Acta Mathematica Academiae Paedagogicae Nyíregyháziensis
31 (2015), 331–340
www.emis.de/journals
ISSN 1786-0091

ON SOME PROPERTIES OF TSALLIS ENTROPY ON MAJORIZATION LATTICE

P. K. BHATIA, SURENDER SINGH, AND VINOD KUMAR

Abstract. In this communication, order preserving property of Tsallis measure of entropy is proved. Further, subdditivity and submodularity of Tsallis measure of entropy on a majorization are established. Additionally, Shannon's [11] little-recognized legacy, an interesting concepts of information elements and information lattices is studied.

1. Introduction and preliminaries

In the seminal paper on Mathematical theory of communications [6], Shannon defined 'information' in context with communication system. The word 'information' has been given different meanings by various writers in the field of information theory. Many of these interpretations acquired the attention of researchers from other fields and significant literature is developed in last 50 years on the interdisciplinary applications. It is likely that study of information measures will prove sufficiently useful in certain applications to deserve further study. To address this issue the connection of information measures with different mathematical concepts and their characterization is essential. In this communication, effort is made to study a generalized information measure in light of combinatorics structures.

Shiva et al. [12] proved order preserving property of Rényi's measure [9] of information of order α. Zografos et al. [16] , Quesada and Taneja [8] studied the order preserving property of two parametric information measures. Here, we study the order preserving property of Tsallis entropy. Supermodularity and submodularity play an important role in fields which are at the crossroads of Information Theory and Combinatorics. More to the point, Fujishige [3] proved that the entropy is submodular on the Boolean lattice of a finite set of random variables, ordered according to subset inclusion, and used this

2010 *Mathematics Subject Classification.* 94A17,05B35.
Key words and phrases. Tsallis entropy, order preserving, majorization, subadditive, submodular.

property of the entropy to derive in a unified way several results in Information Theory. Guardia et al. [5] established new connections among matroids, classical information theory and error-correcting codes. Cicalese and Vaccaro [1] proved that the Shannon entropy is supermodular and subadditive on the majorization lattice. The main result of our communication is to show a kind of generalization to Cicalese and Vaccaro's result on the majorization lattice for Tsallis entropy. In this section, some preliminary concepts required in the proof of main results are presented briefly.

1.1. **Shannon Entropy.** The measure of information was defined Claude. E. Shannon [10] in his treatise paper in 1948.

$$H(P) = \sum_{i=1}^{n} p_i \log p_i \tag{1}$$

where

$$\Gamma_n = \left\{ P = (p_1, p_2, p_3 \ldots p_n) \mid p_i \geq \ 0,\ \sum_{i=1}^{n} p_i = 1,\ n \geq 2 \right\}$$

is the set of all complete finite discrete probability distributions.

1.2. **Order preserving property.** If the information content of a complete finite probability scheme (c.f.p.s.) P is greater than, or equal to or less than that of a c.f.p.s. Q in one measure which is accepted as a standard, as e.g., the Shannon measure, then this order should be preserved in any other measure of information i.e

$$H(P) \leq H(Q) \Rightarrow H_\alpha(P) \leq H_\alpha(Q) \tag{2}$$

Where α lies in some interval.

1.3. **Majorization theory.** In this section, we recall the basic notions of majorization theory [7] which are relevant to our context.

Majorization. Given two probability distributions $P = (p_1, p_2, p_3 \ldots p_n)$ and $Q = (q_1, q_2, q_3 \ldots q_n)$ with $p_1 \geq p_2 \geq p_3 \geq \cdots \geq p_n$ and $q_1 \geq q_2 \geq q_3 \geq \cdots \geq q_n$. We say that P is majorized by Q, in symbols $P \sqsubset Q$, if and only if

$$\sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i,\ \ k = 1, 2, \ldots, n.$$

The main link between the concept of majorization and the theory of inequalities is established by the notion of Schur-convex functions.

