Peter Harremoës

Joint Range of Rényi entropies

Kybernetika, Vol. 45 (2009), No. 6, 901--911

Persistent URL: http://dml.cz/dmlcz/140023

KYBERNETIKA — VOLUME 45 (2009), NUMBER 6, PAGES 901 – 911

JOINT RANGE OF RÉNYI ENTROPIES

Peter Harremoës

The exact range of the joined values of several Rényi entropies is determined. The method is based on topology with special emphasis on the orientation of the objects studied. Like in the case when only two orders of the Rényi entropies are studied, one can parametrize the boundary of the range. An explicit formula for a tight upper or lower bound for one order of entropy in terms of another order of entropy cannot be given.

Keywords: generalized Vandermonde determinant, orientation, Rényi entropies, Shannon entropy

AMS Subject Classification: 94A17, 62B10

1. INTRODUCTION

Let $P = (p_1, p_2, \ldots, p_n)$ be a probability vector. For $\alpha \in \mathbb{R} \backslash \{0, 1\}$ the Rényi entropy of P of order α is defined as a number in $[0; \infty]$ given by the equation

$$H_\alpha(P) = \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right).$$

This definition is extended by continuity so that

$$\begin{aligned} H_{-\infty}(P) &= -\log \min_i p_i \ ; \\ H_0(P) &= \log(\text{number of } p_i \neq 0); \\ H_1(P) &= -\sum_i p_i \log p_i \ ; \\ H_\infty(P) &= -\log \max_i p_i \ . \end{aligned}$$

The Rényi entropy H_0 is essentially the Hartley entropy, and was one among other sources of inspiration to Shannon's information theory. The Rényi entropy of order ∞ is also called the min-entropy and essentially related to the "probability of error". The Rényi entropy H_2 is related to index of coincidence and other quantities used for special purposes in crypto analysis, physics etc. [2, 8].

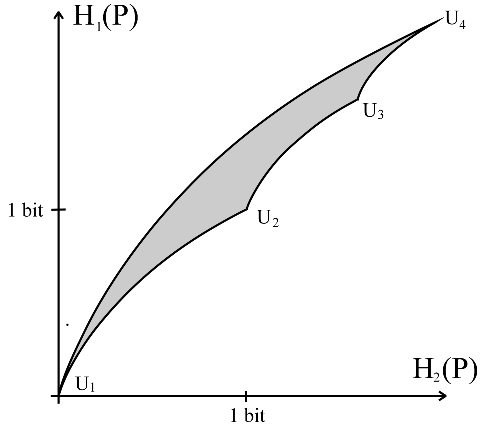

Fig. 1. Range $P \to (H_2(P), H_1(P))$ for a four element set. The uniform distributions are mapped into the diagonal.

For all α the Rényi entropy H_α has the nice property of being additive on product measures. In noiseless source coding for finite systems one wants to avoid very long code words. For such systems the Rényi entropy of some order $\alpha < 1$ (depending on the memory of the system) determines how much the source can be compressed. Rényi entropies are also related to general cut-off rates and "guess-work moments" [1, 4].

The relation between H_0 and H_1 is given by the simple inequality

$$H_1(P) \leq H_0(P).$$

This is a special case of the general result that

$$\alpha \to H_\alpha(P)$$

is a strictly decreasing function except for uniform distributions where it is constant, which follows from a simple application of Jensen's Inequality. The relation between H_1 and H_∞ has been determined independently in various articles [3, 5, 6, 11, 15]. The relation between the Shannon entropy and H_2 has been studied in [7] and in more detail in [8]. The result is illustrated in Figure 1 and by the following theorem.

Theorem 1. The lower bound on $H_1(P)$ given $H_2(P)$ is attained by a mixture of uniform distributions on k and $k+1$ points where k is determined by the condition $\log k \leq H_2(P) < \log(k+1)$. The upper bound on $H_1(P)$ is attained by a mixture of the uniform distribution on n points and a uniform distribution on a singleton.

We shall generalize this result and determine the joint range of several Rényi entropies. In general the boundary can be parametrized, but upper and lower bounds of entropy of one order in terms of entropy of another order cannot be given by explicit formulas. The reason is that the inverse of the function $s \to$

$H_\alpha(sU_k + (1-s)U_{k+1})$, where U_k and U_{k+1} are uniform distributions, is in general not an elementary function.

