MODULUS OF CONTINUITY OF THE QUANTUM
f–ENTROPY WITH RESPECT TO THE TRACE DISTANCE

IOSIF PINELIS

(Communicated by J. Pečarić)

Abstract. A well-known result due to Fannes is a certain upper bound on the modulus of continuity of the von Neumann entropy with respect to the trace distance between density matrices; this distance is the maximum probability of distinguishing between the corresponding quantum states. Much more recently, Audenaert obtained an exact expression of this modulus of continuity.

In the present note, Audenaert's result is extended to a broad class of entropy functions indexed by arbitrary continuous convex functions f in place of the Shannon–von Neumann function $x \mapsto x \log_2 x$. The proof is based on the Schur majorization.

1. Summary and discussion

Let ρ and σ be two density matrices of a finite dimension $d \geqslant 2$, that is, two positive-semidefinite Hermitian linear operators of trace 1 acting on a d-dimensional Hilbert space $\mathscr{H}$. The trace distance between ρ and σ is

$$T(\rho, \sigma) := \frac{1}{2} \operatorname{tr} |\rho - \sigma| = \frac{1}{2} \sum_{i=1}^{d} |\lambda_i(\rho - \sigma)| = \sum_{i=1}^{d} \max (0, \lambda_i(\rho - \sigma)). \qquad (1.1)$$

Here and in what follows, as usual, tr denotes the trace, $\lambda_1(\tau) \geqslant \cdots \geqslant \lambda_d(\tau)$ are the eigenvalues of a Hermitian linear operator τ on $\mathscr{H}$, and, for any real-valued function f defined on the spectrum of τ, the action of f on τ is defined by the formula

$$f(\tau) := \sum_{i=1}^{d} f(\lambda_i(\tau)) P_i(\tau),$$

given the spectral decomposition $\tau = \sum_{i=1}^{d} \lambda_i(\tau) P_i(\tau)$ of τ, with appropriate orthogonal projectors $P_i(\tau)$, so that $\sum_{i=1}^{d} P_i(\tau)$ is the identity operator on $\mathscr{H}$; in particular, $|\tau| = \sum_{i=1}^{d} |\lambda_i(\tau)| P_i(\tau)$.

One may also write

$$T(\rho,\sigma) = \sup_{P} \mathrm{tr}\left(P(\rho - \sigma)\right), \tag{1.2}$$

where the supremum is taken over all orthogonal projectors P of the Hilbert space $\mathscr{H}$; cf. e.g. [13, Ch. 9]. Thus, in view of Gleason's theorem [7], which states that any natural assignment of probabilities to measurement outcomes must follow the Born rule, one sees that the trace distance $T(\rho,\sigma)$ is the maximum probability of distinguishing between the two quantum states given by ρ and σ.

By the norm inequality or by (1.2), $0 \leqslant T(\rho,\sigma) \leqslant 1$. Moreover, it is easy to see that $T(\rho,\sigma)$ can take any value in the interval $[0,1]$. Indeed, suppose e.g. that the operators ρ and σ are commuting, with the same eigenbasis and the corresponding eigenvalues p_j and q_j given by the formulas $p_1 := t$, $p_2 = 1-t$, $q_1 := 1-t$, $q_2 = t$, and $p_3 = q_3 = \ldots = p_d = q_d = 0$. Then, letting t vary from 0 to $1/2$, we see that $T(\rho,\sigma) = \frac{1}{2} \sum_{k=1}^{d} |p_k - q_k|$ will continuously vary from 1 to 0.

The von Neumann entropy of a density matrix ρ is

$$S(\rho) := -\mathrm{tr}(\rho \log_2 \rho),$$

with $0\log_2 0 := 0$. Audenaert [2] showed that

$$|S(\rho) - S(\sigma)| \leqslant \Delta_d(\varepsilon) := h(\varepsilon) + \varepsilon \log_2(d-1), \tag{1.3}$$

where

$$\varepsilon := T(\rho,\sigma) \tag{1.4}$$

and

$$h(\varepsilon) := -\varepsilon \log_2 \varepsilon - (1-\varepsilon) \log_2 (1-\varepsilon).$$

As pointed out in [2], the upper bound $\Delta_d(\varepsilon)$ on $|S(\rho) - S(\sigma)|$ is better (that is, smaller) than the well-known bound due to Fannes [5, 13]. Moreover, as was noted in [2], the bound $\Delta_d(\varepsilon)$ on $|S(\rho) - S(\sigma)|$ is the best possible one in terms of ε and d.