Schur-convex functions. A real-valued function ϕ defined on the set of n-dimensional probability vectors is said to be Schur-convex if it is order preserving with respect to the partial order $\sqsubset$, that is , if

$$P \sqsubset Q \Rightarrow \phi(P) \leq \phi(Q)$$

Lemma 1.1 ([7]). *If ϕ is invariant with respect to any permutation of its inputs and convex then ϕ is Schur-convex.*

By this lemma, we get that entropy function H(P) as defined in Eq.(1) is a Schur-convex function.

Majorization lattice. We first recall [7] that a lattice is a quadruple $(L, \sqsubset, \vee, \wedge)$ where L is a set, $\sqsubset$ is partial ordering on L , and for all $a, b \in L$ there is a unique greatest lower bound (g.l.b.) $a \wedge b$ and a unique least upper bound (l.u.b.) $a \vee b$. More precisely,

$$a \wedge b \sqsubset a, \ a \wedge b \sqsubset b \text{ and } a \vee b \sqsupset a, \ a \vee b \sqsupset b.$$

Now for $n \geq 2$, let

$$\triangle_n = \{P = (p_1, p_2, p_3 \dots p_n); p_i \in [0, 1]; \sum_{i=1}^{n} p_i = 1; p_i \geq p_{i+1}\}$$

be the set of all n-dimensional probability distributions , with components in non decreasing order. Then,$(\triangle_n, \sqsubset, \vee, \wedge)$ is a partially ordered set such that

$$(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) \sqsubset \ (\frac{1}{n-1}, \frac{1}{n-1}, \dots, 0) \sqsubset \ \dots (\frac{1}{2}, \frac{1}{2}, \dots, 0) \sqsubset (1, 0, \dots, 0)$$

i.e.

$$(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) \sqsubset P \sqsubset (1, 0, \dots, 0).$$

Infimum and supremum in majorization lattice. Let $P, Q \in \triangle_n$ and $\alpha(P, Q) = (a_1, a_2, \dots, a_n)$ with

$$\begin{aligned} a_i &= \min\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\} - \sum_{j=1}^{i-1} a_j \\ &= \min\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\} - \min\{\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i-1} q_j\}. \end{aligned}$$

Then $\alpha(P, Q) = P \wedge Q$.

Now let $P, Q \in \triangle_n$ and $\beta(P, Q) = (b_1, b_2, \dots, b_n)$ with

$$\begin{aligned} b_i &= \max\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\} - \sum_{j=1}^{i-1} b_j \\ &= \max\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\} - \max\{\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i-1} q_j\}. \end{aligned}$$

Note that the components of $\beta(P,Q)$ might not be in non increasing order i.e. it is not true in general that $\beta(P,Q) \in \triangle_n$.

Now on rearranging the components of $\beta(P,Q)$ in non increasing order we have

$$\beta^{'}(P,Q) = (b_1^{'}, b_2^{'}, \ldots, b_n^{'})$$

Clearly, $\beta^{'}(P,Q)$ is an upper bound of P and Q in $\triangle_n$.

Lemma 1.2 ([1]). *Let $P = (p_1, p_2, p_3 \ldots p_n)$ be the given probability distribution and j be the smallest integer in $\{2, 3, \ldots, n\}$ such that $p_j > p_{j-1}$. Moreover, let i be the greatest integer in $\{1, 2, \ldots, j-1\}$ such that*

$$p_{i-1} \geq \frac{\sum_{r=1}^{j} p_r}{j-i+1} = a.$$

Let the distribution $Q = (q_1, q_2, q_3 \ldots q_n)$ be defined as

$$q_r = \begin{cases} a, & r=i,\ i+1,\ \ldots,\ j; \\ p_r, & otherwise. \end{cases}$$

Then for probability distribution Q we have that $q_{r-1} \geq q_r$ for all $r = 2, \ldots, j$ and $\sum_{s=1}^{k} q_s \geq \sum_{s=1}^{k} p_s$ for $k = 1, 2, \ldots, n$.

