambda x}$ . We obtain  $\bar{F}(x) = \int_x^\infty f(t)dt = e^{-\lambda x}$ , and

$$\begin{aligned} \bar{\mathcal{H}}(X) &= \frac{1}{\lambda}, \\ \bar{\mathcal{R}}_\alpha(X) &= \frac{1}{\alpha - 1} \log \lambda \alpha, \\ \bar{\mathcal{F}}_\alpha(X) &= \frac{1}{\lambda \alpha}, \\ \bar{\mathcal{E}}_\alpha(X) &= (\lambda \alpha)^{1/(\alpha-1)}. \end{aligned} \quad (68)$$

The survival Shannon entropy, survival Rényi entropy, survival Tsallis entropy, and survival exponential entropy of r.v.  $X$  are plotted in Figure 1(a) with  $0.5 \leq \alpha \leq 0.9$  and  $\lambda = 1.1$ .

If  $\lambda = 1$ , then  $\bar{F}(x) = e^{-x} = f(x)$ . Hence, the survival entropies become the corresponding density entropies. Figure 1(b) shows Shannon entropy, Rényi entropy, Tsallis entropy, and the exponential entropy of an r.v.  $X$  with  $0.5 \leq \alpha \leq 1.5$ . We can see that some properties, such as concavity and monotonicity, may be changed when we generalize entropies by extending their definitions from the density function to the survival function.

*Example 2.* Many literatures have pointed out that Shannon entropy has drawbacks: each frequency of the occurrence event contributes equally in summation or integral of its functional and, simultaneously, the spatial information is neglected. In fact, those entropies that are defined on density functions have similar drawbacks to Shannon entropy. However, the survival entropy can overcome this drawback. For instance, as shown in Figure 2, Lena is used as a fixed image with size of  $256 \times 256$ . We exchange those pixels with the same occurrence frequency in Figure 2(a) and then generate the negative image as shown in Figure 2(b). Images (c) and (d) are their histograms, respectively. Thereby images (a) and (b) share the same occurrence frequency but specified to

TABLE 3: The values acquired from different types of entropy for testing images.

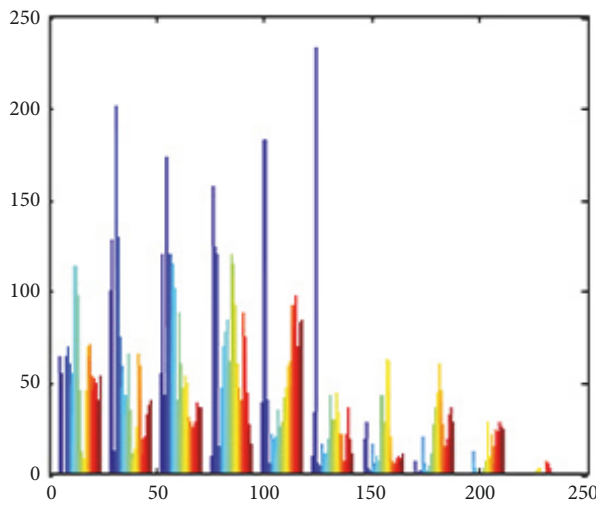
Image	SE	RE	TE	SSE	SRE	STE
(a)	5.1418	5.1755	9.0768	51.0285	22.9235	58.8816
(b)	5.1418	5.1755	9.0768	46.9452	25.9032	51.0613



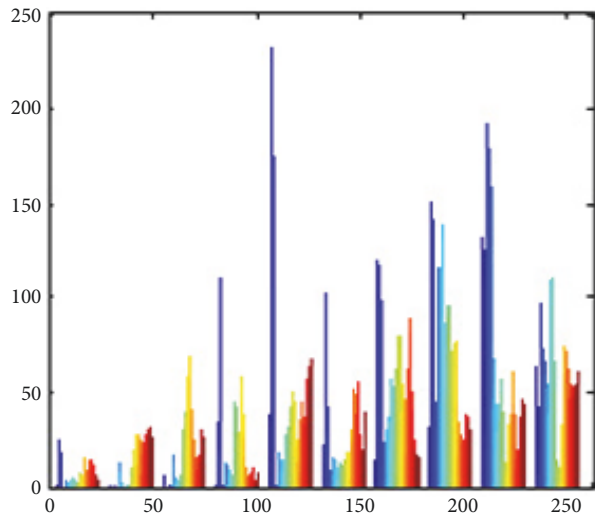
(a)



(b)



(c)