In this note, inequality (1.3) is extended to general continuous convex functions of density matrices instead of the convex function $p \mapsto p\log_2 p$, as follows:

THEOREM 1. *Take any continuous convex function* $f \colon [0,1] \to \mathbb{R}$ *and consider the corresponding generalized f-entropy*

$$S_f(\rho) := -\mathrm{tr} f(\rho) \tag{1.5}$$

of a density matrix ρ. Then

$$|S_f(\rho) - S_f(\sigma)| \leqslant \Delta_{f;d}(\varepsilon) \tag{1.6}$$
$$:= f(1) - f(1-\varepsilon) - (d-1)\left(f\left(\frac{\varepsilon}{d-1}\right) - f(0)\right).$$

The bound $\Delta_{f;d}(\varepsilon)$ on $|S_f(\rho) - S_f(\sigma)|$ is exact for each $\varepsilon \in [0,1]$, as it is attained by $|S_f(\rho) - S_f(\sigma)|$ for some density matrices ρ and σ such that $T(\rho,\sigma) = \varepsilon$.

The necessary proofs will be given in Section 2.

In the particular case when $f(p) = p \ln p$ (with $f(0) = 0$), the exact bound $\Delta_{f;d}(\varepsilon)$ coincides with the Audenauert bound $\Delta_d(\varepsilon)$.

Almost immediately from Theorem 1, we obtain the following.

COROLLARY 2. *The modulus of continuity of the generalized f-entropy $S_f(\cdot)$ with respect to the trace distance is given by the formula*

$$\sup_{(\rho,\sigma)\colon T(\rho,\sigma)\leqslant\varepsilon} |S_f(\rho) - S_f(\sigma)| = \max_{(\rho,\sigma)\colon T(\rho,\sigma)\leqslant\varepsilon} |S_f(\rho) - S_f(\sigma)|$$
$$= \Delta_{f;d}\left(\min(\varepsilon, 1 - \tfrac{1}{d})\right)$$

for all $\varepsilon \in [0,1]$.

The von Neumann entropy is a special case of the generalized f-entropy, with $f(x) = x\log_2 x$. Another special case of the generalized f-entropy is the Tsallis entropy [17, 6, 18]

$$\mathrm{Ts}_\alpha(\rho) := \frac{1 - \mathrm{tr}(\rho^\alpha)}{\alpha - 1}$$

for real $\alpha > 1$, corresponding to the continuous convex function f given by the formula $f(x) := (x^\alpha - x)/(\alpha - 1)$ for real $x \in [0,1]$. The special case of the Tsallis entropy with $\alpha = 2$ is also known as the Gini–Simpson index [11] or the Gibbs–Martin/Blau index [4] or the expected heterozygosity [16]. The Tsallis entropy $\mathrm{Ts}_\alpha(\rho)$ is related with the Renyi entropy

$$\mathrm{R}_\alpha(\rho) := \frac{\log_2 \mathrm{tr}(\rho^\alpha)}{1 - \alpha}$$

by the strictly increasing one-to-one transformation given by the formula

$$\mathrm{Ts}_\alpha(\rho) = \frac{1 - 2^{(1-\alpha)\mathrm{R}_\alpha(\rho)}}{\alpha - 1}.$$

In the non-quantum case, that is, for a "probability distribution" $(p_1, \ldots, p_d)$ with nonnegative $p_1, \ldots, p_d$ such that $p_1 + \ldots + p_d = 1$, the notion of the f-entropy $-\sum_{k=1}^d f(p_k)$, again for a convex function f, was considered in [1].

The special case of Theorem 1 corresponding to the Tsallis entropy was obtained in [19], by a different method, involving probabilistic coupling.

In the following corollary, obtained from Theorem 1 mainly by simple rescaling, the unit-trace condition on ρ and σ is dropped. The resulting statement may be of some convenience. It will actually be used in the proof of Theorem 1.

