RÉNYI EXTRAPOLATION OF SHANNON ENTROPY

KAROL ŻYCZKOWSKI

ABSTRACT. Relations between Shannon entropy and Rényi entropies of integer order are discussed. For any N–point discrete probability distribution for which the Rényi entropies of order two and three are known, we provide an lower and an upper bound for the Shannon entropy. The average of both bounds provide an explicit extrapolation for this quantity. These results imply relations between the von Neumann entropy of a mixed quantum state, its linear entropy and traces.

ver. 2 with corrigendum added,   February 17, 2005

1. INTRODUCTION

We are going to analyze discrete probability distributions $\vec{x} = \{x_1, ..x_N\}$, which consist of non–negative numbers summing to unity. $\sum_{i=1}^{N} x_i = 1$. To characterize quantitatively such probability vectors one uses Shannon (information) entropy [1]

$$(1) \qquad H(\vec{x}) := -\sum_{i=1}^{N} x_i \ln x_i$$

where we adopt the convention that $0 \ln 0 = 0$, if necessary.

The Shannon entropy is distinguished by several unique properties [2], but it is often convenient to introduce generalized *Rényi entropies* [3] parametrized by a continuous parameter q,

$$(2) \qquad H_q(\vec{x}) := \frac{1}{1-q} \ln \Big[\sum_{i=1}^{N} x_i^q \Big].$$

The Rényi entropies are well defined for positive $q \neq 1$, but is is not difficult to show that for any probability distribution $\vec{x}$ one has $\lim_{q \to 1} H_q(\vec{x}) = H(\vec{x})$. For consistency the Shannon entropy H will thus be denoted by H_1. This method of generalizing the Shannon entropy is by far not the only one – for reviews of other generalizations see books by Kapur [4] and Arndt [5].

In this work we discuss relations between Rényi entropies of different orders, and in particular between H_1, H_2 and H_3. Physical motivation for such a study is twofold. First, we may not know the entire vector $\vec{x}$, but only a few of its L_q norms, so knowing the Renyi entropies, say H_2 and H_3 we want to estimate the unknown Shannon entropy H_1. Such a situation occurs if one studies an N dimensional quantum mechanical mixed state ρ according to the scheme recently proposed by P. Horodecki et al. [6, 7] and measures directly the traces $\text{Tr}\rho^k$ for $k = 2, 3, ...M$. If $M < N$ the entire spectrum of ρ remains unknown, and it is not possible to find its von Neumann entropy, $S(\rho) = -\text{Tr}\rho \ln \rho$, (i.e. the Shannon entropy of the spectrum), but the generalized Renyi entropies H_k, including the linear entropy, which is a function of H_2 may be readily obtained. Similar problems arise in many

different branches of physics, for instance by the study of the statistics between particles created in high–energy collisions [8, 9]. Measuring probabilities of that two independent collisions give rise to the same distribution of particles allows one to obtain the Rényi entropy H_2, but not directly the Shannon entropy H_1.

Another reason to study relations between H_1 and H_2 has been provided by the work by Pipek and Varga [10]. They assumed that both these quantities are known, and observed that their difference, $S_{str} := H_1 - H_2$ called *structural entropy*, provides an important characterization of the analyzed probability vector $\vec{x}$. For instance, an increase of the structural entropy charactering an eigenstate of a tight binding model indicates the Anderson transition. Several other applications of structural entropy include also quantum chemistry and statistical analysis of quantum spectra, (see [11] and references therein).

The von Neumann entropy of a mixed state obtained by partial trace of a bi–partite pure state, $\rho = \text{Tr}_B(|\psi\rangle\langle\psi|)$, measures the degree of entanglement of the pure state $|\psi\rangle$. Alternatively one can measure the entanglement by generalized entropies (see e.g. [12]), so relations between entropies analyzed in this work provide bounds between different measures of entanglement. This very point has recently been discussed in the paper by Wei et al. [13], which provides an additional motivation for the present work.

This paper is ogranized as follows. In section 2 the basic properties of the Rényi entropies are reviewed. In section 3 we present recent results of Topsøe and Harremoës [14, 15], which allow us to propose lower and upper bounds on the Shannon entropy obtained out of the Rényi entropies of order two and three, provided the length N of the vector is known. They are derived in section 4, while in section 5 we propose and analyze en estimation of the Shannon entropy.

2. SHANNON AND RÉNYI ENTROPIES

Consider a random variable ξ attaining not more than N different values with probabilities x_i, $i = 1, \ldots, N$. Such discrete probability distribution P may be cahracterized by the Shannon entropy (1) or generalized Rényi entropies (2).