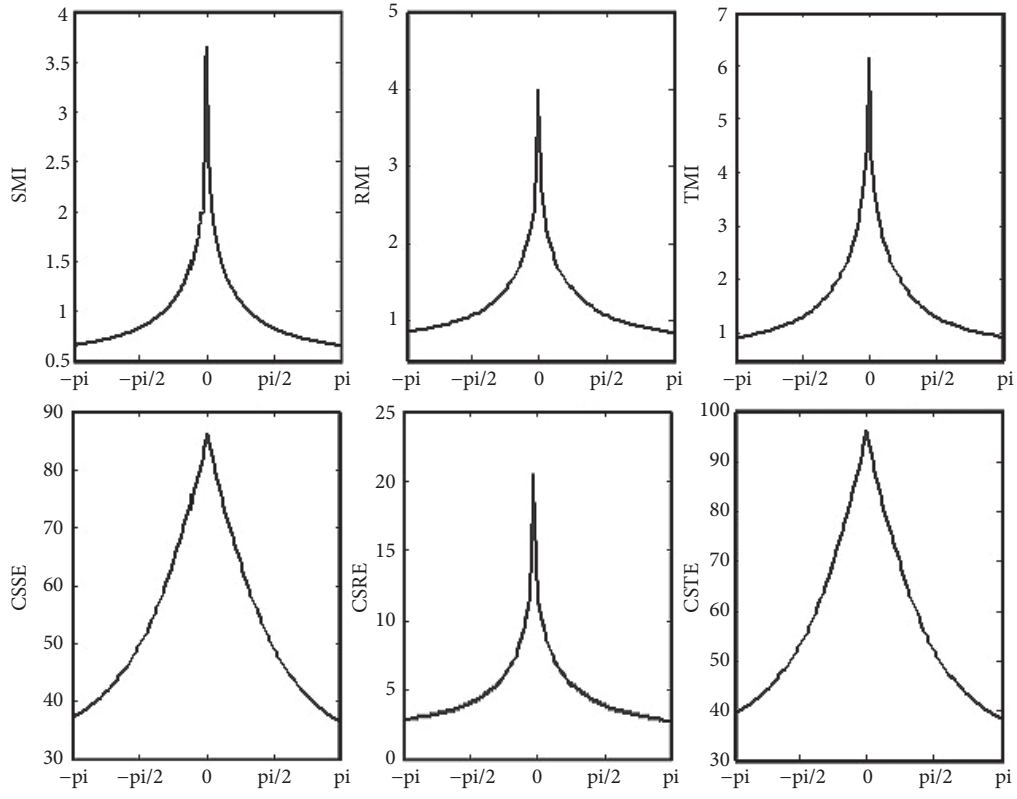
(d)

FIGURE 2: Testing images ((a) and (b)) and their histograms ((c) and (d)) accordingly.

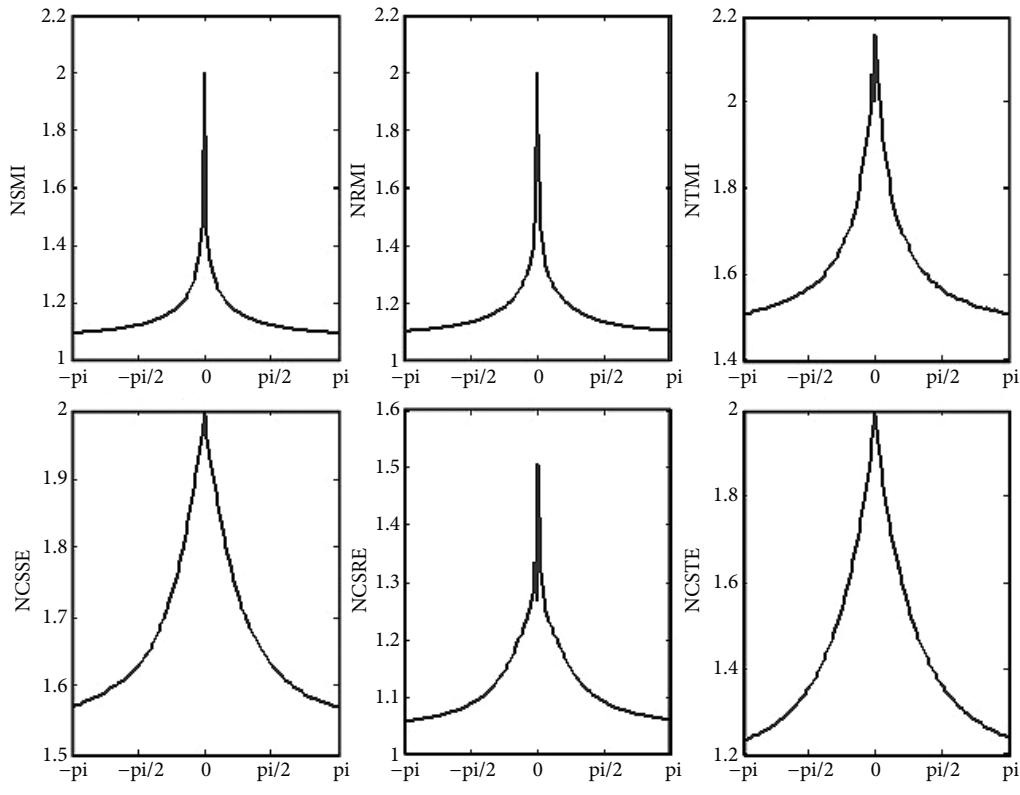
different pixel values or, equivalently, to different random variables.

