to appear in
Linear Algebra and its Applications

Partial subadditivity of entropies

Ádám Besenyei[a,1] **and Dénes Petz**[b,2]

[a]Department of Applied Analysis, Eötvös University
H-1117 Budapest, Pázmány P. 1/c, Hungary

and

[b]Alfréd Rényi Institute of Mathematics
H-1364 Budapest, POB 127, Hungary

In this paper a kind of partial subadditivity of the entropy is investigated in probability theory and in matrix theory. We show some new inequalities related to the subadditivity of the standard entropy and its one-parameter generalization in the discrete case. The matrix analogues of the inequalities are also studied. The standard entropy is related to the function $-x\log x$, but here some other functions are also used. A new version of the subadditivity of the Tsallis entropy is included in the probabilistic case and a conjecture in the matrix case.

1 Introduction

Let X be a discrete random variable with possible values $\{x_1,\dots,x_m\}$ and probability distribution $p(x)=P(X=x)$. The Shannon entropy of the discrete variable X is defined as

$$H(X)=-\sum_{i=1}^{m} p(x_i)\log p(x_i)$$

with the convention that $0\log 0=0$, see [2, 4] and other terminologies in [13, 14]. If Y is another discrete random variable with values $\{y_1,\dots,y_n\}$ and probability distribution $p(y)$, then the joint entropy of the pair (X,Y) with joint distribution $p(x,y)$ is

$$H(X,Y)=-\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i,y_j)\log p(x_i,y_j).$$

[1]E-mail: badam@cs.elte.hu
[2]E-mail: petz@math.bme.hu

It is well known that the joint entropy is not greater than the sum of the individual entropies:

$$H(X,Y) \le H(X) + H(Y) \tag{1}$$

which is called the subadditivity of the entropy. This property is a special case of a more general but also well-known inequality for three random variables, the strong subadditivity of the entropy:

$$H(X,Y,Z) + H(Y) \le H(X,Y) + H(Y,Z). \tag{2}$$

A one-parameter extension of the Shannon entropy is the Tsallis entropy [6, 7, 12]. Define the q-logarithm function as

$$\log_q x = \frac{x^{q-1}-1}{q-1} \quad (q \ne 1), \tag{3}$$

then the Tsallis entropy is

$$H_q(X) = -\sum_{i=1}^{m} p(x_i)\log_q p(x_i) = \frac{\sum_{i=1}^{m}\left(p(x_i)^q - p(x_i)\right)}{1-q}.$$

It is known that the strong subadditivity holds also for the Tsallis entropy in the discrete probabilistic case [6].

The quantum analogues of the above entropies are the following. In the quantum case one has a density matrix ρ which means that ρ is positive and $\mathrm{Tr}\,\rho = 1$. Then the von Neumann entropy of ρ is given by

$$S(\rho) = -\mathrm{Tr}\,\rho\log\rho.$$

For some applications, see [13, 14]. The Tsallis entropy is

$$S_q(\rho) = \frac{\mathrm{Tr}\,\rho^q - 1}{1-q} \quad (q > 1).$$

The strong subadditivity inequality is true in the quantum case for the standard entropy. The aim of the paper is to provide some new inequalities related to the (strong) subadditivity in the discrete case and study their analogues in the matrix case.

Below, we first concentrate on the probability theory and then the matrix theory. The probability theory is a special case where the matrices are diagonal. There are some inequalities which are true only in the probabilistic case.

2 Probability theory

We first reformulate the strong subadditivity inequality in the discrete case. If X, Y, Z are random variables with values $\{x_1,\dots,x_m\}$, $\{y_1,\dots,y_n\}$, $\{z_1,\dots,z_r\}$, then the joint distribution is

$$\{p_{ijk} := p(x_i,y_j,z_k) : 1 \le i \le m, 1 \le j \le n, 1 \le k \le r\}.$$

Denote the marginal distributions as

$$p_{ij-} := p(x_i, y_j) = \sum_{k=1}^{r} p_{ijk},\ \ p_{-j-} := p(y_j) = \sum_{i=1}^{m}\sum_{k=1}^{r} p_{ijk},\ \ p_{-jk} := p(y_j, z_k) = \sum_{i=1}^{m} p_{ijk}.$$