All generalized entropies H_q vary from zero for a certain event (the distribution $Q_1 := \{1, 0, ..., 0\}$) to $\ln N$, for the uniform distribution, (the distribution $Q_N := \{1/N, 1/N, ..., 1/N\}$). For the distributions with k equal elements, $Q_k := \{1/k, ..., 1/k, 0, ..., 0\}$, the entropies admit intermediate values, $\ln k$.

The Rényi entropy H_q converges to the Shannon entropy in the limit $q \to 1$. It is also useful to express the Shannon entropy as the limit of the derivative,

$$(3) \qquad H(\vec{x}) = -\lim_{q \to 1} \frac{\partial[(1-q)\exp(H_q(\vec{x}))]}{\partial q}.$$

Some special cases of H_q are of special interest. For $q = 2$ we have $H_2(\vec{x}) = -\ln[\sum_{i=1}^{N} x_i^2]$. The Rényi entropy of order two, called *extension entropy* [10], is closely related to the *inverse participation ratio*,

$$(4) \qquad R(\vec{x}) := \frac{1}{\sum_{i=1}^{N} x_i^2} = \exp[H_2(\vec{x})].$$

This quantity characterizes the "effective number of different events" which the stochastic variable may admit, and varies from unity for Q_1, to N for the uniform distribution Q_N. Another quantity $r = 1/R = \sum_{i=1}^{N} x_i^2$, called *index of coincidence* [15], in quantum mechanical problems is called *purity*, since the larger r the more

pure, the state it describes. The quantity $L := 1 - r = 1 - \exp(-H_2)$ is called *linear entropy* since in analogy to Shannon entropy it achieves its maximum for the uniform distribution Q_N.

In the case $q = 0$ the Rényi entropy is a function of the number m of positive components of the vector, $H_0(\vec{x}) = \ln m$. In the limit $q \to \infty$ we obtain a quantity analogous to the Chebyshev norm: $H_\infty = -\ln x_{\max}$, where $x_{\max}$ is the largest component of $\vec{x}$.

The Rényi entropy (2) is a sum of N terms so for any finite N the function of H_q on q is differentiable. The functional dependence of the Rényi entropy on its parameter was investigated in detail by Back and Schlögl [16]). Making use of the fact that the function x^s is convex for $s > 1$ and concave for $0 \le s \le 1$ they have proved several inequalities[1], which we recall in the case $q > 0$,

$$(5) \qquad \frac{\partial}{\partial q} H_q \le 0,$$

$$(6) \qquad \frac{\partial}{\partial q} \frac{q-1}{q} H_q \ge 0,$$

$$(7) \qquad \frac{\partial}{\partial q} (1-q) H_q \le 0,$$

$$(8) \qquad \frac{\partial^2}{\partial q^2} (1-q) H_q \ge 0.$$

The first inequality (5) means that the Rényi entropy is a non increasing function of its parameter,

$$(9) \qquad H_q(\vec{x}) \ge H_s(\vec{x}) \qquad \text{for} \qquad \text{any} \qquad s > q$$

and this statement is valid also for infinite probability vectors and the cases of nondifferentiable H_q [16]. Hence the structural entropy $S_{str} := H_1 - H_2$ is non–negative [10].

Inequality (8) implies that the dependence of the Rényi entropy on its parameter is convex.[2] Thus knowing the 0 and 2–entropies one obtains by linear interpolation an upper bound for the Shannon entropy

$$(10) \qquad H_1(\vec{x}) \le H_{u0} := \frac{1}{2}(H_0(\vec{x}) + H_2(\vec{x})).$$

This relation gives us an upper bound for the structural entropy

$$(11) \qquad S_{str}(\vec{x}) := H_1(\vec{x}) - H_2(\vec{x}) \le \frac{1}{2}\big(H_0(\vec{x}) - H_2(\vec{x})\big)$$

valid for any vector $\vec{x}$ of a finite length N.

In an analogous way, if the Rényi entropies of order 2 and 3 are known, the linear extrapolation provides a lower bound for the Shannon entropy

$$(12) \qquad H_1(\vec{x}) \ge H_{d23}(\vec{x}) = 2H_2(\vec{x}) - H_3(\vec{x}),$$

---

[1]In the book[16] the quantity $-H_q$ called *Rényi information* was analyzed, so the direction of the inequalities derived there is inverted.

