IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 August 2022, accepted 1 September 2022, date of publication 8 September 2022, date of current version 16 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3205328

RESEARCH ARTICLE

Entropies and Their Concavity and Schur-Concavity Conditions

TARALD O. KVÅLSETH, (Life Member, IEEE)
Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN 55455, USA
Department of Industrial and Systems Engineering, University of Minnesota, Minneapolis, MN 55455, USA
e-mail: kvals001@umn.edu

ABSTRACT Concavity and Schur-concavity are two of the important properties of any entropy. Since Shannon's classical entropy formulation, a number of generalized entropies have been proposed as parameterized generalizations of Shannon's entropy. For such generalized entropies, the conditions under which they are concave and/or Schur-concave have not always been determined or have been incompletely and incorrectly reported in a variety of publications. This paper provides proofs of those two properties for the various proposed generalized entropies using a unifying approach. First, a new three-parameter entropy is introduced of which other proposed generalized entropies are particular members. Second, a proof is derived for the concavity and Schur-concavity of the new entropy and the underlying conditions. Those results are then applied to the particular one-parameter and two-parameter members. Some new such members are also discussed as are some related inequalities. The various derivations are based on so-called generalized probability distributions when the sum of component probabilities may be less than 1.

INDEX TERMS Entropy, concavity, Schur-concavity, generalized entropies, Gini's means.

I. INTRODUCTION

Since Shannon (1948) introduced entropy as part of his theory of information and communication, entropy as a measure of a variety of attributes has had a profound influence on research in a diverse range of scientific areas. Its popularity is partly attributable to the various desirable properties of Shannon's entropy. Two of those important properties are concavity and Schur-concavity as is the focus of the present paper.

The concavity property is essential for entropy maximization problems (e.g., [1]) and Schur-concavity ensures that, among all probability distributions, no distribution can have a higher entropy value than the uniform distribution [2, pp. 101]. Schur-concavity also reflects the important property that the entropy value increases as the components of a probability distribution become "more nearly equal" or "less spread out" as formalized by *majorization theory* [2, Ch. 1 and 3].

The influence of Shannon's entropy is also reflected by its various generalizations that have been proposed over the years [3], [4], [5], [6], [7]. Such generalizations typically involve parameterized families of entropies such that Shannon's entropy is a member for a particular parameter value. The most popular of these are the one-parameter entropies by Rényi [8] and Tsallis [9]. While most of the proposed entropies have the Schur-concavity property, some lack the concavity property because of its more restrictive parameter constraints.

The concavity and Schur-concavity properties have often been overlooked when new entropies have been proposed. Subsequent discussions of these properties have not always been correct or complete, some have been incorrectly reported, and their proofs are simply lacking for some of the entropies. The presentations of these properties for some of the entropies have also been scattered between a wide range of publications and sources. It is the purpose of the present paper to carefully and systematically derive the conditions or constraints under which various entropies have the concavity and/or Schur-concavity properties.

The systematic approach of this paper is based on the introduction of a most general three-parameter entropy or family of entropies of which other proposed entropies are

The associate editor coordinating the review of this manuscript and approving it for publication was Md Asaduzzaman.

particular members. This new entropy is simply defined as a power function of the well-known mean by Gini [10] applied to the components of a probability distribution. The parameter space or conditions for the concavity and Schur-concavity of this new entropy are then derived and applied to other established member entropies. Some novel members of the new entropy are identified as being of potential interest. Some interesting inequalities between entropies can also be derived from the three-parameter entropy. All derivations are based on the general case when a probability distribution may possibly be incomplete in the sense that the probability components may not necessarily sum to 1 (e.g., [8]).

II. DEFINITIONS

Let $P_n = (p_1, \ldots, p_n)$ be some probability distribution with each $p_i \geq 0$ and $\sum_{i=1}^n p_i \leq 1$ in case P_n is possibly an incomplete distribution. An incomplete distribution with $\sum_{i=1}^n p_i < 1$ may occur if some events are not observable or are ignored or if some data are not available (e.g., [8], [11, Ch. IX], and [12]). For some generic entropy H taking on the value $H(P_n)$ for the distribution P_n, *Schur-concavity* and *majorization* may be defined as follows [2]. If the p_i's are ordered such that $p_{[1]} \geq p_{[2]} \geq \ldots \geq p_{[n]}$, with $q_{[1]} \geq q_{[2]} \geq \ldots \geq q_{[n]}$ for another distribution $Q_n = (q_1, \ldots, q_n)$, P_n is *majorized by* Q_n, denoted by $P_n \prec Q_n$, if

$$\sum_{i=1}^{j} p_{[i]} \leq \sum_{i=1}^{j} q_{[i]}, j = 1, \ldots, n-1 \text{ and } \sum_{i=1}^{n} p_{[i]} = \sum_{i=1}^{n} q_{[i]}. \tag{1}$$

Then, an entropy H is *Schur-concave* if

$$P_n \prec Q_n \text{ implies } H(P_n) \geq H(Q_n). \tag{2}$$

If the inequality in (2) is strict and P_n is not simply a permutation of Q_n, then H is *strictly Schur-concave*.

The Schur-concavity of H ensures that its value $H(P_n)$ increases as the components of P_n become increasingly equal or uniform, attaining its maximum value $H(1/n, \ldots 1/n)$ for the uniform discrete distribution $(1/n, \ldots, 1/n)$ or for

$$H(\sum_{i=1}^{n} p_i/n, \ldots, \sum_{i=1}^{n} p_i/n)$$

when $\sum_{i=1}^n p_i < 1$ since this uniform distribution is majorized by any $P_n = (p_1, \ldots, p_n)$ as can be seen from (1). Clearly, Schur-concavity is a necessary property of any H.

Schur-concavity is a somewhat milder condition than the usual (in the sense of Jensen) concavity. If H is (permutation) symmetric in its arguments and if H is concave, then it is implied that H is also Schur-concave. However, the converse implication does not necessarily hold.

In order to determine if some H is concave and under what necessary conditions, it will be convenient to be able to refer to the following set of composite functions $h(P_n) = g[f(P_n)]$:

If f is concave and g is concave and nondecreasing, then h is concave. (3a)

If f is convex and g is concave and nonincreasing, then h is concave. (3b)

If f is concave and g is convex and nonincreasing, then h is convex. (3c)

If f is convex and g is convex and nondecreasing, then h is convex. (3d)

Some of these relationships follow from the fact that f is concave if and only if $-f$ is convex. Similarly, f is Schur-concave if and only if $-f$ is Schur-convex (when the inequality in (2) is reversed) [2, Ch. 3].

III. GENERALIZED ENTROPIES

A. ONE-PARAMETER ENTROPIES

Among the entropies that involve one parameter, those of Rényi [8] and Tsallis [9] are the best known and most frequently referenced ones. Rényi's entropy is defined as:

$$H_{R\alpha}(P_n) = \frac{1}{1-\alpha} \log \sum_{i=1}^{n} p_i^\alpha, \ \alpha > 0 \tag{4a}$$

where only positive values of the arbitrary parameter α is considered in order for $H_{R\alpha}(P_n)$ to be defined for all $p_i \geq 0$. For the more general case when P_n is a possibly incomplete distribution with $\sum_{i=1}^n p_i \leq 1$, Rényi proposed the entropy

$$H_{R\alpha}(P_n) = \frac{1}{1-\alpha} \log \Big(\sum_{i=1}^{n} p_i^\alpha / \sum_{i=1}^{n} p_i \Big), \ \alpha > 0. \tag{4b}$$

Tsallis' entropy is defined as

$$H_{T\alpha}(P_n) = \frac{1}{1-\alpha} \Big(\sum p_i^\alpha - 1 \Big), \ \alpha \in \mathbb{R}. \tag{5}$$

The entropy of Shannon [13] is the limiting case of both $H_{R\alpha}(P_n)$ and $H_{T\alpha}(P_n)$, i.e.,

$$\lim_{\alpha \to 1} H_{R\alpha}(P_n) = \lim_{\alpha \to 1} H_{T\alpha}(P_n) = -\sum_{i=1}^{n} p_i \log p_i = H_S(P_n) \tag{6}$$

where the natural (base-e) logarithm is used throughout this paper. Havrda and Charvat [14] introduced an entropy that, instead of the term $1/(1-\alpha)$ in (5), used the term $1/(1-2^{\alpha-1})$.

Arimoto [15] proposed the following entropy:

$$H_{A\alpha}(P_n) = \frac{1}{\alpha - 1} \Big[\Big(\sum_{i=1}^{n} p_i^{\frac{1}{\alpha}} \Big)^\alpha - 1 \Big], \ \alpha > 0. \tag{7}$$

Another form of this entropy, the *R-norm* entropy with $R = 1/\alpha$ in (7), has been studied by [16].