Then the strong subadditivity inequality can be written as

$$\sum_{ijk} -p_{ijk}\log p_{ijk} - \sum_{j} p_{-j-}\log p_{-j-} \le -\sum_{ij} p_{ij-}\log p_{ij-} - \sum_{jk} p_{-jk}\log p_{-jk}.$$

or equivalently

$$\sum_{ijk} p_{ijk}\left(\log p_{ijk} + \log p_{-j-} - \log p_{ij-} - \log p_{-jk}\right) \ge 0.$$

The new version of the strong subadditivity is the property that the parameters j and k can be fixed and the sum is over $1 \le i \le m$. So this is a kind of partial strong subadditivity.

Theorem 1 *For fixed* $1 \le j \le n$ *and* $1 \le k \le r$ *the following inequality holds:*

$$\sum_{i=1}^{n} p_{ijk}\left(\log p_{ijk} + \log p_{-j-} - \log p_{ij-} - \log p_{-jk}\right) \ge 0. \tag{4}$$

Proof: The inequality is equivalent to

$$\sum_{i} \frac{p_{ijk}}{p_{-jk}} \log\left(\frac{p_{ij-}p_{-jk}}{p_{ijk}p_{-j-}}\right) \le 0$$

where

$$\sum_{i} \frac{p_{ijk}}{p_{-jk}} = 1.$$

Thus, due to the concavity of the logarithm function,

$$\begin{aligned}\sum_{i} \frac{p_{ijk}}{p_{-jk}} \log\left(\frac{p_{ij-}p_{-jk}}{p_{ijk}p_{-j-}}\right) &\le \log\left(\sum_{i} \frac{p_{ijk}}{p_{-jk}}\frac{p_{ij-}p_{-jk}}{p_{ijk}p_{-j-}}\right)\\ &= \log\left(\sum_{i} \frac{p_{ij-}}{p_{-j-}}\right) = \log 1 = 0.\end{aligned}$$

So the theorem is proved. □

The sum of the inequality (4) with respect to j and k yields the ordinary strong subadditivity inequality (2). If $m = 1$, i.e., we have two random variables, then inequality (4) reduces to

$$\sum_{i=1}^{m} p_{ik}\left(\log p_{ik} - \log p_{i-} - \log p_{-k}\right) \ge 0$$

which can be regarded as a new version of subadditivity, a partial subadditivity. By taking the sum with respect to k, we obtain the subadditivity of the entropy (1).

The strong subadditivity of the Tsallis entropy has the form

$$\sum_{ijk} p_{ijk}\left(\log_q p_{ijk} + \log_q p_{-j-} - \log_q p_{ij-} - \log_q p_{-jk}\right) \geq 0.$$

We show that this inequality remains true if j and k are fixed and the sum is with respect to i, which can be regarded as the q-analogue of Theorem 1. We need a lemma, which also contains a remarkable inequality.

Lemma 1 *Let a_i, b_i $(i = 1, \ldots, m)$ be positive numbers and $q \geq 1$. Then*

$$\sum_{i=1}^m a_i b_i (a_i + b_i)^{q-2} \leq \sum_{i=1}^m a_i \cdot \sum_{i=1}^m b_i \cdot \left(\sum_{i=1}^m (a_i + b_i)\right)^{q-2}.$$

Proof: For simplicity, denote $A = \sum_{i=1}^m a_i$ and $B = \sum_{i=1}^m b_i$. By the concavity of the function $g(x) = x/(1+x)$, Jensen's inequality implies

$$\sum_{i=1}^m \frac{a_i}{A} g\left(\frac{b_i}{a_i}\right) \leq g\left(\sum_{i=1}^m \frac{a_i}{A}\frac{b_i}{a_i}\right) = g\left(\frac{B}{A}\right),$$

equivalently

$$\sum_{i=1}^m \frac{a_i b_i}{a_i + b_i} \leq \frac{AB}{A+B}$$

which is the desired inequality for $q = 1$ (this is called Milne's inequality, see [8]). If $q > 1$ then,

$$\begin{aligned}
\sum_{i=1}^m a_i b_i (a_i + b_i)^{q-2} &= \sum_{i=1}^m \frac{a_i b_i}{a_i + b_i}(a_i + b_i)^{q-1} \\
&\leq \sum_{i=1}^m \frac{a_i b_i}{a_i + b_i}(A + B)^{q-1} \\
&\leq \frac{AB}{A+B}(A+B)^{q-1} = AB(A+B)^{q-2}
\end{aligned}$$

which has to be proved. □

Now we can prove the q-analogue of Theorem 1.