Recently the joint range of Rényi entropies has been used to determine the relative Bahadur efficiency of various power divergence statistics [9, 10]. In these papers the joint range of H_1 and H_α was used with a reference to [8] where the general result for comparison of two Rényi entropies was mentioned without proof. In some cases in physics, joint values of $H_2(P)$ and $H_3(P)$ can be measured or computed and one is interested in bounds on H_1 [16]. In order to get bounds on H_1 one is interested in the exact range of the mapping

$$\Psi : P \to (H_3(P), H_2(P), H_1(P)).$$

The methods developed in [8] will be refined in order to be able to describe the joint range of in principle any number of Rényi entropies of positive order. We restrict our attention to non-negative orders because these are the most important for applications and because Rényi entropies of negative orders are not continuous near uniform distributions. Although the method is very general, we shall only go into details in the cases where two or three Rényi entropies are compared. The main result is that the range has a boundary, which can be parametrized by certain mixtures of uniform distributions.

2. REDUCTION TO MIXTURES OF UNIFORM DISTRIBUTIONS

A probability vector P on a set with n elements can be parametrized by its point probabilities as $(p_1, p_2, \ldots, p_n)$ where $p_j \geq 0$ and

$$\sum_{j=1}^{n} p_j = 1.$$

The Rényi entropies are symmetric in their entries. Therefore we may restrict our attention to probability vectors with decreasing entries, i.e. $p_1 \geq p_2 \geq \cdots \geq p_n \geq 0$. Here we shall assume that n is fixed so that that $H_0(P) \leq \log n$. In order to study the range of $P \to (H_{\alpha_1}(P), H_{\alpha_2}(P), \cdots, H_{\alpha_m}(P))$ we first consider the related map

$$P \to \begin{pmatrix} \frac{1}{1-\alpha_1}\log\left(\sum p_j^{\alpha_1}\right) \\ \frac{1}{1-\alpha_2}\log\left(\sum p_j^{\alpha_2}\right) \\ \vdots \\ \frac{1}{1-\alpha_m}\log\left(\sum p_j^{\alpha_m}\right) \\ \sum p_j \end{pmatrix}. \tag{1}$$

We will assume that the orders are chosen in decreasing order like $\alpha_1 > \alpha_2 > \ldots >$

α_n. The matrix of partial derivatives with respect to $p_1, p_2, \ldots, p_n$ is

$$\begin{pmatrix} \frac{\alpha_1}{1-\alpha_1}\frac{p_1^{\alpha_1-1}}{\sum p_j^{\alpha_1}} & \frac{\alpha_1}{1-\alpha_1}\frac{p_2^{\alpha_1-1}}{\sum p_j^{\alpha_1}} & \cdots & \frac{\alpha_1}{1-\alpha_1}\frac{p_{n-1}^{\alpha_1-1}}{\sum p_j^{\alpha_1}} & \frac{\alpha_1}{1-\alpha_1}\frac{p_n^{\alpha_1-1}}{\sum p_j^{\alpha_1}} \\ \frac{\alpha_2}{1-\alpha_2}\frac{p_1^{\alpha_2-1}}{\sum p_j^{\alpha_2}} & \frac{\alpha_2}{1-\alpha_2}\frac{p_2^{\alpha_2-1}}{\sum p_j^{\alpha_2}} & \cdots & \frac{\alpha_2}{1-\alpha_2}\frac{p_{n-1}^{\alpha_2-1}}{\sum p_j^{\alpha_2}} & \frac{\alpha_2}{1-\alpha_2}\frac{p_n^{\alpha_2-1}}{\sum p_j^{\alpha_2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\alpha_m}{1-\alpha_m}\frac{p_1^{\alpha_m-1}}{\sum p_j^{\alpha_m}} & \frac{\alpha_m}{1-\alpha_m}\frac{p_2^{\alpha_m-1}}{\sum p_j^{\alpha_m}} & \cdots & \frac{\alpha_m}{1-\alpha_m}\frac{p_{n-1}^{\alpha_m-1}}{\sum p_j^{\alpha_m}} & \frac{\alpha_m}{1-\alpha_m}\frac{p_n^{\alpha_m-1}}{\sum p_j^{\alpha_m}} \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}.$$

If this matrix has rank $m+1$ in a neighborhood of a point $P = (p_1, p_2, \ldots, p_n)$, then the map (1) restricted to such a neighborhood is open, i. e. it maps open sets into open sets and a neighborhood of P is mapped into a neighborhood of the image. This follows from the *inverse map theorem* [13], pp. 221–223 and is often termed the *open map theorem*[1].

