Upper continuity bound on the quantum quasi-relative entropy

Anna Vershynina

Department of Mathematics, Philip Guthrie Hoffman Hall, University of Houston, 3551 Cullen Blvd., Houston, TX 77204-3008, USA

April 2, 2019

Abstract

We provide an upper bound on the quasi-relative entropy in terms of the trace distance. The bound is derived for several types of the function, as well for any operator monotone decreasing function and mixed qubit states. We apply the result to the Umegaki relative entropy and the $q$-entropy.

1 Introduction

Quantum quasi-relative entropy was introduced by Petz [8, 9] as a quantum generalization of a classical Csiszár’s $f$-divergence [4]. It is defined in the context of von Neumann algebras, but we consider only the Hilbert space setup. Let $\mathcal{H}$ be a finite-dimensional Hilbert space, $\rho$ and $\sigma$ be two states (given by density operators), $K$ be an operator on $\mathcal{H}$, and $f : (0, \infty) \to \mathbb{R}$ be an operator convex function. Then the quasi-relative entropy is defined as

$$S_f(\rho||\sigma) = \text{Tr}(f(\Delta_{\sigma,\rho})\rho),$$

where $\Delta_{\sigma,\rho}$ is a relative modular operator defined by Araki [1] that acts as a left and right multiplication

$$\Delta_{A,B}(X) = L_AR_B^{-1}(X) = AXB^{-1}.$$ 

Throughout the paper we consider $\rho$ and $\sigma$ to be strictly positive density operators.

Note that taking the logarithmic function $f(x) = -\log(x)$ reduce quasi-relative entropy to the Umegaki relative entropy [11],

$$S(\rho||\sigma) = \text{Tr}(\rho[\log \rho - \log \sigma]).$$
A famous bound relating quantum relative entropy and trace distance between two quantum states, is called Pinsker inequality. The similar inequality holds for the quasi-relative entropy as well, as was shown by Hiai and Mosonyi [6]:

\[
\frac{f''(1)}{2}\|\rho - \sigma\|_1^2 \leq S_f(\rho\|\sigma).
\]

The questions now, is to obtain the upper bound on the quasi-relative entropy in terms of the trace distance. The upper continuity bound for Umegaki relative entropy was obtained in [2] in the following form:

\[
S(\rho\|\sigma) \leq (\alpha_\sigma + T) \log(1 + T/\alpha_\sigma) - \alpha_\rho \log(1 + T/\alpha_\rho),
\]

where throughout the paper \(\alpha_\omega\) is the minimal non-zero eigenvalue of the state \(\omega\), and \(T := \|\rho - \sigma\|_1/2\).

For \(q > 1\), a \(q\)-entropy is defined by

\[
S_q(\rho\|\sigma) = \frac{1}{1 - q} \left(1 - \text{Tr}(\rho^q \sigma^{1-q})\right),
\]

(1.1)

for \(\text{ker}(\sigma) \subset \text{ker}(\rho)\). A series of upper bounds for \(q\)-entropy in terms of the trace distance were obtained in [10]. There is an easy way to improve the constant in one of the bounds, see Remark 4.1. The derived bounds in [10] are, in particular, the following:

- for \(q > 1\)
  \[
  S_q(\rho\|\sigma) \leq \left\lceil q \right\rceil - 1 - \frac{\lambda_{q-1}}{\lambda_{q-1}} \|\rho - \sigma\|_1,
  \]

  where \(\lambda\) is the maximum eigenvalue in the joint spectrum of \(\rho\) and \(\sigma\), and \(\left\lceil q \right\rceil\) is the smallest integer that is larger than \(q\);

- for \(1 < q \leq 2\), and denoting \(\lambda_\rho\) to be the maximal eigenvalue of \(\rho\), and \(\alpha = \min\{\alpha_\rho, \alpha_\sigma\}\), the following bounds hold
  \[
  S_q(\rho\|\sigma) \leq \frac{1}{q - 1} \frac{\lambda_q}{\alpha^q} \|\rho - \sigma\|_1.
  \]

  (1.2)

  From Remark 4.1 the bound can easily be improved to
  \[
  S_q(\rho\|\sigma) \leq \frac{\lambda_q}{\alpha^q} \|\rho - \sigma\|_1.
  \]

  (1.3)

- for \(0 < q < 1\),
  \[
  S_q(\rho\|\sigma) \leq \frac{1}{1 - q} \frac{\lambda_q}{\alpha^q} \|\rho - \sigma\|_1.
  \]

  (1.4)

Note that a series of other bounds was derived in [10], which for some states could be an improvement of the bounds above.
We investigate the upper continuity bound for a quasi-relative entropy for an operator monotone decreasing function \( f \). We derive an upper bound for several specific forms of the function, see Theorems 3.1 and Theorem 3.3.

Moreover, we obtain a bound for any operator monotone decreasing function when states \( \rho \) and \( \sigma \) are either 2-dimensional states or simultaneously diagonalizable, see Theorem 3.5, where we obtain

\[
S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \left[ \frac{\lambda_\rho}{\lambda_\rho - \alpha_\sigma} f(\lambda_\rho^{-1} \alpha_\sigma) - a \right],
\]

where

- \( \lambda_\rho \in (0,1] \) is the largest eigenvalue of \( \rho \),
- \( \alpha_\sigma \in (0,1] \) is the smallest eigenvalue of \( \sigma \),
- \( a = -\lim_{y\to\infty} \frac{f(iy)}{iy} \).

In the most general case, we obtain an upper bound dependent on the dimension of the Hilbert space, see Theorem 3.7. We conjecture that the bound (1.5) holds in the general case as well, see Conjecture 3.8.