Aczél and Daróczy [17] and Kapur [18] independently introduced the entropy

$$H_{ADK\alpha}(P_n) = -\sum_{i=1}^{n} p_i^\alpha \log p_i \Big/ \sum_{i=1}^{n} p_i^\alpha, \ \alpha > 0. \tag{8}$$

See also [12, pp.192]. However, as pointed out below, this entropy does have an important restriction.

Landsberg and Vedral [19] proposed the following entropy:

$$H_{LV\alpha}(P_n) = \frac{1}{\alpha - 1}\left(\frac{1}{\sum_{i=1}^{n} p_i^{\alpha}} - 1\right), \ \alpha \in \mathbb{R}. \tag{9}$$

This entropy as well as that in (5) have been defined for all real values of the arbitrary parameter α, but require $\alpha > 0$ when $p_i \geq 0$ $(i = 1, \ldots, n)$.

B. TWO-PARAMETER ENTROPIES

Aczél and Daróczy [17] and Kapur [18] independently introduced the following two-parameter entropy:

$$H_{ADK\alpha\beta}(P_n) = \frac{1}{\beta - \alpha}\log\left(\frac{\sum_{i=1}^{n} p_i^{\alpha}}{\sum_{i=1}^{n} p_i^{\beta}}\right), \ \alpha, \beta > 0 \tag{10}$$

where again, as throughout this paper, α and β are arbitrary parameters subject to the given constraints.

Another entropy that has received considerable attention is that of Sharma and Mittal [20] and defined as

$$H_{SM\alpha\beta}(P_n) = \frac{1}{2^{(1-\alpha)} - 1}\left[\left(\sum_{i=1}^{n} p_i^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right], \ \alpha, \beta > 0. \tag{11}$$

By using a slightly different parameter expression, Rathie and Taneja [21] proposed a *unified* entropy as

$$H_{RT\alpha\beta}(P_n) = \frac{1}{(1-\alpha)\beta}\left[\left(\sum_{i=1}^{n} p_i^{\alpha}\right)^{\beta} - 1\right], \ \alpha > 0, \beta \in \mathbb{R}. \tag{12}$$

More recently, Hu and Ye [6] also considered this last entropy.

Other two-parameter entropies have been proposed as extensions of one-parameter entropies by simply replacing a parameter with one involving two parameters. However, such extensions do not change the basic form of the entropy function. Examples of such parametric reformulations include Nath [22] who basically proposed to replace α in (4) with α/β or replace the α in the exponential term in (4) with α^{β}. Rathie [23] suggested replacing p_i^{α} in (4) with $p_i^{\alpha_i}$ for $i = 1, \ldots, n$. More recently, Hooda and Ram [24] proposed an entropy that corresponds to Arimoto's entropy when substituting the two-parameter term $(2 - \beta)/R$ for α in (7). Also, substituting the three-parameter term $(2\alpha - \beta)/R$ for α in (7) produces an entropy proposed by Hooda and Sharma [25].

IV. CONDITIONS FOR CONCAVITY AND SCHUR-CONCAVITY

A. GENERAL APPROACH

In order to consider a unified and systematic approach to determining the conditions under which the various parameterized entropies possess the properties of concavity and Schur-concavity, one could start off by considering an entropy with additional generality. Thus, since various proposed entropies are basically decreasing functions of some mean of the probabilities $p_1, \ldots, p_n$, it would seem reasonable to consider the most general (two-parameter) Gini means defined by

$$G_{\alpha\beta}(P_n) = \left(\frac{\sum_{i=1}^{n} p_i^{\alpha}}{\sum_{i=1}^{n} p_i^{\beta}}\right)^{\frac{1}{\alpha-\beta}}, \ \alpha, \beta \in \mathbb{R} \tag{13}$$

after Gini [10] and as discussed by Bullen [26, pp. 248-251]. As a general decreasing power function of $G_{\alpha\beta}(P_n)$, one could define the following three-parameter entropy:

$$\begin{aligned} &H_{K\alpha\beta\gamma}(P_n) \\ &\quad = \frac{1}{1-\gamma}(G_{\alpha\beta}^{\gamma-1} - 1) \\ &\quad = \frac{1}{1-\gamma}\left[\left(\frac{\sum_{i=1}^{n} p_i^{\alpha}}{\sum_{i=1}^{n} p_i^{\beta}}\right)^{\frac{\gamma-1}{\alpha-\beta}} - 1\right], \ \alpha, \beta \geq 0; \gamma \in \mathbb{R}. \end{aligned} \tag{14}$$

The parameters α and β in (14) are restricted to having nonnegative values in order for this entropy to be defined for $p_i \geq 0$ and $i = 1, \ldots, n$. For generality sake, this entropy is also defined for possibly incomplete distributions when $\sum_{i=1}^{n} p_i \leq 1$. Distributions with $\sum_{i=1}^{n} p_i < 1$ may occur when some events and the associated p_i's are missing, ignored, or simply unobservable. Such incomplete cases include situations when the smallest p_i's are not available individually, but are grouped into an "all others" category. Distributions with $\sum_{i=1}^{n} p_i \leq 1$ are also referred to as *generalized distributions* (e.g., [8], [11], [12], and [27, Ch. 5]). In the extreme case of a single event with $P_1 = (p)$, (14) becomes $(p^{\gamma-1} - 1)/(1-\gamma)$ as the entropy of a single event.

The entropy in (14), or family of entropies, has the following property:

Theorem 1: The $H_{K\alpha\beta\gamma}(P_n)$ in (14) is concave if

$$0 \leq \min\{\alpha, \beta\} \leq 1 \leq \max\{\alpha, \beta\} \tag{15}$$

and $\gamma \geq 2$ and Schur-concave for any $\gamma \in \mathbb{R}$.

Proof: The following result follows from Dresher's inequality ([28]; see also [26, pp. 249]) for $G_{\alpha\beta}$ in (13), for any distributions P_n and Q_n (possibly incomplete), and for any $\lambda \in [0, 1]$:

$$\begin{aligned} &G_{\alpha\beta}[\lambda P_n + (1-\lambda)Q_n] \\ &\quad \leq \lambda G_{\alpha,\beta}(P_n) + (1-\lambda)G_{\alpha\beta}(Q_n), \ 0 \leq \lambda \leq 1 \end{aligned} \tag{16}$$

under the parameter constraint in (15). That is, $G_{\alpha\beta}$ subject to (15) is a convex function of P_n, with $\sum_{i=1}^{n} p_i \leq 1$. Then, since $H_{K\alpha\beta\gamma}(P_n)$ in (14) is clearly a concave and nonincreasing function of $G_{\alpha\beta}(P_n)$, it follows from (3b) that $H_{K\alpha\beta\gamma}(P_n)$ is a concave function of P_n restricted to (15) and $\gamma \geq 2$. Furthermore, since $G_{\alpha\beta}(P_n)$ subject to (15) is symmetric in the p_i's, it is also Schur-concave, which makes $H_{K\alpha\beta\gamma}(P_n)$ Schur-concave since it is decreasing in $G_{\alpha\beta}(P_n)$ for all real values of γ. This completes the proof.

The $H_{K\alpha\beta\gamma}(P_n)$ in (14) contains all the entropies in (4)-(12) as particular members. A two-parameter member of $H_{K\alpha\beta\gamma}(P_n)$ that also incorporates some of the other entropies is obtained for all $\gamma = \alpha - \beta + 1$ as follows:

$$H_{K\alpha\beta}(P_n) = \frac{1}{\beta - \alpha}\Big(\frac{\sum_{i=1}^n p_i^\alpha}{\sum_{i=1}^n p_i^\beta} - 1\Big), \quad \sum_{i=1}^n p_i \le 1,\ 0 < \beta \le 1,\ \alpha \ge 1 + \beta. \tag{17}$$

This entropy has the following property:

Theorem 2: The $H_{K\alpha\beta}(P_n)$ in (17) is concave if $0 < \beta \le 1$ and $\alpha \ge 1 + \beta$ and Schur-concave if $0 < \beta \le 1$ and $\alpha \ge 1$.

Proof: With $0 < \beta \le 1$, the concavity of $H_{K\alpha\beta}(P_n)$ follows from Theorem 1 and the requirement that $\gamma = \alpha - \beta + 1 \ge 2$. The Schur-concavity of $H_{K\alpha\beta}(P_n)$ is implied by $R_{\alpha\beta}(P_n) = \sum_{i=1}^n p_i^\alpha / \sum_{i=1}^n p_i^\beta$ being Schur-convex for $0 < \beta \le 1$ and $\alpha \ge 1$ [2, Ch. 3] and $H_{K\alpha\beta}(P_n)$ being a decreasing function of $R_{\alpha\beta}(P_n)$ for $\alpha > \beta$.

Comment: While $H_{K\alpha\beta\gamma}(P_n)$ in (14) is symmetric in the parameters α and β, this symmetry does not apply to $H_{K\alpha\beta}(P_n)$.

