MONOTONICITY OF A RELATIVE RÉNYI ENTROPY

RUPERT L. FRANK AND ELLIOTT H. LIEB

Abstract. We show that a recent definition of relative Rényi entropy is monotone under completely positive, trace preserving maps. This proves a recent conjecture of Müller–Lennert et al.

Recently, Müller–Lennert et al. [12] and Wilde et al. [15] modified the traditional notion of relative Rényi entropy and showed that their new definition has several desirable properties of a relative entropy. One of the fundamental properties of a relative entropy, namely monotonicity under completely positive, trace preserving maps (quantum operations) was shown only in a limited range of parameters and conjectured for a larger range. Our goal here is to prove this conjecture.

More precisely, the definition of the quantum Rényi divergence [12] or sandwiched Rényi entropy [15] is

\[ D_\alpha(\rho\|\sigma) = \begin{cases} \frac{1}{2} \log (\mathrm{Tr} \sigma^{(1-\alpha)/\alpha} \rho \sigma^{(1-\alpha)/(2\alpha)}) & \text{if } \alpha \in (0, 1) \cup (1, \infty), \\
\frac{1}{2} \log (\rho^{1/\alpha} - \sigma^{1/\alpha}) & \text{if } \alpha = 1, \\
\log \|\sigma^{1/2} \rho^{1/2} - \sigma^{1/2} \rho^{1/2}\|_\infty & \text{if } \alpha = \infty \end{cases} \]

for non-negative operators \(\rho, \sigma\). Here, for \(\alpha \geq 1\), we define \(\mathrm{Tr} (\sigma^{(1-\alpha)/\alpha} \rho \sigma^{(1-\alpha)/(2\alpha)})^\alpha\) if the kernel of \(\sigma\) is not contained in the kernel of \(\rho\). The factor \((\mathrm{Tr} \rho)^{-1}\) is inessential and could be dropped, but we keep it in order to be consistent with [12]. After a first version of our paper appeared (arXiv:1306.5358) we were made aware of the fact that \(D_\alpha(\rho\|\sigma)\) is a special case of a two-parameter family of relative entropies introduced earlier in [7].

Note that \(D_\alpha(\rho\|\sigma)\) is the relative von Neumann entropy for \(\alpha = 1\), the relative max-entropy for \(\alpha = \infty\) and closely related to the fidelity \(\mathrm{Tr} (\sigma^{1/2} \rho^{1/2})^{1/2}\) for \(\alpha = 1/2\). In [12] it is shown that \(D_\alpha(\rho\|\sigma)\) depends continuously on \(\alpha\), in particular, at \(\alpha = 1\) and \(\alpha = \infty\).

The definition of \(D_\alpha(\rho\|\sigma)\) should be compared with the traditional relative Rényi entropy (see e.g. [11]),

\[ D'_\alpha(\rho\|\sigma) = (\alpha - 1)^{-1} \log (\mathrm{Tr} \sigma^{1-\alpha} \rho^{1-\alpha}) \quad \text{if } \alpha \in (0, 1) \cup (1, \infty). \]

© 2013 by the authors. This paper may be reproduced, in its entirety, for non-commercial purposes.

U.S. National Science Foundation grants PHY1347399 (R.F.), PHY-0965859 and PHY-1265118 (E.L.) and the Simons Foundation grant 230207 (E.L.) are acknowledged.

October 1, 2013.
Note that by the Lieb–Thirring trace inequality \[9\]
\[
D_\alpha(\rho\|\sigma) \leq D'_\alpha(\rho\|\sigma)
\]
for \(\alpha > 1\).

Our main results in this paper are the following two theorems.

**Theorem 1** (Monotonicity). Let \(1/2 \leq \alpha \leq \infty\) and let \(\rho, \sigma \geq 0\). Then for any completely positive, trace preserving map \(E\),
\[
D_\alpha(\rho\|\sigma) \geq D_\alpha(E(\rho)\|E(\sigma)).
\]

**Theorem 2** (Joint convexity). Let \(1/2 \leq \alpha \leq 1\). Then \(D_\alpha(\rho\|\sigma)\) is jointly convex on pairs \((\rho, \sigma)\) of non-negative operators with \(\text{Tr} \rho = t\) for any fixed \(t > 0\).