Next we show that if P has $m+1$ different point probabilities then P is mapped into an interior point in the range. Therefore, assume that P has $m+1$ different point probabilities, i. e. $n = m+1$. Then

$$\begin{vmatrix} \frac{\alpha_1}{1-\alpha_1}\frac{p_1^{\alpha_1-1}}{\sum p_j^{\alpha_1}} & \frac{\alpha_1}{1-\alpha_1}\frac{p_2^{\alpha_1-1}}{\sum p_j^{\alpha_1}} & \cdots & \frac{\alpha_1}{1-\alpha_1}\frac{p_m^{\alpha_1-1}}{\sum p_j^{\alpha_1}} & \frac{\alpha_1}{1-\alpha_1}\frac{p_{m+1}^{\alpha_1-1}}{\sum p_j^{\alpha_1}} \\ \frac{\alpha_2}{1-\alpha_2}\frac{p_1^{\alpha_2-1}}{\sum p_j^{\alpha_2}} & \frac{\alpha_2}{1-\alpha_2}\frac{p_2^{\alpha_2-1}}{\sum p_j^{\alpha_2}} & \cdots & \frac{\alpha_2}{1-\alpha_2}\frac{p_m^{\alpha_2-1}}{\sum p_j^{\alpha_2}} & \frac{\alpha_2}{1-\alpha_2}\frac{p_{m+1}^{\alpha_2-1}}{\sum p_j^{\alpha_2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\alpha_m}{1-\alpha_m}\frac{p_1^{\alpha_m-1}}{\sum p_j^{\alpha_m}} & \frac{\alpha_m}{1-\alpha_m}\frac{p_2^{\alpha_m-1}}{\sum p_j^{\alpha_m}} & \cdots & \frac{\alpha_m}{1-\alpha_m}\frac{p_m^{\alpha_m-1}}{\sum p_j^{\alpha_m}} & \frac{\alpha_m}{1-\alpha_m}\frac{p_{m+1}^{\alpha_m-1}}{\sum p_j^{\alpha_m}} \\ 1 & 1 & \cdots & 1 & 1 \end{vmatrix} \tag{2}$$

$$= \left(\prod_{i=1}^{m}\frac{\alpha_i}{1-\alpha_i}\cdot\prod_{i=1}^{m}\frac{1}{\sum_j p_j^{\alpha_1}}\right)\begin{vmatrix} p_1^{\alpha_1-1} & p_2^{\alpha_1-1} & \cdots & p_m^{\alpha_1-1} & p_{m+1}^{\alpha_1-1} \\ p_1^{\alpha_2-1} & p_2^{\alpha_2-1} & \cdots & p_m^{\alpha_2-1} & p_{m+1}^{\alpha_2-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_1^{\alpha_m-1} & p_2^{\alpha_m-1} & \cdots & p_m^{\alpha_m-1} & p_{m+1}^{\alpha_m-1} \\ 1 & 1 & \cdots & 1 & 1 \end{vmatrix}$$

Note that the last row can be written as $\begin{pmatrix} p_1^{\alpha-1} & p_2^{\alpha-1} & \cdots & p_m^{\alpha-1} & p_{m+1}^{\alpha-1} \end{pmatrix}$ with $\alpha = 1$. The last determinant is an *exponential Vandermonde determinant*. By a result of Robbin and Salamon an exponential Vendermonde determinant is non-negative [12]. It is positive if and only if $p_1 > p_2 > \ldots > p_n$.

We see that if $1 > \alpha_m > \cdots > \alpha_m > 0$ then the determinant (2) is positive. It is easy to check that this is also the case with the relaxed condition $\alpha_m > \cdots > \alpha_m > 0$.

The extreme points in the set of ordered probability vectors are the uniform distributions. Let U_k denote the uniform distribution $\left(\frac{1}{k}, \frac{1}{k}, \cdots, \frac{1}{k}, 0, 0, \cdots, 0\right)$. Let $k_1, k_2, \ldots, k_\ell$ be a sequence of different numbers in $\{1, 2, \cdots, n\}$. Then the simplex

[1] There are several different theorems called "The Open Map Theorem". This is just one of them.

formed by convex combinations of $U_{k_1}, U_{k_2}, \ldots, U_{k_\ell}$ will be denoted $\Delta_{k_1,k_2,\cdots,k_\ell}$ and be given an orientation according to the sequence $U_{k_1}, U_{k_2}, \ldots, U_{k_\ell}$. Observe that if $k_1 < k_2 < \ldots < k_{m+1}$ then the mapping $\Delta_{k_1,k_2,\cdots,k_{m+1}} \to \mathbb{R}^m$ defined by

$$P \to \begin{pmatrix} \frac{1}{1-\alpha_1} \log\left(\sum p_j^{\alpha_1}\right) \\ \frac{1}{1-\alpha_2} \log\left(\sum p_j^{\alpha_2}\right) \\ \vdots \\ \frac{1}{1-\alpha_m} \log\left(\sum p_j^{\alpha_m}\right) \end{pmatrix}$$

has positive orientation if $\alpha_1 > \alpha_2 > \cdots > \alpha_m > 0$.