For the case when $\beta = 1$ and $\sum_{i=1}^n p_i = 1$, (14) reduces to

$$H_{K\alpha 1\gamma}(P_n) = \frac{1}{1-\gamma}\Big[\Big(\sum_{i=1}^n p_i^\alpha\Big)^{\frac{\gamma-1}{\alpha-1}}\Big],\ \alpha \ge 0, \gamma \in \mathbb{R} \tag{18}$$

which is seen to be equivalent to the entropy by Sharma and Mittal [20], although they used the denominator $2^{1-\gamma} - 1$ (instead of $1 - \gamma$). This entropy has the following property:

Theorem 3: The $H_{K\alpha 1\gamma}(P_n)$ in (18) is concave if $\alpha > 0$ and $\gamma \ge 2 - 1/\alpha$ and Schur-concave if $\alpha > 0$ and for all $\gamma \in \mathbb{R}$.

Proof: It follows immediately from Minkowski's inequality (e.g., [26, pp. 189]) that, for the power sum $S_\alpha(P_n) = \Big(\sum_{i=1}^n p_i^\alpha\Big)^{1/\alpha}$ and for any $\lambda \in [0, 1]$,

$$S_\alpha[\lambda P_n + (1-\lambda)Q_n] \le \lambda S_\alpha(P_n) + (1-\lambda)S_\alpha(Q_n) \tag{19}$$

if $\alpha > 1$ and for any two (possibly incomplete) distributions $P_n = (p_1, \ldots, p_n)$ and $Q_n = (q_1, \ldots, q_n)$, with the inequality in (19) reversed for $\alpha < 1$. That is, the function S_α is convex for $\alpha > 1$ and concave for $\alpha < 1$. By expressing (18) as

$$H_{K\alpha 1\gamma}(P_n) = \frac{1}{1-\gamma}[(S_\alpha(P_n))^{\frac{\alpha(\gamma-1)}{\alpha-1}} - 1] \tag{20}$$

it is seen from the derivatives of $H_{K\alpha 1\gamma}(P_n)$ with respect to $S_\alpha(P_n)$ in (20) that $H_{K\alpha 1\gamma}(P_n)$ is a nondecreasing and concave function of $S_\alpha(P_n)$ for $0 < \alpha \le 1$ and $\gamma \ge 2 - 1/\alpha$ and nonincreasing and concave for $\alpha \ge 1$ and $\gamma \ge 2 - 1/\alpha$. Since $S_\alpha(P_n)$ is concave for $\alpha < 1$ and $H_{K\alpha 1\gamma}(P_n)$ is a nondecreasing and concave function of $S_\alpha(P_n)$ for $0 < \alpha \le 1$ and $\gamma \ge 2 - 1/\alpha$, it follows from (3a) that $H_{K\alpha 1\gamma}(P_n)$ is concave under these parameter restrictions. Similarly, since $S_\alpha(P_n)$ is convex for $\alpha > 1$ and $H_{K\alpha 1\gamma}(P_n)$ is a nonincreasing and concave function of $S_\alpha(P_n)$ for $\alpha \ge 1$ and $\gamma \ge 2 - 1/\alpha$, it follows from (3b) that $H_{K\alpha 1\gamma}(P_n)$ is concave if $\alpha \ge 1$ and $\gamma \ge 2 - 1/\alpha$. That is, $H_{K\alpha 1\gamma}(P_n)$ a concave function of P_n if $\alpha > 0$ and $\gamma \ge 2 - 1/\alpha$. The Schur-concavity in Theorem 3 follows from majorization theory [2, Ch. 3] as an immediate consequence of the fact that (a) $T_\alpha(P_n) = \sum_{i=1}^n p_i^\alpha$ is Schur-concave for $0 < \alpha \le 1$ and Schur-convex for $\alpha \ge 1$ and (b) $H_{K\alpha 1\gamma}(P_n)$ is an increasing function of $T_\alpha(P_n)$ for $0 < \alpha \le 1$ and decreasing for $\alpha \ge 1$ and for all $\gamma \in \mathbb{R}$. This completes the proof.

Remark: The parameter restrictions in Theorem 3 are seen to be an improvement over those obtained by simply setting $\beta = 1$ in Theorem 1.

B. PARTICULAR CASES

Theorem 4: $H_{R\alpha}(P_n)$ in (4a) is concave if $0 < \alpha \le 1$ and Schur-concave if $\alpha > 0$.

Proof: This property of $H_{R\alpha}(P_n)$ follows from Theorem 3 with $\gamma = 1$ since $H_{R\alpha}(P_n) = \lim_{\gamma\to 1} H_{K\alpha 1\gamma}(P_n)$.

Theorem 5: $H_{R\alpha}(P_n)$ in (4b) is not concave for any α, but is Schur-concave if $\alpha \ge 1$.

Proof: Since the entropy in (4b) is a particular member of $H_{K\alpha\beta\gamma}(P_n)$ in (14) with $\beta = 1$, $\sum_{i=1}^n p_i \le 1$, and $\gamma \to 1$, it follows from Theorem 1 that $H_{R\alpha}(P_n)$ in (4b) cannot be concave since it violates the requirement of $\gamma \ge 2$. However, it meets the Schur-concavity condition for $\alpha \ge 1$ since that of Theorem 1 applies to all real-valued γ. This completes the proof.

Theorem 6: $H_{T\alpha}$ in (5) is concave and Schur-concave if $\alpha > 0$.

Proof: This entropy is the particular member of $H_{K\alpha 1\gamma}(P_n)$ in (18) with $\gamma = \alpha$ so that the condition $\gamma \ge 2 - 1/\alpha$ reduces to $(\alpha - 1)^2 \ge 0$, which is obviously met by all real values of α. Thus, from Theorem 3, $H_{T\alpha}(P_n)$ is both concave and Schur-concave for $\alpha > 0$.

Theorem 7: $H_{A\alpha}(P_n)$ in (7) is concave and Schur-concave if $\alpha > 0$.

Proof: The entropy in (7) can be considered a particular case of $H_{K\alpha 1\gamma}(P_n)$ in (18) by simply substituting $1/\alpha$ for α and $2-\alpha$ for γ. Then, it follows from Theorem 3 that $H_{A\alpha}(P_n)$ is concave and Schur-concave for $\alpha > 0$ since, besides $1/\alpha > 0$, the parameter condition $\gamma \ge 2 - 1/\alpha$ after substitution becomes $(2 - \alpha) = (2 - \alpha)$, which holds for all $\alpha \ge 0$. This completes the proof.

Theorem 8: $H_{ADK\alpha}(P_n)$ in (8) is neither concave nor Schur-concave for all $\alpha > 0$.

Proof: The entropy in (8) is a limiting member of $H_{K\alpha\beta\gamma}(P_n)$ in (14) when $\beta \to \alpha$ and $\gamma \to 1$. However, Theorem 1 requires $\gamma \ge 2$ for concavity of $H_{K\alpha\beta\gamma}(P_n)$ with $\alpha, \beta > 0$ so that $H_{ADK}(P_n) = H_{K\alpha\alpha 1}(P_n)$ cannot be concave or Schur-concave for all $\alpha > 0$, completing the proof.

Comment: Stolarsky [29] and Clausing [30] explored the problem of determining the minimum value α_0 of the parameter α such that

$$H_{ADK\alpha}(P_n) \le H_{ADK\alpha}(1/n, \ldots, 1/n) = \log n \quad \text{for all } P_n \text{ and } \alpha \ge \alpha_0$$

which is a Schur-concavity requirement. Those authors showed that α_0 depends on n.

Theorem 9: $H_{LV\alpha}(P_n)$ in (9) is concave if $0 < \alpha \le 1$ and Schur-concave if $\alpha > 0$.

Proof: This entropy is a particular member of $H_{K\alpha 1\gamma}(P_n)$ in (18) with $\gamma = 2-\alpha$. It then follows from Theorem 3 that $H_{LV\alpha}(P_n)$ is Schur-concave for $\alpha > 0$ and concave if $\alpha > 0$ and $(2-\alpha) \geq 2 - 1/\alpha$, i.e., $\alpha^2 \leq 1$ or $0 < \alpha \leq 1$, which completes the proof.

Theorem 10: $H_{ADK\alpha\beta}(P_n)$ in (10) is not concave for α, $\beta > 0$, but it is Schur-concave subject to the parameter constraint in (15).

Proof: Since this entropy is a limiting case of $H_{K\alpha\beta\gamma}(P_n)$ in (14) as $\gamma \to 1$ and since the concavity of Theorem 1 requires $\gamma \geq 2$, $H_{ADK\alpha\beta}(P_n)$ is not concave. However, it is Schur-concave as a consequence of Theorem 1, which completes the proof.

Theorem 11: $H_{SM\alpha\beta}(P_n)$ in (11) is concave if $\alpha > 0$ and $\beta \geq 2 - 1/\alpha$ and Schur-concave if $\alpha > 0$ and for all $\beta \in \mathbb{R}$.