COROLLARY 3. *Let ρ and σ be two positive-semidefinite Hermitian linear operators of the same trace t acting on a Hilbert space of a finite dimension $d \geqslant 2$. Take any continuous convex function $f\colon [0,t] \to \mathbb{R}$. Then $\varepsilon \in [0,t]$ and*

$$|S_f(\rho) - S_f(\sigma)| \leqslant \Delta_{t;f;d}(\varepsilon) := f(t) - f(t - \varepsilon) - (d - 1)\left(f\left(\frac{\varepsilon}{d-1}\right) - f(0)\right), \quad (1.7)$$

with S_f, T, and ε still as defined in (1.5), (1.1), and (1.4). Also, $\Delta_{t;f;d}(\varepsilon)$ is nondecreasing in t and in d.

REMARK 4. Just as the bound $\Delta_{f;d}(\varepsilon) = \Delta_{1;f;d}(\varepsilon)$ is exact in the setting of Theorem 1, the bound $\Delta_{t;f;d}(\varepsilon)$ is exact in the slightly more general setting of Corollary 3. However, somewhat surprisingly and in contrast with the last sentence in Corollary 3, $\Delta_{t;f;d}(\varepsilon)$ is not monotonic in $\varepsilon \in [0,t]$, for any real $t > 0$, as can be easily seen from the proofs of Corollaries 3 and 2.

2. Proofs

First here, let us deduce Corollary 3 from Theorem 1 or, more specifically, from inequality (1.6):

Proof of Corollary 3. That $\varepsilon \in [0,t]$ follows by the norm inequality. If $t = 0$ then $\rho = \sigma = 0$ and hence $\varepsilon = 0$, so that inequality (1.7) is trivial. Assume now that $t > 0$. Let $\tilde{\rho} := \rho/t$, $\tilde{\sigma} := \sigma/t$, $\tilde{\varepsilon} := T(\tilde{\rho}, \tilde{\sigma}) = \varepsilon/t$, and $\tilde{f}(u) := f(tu)$ for all real u. Then $\operatorname{tr} \tilde{\rho} = 1 = \operatorname{tr} \tilde{\sigma}$, $S_f(\rho) = S_{\tilde{f}}(\tilde{\rho})$, $S_f(\sigma) = S_{\tilde{f}}(\tilde{\sigma})$, and $\Delta_{\tilde{f};d}(\tilde{\varepsilon}) = \Delta_{t;f;d}(\varepsilon)$. So, (1.7) is obtained by using (1.6) with $\tilde{\rho}$, $\tilde{\sigma}$, $\tilde{\varepsilon}$, and $\tilde{f}$ instead of ρ, σ, ε, and f, respectively. The last sentence in Corollary 3 follows by the convexity of f. □

Proof of Theorem 1. We shall prove inequality (1.6) by induction on d.

By approximation, without loss of generality (wlog) the function f is *strictly* convex and differentiable.

The proof uses the powerful majorization tool; cf. e.g. [12, Definition A.1]. For any natural n, we say that a vector $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ majorizes (in the Schur sense) a vector $b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ and write $a \succcurlyeq b$ if for the corresponding decreasing rearrangements $a^{\downarrow} = (a_{n:1}, \ldots, a_{n:n})$ and $b^{\downarrow} = (b_{n:1}, \ldots, b_{n:n})$ of the vectors a and b we have $\sum_{i=1}^{n} a_{n:i} = \sum_{i=1}^{n} b_{n:i}$ and $\sum_{i=1}^{j} a_{n:i} \geqslant \sum_{i=1}^{j} b_{n:i}$ for all $j = 1, \ldots, n-1$. One may note here that

$$a_{n:j} = \max_{J \in \mathscr{J}_{n,j}} \min_{i \in J} a_i \tag{2.1}$$

for $j \in [n] := \{1, \ldots, n\}$, where $\mathscr{J}_{n,j}$ denotes the set of all subsets of cardinality j of the set $[n]$; cf. e.g. [14, formula (8.2)].

