

## Monotonicity of a relative Rényi entropy

Rupert L. Frank and Elliott H. Lieb

Citation: *Journal of Mathematical Physics* **54**, 122201 (2013); doi: 10.1063/1.4838835

View online: <http://dx.doi.org/10.1063/1.4838835>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jmp/54/12?ver=pdfcov>

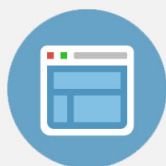
Published by the [AIP Publishing](#)

---



## Re-register for Table of Content Alerts

Create a profile.



Sign up today!





## Monotonicity of a relative Rényi entropy

Rupert L. Frank<sup>1,a)</sup> and Elliott H. Lieb<sup>2,b)</sup>

<sup>1</sup>*Mathematics 253-37, Caltech, Pasadena, California 91125, USA*

<sup>2</sup>*Departments of Mathematics and Physics, Princeton University, Washington Road, Princeton, New Jersey 08544, USA*

(Received 15 July 2013; accepted 10 October 2013; published online 13 December 2013)

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.* [“On quantum Rényi entropies: A new definition, some properties,” *J. Math. Phys.* **54**, 122203 (2013); e-print [arXiv:1306.3142v1](http://arxiv.org/abs/1306.3142v1); see also e-print [arXiv:1306.3142](http://arxiv.org/abs/1306.3142)]. © 2013 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4838835>]

Recently, Müller-Lennert *et al.*<sup>12</sup> and Wilde *et al.*<sup>15</sup> 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*<sup>12</sup> or *sandwiched Rényi entropy*<sup>15</sup> is

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

for non-negative operators  $\rho, \sigma$ . Here, for  $\alpha \geq 1$ , we define  $\text{Tr} \left( \sigma^{(1-\alpha)/\alpha} \rho \sigma^{(1-\alpha)/\alpha} \right)^\alpha = \infty$  if the kernel of  $\sigma$  is not contained in the kernel of  $\rho$ . The factor  $(\text{Tr } \rho)^{-1}$  is inessential and could be dropped, but we keep it in order to be consistent with Ref. 12. After a first version of our paper appeared ([arXiv:1306.5358](http://arxiv.org/abs/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 Ref. 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  $\text{Tr} \left( \sigma^{1/2} \rho \sigma^{1/2} \right)^{1/2}$  for  $\alpha = 1/2$ . In Ref. 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., Ref. 11),

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

Note that by the Lieb–Thirring trace inequality<sup>9</sup>

$$D_\alpha(\rho\|\sigma) \leq D'_\alpha(\rho\|\sigma) \quad \text{for } \alpha > 1.$$

a)E-mail: [rfrank@caltech.edu](mailto:rfrank@caltech.edu)

b)E-mail: [lieb@princeton.edu](mailto:lieb@princeton.edu)

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  $\mathcal{E}$ ,*

$$D_\alpha(\rho\|\sigma) \geq D_\alpha(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)).$$

**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,<sup>10</sup> whose proof is based on Lieb's concavity theorem.<sup>8</sup> Theorem 1 for  $\alpha \in (1, 2]$  is due to Refs. 12 and 15. In a preprint of Ref. 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<sup>3</sup> 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.<sup>8</sup> The proof for  $1/2 \leq \alpha < 1$  uses a close relative of this theorem, namely, Ando's convexity theorem.<sup>1</sup> 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*

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

*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 Refs. 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 Ref. 10, but here we also use ideas from Ref. 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  $\mathcal{E}$  is a completely positive, trace preserving map then by the Stinespring representation theorem<sup>13</sup> 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

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

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

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

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

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

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:

$$\begin{aligned} & \int D_\alpha((1 \otimes u)U(\rho \otimes \tau)U^*(1 \otimes u^*)\|(1 \otimes u)U(\sigma \otimes \tau)U^*(1 \otimes u^*)) du \\ &= \int D_\alpha(\rho \otimes \tau\|\sigma \otimes \tau) du \\ &= D_\alpha(\rho \otimes \tau\|\sigma \otimes \tau) \\ &= D_\alpha(\rho\|\sigma). \end{aligned}$$

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

*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.  $\square$

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  $\alpha > 1$ ,

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

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

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

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

is concave for  $-1 \leq p \leq 1$ ,  $p \neq 0$ .

