On quantum quasi-relative entropy

Anna Vershynina

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

January 21, 2019

Abstract

We consider a quantum quasi-relative entropy S_f^K for an operator K and an operator convex function f. We show how to obtain the error bounds for the monotonicity and joint convexity inequalities from the recent results for the f-divergences (i.e. $K = I$). We also provide an error term for a class of operator inequalities, that generalize operator strong subadditivity inequality. We apply those results to demonstrate explicit bounds for the logarithmic function, that leads to the quantum relative entropy, and the power function, which gives, in particular, a Wigner-Yanase-Dyson skew information. In particular, we provide the remainder terms for the strong subadditivity inequality, operator strong subadditivity inequality, WYD-type inequalities, and the Cauchy-Schwartz inequality.

Contents

1 Introduction

Quantum quasi-relative entropy was introduced by Petz [27, 28] as a quantum generalization of a classical Csiszár's f-divergence [10]. 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, ρ and σ be two states (given by density operators), K be an operator on $\mathcal{H}$, and $f : (0, \infty) \to \mathbf{R}$ be an operator convex function. Then the quasi-relative entropy is defined as

$$S_f^K(\rho||\sigma) = \mathrm{Tr}(\rho^{1/2} K^* f(\Delta_{\sigma,\rho})(K\rho^{1.2})) \ ,$$

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

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

The modular operator can be applied straightforward when ρ is invertible. When ρ is not invertible, we take the generalized inverse of ρ and denote it as ρ^{-1} as well. The generalized inverse is defined as follows: if $\sum(\rho)$ denotes the spectrum of ρ, and P_λ denotes the spectral projection corresponding to the eigenvalue λ, then the generalized inverse $\rho^{-1} := \sum_{\lambda \in \sum(\rho) \backslash \{0\}} \lambda^{-1} P_\lambda$.

Note that taking $f(x) = -\log(x)$ and $K = I$ reduces quasi-relative entropy to the Umegaki relative entropy [39],

$$S(\rho||\sigma) = \mathrm{Tr}(\rho[\log \rho - \log \sigma]) \ .$$

We consider several properties of the relative entropy and their analogue in the case of a quasi-relative entropy.

Monotonicity of relative entropy. The most essential property of the relative entropy is the monotonicity inequality (or data processing inequality). It states that the quasi-relative entropy cannot increase after the states pass through a noisy quantum channel (i.e. a completely-positive, trace-preserving map) $\mathcal{N}$:

$$S(\rho||\sigma) \geq S(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) \ .$$

This inequality was proved by Lindblad [23], building on the work of Lieb and Ruskai [20]. Due to a Stinespring factorization [35], the monotonicity inequality under any quantum channel, is equivalent to the monotonicity inequality under partial traces, which will be used throughout the

paper. Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, and ρ and σ be two states on $\mathcal{H}$. We denote the partial trace over the second system as $\rho_1 := \mathrm{Tr}\rho_{12}$, and others similarly. Then the monotonicity inequality states that

$$S(\rho\|\sigma) \geq S(\rho_1\|\sigma_1) \ .$$

Petz [29, 30] provided a condition on states ρ and σ, for which the monotonicity inequality in saturated. He showed that for a given quantum channel $\mathcal{N}$, two states ρ and σ lead to equality in the monotonicity relation if and only if the channel in noiseless for these states, i.e. the action of the channel can be reversed. The reverse action is implemented by a certain recovery map $\mathcal{R}$, now called a Petz's recovery map, which can be written explicitly for a given channel.

These results motivated the question of stability of the recovery map: if the decrease of relative entropy after states pass through a quantum channel is small, how well can these states be recovered? Work on this question started following a breakthrough result by Fawzi and Renner in 2015 [12]. They proved that if the strong subadditivity inequality (SSA) (which is equivalent to the monotonicity inequality) is nearly saturated, then quantum Markov chain condition, known to be necessary and sufficient for equality in SSA [13], is also nearly satisfied, and they gave a precise quantitative version of this stability result.

Further refinements of the monotonicity relation occurred later in, for example, [3, 5, 6, 16, 36, 37, 45, 46]. Most of these results involve some sort of a recovery channel in the lower bound, and even though it often derived from, or related to, the Petz recovery map, it is in no case the Petz map itself. The quantities provided in those lower bounds are hard to compute.

Recently, Carlen and Vershynina [8] proved a sharp stability result for the monotonicity inequality in terms of the original Petz recovery map:

$$S(\rho\|\sigma) - S(\rho_1\|\sigma_1) \geq \left(\frac{\pi}{8}\right)^4 \|\Delta_{\sigma,\rho}\|^{-2}\|\mathcal{R}_\rho(\sigma_1) - \sigma\|_1^4 \ .$$

It is possible to obtain the lower bound in terms of recovery map $\mathcal{R}_\sigma$, but the constant changes. This bound makes it very easy to see the relation between the saturation of the monotonicity inequality and the perfect recovery of both states.

Monotonicity of quasi-relative entropy. The monotonicity inequality for a quasi-relative entropy S_f^K also holds, but for operators K of the a certain form. As mentioned before, considering quantum channels is equivalent to considering only a partial trace channel. For operators $K = K \otimes I$ we have

$$S_f^{K \otimes I}(\rho\|\sigma) \geq S_f^K(\rho_1\|\sigma_1) \ .$$

This statement was explicitly proved by Sharma [34], building on the work of Nielsen and Petz [25]. The monotonicity can also be proved for operators $K = K \otimes V$, where V is a unitary.

Sharma has also provided a condition on the equality, showing that the equality in the monotonicity inequality holds if and only if, for all $\beta \in \mathbb{C}$,

$$\mathrm{Tr}(K^*\sigma_1 K \rho_1^{1-\beta} - K^*\sigma^\beta K \rho^{1-\beta}) = 0 \ .$$

Building on their previous work, Carlen and Vershynina [9] also recently showed that for f-divergences (i.e. $K = I$), the following sharpening of the monotonicity inequality holds: for a

large class of functions f (that define a constant c), there is constant M depending only on the smallest non-zero eigenvalue of ρ, $\|\sigma_1^{-1}\|$, β, and f, such that

$$\max\{\|\mathscr{R}_\rho(\sigma_1) - \sigma\|_1 , \|\mathscr{R}_\sigma(\rho_1) - \rho\|_1\} \leq M(S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{\frac{1}{4}\frac{1}{1+c}} .$$

Similar to their previous bound, it is very easy to see here that if the monotonicity is saturated, Petz's map recovers both states perfectly. The other way is easily derived.

In fact, Carlen and Vershynina showed that even a stronger bound holds: for all $\beta \in (0, 1)$, and conditions above, there is a constant $\alpha(\beta)$ (that also depends on the function f) such that

$$\|\sigma_1^\beta \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta \rho^{1/2-\beta}\|_2 \leq M(S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{\alpha(\beta)} .$$

This means that both: saturation of the monotonicity inequality and Petz's recovery of both states, is equivalent to the following condition:

$$\text{for all } \beta \in \mathbb{C}: \ \sigma_1^\beta \rho_1^{-\beta} = \sigma^\beta \rho^{-\beta} .$$

We build on Carlen and Vershynina [9] work to improve the monotonicity result for quasi-relative entropies with an operator $K = K \otimes V$, where V is unitary. We show that with the same conditions as above,

$$\|\sigma_1^\beta K \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta K \rho^{1/2-\beta}\|_2 \leq M(S_f^K(\rho||\sigma) - S_f^K(\rho_1||\sigma_1))^{\alpha(\beta)} .$$

In particular, we obtain the bound in terms of the Petz's recovery map

$$\|\mathscr{R}_\rho(K_1^*\sigma_1 K_1) - K^*\sigma K\|_1 \leq M(S_f^K(\rho||\sigma) - S_f^{K_1}(\rho_1||\sigma_1))^{\frac{1}{4(1+c)}} .$$

Moreover, we show that the equality in the monotonicity inequality holds if and only if, for all $\beta \in \mathbb{C}$,

$$\sigma_1^\beta K \rho_1^{-\beta} \rho^{1/2} = \sigma^\beta K \rho^{1/2-\beta} .$$

Strong subadditivity inequality. For a tri-partite state ρ_{ABC} on a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, strong subadditivity inequality (SSA) states

$$0 \leq S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) .$$

This theorem was proved by Lieb and Ruskai [20], using Lieb's theorem that was proved in [18]. Note that this inequality is equivalent to the monotonicity inequality: having monotonicity inequality, to obtain the SSA inequality, one takes $\rho = \rho_{ABC}$ and $\sigma = \rho_{AB} \otimes \rho_C$ and a trace over the system $\mathcal{H}_A$. Having SSA, one chooses ρ_{ABC} to be block-diagonal, which implies that the map $\rho_{12} \to S\rho_1 - S(\rho_{12})$ is convex. Following Lieb and Ruskai [20], this yields the monotonicity inequality.