3. JOINT RANGE OF TWO RÉNYI ENTROPIES

First we consider distributions on a set with n elements. We determine the joint range of H_{α_1} and H_{α_2} where we assume that $\alpha_1 > \alpha_2 > 0$. First we shall also assume that $\alpha_1, \alpha_2 \in \left]0;\infty\right[\setminus \{1\}$. Let Φ denote the map

$$P \to \begin{pmatrix} H_{\alpha_1}(P) \\ H_{\alpha_2}(P) \end{pmatrix}.$$

Assume that $k_1 < k_2 < k_3$. Then $\Phi\left(U_{k_j}\right)$ lies on the diagonal $\{(x,x) : x \geq 0\}$, and these points are ordered,

$$H_\alpha(U_{k_1}) < H_\alpha(U_{k_2}) < H_\alpha(U_{k_3})$$

where $\alpha = \alpha_1$ or $\alpha = \alpha_2$. We know that $H_{a_1}(P) \leq H_{\alpha_2}(P)$ with equality if and only if P is a uniform distribution. We know that Φ maps interior points of Δ_{k_1,k_2,k_3} into interior points of the range of Φ so boundary points of the range of Φ must have preimages that are boundary points of Δ_{k_1,k_2,k_3}. We follow the conventions from homology theory [14] and calculate boundary with orientation. The boundary of $\Phi(\Delta_{k_1,k_2,k_3})$ is

$$\begin{aligned} \partial\Phi(\Delta_{k_1,k_2,k_3}) &= \Phi\partial(\Delta_{k_1,k_2,k_3}) \\ &= \Phi(\Delta_{k_2,k_3} - \Delta_{k_1,k_3} + \Delta_{k_1,k_2}) \\ &= \Phi(\Delta_{k_1,k_2} + \Delta_{k_2,k_3} + \Delta_{k_3,k_1}), \end{aligned}$$

which is just another way to write the closed curve from U_{k_1} to U_{k_2} to U_{k_3} and back to U_{k_1}. Therefore any point on the boundary of the range of Φ must be the image of a mixture of two uniform distributions.

Assume that $k_1 < k_2 < k_3 < k_4$. Then the simplices Δ_{k_1,k_2,k_3} and Δ_{k_1,k_3,k_4} are both positively oriented and

$$\begin{aligned}\partial\Phi\left(\Delta_{k_1,k_2,k_3}+\Delta_{k_1,k_3,k_4}\right) &= \Phi\partial\left(\Delta_{k_1,k_2,k_3}+\Delta_{k_1,k_3,k_4}\right)\\ &= \Phi\begin{pmatrix}\partial\Delta_{k_1,k_2,k_3}\\ +\partial\Delta_{k_1,k_3,k_4}\end{pmatrix}\\ &= \Phi\begin{pmatrix}\Delta_{k_2,k_3}-\Delta_{k_1,k_3}+\Delta_{k_1,k_2}\\ +\Delta_{k_3,k_4}-\Delta_{k_1,k_4}+\Delta_{k_3,k_4}\end{pmatrix}\\ &= \Phi\begin{pmatrix}\Delta_{k_2,k_3}-\Delta_{k_1,k_3}+\Delta_{k_1,k_2}\\ +\Delta_{k_3,k_4}-\Delta_{k_1,k_4}+\Delta_{k_1,k_3}\end{pmatrix}\\ &= \Phi\left(\Delta_{k_1,k_2}+\Delta_{k_2,k_3}+\Delta_{k_3,k_4}+\Delta_{k_4,k_1}\right).\end{aligned}$$

We see that $\Phi\left(\Delta_{k_1,k_3}\right)$ does not contribute to this boundary. Similarly $\Phi\left(\Delta_{k_2,k_4}\right)$ does not contribute to the boundary. We may formulate this result as $\Delta_{a,b}$ does not contribute to the range if it is a diagonal in a quadruple. The non-diagonal simplices are $\Delta_{1,2},\Delta_{2,3},\cdots,\Delta_{n-1,n}$ and $\Delta_{n,1}$. These form a closed curve

$$\Delta_{1,2}+\Delta_{2,3}+\cdots+\Delta_{n-1,n}+\Delta_{n,1}$$

and the boundary is the image of this curve, i. e.

$$\Phi\left(\Delta_{1,2}+\Delta_{2,3}+\cdots+\Delta_{n-1,n}+\Delta_{n,1}\right).$$

This result easily extends to the cases where one or more of the orders equal 1 or ∞. The lower bound does not depend on n so we get the following theorem.

Theorem 2. Assume $\alpha_1 > \alpha_2 > 0$. Then the lower bound on $H_{\alpha_2}(P)$ given $H_{\alpha_1}(P)$ is attained by a mixture of uniform distributions on k and $k+1$ points where k is determined by the condition $\log k \leq H_{\alpha_1}(P) < \log(k+1)$.