Proof: With $\gamma = \beta$ in (18), $H_{SM\alpha\beta}(P_n) = CH_{K\alpha 1\beta}(P_n)$ with $C = (1-\beta)/(2^{1-\beta}-1)$, so that from Theorem 3 it follows that $H_{SM\alpha\beta}(P_n)$ has the same concavity condition as $H_{K\alpha 1\beta}(P_n)$ since concavity is invariant under multiplication with a constant $C \geq 0$.

The Schur-concavity of $H_{SM\alpha\beta}(P_n)$ is implied by Theorem 3 and the fact that $H_{SM\alpha\beta}(P_n)$ is a strictly increasing function of the Schur-concave $H_{K\alpha 1\gamma}(P_n)$ with $\gamma = \beta$ (with $H_{SM\alpha\beta}(P_n) = [(1-\beta)/(2^{1-\beta}-1)]H_{K\alpha 1\beta}(P_n)$).

Theorem 12: $H_{RT\alpha\beta}(P_n)$ in (12) is concave if $\alpha > 0$ and $\alpha(\alpha-1)\beta \geq \alpha - 1$ and Schur-concave if $\alpha > 0$ and for all $\beta \in \mathbb{R}$.

Proof: This property is an immediate consequence of the fact that $H_{RT\alpha\beta}(P_n)$ is equivalent to $H_{K\alpha 1\gamma}(P_n)$ in (18) with $\gamma = (\alpha-1)\beta+1$. The restriction $\gamma \geq 2-1/\alpha$ for $H_{K\alpha 1\gamma}(P_n)$ becomes $\alpha(\alpha-1)\beta \geq \alpha - 1$ for $H_{RT\alpha\beta}(P_n)$.

Remark: Rathie and Taneja [21] proved the concavity of $H_{RT\alpha\beta}(P_n)$ under the following constraints: $0 < \alpha \leq 1$, $\alpha\beta \leq 1$ or $\alpha \geq 1$, $\alpha\beta \geq 1$. However, these two constraints are equivalent to $\alpha(\alpha-1)\beta \geq \alpha - 1$ for all $\alpha > 0$. Hu and Ye [6] presented the same proof as that of Rathie and Taneja [21]. See also Kapur [3, pp. 122-124].

V. EXPLORATIONS OF $H_{K\alpha\beta}$ IN (17)

Besides the two-parameter unifying property of $H_{K\alpha\beta}(P_n)$ in (17), this entropy is also defined for distributions $P_n = (p_1, \ldots, p_n)$ that may possibly be incomplete, i.e., $\sum_{i=1}^n p_i \leq 1$. In the extreme case of a single event $P_1 = (p)$,

$$H_{K\alpha\beta}(p) = \frac{1}{\beta-\alpha}(p^{\alpha-\beta}-1) = \begin{cases} 1-p & \text{if } \alpha = 2, \beta = 1 \\ -\log p & \text{if } \beta \to \alpha. \end{cases} \tag{21}$$

The $H_{K\alpha\beta}(P_n)$ becomes the following weighted mean of (21):

$$H_{K\alpha\beta}(P_n) = \sum_{i=1}^n \Big(\frac{p_i^\beta}{\sum_{i=1}^n p_i^\beta}\Big) H_{K\alpha\beta}(p_i) = \frac{1}{\beta-\alpha}\Big(\frac{\sum_{i=1}^n p_i^\alpha}{\sum_{i=1}^n p_i^\beta} - 1\Big). \tag{22}$$

Some potentially interesting one-parameter members of $H_{K\alpha\beta}(P_n)$ can be outlined as follows:

$$H_{K\alpha 1}(P_n) = (1-\alpha)^{-1}(\sum_{i=1}^n p_i^\alpha / \sum_{i=1}^n p_i - 1), \quad \sum_{i=1}^n p_i \leq 1,\ \alpha > 1\ (\alpha > 2); \tag{23}$$

$$H_{K\alpha 1}(P_n) = (1-\alpha)^{-1}(\sum_{i=1}^n p_i^\alpha - 1), \quad \sum_{i=1}^n p_i = 1,\ \alpha > 0\ (\alpha > 0); \tag{24}$$

$$H_{K1\beta}(P_n) = (\beta-1)^{-1}(\sum_{i=1}^n p_i / \sum_{i=1}^n p_i^\beta - 1), \quad \sum_{i=1}^n p_i \leq 1,\ \beta > 0\ (0 < \beta \leq 1); \tag{25}$$

$$H_{K1\beta}(P_n) = (\beta-1)^{-1}(1/\sum_{i=1}^n p_i^\beta - 1), \quad \sum_{i=1}^n p_i = 1,\ \beta > 0\ (0 < \beta \leq 1); \tag{26}$$

$$H_{K\alpha\alpha}(P_n) = -\sum_{i=1}^n p_i^\alpha \log p_i / \sum_{i=1}^n p_i^\alpha, \quad \sum_{i=1}^n p_i \leq 1,\ \alpha > \alpha_0(n); \tag{27}$$

$$H_{K\alpha,\alpha-1}(P_n) = 1 - \sum_{i=1}^n p_i^\alpha / \sum_{i=1}^n p_i^{\alpha-1}, \quad \sum_{i=1}^n p_i \leq 1,\ 1 < \alpha \leq 2\ (1 < \alpha \leq 2). \tag{28}$$

The parameter restrictions in (23)-(28) are those for which the entropies are Schur-concave (and concave in parentheses). Those parameter restrictions follow from Theorem 2 in the case of (23) and (28), from Theorem 6 for (24), from Theorem 9 for (26), from Theorem 8 for (27), and from Theorem 2 for the concavity condition of (25). The Schur-concavity condition for (25) results from $\sum_{i=1}^n p_i / \sum_{i=1}^n p_i^\beta$ being Schur-convex and $(\beta-1)^{-1} < 0$ for $0 < \beta < 1$ and $\sum_{i=1}^n p_i / \sum_{i=1}^n p_i^\beta$ being Schur-concave and $(\beta-1)^{-1} > 0$ for $\beta > 1$.

It is apparent from the majorization definition in (1) that, for any possibly incomplete or generalized distribution P_n,

$$P_n^1 = (\sum_{i=1}^n p_i/n, \ldots \sum_{i=1}^n p_i/n) \prec P_n \prec (\sum_{i=1}^n p_i, 0, \ldots, 0) = P_n^0,\ \sum_{i=1}^n p_i \leq 1. \tag{29}$$

Consequently, from the Schur-concavity of $H_{K\alpha\beta}(P_n)$ together with (29),

$$H_{K\alpha\beta}(P_n^0) = \frac{1}{\beta-\alpha}\Big[\Big(\sum_{i=1}^n p_i\Big)^{\alpha-\beta} - 1\Big]$$

$$\begin{aligned} &\leq H_{K\alpha\beta}(P_n) \\ &\leq H_{K\alpha\beta}(P_n^1) \\ &= \frac{1}{\beta-\alpha}\Big[\Big(\frac{\sum_{i=1}^n p_i}{n}\Big)^{\alpha-\beta} - 1\Big]. \end{aligned} \quad (30)$$

The corresponding upper and lower bounds on members of $H_{K\alpha\beta}(P_n)$ follow from (30) and the Schur-concavity parameter values in (23)-(28).

The $H_{K\alpha\beta}(P_n)$ in (17) is basically an immediate generalization of Tsallis' entropy $H_{T\alpha}(P_n)$ in (5), i.e., $H_{K\alpha 1}(P_n)$ in (24) and of the Landsberg-Vedral entropy $H_{LV\alpha}$ in (9), i.e., $H_{K1\beta}(P_n)$ in (26). In fact, $H_{K\alpha\beta}(P_n)$ can be expressed as the following combination of the two:

$$\begin{aligned} &H_{K\alpha\beta}(P_n) \\ &= \Big(\frac{1}{\beta-\alpha}\Big) \\ &\cdot\Big[\Big\{(1-\alpha)H_{T\alpha}(P_n)+1\Big\}\Big\{(\beta-1)H_{LV\beta}(P_n)+1\Big\}-1\Big]. \end{aligned} \quad (31)$$

VI. EXPLORATIONS OF $H_{K\alpha\beta\gamma}(P_n)$ IN (14) FOR $\beta = 1$

For $\beta = 1$ and setting $\gamma = (\alpha-1)\delta + 1$, the three-parameter entropy in (14) reduces to

$$H_{K\alpha 1\delta}(P_n) = \frac{1}{(1-\alpha)\delta}\Big[\Big(\frac{\sum_{i=1}^n p_i^\alpha}{\sum_{i=1}^n p_i}\Big)^\delta - 1\Big], \quad \sum_{i=1}^n p_i \leq 1. \quad (32)$$

This entropy represents an immediate extension of that in (12) to include distributions with $\sum_{i=1}^n p_i \leq 1$. For $\delta = 1$, it reduces to the entropy in (23), which is an extension of Tsallis' entropy in (5). It has the following property:

Theorem 13: The entropy in (32) is concave if $\alpha \geq 0$ and $(\alpha-1)\delta \geq 1$ and Schur-concave if $\alpha \geq 0$ and $\delta \in \mathbb{R}$.