We apply previous results in [8], to obtain the sharpening of the SSA: for any $\beta \in (0, 1)$, there are some constants N (depending on the minimal eigenvalue of ρ_{ABC}, β and f) and $\alpha(\beta)$ (depending on the function f) such that

$$N\|\rho_B^\beta \otimes \rho_C^\beta \rho_{BC}^{-\beta}\rho^{1/2} - \rho_{AB}^\beta \otimes \rho_C^\beta \rho^{1/2-\beta}\|_2^{1/\alpha(\beta)} \leq S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) .$$

In particular we show that the equality in SSA holds if and only if Petz's map $\mathscr{R}$ recovers the reduced state perfectly:

$$\mathscr{R}_\rho(\rho_B \otimes \rho_C) = \rho_{AB} \otimes \rho_C .$$

Operator inequalities of the strong subadditivity type. In 2012 Kim [17] proved the following operator version of the strong subadditivity inequality

$$0 \leq \mathrm{Tr}_{AB}\rho_{ABC}[\log \rho_{ABC} - \log \rho_{AB} - \log \rho_{BC} + \log \rho_B] .$$

Note that this inequality leads to the strong subadditivity inequality after taking the trace over system $\mathcal{H}_C$.

Building on this work, Ruskai [33] provided a class of operator inequalities: for an operator monotone decreasing function f,

$$0 \leq \mathrm{Tr}_{AB}\left([f(L_{\sigma_{AB}}R_{\rho_{ABC}}^{-1}) - f(L_{\sigma_B}R_{\rho_{BC}}^{-1})]\rho_{ABC}\right) . \tag{1.1}$$

We provide the error term for these inequalities of the type: for a large class of operator convex functions f, there are some constants N (depending on the minimal eigenvalue of ρ_{ABC}, β and f) and $\alpha(\beta)$ (depending on the function f) such that

$$N[\mathrm{Tr}_{AB}(P_{ABC}(\rho, \sigma)P_{ABC}^*(\rho, \sigma))]^{1/\alpha(\beta)} \leq \mathrm{Tr}_{AB}\left([f(L_{\sigma_{AB}}R_{\rho_{ABC}}^{-1}) - f(L_{\sigma_B}R_{\rho_{BC}}^{-1})]\rho_{ABC}\right) ,$$

where

$$P_{ABC}(\rho, \sigma) := \sigma_B^\beta \rho_{BC}^{-\beta} \rho_{ABC}^{1/2} - \sigma_{AB}^\beta \rho_{ABC}^{1/2-\beta} .$$

Note that the left-hand side of this inequality is the operator on $\mathcal{H}_C$, as is the right-hand side. Therefore, this inequality holds between operators on $\mathcal{H}_C$.

Moreover, we show that the equality in (1.1) holds if and only if, for all $\beta \in (0,1)$

$$\sigma_B^\beta \rho_{BC}^{-\beta} = \sigma_{AB}^\beta \rho_{ABC}^{-\beta} ,$$

which is in turn equivalent to the recovery condition

$$\mathscr{R}_{\rho_{ABC}}(\sigma_B) = \sigma_{AB} .$$

Additionally, by taking f to be a power function, Ruskai [33] showed that (1.1) leads to

$$0 \leq \frac{1}{p(1-p)}[-\mathrm{Tr}_{AB}\rho_{ABC}^{1-p}\sigma_{AB}^p + \mathrm{Tr}_B\rho_{BC}^{1-p}\sigma_B^p] ,$$

where $p \in (-1, 2)$. We apply our results to show the operator strengthening of this inequality:

$$N\left[\mathrm{Tr}_{AB}(Q_{ABC}^*(\sigma_{AB}, \rho_{ABC})Q_{ABC}(\sigma_{AB}, \rho_{ABC}))\right]^{1/\alpha(\beta)} \leq \frac{1}{p(1-p)}\left[-\mathrm{Tr}_{AB}\rho_{ABC}^{1-p}\sigma_{AB}^p + \mathrm{Tr}_B\rho_{BC}^{1-p}\sigma_B^p\right] ,$$

where

$$Q_{ABC}(\sigma_{AB}, \rho_{ABC}) := \rho_{BC}^\beta \sigma_B^{-\beta} \sigma_{AB}^{1/2} - \rho_{ABC}^\beta \sigma_{AB}^{1/2-\beta} .$$

Joint convexity of the relative entropy. It was noted by Lindblad [22] and Ulhmann [38], the relative entropy is jointly convex, i.e., if $\rho = \sum_j p_j \rho_j$ and $\sigma = \sum_j p_j \sigma_j$ (with $\sum_j p_j = 1$ and $p_j \geq 0$), then

$$0 \leq \sum_j p_j S^K(\rho_j \| \sigma_j) - S^K(\rho \| \sigma) \ .$$

This inequality is also equivalent to the monotonicity inequality: having monotonicity inequality, one chooses both ρ and σ to be block-diagonal, which immediately leads to the joint convexity (see Section 5). One the other hand, having the joint convexity inequality, Araki and Lieb [2, 19] used a purification process to show that on the set of pure states (the extreme points of the set of density matrices) the SSA holds with equality.

Joint convexity of the quasi-relative entropy. If was shown by Petz [26, 28] or [31, Theorem 2] that the quasi relative entropy is jointly convex in ρ and σ:

$$0 \leq \sum_j p_j S_f^K(\rho_j \| \sigma_j) - S_f^K(\rho \| \sigma).$$

Using our strengthening of the monotonicity inequality, we obtain the error term for the joint convexity inequality: for a large class of operator convex functions f, and any $\beta \in (0, 1)$, there are some constants M and $\alpha(\beta)$ (see above for the dependence), the following holds

$$\sum_j p_j^{1/2} \| \sigma^\beta K \rho^{-\beta} \rho_j^{1/2} - \sigma_j^\beta K \rho_j^{1/2-\beta} \|_2 \leq M \left(\sum_j p_j S_f^K(\rho_j \| \sigma_j) - S_f^K(\rho \| \sigma) \right)^{\alpha(\beta)} \ .$$

Moreover, we show that the equality in the joint convexity inequality holds if and only if, for all j and all $\beta \in \mathbb{C}$,

$$\sigma^\beta K \rho^{-\beta} = \sigma_j^\beta K \rho_j^{-\beta} \ .$$

We apply these results to show the error term in the concavity inequality of the term $\mathrm{Tr}(K^* \sigma^p K \rho^{1-p})$. The concavity of this term was shown by Lieb [18] for powers of ρ and σ that sum up to a number no greater than one.