Moreover, for all $t = (t_1, t_2, t_3 \ldots t_n) \in \triangle_n$ such that $\displaystyle\sum_{s=1}^{k} t_s \geq \sum_{s=1}^{k} p_s$ and $\displaystyle\sum_{s=1}^{k} t_s \geq \sum_{s=1}^{k} q_s$ for $k = 1, 2, \ldots, n$.

Now the vector $\beta(P,Q) = (b_1, b_2, \ldots, b_n)$, by iteratively applying the transformation in lemma 1.2, actually with no more (n-1) iterations, we eventually obtain a vector $r = (r_1, r_2, \ldots, r_n) = P \vee Q$, the least upper bound of P and Q in $\triangle_n$.

Supermodularity on majorization lattice. A real valued function defined on a lattice L is supermodular if and only if for any $a, b \in L$ we have that, $f(a \wedge b) + f(a \vee b) \geq f(a) + f(b)$.

Note that f is submodular if and only if $-f$ is supermodular.

2. Tsallis entropy

Tsallis entropy introduced by Tsallis [13] is defined as

$$S_q(P) = k\frac{1 - \sum_{i=1}^{n} p_i^q}{q-1}$$

For a given probability distribution $P = (p_1, p_2, p_3 \ldots p_n)$, where k is a positive constant and $q \in R^+$. Tsallis entropy is one parameter generalization of

Shannon entropy in the sense that

$$\lim_{q \to 1} S_q = S_1 = -k \sum_{i=1}^{n} p_i \log p_i$$

The characteristic property of Tsallis entropy is pseudoadditivity. Introduction of Tsallis entropy opened a new research area in statistical physics, providing an advantageous generalization of traditional Boltzmann Gibbs statistical mechanics. The generalization enables us to find a consistent treatment of dynamics in many non extensive physical systems such as long-range interactions, long-time memories, and multifractal structures, which cannot be coherently explained within the conventional Boltzmann Gibbs statistics [14]. In particular, the successful applications of Tsallis entropy can be often observed in dynamical chaotic systems. For example, the results derived from the generalized Kolmogorov Sinai entropy using Tsallis entropy have the perfect matching with nonlinear dynamical behavior such as the sensitivity to the initial conditions in chaos. On the other hand, the usual Kolmogorov Sinai entropy using Shannon entropy does not have these convenient properties; see [15], [6], [2] for further details. Thus, Tsallis entropy inspires many physicists to establish a generalized Boltzmann Gibbs statistical mechanics leading to numerous applications [14].

3. Properties of Tsallis entropy on majorization lattice

Some information theoretic properties have been proved by Furuichi [4]. In this section, we prove some properties of Tsallis entropy on a majorization lattice.

Theorem 3.1. *Tsallis entropy is order preserving.*

Proof. If $H(P) \leq H(Q)$ for any $P, Q \in \Gamma_n$ then it can obviously seen that $S_q(P) \leq S_q(Q)$ for all $q \in R^+$. □

Theorem 3.2. *The Tsallis entropy $S_q(P)$ is submodular on majorization lattice $(\triangle_n, \sqsubset, \vee, \wedge)$, i.e for all $P, Q \in \Gamma_n$*

$$S_q(P \wedge Q) + S_q(P \vee Q) \leq S_q(P) + S_q(Q).$$

Proof. We shall prove slightly stronger result:

$$S_q(P \wedge Q) + S_q(\beta(P, Q)) \leq S_q(P) + S_q(Q)$$

Recalling the definition of $(b_1^{'}, b_2^{'}, \ldots, b_n^{'})$ as re-ordered version of $\beta(P, Q)$; we have

$$P \vee Q \leq (b_1^{'}, b_2^{'}, \ldots, b_n^{'})$$

and by order preserving property we have,

$$S_q(P \vee Q) \leq S_q(b_1^{'}, b_2^{'}, \ldots, b_n^{'}).$$

Also,

$$S_q(\beta(P \vee Q)) \leq S_q(b_1^{'}, b_2^{'}, \ldots, b_n^{'}).$$

Therefore,

$$S_q(P \vee Q) \leq S_q(\beta(P \vee Q)).$$

Obviously, $S_q(P \wedge Q) \leq S_q(P)$ and $S_q(P \wedge Q) \leq S_q(Q)$. Here, we will also assume that P is not majorized by Q and,Q is not majorized by P otherwise there is obvious equality.