Proof: This theorem is an immediate consequence of Theorem 1 for $\beta = 1$ and the condition that $\gamma = (\alpha-1)\delta + 1 \geq 2$.

The entropy in (32) includes some potentially interesting members with $\alpha = 2$, including the limiting case when $\delta \to 0$, i.e.,

$$H_{K210}(P_n) = -\log\sum_{i=1}^n p_i^2 \Big/ \sum_{i=1}^n p_i. \quad (33)$$

However, from Theorem 13, this entropy is not concave since it does not meet the parameter condition $(\alpha-1)\delta \geq 1$. It is, however, Schur-concave from Theorem 13.

For $\alpha = 2$ and $\delta = 1$, (32) reduces to

$$H_{K211}(P_n) = 1 - \sum_{i=1}^n p_i^2 \Big/ \sum_{i=1}^n p_i \quad (34)$$

which, from Theorem 13, is both concave and Schur-concave. Also, consider the case when $\alpha = 2$ and $\delta = -1$ when (32) becomes

$$H_{K21(-1)}(P_n) = \sum_{i=1}^n p_i \Big/ \sum_{i=1}^n p_i^2 - 1 \quad (35)$$

which, however, from Theorem 13, is not concave, but is Schur-concave. The entropies in (33)-(35), with the assumption that $\sum_{i=1}^n p_i = 1$, and their normalized forms are well-known measures of diversity and evenness used in biology (e.g., [31]). The entropy in (34) with $\sum_{i=1}^n p_i \leq 1$ is a simple extension of the so-called quadratic entropy (e.g., [32], [4, pp. 174-176]). The entropy in (34) with $\sum_{i=1}^n p_i = 1$ has also been called the *logical entropy* and studied extensively by Ellerman [33], [34].

Another particular case of (32) is the following entropy for $\alpha = 1/2$, $\beta = 1$, and $\delta = 2$:

$$H_{K,1/2,1,2}(P_n) = \Big(\frac{\sum_{i=1}^n \sqrt{p_i}}{\sum_{i=1}^n p_i}\Big)^2 - 1 \quad (36)$$

which, for $\sum_{i=1}^n p_i = 1$, Arimoto [15] discussed as a member of this entropy in (7). From Theorem 13, the entropy in (36) is not concave since the condition $(\alpha-1)\delta \geq 1$ is not met, but it is Schur-concave. However, when $\sum_{i=1}^n p_i = 1$, this entropy is concave from Theorem 7. A related entropy has also been advocated by [35].

As another particular member of $H_{K\alpha\beta\gamma}(P_n)$ in (14), consider the so-called min-entropy defined as

$$H_m(P_n) = -\log p_{\max}, \; p_{\max} = \max_i\{p_1, \ldots, p_n\}. \quad (37)$$

This entropy is typically identified as the limit of Rényi's entropy $H_{R\alpha}(P_n)$ in (4a) as $\alpha \to \infty$ (see, e.g., [36], [37], [38]). Since, for $\beta = 1$ (or for any $0 \leq \beta \leq 1$), $\lim_{\alpha\to\infty} G_{\alpha 1}(P_n) = p_{\max}$, it follows from (14) that

$$\lim_{\alpha\to\infty} H_{K\alpha 1\gamma}(P_n) = H_{m\gamma}(P_n) = \frac{1}{1-\gamma}(p_{\max}^{\gamma-1} - 1) \quad (38)$$

and defined for $\sum_{i=1}^n p_i \leq 1$. The entropy in (37) is the limiting case $H_{m1}(P_n)$.

It follows from Theorem 1 that, with $\beta = 1$ and $\alpha \to \infty$, $H_{m\gamma}(P_n)$ in (38) is concave for $\gamma \geq 2$ and Schur-concave for all $\gamma \in \mathbb{R}$. However, with $\gamma = 1$, the min-entropy in (37) is not concave. An interesting member of (38) is the one for $\gamma = 2$, i.e.,

$$H_{m2}(P_n) = 1 - p_{\max} \quad (39)$$

which is concave and Schur-concave and known as the *variation ratio* in statistics and used as a measure of qualitative variation (e.g., [39, pp. 68]). Also, since $H_{m\gamma}(P_n)$ is strictly decreasing in γ for any given P_n, the following inequality may be pointed out:

$$1 - p_{\max} \leq -\log p_{\max} \leq p_{\max}^{-1} - 1$$

where the only concave term is $1 - p_{\max}$.

Consider also the case of $\alpha = \beta = 1$ when $G_{\alpha\beta}(P_n)$ in (13) becomes the self-weighted geometric mean and the entropy in (14) for the generalized distribution with $\sum_{i=1}^n p_i \leq 1$ becomes

$$H_{K11\gamma}(P_n) = \frac{1}{1-\gamma}\Big[\Big(\prod_{i=1}^n p_i^{p_i}\Big)^{\frac{\gamma-1}{\sum_{i=1}^n p_i}} - 1\Big], \; \gamma \in \mathbb{R}. \quad (40)$$

From Theorem 1, this entropy is concave if $\gamma \geq 2$ and Schur-concave for all real-valued γ. However, for the limiting case when $\gamma = 1$ and $\sum_{i=1}^{n} p_i = 1$, (40) reduces to Shannon's entropy $-\sum_{i=1}^{n} p_i \log p_i$, which (being a sum of concave terms) is concave. Since $H_{K11\gamma}(P_n)$ is strictly decreasing in γ, the following interesting inequality follows from (40) when $\sum_{i=1}^{n} p_i = 1$:

$$1 - \prod_{i=1}^{n} p_i^{p_i} \leq -\sum_{i=1}^{n} p_i \log p_i \leq \left(\prod_{i=1}^{n} p_i^{p_i}\right)^{-1} - 1. \quad (41)$$

The upper bound in (41), although Schur-concave, is not concave when $\gamma = 0$. For the discrete uniform distribution $(1/n, \ldots, 1/n)$, (41) becomes the following well-known inequality:

$$1 - 1/n \leq \log n \leq n - 1.$$

VII. COMBINATIONS OF ENTROPIES

Besides the various entropies discussed above, one could also consider combinations of those individual entropies. This was done by Wondie and Kumar [40] for their proposed entropy

$$H_{WK\alpha} = \frac{1}{\alpha^{-1} - \alpha}\left[\log \sum_{i=1}^{n} p_i^{\alpha} + \sum_{i=1}^{n} p_i^{\alpha} - 1\right], \ \alpha > 0 \quad (42)$$

as a combination of Rényi's entropy in (4a) and Tsallis' entropy in (5). Alternatively, one could consider the mean or the sum of those two entropies as

$$H_{RT\alpha}(P_n) = \frac{1}{1-\alpha}\left(\log \sum_{i=1}^{n} p_i^{\alpha} + \sum_{i=1}^{n} p_i^{\alpha} - 1\right). \quad (43)$$

Both (42) and (43) relate to Shannon's entropy as

$$\begin{aligned} -\sum_{i=1}^{n} p_i \log p_i &= \lim_{\alpha \to 1} H_{WK\alpha}(P_n) \\ &= (1/2) \lim_{\alpha \to 1} H_{RT\alpha}(P_n). \end{aligned} \quad (44)$$

As a consequence of Theorems 4 and 6, the entropies in (42) and (43) are concave for $0 < \alpha \leq 1$ and Schur-concave for all $\alpha > 0$.

Another potentially interesting entropy could be based on Tsallis' entropy in (5) and Arimoto's entropy in (7) as follows:

$$H_{TA\alpha}(P_n) = \frac{1}{1-\alpha}\left[\sum_{i=1}^{n} p_i^{\alpha} - \left(\sum_{i=1}^{n} p_i^{1/\alpha}\right)^{\alpha}\right], \ \alpha > 0 \quad (45)$$

which, from Theorems 6 and 7 and since sums of concave functions are concave, is concave and Schur-concave for all $\alpha > 0$. Note that $H_{TA1}(P_n) = 2(-\sum_{i=1}^{n} p_i \log p_i)$. Two notable members of (45) are

$$H_{TA2}(P_n) = \left(\sum_{i=1}^{n} \sqrt{p_i}\right)^2 - \sum_{i=1}^{n} p_i^2;$$

$$H_{TA(1/2)}(P_n) = 2\left(\sum_{i=1}^{n} \sqrt{p_i} - \sqrt{\sum_{i=1}^{n} p_i^2}\right). \quad (46)$$

An interesting comparison can be made between the entropies in (46) and those in (34)-(36) for the general case when $\sum_{i=1}^{n} p_i \leq 1$.