[2]This claim is withdrawn - see into corrigendum Sec. 7, in which consequences of this error are pointed out.

which combined with (10), allows one to write down a simple estimation $H_{023} = (H_d + H_{u0})/2$, (see Fig 3.a),

$$(13) \qquad H_1(\vec{x}) \approx H_{023}(\vec{x}) := \frac{1}{4}[H_0(\vec{x}) + 5H_2(\vec{x}) - 2H_3(\vec{x})].$$

Making use of the inequality (6) we obtain the relation

$$(14) \qquad \frac{q-1}{q}H_q(\vec{x}) \leq \frac{s-1}{s}H_s(\vec{x}) \qquad \text{for} \quad \text{any} \qquad s \geq q \,,$$

which is equivalent to the statement that the L_q–norm is a non–increasing function, $||\vec{x}||_s \leq ||\vec{x}||_q$. This result provides another upper bound, $H_q \leq q(s-1)H_s/s(q-1)$. Although it is not applicable for the Shannon entropy, for which $q = 1$ so the inequality becomes trivial, but it gives an usefull bound on H_q with $g > 1$ by the limiting value H_∞,

$$(15) \qquad H_q(\vec{x}) \leq \frac{q}{q-1}H_\infty(\vec{x}).$$

In further sections of this work we shall discuss possibilities of finding more precise bounds and estimations for the Shannon entropy, provided the dimension N of the probability vector is known.

3. Bounds between Rényi entropies

For any value $q \geq 0$ the generalized entropy H_q is equal zero for certain events described by the distribution Q_1, and achieves its maximum for the uniform distribution, $S(Q_N) = \ln N$.

To investigate further relations between the Rényi entropies of different order we have chosen to analyze the case of $N = 3$ dimensional vectors $\vec{x}$. The space of all possible probability vectors, plotted in the the plane $x_3 = 1 - x_1 - x_2$ forms an equilateral triangle of side $\sqrt{2}$ measured in the Euclidean distance. Its three corners: (100), (010) and (001) represent certain events, while the center of the triangle corresponds to the uniform distribution Q_3.

Fig. 1 shows sets of points characterized by the same Rényi entropy of order q, which may be called iso-entropy curves. Independently of the value of the parameter q the generalized entropy attains its minimum, $H_q = 0$, at the corners of the triangle, while the maximum $H_q = \ln 3$ is achieved at the point Q_3 at the center of the triangle. As shown in Fig. 1a the maximum is rather flat for $q = 1/4$. The case shown in this panel resembles the limiting case H_0, for which the entropy reflects the number of events which may occur: it vanish at the corners of the triangle, is equal to $\ln 2$ at its sides and equals to $\ln 3$ for any point inside the triangle. The other example, $q = 8$, presented in Fig. 1d. is similar to the limiting case H_∞, for which the iso-entropy curves are perpendicular to the lines joining Q_3 with the corners.

Superimposing some of the above pictures on one graph allows one to understand further relations between the Rényi entropies. The generalized entropies are correlated; e.g. for the distributions Q_k the entropies are equal to $\ln k$ independently of the value of q.

The problem, which values the entropy H_q may admit, provided H_s is given, has been solved by Harremoës and Topsøe [15]. For any distribution $P \in \mathbb{R}^N$ they proved a simple (but not very sharp) upper bound on H_1 by H_2,

$$(16) \qquad H_2(\vec{x}) \leq H_1(\vec{x}) \leq \ln N + 1/N - \exp(-H_2(\vec{x}))$$
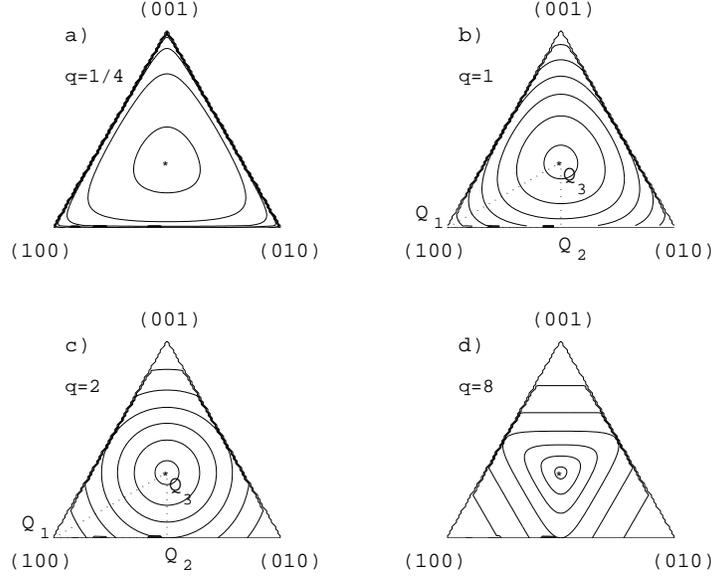
FIGURE 1. Iso-entropy curves in the space of probability distributions with $N = 3$. The generalized Rényi entropy is constant along the curves plotted for (a) $q = 1/4$, (b) $q = 1$ (Shannon entropy), (c) $q = 2$ (Euclidean circles–distance D_2), and (d) $q = 8$. Dotted lines form the triangle $\triangle(Q_1, Q_2, Q_3)$.