Applying the result of Theorems 3.3 to the function \( f(x) = -\log x \), we obtain an upper bound on the relative entropy

\[
S(\rho\|\sigma) \leq \frac{\lambda_\rho}{\alpha} \|\rho - \sigma\|_1,
\]

where \( \alpha = \min\{\alpha_\rho, \alpha_\sigma\} \). If states are two-dimensional, from (1.5) we obtain

\[
S(\rho\|\sigma) \leq \max\left\{1, \frac{\lambda_\rho}{\alpha_\sigma} \right\} \|\rho - \sigma\|_1.
\]

Moreover, for \( p \in (-1,1) \) and \( p \neq 0 \) taking the function \( f_p(x) := \frac{1}{p(1-p)}(1 - x^p) \) in Theorems 3.1 results in the quasi-relative entropy of the form

\[
S_{f_p}(\rho\|\sigma) = \frac{1}{p(1-p)} \left(1 - \text{Tr}(\rho^{1-p} \sigma^p) \right).
\]

This is a renormalized \( q \)-entropy (1.1) for \( q = 1 - p \in (0,2), q \neq 1 \). Note that the factor \( p(1-p) \) is important here, as it changes the sign at \( p = 0 \). As a corollary, we obtain, for \( p \in (-1,1) \)

\[
S_{f_p}(\rho\|\sigma) \leq \frac{1}{1-p} \|\rho - \sigma\|_1 \frac{\lambda_\rho^{1-p}}{\alpha_\sigma^{1-p}}.
\]

This bound matches the bound (1.3) exactly for \( q \in (1,2) \) or \( p \in (-1,0) \). For \( q \in (0,1) \) or \( p \in (0,1) \), if \( \alpha = \alpha_\sigma \), our bound improves the bound (1.4) as it improves the constant, since the \( q \)-entropy (1.1) does not contain a pre-factor \( \frac{1}{q} = \frac{1}{1-p} \geq 1 \).

If the states \( \rho \) and \( \sigma \) are two-dimensional, then from Theorem 3.7 we obtain

\[
S_{f_p}(\rho\|\sigma) \leq \frac{1}{1-p} \|\rho - \sigma\|_1 \max\left\{1, \frac{\lambda_\rho^{1-p}}{\alpha_\sigma^{1-p}} \right\}.
\]
2 Preliminaries

2.1 Operator monotone functions

2.1 Definition. A function \( f : (a, b) \to \mathbb{R} \) is operator monotone if for any pair of self-adjoint operators \( A \) and \( B \) on some Hilbert space that have spectrum in \( (a, b) \), the operator

\[
f(A) - f(B) \geq 0
\]

is positive semidefinite whenever \( A - B \geq 0 \) is positive semidefinite. We say that \( f \) is operator monotone decreasing on \( (a, b) \) in case \( -f \) is operator monotone.

2.2 Definition. A function \( f \) is operator concave on the positive operators, when for all positive semidefinite operators \( A \) and \( B \), and all \( \lambda \) in \( (0, 1) \),

\[
f((1 - \lambda)A + \lambda B) - (1 - \lambda)f(A) - \lambda f(B) \geq 0
\]

is positive semidefinite. A function \( f \) is operator convex if \( -f \) is operator concave.

2.3 Theorem (Bhatia ’97). Every operator monotone function \( f : [0, \infty) \to \mathbb{R} \) is operator concave. Moreover, every continuous function \( f \) mapping \( [0, \infty) \to [0, \infty) \) into itself is operator monotone if and only if it is operator concave.

2.4 Example. Note that

- \( f(x) = \log x \) is operator monotone;
- \( f(x) = x \log x \) is operator convex.

2.5 Example. Let \( f(x) = x^p \), where \( p \in \mathbb{R} \). Then by Theorem V.2.10] the function \( f \) is

1. operator monotone and operator concave if and only if \( p \in [0, 1] \);
2. operator convex if and only if \( p \in [-1, 0] \cup [1, 2] \);
3. operator monotone decreasing and operator convex if and only if \( p \in [-1, 0] \).

2.6 Definition. A Pick function is a function \( f \) that is analytic on the upper half plane and has a positive imaginary part. The set of Pick functions on \( (a, b) \) is denoted as \( \mathcal{P}_{(a,b)} \).

2.7 Theorem (Löwner ’34). A function \( f \) on \( (a, b) \) is operator monotone if and only if \( f \) is a restriction of a pick function \( f \in \mathcal{P}_{(a,b)} \) to \( (a, b) \).

2.8 Corollary. A function \( f \) on \( (0, \infty) \) is operator monotone decreasing if and only if \( -f \in \mathcal{P}_{(0,\infty)} \).

Denote the set of operator monotone decreasing functions \( f \) (i.e.\( -f \in \mathcal{P}_{(0,\infty)} \)) as \( \mathcal{Q}_{(0,\infty)} \).

2.9 Example. From Exercise V.4.8] The following functions belong to \( \mathcal{Q}_{(0,\infty)} \):
• $f(x) = -\log x$,
• $f(x) = -x^p$ for $p \in [0, 1]$,
• $f(x) = x^p$ for $p \in [-1, 0]$.

According to [5, Chapter II, Theorem I] every function $f \in \mathcal{Q}(0, \infty)$, has a canonical integral representation

$$f(x) = -ax - b + \int_0^\infty \left( \frac{1}{t + x} - \frac{t}{t^2 + 1} \right) \, d\mu_f(t),$$

where $a := -\lim_{y \to \infty} \frac{f(iy)}{iy} \geq 0$, $b := -\text{Re}(f(i)) \in \mathbb{R}$ and $\mu$ is a positive measure on $(0, \infty)$ such that

$$\int_0^\infty \frac{1}{t^2 + 1} \, d\mu_f(t) < \infty,$$

and

$$\mu_f(x_1) - \mu_f(x_0) = -\lim_{y \downarrow 0} \frac{1}{\pi} \int_{x_0}^{x_1} \text{Im}(f(-x + iy)) \, dx.$$  \hspace{1cm} (2.2)

Conversely, every such function belongs in $\mathcal{Q}(0, \infty)$.