As in [2], we now invoke the fundamental double inequality

$$\tfrac{1}{2} \|p^{\downarrow} - q^{\downarrow}\|_1 =: \varepsilon^{\downarrow\downarrow} \leqslant \varepsilon = T(\rho, \sigma) \leqslant \varepsilon^{\uparrow\downarrow} := \tfrac{1}{2} \|p^{\uparrow} - q^{\downarrow}\|_1, \tag{2.2}$$

where $\|\cdot\|_1$ is the ℓ^1 norm on $\mathbb{R}^d$, $p^{\uparrow}$ is the vector of the eigenvalues $p_1, \ldots, p_d$ of ρ sorted in the ascending order, and $q^{\downarrow}$ is the vector of the eigenvalues $q_1, \ldots, q_d$ of σ sorted in the descending order. The double inequality (2.2) is a special, trace-norm case of the corresponding more general result for unitarily invariant norms; see e.g. the double inequality [3, (IV.62)].

Note that for $\varepsilon \in (0,1)$ the derivative in ε of $\Delta_{f;d}(\varepsilon)$ is $f'(1-\varepsilon) - f'(\frac{\varepsilon}{d-1}) \geqslant 0$ if $1 - \varepsilon \geqslant \frac{\varepsilon}{d-1}$, that is, if $\varepsilon \leqslant 1 - \frac{1}{d}$. So, $\Delta_{f;d}(\varepsilon)$ is nondecreasing in $\varepsilon \in [0, 1 - \frac{1}{d}]$. Similarly, $\Delta_{f;d}(\varepsilon)$ is nonincreasing in $\varepsilon \in [1 - \frac{1}{d}, 1]$. Now it follows by (2.2) that

$$\Delta_{f;d}(\varepsilon) \geqslant \min\left(\Delta_{f;d}(\varepsilon^{\downarrow\downarrow}), \Delta_{f;d}(\varepsilon^{\uparrow\downarrow})\right). \tag{2.3}$$

(The argument presented in this paragraph is missing in [2].)

> At this point, let us "forget" the definition of ε in (1.4) and, instead, take any $\varepsilon \in (0,1)$.

In view of (2.3), to prove inequality (1.6), it is enough to show that

$$D_f(p,q) \overset{(?)}{\leqslant} \Delta_{f;d}(\varepsilon) \tag{2.4}$$

for all $(p,q) \in P_{d;\varepsilon}$, where

$$D_f(p,q) := \sum_1^d f(p_i) - \sum_1^d f(q_i), \tag{2.5}$$

$$P_{d;\varepsilon} := \{(p,q) \in P_d : E(p,q) = \varepsilon\},$$

$$P_d := S_d \times S_d,$$

$$S_d := \left\{ p = (p_1,\ldots,p_d) \in [0,\infty)^d : \sum_1^d p_i = 1 \right\},$$

$$E(p,q) := \frac{1}{2} \sum_1^d |p_i - q_i|.$$

In the case $d = 2$, take any $(p,q) \in P_{d;\varepsilon}$, so that $p = (a, 1-a) \in S_d$ and $q = (b, 1-b) \in S_d$ for some a and b in $[0,1]$ such that $\varepsilon = E(p,q) = |a-b|$. Wlog, $a \geqslant b$, and hence $\varepsilon = E(p,q) = a - b \geqslant 0$. Therefore and because f is convex,

$$\begin{aligned} D_f(p,q) &= [f(a) - f(b)] - [f(1-b) - f(1-a)] \\ &= [f(a) - f(a-\varepsilon)] - [f(1-a+\varepsilon) - f(1-a)] \\ &\leqslant [f(1) - f(1-\varepsilon)] - [f(\varepsilon) - f(0)] = \Delta_{f;d}(\varepsilon). \end{aligned}$$

Thus, in the case $d = 2$, (2.4) holds and hence (1.6) holds. This establishes the basis of the induction mentioned in the very beginning of the proof of Theorem 1. Assume now that $d \geqslant 3$. By continuity and compactness, there is a maximizer $(p,q) \in P_{d;\varepsilon}$ of D_f.

> In what follows, it is assumed by default that $(p,q) \in P_{d;\varepsilon}$ is such a maximizer; in particular, we have $E(p,q) = \varepsilon$.

Wlog, for some $k \in [d]$ we have

$$p_i \geqslant q_i \text{ for } i = 1,\ldots,k, \tag{2.6}$$

$$p_i \leqslant q_i \text{ for } i = k+1,\ldots,d, \tag{2.7}$$

$$q_1 \geqslant \cdots \geqslant q_k. \tag{2.8}$$

So,

$$\varepsilon = \sum_{i=1}^{k}(p_i - q_i) = \sum_{i=k+1}^{d}(q_i - p_i).$$

Let now

$$p_1^* := q_1 + \varepsilon \text{ and } p_i^* := q_i \text{ for } i = 2, \dots, k. \tag{2.9}$$

Then the vector $(p_1^*, \dots, p_k^*)$ majorizes (in the Schur sense) the vector $(p_1, \dots, p_k)$.