Theorem 2 *For $q > 1$ and fixed $1 \leq j \leq n$, $1 \leq k \leq r$ the following inequality holds:*

$$\sum_{i=1}^n p_{ijk}\left(\log_q p_{ijk} + \log_q p_{-j-} - \log_q p_{ij-} - \log_q p_{-jk}\right) \geq 0.$$

Proof: The inequality is equivalent to

$$\sum_{i=1}^{n} p_{ijk}\left(p_{ij-}^{q-1} - p_{ijk}^{q-1}\right) \le p_{-jk}\left(p_{-j-}^{q-1} - p_{-jk}^{q-1}\right).$$

Denote $a_i = p_{ijk}$, $b_i = p_{ij-} - p_{ijk}$ $(i = 1, \dots, m)$ and $A = \sum_{i=1}^m a_i$, $B = \sum_{i=1}^m b_i$. Then $A = p_{-jk}$ and $A + B = p_{-j-}$ so the inequality reduces to the form

$$\sum_{i=1}^{m} a_i\left((a_i + b_i)^{q-1} - a_i^{q-1}\right) \le A\left((A+B)^{q-1} - A^{q-1}\right).$$

Since

$$(x+y)^{q-1} - x^{q-1} = (q-1)\int_0^1 y(x+ty)^{q-2}\, dt,$$

we have to prove that

$$(q-1)\int_0^1 \sum_{i=1}^{m} a_i b_i (a_i + tb_i)^{q-2}\, dt \le (q-1)\int_0^1 AB(A+tB)^{q-2}\, dt.$$

Applying Lemma 1 to the numbers a_i and tb_i $(i = 1, \dots, m)$ we obtain

$$t\sum_{i=1}^{m} a_i b_i (a_i + tb_i)^{q-2}\, dt \le tAB(A+tB)^{q-2},$$

and the desired inequality follows after division by t and integration over $[0, 1]$. □

As a special case of Theorem 2 we obtain a new version of the subadditivity of the Tsallis entropy.

Theorem 3 *Let p_{ij} $(i = 1, \dots, m; j = 1, \dots, n)$ be a probability distribution and denote the marginal distributions as $p_{i-} = \sum_{j=1}^n p_{ij}$ $(i = 1, \dots, m)$, $p_{-j} = \sum_{i=1}^m p_{ij}$ $(j = 1, \dots, n)$. Then for $q > 1$ and for every fixed $1 \le j \le n$,*

$$\sum_{i=1}^{m} p_{ij}\left(\log_q p_{ij} - \log_q p_{i-} - \log_q p_{-j}\right) \ge 0.$$

If the sum is also over j, then the result was already proved in 1970 by Daróczy [5].

3 Matrix theory

Now we consider matrices [9, 13, 14, 15]. Let $M_n(\mathbb{C})$ be the algebra of $n \times n$ matrices. If $A, B \in M_n(\mathbb{C})$, then $\langle A, B\rangle = \operatorname{Tr} A^*B$. The linear mappings $L_A, R_B : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ are defined as $L_A X = AX$ and $R_B X = XB$. They are commuting operators.

For a function $g : (0,\infty) \to \mathbb{R}$, we define the operator $m_g(P,Q)$ as follows:

$$\langle K, m_g(P,Q)K\rangle = \mathrm{Tr}\, K^* g(L_P R_Q^{-1}) R_Q(K) = \mathrm{Tr}\, K^* g(L_P R_Q^{-1}) KQ\,, \tag{5}$$

where $P, Q > 0$ are positive definite matrices. It is well-known that the map $(P,Q) \mapsto \langle K, m_g(P,Q)K\rangle$ is jointly convex for any pair of positive definite P, Q and any fixed K if g is operator convex, see Theorem 2 in [15]. There are different notation for $m_g(P,Q)$, for example $m_g(P,Q) = \mathbb{J}^g_{P,Q}$ and $\langle K, m_g(P,Q)K\rangle = S^K_g(P||Q)$.