where the lower bound is a special case of (9). Moreover, they showed that the set $\Delta_{q,s}$ of possible probability distributions plotted in the plane H_q versus H_s is not convex (see Fig. 2), and its boundaries are formed of arcs corresponding to the interpolating probability distributions

(17) $$Q_{k,l}(a) := aQ_k + (1 - a)Q_l \qquad \text{with} \qquad a \in [0, 1].$$

More precisely, for any probability distribution P consisting of N components and arbitrary $s > q > 0$ the following bounds hold [15]

(18) $$H_q(Q_{k-1,k}(a)) \leq H_q(P) \leq H_q(Q_{1,N}(a)),$$

where a is a function of the known value of $H_s(P)$ and the natural number k is selected by the inequality $\ln(k - 1) \leq H_s(P) \leq \ln k$.

The above results, crucial for the main body of this work, are easy to understand. Let us discuss the simplest nontrivial case with $N = 3$. The two dimensional simplex of probability distributions may be divided into 6 identical parts, equivalent to the triangle $\triangle(Q_1, Q_2, Q_3)$, as shown in Fig. 1c. Three sides of the triangle are formed of the interpolating distributions $Q_{1,2}$, $Q_{1,3}$ and $Q_{2,3}$ and these distinguished probabability distributions are extreme in a sense that they lead to the bounds (18). The bounds between H_1 and H_2 for $N = 3$ are presented in Fig. 2a. To obtain them it is sufficient to travel along the sides of the triangle $\triangle(Q_1, Q_2, Q_3)$, computing H_1 and H_2 at each point and to plot the data obtained in the plane H_1 versus H_2.

More formally, the upper boundary of the set $\Delta_{q,s}$ consist of one arc derived from the family of distributions $Q_{1,N}(a)$; for any value of a we compute $H_s(a)$, invert it to obtain $a(H_s)$ and plot $H_q(a(H_s))$. In the case $N = 3$ the upper bounds plotted

in Fig. 2a, c and e arise from the hypotenuse Q_1, Q_3 of the triangle $\triangle(Q_1, Q_2, Q_3)$ from Fig. 1.b and c.

In the similar way the lower bound may be derived from the distributions $Q_{k,k+1}(a)$ for $k = 1, ...N - 1$. It consists of $(N - 1)$ arcs forming an cascade $H_q(Q_{1,2}(a)) \smile H_q(Q_{2,3}(a)) \smile \cdots \smile H_q(Q_{N-1,N}(a))$. Note that the distributions Q_k are represented in each plot by the points $(\ln k, \ln k)$, which connect the neighbouring arcs. For $N = 3$ the lower bound consists of two arcs, corresponding to the adjacent sides Q_1, Q_2 and Q_2, Q_3 of $\triangle(Q_1, Q_2, Q_3)$.

The shape of the set $\Delta_{q,s}$ requires a comment. The $N - 1$ dimensional simplex – the set of all N–points probability distributions is convex and any of its projections onto a plane forms a convex set. However, its image at the plane H_q versus H_s needs not to be convex, since the transformations $H_q(\vec{x})$ and $H_s(\vec{x})$ are nonlinear. The boundaries of $\Delta_{q,s}$ are obtained as the image of an appropriately chosen path on the boundary of the simplex. In the case considered it is the path $Q_1 \rightarrow Q_2 \rightarrow \cdots \rightarrow Q_N \rightarrow Q_1$, independently of the values of q and s. Observe that the general structure of the set $\Delta_{q,s}$ does not depend on s. However, the larger difference $s - q$, the larger area of the set: the less information on H_q is provided by H_s.
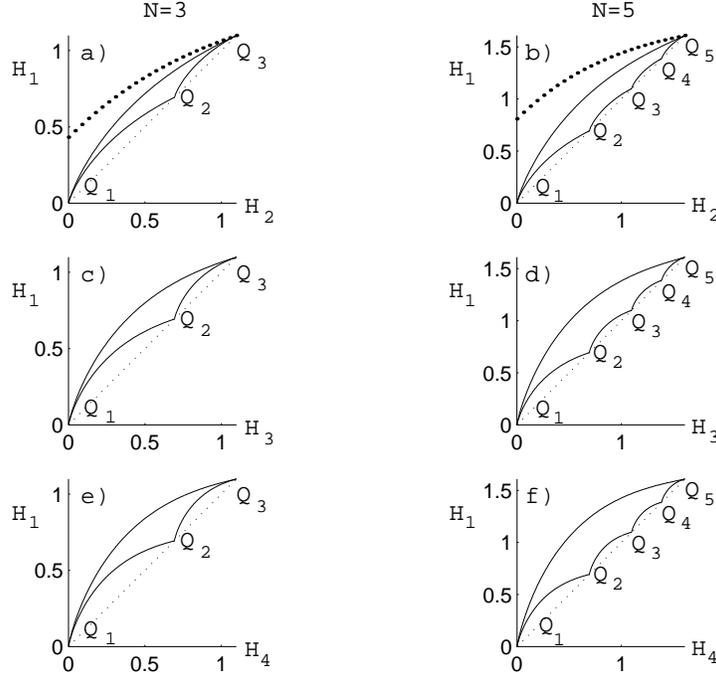