To see why this is true, note first that, by (2.9), (2.8), and (2.6), $p_{k:1}^* = p_1^* = q_1 + \varepsilon$ and $p_{k:i}^* = q_{k:i} = q_i \leqslant p_i$ for $i = 2, \dots, k$. Moreover, by (2.1) and (2.6), $q_{k:i} \leqslant p_{k:i}$ for all $i \in [k]$. So,

$$\sum_{i=j+1}^{k} p_{k:i}^* = \sum_{i=j+1}^{k} q_{k:i} \leqslant \sum_{i=j+1}^{k} p_{k:i} \text{ for all } j \in [k]. \tag{2.10}$$

Also,

$$\sum_{i=1}^{k} p_{k:i}^* = \sum_{i=1}^{k} p_i^* = q_1 + \varepsilon + \sum_{i=2}^{k} q_i = \sum_{i=1}^{k} p_i. \tag{2.11}$$

By (2.11) and (2.10), $\sum_{i=1}^{k} p_{k:i}^* = \sum_{i=1}^{k} p_i$ and $\sum_{i=1}^{j} p_{k:i}^* \geqslant \sum_{i=1}^{j} p_{k:i}$ for all $j \in [k]$. Thus, indeed $(p_1^*, \dots, p_k^*) \succcurlyeq (p_1, \dots, p_k)$.

Therefore and because f is continuous and convex, we have $\sum_1^k f(p_i^*) \geqslant \sum_1^k f(p_i)$, in view of the equivalence of items (i) and (iv) in [12, Theorem A.3]. Also, by (2.9), $p_i^* \geqslant q_i$ for all $i \in [k]$. So, if we replace $p_1, \dots, p_k$ respectively by $p_1^*, \dots, p_k^*$, then condition (2.6) will continue to hold, as well as the other conditions imposed above on $p = (p_1, \dots, p_d)$, whereas the value of D_f, as defined in (2.5), may only increase after this replacement.

So, wlog $(p_1, \dots, p_k) = (p_1^*, \dots, p_k^*)$. Then, by (2.9) and (2.7),

$$p_1 > q_1 \text{ and } p_i \leqslant q_i \text{ for all } i = 2, \dots, d, \tag{2.12}$$

and hence

$$\varepsilon = p_1 - q_1 = \sum_{2}^{d}(q_i - p_i). \tag{2.13}$$

Take any permutation $\pi\colon [d] \to [d]$ such that $\pi(1) = 1$. Let $p_\pi := (p_{\pi(1)}, \dots, p_{\pi(d)})$ and similarly define q_π. Then clearly $E(p_\pi, q_\pi) = E(p, q)$, $D_f(p_\pi, q_\pi) = D_f(p, q)$, and condition (2.12) holds with (p_π, q_π) in place of (p, q) whenever it holds for (p, q). Let us refer to this as the permutation invariance (of E, D_f, and (2.12)).

Suppose now for a moment that $p_i = q_i$ for some $i \in [d]$; by the permutation invariance, wlog $i = d$, so that $p_d = q_d =: c$. Let ρ and σ be two positive-semidefinite Hermitian linear operators acting on a Hilbert space of the finite dimension $d - 1 \geqslant 2$, with the same eigenbasis and with the eigenvalues $p_1, \dots, p_{d-1}$ for ρ and $q_1, \dots, q_{d-1}$ for σ. Then $\mathrm{tr}\,\rho = \mathrm{tr}\,\sigma = 1 - c$ and $T(\rho, \sigma) = p_1 - q_1 = \sum_2^{d-1}(q_i - p_i) = \sum_2^{d}(q_i - p_i) = \varepsilon$, with ε as in (2.13). So, by Corollary 3 (which is, specifically, a corollary of inequality

(1.6)) and the induction mentioned in the beginning of the proof of Theorem 1,

$$|D_f(p,q)| = \left| \sum_1^d f(p_i) - \sum_1^d f(q_i) \right| = \left| \sum_1^{d-1} f(p_i) - \sum_1^{d-1} f(q_i) \right|$$

$$= |S_f(\rho) - S_f(\sigma)| \leqslant \Delta_{1-c;f;d-1}(\varepsilon) \leqslant \Delta_{1;f;d}(\varepsilon) = \Delta_{f;d}(\varepsilon).$$

So, (2.4) follows if $p_i = q_i$ for some $i \in [d]$.