For the relative von Neumann entropy \((\alpha = 1)\) both theorems are due to Lindblad \[10\], whose proof is based on Lieb's concavity theorem \[8\]. Theorem \[1\] for \(\alpha \in (1, 2]\) is due to \[12\] and \[15\]. In a preprint of \[12\] its validity was conjectured for all values \(\alpha \geq 1/2\). Shortly after the first version of our paper appeared (arXiv:1306.5358v1) which proved this conjecture for all \(\alpha \geq 1/2\), Beigi independently posted (arXiv:1306.5920) an alternative proof of Theorem \[1\] in the range \(\alpha \in (1, \infty)\).

Just as in Lindblad's monotonicity proof for \(\alpha = 1\), we will deduce Theorem \[1\] for \(\alpha > 1\) from Lieb's concavity theorem \[8\]. The proof for \(1/2 \leq \alpha < 1\) uses a close relative of this theorem, namely, Ando's convexity theorem \[1\]. These theorems enter in the proof of Proposition \[3\] below.

Let us turn to the proofs of the theorems. Both of them are based on the following proposition.

**Proposition 3.** The following map on pairs of non-negative operators
\[
(\rho, \sigma) \mapsto \text{Tr} \left( (1-\alpha)/(2\alpha) \rho(1-\alpha)/(2\alpha) \right)^\alpha
\]
is jointly concave for \(1/2 \leq \alpha < 1\) and jointly convex for \(\alpha > 1\).

We note that this proposition implies that \(\exp((\alpha - 1)D_\alpha(\rho\|\sigma))\) is jointly concave for \(1/2 \leq \alpha < 1\) and jointly convex for \(\alpha > 1\) on pairs \((\rho, \sigma)\) of non-negative operators with \(\text{Tr} \rho = t\) for any fixed \(t > 0\). Since \(x \mapsto x^{1/(\alpha-1)}\) is increasing and convex for \(1 < \alpha \leq 2\), we deduce that \(\exp(D_\alpha(\rho\|\sigma))\) is jointly convex for \(1 < \alpha \leq 2\) on pairs \((\rho, \sigma)\) of non-negative operators with \(\text{Tr} \rho = t\) for any fixed \(t > 0\). This fact is also proved in \[12\] and \[15\].

The argument to derive Theorem \[1\] from Proposition \[3\] is well known, but we include it for the sake of completeness. The fact that joint convexity implies monotonicity appears in \[10\], but here we also use ideas from \[14\].

**Proof of Theorem \[1\] given Proposition \[3\]** We prove the assertion for \(\alpha \in [1/2, 1) \cup (1, \infty)\). The remaining two cases follow by continuity in \(\alpha\). By a limiting argument we may assume that the underlying Hilbert space is \(\mathbb{C}^N\) for some finite \(N\). If \(E\) is a completely positive, trace preserving map then by the Stinespring representation
There is an integer $N' \leq N^2$, a density matrix $\tau$ on $\mathbb{C}^{N'}$ (which can be chosen to be pure) and a unitary $U$ on $\mathbb{C}^N \otimes \mathbb{C}^{N'}$ such that

$$E(\gamma) = \text{Tr}_2 U (\gamma \otimes \tau) U^*.$$ 

Thus, if $du$ denotes normalized Haar measure on all unitaries on $\mathbb{C}^{N'}$, then

$$E(\gamma) \otimes (N')^{-1} 1_{\mathbb{C}^{N'}} = \int (1 \otimes u) U (\gamma \otimes \tau) U^* (1 \otimes u^*) du.$$ (1)