We consider functions $f \in \mathcal{Q}(0, \infty)$ such that $f(1) = 0$. The last condition is equivalent to

$$0 = f(1) = -a - b + \int_0^\infty \left( \frac{1}{t + 1} - \frac{t}{t^2 + 1} \right) \, d\mu_f(t),$$

in other words,

$$a + b = \int_0^\infty \left( \frac{1}{t + 1} - \frac{t}{t^2 + 1} \right) \, d\mu_f(t).$$ \hspace{1cm} (2.3)

Therefore, the operator monotone decreasing function $f$ such that $f(1) = 0$ has the following integral representation

$$f(x) = a(1-x) + \int_0^\infty \left( \frac{1}{t + x} - \frac{1}{t + 1} \right) \, d\mu_f(t).$$ \hspace{1cm} (2.4)

2.10 Example. Consider the power function $f(x) = -x^p$ for $p \in (0, 1)$. It is operator monotone decreasing. Then

$$a = -\lim_{y \to \infty} \frac{f(iy)}{iy} = 0, \hspace{0.5cm} \text{and} \hspace{0.5cm} b = \cos(p\pi/2).$$

For $x > 0$, $\lim_{y \downarrow 0} \, \text{Im}(f(-x + iy)) = -x^p \sin(p\pi)$ so that

$$d\mu(x) = \pi^{-1} \sin(p\pi) x^p dx.$$ 

This yields the representation

$$-x^p = -\cos(p\pi/2) + \frac{\sin(p\pi)}{\pi} \int_0^\infty t^p \left( \frac{1}{t + x} - \frac{t}{t^2 + 1} \right) \, dt.$$ \hspace{1cm} (2.5)
2.11 Example. Let \( f(x) = -\log(x) \). It is operator monotone decreasing. Then
\[
\begin{align*}
  b &= \Re(\log(i)) = 0 , \\
  a &= \lim_{y \to \infty} \log(iy)/(iy) = \lim_{y \to \infty} (\log y + i\pi/2)/(iy) = 0 .
\end{align*}
\]
It is clear from (2.2) that
\[
\text{d}\mu(x) = \frac{1}{\pi} \lim_{y \to 0} \Im \log(-x + iy) \text{d}x = \text{d}x .
\]
Then the integral representation (2.1) gives the following formula for the logarithmic function
\[
-\log x = \int_0^\infty \left( \frac{1}{t + x} - \frac{t}{t^2 + 1} \right) \text{d}t .
\] (2.6)

2.2 Quasi-relative entropy

2.12 Definition. For an operator convex function \( f \), such that \( f(1) = 0 \), and strictly positive states \( \rho \) and \( \sigma \) acting on a finite-dimensional Hilbert space \( \mathcal{H} \), \( f \)-devergence is defined as
\[
S_f(\rho||\sigma) = \text{Tr}(f(\Delta_{\sigma,\rho})\rho) ,
\]
where the modular operator, introduced by Araki \[1\],
\[
\Delta_{A,B}(X) = L_A R_B^{-1}(X) = AXB^{-1}
\]
is a product of left and right multiplication operators, \( L_A(X) = AX \) and \( R_B(X) = XB \).

There right and left multiplication operators have the following properties \[7\]
1. They commute, i.e.
\[
[L_A, R_B] = 0 ,
\]
since
\[
L_A R_B(X) = AXB = R_B L_A(X) .
\]
2. The operators \( L_A \) and \( R_A \) are invertible if and only if \( A \) is non-singular, giving \( L_A^{-1} = L_A^{-1} \) and \( R_A^{-1} = R_A^{-1} \).
3. If \( A \) is self-adjoint, then \( L_A \) and \( R_A \) are both self-adjoint with respect to the Hilbert Schmidt inner product.
4. If \( A \geq 0 \), then \( L_A \) and \( R_A \) are positive semi-definite, i.e.
\[
\text{Tr}X^* L_A(X) = \text{Tr}X^* AX \geq 0
\]
and
\[
\text{Tr}X^* R_A(X) = \text{Tr}X^* AX = \text{Tr}X^* AX \geq 0
\]
5. If $A > 0$, for any function $f(0, \infty) \to \mathbb{R}$, we have $f(L_A) = L_{f(A)}$ and $f(R_A) = R_{f(A)}$. This follows from the spectral decomposition of $A$, denoted as $A = \sum_{j=1}^{d} \lambda_j |j\rangle \langle j|$. Then for any $j, k = 1, \ldots, d$ the operator $|j\rangle \langle k|$ is the eigenstate of the operator $L_A$ (and $R_A$) with eigenvalue $\lambda_j$ (or $\lambda_k$). The later has degeneracy $d$

$$L_A |j\rangle \langle k| = \lambda_j |j\rangle \langle k|, \quad R_A |j\rangle \langle k| = \lambda_k |j\rangle \langle k|.$$ 

Therefore,

$$f(L_A) |j\rangle \langle k| = f(\lambda_j) |j\rangle \langle k|, \quad f(R_A) |j\rangle \langle k| = f(\lambda_k) |j\rangle \langle k|.$$ 

There is a straightforward way to calculate the quasi-relative entropy from the spectral decomposition of states. Let $\rho$ and $\sigma$ have the following spectral decomposition

$$\rho = \sum_{j} \lambda_j |\psi_j\rangle \langle \psi_j|, \quad \sigma = \sum_{k} \mu_k |\phi_k\rangle \langle \phi_k|,$$

where the eigenvalues are ordered:

$$\lambda_n \leq \cdots \leq \lambda_1, \quad \mu_n \leq \cdots \leq \mu_1.$$

the set $\{|\phi_k\rangle \langle \psi_j|\}_{j,k}$ forms an orthonormal basis of $\mathcal{B}(\mathcal{H})$, the space of bounded linear operators, with respect to the Hilbert-Schmidt inner product defined as $\langle A, B \rangle = \text{Tr}(A^*B)$. By [12], the modular operator can be written as

$$\Delta_{\sigma,\rho} = \sum_{j,k} \frac{\mu_k}{\lambda_j} P_{j,k}, \quad (2.8)$$

where $P_{j,k} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is defined by

$$P_{j,k}(X) = |\phi_k\rangle \langle \phi_k| X |\psi_j\rangle.$$

The quasi-relative entropy is calculated as follows

$$S_f(\rho||\sigma) = \sum_{j,k} \lambda_j f \left( \frac{\mu_k}{\lambda_j} \right) |\langle \phi_k| |\psi_j\rangle|^2. \quad (2.9)$$

2.13 Example. For $f(x) = -\log x$, the quasi-relative entropy becomes the Umegaki relative entropy

$$S_{\log}(\rho||\sigma) = S(\rho||\sigma) = \text{Tr}(\rho \log \rho - \rho \log \sigma).$$

2.14 Example. For $p \in (-1, 2)$ and $p \neq 0, 1$ let us take the function

$$f_p(x) := \frac{1}{p(1-p)} (1-x^p),$$

which is operator convex. The quasi-relative entropy for this function is calculated to be

$$S_{f_p}(\rho||\sigma) = \frac{1}{p(1-p)} (1 - \text{Tr}(\sigma^p \rho^{1-p})).$$
3 Upper continuity bound

We consider three types of functions for the quasi-relative entropy, and we derive an upper bound for each of these types. The last type is any operator monotone function decreasing function.