VIII. STRICT SCHUR-CONCAVITY CONDITIONS

A. GENERAL CASE

The discussion in this paper so far has focused on the concavity and Schur-concavity of an entropy H, with the Schur-concavity being defined by (2). In order for H to be strictly Schur-concave, the inequality in (2) has to be a strict one. In that case, if one probability distribution is majorized by another one, their entropy values will necessarily be different. As an extreme consequence, the maximum value of H is then taken on by and only by the uniform distribution $(\sum_{i=1}^{n} p_i/n, \ldots, \sum_{i=1}^{n} p_i/n)$ for the possibly incomplete (generalized) case of $\sum_{i=1}^{n} p_i \leq 1$ or $(1/n, \ldots, 1/n)$ when $\sum_{i=1}^{n} p_i = 1$.

It is to be expected that the conditions for strict Schur-concavity will differ slightly from those of Schur-concavity and that those for $\sum_{i=1}^{n} p_i \leq 1$ will differ from those when $\beta = 1$ and $\sum_{i=1}^{n} p_i = 1$. For the most general entropy in (14), the strict Schur-concavity property can be expressed as follows:

Theorem 14: The entropy $H_{K\alpha\beta\gamma}(P_n)$ in (14) with $\sum_{i=1}^{n} p_i \leq 1$ is strictly Schur-concave if

$$\begin{aligned} &0 \leq \min\{\alpha, \beta\} \leq 1 < \max\{\alpha, \beta\} \text{ or} \\ &0 < \min\{\alpha, \beta\} < 1 \leq \max\{\alpha, \beta\} \end{aligned} \quad (47)$$

and for any $\gamma \in \mathbb{R}$.

Theorem 15: The entropy $H_{K\alpha\beta\gamma}(P_n)$ in (14) with $\beta = 1$ and $\sum_{i=1}^{n} p_i = 1$ is strictly Schur-concave if $\alpha > 0$ and for any $\gamma \in \mathbb{R}$.

Proofs: It follows from the theory of majorization [2, Ch. 3] that, for the ratio $R_{\alpha\beta}(P_n) = \sum_{i=1}^{n} p_i^{\alpha} / \sum_{i=1}^{n} p_i^{\beta}$,

$$\begin{aligned} &R_{\alpha\beta} \text{ is strictly Schur-convex if } 0 \leq \beta \leq 1 < \alpha \\ &\qquad \text{or } 0 < \beta < 1 \leq \alpha, \quad (48a) \\ &R_{\alpha\beta} \text{ is strictly Schur-concave if } 0 \leq \alpha \leq 1 < \beta \\ &\qquad \text{or } 0 < \alpha < 1 \leq \beta. \quad (48b) \end{aligned}$$

From (48a)-(48b) and the fact that $H_{K\alpha\beta\gamma}(P_n)$ as a function of $R_{\alpha\beta}(P_n)$ is seen to be strictly increasing for $\beta > \alpha$ and strictly decreasing for $\beta < \alpha$ and for all $\gamma \in \mathbb{R}$, it follows that $H_{K\alpha\beta\gamma}(P_n)$ is strictly Schur-concave under the parameter constraints in (47). This completes the proof of Theorem 14.

Consider now the case when $\beta = 1$ and $\sum_{i=1}^{n} p_i = 1$ when $\sum_{i=1}^{n} p_i^{\alpha}$ is strictly Schur-concave for $0 < \alpha < 1$ and strictly Schur-convex for $\alpha > 1$ [2, pp. 138-139]. Then, since $H_{K\alpha 1\gamma}(P_n)$ as a function of $\sum_{i=1}^{n} p_i^{\alpha}$ is strictly increasing for $0 < \alpha < 1$ and strictly decreasing for $\alpha > 1$ and for all $\gamma \in \mathbb{R}$, $H_{K\alpha 1\gamma}(P_n)$ is necessarily strictly Schur-concave for $\alpha > 0$ and $\gamma \in \mathbb{R}$, completing the proof of Theorem 15.

B. PARTICULAR CASES

It follows immediately from Theorem 15, with the obvious parameter conversions, that the following one-parameter entropies are strictly Schur-concave for $\alpha > 0$: $H_{R\alpha}(P_n)$ in (4), $H_{T\alpha}(P_n)$ in (5), $H_{A\alpha}(P_n)$ in (7), and $H_{LV}(P_n)$ in (9). With respect to the two-parameter entropies, it follows from Theorem 15 that $H_{SM\alpha\beta}(P_n)$ in (11), and $H_{RT\alpha\beta}(P_n)$ in (12) are strictly Schur-concave if $\alpha > 0$ and for all real-valued β. In the case of $H_{ADK\alpha\beta}(P_n)$ in (10), which is the limiting case of $H_{K\alpha\beta\gamma}(P_n)$ in (14) as $\gamma \to 1$, it follows from Theorem 14 that the entropy in (10) is strictly Schur-concave when the parameters meet the condition in (47). Similarly, from Theorem 14 with $\gamma = \alpha-\beta+1$ it follows that $H_{K\alpha\beta}(P_n)$ in (17) is strictly Schur-concave subject to (47).

For the two-parameter entropy $H_{K\alpha 1\delta}(P_n)$ in (32) with $\sum_{i=1}^n p_i \leq 1$, it follows from Theorem 14 with $\beta = 1$ and $\delta = (\gamma - 1)/(\alpha - 1)$ that $H_{K\alpha 1\delta}(P_n)$ is strictly Schur-concave if $\alpha > 0$ and for all $\delta \in \mathbb{R}$. This result implies that the particular member entropies in (33)-(36) are all strictly Schur-concave. Also, the entropies in (42)-(46) are strictly Schur-concave for $\alpha > 0$ since they are sums of strictly Schur-concave functions.

IX. DISCUSSION

A. SOME ENTROPY INEQUALITIES

Besides using the generalized entropy in (14) to derive conditions for concavity and Schur-concavity of individual members of $H_{K\alpha\beta\gamma}(P_n)$, this family of entropies may also serve other purposes such as deriving inequalities between family members. Such derivations can conveniently be based on the fact that $H_{K\alpha\beta\gamma}(P_n)$ is strictly decreasing in α, β, and γ for any given P_n. In the case of α and β (with γ fixed), this property of $H_{K\alpha\beta\gamma}(P_n)$ follows from $G_{\alpha\beta}(P_n)$ in (13) being a strictly increasing function of α and β for any given P_n [26, pp. 249] and from $H_{K\alpha\beta\gamma}(P_n)$ being a strictly decreasing function of $G_{\alpha\beta}(P_n)$ for any given real value of γ. The effect of varying γ on $H_{K\alpha\beta\gamma}(P_n)$ for fixed α, β, and P_n is determined from the following partial derivative:

$$\frac{\partial H_{K\alpha\beta\gamma}(P_n)}{\partial\gamma} = (1-\gamma)^{-2} \cdot \left[G_{\alpha\beta}^{\gamma-1}(P_n) - 1 - G_{\alpha\beta}^{\gamma-1}(P_n)\log G_{\alpha\beta}^{\gamma-1}(P_n)\right] \leq 0$$

with the bracketed term being nonpositive from the well-known inequality $x-1-x\log x \leq 0$ for $x > 0$. Consequently,

$$H_{K\alpha_1\beta_1\gamma_1}(P_n) \geq H_{K\alpha_2\beta_2\gamma_2}; \quad \alpha_1 \leq \alpha_2, \beta_1 \leq \beta_2, \gamma_1 \leq \gamma_2. \tag{49}$$

An interesting inequality from (49) may be between the entropies of Shannon, Rényi, and Tsallis for the general case when $\sum_{i=1}^n p_i \leq 1$. Since those entropies are the respective members of $H_{K\alpha 1\gamma}(P_n)$ with $\gamma = \alpha = 1$, $\gamma = 1$, and $\gamma = \alpha$, the following inequalities follow directly from (49):

$$-\sum_{i=1}^n p_i \log p_i \Big/ \sum_{i=1}^n p_i \leq \frac{1}{1-\alpha}\log\left(\sum_{i=1}^n p_i^\alpha \Big/ \sum_{i=1}^n p_i\right) \leq \frac{1}{1-\alpha}\left(\sum_{i=1}^n p_i^\alpha \Big/ \sum_{i=1}^n p_i - 1\right), \ 0 \leq \alpha \leq 1 \tag{50}$$

with the inequalities reversed if $\alpha \geq 1$. Equalities in (50) occur only in the limiting case of $\alpha \to 1$ when the three entropies are the same. Similarly, for Tsallis entropy (5) and the Landsberg-Vedral entropy in (9) and for $\sum_{i=1}^n p_i \leq 1$ when those entropies become (23) and (25) corresponding to $\gamma = \alpha$ and $\gamma = 2-\alpha$, it follows from (49) with $\beta_1 = \beta_2 = 1$ that

$$\frac{1}{\alpha-1}\left(\sum_{i=1}^n p_i \Big/ \sum_{i=1}^n p_i^\alpha - 1\right) \leq \frac{1}{1-\alpha}\left(\sum_{i=1}^n p_i^\alpha \Big/ \sum_{i=1}^n p_i - 1\right), \ 0 \leq \alpha \leq 1, \tag{51}$$

with the reverse inequality if $\alpha \geq 1$.