FIGURE 2. The set of all possible discrete distributions for $N = 3$ and $N = 5$ at the Rényi entropies plane H_1 and H_q: $q = 2$ (a,b); $q = 3$ (c,d) and $q = 4$ (e,f). Thin dotted lines in each panel represent the monotonicity lower bounds (9) while bold dotted curves in panel (a) and (b) denote upper bound (16).

Let us emphasize in this point that results presented in [15] do not close the issue of finding bounds and relations between different entropies. Results analogous to

(18) for a more general class of entropy functions were recently obtained by Berry and Sanders [17]. A more precise lower bound for Shannon entropy, quadaratic in terms of index of coincidence (purity) was found by Topsøe [18].

4. N–DEPENDENT BOUNDS FOR SHANNON ENTROPY

4.1. Bounds based on H_2 and the length N of the probability vector.
Results (18) allow us to obtain bounds for a value of the entropy H_q, provided the value H_s is known. Let us first assume that the entropy H_2 is known and we want to extract some information on H_1. We start computing the Rényi entropy of order two,

$$(19) \qquad H_2(Q_{1,N}(a)) = -\ln\Big[\frac{(1 + (N-1)a)^2}{N^2} + (N-1)\frac{(1-a)^2}{N^2}\Big]$$

and invert it to obtain

$$(20) \qquad a = \sqrt{\frac{N\exp(-H_2) - 1}{N - 1}}.$$

In this way we receive sharp upper bounds for $q \in (0, 2)$

$$(21) \qquad H_q(P) \leq \frac{1}{1-q}\ln\Big[\Big(\frac{1 + (N-1)a}{N}\Big)^q + (N-1)\Big(\frac{1-a}{N}\Big)^q\Big],$$

which for $q \to 1$ reduces to

$$(22) \quad H_1(P) \leq H_{12}^u := (1-N)\frac{1-a}{N}\ln\frac{1-a}{N} - \frac{1 + a(N-1)}{N}\ln\frac{1 + a(N-1)}{N} ,$$

with a given by (20).

To obtain analogous lower bound we find k such that $\ln(k-1) \leq H_2 \leq \ln k$ and compute $H_2(Q_{k-1,k}(a))$. Also this relation may be easily inverted providing $a = \sqrt{k(k-1)\exp(-H_2) + 1 - k}$. Thus we arrive at a lower bound for the Rényi entropy

$$(23) \qquad H_q(P) \geq \frac{1}{1-q}\ln\Big[(k-1)z^q + y^q\Big] ,$$

and in particular case, for the Shannon entropy

$$(24) \qquad H_1(P) \geq H_{12}^d := (1-k)z\ln z - y\ln y ,$$

with $z = (k + a - 1)/(k^2 - k)$ and $y = (1-a)/k$.

4.2. Bounds based on H_3 and N.
Let us now assume, we know the value of the Rényi entropy H_3. As in (19) we compute $H_3(Q_{1,N}(a))$, and invert it finding $a \in [0,1]$ as the largest (real) root of the polynomial

$$(25) \qquad W_u(a) = a^3 + a^2\frac{3}{N-2} - \frac{N^2\exp(-2H_3) - 1}{(N-1)(N-2)} = 0.$$

Then the upper bound valid for $q \in (0,3)$ is given by the same formula (21) with a given by the root of (25) instead of (20). For $q \to 1$ one obtains then the upper bound for the Shannon entropy

$$(26) \quad H_1(P) \leq H_{13}^u := (1-N)\frac{1-a}{N}\ln\frac{1-a}{N} - \frac{1 + a(N-1)}{N}\ln\frac{1 + a(N-1)}{N} ,$$

with a determined by (25).

To get the lower bound we look for k' such that $\ln(k'-1) \leq H_3 \leq \ln k'$. The relation $H_3(Q_{k'-1,k'}(a))$ may be inverted explicitly for $k' = 2$ providing $a = \sqrt{[4\exp(-2H_3)-1]/3}$. For $k' > 2$ a is given by the only root of the polynomial

$$(27) \qquad W_d(a) = a^3 + a^2 \frac{3(k'-1)}{2-k'} + \frac{(k'-1)^2}{2-k'}[1 - k'^2 \exp(-2H_3)] = 0.$$

in the interval $[0,1]$. The lower bound for the Rényi entropies has the same form as (23) and gives for the Shannon entropy

$$(28) \qquad\qquad H_1(P) \geq H_{13}^d := (1-k')z'\ln z' - y'\ln y'$$

with $z' = (k'+a-1)/(k'(k'-1))$ and $y' = (1-a)/k'$.