By the tensor property of $D_\alpha(\cdot \| \cdot)$,

$$D_\alpha(E(\rho) \| E(\sigma)) = D_\alpha(E(\rho) \otimes (N')^{-1} 1_{\mathbb{C}^{N'}} \| E(\sigma) \otimes (N')^{-1} 1_{\mathbb{C}^{N'}}).$$ (2)

By (1) and Proposition 3 the double, normalized $u$ integral in (2) is bounded from below (if $1/2 \leq \alpha < 1$) or above (if $\alpha > 1$) by a single integral:

$$\int D_\alpha((1 \otimes u) U (\rho \otimes \tau) U^* (1 \otimes u^*)) \| (1 \otimes u) U (\sigma \otimes \tau) U^* (1 \otimes u^*)) du$$

$$= \int D_\alpha(\rho \otimes \tau \| \sigma \otimes \tau) du$$

$$= D_\alpha(\rho \otimes \tau).$$

Here, we used the unitary invariance of $D_\alpha(\cdot \| \cdot)$, the normalization of the Haar measure and the tensor property of $D_\alpha(\cdot \| \cdot)$.

Dividing the inequality we have obtained by $\text{Tr} E(\rho) = \text{Tr} \rho$, taking logarithms and multiplying by $\alpha - 1$ we obtain the monotonicity stated in the theorem. □

**Proof of Theorem 2 given Proposition 3.** This follows immediately from Proposition 3 together with the fact that $x \mapsto \log x$ is increasing and concave. □

Thus, we have reduced the proofs of Theorems 1 and 2 to the proof of Proposition 3. The latter, in turn, is based on two ingredients. The first one is a representation formula for $\text{Tr} (\sigma^{(1-\alpha)/(2\alpha)} \rho \sigma^{(1-\alpha)/(2\alpha)})^\alpha$.

**Lemma 4.** Let $\rho, \sigma \geq 0$ be operators. Then, if $\alpha > 1$,

$$\text{Tr} (\sigma^{(1-\alpha)/(2\alpha)} \rho \sigma^{(1-\alpha)/(2\alpha)})^\alpha = \sup_{H \geq 0} \left( \alpha \text{Tr} H \rho - (\alpha - 1) \text{Tr} \left( H^{1/2} \sigma^{(\alpha-1)/\alpha} H^{1/2} \right)^{\alpha/(\alpha-1)} \right).$$

The same equality holds for $0 < \alpha < 1$, provided $\sup$ is replaced by inf.

The second ingredient in the proof of Proposition 3 is a concavity result for $\text{Tr} (B^* A^p B)^{1/p}$.

**Lemma 5.** For a fixed operator $B$, the map on positive operators

$$A \mapsto \text{Tr} (B^* A^p B)^{1/p}$$

is concave for $-1 \leq p \leq 1$, $p \neq 0$. 
The case $0 < p \leq 1$ in this lemma is due to Epstein [6], with an alternative proof due to Carlen–Lieb [5] based on the Lieb concavity theorem [8]. Legendre transforms, similar to Lemma 4, are also used in [5].

The remaining case $-1 \leq p < 0$ can be proved similarly, using Ando’s convexity theorem [1], as in [5]. (For an introduction to both theorems we refer to [4].) While this case could easily have been included in [5], it was not, and for the benefit of the reader we explain the argument below. Alternatively, one could probably follow Bekjan’s adaption [2] of Epstein’s proof to establish the $-1 \leq p < 0$ case.

Proof of Proposition 3 given Lemmas 4 and 5. Lemma 5 implies that

$$
\sigma \mapsto (1 - \alpha) \text{Tr} \left( H^{1/2} \sigma^{(\alpha-1)/\alpha} H^{1/2} \right)^{\alpha/(\alpha-1)}
$$

is concave for $1/2 \leq \alpha < 1$ and convex for $\alpha > 1$. The claim of the proposition now follows from the representation formula in Lemma 4.

It remains to prove the lemmas.