As another example, consider the two entropies in (8) and (10). Since the entropy in (8) is the member of $H_{K\alpha\beta\gamma}(P_n)$ in (14) for $\gamma = 1$ and $\alpha = \beta$ and that of (10) corresponds to $\gamma = 1$, the following inequality is an immediate consequence of (49):

$$-\sum_{i=1}^n p_i^\alpha \log p_i \Big/ \sum_{i=1}^n p_i^\alpha \leq \frac{1}{\beta-\alpha}\log\left(\sum_{i=1}^n p_i^\alpha \Big/ \sum_{i=1}^n p_i^\beta\right), \ 0 \leq \beta \leq \alpha. \tag{52}$$

B. SOME APPLICATION IMPLICATIONS

Since entropies have no fixed upper bounds unless n is fixed, interpretations of results when using entropies as summary measures become difficult. It is easier to interpret the extent of a characteristic (attribute) reflected by the distribution $P_n = (p_1, \ldots, p_n)$ and represented by a summary measure if the measure has a fixed range such as the [0, 1]-interval. Consequently, normalized entropies are sometimes being used.

Examples of normalized entropies include the following form of Shannon's entropy [13]:

$$H_s^*(P_n) = -\sum_{i=1}^n p_i \log p_i / \log n \in [0, 1] \tag{53}$$

and the normalized form of the quadratic entropy in (34):

$$H_Q^*(P_n) = \left(1 - \sum_{i=1}^n p_i^2\right)\Big/(1 - 1/n) \in [0, 1] \tag{54}$$

for complete distributions with $\sum_{i=1}^n p_i = 1$. Both of these measures have been used for measuring evenness (uniformity) among biological species (e.g., [31]) and for measuring variation of nominal categorical data (e.g., [41]).

The most general entropy introduced in (14) can similarly be normalized. From Theorem 1, $H_{K\alpha\beta\gamma}$ is Schur-concave so that, from the majorization in (29), the bounds on $H_{K\alpha\beta\gamma}(P_n)$ are as follows:

$$0 \leq H_{K\alpha\beta\gamma}(P_n) \leq \frac{1}{1-\gamma}\left[\left(\sum_{i=1}^n p_i/n\right)^{\gamma-1} - 1\right]. \tag{55}$$

Then, the normalized form becomes

$$H^*_{K\alpha\beta\gamma}(P_n) = \frac{\left(\sum_{i=1}^n p_i^\alpha)/\sum_{i=1}^n p_i^\beta\right)^{\frac{\gamma-1}{\alpha-\beta}} - 1}{\left(\sum_{i=1}^n p_i/n\right)^{\gamma-1} - 1} \in [0, 1] \quad (56)$$

subject to the parameter constraints on α and β in (15), but for any real-valued γ. Various other normalized entropies are then particular cases of (56). For example, it can be verified (using L'Hospital's rules) that (53) is the member of (56) for $\beta = 1, \gamma = 2 - \alpha, \sum_{i=1}^n p_i = 1$, and $\alpha = 1$ (in the limit as $\alpha \to 1$). Similarly, $H^*_Q(P_n)$ in (54) is the particular case of (56) when $\alpha = 2, \beta = 1, \sum_{i=1}^n p_i = 1$, and $\gamma = 2$. The evenness index proposed by Chao and Ricotta [42] is a member of (56) for $\beta = 1, \sum_{i=1}^n p_i = 1$, and $\gamma = 2 - \alpha$.

The parameters α and β affect the weights given to the different p_i's, emphasizing the larger p_i's over the smaller ones or vice versa. Consider, for example, the entropy in (17) that can be expressed as follows:

$$H_{K\alpha\beta}(P_n) = \sum_{i=1}^n w_i p_i,\ w_i = \frac{1}{\beta - \alpha}\left(p_i^{\alpha-1}/\sum_{i=1}^n p_i^\beta - 1\right) \quad (57)$$

indicating the effect of α and β on the set of weights $\{w_i\}$. For some real data P_n, instead of using a single entropy value based on fixed parameter values, entropy values can be computed as functions of the varying parameters, producing *entropy profiles*.

In the case of a two-parameter entropy such as $H_{K\alpha\beta}(P_n)$ in (17) or (57), the graph of $H_{K\alpha\beta}(P_n)$ as a function of α and β for given P_n would be a surface whereas for one-parameter entropies such as those in (4a)-(9), the graph would be a curve. For the measurement of biological diversity and evenness, for example, the application of such one-parameter profiles have been emphasized by some (e.g., [42] and [43]). However, a limitation on the use of such profiles arises if the profiles of $P_n = (p_1, \ldots, p_n)$ and $Q_m = (q_1, \ldots, q_m)$ cross, making comparisons difficult or meaningless.

Besides concavity and Schur-concavity, there are, of course, other important properties required of an entropy, especially when a particular entropy is used as a summary measure for real data. For the generic entropy H, such additional properties include *symmetry*: $H(P_n)$ is (permutation) symmetric in its arguments; *zero-indifference (expansibility)*: $H(p_1, \ldots, p_n, 0, \ldots, 0) = H(p_1, \ldots, p_n)$; *non-negativity*: $H(P_n) \geq 0$ for all P_n; *continuity*: H is a continuous function of all $p_i (i = 1, \ldots, n)$; *maximality*: $H(p_1, \ldots, p_n) \leq H(1/n, \ldots, 1/n)$ for all complete P_n; *monotonicity:* $H(1/n, \ldots, 1/n)$ is strictly increasing in n. Many of the entropies discussed above can be verified as having such additional properties. See also [44].

C. QUANTUM ENTROPIES

All the entropies discused in this paper are viewed as functions of a probability distribution $P_n = (p_1, \ldots, p_n)$ where $p_i \geq 0$ for $i = 1, \ldots, n$ and $\sum_{i=1}^n p_i = 1$ or more generally $\sum_{i=1}^n p_i \leq 1$. While not the subject of this paper, a brief mention of quantum entropies may be appropriate. The equivalent quantum entropies could also be formulated by substituting traces of density matrices for the probability summations in the entropies discussed above as done, for example, by Hu and Ye [6] when expressing the quantum equivalent of the Rathie-Taneja entropy in (12). See also [36] for the quantum equivalents of some other entropies. Thus, if ρ is a density matrix of a system of interest involving a finite dimensional Hilbert space, one could define the quantum equivalent to the most general entropy in (14) as follows:

$$H_{K\alpha\beta\gamma}(\rho) = \frac{1}{1-\gamma}\left[\left(\frac{\mathrm{Tr}(\rho^\alpha)}{\mathrm{Tr}(\rho^\beta)}\right)^{\frac{\gamma-1}{\alpha-\beta}} - 1\right],\ \alpha, \beta \geq 0, \gamma \in \mathbb{R} \quad (58)$$

where Tr is the trace.

The quantum entropy studied by Hu and Ye [6] would be a member of $H_{K\alpha 1\gamma}(\rho)$ in (58) for $\gamma = \beta(\alpha - 1) + 1$ if $\mathrm{Tr}(\rho) = 1$. Those authors gave a proof of the concavity of their quantum entropy with parameter constraints equivalent to those of Theorem 12. Ultimately, if $\beta = 1, \gamma = \alpha$, and $Tr(\rho) = 1$, then (58) reduces to the following limitating case:

$$H_{K111}(\rho) = \lim_{\alpha\to 1} H_{K\alpha 1\alpha}(\rho) = -Tr(\rho \log \rho)$$

which is the von Neumann entropy and the quantum version of Shannon's entropy.

X. CONCLUSION

The focus of this paper is on two important properties of an entropy: concavity and Schur-concavity. In order to make the analysis comprehensive and systematic, a new three-parameter entropy that includes other entropies as particular cases is being introduced. The parameter conditions under which this most general entropy is concave, Schur-concave, or both can then be used as a basis for exploring those properties for other entropies. The analysis throughout this paper is sufficiently general to include the potential of a probability distribution $P_n = (p_1, \ldots, p_n)$ being incomplete with $\sum_{i=1}^n p_i < 1$.

An argument in favor of generalized entropies is that they offer flexibility by means of the choice of parameter values appropriate for different situations. Furthermore, such generalizations serve to systematize or unify entropies and their properties as in the case of the three-parameter entropy in (14) or (58). The popularity of generalized entropies has been demonstrated by the large number of relevant publications and citations (e.g., Google Scholar lists about 10,000 citations to Tsallis [9] and 6,500 to Rényi [8]).