Structure of the paper. In the next Section 2 we review known results for the operator monotone functions, in particular its integral representation. In Section 3 we introduce the quantum quasi-relative entropy and present few simples, but important facts about it. In Section 4 we review the previous strengthening of the monotonicity inequality for relative entropy and f-divergences, and present the error term for the quasi-relative entropy, with the condition for the equality. In Section 5 we apply the strengthening of the monotonicity inequality established in the previous section to the joint convexity inequality and provide the condition for the equality as well. In Section 6 we provide the error term for a class of operator inequalities established by Ruskai [33] with the condition for equality. In Section 7 we apply all previous results to obtain new inequalities for the relative entropy by taking the function $f(x) = -\log(x)$ and $K = I$. In particular, we obtain the error terms for the joint convexity inequality, strong subadditivity inequality, and the operator strong subadditivity inequalities. In Section 8 we apply previous results to the power function. The quasi-relative entropy for the power functions gives a term

$\text{Tr}(K^*\sigma^p K\rho^{1-p})$, concavity of which leads to the concavity of the Wigner-Yanase-Dyson p-skew information. In particular, we provide a Pinsker inequality for such a term, and the error terms for the joint concavity and the operator version of WYD inequalities. At the end, we apply these results to show the error term for the operator Cauchy-Schwartz inequality.

2 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 λ 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* when $-f$ is operator concave.

2.3 Theorem (Bhatia '97). *[4, Theorem V.2.5] Every continuous function f mapping $[0, \infty)$ into itself is operator monotone if and only if it is operator concave.*

2.4 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 $\mathcal{P}_{(a,b)}$.

2.5 Theorem (Löwner '34). *[4, Theorem V.4.7] 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).*

According to [11, Chapter II, Theorem I] every operator monotone decreasing function f (i.e. $-f \in \mathcal{P}_{(0,\infty)}$) has a canonical integral representation

$$f(x) = ax + b - \int_0^\infty \left(\frac{1}{t + x} - \frac{t}{t^2 + 1} \right) \mathrm{d}\mu_f(t) , \tag{2.1}$$

where $a \leq 0$, $b \in \mathbb{R}$ and μ is a positive measure on $(0, \infty)$ such that $\int_0^\infty \frac{1}{t^2 + 1} \mathrm{d}\mu_f(t) < \infty$. Conversely, every such function is operator monotone decreasing.

The following formulas [11, Chapter II, p. 24] determine a, b and μ corresponding to f.

$$a = \lim_{y \uparrow \infty} \frac{f(iy)}{iy} \qquad \text{and} \qquad b = \text{Re}\,(f(i)) . \tag{2.2}$$

Define the monotone increasing function $\mu(x) := \frac{1}{2}\mu(\{x\}) + \mu((-\infty, x))$, then according to [11, Chapter II, Lemma 2] we have that

$$\mu(x_1) - \mu(x_0) = \lim_{y \downarrow 0} \frac{1}{\pi} \int_{x_0}^{x_1} \text{Im}\, f(-x + iy)\mathrm{d}x . \tag{2.3}$$

2.6 Definition. A operator monotone function f is *regular* in case the measure μ in the canonical integral representation of f is absolutely continuous with respect to Lebesgue measure, and for each $S, T > 0$, there is a finite constant $C_{S,T}^f$ such that

$$\mathrm{d}t \le C_{S,T}^f \mathrm{d}\mu(t) \tag{2.4}$$

on the interval $[1/S, T]$. An operator monotone decreasing function is *regular* if and only if $-f$ is regular.

2.7 Example. Let $f(x) = -\log(x)$. It is operator monotone decreasing. Then

$$b = \mathrm{Re}\left(\log(i)\right) = 0 \ ,$$

and

$$a = \lim_{y \uparrow \infty} \log(iy)/(iy) = \lim_{y \uparrow \infty} (\log y + i\pi/2)/(iy) = 0 \ .$$

It is clear from (2.3) that

$$\mathrm{d}\mu(x) = \frac{1}{\pi} \lim_{y \downarrow 0} \mathrm{Im} \, \log(-x + iy) \mathrm{d}x = \mathrm{d}x \ .$$

Then the integral representation (2.1) gives the following formula for the logarithmic function

$$-\log x = \int_0^\infty \left(\frac{1}{t+x} - \frac{t}{t^2+1} \right) \mathrm{d}t \ . \tag{2.5}$$

2.8 Example. Let $f(x) = x^p$, where $p \in \mathbb{R}$. Then by [4, Theorem V.2.10] the function f is

1. operator monotone if and only if $p \in [0, 1]$;

2. operator convex if and only if $p \in [-1, 0] \cup [1, 2]$;

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

$$a = \lim_{y \uparrow \infty} f(iy)/(iy) = 0 \ , \quad \text{and} \quad b = \cos(p\pi/2) \ .$$

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

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

This yields the representation

$$-x^p = -\cos(p\pi/2) + \frac{\sin(p\pi)}{\pi} \int_0^\infty t^p \left(\frac{1}{t+x} - \frac{t}{t^2+1} \right) \mathrm{d}t \ . \tag{2.6}$$

3   Quantum quasi-relative entropy

The notion of quantum quasi-relative entropy was introduced by Petz [27, 28] or [26, Chapter 7]. Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and ρ and σ be states on $\mathcal{H}$ (i.e. trace-one, positive semi-definite operators). Note that notions and results in this paper can be formulated for von Neumann algebras, as it was done in multiple references.

3.1 Definition. For an operator monotonically decreasing function f (which implies that f is also operator convex), such that $f(1) = 0$, states ρ and σ, and a bounded operator K on $\mathcal{H}$, *the quasi-relative entropy* is defined as

$$S_f^K(\rho||\sigma) = \mathrm{Tr}(\sqrt{\rho}K^* f(\Delta_{\sigma,\rho})K\sqrt{\rho}),$$

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 is a straightforward way to calculate the quasi-relative entropy from the spectral decomposition of states [41]. Let ρ and σ have the following spectral decomposition

$$\rho = \sum_j \lambda_j |\psi_j\rangle \langle\psi_j|, \quad \sigma = \sum_k \mu_k |\phi_k\rangle \langle\phi_k|.$$

Then the quasi-relative entropy is calculated as follows

$$S_f^K(\rho||\sigma) = \sum_{j,k} \lambda_j f\left(\frac{\mu_k}{\lambda_j}\right) |\langle\phi_k| K |\psi_j\rangle|^2. \tag{3.1}$$

3.2 Example.   1. For $K = I$, the quasi-relative entropy

$$S_f(\rho||\sigma) = \mathrm{Tr}(\rho^{1/2} f(\Delta_{\sigma,\rho})\rho^{1/2}),$$

is sometimes referred to as an f-divergence.

2. For $-1 \leq p < 1$ define a function

$$f_p(x) := \begin{cases} \frac{1}{p(1-p)}(1 - x^p) & p \neq 0 \\ -\log x & p = 0 \end{cases}.$$

Note that from Example 2.8 the function is convex. The quasi-relative entropy for this function

$$S_{f_{1-p}}^{K^*}(\rho||\sigma) = J_p(K, \rho, \sigma) := \mathrm{Tr}(\sqrt{\sigma}K^* g_p(\Delta_{\rho,\sigma})K\sqrt{\sigma})$$

is the function J_p defined by Jencova and Ruskai in [15], here for $0 < p \leq 2$ the function g_p is defined as

$$g_p(x) := x f_{1-p}(x^{-1}) = \begin{cases} \frac{1}{p(1-p)}(x - x^p) & p \neq 1 \\ x \log x & p = 1 \end{cases}.$$

3. In the above example, if $p = 0$ (i.e. $f(x) = -\log x$) and $K = I$, we obtain the Umegaki relative entropy [39]

$$S^I_{f_0}(\rho||\sigma) = J_1(I, \rho, \sigma) = S(\rho||\sigma) = \text{Tr}(\rho \log \rho - \rho \log \sigma).$$

4. In example 2, the quasi-relative entropy can be calculated to be

$$S^K_{f_p}(\rho||\sigma) = \frac{1}{p(1-p)} \text{Tr}(K^*\rho K - K^*\sigma^p K \rho^{1-p}) \ .$$

This expression has the term $\text{Tr}(K^*\sigma^p K\rho^{1-p})$, concavity of which was proved by Lieb in [18] with more general powers.