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

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

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

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.  $\square$

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

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

Let us show that the supremum is given by  $\operatorname{Tr} (\sigma^{-\beta}\rho\sigma^{-\beta})^\alpha$ . 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

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

that is,

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

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

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

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

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.<sup>5,6</sup> Therefore, we only give the proof for  $-1 \leq p < 0$  and for this we adapt the argument of Ref. 5. We note that

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

(The proof is similar to the proof of Lemma 4.) If we can prove that

$$(A, X) \mapsto \operatorname{Tr} A^{p/2} B X^{1-p} B^* A^{p/2} \tag{3}$$

is jointly convex on pairs of non-negative operators, then  $p \operatorname{Tr} (B^* A^p B)^{1/p}$  as an infimum over jointly convex functions is convex (see Lemma 2.3 of Ref. 5), which implies the lemma.

To prove that (3) is jointly convex, we write, as in Ref. 8,

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

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,

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

By Ando’s convexity theorem,<sup>1</sup> 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  $B$ ,  $A \mapsto \operatorname{Tr} (B^* A^p B)^{q/p}$  is concave on non-negative operators for  $0 < |p| \leq q \leq 1$ . The case  $p > 0$  is due to Carlen–Lieb<sup>5</sup> and the case  $p < 0$

follows from similar arguments. More precisely, we can write

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

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

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

with  $Z$  and  $K$  as in the previous proof, the more general assertion again follows from Ando's convexity theorem.<sup>1</sup>

## ACKNOWLEDGMENTS

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. U.S. National Science Foundation Grants Nos. PHY-1347399 (R.L.F.), PHY-0965859 and PHY-1265118 (E.H.L.), and the Simons Foundation Grant No. 230207 (E.H.L.) are acknowledged.

- <sup>1</sup>T. Ando, "Convexity of certain maps on positive definite matrices and applications to Hadamard products," *Linear Algebra Appl.* **26**, 203–241 (1979).
- <sup>2</sup>T. Bekjan, "On joint convexity of trace functions," *Linear Algebra Appl.* **390**, 321–327 (2004).
- <sup>3</sup>S. Beigi, "Sandwiched Rényi divergence satisfies data processing inequality," *J. Math. Phys.* **54**, 122202 (2013); e-print [arXiv:1306.5920](https://arxiv.org/abs/1306.5920).
- <sup>4</sup>E. A. Carlen, "Trace inequalities and quantum entropy. An introductory course," in *Entropy and the Quantum*, Contemporary Mathematics Vol. 529 (American Mathematical Society, Providence, RI, 2010), pp. 73–140.
- <sup>5</sup>E. A. Carlen and E. H. Lieb, "A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity," *Lett. Math. Phys.* **83**, 107–126 (2008).
- <sup>6</sup>H. Epstein, "Remarks on two theorems of E. Lieb," *Commun. Math. Phys.* **31**, 317–325 (1973).
- <sup>7</sup>V. Jaksic, Y. Ogata, Y. Pautrat, and 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 Vol. 95 (Oxford University Press, 2012).
- <sup>8</sup>E. H. Lieb, "Convex trace functions and the Wigner-Yanase-Dyson conjecture," *Adv. Math.* **11**, 267–288 (1973).
- <sup>9</sup>E. H. Lieb and 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*, edited by E. H. Lieb, B. Simon, and A. S. Wightman (Princeton University Press, Princeton, 1976), pp. 269–297.
- <sup>10</sup>G. Lindblad, "Expectations and entropy inequalities for finite quantum systems," *Commun. Math. Phys.* **39**, 111–119 (1974).
- <sup>11</sup>M. Mosonyi and F. Hiai, "On the quantum Rényi relative entropies and related capacity formulas," *IEEE Trans. Inf. Theory* **57**(4), 2474–2487 (2011).
- <sup>12</sup>M. Müller-Lennert *et al.*, "On quantum Rényi entropies: a new definition, some properties," *J. Math. Phys.* **54**, 122203 (2013); e-print [arXiv:1306.3142v1](https://arxiv.org/abs/1306.3142v1); see also e-print [arXiv:1306.3142](https://arxiv.org/abs/1306.3142).
- <sup>13</sup>W. F. Stinespring, "Positive functions on C\*-algebras," *Proc. Am. Math. Soc.* **6**, 211–216 (1955).
- <sup>14</sup>A. Uhlmann, "Endlich-dimensionale Dichtematrizen II," *Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R.* **22**, 139–177 (1973).
- <sup>15</sup>M. M. Wilde, A. Winter, and D. Yang, "Strong converse for the classical capacity of entanglement-breaking channels," e-print [arXiv:1306.1586](https://arxiv.org/abs/1306.1586).