Thus, we may and will henceforth assume that all inequalities in (2.12) are strict:

$$p_1 > q_1 \text{ and } p_i < q_i \text{ for all } i \in [d] \setminus \{1\}. \tag{2.14}$$

Suppose next that there are two distinct numbers j and k in $[d] \setminus \{1\}$ such that $p_j > 0$ and $p_k > 0$. By the permutation invariance, wlog $p_2 \geqslant p_3 > 0$. Replace now $p = (p_1, \ldots, p_d)$ by $\tilde{p} := (p_1, p_2 + t, p_3 - t, p_4, \ldots, p_d)$, where $t > 0$ is close enough to 0 – more specifically, one may take here any $t \in (0, \min[p_3, q_2 - p_2])$. Then $(\tilde{p}, q) \in P_d$ and, by the condition $p_i < q_i$ for $i \in [d] \setminus \{1\}$ in (2.14), $E(\tilde{p}, q) = E(p, q) = \varepsilon$, so that $(\tilde{p}, q) \in P_{d;\varepsilon}$. Also, by the strict convexity of f, we have $D_f(\tilde{p}, q) > D_f(p, q)$, which contradicts the assumption that $(p, q) \in P_{d;\varepsilon}$ is a maximizer of D_f.

So, $p_{i_*} > 0$ for at most one $i_* \in [d] \setminus \{1\}$, and, by the permutation invariance, wlog $i_* = 2$, so that

$$p = (p_1, p_2, 0, \ldots, 0). \tag{2.15}$$

Further, using the convexity of the function f and Jensen's inequality, we see that $\sum_{i=3}^d f(q_i) \geqslant (d-2) f(\frac{1}{d-2} \sum_{i=3}^d q_i)$. So, wlog $q_3 = \cdots = q_d$.

Furthermore, if $q_2 \neq q_3$, replace q by $\tilde{q} := (q_1, (1-t)q_2 + tq_3, (1-t)q_3 + tq_2, q_4, \ldots, q_d)$ for a small enough $t \in (0, 1)$. Then, in view of the condition $p_i < q_i$ for $i \in [d] \setminus \{1\}$ in (2.14), $(p, \tilde{q}) \in P_{d;\varepsilon}$, but $D_f(p, \tilde{q}) > D_f(p, q)$, a contradiction.

Thus, $q_2 = q_3 = \cdots = q_d$, so that, in view of (2.15) and (2.13), for some $x \in [0, 1]$ we have

$$p = P(x) := (x, 1-x, 0, \ldots, 0), \quad q = Q(x) := \left(x - \varepsilon, \frac{1-x+\varepsilon}{d-1}, \ldots, \frac{1-x+\varepsilon}{d-1} \right). \tag{2.16}$$

Moreover, condition $p_2 < q_2$ in (2.14) can now be rewritten as

$$\frac{1-x+\varepsilon}{d-1} > 1-x \tag{2.17}$$

and inequality (2.4) can be rewritten as

$$h(x) := f(x) - f(x-\varepsilon) + f(1-x) - (d-1)f\left(\frac{1-x+\varepsilon}{d-1}\right) \overset{(?)}{\leqslant} h(1), \tag{2.18}$$

which follows because

$$h'(u) := [f'(u) - f'(u-\varepsilon)] + \left[f'\left(\frac{1-u+\varepsilon}{d-1}\right) - f'(1-u) \right] \geqslant 0 \tag{2.19}$$

for $u \in [x, 1]$, in view of the convexity of f and condition (2.17) (which implies $\frac{1-u+\varepsilon}{d-1} > 1 - u$ for all $u \in [x, 1]$).

This completes the proof of the inequality in (1.6).