5. In the above example, taking $\sigma = \rho$ and $K^* = K$, results in

$$S^K_{f_p}(\rho||\rho) = -\frac{1}{2p(1-p)} \text{Tr}[K, \rho^p][K, \rho^{1-p}] \geq 0 \ .$$

Up to a constant, this is the Wigner-Yanase-Dyson p-skew information [43, 44], for $p \in (0, 1)$.

Now we state a few simple properties of quasi-relative entropy, some of which have been noted before.

3.3 Proposition. *The quasi-relative entropy scales as follows: for any constant c,*

$$S^K_f(c\rho||c\sigma) = cS^K_f(\rho||\sigma) \ .$$

Proof. This following directly from the formula (3.1). $\qquad\square$

3.4 Proposition. *For a unitary U, and states ρ and σ*

$$S^U_f(\rho||\sigma) = S^I_f(U\rho U^*||\sigma) = S^I_f(\rho||U^*\sigma U).$$

Proof. If ρ and σ have the following spectral decomposition

$$\rho = \sum_j \lambda_j |\psi_j\rangle \langle\psi_j|, \quad \sigma = \sum_j \beta_j |\phi_j\rangle \langle\phi_j|,$$

then by [41]

$$f(\Delta_{\sigma,\rho}) = \sum_{j,k} f\left(\frac{\beta_j}{\lambda_k}\right) \Delta_{|\phi_j\rangle\langle\phi_j|, |\psi_k\rangle\langle\psi_k|}.$$

Therefore, the quasi-relative entropy can be written as

$$S^U_f(\rho||\sigma) = \sum_{j,k} f\left(\frac{\beta_j}{\lambda_k}\right) \text{Tr}(\sqrt{\rho}\, U^* |\phi_j\rangle \langle\phi_j| U \sqrt{\rho} |\psi_k\rangle \langle\psi_k|). \qquad (3.2)$$

On the other hand, using the spectral decomposition of ρ and the fact that U is unitary, we obtain

$$S_f^I(U\rho U^*||\sigma) = \text{Tr}(U\sqrt{\rho}\,U^* f(\Delta_{\sigma,U\rho U^*})(U\sqrt{\rho}\,U^*)) \tag{3.3}$$

$$= \sum_{j,k} f\left(\frac{\beta_j}{\lambda_k}\right) \text{Tr}(U\sqrt{\rho}\,U^*\,|\phi_j\rangle\,\langle\phi_j|\,U\sqrt{\rho}\,U^*U\,|\psi_k\rangle\,\langle\psi_k|\,U^*) \tag{3.4}$$

$$= \sum_{j,k} f\left(\frac{\beta_j}{\lambda_k}\right) \text{Tr}(\sqrt{\rho}\,U^*\,|\phi_j\rangle\,\langle\phi_j|\,U\sqrt{\rho}\,|\psi_k\rangle\,\langle\psi_k|) \tag{3.5}$$

$$= S_f^U(\rho||\sigma). \tag{3.6}$$

Similarly, $S_f^U(\rho||\sigma) = S_f^I(\rho||U^*\sigma U)$. $\qquad\square$

3.5 Proposition. *For a unitary U, and states ρ and σ the quasi-relative entropy is non-negative*

$$S_f^U(\rho||\sigma) \geq 0.$$

Proof. Form the previous proposition it is evident that it is enough to consider $U = I$. Fix a basis $\{|j\rangle\,\langle j|\}_j$ of the Hilbert space $\mathcal{H}$ the states act on, and consider the map

$$\Phi(\omega) = \sum_j \langle j|\,\omega\,|j\rangle\,|j\rangle\,\langle j|.$$

It is a Schwarz map, so from the monotonicity of the quasi-relative entropy (4.8), which we will discuss in Section 4.1, we have

$$S_f^I(\rho||\sigma) \geq S_f^I(\Phi(\rho)||\Phi(\sigma)).$$

The last expression is the classical quasi-relative entropy, which is non-negative. $\qquad\square$

3.6 Proposition. *For a unitary U, and states ρ and σ*

$$S_f^U(\rho||\sigma) = 0 \quad \text{if and only if} \quad \rho = U^*\sigma U.$$

Proof. According to Proposition 3.4, it is enough to consider $U = I$.

Then, one way: if $\rho = \sigma$, then $S_f(\rho,\sigma) = 0$ by definition, since $f(1) = 0$.

The other way: from the proof of last proposition,

$$0 = S_f^I(\rho||\sigma) \geq S_f^I(\Phi(\rho)||\Phi(\sigma)).$$

Therefore, for any basis $\{|j\rangle\}_j$, $\Phi(\rho) = \Phi(\sigma)$, which implies that $\rho = \sigma$. $\qquad\square$

A famous bound relating quantum relative entropy and trace distance between two quantum states, called Pinsker inequality. The similar inequality holds for the quasi-relative entropy as well, as was shown by Hiai and Mosonyi [14] for $U = I$.

3.7 Proposition. *For an operator convex function f on $(0,\infty)$ with $f(1) = 0$, a unitary U, and states ρ and σ, the following holds*

$$\frac{f''(1)}{2}||\rho - U^*\sigma U||_1^2 \leq S_f^U(\rho||\sigma) .$$

It's worthwhile to point out that any inequality that holds in the classical case for probability distributions, also holds in the quantum case between density operators. Let us show how to obtain quantum inequality from the classical one. This relation can be found in many quantum information books, e.g. [42], for relative entropies, i.e. when $f(x) = -\log(x)$. For a general function f the proof is similar.

3.8 Lemma. *If the following inequality holds for all probability distributions p and q:*

$$S_f^{cl}(p\|q) \geq F(\|p - q\|_1),$$

for some function F, then for all states ρ and σ the following holds as well

$$S_f(\rho\|\sigma) \geq F(\|\rho - \sigma\|_1) .$$

Proof. For an operator $\rho - \sigma$ consider its Jordan-Hahn decomposition

$$\rho - \sigma = P - Q,$$

where $P, Q > 0$. Define a projector onto the image of P as $\Pi = \Pi_{im(P)}$. Then, a measurement with operators $\{\Pi, I - \Pi\}$ is a projective measurement. It holds that the trace distance between states ρ and σ equals to the L^1 distance between probability distributions

$$\|\rho - \sigma\|_1 = \text{Tr}(P) + \text{Tr}(Q) = 2\text{Tr}(P) = 2|\text{Tr}(\Pi(\rho - \sigma))| = \|p - q\|_1,$$

where $p = \{\text{Tr}(\Pi\rho), 1 - \text{Tr}(\Pi\rho)\}$, $q = \{\text{Tr}(\Pi\sigma), 1 - \text{Tr}(\Pi\sigma)\}$ are probability ensembles.

For above Π, define a quantum-to-classical channel as follows

$$\Phi(\omega) = \text{Tr}(\Pi\omega) |0\rangle \langle 0| + (1 - \text{Tr}(\Pi\omega)) |1\rangle \langle 1| .$$

Then from the monotonicity inequality for the quasi-relative entropy, and for above p and q, the quantum quasi-relative entropy is lower bounded by the classical quasi-relative entropy

$$S_f(\rho\|\sigma) \geq S_f(\Phi(\rho)\|\rho(\sigma)) = \sum_j p_j f(p_j^{-1} q_j) = S_f^{cl}(p\|q).$$

Therefore, any inequality that holds between classical f-divergence and the L^1 distance, also holds in the quantum case between quantum f-divergence and trace distance between two states. $\square$

4   Monotonicity inequality

In this section we assume a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Two states ρ and σ act on this Hilbert space. Denote $\rho_1 := \text{Tr}_2\rho$ and similarly, $\sigma_1 = \text{Tr}_2\sigma$. Here we build on results in [9] to provide a strengthening of the monotonicity inequality for quasi-relative entropies for a large class of operators K.