Let index $i \in \{2, 3, \ldots, n\}$ be an inversion point of P and Q if either

$$\sum_{t=1}^{i-1} p_t < \sum_{t=1}^{i-1} q_t \text{ and } \sum_{t=1}^{i} p_t > \sum_{t=1}^{i} q_t;$$

vice-versa, if

$$\sum_{t=1}^{i-1} p_t > \sum_{t=1}^{i-1} q_t \text{ and } \sum_{t=1}^{i} p_t < \sum_{t=1}^{i} q_t.$$

Let $2 \leq i_1 < i_2 < \cdots < i_k \leq n$ be all the inversion points of for P and Q. Moreover, let

$$L = (l_1, l_2, \ldots, l_n) = P \wedge Q$$

and

$$S = (s_1, s_2, \ldots, s_n) = \beta(P, Q).$$

We assume that $i_0 = 0$ and $i_{k+1} = n + 1$ for sake of definiteness. We establish two claims.

Claim 1. For each inversion point i_r, $r = 0, 1, \ldots, k$, we have

$$\sum_{t=i_r+1}^{i_{r+1}} (p_t^q + q_t^q) = \sum_{t=i_r+1}^{i_{r+1}} (s_t^q + l_t^q)$$

To prove this claim, Without loss of generality, we assume that

$$\sum_{t=1}^{i_r} p_t > \sum_{t=1}^{i_r} q_t$$

Then

$$\sum_{t=1}^{i_r} s_t = \sum_{t=1}^{i_r} p_t$$

and for all $d = i_r + 1, \ldots, i_r - 1$, we have

$$s_d = \sum_{t=1}^{d} p_t - \sum_{t=1}^{d-1} s_t = p_d.$$

Accordingly, we have

$$\sum_{t=1}^{i_r} l_t = \sum_{t=1}^{i_r} q_t$$

and for all $d = i_r + 1, \ldots, i_{r+1} - 1$, we have

$$t_d = \sum_{t=1}^{d} q_t - \sum_{t=1}^{d-1} l_t = q_d$$

Consequently, we have

$$\sum_{t=i_r+1}^{i_{r+1}} (p_t^q + q_t^q - s_t^q - l_t^q) = 0.$$

Claim 2. For each inversion point i_r, $r = 0, 1, \ldots, k$, we have

$$p_{i_r}^q + q_{i_r}^q - s_{i_r}^q - l_{i_r}^q < 0$$

Let us write p, q, s, l instead of $p_{i_r}, q_{i_r}, s_{i_r}, l_{i_r}$ respectively. Without loss of generality let us assume that for the inversion point i_r, it holds that

$$\sum_{t=1}^{i_r-1} p_t > \sum_{t=1}^{i_r-1} q_t \text{ and } \sum_{t=1}^{i_r} p_t < \sum_{t=1}^{i_r} q_t.$$

It follows that $l, s > p$ and $q > l, s$. Moreover, it is not hard to see that $s + l = p + q$.

Let $s > l$, then write $s = q - \delta$, and $l = p + \delta$. Then we have

$$\begin{aligned} s^q + l^q - p^q - q^q &= (q-\delta)^q + (p+\delta)^q - p^q - q^q \\ &= q^q(1 - \frac{\delta}{q})^q + p^q\big(1 - \frac{\delta}{p})^q - p^q - q^q \\ &= q\frac{\delta}{q} - q\frac{\delta}{p} < 0, \end{aligned}$$

neglecting the higher powers. Consequently,

$$s^q + l^q - p^q - q^q < 0$$

The remaining case $s < l$, is completely symmetric, provided that we set $s = p + \delta$, and $l = q - \delta$. This concludes the proof of claim 2.