For distributions on sets with n elements we also get a tight upper bound, but if we have no restriction on n the situation is a little more complicated.

Theorem 3. Assume $\alpha_1 > \alpha_2 > 0$. If P is a distribution on a set with n elements and $H_{\alpha_1}(P)$ is fixed then a upper bound on H_{α_2} is attained for a mixture of the uniform distributions U_1 and U_n. If no restriction on n is given and if $H_{\alpha_1}(P) > 0$ is fixed then a tight upper bound on $H_{\alpha_2}(P)$ is given by

$$H_{a_2}(P) < \begin{cases}\infty, & \text{if } \alpha_1 \leq 1;\\ \frac{\alpha_2}{\alpha_2-1}\frac{\alpha_1-1}{\alpha_1}H_{\alpha_1}(P), & \text{if } \alpha_1 > 1.\end{cases}$$

Proof. If we have no restriction on n then the range is

$$\bigoplus_{n=2}^{\infty}\Phi\left(\Delta_{1,n,n+1}\right)$$

so we just have to determine the asymptotics of $\Phi(\Delta_{1,n})$. The curve $\Delta_{1,n}$ has the parametrization $P_t = \left(\frac{t}{n}, \frac{t}{n}, \cdots, \frac{t}{n}, \frac{t}{n} + 1 - t\right)$, $t \in [0;1]$. Therefore the curve $\Phi(\Delta_{n,1})$ has the parametrization

$$\begin{pmatrix} \frac{1}{1-\alpha_1} \log\left((n-1)\left(\frac{t}{n}\right)^{\alpha_1} + \left(\frac{t}{n} + 1 - t\right)^{\alpha_1}\right) \\ \frac{1}{1-\alpha_2} \log\left((n-1)\left(\frac{t}{n}\right)^{\alpha_2} + \left(\frac{t}{n} + 1 - t\right)^{\alpha_2}\right) \end{pmatrix}.$$

We have to study the asymptotics of this curve for n tending to infinity. There are several cases and they need separate analysis.

Case $\alpha_2 > 1$. We also have $\alpha_1 > 1$ so for a fixed value of t we get

$$\begin{pmatrix} \frac{1}{1-\alpha_1} \log\left((n-1)\left(\frac{t}{n}\right)^{\alpha_1} + \left(\frac{t}{n} + 1 - t\right)^{\alpha_1}\right) \\ \frac{1}{1-\alpha_2} \log\left((n-1)\left(\frac{t}{n}\right)^{\alpha_2} + \left(\frac{t}{n} + 1 - t\right)^{\alpha_2}\right) \end{pmatrix} \to \begin{pmatrix} \frac{\alpha_1}{1-\alpha_1} \log(1-t) \\ \frac{\alpha_2}{1-\alpha_2} \log(1-t) \end{pmatrix}$$

for n tending to infinity. Hence the straight line with slope $\frac{\alpha_2}{\alpha_2-1}\frac{\alpha_1-1}{\alpha_1}$ is the boundary of the range.

Case $\alpha_1 \geq 1$ and $\alpha_2 \leq 1$. First we assume that $\alpha_1 < 1$. For a fixed value of the parameter t the Rényi entropy H_{a_2} tends to a constant as above but H_{α_1} tends to infinity. For a fixed value of $H_{\alpha_1}(P) > 0$ the lower bound $H_{\alpha_2}(P) > 0$ is tight. This bound is also tight for $\alpha_1 = 1$ and can be obtained by letting α_1 tend to 1 from above or below.

Case $0 < \alpha_1 \leq 1$. First assume that $\alpha_2 < 1$. If $t = n^{1-1/\alpha_2}$ then

$$\begin{pmatrix} \frac{1}{1-\alpha_1} \log\left((n-1)\left(\frac{t}{n}\right)^{\alpha_1} + \left(\frac{t}{n} + 1 - t\right)^{\alpha_1}\right) \\ \frac{1}{1-\alpha_2} \log\left((n-1)\left(\frac{t}{n}\right)^{\alpha_2} + \left(\frac{t}{n} + 1 - t\right)^{\alpha_2}\right) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{1-\alpha_1} \log\left(n^{-\frac{\alpha_1}{\alpha_2}} \cdot \frac{n-1}{n} + \left(n^{-1/\alpha_2} + 1 - n^{1-1/\alpha_2}\right)^{\alpha_1}\right) \\ \frac{1}{1-\alpha_2} \log\left(\frac{n-1}{n} + \left(n^{-1/\alpha_2} + 1 - n^{1-1/\alpha_2}\right)^{\alpha_2}\right) \end{pmatrix}.$$

We see that the second coordinate tends to $\frac{1}{1-\alpha_2}\log 2$, while the first coordinate tends to 0. Therefore for a fixed value of $H_{\alpha_1}(P) > 0$ the upper bound $H_{\alpha_2}(P) > 0$ is tight. Tightness of this bound also holds for $\alpha_2 = 1$, which can be seen by letting α_2 tend to 1 from above or below.