5. Combined extrapolation

In previous section we obtained two upper bounds for the Shannon entropy: (23) stemming from the Rényi entropy H_2, and (12) obtained from H_3. The latter is in general a worse one[3], but it allows for a linear extrapolation, which gives

$$(29) \qquad\qquad H_{up}(\vec{x}) := 2H_{12}^u(\vec{x}) - H_{13}^u(\vec{x}).$$

Our numerical results allow us to advance the following

Conjecture.

For any probability distribution $\vec{x}$ the bound

$$(30) \qquad\qquad H_1(\vec{x}) \leq H_{up}(\vec{x})$$

holds.

In the same way one may try to extrapolate lower bounds defining $H_d(\vec{x}) := 2H_{12}^d(\vec{x}) - H_{13}^d(\vec{x})$. For certain probability vectors this quantity may give a useful approximation for the Shannon entropy. Interestingly, a relation analogous to (30), $H_1(\vec{x}) \geq H_d(\vec{x})$ is not true: it is violated e.g. if $k \neq k'$.

Making use of the rigorous bound (12) and the conjecture (30) we may suggest to estimate the unknown value of the Shannon entropy by the mean value

$$(31) \qquad H'_*(\vec{x}) := \frac{1}{2}[H_{up}(\vec{x}) + H_{d23}(\vec{x})] = H_{12}^u(\vec{x}) + H_2(\vec{x}) - \frac{1}{2}[H_{13}^u(\vec{x}) + H_3(\vec{x})],$$

which is obtained by combined methods based as well on the lower as well as on the upper bounds. Observe that this explicitly computable quantity involves only Rényi entropies H_2 and H_3 and the dimension N. This estimation may be improved noting that the rigorous lower bounds (12) and (24) are not equivalent. Since for some distributions the latter bound gives better (higher) results, we may improve (31) writing

$$(32) \qquad\qquad H_*(\vec{x}) := \frac{1}{2}H_u(\vec{x}) + \frac{1}{2}\max\{H_{d23}(\vec{x}), H_{12}^d(\vec{x})\}.$$

If the value of the zero-entropy, H_0, is known, one may replace the upper bound $H_u(\vec{x})$ used above by the minimum, $\min\{H_{u0}(\vec{x}), H_{up}(\vec{x})\}$.

Figure 3 shows the bounds and the estimations described above for a randomly chosen probability distributions with $N = 15$ components. The overall quality of the proposed estimations for the Shannon entropy may be judged from Fig. 4, which shows the histogram of the deviations $\delta_1 = H_* - H_1$ and $\delta_2 = H_d - H_1$ for a sample

---

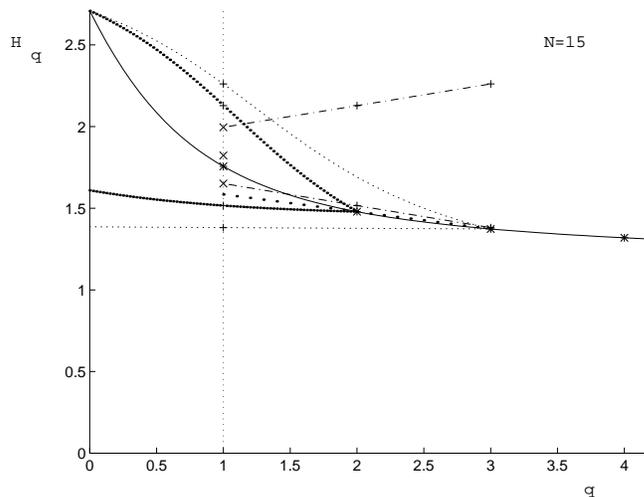[3]Knowing the function $H(q)$ at $q = 3$ we have less information on H_1, than knowing it at $q = 2$.

FIGURE 3. Rényi entropies H_q for a random probability distribution of size $N = 15$ (solid line). Dense dotted curves represent H_2 bounds: upper (21) and lower (23), while faint dotted curve denote analogous bounds based on H_3. Straight dashed–dotted lines show lower (12) and upper (29) extrapolations. The exact value of H_1 is denoted by $(*)$, while the estimation (32) by the middle $(\times)$.

of 10^4 probability vectors generated randomly according to the statistical (Fisher–Rao) measure on the $N - 1$ dimensional simplex [19]. This measure has a simple geometric interpretation: is suffices to consider a unit vector $\vec{t}$ distributed uniformly on the sphere S^{N-1}, and to define the probability vector $\vec{x} = \{t_1^2, ..., t_N^2\}$. Numerical results obtained in this way allow us to conclude that the proposed estimation (32) provides a useful all-purpose approximation of the Shannon entropy. Note that the precision of this approximation decreases with the length N of the probability vector.