3.1 Theorem. Let $f \in Q_{(0,\infty)}$ be an operator monotone decreasing function such that for any $x, y > 0$ the following holds

$$f(xy) = 1 - g(x)h(y),$$

where $g$ is an operator monotone function or an operator monotone decreasing function, and $g(1) = h(1) = 1$. Let $\rho$ and $\sigma$ be two strictly positive density operators on a finite-dimensional Hilbert space. Then

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \max_j \left\{ \lambda_j |h(\lambda_j^{-1})| \right\} \left| \frac{g(\alpha_\rho) - g(\alpha_\sigma)}{\alpha_\rho - \alpha_\sigma} \right|, \quad (3.1)$$

where

- $\{\lambda_j\}_j$ are the eigenvalues of $\rho$,
- $\alpha_\rho \in (0, 1]$ is the smallest eigenvalue of $\rho$,
- $\alpha_\sigma \in (0, 1]$ is the smallest eigenvalue of $\sigma$,

3.2 Corollary. Under the same conditions as in theorem, if we additionally assume that $g$ is concave if it is operator monotone, and convex if it is operator monotone decreasing, the quasi-relative entropy is bounded by

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \max_j \left\{ \lambda_j |h(\lambda_j^{-1})| \right\} \left| g'(\alpha) \right|, \quad (3.2)$$

where $\alpha = \min\{\alpha_\rho, \alpha_\sigma\}$.

Proof. (of Corollary 3.2) Consider an operator monotone function $g$. Without loss of generality assume that $\alpha \leq \alpha_\sigma$. By the Mean Value Theorem, there exists $c \in [\alpha_\rho, \alpha_\sigma]$ such that

$$\frac{g(\alpha_\rho) - g(\alpha_\sigma)}{\alpha_\rho - \alpha_\sigma} g'(c).$$

Since $g$ is concave, the derivative $g'$ is monotonically decreasing, therefore

$$g'(c) \leq g'(\alpha_\rho).$$

The proof is similar for an operator monotone decreasing function $g$. \qed
Proof. (of Theorem 3.1)

Since $S_f(\rho\|\rho) = 0 = \text{Tr}(f(\Delta_{\rho,\rho})\rho)$, we have

$$S_f(\rho\|\sigma) = \text{Tr}\{f(\Delta_{\sigma,\rho})\rho\} - \text{Tr}(f(\Delta_{\rho,\rho})\rho)$$

(3.3)

$$= \text{Tr}\{(f(L_\sigma R_{\rho^{-1}}) - f(L_\rho R_{\rho^{-1}}))\rho\}$$

(3.4)

$$= \text{Tr}\{(g(L_\sigma) - g(L_\rho))h(R_{\rho^{-1}})\rho\}$$

(3.5)

$$= \text{Tr}\{(g(\sigma) - g(\rho))h(\rho^{-1})\rho\}$$  

(3.6)

Here we used that the modular operator is the product of left and right multiplications, $\Delta_{\sigma,\rho} = L_\sigma R_{\rho^{-1}}$.

Let $g(x)$ be an operator monotone function, such that $g(1) = 1$, then it admits an integral representation that follows from (2.4):

$$g(x) = 1 + a_g(x - 1) - \int_0^\infty \left(\frac{1}{t + x} - \frac{1}{t + 1}\right) d\mu_g(t)$$

(3.7)

where $a_g \geq 0$. Therefore,

$$S_f(\rho\|\sigma) = a_g \text{Tr}\{(\sigma - \rho)h(\rho^{-1})\rho\} + \int_0^\infty \text{Tr}\left\{\left(\frac{1}{t + \rho} - \frac{1}{t + \sigma}\right) h(\rho^{-1})\rho\right\} d\mu_g(t)$$

(3.8)

By (13), for any operators $X, Y, Z$ the following bound holds

$$|\text{Tr}(XYZ)| \leq \|X\|\|Y\|\|Z\|\text{Tr}|Y|$$

(3.9)

For the first term in (3.8), we use the bound above for two operators:

$$|\text{Tr}\{(\sigma - \rho)h(\rho^{-1})\rho\}| \leq \|\rho - \sigma\|\|h(\rho^{-1})\rho\| = \|\rho - \sigma\|\max_j \{\lambda_j|h(\lambda_j^{-1})|\}.$$  

(3.10)

For the second term, we note that the formula $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$ holds for any invertible operators $A$ and $B$. Therefore,

$$\left|\int_0^\infty \text{Tr}\left\{\left(\frac{1}{t + \rho} - \frac{1}{t + \sigma}\right) h(\rho^{-1})\rho\right\} d\mu_g(t)\right| \leq \int_0^\infty \text{Tr}\left|(t + \rho)^{-1}(\sigma - \rho)(t + \sigma)^{-1}h(\rho^{-1})\rho\right| d\mu_g(t).$$

(3.11)

Applying (3.9) and the fact that for two operators $\|XY\| \leq \|X\|\|Y\|$, we have that the last line can be bounded as

$$\leq \int_0^\infty \|\rho - \sigma\|\|t + \rho\|^{-1}\|\sigma - \rho\|^{-1}\|\rho\|d\mu_g(t)$$

(3.12)

$$= \|\rho - \sigma\|\max_j \{\lambda_j|h(\lambda_j^{-1})|\} \int_0^\infty \frac{1}{t + \alpha_\rho} d\mu_g(t)$$

(3.13)

$$= \|\rho - \sigma\|\max_j \{\lambda_j|h(\lambda_j^{-1})|\} \int_0^\infty \left(\frac{1}{t + \alpha_\sigma} - \frac{1}{t + \alpha_\rho}\right) d\mu_g(t).$$

(3.14)
Note that from (3.7) we have

$$\int_0^\infty \left( \frac{1}{t + \alpha_\sigma} - \frac{1}{t + \alpha_\rho} \right) d\mu_g(t) = g(\alpha_\rho) - g(\alpha_\sigma) - a_g(\alpha_\rho - \alpha_\sigma),$$

(3.15)

and therefore the second term can be bounded by

$$\|\rho - \sigma\|_1 \max_j \lambda_j |h(\lambda_j^{-1})| \frac{g(\alpha_\rho) - g(\alpha_\sigma)}{\alpha_\rho - \alpha_\sigma} - a_g\|\rho - \sigma\|_1 \max_j \lambda_j |h(\lambda_j^{-1})|. \quad (3.16)$$

Putting (3.10) and (3.30) together, we find

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \max_j \lambda_j |h(\lambda_j^{-1})| \frac{g(\alpha_\rho) - g(\alpha_\sigma)}{\alpha_\rho - \alpha_\sigma}.$$

The proof is similar for the operator monotone decreasing function $g$. \hfill \square

Now we discuss the second type of the function.