The values of entropies of SE, RE, TE, SSE, SRE, and STE for Figures 2(a) and 2(b) are shown in Table 3, respectively, where  $\alpha = 0.8$ . We can see that density entropies provide the same values, whereas survival entropies provide different values for these two images. In other words, these two images contain the same information in the view of the density entropy. But, to the survival entropy, the amount of information there is different. It is demonstrated that the survival entropy is capable of distinguishing the two images, whereas the density entropy is not. The reason is that spatial information is taken into account in survival entropy formulas.

*Example 3.* We use a computerized tomography (CT) of brain slice as a testing image to evaluate the properties of similarity measures and normalized similarity measures. The density of CT ranges from 0 to 255 with the mean of 92.683 and size of  $256 \times 256$ . It is originally fixed and viewed as a template. A series of rotation transformations are applied to it so as to obtain floating images for comparison with the template. The range of rotation angle is set as  $[-\pi, \pi]$ . We calculate the values of SMI, RMI, TMI, CSSE, CSRE, and CSTE between the template and floating images, as shown in Figure 3(a). Accordingly, as shown in Figure 3(b), the values of the normalized similarity measures of NSMI, NRMI, NTMI, NCSSE, NCSRE, and NCSTE acquired from the template and floating images in this experiment are also



(a)



(b)

FIGURE 3: The values of similarity measures (a) and normalized similarity measures (b) for image rotation transformations with angle range  $[-\pi, \pi]$ .

calculated with the same parameter settings,  $\alpha = 0.8$ , as Figure 3(a).

Figure 3 demonstrates that when the rotation angle is zero or, equivalently, the template and the floating image are perfectly aligned, all the values of the considered similarity metrics reach their maximums. In other words, they all can be adopted as similarity metrics in image processing.

Generally, the similarity measures defined on survival function are more regular than those defined on density function, since the density is computed as the derivative of the survival function [19]. Therefore, the similarity measures defined on survival function are more robust in matching problems.

A kernel problem arising from matching problem is the choice of similarity metrics. In Table 2, sixteen different types of similarity measures and normalized similarity measures are provided with unifying presentations by (64) and (65). Their symmetric versions are given by (66) and (67) simultaneously. The proposed similarity measures are capable of developing the means of matching in signal processing to some extent.

## 6. Conclusions

This research is conducted in two dimensions. On one hand, we extend KN averages to  $q$ -generalized KN averages to redefine the existing classical entropies and to define the survival Rényi entropy and survival Tsallis entropy on the survival function. On the other hand, for each type of entropies discussed, a similarity measure and a normalized similarity measure are proposed accordingly. Some properties of the information measures are researched.

We make an inventory of sixteen types of entropies, similarity measures, and normalized similarity measures which exhibit diversity and unity. This leads to the proposition of the unifying frameworks for information measures. Therefore, our work addresses a broad spectrum of information measures as a whole through the unifying frameworks. Undoubtedly, some of them will be conceptual if they are not adapted for applications. More applications of the proposed information measures will dominate our research in the near future.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research is supported by 973 Program (2013CB329404), the Fundamental Research Funds for the Central Universities (ZYGX2013Z005), NSFC (61370147 and 61170311), and Sichuan Province Science and Technology Research Project (2012GZX0080).