□

4. JOINT RANGE OF THREE RÉNYI ENTROPIES

Determining the range of three Rényi entropies is done in a similar way as in the previous section. We consider the map Ψ given by

$$P \to \begin{pmatrix} H_{\alpha_1}(P) \\ H_{\alpha_2}(P) \\ H_{\alpha_3}(P) \end{pmatrix}$$

where $\alpha_1 > \alpha_2 > \alpha_3 > 0$. First we consider the situation where the domain consist of distributions on n points. The boundary points of Ψ must be images of mixtures of three uniform distributions, i. e. in the range of Ψ restricted to some simplex $\Delta_{k,\ell,m}$. As we shall see most of these simplices do not contribute to the boundary but are "interior walls" in the range.

If $1 < k < \ell < m < n$ then the restriction of Ψ to the simplex $\Delta_{1,k,\ell,m}$ or to the simplex $\Delta_{k,\ell,m,n}$ conserves orientation. Therefore

$$\begin{aligned}
\partial\Psi\left(\Delta_{1,k,\ell,m} + \Delta_{k,\ell,m,n}\right) &= \partial\Psi\left(\partial\Delta_{1,k,\ell,m} + \partial\Delta_{k,\ell,m,n}\right) \\
&= \partial\Psi\begin{pmatrix} \Delta_{k,\ell,m} - \Delta_{1,\ell,m} + \Delta_{1,k,m} - \Delta_{1,k,\ell} \\ +\Delta_{\ell,m,n} - \Delta_{k,m,n} + \Delta_{k,\ell,n} - \Delta_{k,\ell,m} \end{pmatrix} \\
&= \partial\Psi\begin{pmatrix} -\Delta_{1,\ell,m} + \Delta_{1,k,m} - \Delta_{1,k,\ell} \\ +\Delta_{\ell,m,n} - \Delta_{k,m,n} + \Delta_{k,\ell,n} \end{pmatrix}.
\end{aligned}$$

We see that $\Delta_{k,\ell,m}$ gives no contribution to the boundary and therefore only simplices $\Delta_{k,\ell,m}$ with either $k = 1$ or $m = n$ give a contributions to the boundary.

If $1 < k < \ell < m < n$ then the restriction of Ψ to the simplices $\Delta_{1,k,\ell,m}$ or to $\Delta_{1,k,m,n}$ conserves orientation. Therefore

$$\begin{aligned}
\partial\Psi\left(\Delta_{1,k,\ell,m} + \Delta_{1,k,m,n}\right) &= \partial\Psi\left(\partial\Delta_{1,k,\ell,m} + \partial\Delta_{1,k,m,n}\right) \\
&= \partial\Psi\begin{pmatrix} \Delta_{k,\ell,m} - \Delta_{1,\ell,m} + \Delta_{1,k,m} - \Delta_{1,k,\ell} \\ +\Delta_{k,m,n} - \Delta_{1,m,n} + \Delta_{1,k,n} - \Delta_{1,k,m} \end{pmatrix} \\
&= \partial\Psi\begin{pmatrix} \Delta_{k,\ell,m} - \Delta_{1,\ell,m} - \Delta_{1,k,\ell} \\ +\Delta_{k,m,n} - \Delta_{1,m,n} + \Delta_{1,k,n} \end{pmatrix}.
\end{aligned}$$

We see that the simplex $\Delta_{1,k,m}$ gives no contribution to the boundary of the range of Ψ if $m < n$ and if $k < m - 1$.

If $1 < k < \ell < m < n$ then the restriction of Ψ to the simplices $\Delta_{1,k,m,n}$ or to $\Delta_{k,\ell,m,n}$ conserves orientation. Therefore

$$\begin{aligned}
\partial\Psi\left(\Delta_{1,k,m,n} + \Delta_{k,\ell,m,n}\right) &= \partial\Psi\left(\partial\Delta_{1,k,m,n} + \partial\Delta_{k,\ell,m,n}\right) \\
&= \partial\Psi\begin{pmatrix} \Delta_{k,m,n} - \Delta_{1,m,n} + \Delta_{1,k,n} - \Delta_{1,k,m} \\ +\Delta_{\ell,m,n} - \Delta_{k,m,n} + \Delta_{k,\ell,n} - \Delta_{k,\ell,m} \end{pmatrix} \\
&= \partial\Psi\begin{pmatrix} -\Delta_{1,m,n} + \Delta_{1,k,n} - \Delta_{1,k,m} \\ +\Delta_{\ell,m,n} + \Delta_{k,\ell,n} - \Delta_{k,\ell,m} \end{pmatrix}.
\end{aligned}$$

We see that the simplex $\Delta_{k,m,n}$ gives no contribution to the boundary of the range of Ψ if $k > 1$ and if $k < m - 1$.