### 3.3 Theorem

Let $f \in Q_{(0,\infty)}$ be an operator monotone decreasing function such that for any $x, y > 0$ the following holds

$$f(xy) = g(x) + h(y),$$

where $g$ and $h$ are operator monotone decreasing function, and $g(1) = h(1) = 0$. Let $\rho$ and $\sigma$ be two strictly positive density operators on a finite-dimensional Hilbert space. Then

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \lfloor \lambda \rfloor |h(\lambda^{-1})| \frac{g(\alpha_\rho) - g(\alpha_\sigma)}{\alpha_\rho - \alpha_\sigma},$$

(3.17)

where

- $\alpha_\rho \in (0, 1]$ is the smallest eigenvalue of $\rho$,
- $\alpha_\sigma \in (0, 1]$ is the smallest eigenvalue of $\sigma$.

The proof of the following corollary is the same as for Corollary 3.2.

### 3.4 Corollary

Under the same conditions as in theorem, if we additionally assume that $g$ is convex, the quasi-relative entropy is bounded by

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \lambda \rho \frac{g(\alpha_\sigma) - g(\alpha_\rho)}{\alpha_\rho - \alpha_\sigma},$$

(3.18)

where $\alpha = \min\{\alpha_\rho, \alpha_\sigma\}$.

**Proof.** (of Theorem 3.3) The proof is very similar to the previous proof.

By the assumption for every $x$,

$$f(1) = 0 = f(xx^{-1}) = g(x) + h(x^{-1}).$$
Therefore, \( h(x^{-1}) = -g(x) \). The quasi-relative entropy can be written as

\[
S_f(\rho||\sigma) = \text{Tr}\{f(\Delta_{\sigma,\rho})\rho\} = \text{Tr}\{(g(L_\sigma) + h(R_{\rho^{-1}}))\rho\} = \text{Tr}\{(g(\sigma) + h(\rho^{-1}))\rho\} = \text{Tr}\{(g(\sigma) - g(\rho))\rho\}.
\]

(3.19) (3.20) (3.21) (3.22)

Since \( g(x) \) is an operator monotone decreasing function, such that \( g(1) = 0 \), it admits an integral representation \[24\]:

\[
g(x) = a_g(1 - x) + \int_0^\infty \left( \frac{1}{t + x} - \frac{1}{t + 1} \right) d\mu_g(t),
\]

(3.23)

where \( a_g \geq 0 \). Therefore,

\[
S_f(\rho||\sigma) = a_g \text{Tr}\{(\rho - \sigma)\rho\} - \int_0^\infty \text{Tr}\left\{ \left( \frac{1}{t + \rho} - \frac{1}{t + \sigma} \right)\rho \right\} d\mu_g(t).
\]

(3.24)

The first term in (3.24) is bounded by \( a_g\|\rho - \sigma\|_1\|\rho\|_\infty = a_g\|\rho - \sigma\|_1\lambda_\rho \), where \( \lambda_\rho \) is the maximal eigenvalue of \( \rho \).

For the second term, we note that the formula \( A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1} \) holds for any invertible operators \( A \) and \( B \). Therefore,

\[
\left| \int_0^\infty \text{Tr}\left\{ \left( \frac{1}{t + \rho} - \frac{1}{t + \sigma} \right)\rho \right\} d\mu_g(t) \right| \leq \int_0^\infty \text{Tr}\left| (t + \rho)^{-1}(\sigma - \rho)(t + \sigma)^{-1}\rho \right| d\mu_g(t).
\]

(3.25)

Applying (3.9) and the fact that for two operators \( \|XY\|_\infty \leq \|X\|_\infty \|Y\|_\infty \), we have that the last line can be bounded as

\[
\leq \int_0^\infty \|\rho - \sigma\|_1(t + \rho)^{-1}\|\sigma + \rho\|_\infty(t + \sigma)^{-1}\|\rho\|_\infty d\mu_g(t)
\]

(3.26)

\[
= \|\rho - \sigma\|_1\lambda_\rho \int_0^\infty \frac{1}{t + \alpha_\sigma} \frac{1}{t + \alpha_\rho} d\mu_g(t)
\]

(3.27)

\[
= \|\rho - \sigma\|_1 \frac{1}{\alpha_\rho - \alpha_\sigma} \lambda_\rho \int_0^\infty \left( \frac{1}{t + \alpha_\sigma} - \frac{1}{t + \alpha_\rho} \right) d\mu_g(t).
\]

(3.28)

Note that from (3.23) we have

\[
\int_0^\infty \left( \frac{1}{t + \alpha_\sigma} - \frac{1}{t + \alpha_\rho} \right) d\mu_g(t) = g(\alpha_\sigma) - g(\alpha_\rho) - a_g(\alpha_\rho - \alpha_\sigma),
\]

(3.29)

and therefore the second term can be bounded by

\[
\|\rho - \sigma\|_1\lambda_\rho \frac{g(\alpha_\sigma) - g(\alpha_\rho)}{\alpha_\rho - \alpha_\sigma} - a_g\|\rho - \sigma\|_1\lambda_\rho.
\]

(3.30)

Putting (3.10) and (3.30) together, we find

\[
S_f(\rho||\sigma) \leq \|\rho - \sigma\|_1\lambda_\rho \frac{g(\alpha_\sigma) - g(\alpha_\rho)}{\alpha_\rho - \alpha_\sigma}.
\]
Consider a case when $\rho$ and $\sigma$ are 2-dimensional states, or diagonalizable in the same basis states on $d$-dimensional Hilbert space. Then for any operator monotone decreasing function $f$ the following upper bound holds.