Proof of Lemma 4. Let $\alpha > 1$ and abbreviate $\beta = (\alpha - 1)/(2\alpha)$. Since $H^{1/2} \sigma^{2\beta} H^{1/2}$ and $\sigma^\beta H \sigma^\beta$ have the same non-zero eigenvalues, the right side of the lemma is the same as

$$
\sup_{H \geq 0} \left( \alpha \text{Tr} \ H \rho - (\alpha - 1) \text{Tr} \left( \sigma^\beta H \sigma^\beta \right)^{1/(2\beta)} \right).
$$

Let us show that the supremum is given by $\text{Tr} \left( \sigma^{-\beta} \rho \sigma^{-\beta} \right)^{\alpha}$. To prove this, we may assume (by continuity) that $\sigma$ is positive and we observe that the supremum is attained (at least if the underlying Hilbert space is finite-dimensional, which we may assume again by an approximation argument). The Euler–Lagrange equation for the optimal $\hat{H}$ reads

$$
\alpha \rho - \alpha \sigma^\beta \left( \sigma^\beta \hat{H} \sigma^\beta \right)^{1/(\alpha-1)} \sigma^\beta = 0,
$$

that is,

$$
\hat{H} = \sigma^{-\beta} \left( \sigma^{-\beta} \rho \sigma^{-\beta} \right)^{\alpha-1} \sigma^{-\beta}.
$$

By inserting this into the expression we wish to maximize, we obtain $\text{Tr} \left( \sigma^{-\beta} \rho \sigma^{-\beta} \right)^{\alpha}$, as claimed. The proof for $0 < \alpha < 1$ is similar.

We are grateful to the referee for suggesting the following alternative proof of Lemma 4 for $\alpha > 1$. Recall that for positive operators $X$ and $Y$ and $1 < p, q < \infty$ with $1/p + 1/q = 1$ one has

$$
\text{Tr} \ XY \leq \frac{1}{p} \text{Tr} \ X^p + \frac{1}{q} \text{Tr} \ Y^q,
$$

with equality if $X^p = Y^q$. This implies the statement of the lemma, if we set $X = \sigma^{-\beta} \rho \sigma^{-\beta}$, $Y = \sigma^\beta H \sigma^\beta$ and $p = \alpha$, $q = \alpha/(\alpha - 1)$. 

□
Proof of Lemma 5. As we have already mentioned, the result for $0 < p \leq 1$ is known [6, 5]. Therefore, we only give the proof for $-1 \leq p < 0$ and for this we adapt the argument of [5]. We note that

$$ p \text{Tr} \left( B^* A^p B \right)^{1/p} = \inf_{X \geq 0} \left( \text{Tr} A^{p/2} B X^{1-p} B^* A^{p/2} - (1 - p) \text{Tr} X \right). $$

(The proof is similar to that of Lemma 4.) If we can prove that $(A, X) \mapsto \text{Tr} A^{p/2} B X^{1-p} B^* A^{p/2}$ is jointly convex on pairs of non-negative operators, then $p \text{Tr} \left( B^* A^p B \right)^{1/p}$ as an infimum over jointly convex functions is convex, (see [5, Lemma 2.3]) which implies the lemma.

To prove that (3) is jointly convex, we write, as in [8],

$$ \text{Tr} A^{p/2} B X^{1-p} B^* A^{p/2} = \text{Tr} Z^p K^* Z^{1-p} K, $$

where

$$ K = \begin{pmatrix} 0 & 0 \\ B^* & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} A & 0 \\ 0 & X \end{pmatrix}. $$

We can consider $K$, which is an operator in $\mathcal{H} \oplus \mathcal{H}$, as a vector in $(\mathcal{H} \oplus \mathcal{H}) \otimes (\mathcal{H} \oplus \mathcal{H})$ and write $\tilde{K}$. Thus,

$$ \text{Tr} Z^p K^* Z^{1-p} K = \langle \tilde{K}, Z^p \otimes Z^{1-p} \tilde{K} \rangle. $$

By Ando’s convexity theorem [1], the right side is a convex function of $Z$. This is equivalent to (3) being jointly convex, as we set out to prove. \(\square\)