Now we have

$$\begin{aligned} S_q(S) + S_q(L) - S_P(S) - S_q(Q) &= \sum_{t=1}^{n} (s_t^q + l_t^q - p_t^q - q_t^q) \\ &= \sum_{r=1}^{k} (s_{i_r}^q + l_{i_r}^q - p_{i_r}^q - q_{i_r}^q) \\ &\quad + \sum_{t=i_{r-1}+1}^{i_r-1} (s_t^q + l_t^q - p_t^q - q_t^q) \\ &\quad + \sum_{t=i_k+1}^{n} (s_t^q + l_t^q - p_t^q - q_t^q) \end{aligned}$$

$$= \sum_{r=1}^{k} (s_{i_r}^q + l_{i_r}^q - p_{i_r}^q - q_{i_r}^q) < 0,$$

by claim 1 and 2. That is,

$$S_q(S) + S_q(L) - S_P(S) - S_q(Q) < 0$$
$$S_q(S) + S_q(L) < S_P(S) + S_q(Q).$$

Consequently,

$$S_q(P \wedge Q) + S_q(\beta(P, Q)) < S_P(S) + S_q(Q).$$

Thus the submodularity of Tsallis entropy follows. □

Theorem 3.3. *The Tsallis entropy $S_q(P)$, $(q > 1)$ is subadditive on majorization lattice $(\triangle_n, \sqsubset, \vee, \wedge)$, i.e for all $P, Q \in \Gamma_n$*

$$S_q(P \wedge Q) \leq S_q(P) + S_q(Q).$$

Proof. Let $L = (l_1, l_2, \ldots, l_n) = P \wedge Q$ and $S = (s_1, s_2, \ldots, s_n) = \beta(P, Q)$.

If $P \sqsubset Q$, by definition of g.l.b $P \wedge Q = P$ and the result is obvious, similar is the case for $Q \sqsubset P$.

Now, we assume, P is not majorized by Q and Q is not majorized by P. This also implies that $n \geq 2$.

For each $i = 2, 3, 4, \ldots, n$, it holds that

$$p_i^q + q_i^q - l_i^q \geq 0$$

By definition of L, we have

$$\min\{p_i^q, q_i^q\} \leq l_i^q \leq \max\{p_i^q, q_i^q\}$$

Now, we have

$$\sum_{i=1}^{n} p_i^q + \sum_{i=1}^{n} q_i^q - \sum_{i=1}^{n} l_i^q \geq 0$$

It gives,

$$\frac{(\sum_{i=1}^n p_i^q - 1)}{1-q} + \frac{(\sum_{i=1}^n q_i^q - 1)}{1-q} - \frac{(\sum_{i=1}^n l_i^q - 1)}{1-q} \geq \frac{-1}{1-q}$$

Therefore, for $q > 1$,

$$S_q(P \wedge Q) \leq S_q(P) + S_q(Q)$$

Hence, Tsallis entropy is subadditive on majorization lattice $\triangle_n$ when parameter $q > 1$. □

4. Some open problems

For a random variable X, let $H(X)$ is Shannon's entropy. Shannon [11] considered a random variable as an information element and for two information elements X and Y defined a relation '$\leq$' as follows: '$X \geq Y$ if and only if $H(Y|X) = 0$' i.e. this essentially requires that Y can be obtained by suitable finite state operation (or limit of such operations) on X. If $X \geq Y$ the Y is called 'abstraction' of X. Let $\Im$ be the set of all information elements then it is observed that $(\Im, \geq)$ is a partially ordered set. Further, if $X + Y =$ the total information of X and Y=l.u.b{X,Y}; $X * Y =$ common information of X and Y=g.l.b{X,Y}. Then $(\Im, \geq, +, *)$ is a lattice and is called an information lattice.

The notions of information elements and information lattices seems to have some interdisciplinary applications. This little exploited legacy of Shannon [11] needs further investigations. In the present context some open problems arise:

- Possibility of existence of an isomorphism between information lattice and majorization lattice.
- The information lattice as defined by Shannon [11] is neither distributive nor modular but can be made relatively complimented. Information lattice needs a parallel characterization as that of conventional lattice, and their is possibility of development of Information lattice theory.
- Exploration of applications of information lattice theory.