**3.5 Theorem.** Let $f \in \mathcal{Q}_{(0,\infty)}$ be an operator monotone decreasing function such that $f(1) = 0$. Let $\rho$ and $\sigma$ be two strictly positive density operators on a 2-dimensional Hilbert space, or states diagonalizable in the same basis on any finite-dimensional Hilbert space. Then

\[
S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \left[ \frac{\lambda_\rho}{\lambda_\rho - \alpha_\sigma} f(\lambda_\rho^{-1}\alpha_\sigma) - a \right],
\]

where

- $\lambda_\rho \in (0,1]$ is the largest eigenvalue of $\rho$,
- $\alpha_\sigma \in (0,1]$ is the smallest eigenvalue of $\sigma$,
- $a = -\lim_{y \to \infty} \frac{f(iy)}{iy}$.

**Proof.** Every function $f \in \mathcal{Q}_{(0,\infty)}$ (i.e. operator monotone decreasing function), such that $f(1) = 0$ admits an integral representation (2.4). Since $S_f(\rho\|\rho) = 0 = \text{Tr}(f(\Delta_{\rho,\rho})\rho)$, we have

\[
S_f(\rho\|\sigma) = \text{Tr}\{(f(\Delta_{\sigma,\rho}) - f(\Delta_{\rho,\rho}))\rho\}
\]

\[
= a_f \text{Tr}\{\Delta_{\sigma,\rho}\rho\} - a_f \text{Tr}\{\Delta_{\rho,\rho}\rho\} + \int_0^\infty d\mu_f(t) \text{Tr}\{((t I + \Delta_{\sigma,\rho})^{-1} - (t I + \Delta_{\rho,\rho})^{-1})\rho\}
\]

\[
= \int_0^\infty d\mu_f(t) \text{Tr}\{((t I + \Delta_{\sigma,\rho})^{-1} - (t I + \Delta_{\rho,\rho})^{-1})\rho\}.
\]

The formula $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$ holds for any invertible operators $A$ and $B$. Using the fact that the modular operator is the product of left and right multiplications, $\Delta_{\sigma,\rho} = L_{\sigma}R_{\rho}^{-1}$, we obtain

\[
S_f(\rho\|\sigma) = \int_0^\infty d\mu_f(t) \text{Tr}\{((t I + \Delta_{\sigma,\rho})^{-1}(L_{\rho} - L_{\sigma})R_{\rho}^{-1}(t I + \Delta_{\rho,\rho})^{-1})\rho\}
\]

\[
= \int_0^\infty d\mu_f(t) \text{Tr}\{((t I + \Delta_{\sigma,\rho})^{-1}(L_{\rho} - L_{\sigma})(t I + \Delta_{\rho,\rho})^{-1})I\}.
\]

From (2.8), the last trace can be written as a trace of a product of two matrices:

\[
S_f(\rho\|\sigma) = \int_0^\infty d\mu_f(t) (t + 1)^{-1}\text{Tr}\{D_t(\rho - \sigma)\},
\]

where, with the spectral decomposition (2.7) of $\rho$ and $\sigma$,

\[
D_t = \sum_{jk} \left(t + \frac{\mu_k}{\lambda_j}\right)^{-1} \langle \psi_j| |\phi_k\rangle |\psi_j\rangle \langle \phi_k|.
\]
In Lemma 3.6 take $X = \rho - \sigma$, $D = D_t$ and $C = \max_{kj} \left( t + \frac{\mu}{\lambda_j} \right)^{-1} = (t + \lambda_\rho^{-1} \alpha_\sigma)^{-1}$. Then in both cases for states $\rho$ and $\sigma$ specified in Theorem, we obtain

$$|\text{Tr}\{D_t(\rho - \sigma)\}| \leq (t + \lambda_\rho^{-1} \alpha_\sigma)^{-1} \|\rho - \sigma\|_1.$$  

Therefore, in both cases,

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \int_0^\infty \frac{1}{t + \lambda_\rho^{-1} \alpha_\sigma} \cdot \frac{1}{t + 1} d\mu_f(t),$$

Note that

$$\frac{1}{t + \lambda_\rho^{-1} \alpha_\sigma} \cdot \frac{1}{t + 1} = \frac{\lambda_\rho}{\lambda_\rho - \alpha_\sigma} \left\{ \frac{1}{t + \lambda_\rho^{-1} \alpha_\sigma} - \frac{1}{t + 1} \right\}.$$  

Therefore,

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \frac{\lambda_\rho}{\lambda_\rho - \alpha_\sigma} \left[ f(\lambda_\rho^{-1} \alpha_\sigma) - a_f(1 - \lambda_\rho^{-1} \alpha_\sigma) \right],$$

where the last equation is obtained from the integral representation (2.4) of function $f$. Here, recall, $a_f = -\lim_{y \to \infty} \frac{f(iy)}{iy}$.

3.6 Lemma. For orthogonal bases $\{|\psi_j\rangle\}$ and $\{|\phi_k\rangle\}$, let

$$D = \sum_{kj} C_{kj} \langle \psi_j | \phi_k \rangle \langle \psi_j | \phi_k \rangle,$$  

such that $0 \leq C_{kj} \leq C$ for all $k, j$ and some $C$. Consider two cases:

- Let $X$ be a diagonal matrix in either basis: without loss of generality let $X = \sum_k x_k |\phi_k\rangle \langle \phi_k|$.  
- Let $X$ be a $2 \times 2$ Hermitian traceless matrix, i.e. $X^* = X$ and $\text{Tr}(X) = 0$.