To judge about possible application of the estimate H_* in the analysis of physical data, one should perform analogous numerical simulations with random vectors $\vec{x}$ generated according to a specific probability distribution adjusted to a given physical problem.

6. CONCLUDING REMARKS

In this work we considered the problem of finding the bounds and extrapolations for the Shannon (entropy), provided some of the Renyi entropies are known. In general, generalized entropies of integer order $q = 2$ and $q = 3$ are easiest to calculate, and they are sufficient to obtain bounds (10) and (12) for the Shannon entropy. The quality of the bound may be improved if the length N of the probability vector is known. Then an explicit extrapolation (32) allows us to estimate the actual value of the entropy H.
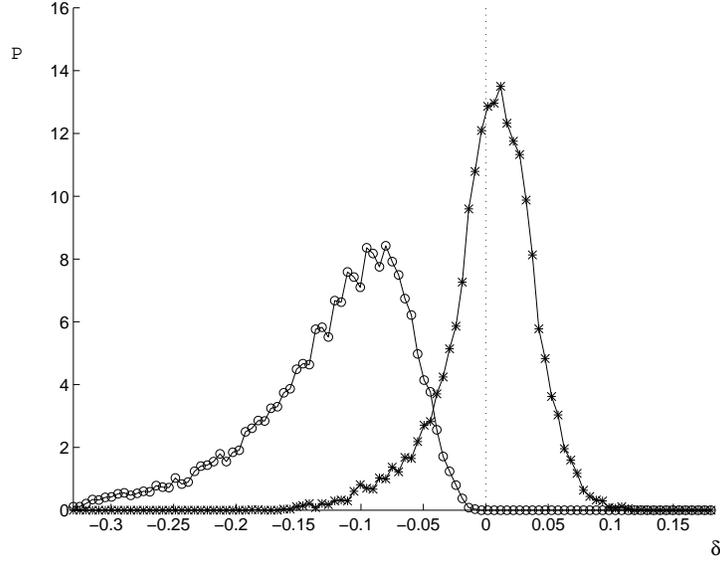
FIGURE 4. Histogram of deviations of the extrapolations form the
real value of the Shannon entropy H_1: ($*$) denotes the probability
density of error $\delta_1 = H_* - H_1$ of the estimation (32), while (o)
denotes results for the error $\delta_2 = H_d - H_1$ of the lower bound (12)
for a sample of 10^4 random probability vectors of size $N = 10$.

Note that the bounds and extrapolations discussed may be easily rewritten in
terms of a non-extensive entropy

$$(33) \qquad S_q(\vec{x}) := \frac{1}{1-q}\Big[\sum_{i=1}^{n} x_i^q - 1\Big]$$

used by Havrda and Charvat [20] and Daroczy [21], which became often used in
statistical physics after the seminal work of Tsallis [22]. In particular, the linear
entropy L is just the nonextensive entropy of order two, $L(\vec{x}) = S_2(\vec{x})$, and the
bounds between H_q and H_1 imply analogous relations between S_q and H_1. In fact
the plot presented in [13] shows bounds between von Neumann entropy and the
linear entropy and they follow directly from relation (18) proven in [15].

The issue of comparing the Rényi entropies H_0, H_1 and H_2 is closely related to
the problem of describing the degree of chaos of an analyzed classical dynamical
system by the topological entropy K_0, the Kolmogorov–Sinai (metric) entropy K_1
and the correlation entropy K_2. These dynamical entropies are defined as the rate
of the increase of the Rényi entropies in time [16], but since the length of the
probability vector is not finite the N–dependent bounds discussed in this work are
not applicable. The same concerns comparison of generalized fractal dimensions D_q
of fractal measures: the box–counting dimension D_0, the information dimension D_1
and the correlation dimension D_2, which also form a non–increasing function of the
Rényi parameter q [16], are defined by the limit $N \to \infty$.

The generalized entropies may also be used to characterize localization properties
of continuous probability distributions. For instance, any pure quantum state may

be represented in the phase space by the Husimi distribution. Its localization can be measured by the *Wehrl entropy* defined as the continuous (Boltzmann–Gibbs) entropy of the Husimi distribution [23]. In an analogous way one may define generalized Rényi–Wehrl entropies [24, 11], and above results may be used to obtain similar bounds for the Wehrl entropy.