If $1 < k < \ell < m < n$ then the restriction of Ψ to the simplices $\Delta_{1,k,\ell,n}$ or to

$\Delta_{1,\ell,m,n}$ conserves orientation. Therefore

$$\begin{aligned}\partial\Psi\left(\Delta_{1,k,\ell,n}+\Delta_{1,\ell,m,n}\right) &= \partial\Psi\left(\partial\Delta_{1,k,\ell,n}+\partial\Delta_{1,\ell,m,n}\right)\\ &= \partial\Psi\left(\begin{array}{c}\Delta_{k,\ell,n}-\Delta_{1,\ell,n}+\Delta_{1,k,n}-\Delta_{1,k,\ell}\\ +\Delta_{\ell,m,n}-\Delta_{1,m,n}+\Delta_{1,\ell,n}-\Delta_{1,\ell,m}\end{array}\right)\\ &= \partial\Psi\left(\begin{array}{c}\Delta_{k,\ell,n}+\Delta_{1,k,n}-\Delta_{1,k,\ell}\\ +\Delta_{\ell,m,n}-\Delta_{1,m,n}-\Delta_{1,\ell,m}\end{array}\right).\end{aligned}$$

We see that the simplex $\Delta_{1,\ell,n}$ gives no contribution to the boundary of the range of Ψ except if $\ell=2$ or $\ell=n-1$.

Thus the boundary of the range consist of images of the simplices $\Delta_{1,m,m+1}$ and of the form $\Delta_{m-1,m,n}$, where $m=2,3,\cdots,n-1$. Here we notice that

$$\begin{aligned}\partial\left(\bigoplus_{m=2}^{n-1}\Delta_{m-1,m,n}+\bigoplus_{m=2}^{n-1}\Delta_{1,m+1,m}\right) &= \bigoplus_{m=2}^{n-1}\partial\Delta_{m-1,m,n}+\bigoplus_{m=2}^{n-1}\partial\Delta_{1,m+1,m}\\ &= \bigoplus_{m=2}^{n-1}\left(\Delta_{m,n}-\Delta_{m-1,n}+\Delta_{m-1,m}\right)\\ &+\bigoplus_{m=2}^{n-1}\left(\Delta_{m+1,m}-\Delta_{m+1,1}+\Delta_{m,1}\right)\\ &= 0,\end{aligned}$$

so that

$$\bigoplus_{m=2}^{n-1}\Delta_{m-1,m,n}+\bigoplus_{m=2}^{n-1}\Delta_{1,m+1,m}$$

is a closed surface, and that the range of Ψ has the image of this surface as boundary.

It is possible to describe the situation in more detail. Let Φ denote the map

$$P\to\left(\begin{array}{c}H_{\alpha_1}(P)\\ H_{\alpha_2}(P)\end{array}\right).$$

Then Φ restricted to $\bigoplus_{m=2}^{m}\Delta_{1,m,m+1}$ is a homeomorphism. If

$$\Phi(P)=\left(\begin{array}{c}a\\ b\end{array}\right)$$

then there exists a unique m and unique weights $x,y,z\geq 0$ that sum up to 1 such that $P=x\cdot U_1+y\cdot U_m+z\cdot U_{m+1}$. For any distribution Q with $\Phi(Q)=\left(\begin{array}{c}a\\ b\end{array}\right)$ we have $H_{\alpha_3}(Q)\leq H_{\alpha_3}(P)$. Thus,