## References

- [1] R. T. V. Hartley, "Transmission of information," *Bell System Technical Journal*, vol. 7, no. 3, pp. 535–563, 1928.
- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [3] M. Banerjee and N. R. Pal, "Feature selection with SVD entropy: some modification and extension," *Information Sciences*, vol. 264, pp. 118–134, 2014.
- [4] S. Kullback, *Information Theory and Statistics*, John Wiley & Sons, New York, NY, USA, 1959.
- [5] M. Ali, C. W. Ahn, M. Pant, and P. Siarry, "An image watermarking scheme in wavelet domain with optimized compensation of singular value decomposition via artificial bee colony," *Information Sciences*, vol. 301, pp. 44–60, 2015.
- [6] V. M. Ilić and M. S. Stanković, "Generalized Shannon-Khinchin axioms and uniqueness theorem for pseudo-additive entropies," *Physica A*, vol. 411, pp. 138–145, 2014.
- [7] S. Yu, T.-Z. Huang, X. Liu, and W. Chen, "Information measures based on fractional calculus," *Information Processing Letters*, vol. 112, no. 23, pp. 916–921, 2012.
- [8] J.-F. Bercher, "A simple probabilistic construction yielding generalized entropies and divergences, escort distributions and  $q$ -Gaussians," *Physica A*, vol. 391, pp. 4460–4469, 2012.
- [9] G. J. Klir, *Foundations of Generalized Information Theory*, John Wiley Sons, Inc, New Jersey, NJ, USA, 2006.
- [10] A. I. Khinchin, *Mathematical Foundation of Information Theory*, Mathematical Foundations of Information Theory, New York, NY, USA, 1956.
- [11] A. Ben-Tal, "On generalized means and generalized convex functions," *Journal of Optimization Theory and Applications*, vol. 21, no. 1, pp. 1–13, 1977.
- [12] A. N. Kolmogorov, "Three approaches to the quantitative definition of information," *International Journal of Computer Mathematics. Section A. Programming Theory and Methods. Section B. Computational Methods*, vol. 2, pp. 157–168, 1968.
- [13] C. Tsallis, "What are the numbers that experiments provide?" *Química Nova*, vol. 17, p. 468, 1994.
- [14] A. Rényi, "On measures of entropy and information," in *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 547–561, University of California, Berkeley, Calif, USA, 1961.
- [15] C. E. Shannon, "A mathematical theory of communication," *Bell Labs Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- [16] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.
- [17] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, vol. 52, no. 1-2, pp. 479–487, 1988.
- [18] L. L. Campbell, "Exponential entropy as a measure of extent of a distribution," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 5, no. 3, pp. 217–225, 1966.
- [19] M. Rao, Y. Chen, B. C. Vemuri, and F. Wang, "Cumulative residual entropy: a new measure of information," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 50, no. 6, pp. 1220–1228, 2004.
- [20] F. E. Wang, B. C. Vemuri, M. Rao, and Y. Chen, "A new and robust information theoretic measure and its application to image alignment," in *Proceedings of the International Conference of Information Processing in Medical Imaging*, pp. 388–400, Ambleside, UK, 2003.
- [21] M. Asadi and Y. Zohrevand, "On the dynamic cumulative residual entropy," *Journal of Statistical Planning and Inference*, vol. 137, no. 6, pp. 1931–1941, 2007.

- [22] A. Di Crescenzo and M. Longobardi, "On cumulative entropies," *Journal of Statistical Planning and Inference*, vol. 139, no. 12, pp. 4072–4087, 2009.
- [23] K. Zografos and S. Nadarajah, "Survival exponential entropies," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 51, no. 3, pp. 1239–1246, 2005.
- [24] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.
- [25] N. Dowson, T. Kadir, and R. Bowden, "Estimating the joint statistics of images using nonparametric windows with application to registration using mutual information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1841–1857, 2008.
- [26] Z. Zenadic, "Information discriminant analysis: feature extraction with an information-theoretic objective," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1394–1407, 2007.
- [27] S. M. H. Anvar, W.-Y. Yau, and E. K. Teoh, "Multiview face detection and registration requiring minimal manual intervention," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, pp. 2484–2497, 2013.
- [28] P. Jizba, "The world according to Rnyi: thermodynamics of multifractal systems," *Annals of Physics*, vol. 312, no. 1, pp. 17–59, 2004.
- [29] J. L. W. V. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Mathematica*, vol. 30, no. 1, pp. 175–193, 1906.
- [30] D. A. Hahn, V. Daum, and J. Hornegger, "Automatic parameter selection for multimodal image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 5, pp. 1140–1155, 2010.
- [31] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 20, no. 2, pp. 189–201, 2009.
- [32] Z. Zhang and K. Zhao, "Low-rank matrix approximation with manifold regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1717–1729, 2013.
- [33] S. Abramovich, B. Mond, and J. E. Pecaric, "Sharpening Hölders inequality," *Journal of Mathematical Analysis and Applications*, vol. 196, pp. 1131–1134, 1995.
- [34] W. Pawłucki and W. Pleśniak, "Markov's inequality and  $C^\infty$  functions on sets with polynomial cusps," *Mathematische Annalen*, vol. 275, no. 3, pp. 467–480, 1986.





**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