4.1 Review

For a regular relative entropy, i.e. $S(\rho||\sigma) = \text{Tr}[\rho(\log \rho - \log \sigma)]$, Lindblad proved [22] the following monotonicity inequality

4.1 Theorem (Monotonicity of relative entropy)**.**

$$S(\rho||\sigma) - S(\rho_1||\sigma_1) \geq 0. \tag{4.1}$$

Lindblad showed that the monotonicity inequality (4.1) is equivalent to the joint convexity of the relative entropy $(\rho, \sigma) \mapsto S(\rho||\sigma)$, see Proposition 5.1, and this in turn is an immediate consequence of Lieb's Concavity Theorem [18], which was proved by Lieb and Ruskai [20], who showed it to be equivalent to the Strong Subadditivity (SSA) of the von Neumann entropy, see Theorem 6.1.

Note that the monotonicity inequality also holds for CPTP (completely-positive trace-preserving) maps, not just a partial trace. This fact was proved by Lindblad [23] by using Stinesping's Dialtion Theorem [35] that relates general CPTP maps to a partial trace. Therefore, we will focus only on partial traces, not general CPTP maps.

Petz has proved [29, 30] that for a relative entropy there is an equality in the monotonicity inequality (4.1) if and only if both states ρ and σ can be recovered perfectly. The recovery map $\mathscr{R}_\rho$ is known as *Petz recovery map* and is defined as $\mathscr{R}_\rho : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2)$

$$\mathscr{R}_\rho(\gamma) = \rho^{1/2} \rho_1^{-1/2} \gamma \rho_1^{-1/2} \rho^{1/2} . \tag{4.2}$$

There has been several results that provide a lower bound in (4.8) other than zero [12, 16, 45, 46], but the lower bounds provided there involve quantities that are hard to compute, e.g. rotated and twirled Petz recovery maps. For another fidelity type bound not explicitly involving the recovery map, see [7, Theorem 2.2].

In 2017, Carlen and Vershynina [8] provided the following sharpening of the monotonicity inequality

$$S(\rho||\sigma) - S(\rho_1||\sigma_1) \geq \left(\frac{\pi}{4}\right)^4 \|\Delta_{\sigma,\rho}\|^{-2} \|\sigma_1^{1/2} \rho_1^{-1/2} \rho^{1/2} - \sigma^{1/2}\|_2^4 , \tag{4.3}$$

with $\|\cdot\|_2$ denoting the Hilbert-Schmidt norm

$$\|A\|_2 := \text{Tr}(A^*A). \tag{4.4}$$

As a corollary they provided a bound that explicitly involves Petz's recovery map:

$$S(\rho||\sigma) - S(\rho_1||\sigma_1) \geq \left(\frac{\pi}{8}\right)^4 \|\Delta_{\sigma,\rho}\|^{-2} \|\mathscr{R}_\rho(\sigma_1) - \sigma\|_1^4 , \tag{4.5}$$

and

$$S(\rho||\sigma) - S(\rho_1||\sigma_1) \geq \left(\frac{\pi}{8}\right)^4 \|\Delta_{\sigma,\rho}\|^{-2} \|\rho_1\|^{-2} \|\sigma_1^{-1}\|^{-2} \|\mathscr{R}_\sigma(\rho_1) - \rho\|_1^4 , \tag{4.6}$$

with $\|\cdot\|_1$ denoting the trace norm

$$\|A\|_1 := \text{Tr}|A|. \tag{4.7}$$

This is the first time when the original Petz recovery map $\mathscr{R}$ appeared in the lower bound of the monotonicity or equivalent inequality.

For quasi-relative entropies S_f^K the monotonicity inequality holds for operators K such that $K = K_1 \otimes V_2$ for a unitary V. This was explicitly proved by Sharma [34], building on the work of Nielsen and Petz [25].

4.2 Theorem (Monotonicity of quasi-relative entropy, Sharma '14). *For every operator convex function f and operator $K = K_1 \otimes V$, where V is a unitary, we have*

$$S_f^K(\rho||\sigma) - S_f^{K_1}(\rho_1||\sigma_1) \geq 0. \tag{4.8}$$

In Section 4.3 we will recall the proof of this theorem, as we rely on it later.

In [9] Carlen and Vershynina, generalized their previous result (4.3) for f-divergences for regular functions f, see Definition 2.6. In results below, assume that $T > 0$, $\beta \in (0,1)$, and

(1) for $\beta \leq 1/2$, define $T_L(\beta) := T$ and $T_R(\beta) := T^{\beta/(1-\beta)}$;

(2) for $\beta \geq 1/2$, define $T_L(\beta) := T^{(1-\beta)/\beta}$ and $T_R(\beta) := T$.

Moreover, define $C_{T,\beta}^f$ to be the least positive constant such that $dt \leq C_{T,\beta}^f \, d\mu_f(t)$ for $t \in [T_L(\beta)^{-1}, T_R(\beta)]$, noting that $C_{T,\beta}^f > 0$ since f is regular.

In [9] it is proved that under the above conditions for a regular operator monotone decreasing function f the following holds

$$\frac{\pi}{\sin \beta\pi} \|\sigma_1^\beta \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta \rho^{1/2-\beta}\|_2 \tag{4.9}$$

$$\leq \ 2\left(\frac{1}{\beta} + \frac{\|\Delta_{\sigma,\rho}\|}{1-\beta}\right) \frac{1}{T^{\alpha_1(\beta)}} + T^{\alpha_2(\beta)} \left(C_{T,\beta}^f\right)^{1/2} (S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{1/2} \ . \tag{4.10}$$

where

$$\alpha_1(\beta) = \begin{cases} \beta & \text{when } \beta \leq 1/2 \\ 1-\beta & \text{when } \beta \geq 1/2. \end{cases} \quad \text{and} \quad \alpha_2(\beta) = \begin{cases} \frac{1-\beta}{2} + \frac{\beta^2}{2(1-\beta)} & \text{when } \beta \leq 1/2 \\ \beta & \text{when } \beta \geq 1/2. \end{cases}$$

Optimizing in T for functions, for which $C_{T,\beta}^f$ scales as a power of T^c, it was straightforward to prove [9, Corollary 4.4] that there is a constant M, depending only on the smallest non-zero eigenvalue of ρ, β, C and c, such that

$$\|\sigma_1^\beta \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta \rho^{1/2-\beta}\|_2 \leq M(S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{\alpha(\beta)} \ , \tag{4.11}$$

where

$$\alpha(\beta) = \begin{cases} \frac{\beta(1-\beta)}{1+2c(1-\beta)} & \text{when } \beta \leq 1/2 \\ \frac{1}{2}\frac{1-\beta}{1+c} & \text{when } \beta \geq 1/2. \end{cases}$$

Moreover, taking $\beta = 1/2$, for a constant M depending only on the smallest non-zero eigenvalue of ρ, $\|\sigma_1^{-1}\|$, β, C and c, we have

$$\max\{\|\mathscr{R}_\rho(\sigma_1) - \sigma\|_1 \ , \ \|\mathscr{R}_\sigma(\rho_1) - \rho\|_1\} \leq M(S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{\frac{1}{4}\frac{1}{1+c}} \ . \tag{4.12}$$

From this expression it is evident that the monotonicity inequality is saturated for a broad class of operator monotone decreasing functions f, if and only if the Petz recovery map recovers both states ρ and σ perfectly well.

4.2 Monotonicity inequality

We generalize the result in [9] to include quasi-relative entropies with a large class of operators K.

4.3 Definition. For an operator monotone decreasing function f assume that $T > 0$, $\beta \in (0, 1)$, and

(1) for $\beta \leq 1/2$, define $T_L(\beta) := T$ and $T_R(\beta) := T^{\beta/(1-\beta)}$;

(2) for $\beta \geq 1/2$, define $T_L(\beta) := T^{(1-\beta)/\beta}$ and $T_R(\beta) := T$.