In both cases,

$$|\text{Tr}(DX)| \leq C\|X\|_1.$$  

Proof. 1. Since $X$ is diagonal matrix, the trace norm is the sum of the absolute values of the eigenvalues

$$\|X\|_1 = \text{Tr}|X| = \sum_j |x_j|.$$  

The trace can be calculated

$$\text{Tr}\{DX\} = \sum_{kj} C_{kj} x_k \langle \psi_j | \phi_k \rangle \langle \psi_j | \phi_k \rangle.$$
Therefore,

\[ |\text{Tr}\{DX\}| \leq \sum_{kj} C_{kj} |x_k||\langle \psi_j | \phi_k \rangle|^2 \leq C \sum_{kj} |x_k||\langle \psi_j | \phi_k \rangle|^2 \leq C \sum_k |x_k| = C\|X\|_1. \]

2. Assume that X is a $2 \times 2$ Hermitian traceless matrix, such that

\[ X = \sum_{ij} x_{ij} |\psi_i\rangle \langle \psi_j|. \]

First, let us compute the trace norm of X. Let $\omega_1$ and $\omega_2$ be the singular values of X, then

\[ \|X\|_1^2 = (\omega_1 + \omega_2)^2 = \text{Tr}(X^*X) + 2|\text{det}(X)| \]

(3.40)

\[ = x_{11}^2 + x_{22}^2 + |x_{12}|^2 + |x_{21}|^2 + 2|x_{11}x_{22} - x_{12}x_{21}| \]

(3.41)

\[ = x_{11}^2 + x_{22}^2 + |x_{12}|^2 + |x_{21}|^2 + 2(x_{11}^2 + |x_{12}|^2) \]

(3.42)

\[ = 2(x_{11}^2 + x_{22}^2 + |x_{12}|^2 + |x_{21}|^2) \]

(3.43)

\[ = 2 \sum_{ij} |x_{ij}|^2. \]

(3.44)

Here we used that $0 = \text{Tr}(X) = x_{11} + x_{22}$, and $X^* = X$, so $x_{12} = x_{21} \overline{x_{21}}$. On the other hand, let us denote a diagonal matrix

\[ \Gamma^j = \sum_k C_{kj} |\phi_k\rangle \langle \phi_k|. \]

Then $D = \sum_j \Gamma^j |\psi_j\rangle \langle \psi_j|$, and therefore by Cauchy-Schwartz inequality

\[ |\text{Tr}\{DX\}|^2 = \left| \sum_{ji} x_{ij} |\psi_i| \Gamma^j |\psi_j| \right|^2 \]

(3.46)

\[ \leq \left( \sum_{ji} |x_{ij}|^2 \right) \left( \sum_{ij} |\langle \psi_i | \Gamma^j |\psi_j \rangle|^2 \right) \]

(3.47)

\[ = \frac{1}{2} \|X\|_1^2 \sum_{ij} |\langle \psi_j | \Gamma^j |\psi_i \rangle|^2 \]

(3.48)

\[ = \frac{1}{2} \|X\|_1^2 \sum_{j} |\langle \psi_j | (\Gamma^j)^2 |\psi_j \rangle| \]

(3.49)

\[ \leq \|X\|_1^2 C^2. \]

(3.50)

The last inequality follows from the fact that $\Gamma^j \leq CI$. And therefore,

\[ |\text{Tr}\{DX\}| \leq C\|X\|_1. \]
In the most general case, unfortunately, we are picking up a factor of $\sqrt{d}$ in the upper bound. Note that the only instance where the conditions on $\rho$ and $\sigma$ were used in the proof of Theorem 3.5 are in the proof of the Lemma 3.6. In the most general case,

$$|\text{Tr}(DX)| \leq \|DX\|_1 \leq \|X\|_1 \|D\|_\infty \leq \|X\|_1 \|D\|_2.$$  

And from the structure of $D$ in (3.52),

$$\|D\|_2^2 = \text{Tr}(D^*D) = \sum_{kj} C_{kj}^2 |\langle \psi_j | \phi_k \rangle|^2 \leq C^2 d.$$  

And therefore, using this result in the proof of Theorem 3.5, we obtain the following upper bound.

3.7 Theorem. Let $f \in Q_{(0,\infty)}$ be an operator monotone decreasing function such that $f(1) = 0$. Let $\rho$ and $\sigma$ be two strictly positive density operators on a $d$-dimensional Hilbert space. Then

$$S_f(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \sqrt{d} \left[ \frac{\lambda_{\rho}}{\lambda_{\rho} - \alpha_{\sigma}} f(\lambda_{\rho}^{-1} \alpha_{\sigma}) - a_f \right],$$

(3.51)

where

- $\lambda_{\rho} \in (0,1]$ is the largest eigenvalue of $\rho$,
- $\alpha_{\sigma} \in (0,1]$ is the smallest eigenvalue of $\sigma$,
- $a_f = -\lim_{y\uparrow\infty} \frac{f(iy)}{i\gamma}$.

We conjecture that the dimensionless bound holds in any dimension without any restriction on the states.

3.8 Conjecture. Let $\rho = \sum_j \lambda_j |\psi_j\rangle \langle \psi_j|$, and $\sigma = \sum_k \mu_k |\phi_k\rangle \langle \phi_k|$ be written in their spectral decomposition. Let

$$D = \sum_{kj} C_{kj} |\psi_j\rangle \langle \psi_j| \langle \phi_k| \langle \phi_k|,$$

(3.52)

such that $0 \leq C_{kj} \leq C$ for all $k, j$ and some $C$. Then

$$|\text{Tr}(D(\rho - \sigma))| \leq C\|\rho - \sigma\|_1.$$  

Note that with the above notations,

$$\text{Tr}(D(\rho - \sigma)) = \sum_{kj} C_{kj} (\lambda_j - \mu_k) |\langle \phi_k| \langle \psi_j\rangle|^2.$$
4 Examples

4.1 Logarithmic function

The function $f(x) = -\log x$ is operator monotone decreasing and convex. Moreover, $f(xy) = f(x) + f(y)$, for any positive $x, y$. Then by Corollary 3.4 we obtain

$$S(\rho\|\sigma) \leq \frac{\lambda_\rho}{\alpha}\|\rho - \sigma\|_1.$$  \hfill (4.1)

If $\rho$ and $\sigma$ are two-dimensional, then from Theorem 3.5 we obtain

$$S(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \frac{\log \lambda_\rho - \log \alpha_\sigma}{\lambda_\rho - \alpha_\sigma}.$$  

First, assume that $\alpha_\sigma \leq \lambda_\rho$. By the Mean Value Theorem for the logarithmic function, there exists $c \in [\alpha_\sigma, \lambda_\rho]$ such that

$$\frac{\log \lambda_\rho - \log \alpha_\sigma}{\lambda_\rho - \alpha_\sigma} = c^{-1} \leq \alpha_\sigma^{-1}.$$  

Therefore,

$$S(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \frac{\lambda_\rho}{\alpha_\sigma}.$$  

Second, assume that $\alpha_\sigma \geq \lambda_\rho$. By the Mean Value Theorem for the logarithmic function, there exists $c \in [\lambda_\rho, \alpha_\sigma]$ such that

$$\frac{\log \lambda_\rho - \log \alpha_\sigma}{\lambda_\rho - \alpha_\sigma} = c^{-1} \leq \lambda_\rho^{-1}.$$  