To complete the entire proof of Theorem 1, it remains to note that the bound $\Delta_{f;d}(\varepsilon)$ on $|S_f(\rho) - S_f(\sigma)|$ is attained, for each $\varepsilon \in [0, 1]$, when the density matrices ρ and σ have a common eigenbasis with respective d-tuples of eigenvalues $p = (p_1, \ldots, p_d) = P(1)$ and $q = (q_1, \ldots, q_d) = Q(1)$ with P and Q as defined in (2.16). □

Proof of Corollary 2. This follows immediately from Theorem 1 and the observation, made in the paragraph containing inequality (2.3), that $\Delta_{f;d}(\varepsilon)$ is nondecreasing in $\varepsilon \in [0, 1 - \frac{1}{d}]$ and nonincreasing in $\varepsilon \in [1 - \frac{1}{d}, 1]$. □

After this paper was accepted for publication and posted on arXiv [15], the author has been informed by N. Datta about the paper [10], where results more general than Theorem 1 and Corollary 2 were obtained, in a somewhat different form, as well as about the related papers [8, 9]. However, the proofs of Theorem 1 and Corollary 2 in this note seem to be shorter and more self-contained.

REFERENCES

[1] S. ARIMOTO, *Information-theoretical considerations on estimation problems*, Information and Control, 19: 181–194, 1971.

[2] K. M. R. AUDENAERT, *A sharp continuity estimate for the von Neumann entropy*, Journal of Physics A: Mathematical and Theoretical, 40 (28): 8127–8136, Jun 2007.

[3] R. BHATIA, *Matrix analysis*, volume 169 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[4] P. M. BLAU AND J. E. SCHWARTZ, *Crosscutting social circles: testing a macrostructural theory of intergroup relations*, Routledge, London, 1st edition, 2018., Description based on CIP data; resource not viewed.

[5] M. FANNES, *A continuity property of the entropy density for spin lattice systems*, Comm. Math. Phys., 31 (4): 291–294, 1973.

[6] M. GELL-MANN AND C. TSALLIS, editors, *Nonextensive entropy – interdisciplinary applications*, Santa Fe Institute Studies in the Sciences of Complexity, Oxford University Press, New York, 2004.

[7] A. GLEASON, *Measures on the closed subspaces of a Hilbert space*, Indiana Univ. Math. J., 6: 885–893, 1957.

[8] E. P. HANSON AND N. DATTA, *Tight uniform continuity bound for a family of entropies*, 2017., arXiv:1707.04249 [quant-ph].

[9] E. P. HANSON AND N. DATTA, *Maximum and minimum entropy states yielding local continuity bound*, J. Math. Phys., 59 (4): 042204, 31, 2018.

[10] E. P. HANSON AND N. DATTA, *Universal proofs of entropic continuity bounds via majorization flow*, 2019, arXiv:1909.06981 [quant-ph].

[11] L. JOST, *Entropy and diversity*, Oikos, 113 (2): 363–375, 2006.

[12] 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, 2009.

[13] M. A. NIELSEN AND I. L. CHUANG, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2000.

[14] I. PINELIS, *Generalized semimodularity: Order statistics*, In *High dimensional probability VIII*, volume 74 of *Progress in Probability*, pages 99–119, Springer, 2019.

[15] I. PINELIS, *Modulus of continuity of the quantum f-entropy with respect to the trace distance*, 2021., arXiv:2107.10112 [quant-ph].

[16] V. SBORDONI, G. ALLEGRUCCI, AND D. CESARONI, *Population structure*, In W. B. White and D. C. Culver, editors, *Encyclopedia of Caves (Second Edition)*, pages 608–618, Academic Press, Amsterdam, second edition, 2012.

[17] C. TSALLIS, *Possible generalization of Boltzmann–Gibbs statistics*, Journal of Statistical Physics, 52 (1): 479–487, Jul 1988.

[18] C. TSALLIS, *Introduction to nonextensive statistical mechanics: approaching a complex world*, Springer, New York, NY, 2010.

[19] Z. ZHANG, *Uniform estimates on the Tsallis entropies*, Lett. Math. Phys., 80 (2): 171–181, 2007.

Iosif Pinelis
Department of Mathematical Sciences
Michigan Technological University
Houghton, Michigan 49931, USA
e-mail: ipinelis@mtu.edu

Mathematical Inequalities & Applications
www.ele-math.com
mia@ele-math.com