Next we have $g(x) = -\log x$, then

$$\langle K, m_{-\log x}(P,Q)K\rangle = \mathrm{Tr}\,\Big(K^*KQ\log Q - KQK^*\log P\Big)$$

and this reduces to the usual relative entropy when $K = I$.

In the next two lemmas the matrix space is $M_m(\mathbb{C})\otimes M_n(\mathbb{C})$. If $\rho_{12} \in M_m(\mathbb{C})\otimes M_n(\mathbb{C})$, then the reduced densities $\rho_1 \in M_m(\mathbb{C})$ and $\rho_2 \in M_n(\mathbb{C})$ are defined as

$$\mathrm{Tr}\,(Y \otimes I_2)\rho_{12} = \mathrm{Tr}\, Y\rho_1, \quad \mathrm{Tr}\,(I_1 \otimes X)\rho_{12} = \mathrm{Tr}\, X\rho_2 \quad (Y \in M_m, \quad X \in M_n).$$

The general reduction is not simple, but for $n = 2$ and

$$\rho_{12} = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \qquad (A, B, C \in M_m)$$

we have

$$\rho_1 = \begin{bmatrix} \mathrm{Tr}\, A & \mathrm{Tr}\, B \\ \mathrm{Tr}\, B^* & \mathrm{Tr}\, C \end{bmatrix}, \qquad \rho_2 = A + C.$$

Lemma 2 *There are unitaries $\{W_t \in M_m(\mathbb{C}) : 1 \le t \le m^2\}$ such that for a unitary $U \in M_m(\mathbb{C})$ and for an operator $A_{12} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$ the formula*

$$\frac{1}{m}\sum_{t=1}^{m^2} (W_t U^* \otimes I_2) A_{12} (W_t U^* \otimes I_2)^* = I_1 \otimes A_2 \tag{6}$$

holds.

Proof: Let $|e_1\rangle, |e_2\rangle, \dots, |e_m\rangle$ be an orthonormal basis in $\mathbb{C}^m$. We define the unitary operators $X, Z \in M_m(\mathbb{C})$ as

$$X|e_k\rangle = |e_{k+1}\rangle \quad \text{for} \quad k < m \quad \text{and} \quad X|e_m\rangle = |e_1\rangle,$$

$$Z|e_k\rangle = \alpha^k|e_k\rangle \quad \text{for} \quad \alpha = e^{2\pi \mathrm{i}/m} \quad \text{and} \quad 1 \le k \le m.$$

Now

$$W_{j+m(k-1)} = X^j Z^k \qquad (1 \le j \le m, 1 \le k \le m)$$

gives m^2 unitaries.

We first note that for a unitary W the simple formula

$$W|f_1\rangle\langle f_2|W^* = |Wf_1\rangle\langle Wf_2|$$

can be used. Therefore,

$$(X^j Z^k)|e_u\rangle\langle e_v|(X^j Z^k)^* = |X^j Z^k e_u\rangle\langle X^j Z^k e_v| = |\alpha^{ku} e_{u+j}\rangle\langle\alpha^{kv} e_{v+j}|,$$

where the sums $u+j, v+j$ and the products ku, kv are modulo m. So

$$(X^j Z^k)|e_u\rangle\langle e_v|(X^j Z^k)^* = \alpha^{kv-ku}|e_{u+j}\rangle\langle e_{v+j}|.$$

The $u = v$ case is rather simple:

$$\sum_{j,k}(X^j Z^k)|e_u\rangle\langle e_u|(X^j Z^k)^* = \sum_k \left(\sum_j |e_{u+j}\rangle\langle e_{u+j}|\right) = \sum_k I = mI.$$

When $u \neq v$, then

$$\sum_{j,k} \alpha^{k(v-u)}|e_{u+j}\rangle\langle e_{v+j}| = 0.$$

Therefore

$$\frac{1}{m}\sum_{i=1}^{m^2} W_i A W_i^* = (\mathrm{Tr}\, A)I \qquad (A \in M_m(\mathbb{C}))$$

and

$$\frac{1}{m^2}\sum_{i=1}^{m^2}(W_i \otimes I)A_{12}(W_i \otimes I)^* = \frac{I}{m} \otimes \mathrm{Tr}_1 A_{12}.$$