5. Concluding remarks

Along with proposal of some open problems in section 4, we have proved that the Tsallis entropy is order preserving, submodular and subadditive on the majorization lattice. Likewise to the classical case, one could use the subadditivity property as a basis to define a new information-theoretical calculus in $\triangle_n$. In such a calculus, the entropy $S_q(P \wedge Q)$ plays analogous role of $S_q(P * Q)$ the classical joint entropy of two random variables with marginal probability distributions P and Q, respectively. Proceeding with the analogy, one could define a conditional Tsallis entropy

$$S_q(P|Q) = S_q(P \wedge Q) - S_q(Q)$$

and, because of Theorem 3.3, one would get that $S_q(P|Q) \leq S_q(Q)$ for all $P, Q \in \triangle_n$.

References

[1] F. Cicalese and U. Vaccaro. Supermodularity and subadditivity properties of the entropy on the majorization lattice. *IEEE Transactions on Information Theory*, 48(4):933–938, 2002.

[2] U. M. S. Costa, M. L. Lyra, A. R. Plastino, and C. Tsallis. Power-law sensitivity to initial conditions within a logisticlike family of maps: Fractality and nonextensivity. *Phys. Rev. E*, 56:245–250, Jul 1997.

[3] S. Fujishige. Polymatroidal dependence structure of a set of random variables. *Inform. and Control*, 39(1):55–72, 1978.

[4] S. Furuichi. Information theoretical properties of Tsallis entropies. *J. Math. Phys.*, 47(2):023302, 18, 2006.

[5] G. G. La Guardia, L. Grossi, R. C. A. Vieira, and M. A. V. Pinto. A note on matroids, codes and information theory. *Int. J. Pure Appl. Math.*, 70(3):307–316, 2011.

[6] M. L. Lyra and C. Tsallis. Nonextensivity and multifractality in low-dimensional dissipative systems. *Phys. Rev. Lett.*, 80:53–56, Jan 1998.

[7] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: theory of majorization and its applications*. Springer Series in Statistics. Springer, New York, second edition, 2011.

[8] V. Quesada and I. J. Taneja. Order preserving property of unified (r, s)-information measures. *Soochow J. Math.*, 18(4):379–395, 1992.

[9] A. Rényi. On measures of entropy and information. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I*, pages 547–561. Univ. California Press, Berkeley, Calif., 1961.

[10] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.

[11] C. E. Shannon. The lattice theory of information. *IEEE Transactions on Information Theory*, 1:105–107, 1953.

[12] S. G. S. Shiva, N. U. Ahmed, and N. D. Georganas. Order preserving measures of information. *J. Appl. Probability*, 10:666–670, 1973.

[13] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Statist. Phys.*, 52(1-2):479–487, 1988.

[14] C. Tsallis, S. Abe, and Y. Okamoto. Nonextensive statistical mechanics and its applications. In *Nonextensive Generalization of Boltzmann-Gibbs Statistical Mechanics and its Applications*, volume 560 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 2001.

[15] C. Tsallis, A. R. Plastino, and W.-M. Zheng. Power-law sensitivity to initial conditions—new entropic representation. *Chaos Solitons Fractals*, 8(6):885–891, 1997.

[16] K. Zografos, K. Ferentinos, and T. Papaioannou. Order preserving property of measures of information. *Comm. Statist. Theory Methods*, 18(7):2647–2656, 1989.

Received October 24, 2013.

P. K. Bhatia
Depatment of Mathematics,
DCR University of Science and Technology,
Murthal-131039, Haryana, India
E-mail address: bhatiapk@rediffmail.com

Surender Singh (corresponding author)
School of Mathematics,
Shri Mata Vaishno Devi University,
Sub-post office, Katra-182320, J & K, India
E-mail address: surender1976@gmail.com

Vinod Kumar
Department of Mathematics,
Pt. C. L. S. Government (P. G.) College,
Karnal-132001, Haryana, India
E-mail address: vkmajra@gmail.com