Moreover, define $C_{T,\beta}^f$ to be the least positive constant such that $\mathrm{d}t \leq C_{T,\beta}^f \, \mathrm{d}\mu_f(t)$ for $t \in [T_L(\beta)^{-1}, T_R(\beta)]$, noting that $C_{T,\beta}^f > 0$ since f is regular. We will call such functions $C_{T,\beta}^f$-*regular.*

4.4 Theorem. *Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, and let an operator K be such that $K = K_1 \otimes V_2$, where V is a unitary operator. Let f be a $C_{T,\beta}^f$-regular function. Then for any states ρ, σ on $\mathcal{H}$ and any $\beta \in (0, 1)$*

$$\frac{\pi}{\sin \beta\pi}\|\sigma_1^\beta K \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta K \rho^{1/2-\beta}\|_2 \tag{4.13}$$

$$\leq 2\left(\frac{\|K\|}{\beta} + \frac{\|\Delta_{\sigma,\rho}\|}{1-\beta}\right)\frac{1}{T^{\alpha_1(\beta)}} + T^{\alpha_2(\beta)}\left(C_{T,\beta}^f\right)^{1/2}(S_f^K(\rho\|\sigma) - S_f^{K_1}(\rho_1\|\sigma_1))^{1/2} . \tag{4.14}$$

where

$$\alpha_1(\beta) = \begin{cases} \beta & when \ \beta \leq 1/2 \\ 1-\beta & when \ \beta \geq 1/2. \end{cases} \quad and \quad \alpha_2(\beta) = \begin{cases} \frac{1-\beta}{2} + \frac{\beta^2}{2(1-\beta)} & when \ \beta \leq 1/2 \\ \beta & when \ \beta \geq 1/2. \end{cases}$$

The proof of this theorem is given in Section 4.3. Note that if $K = I$ we arrive precisely at the statement in [9], i.e. (4.9). Similarly, in order to optimize in T one would need more information about the function f. For instance, when C_T^f grows like a power of T, the optimization is very straightforward. Using [9, Lemma 4.3], we obtain the following corollary:

4.5 Corollary. *Let $\beta \in (0, 1)$ and f be a $C_{T,\beta}^f$-regular function. Let $K = K_1 \otimes V_2$ with a unitary V. Suppose that $C_{T,\beta}^f \leq C \, T^{2c}$ for some $c, C > 0$. Then there is an explicitly computable constant M depending only on the smallest non-zero eigenvalue of ρ, β, $\|K\|$, C and c, such that,*

$$\|\sigma_1^\beta K \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta K \rho^{1/2-\beta}\|_2 \leq M(S_f^K(\rho\|\sigma) - S_f^{K_1}(\rho_1\|\sigma_1))^{\alpha(\beta)} , \tag{4.15}$$

where

$$\alpha(\beta) = \begin{cases} \frac{\beta(1-\beta)}{1+2c(1-\beta)} & when \ \beta \leq 1/2 \\ \frac{1}{2}\frac{1-\beta}{1+c} & when \ \beta \geq 1/2. \end{cases}$$

In particular, for $\beta = 1/2$,

$$\|\sigma_1^{1/2} K \rho_1^{-1/2} \rho^{1/2} - \sigma^{1/2} K\|_2 \leq M(S_f^K(\rho\|\sigma) - S_f^{K_1}(\rho_1\|\sigma_1))^{\frac{1}{4(1+c)}} . \tag{4.16}$$

Moreover, for $\beta = 1/2$ we can relate the expression $\|\sigma_1^{1/2}K\rho_1^{-1/2}\rho^{1/2} - \sigma^{1/2}K\|_2$ with the one involving Petz recovery map:

4.6 Corollary. *Let f be a $C_{T,1/2}^f$-regular function. Suppose that $C_{T,1/2}^f \leq C\,T^{2c}$ for some $c, C > 0$. Then there is an explicitly computable constant M depending only on the smallest non-zero eigenvalue of ρ, $\|K\|$, C and c, such that,*

$$\|\mathscr{R}_\rho(K_1^*\sigma_1 K_1) - K^*\sigma K\|_1 \leq M(S_f^K(\rho\|\sigma) - S_f^{K_1}(\rho_1\|\sigma_1))^{\frac{1}{4(1+c)}} . \tag{4.17}$$

Proof. In Lemma 2.2 in [8] take

$$X = \sigma_1^{1/2}K\rho_1^{-1/2}\rho^{1/2}$$

and

$$Y = \sigma^{1/2}K.$$

Since V is unitary, $K^*\sigma_1 K = K_1\sigma_1 K_1 \otimes I$. Therefore, we have that

$$\|\sigma_1^{1/2}K\rho_1^{-1/2}\rho^{1/2} - \sigma^{1/2}K\|_2 \geq \frac{1}{2}\|\mathscr{R}_\rho(K_1^*\sigma_1 K_1) - K^*\sigma K\|_1,$$

where $\mathscr{R}_\rho$ is a Petz recovery map. □

Note that if K is invertible, we may interchange the roles of ρ and σ:

4.7 Corollary. *Let an invertible operator K be such that $K = K_1 \otimes V$, where V is a unitary operator. Let f be a $C_{T,1/2}^f$-regular function. Suppose that $C_{T,1/2}^f \leq C\,T^{2c}$ for some $c, C > 0$. Then there is an explicitly computable constant M depending only on the smallest non-zero eigenvalue of ρ, $\|K\|$, C and c, such that,*

$$\frac{1}{2}\|\rho_1\|^{-1/2}\|K^{-1}\|^{-1}\|K^*\mathscr{R}_\sigma\left((K_1^{-1})^{-1}\rho_1 K_1^{-1}\right)K - \rho\|_1 \leq M(S_f^K(\rho\|\sigma) - S_f^{K_1}(\rho_1\|\sigma_1))^{\frac{1}{4(1+c)}} . \tag{4.18}$$

Proof. Recalling L_A being a left multiplication operation,

$$L_{\rho_1^{1/2}}L_{K_{-1}}L_{\sigma_1^{-1/2}}(\sigma_1^{1/2}K\rho_1^{-1/2}\rho^{1/2} - \sigma^{1/2}K) = \rho^{1/2} - \rho_1^{1/2}K^{-1}\sigma_1^{-1/2}\sigma^{1/2}K ,$$

and hence

$$\|\rho^{1/2} - \rho_1^{1/2}K^{-1}\sigma_1^{-1/2}\sigma^{1/2}K\|_2 \leq \|L_{\rho_1^{1/2}}\|\|L_{K^{-1}}\|\|L_{\sigma_1^{-1/2}}\|\|\sigma_1^{1/2}K\rho_1^{-1/2}\rho^{1/2} - \sigma^{1/2}K\|_2 . \tag{4.19}$$

Since $\|L_{\rho_1^{1/2}}\| = \|\rho_1\|^{1/2}$, $\|L_{\sigma_1^{-1/2}}\| = \|\sigma_1^{-1}\|^{1/2}$, and $\|L_{K^{-1}}\| = \|K^{-1}\|$ we may combine (4.19) with Theorem 4.4 to obtain

$$\|\rho_1\|^{-1/2}\|K^{-1}\|^{-1}\|\sigma_1^{-1}\|^{1/2}\|\rho^{1/2} - \rho_1^{1/2}K^{-1}\sigma_1^{-1/2}\sigma^{1/2}K\|_2 \leq \tag{4.20}$$

$$\frac{1}{\pi}T^{1/2}\left(C_T^f\right)^{1/2}(S_f^K(\rho\|\sigma) - S_f^{K_1}(\rho_1\|\sigma_1))^{1/2} + \frac{4}{\pi T^{1/2}}(\|\Delta_{\sigma,\rho}\| + \|K\|) . \tag{4.21}$$

which is the analog of Theorem 4.4 with a drfferent constant on the right, but the roles of ρ and σ interchanged there. Note that since V in unitary, K is invertible if and only if K_1 is. Moreover, $(K^{-1})^{-1}\rho_1 K^{-1} = (K_1^{-1})^{-1}\rho_1 K_1^{-1} \otimes I$. Then using Lemma 2.2 in [8] once more, we obtain

$$\|\rho^{1/2} - \rho_1^{1/2} K^{-1}\sigma_1^{-1/2}\sigma^{1/2}K\|_2 \geq \frac{1}{2}\|K^*\mathscr{R}_\sigma\left((K_1^{-1})^{-1}\rho_1 K_1^{-1}\right)K - \rho\|_1.$$

$\square$

4.3   Proof

We are inspired by the proof of (4.9) in [9]. And for completeness sake we provide all statements here as well. First, we recall Sharma's proof [34] of the monotonicity of the quasi relative entropies S_f^K for operator convex f and $K = K_1 \otimes I_2$, and modify it accordingly for $K = K \otimes V$, where V is unitary.