If A_{12} is replaced with $(U^* \otimes I)A_{12}(U \otimes I)$, then (6) follows. □

Lemma 3 *Let A_{12}, B_{12} be strictly positive in $M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$ and let $K_{12} = V_1 \otimes K_2$ with V_1 unitary in $M_m(\mathbb{C})$. Then*

$$\langle K_2, m_g(A_2, B_2)K_2\rangle \le \langle K_{12}, m_g(A_{12}, B_{12})K_{12})\rangle.$$

Proof: Writing $\mathcal{W}_n$ for $W_n \otimes I_2$ and using (6) we have

$$\langle K_2, m_g(A_2, B_2)K_2\rangle = \frac{1}{m}\langle I_1 \otimes K_2, m_g(I_1 \otimes A_2, I_1 \otimes B_2)I_1 \otimes K_2\rangle$$
$$= \frac{1}{m}\langle I_1 \otimes K_2, m_g\Big(\frac{1}{m}\sum_n \mathcal{W}_n(V_1^* \otimes I_2)A_{12}(V_1 \otimes I_2)\mathcal{W}_n^*, \frac{1}{m}\sum_n \mathcal{W}_n B_{12}\mathcal{W}_n^*\Big)(I_1 \otimes K_2)\rangle$$

Now we use the joint convexity of $\langle K, m_g(A, B)K\rangle \equiv S_g^K(A||B)$ in A, B:

$$\le \frac{1}{m^2}\sum_n \langle I_1 \otimes K_2, m_g\Big(\mathcal{W}_n(V_1^* \otimes I_2)A_{12}(V_1 \otimes I_2)\mathcal{W}_n^*, \mathcal{W}_n B_{12}\mathcal{W}_n^*\Big)(I_1 \otimes K_2)\rangle$$
$$= \langle V_1 \otimes K_2, m_g(A_{12}, B_{12})V_1 \otimes K_2\rangle$$

where the equality follows from the unitary invariance of the trace. □

From the previous lemma we have the inequality

$$\langle K_{23}, m_g(A_{23}, B_{23})K_{23}\rangle \le \langle I_1 \otimes K_{23}, m_g(A_{123}, B_{123})I_1 \otimes K_{23}\rangle$$

when $A_{123}, B_{123} \in M_m(\mathbb{C})\otimes M_n(\mathbb{C})\otimes M_r(\mathbb{C})$ are reduced to $M_n(\mathbb{C})\otimes M_r(\mathbb{C})$. An essential case is $A_{123} = A_{12} \otimes I_3$ and $K_{23} = I_2 \otimes K_3$. Now this is rewritten as

$$\mathrm{Tr}\,(I_{12} \otimes K_3^*K_3)g(L_{A_{12}}R^{-1}_{B_{123}})\,B_{123} \ge \mathrm{Tr}(I_{12} \otimes K_3^*K_3)\,g(L_{A_2}R^{-1}_{B_{23}})\,B_{123}.$$

To have the result $\rho_{123} = B_{123}$, $\rho_2 = A_2$, $\rho_{12} = A_{12}$, $\rho_{23} = B_{23}$ and $T_3 = K_3^*K_3$ we have:

Theorem 4 *The density $\rho_{123} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C}) \otimes M_r(\mathbb{C})$ and reduced densities are ρ_{12}, ρ_2, ρ_{23}. If $0 \le T_3 \in M_r(\mathbb{C})$, then*

$$\mathrm{Tr}(I_{12} \otimes T_3)\rho_{123}(\log\rho_{123} - \log\rho_{12} + \log\rho_2 - \log\rho_{23}) \ge 0.$$

The previous theorem was obtained recently in [11].