Remark 6. More generally, for a fixed operator $B$, $A \mapsto \text{Tr} \left( B^* A^p B \right)^{q/p}$ is concave on non-negative operators for $0 < |p| \leq q \leq 1$. The case $p > 0$ is due to Carlen–Lieb [5] and the case $p < 0$ follows from similar arguments. More precisely, we can write

$$ r \text{Tr} \left( B^* A^p B \right)^{q/p} = \inf_{X \geq 0} \left( \text{Tr} A^{p/2} B X^{1-r} B^* A^{p/2} - (1 - r) \text{Tr} X \right) $$

with the notation $r = p/q < 0$. Since

$$ \text{Tr} A^{p/2} B X^{1-r} B^* A^{p/2} = \text{Tr} Z^p K^* Z^{1-r} K $$

with $Z$ and $K$ as in the previous proof, the more general assertion again follows from Ando’s convexity theorem [1].

Acknowledgements. We thank E. Carlen, V. Jaksic, C.-A. Pillet and A. Vershynina for valuable comments on a first draft of this paper. We are grateful to the referee for various suggestions that helped to improve this paper.
References

[1] T. Ando, Convexity of certain maps on positive definite matrices and applications to Hadamard products. Lin. Alg. and Appl. 26 (1979), 203–241.

[2] T. Bekjan, On joint convexity of trace functions. Lin. Alg. and Appl. 390 (2004), 321–327.

[3] S. Beigi, Quantum Rényi divergence satisfies data processing inequality. arXiv:1306.5920.

[4] E. A. Carlen, Trace inequalities and quantum entropy. An introductory course. In: Entropy and the quantum, R. Sims and D. Ueltschi (eds.), 73–140, Contemp. Math. 529, Amer. Math. Soc., Providence, RI, 2010.

[5] E. A. Carlen, E. H. Lieb, A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity. Lett. Math. Phys. 83 (2008), 107–126.

[6] H. Epstein, Remarks on two theorems of E. Lieb. Commun. Math. Phys. 31 (1973), 317–325.

[7] V. Jaksic, Y. Ogata, Y. Pautrat, C.-A. Pillet, Entropic fluctuations in quantum statistical mechanics. An Introduction. In: Quantum Theory from Small to Large Scales: Lecture Notes of the Les Houches Summer School: Volume 95, August 2010, Oxford University Press, 2012

[8] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture. Adv. in Math. 11 (1973), 267–288.

[9] E. H. Lieb, W. Thirring, Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities. In: Studies in mathematical physics: essays in honor of Valentine Bargman (E. H. Lieb, B. Simon, A. S. Wightman, editors), pp. 269–297. Princeton University Press, Princeton, 1976.

[10] G. Lindblad, Expectations and entropy inequalities for finite quantum systems. Comm. Math. Phys. 39 (1974), 111–119.

[11] M. Mosonyi, F. Hiai, On the quantum Rényi relative entropies and related capacity formulas. IEEE Transactions on Information Theory, 57 (2011), no. 4, 2474–2487.

[12] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, M. Tomamichel, On quantum Rényi entropies: a new definition, some properties and several conjectures. arXiv:1306.3142 see also arXiv:1306.3142v1.

[13] W. F. Stinespring, Positive functions on C*-algebras. Proc. Amer. Math. Soc. 6 (1955), 211–216.

[14] A. Uhlmann, Endlich-dimensionale Dichtematrizen II. Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R. 22 (1973), 139–177.

[15] M. M. Wilde, A. Winter, D. Yang, Strong converse for the classical capacity of entanglement-breaking channels. arXiv:1306.1586

Rupert L. Frank, Mathematics 253-37, Caltech, Pasadena, CA 91125, USA
E-mail address: rlfrank@caltech.edu

Elliott H. Lieb, Departments of Mathematics and Physics, Princeton University, Washington Road, Princeton, NJ 08544, USA
E-mail address: lieb@princeton.edu