Proof of Theorem 4.2 (Sharma '14). Define the operator U mapping on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$U(X) = (X_1 \otimes V)\rho_1^{-1/2}\rho^{1/2} . \tag{4.22}$$

The adjoint operator on $\mathcal{H}$ is given by

$$U^*(Y) = \text{Tr}_2(Y\rho^{1/2}(I_1 \otimes V^*))\rho_1^{-1/2} \tag{4.23}$$

for all Y on $\mathcal{H}$.

Then note that U is an isometry on $\mathcal{B}(\mathcal{H}_1)$

$$\langle U(X_1), U(Y_1)\rangle = \text{Tr}\left(\rho^{1/2}\rho_1^{1/2}(X_1^* \otimes V^*)(Y_1 \otimes V)\rho_1^{-1/2}\rho^{1/2}\right) \tag{4.24}$$

$$= \text{Tr}\left(\rho\left(\rho_1^{-1/2}X_1^*Y_1^*\rho_1^{-1/2} \otimes V^*V\right)\right) \tag{4.25}$$

$$= \text{Tr}\left(\rho_1^{1/2}X_1^*Y_1^*\rho_1^{-1/2}\right) \tag{4.26}$$

$$= \langle X_1, Y_1\rangle . \tag{4.27}$$

Now observe that for all X_1 on $\mathcal{H}_1$,

$$U^*\Delta_{\sigma,\rho}U(X_1) = \text{Tr}_2\left(\sigma(X_1\rho_1^{-1/2} \otimes V)\rho^{1/2}\rho^{-1}\rho^{1/2}(I \otimes V^*)\rho_1^{-1/2}\right) \tag{4.28}$$

$$= \text{Tr}(\sigma(X_1\rho_1^{-1} \otimes VV^*)) \tag{4.29}$$

$$= \sigma_1 X_1 \rho_1^{-1} \tag{4.30}$$

$$= \Delta_{\sigma_1,\rho_1}(X_1). \tag{4.31}$$

By the operator Jensen inequality

$$f\left(U^*\Delta_{\sigma,\rho}U\right) \leq U^*f(\Delta_{\sigma,\rho})U . \tag{4.32}$$

Combining (4.31) and (4.32), and using the fact that

$$U(K_1\rho_1^{1/2}) = K\rho^{1/2},\qquad(4.33)$$

we obtain

$$
\begin{aligned}
S_f^{K_1}(\rho_1||\sigma_1) &= \langle K_1\rho_1^{1/2}, f(\Delta_{\sigma_1,\rho_1})\left(K_1\rho_1^{1/2}\right)\rangle \\
&\leq \langle U\left(K_1\rho_1^{1/2}\right), f(\Delta_{\sigma,\rho})U\left(K_1\rho_1^{1/2}\right)\rangle \\
&= \langle K\rho^{1/2}, f(\Delta_{\sigma,\rho})K\rho^{1/2}\rangle = S_f^K(\rho||\sigma) \ .
\end{aligned}
$$

This proves the monotonicity theorem for the quasi relative entropy S_f^K for every operator convex function f and operator $K = K_1 \otimes V$ for any unitary V. $\qquad\square$

We will use [8, Lemma 2.1], the statement of which is the following:

4.8 Lemma (Carlen, Vershynina '17). *Let U be a partial isometry embedding a Hilbert space $\mathcal{K}$ into a Hilbert space $\mathcal{H}$. Let B be an invertible positive operator on $\mathcal{K}$, A be an invertible positive operator on $\mathcal{H}$, and suppose that $U^*AU = B$. Then for all $v \in \mathcal{K}$,*

$$\langle v, U^*A^{-1}Uv\rangle = \langle v, B^{-1}v\rangle + \langle w, Aw\rangle \ ,\qquad(4.34)$$

where

$$w := UB^{-1}v - A^{-1}Uv \ .\qquad(4.35)$$

Proof of Theorem 4.4. For an operator monotone decreasing function f, according to the integral representation (2.1) the quasi-relative entropy S_f^K can we written as

$$S_f^K(\rho||\sigma) = -a\mathrm{Tr}(K^*\sigma K) - b\mathrm{Tr}(K\rho K^*) + \int_0^\infty \left(S_{(t)}^K(\rho||\sigma) - \frac{t}{t^2+1}\mathrm{Tr}(K^*K\rho)\right)d\mu_f(t) \ ,$$

for $a \geq 0$ and $b \in \mathbb{R}$. Here

$$S_{(t)}^K(\rho||\sigma) = \mathrm{Tr}(\sqrt{\rho}K^*(t\mathbf{1} + \Delta_{\sigma,\rho})^{-1}K\sqrt{\rho}).$$

Because $\mathrm{Tr}((K_1 \otimes V)A(K_1 \otimes V)^*) = \mathrm{Tr}(K_1 A_1 K_1)$, it is clear that the difference between relative entropies can be written in terms of the $S_{(t)}$-family,

$$S_f^K(\rho||\sigma) - S_f^{K_1}(\rho_1||\sigma_1) = \int_0^\infty \left(S_{(t)}^K(\rho||\sigma) - S_{(t)}^{K_1}(\rho_1||\sigma_1)\right)d\mu_f(t) \ .\qquad(4.36)$$

We apply Lemma 4.8 with $A := (t\mathbf{1} + \Delta_{\sigma,\rho})$, $B = (t\mathbf{1} + \Delta_{\sigma_1,\rho_1})$ and $v := K_1\rho_1^{1/2}$, and with U defined as above. The lemma's condition, $U^*AU = B$, follows from (4.31) and the fact that $U^*U(X_1) = X_1$ for any X_1 on $\mathcal{H}_1$.

Therefore, applying Lemma 4.8 with (4.33),

$$
\begin{aligned}
S_{(t)}^K(\rho||\sigma) - S_{(t)}^{K_1}(\rho_1||\sigma_1) &= \langle K\rho^{1/2}, (t\mathbf{1} + \Delta_{\sigma,\rho})^{-1}K\rho^{1/2}\rangle - \langle K_1\rho_1^{1/2}, (t\mathbf{1} + \Delta_{\sigma_1,\rho_1})^{-1}K_1\rho_1^{1/2}\rangle \\
&= \langle w_t, (t\mathbf{1} + \Delta_{\sigma,\rho})w_t\rangle \geq t\|w_t\|^2,
\end{aligned}
\tag{4.37}
$$

where,

$$
w_t := U(t\mathbf{1} + \Delta_{\sigma_1,\rho_1})^{-1}(K_1\rho_1^{1/2}) - (t\mathbf{1} + \Delta_{\sigma,\rho})^{-1}K\rho^{1/2} .
\tag{4.38}
$$

Notice that by definition of U (4.22) and (4.33)

$$
-w_t = U[t^{-1}\mathbf{1} - (t\mathbf{1} + \Delta_{\sigma_1,\rho_1})^{-1}](K_1\rho_1^{1/2}) - [t^{-1}\mathbf{1} - (t\mathbf{1} + \Delta_{\sigma,\rho})^{-1}]K\rho^{1/2} .
$$

Since U is an isometry on $\mathcal{B}(\mathcal{H}_1)$,

$$
\|w_t\| \leq \|[t^{-1}\mathbf{1} - (t\mathbf{1} + \Delta_{\sigma_1,\rho_1})^{-1}](K_1\rho_1^{1/2})\| + \|[t^{-1}\mathbf{1} - (t\mathbf{1} + \Delta_{\sigma,\rho})^{-1}]K\rho^{1/2}\| .
$$