Besides introducing parameters to entropy functions, an alternative way of generalization would be to generalize the probability distribution itself by the use of so-called *escort distributions* introduced by Beck and Schögle [45]. By definition, such a distribution is given by

$$p_{\epsilon i} = \frac{p_i^\epsilon}{\sum_{i=1}^n p_i^\epsilon},\ i = 1, \ldots, n, -\infty < \epsilon < \infty$$

with negative values of the parameter ϵ requiring all $p_i > 0$. While beyond the scope of the present paper, the effect of substituting the distribution $\{p_{\epsilon i}\}$ for the original distribution $\{p_i\}$ in entropy formulations may be a worthwhile analysis.

The flexibility offered by generalized entropies and generalized probability distributions ($\sum_{i=1}^{n} p_i \leq 1$), assuming such important properties as Schur-concavity and concavity, provides for interesting and potentially important theoretical explorations as indicated by the extensive published literature. In terms of real applications, however, the utility of such generalization may so far seem less convincing. Further applied work is warranted.

REFERENCES

[1] J. N. Kapur and H. K. Kesavan, *Entropy Optimization Principles With Applications*. Boston, MA, USA: Academic Press, 1992.

[2] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*, 2nd ed. New York, NY, USA: Springer, 2011.

[3] J. N. Kapur, *Measures of Information and Their Applications*. New Delhi, India: Wiley Eastern, 1994.

[4] C. Arndt, *Information Measures*. Berlin, Germany: Springer, 2004.

[5] C. Beck, "Generalized information and entropy measures in physics," *Contemp. Phys.*, vol. 50, no. 4, pp. 495–510, 2009.

[6] X. Hu and Z. Ye, "Generalized quantum entropy," *J. Math. Phys.*, vol. 47, pp. 1–7, Feb. 2006.

[7] J. M. Amigó, S. G. Balogh, and S. Hernández, "A brief review of generalized entropies," *Entropy*, vol. 20, no. 813, pp. 1–21, 2018.

[8] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Math., Statist., Probab.* J. Neyman, Ed. Berkeley, CA, USA: Univ. California Press, 1961, pp. 547–561.

[9] C. Tsallis, "Possible generalizations of Boltzman–Gibbs statistics," *J. Stat. Phys.*, vol. 52, pp. 479–487, Jul. 1988.

[10] C. Gini, "Di una formula comprensiva delle medie," *Metron*, vol. 13, no. 2, pp. 3–22, 1938.

[11] A. Rényi, *Probability Theory*. Amsterdam, The Netherlands: North-Holland Publishing, 1970.

[12] J. Aczél and Z. Daróczy, *On Measures of Information and Their Applications*. New York, NY, USA: Academic, 1975.

[13] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 279–423, Oct. 1948.

[14] J. H. Havrda and F. Charvat, "Quantification methods of classification processes: Concepts of structural *a*-entropy," *Kybernetica*, vol. 3, no. 1, pp. 30–35, 1967.

[15] S. Arimoto, "Information-theoretical considerations on estimation problems," *Inf. Control*, vol. 19, no. 3, pp. 181–194, Oct. 1971.

[16] D. E. Boekee and J. C. A. Van der Lubbe, "The R-norm information measure," *Inf. Control*, vol. 45, no. 2, pp. 136–155, May 1980.

[17] J. Aczél and Z. Daróczy, "Über verallgemeinerte grasilineare mittelwerte, die mit Gewichtsfunktionen gebildet sind," *Publ Math. Debrecen*, vol. 10, pp. 171–190, Jan. 1963.

[18] J. N. Kapur, "On some properties of generalized entropies," *Indian J. Math.*, vol. 9, pp. 427–442, Jan. 1967.

[19] P. T. Landsberg and V. Vedral, "Distributions and channel capacities in generalized statistical mechanics," *Phys. Lett. A*, vol. 247, no. 3, pp. 211–217, Oct. 1998.

[20] B. D. Sharma and D. P. Mittal, "New nonadditive measures of entropy for discrete probability distributions," *J. Math. Sci.*, vol. 10, pp. 28–40, Aug. 1975.

[21] P. N. Rathie and I. J. Taneja, "Unified (r, s)-entropy and its bivariate measures," *Inf. Sci.*, vol. 54, nos. 1–2, pp. 23–39, Mar. 1991.

[22] P. Nath, "Entropy, inaccuracy and information," *Metrika*, vol. 13, no. 1, pp. 136–148, Dec. 1968.

[23] P. N. Rathie, "On a generalized entropy and a coding theorem," *J. Appl. Probab.*, vol. 7, no. 1, pp. 124–133, Apr. 1970.

[24] D. S. Hooda and A. Ram, "Characterization of a generalized measure of R-norm entropy," *Carribean J. Math. Comput. Sci.*, vol. 8, no. 1, pp. 18–31, 2002.

[25] D. S. Hooda and D. K. Sharma, "Generalized R-norm information measures," *J. Appl. Math., Statist., Informat.*, vol. 4, no. 2, pp. 153–168, 2008.

[26] P. S. Bullen, *Handbook of Means and Their Inequalities*. Dordrecht, The Netherlands: Kluwer, 2003.

[27] A. M. Mathai and P. N. Rathie, *Basic Concepts in Information Theory and Statistics*. New Delhi, India: Wiley Eastern, 1975.

[28] M. Dresher, "Moment spaces and inequalities," *Duke Math. J.*, vol. 20, no. 2, pp. 261–271, Jun. 1953.

[29] K. B. Stolarsky, "A stronger logarithmic inequality suggested by the entropy inequality," *SIAM J. Math. Anal.*, vol. 11, no. 2, pp. 242–247, Mar. 1980.

[30] A. Clausing, "Type t entropy and majorization," *SIAM J. Math. Anal.*, vol. 14, no. 1, pp. 203–208, Jan. 1983.

[31] T. O. Kvålseth, "Evenness indices once again: Critical analysis of properties," *SpringerPlus*, vol. 4, no. 1, pp. 1–12, Dec. 2015.

[32] I. Vajda, "Bounds on the minimal error probability in checking a finite or countable number of hypotheses," *Inform. Transmiss. Problems*, vol. 4, no. 1, pp. 9–17, 1968.

[33] D. Ellerman, "An introduction to logical entropy and its relation to Shannon entropy," *Int. J. Semantic Comput.*, vol. 7, no. 2, pp. 121–145, Jun. 2013.

[34] D. Ellerman, *New Foundations for Information Theory: Logical Entropy and Shannon Entropy*. Cham, Switzerland: Springer, 2021.

[35] T. O. Kvålseth, "On the measurement of randomness (uncertainty): A more informative entropy," *Entropy*, vol. 18, pp. 1–15, Apr. 2016.

[36] S. S. Chehade and A. Vershynina, "Quantum entropies," *Scholarpedia*, vol. 14, no. 2, pp. 1–19, 2019.

[37] R. König, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4337–4347, Sep. 2009.

[38] J. Liu, J. Xie, L. Chao, H. Wu, P. Ding, X. Chen, and H. Feng, "Min-entropy estimation for semiconductor superlattice true random number generators," *Sci. Rep.*, vol. 12, no. 1, pp. 1–9, Dec. 2022.

[39] D. Bors, *Data Analysis for the Social Sciences*. Thousand Oaks, CA, USA: Sage, 2018.

[40] L. Wondie and S. Kumar, "A joint representation of Rényi's and Tsallis' entropy with application in coding theory," *Int. J. Math. Math. Sci.*, vol. 2017, pp. 1–5, Oct. 2017.

[41] H. F. Weisberg, *Central Tendency and Variability*. Thousand Oaks, CA, USA: Sage, 1992.

[42] A. Chao and C. Ricotta, "Quantifying evenness and linking it to diversity, beta diversity, and similarity," *Ecology*, vol. 100, no. 12, pp. 1–15, Dec. 2019.

[43] A. J. Daly, J. M. Baetens, and B. DeBaets, "Ecological diversity: Measuring the unmeasurable," *Mathematics*, vol. 6, no. 7, p. 119, 2018.

[44] T. Kvålseth, "Entropy evaluation based on value validity," *Entropy*, vol. 16, no. 9, pp. 4855–4873, Sep. 2014.

[45] C. Beck and F. Schögle, *Thermodynamics of Chaotic Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

TARALD O. KVÅLSETH (Life Member, IEEE) was born in Brunkeberg, Telemark, Norway. He received the B.Sc. degree in mechanical engineering from the King's College, University of Durham, U.K., in 1963, and the M.S. and Ph.D. degrees in industrial engineering and operations research from the University of California, Berkeley, in 1966 and 1971, respectively.

He was a Faculty Member at the Georgia Institute of Technology, Atlanta, Georgia, from 1971 to 1974, and the Norwegian Institute of Technology, Trondheim, Norway, from 1974 to 1979. He has been a Faculty Member with the University of Minnesota, Minneapolis, MN, USA, since 1979. He is currently a Professor Emeritus with the Department of Mechanical Engineering and the Department of Industrial and Systems Engineering, University of Minnesota. He has published extensively in journals, conference proceedings, and book chapters and has edited two books. His research interests include information theory, statistical methods, and measurement theory.