Now we turn to the case of the Tsallis entropy. Let $\rho_{12} \in M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ be a density matrix and let $\rho_1 \in M_n(\mathbb{C})$ and $\rho_2 \in M_m(\mathbb{C})$ be the reduced densities. Then the subadditivity of the Tsallis entropy is

$$\mathrm{Tr}\,\rho_{12}\log_q\rho_{12} \ge \mathrm{Tr}\,\rho_1\log_q\rho_1 + \mathrm{Tr}\,\rho_2\log_q\rho_2.$$

or equivalently

$$\mathrm{Tr}\,(\rho_{12}^q - \rho_{12}) \ge \mathrm{Tr}\,(\rho_1^q - \rho_1) + \mathrm{Tr}\,(\rho_2^q - \rho_2).$$

In the paper [1] the inequality of Audenaert

$$1 + \mathrm{Tr}\,\rho_{12}^q \ge \mathrm{Tr}\,\rho_1^q + \mathrm{Tr}\,\rho_2^q \tag{7}$$

is proved. In the previous formulas there are different traces, to have one trace we can take

$$\mathrm{Tr}\,\rho_{12}\left(\log_q\rho_{12} - \log_q\rho_1 \otimes I_2 - I_1 \otimes \log_q\rho_2\right) \ge 0\,.$$

Motivated by Theorem 4 and numeical examples we conjecture the inequality

$$\mathrm{Tr}\,(T \otimes I_2)\rho_{12}\left(\log_q\rho_{12} - \log_q\rho_1 \otimes I_2 - I_1 \otimes \log_q\rho_2\right) \ge 0 \tag{8}$$

for $0 \le T \in M_n(\mathbb{C})$. Actually this has some similarity to (7):

$$\mathrm{Tr}\,T\rho_1 + \mathrm{Tr}\,(T \otimes I_2)\rho_{12}^q \ge \mathrm{Tr}\,T\rho_1^q + \mathrm{Tr}\,(T \otimes \rho_2^{q-1})\rho_{12}. \tag{9}$$

The above inequalites seem numerically true. Moreover, we can prove some special cases: **1.** $q > 1$ and $\rho_{12} = \rho_1 \otimes \rho_2$. **2.** $q = 2$ and T, ρ_1, ρ_2 are 2×2 matrices. The second case is detailed in the next example.

Example 1 We concentrate on the space $M_2(\mathbb{C})\otimes M_2(\mathbb{C})$. So the essential formulas are the following:

$$\rho_{12}=\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \qquad \rho_2=A+C, \qquad \rho_1=\begin{bmatrix} \mathrm{Tr}\, A & \mathrm{Tr}\, B \\ \mathrm{Tr}\, B^* & \mathrm{Tr}\, C \end{bmatrix},$$

$$\rho_{12}^2=\begin{bmatrix} A^2+BB^* & AB+BC \\ B^*A+CB^* & B^*B+C^2 \end{bmatrix}, \quad T=\begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad T\otimes I=\begin{bmatrix} aI & bI \\ bI & cI \end{bmatrix}.$$

By simple calculation we have:

$$\begin{gathered}
\mathrm{Tr}\,(T\otimes I)(\rho_{12}^2-\rho_{12})=a\mathrm{Tr}\,(A^2+BB^*)+b\mathrm{Tr}\,(B^*A+CB^*)+b\mathrm{Tr}\,(AB+BC) \\
+c\mathrm{Tr}\,(B^*B+C^2)-a\mathrm{Tr}\,A-b\mathrm{Tr}\,B^*-b\mathrm{Tr}\,B-c\mathrm{Tr}\,C, \\
\mathrm{Tr}\,(T\otimes I)(\rho_{12}(\log_2\rho_1\otimes I))=a\mathrm{Tr}\,A(\mathrm{Tr}\,A-1)+a\mathrm{Tr}\,B\mathrm{Tr}\,B^*+b\mathrm{Tr}\,B^*(\mathrm{Tr}\,A-1) \\
+b\mathrm{Tr}\,C\mathrm{Tr}\,B^*+c\mathrm{Tr}\,B^*\mathrm{Tr}\,B+c\mathrm{Tr}\,C(\mathrm{Tr}\,C-1), \\
\mathrm{Tr}\,(T\otimes I)(\rho_{12}(I\otimes\log_2\rho_2))=a\mathrm{Tr}\,(A^2+AC-A)+b\mathrm{Tr}\,(BA+BC-B) \\
+b\mathrm{Tr}\,(B^*A+B^*C-B^*)+c\mathrm{Tr}\,(CA+C^2-C).
\end{gathered}$$