7. Corrigendum – *February 17, 2005*

Eq. (8) implies that $(1 - q)H_q$ is a convex function of q. However, it does not imply that the Rényi entropy H_q is a convex function of q. For instance, the Rényi entropy for a $N = 20$ probability vector $P = \frac{1}{100}(43, 3, \ldots, 3)$ is not a convex function of q. I am deeply obliged to Christian Schaffner for drawing my attention to this fact.

Therefore, equations (10), (11), and (12) are not satisfied and conjecture (30) cannot hold. Furthermore, equations (13), (31), (32) may only be used to extrapolate the unknown value of the Shannon entropy, from the available data on H_2 and H_3.

On the other hand, we would like to mention that the lack of convexity of H_q in q does not influence the bounds between particular values of the Rényi entropy presented in sections 3 and 4 of the present paper.

References

1. C. Shannon, Mathematical theory of communication, *Bell System Tech. J.* **27**, 379 (1948)
2. A. I. Khinchin, *Mathematical foundations of information theory*, Dover, 1957,
3. A. Rényi, On measures of entropy and information, in. *Proc. Fourth. Berkeley Symp. Math. Stat. Prob. 1960*, Vol. I, p.547, (University of California Press, Berkeley, 1961).
4. J. N. Kapur, *Measures of Information and Their Applications* (John Wiley & Sons, New York, 1994).
5. C. Arndt. *Information Measures: Information and its Description in Science and Engineering* (Springer, Berlin 2001).
6. P. Horodecki and A. Ekert, Direct detection of quantum entanglement, *Phys. Rev. Lett.* **89**, 127902 (2002).
7. C. M. Alves, P. Horodecki, D. K. L. Oi, L. C. Kwek, and A. K. Ekert, Direct estimation of functionals of density operators by local operations and classical communication, preprint quant-ph/0304123
8. A. Białas and W. Czyż, Renyi Entropies in Multiparticle Production, *Acta Phys. Pol.* **B 31**, 2803 (2000).
9. A. Białas, W. Czyż and A. Ostruszka, Renyi entropies in particle cascades, *Acta Phys. Pol.* **B 34**, 69 (2003).
10. J. Pipek and I. Varga, Universal scheme for the spacial–localization properties of one–particle states in finite d–dimensional systems, *Phys. Rev.* **A 46**, 3148 (1992).
11. I. Varga and J. Pipek, On Rényi entropies characterizing the shape and the extension of the phase space representation of quantum wave functions in disordered systems, arXiv preprint cond-mat/0204041 (2002).

12. K. Życzkowski, I. Bengtsson, Relativity of pure states entanglement, *Ann. Phys. (N.Y.)* **295**, 115-135 (2002)

13. T.-C. Wei, K. Nemoto, P.M. Goldbart, P.G. Kwiat, W.J. Munro and F. Verstraete, Maximal entanglement versus entropy for mixed quantum states, *Phys. Rev.* **A 67**, 022110 (2003)

14. F. Topsøe, Some inequalities for information divergence and Related measures of discrimination *IEEE Trans. Inform. Theory* **46**, 1602 (2000).

15. P. Harremoës and F. Topsøe, Inequalities between Entropy and Index of Coincidence derived from Information Diagrams, *IEEE Trans. Inform. Theory* **47**, 2944-2960 (2001)

16. C. Beck and F. Schlögl, *Thermodynamics of chaotic systems*, (Cambridge University Press, Cambridge, 1993).

17. D. W. Berry and B. C. Sanders, Bounds on entropy, preprint arXiv quant-ph/0305059, (2003)

18. F. Topsøe, Entropy and Index of Coincidence, lower bounds, preprint, Copenhagen, 2003

19. R. A. Fisher, Theory of Statistical Estimation *Proc. Cambridge Philos. Soc.* **22**, 700 (1925).

20. J. Havrda and F. Charvat, Quantification methods of classification Processes: Concept of structural α–Entropy, *Kybernetica* **3**, 30 (1967)

21. Z. Daroczy, *Inf. Control* **16**, 36 (1970)

22. C. Tsallis, Possible generalization of Boltzmann–Gibs statistics, *J. Stat. Phys.* **52**, 479 (1988)

23. A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50**, 221 (1978)

24. S. Gnutzmann and K. Życzkowski, Renyi-Wehrl entropies as measures of localization in phase space, *J. Phys.* **A 34**, 10123-10139 (2001).

Instytut Fizyki im. Smoluchowskiego, Uniwersytet Jagielloński, ul. Reymonta 4, 30-059 Kraków, Poland, and, Centrum Fizyki Teoretycznej, Polska Akademia Nauk, Al. Lotników 32/44, 02-668 Warszawa, Poland

E-mail address: `karol@tatry.if.uj.edu.pl`