$$\bigoplus_{m=2}^{n-1}\Delta_{1,m,m+1}$$

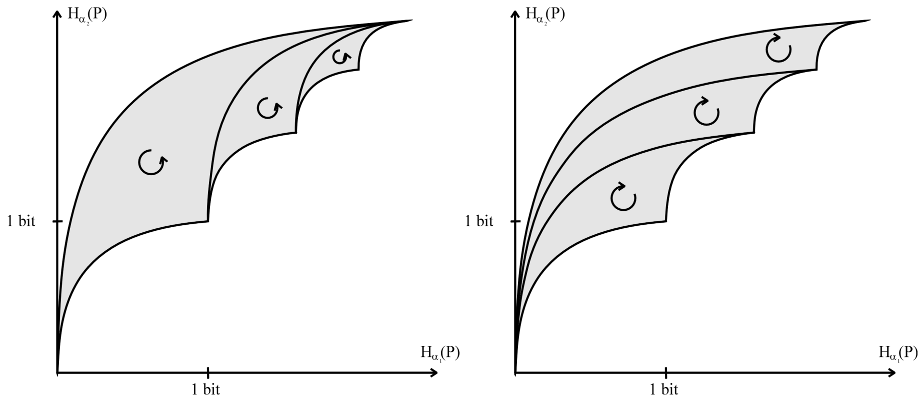

Fig. 2. The left diagram depictures the image of $\Delta_{125}+\Delta_{235}+\Delta_{345}$. The right diagram is the image of $\Delta_{132}+\Delta_{143}+\Delta_{154}$. Orientations are indicated. If an extra dimension H_{α_3} is added these two surfaces form the boundary of the image of Δ_{12345}. In these plots $\alpha_1 = 5$ and $\alpha_2 = 1/2$ were used.

gives a tight lower bound on H_{α_3} in terms of H_{α_1} and H_{α_2}. We notice that this lower bound does not depend on n. Similarly, the upper bound on H_{α_3} for fixed H_{α_1} and H_{α_2} is determined by the surface

$$\bigoplus_{m=2}^{n-1} \Delta_{m-1,m,n}$$

and just as in the case of two Rényi entropies the upper bound does depend on n.

5. DISCUSSION

The result can be seen as a generalization of the result in [8]. The essential step in the whole construction is the positivity of the exponential Vandermonde determinant. Therefore the construction can be iterated so that one in principle can determine the boundary of the range of any number of Rényi entropies of positive order.

ACKNOWLEDGEMENT

I thank Karol Zyczkowski for useful discussions. His paper [16] was an important inspiration for this article. He, Flemming Topsøe, and Christian Schaffner have also contributed with several important remarks to this paper. The author was supported by grants from Villum Kann Rasmussen Foundation, The Banach Center, INTAS (project 00-738), Danish Natural Research Council, and the European Pascal Network of Excellence.

(Received April 16, 2009.)

REFERENCES

[1] E. Arikan: An inequality on guessing and its application to sequential decoding. IEEE Trans. Inform. Theory *42* (1996), 1, 99–105.

[2] C. Arndt: Information Measures. Springer, Berlin 2001.

[3] M. Ben-Bassat: f-entropies, probability of error, and feature selection. Inform. and Control *39* (1978), 227–242.

[4] I. Csiszár: Generalized cutoff rates and Rényi information measures. IEEE Trans. Inform. Theory *41* (1995), 1, 26–34.

[5] M. Feder and N. Merhav: Relations between entropy and error probability. IEEE Trans. Inform. Theory *40* (1994), 259–266.

[6] J. D. Golić: On the relationship between the information measures and the Bayes probability of error. IEEE Trans. Inform. Theory *35* (1987), 5, 681–690.

[7] A. György and T. Linder: Optimal entropy-constrained scalar quantization of a uniform source. IEEE Trans. Inform. Theory *46* (2000), 7, 2704–2711.

[8] P. Harremoës and F. Topsøe: Inequalities between entropy and index of coincidence derived from information diagrams. IEEE Trans. Inform. Theory *47* (2001), 7, 2944–2960.

[9] P. Harremoës and I. Vajda: Efficiency of entropy testing. In: Internat. Symposium on Information Theory, pp. 2639–2643. IEEE 2008.

[10] P. Harremoës and I. Vajda: On the Bahadur-efficient testing of uniformity by means of the entropy. IEEE Trans. Inform. Theory *54* (2008), 1, 321–331.

[11] V. A. Kovalevskij: The Problem of Character Recognition from the Point of View of Mathematical Statistics. Spartan, New York 1967, pp. 3–30.

[12] J. W. Robbin and D. A. Salamon: The exponential Vandermonde matrix. Linear Algebra Appl. *317* (2000), 1–3, 225 – 226.

[13] W. Rudin: Principles of Mathematical Analysis. (Internat. Series in Pure and Applied Mathematics.) Third edition. McGraw-Hill, New York 1976.

[14] E. H. Spanier: Algebraic Topology. Springer, Berlin 1982.

[15] D. L. Tebbe and S. J. Dwyer: Uncertainty and the probability of error. IEEE Trans. Inform. Theory *14* (1968), 14, 516–518.

[16] K. Zyczkowski: Rényi extrapolation of Shannon entropy. Open Systems and Information Dynamics *10* (2003), 297–310.

Peter Harremoës, Centrum Wiskunde and Informatica, Science Park 123, 1090 GB Amsterdam, Noord-Holland. The Netherlands.
e-mail: harremoes@ieee.org