For the simplification we used several times the formula $\mathrm{Tr}\,A+\mathrm{Tr}\,C=1$. So inequality reduces to (8)

$$(a+c)(\mathrm{Tr}\,BB^*-\mathrm{Tr}\,B\mathrm{Tr}\,B^*-\mathrm{Tr}\,AC+\mathrm{Tr}\,A\mathrm{Tr}\,C)\geq 0$$

and we need to prove

$$\mathrm{Tr}\,AC-\mathrm{Tr}\,BB^*\leq(\mathrm{Tr}\,A)(\mathrm{Tr}\,C)-|\mathrm{Tr}\,B|^2\,.$$

We may suppose that A is diagonal, then this inequality has the form

$$\sum_i A_{ii}C_{ii}-\sum_{ij}|B_{ij}|^2\leq\sum_i A_{ii}\sum_i C_{ii}-\left|\sum_i B_{ii}\right|^2$$

or

$$\left|\sum_i B_{ii}\right|^2-\sum_{ij}|B_{ij}|^2\leq\sum_i A_{ii}\sum_i C_{ii}-\sum_i A_{ii}C_{ii}=\sum_{i>j}(A_{ii}C_{jj}+A_{jj}C_{ii}).$$

The left-hand side can be written as

$$\sum_{ij}B_{ii}\overline{B_{jj}}-\sum_{ij}|B_{ij}|^2=2\sum_{i>j}\mathrm{Re}\,B_{ii}\overline{B_{jj}}-\sum_{i\neq j}|B_{ij}|^2\,.$$

So the inequality reduces to

$$2\sum_{i>j}\mathrm{Re}\,B_{ii}\overline{B_{jj}}-\sum_{i\neq j}|B_{ij}|^2\leq\sum_{i>j}(A_{ii}C_{jj}+A_{jj}C_{ii}).$$

We show that

$$2\sum_{i>j} \mathrm{Re}\, B_{ii}\overline{B_{jj}} \le A_{ii}C_{jj} + A_{jj}C_{ii}$$

for $i > j$. This follows from the inequality of arithmetic and geometric means:

$$A_{ii}C_{jj} + A_{jj}C_{ii} \ge 2\sqrt{A_{ii}C_{jj}A_{jj}C_{ii}} \ge 2\sqrt{|B_{ii}|^2|B_{jj}|^2} \ge 2\mathrm{Re}\,(B_{ii}B_{jj}).$$

So we the inequality (9) holds for $q = 2$. The case of equality is very special. If A is diagonal, then A and C should be proportional. □

Acknowledgments. This work was partially supported by the Hungarian Research Grant OTKA K104206.

References

[1] K.M. R. Audenaert, Subadditivity of q-entropies for $q > 1$, J. Math. Phys. **48**(2007), 083507.

[2] T.M. Cover and J. A. Thomas, *Elements of information theory*, second edition, Wiley-Interscience, 2006.

[3] I. Csiszár, Information type measure of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar. **2**(1967), 299–318.

[4] I. Csiszár and J. Körner, *Information theory. Coding theorems for discrete memoryless systems*, second edition, Cambridge University Press, Cambridge, 2011.

[5] Z. Daróczy, Generalized information functions, Information and Control **16**(1970), 36–51.

[6] S. Furuichi, *Tsallis entropies and their theorems, properties and applications*, Aspects of Optical Sciences and Quantum Information, 2007.

[7] S. Furuichi, N. Minculete and F-C. Mitroi, Some inequalities on generalized entropies, Journal of Inequalities and Applications, **226**(2012), 1–17.

[8] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1934.

[9] F. Hiai, M. Mosonyi D. Petz and C. Beny, Monotonicity of f-divergences and error correction, Rev. Math. Phys. **23**(2011), 691–747.

[10] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*, to appear in Hindustan Book Agency.

[11] I.H. Kim, Operator extension of strong subadditivity of entropy, arXiv:1210.5190

[12] J. Naudts, Generalised exponential families and associated entropy functions, Entropy **10**(2008), 131–149.

[13] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Springer-Verlag, Heidelberg, 1993. Second edition 2004.

[14] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer, Berlin, Heidelberg, 2008.

[15] D. Petz, From quasi-entropy, Annales Univ. Sci. Budapest, **55**(2012), 81–93.