Therefore,

$$S(\rho\|\sigma) \leq \|\rho - \sigma\|_1.$$  

Therefore, putting these bounds together,

$$S(\rho\|\sigma) \leq \|\rho - \sigma\|_1 \max \left\{1, \frac{\lambda_\rho}{\alpha_\sigma}\right\}.$$  

4.2 Power function

For $p \in (-1, 1)$ and $p \neq 0$ let us take the function

$$f_p(x) := \frac{1}{p(1-p)}(1 - x^p),$$

which is operator monotone decreasing and operator convex. The quasi-relative entropy for this function is calculated to be

$$S_{f_p}(\rho\|\sigma) = \frac{1}{p(1-p)} \left(1 - \text{Tr}(\sigma^p \rho^{1-p})\right).$$
The function $f$ can be written in a form as in Theorem 3.1 for two points $x, y$

$$f(xy) = \frac{1}{p(1-p)} - g(x)h(y),$$

where we define

$$g(x) = \frac{1}{p(1-p)} x^p, \quad h(y) = y^p.$$ 

Function $g$ is operator monotone and concave for $p \in (-1, 1)$. The proof of the Theorem can be easily modified for all functions of the form

$$f(xy) = K - g(x)h(y),$$

where $g$ is operator monotone and $g(1) = K$ and $h(1) = 1$, for some $K \geq 0$. The upper bound in the theorem or Corollary 3.2 will not change. Therefore, from Corollary 3.2, we obtain

$$S_{f_p}(\rho\|\sigma) \leq \frac{1}{1-p} \|\rho - \sigma\|_1 \max_j \left\{ \lambda_j^{1-p} \right\} \alpha^{p-1} \leq \frac{1}{1-p} \|\rho - \sigma\|_1 \frac{\lambda_1^{1-p}}{\alpha^{1-p}}. \quad (4.2)$$

If $\rho$ and $\sigma$ are two-dimensional states, then from Theorem 3.7, we obtain

$$S_{f_p}(\rho\|\sigma) \leq \frac{1}{p(1-p)} \|\rho - \sigma\|_1 \frac{\lambda_\rho}{\lambda_\rho - \alpha_\sigma} (1 - \lambda_\rho^{-p} \alpha_\rho^p) \quad (4.3)$$

$$= \frac{1}{p(1-p)} \lambda_\rho^{1-p} \|\rho - \sigma\|_1 \frac{\lambda_\rho - \alpha_\rho}{\lambda_\rho - \alpha_\sigma}. \quad (4.4)$$

By the Mean Value Theorem, there exists $c$ between points $\alpha_\rho$ and $\lambda_\rho$ such that

$$\frac{\lambda_\rho - \alpha_\rho}{\lambda_\rho - \alpha_\sigma} = pc^{p-1},$$

and

$$c^{p-1} \leq \min\{\alpha_\rho^{p-1}, \lambda_\rho^{p-1}\}. \quad (4.5)$$

Therefore,

$$S_{f_p}(\rho\|\sigma) \leq \frac{1}{1-p} \|\rho - \sigma\|_1 \max \left\{ 1, \frac{\lambda_\rho^{1-p}}{\alpha_\rho^{1-p}} \right\}. \quad (4.6)$$

4.1 Remark. Let us look at the inequality (3.7) in [10], which is used in the derivation of the bound (1.2): the inequality states

$$\frac{\alpha_\rho^{-r} - \alpha_\sigma^{-r}}{\alpha_\sigma - \alpha_\rho} \leq \frac{1}{\alpha^q}, \quad (4.5)$$

where $r := q - 1 \in (0, 1]$ for $q \in (1, 2]$.

The function $h(x) := -x^{-r}$ is convex and monotonically decreasing for $r \in (0, 1]$. Then by the Mean Value Theorem, there exists a point $c$ between points $\alpha_\rho$ and $\alpha_\sigma$, such that

$$f(\alpha_\rho) - f(\alpha_\sigma) = -h'(c) \leq -h'(\alpha) = r \alpha^{-r-1} = (q-1) \alpha^{-q},$$

where $\alpha = \min\{\alpha_\rho, \alpha_\sigma\}$. This improves the constant in (4.5), leading to the bound (1.3).

Acknowledgments. A. V. is partially supported by NSF grant DMS-1812734.
References

[1] H. Araki, “Relative Entropy of States of von Neumann Algebras”, Publ. RIMS Kyoto Univ. 9: 809, (1976)

[2] K. Audenaert, J. Eisert, “Continuity bounds on the quantum relative entropy -II,” J Math Phys 52.11: 112201, (2011)

[3] Bhatia, Matrix analysis, Springer-Verlag, New York, 1997

[4] I. Csiszár, “Information type measure of difference of probability distributions and indirect observations”, Studia Sci. Math. Hungar. 2: 299, (1967)

[5] W. F. Donoghue, “Monotone Matrix Functions and analytic Continuation”, Grundlehren der Math. Wiss., vol. 287. Springer-Verlag, Berlin, 1974.

[6] F. Hiai, M. Mosonyi, “Different quantum f-divergences and the reversibility of quantum operations,” Reviews in Mathematical Physics, 29.07: 1750023, (2017)

[7] A. Jencova, M. B. Ruskai, “A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality”, Rev. Math. Phys., 22.09: 1099, (2010)

[8] D. Petz, “Quasi-entropies for states of a von Neumann algebra”, Publ. RIMS. Kyoto Univ. 21: 781, (1985)

[9] D. Petz, “Quasi-Entropies for Finite Quantum Systems”, Rep. Math. Phys. 23: 57, (1986)

[10] A. E. Rastegin, “Upper continuity bounds on the relative $q$-entropy for $q > 1$”, J Math. Phys. 52.6: 062203, (2011)

[11] H. Umegaki, “Conditional Expectation in an Operator Algebra. IV. Entropy and Information”, Kodai Math. Sem. Rep. 14: 59, (1962)

[12] D. Virosztek, “Quantum entropies, relative entropies, and related preserver problems”, Ph.D. Thesis, (2016)

[13] J. Watrous, “The theory of quantum information”, Cambridge University Press, 2018