Since the the modular operator is non-negative, $0 \leq t^{-1}\mathbf{1} - (t\mathbf{1} + \Delta_{\sigma_1,\rho_1})^{-1} \leq t^{-1}\mathbf{1}$, with the analogous estimate valid with $\Delta_{\sigma,\rho}$ in place of Δ_{σ_1,ρ_1}, Therefore,

$$
\|w_t\| \leq 2t^{-1}\|K\| .
\tag{4.39}
$$

Now using the integral representation of the power function (recall that $\beta \in (0,1)$)

$$
X^\beta = \frac{\sin\beta\pi}{\pi} \int_0^\infty t^\beta \left(\frac{1}{t}\mathbf{1} - \frac{1}{t+X}\right) \mathrm{d}t,
$$

and (4.33) once more, we conclude that

$$
U(\Delta_{\sigma_1,\rho_1})^\beta(K_1\rho_1^{1/2}) - (\Delta_{\sigma,\rho})^\beta K\rho^{1/2} = -\frac{\sin\beta\pi}{\pi} \int_0^\infty t^\beta w_t \mathrm{d}t .
\tag{4.40}
$$

On the other hand,

$$
\begin{aligned}
U(\Delta_{\sigma_1,\rho_1})^\beta(K_1\rho_1^{1/2}) - (\Delta_{\sigma,\rho})^\beta K\rho^{1/2} &= U(\sigma_1^\beta K_1\rho_1^{1/2-\beta}) - \sigma^\beta K\rho^{1/2-\beta} \\
&= \sigma_1^\beta K\rho_1^{-\beta}\rho^{1/2} - \sigma^\beta K\rho^{1/2-\beta} .
\end{aligned}
$$

Combining the last two equalities, and taking the Hilbert space norm associated with $\mathcal{H}$, for any $T_L, T_R > 0$,

$$
\begin{aligned}
\|\sigma_1^\beta K\rho_1^{-\beta}\rho^{1/2} - \sigma^\beta K\rho^{1/2-\beta}\|_2 &= \frac{\sin\beta\pi}{\pi} \left\|\int_0^\infty t^\beta w_t \mathrm{d}t\right\|_2 \\
&\leq \frac{\sin\beta\pi}{\pi} \int_0^{1/T_L} t^\beta \|w_t\|_2 \mathrm{d}t + \frac{\sin\beta\pi}{\pi} \int_{1/T_L}^{T_R} t^\beta \|w_t\|_2 \mathrm{d}t + \frac{\sin\beta\pi}{\pi} \left\|\int_{T_R}^\infty t^\beta w_t \mathrm{d}t\right\|_2 .
\end{aligned}
\tag{4.41}
$$

Let us look at these three terms separately. The first term can be bounded using (4.39):

$$
\int_0^{1/T_L} t^\beta \|w_t\|_2 \mathrm{d}t \leq 2 \int_0^{1/T_L} t^{\beta-1} \mathrm{d}t = \frac{2}{\beta}\frac{\|K\|}{T_L^\beta} .
\tag{4.42}
$$

The third term in (4.41) can be bounded the following way: For any positive operator $X > 0$,

$$t^\beta \left(\frac{1}{t} \mathbf{1} - \frac{1}{t+X} \right) \leq t^\beta \left(\frac{1}{t} - \frac{1}{t + \|X\|} \right) \mathbf{1} = \frac{t^{\beta-1} \|X\|}{(t + \|X\|)} \mathbf{1},$$

and hence

$$\int_T^\infty t^\beta \left(\frac{1}{t} \mathbf{1} - \frac{1}{t+X} \right) \mathrm{d}t \leq \|X\|^\beta \left(\int_{T/\|X\|}^\infty \frac{t^{\beta-1}}{1+t} \mathrm{d}t \right) \mathbf{1} \leq \frac{\|X\|}{(1-\beta)T^{1-\beta}} \mathbf{1} \ .$$

Since spectra of σ_1 and ρ_1 lie in the convex hulls of the spectra of σ and ρ respectively, it follows that $\|\Delta_{\sigma_1,\rho_1}\| \leq \|\Delta_{\sigma,\rho}\|$. Therefore, recalling the definition of w_t, we obtain

$$\left\| \int_{T_R}^\infty t^\beta w_t \mathrm{d}t \right\|_2 \leq \frac{2\|\Delta_{\sigma,\rho}\|}{(1-\beta) T_R^{1-\beta}} \ . \tag{4.43}$$

The second term can be bounded using Cauchy-Schwartz inequality and the fact that f is regular, i.e. there is a constant C_{T_L,T_R}^f such that $dt \leq C_{T_L,T_R}^f \mathrm{d}\mu_f(t)$ for $t \in [1/T_L, T_R]$.

Case 1: $\beta \leq 1/2$.

$$
\begin{aligned}
\left(\int_{1/T_L}^{T_R} t^\beta \|w_t\|_2 \mathrm{d}t \right)^2 &\leq T_R \int_{1/T_L}^{T_R} t^{2\beta} \|w_t\|_2^2 \mathrm{d}t \\
&\leq T_R T_L^{1-2\beta} \int_{1/T_L}^{T_R} t \|w_t\|_2^2 \mathrm{d}t \\
&\leq T_R T_L^{1-2\beta} \int_{1/T_L}^{T_R} S_{(t)}(\rho\|\sigma) - S_{(t)}(\rho_1\|\sigma_1) \mathrm{d}t \\
&\leq T_R T_L^{1-2\beta} C_{T_L,T_R}^f \int_{1/T_L}^{T_R} S_{(t)}(\rho\|\sigma) - S_{(t)}(\rho_1\|\sigma_1) \mathrm{d}\mu_f(t) \\
&\leq T_R T_L^{1-2\beta} C_{T_L,T_R}^f \left(S_f(\rho\|\sigma) - S_f(\rho_1\|\sigma_1) \right) \ . \tag{4.44}
\end{aligned}
$$

Therefore, combining (4.41), (4.42), (4.43), and (4.44) we have

$$
\begin{aligned}
\frac{\pi}{\sin \beta\pi} \|\sigma_1^\beta K \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta K \rho^{1/2-\beta}\|_2 &\leq \frac{2\|K\|}{\beta T_L^\beta} + \frac{2\|\Delta_{\sigma,\rho}\|}{(1-\beta) T_R^{1-\beta}} \\
&\quad + T_R^{1/2} T_L^{1/2-\beta} \left(C_{T_L,T_R}^f \right)^{1/2} (S_f(\rho\|\sigma) - S_f(\rho_1\|\sigma_1))^{1/2} \ .
\end{aligned}
$$

Taking $T_L := T$ and $T_R := T^{\beta/(1-\beta)}$ we obtain

$$
\begin{aligned}
&\frac{\pi}{\sin \beta\pi} \|\sigma_1^\beta K \rho_1^{-\beta} \rho^{1/2} - \sigma^\beta K \rho^{1/2-\beta}\|_2 \\
&\leq 2 \left(\frac{\|K\|}{\beta} + \frac{\|\Delta_{\sigma,\rho}\|}{1-\beta} \right) \frac{1}{T^\beta} + T^{\frac{1-\beta}{2} + \frac{\beta^2}{2(1-\beta)}} \left(C_{T,\beta}^f \right)^{1/2} (S_f(\rho\|\sigma) - S_f(\rho_1\|\sigma_1))^{1/2} \ .
\end{aligned}
$$

Case 2: $\beta > 1/2$.

$$\left(\int_{1/T_L}^{T_R} t^\beta \|w_t\|_2 \mathrm{d}t\right)^2 \leq T_R \int_{1/T_L}^{T_R} t^{2\beta} \|w_t\|_2^2 \mathrm{d}t$$

$$\leq T_R^{2\beta} \int_{1/T_L}^{T_R} t\|w_t\|_2^2 \mathrm{d}t$$

$$\leq T_R^{2\beta} \int_{1/T_L}^{T_R} (S_{(t)}(\rho||\sigma) - S_{(t)}(\rho_1||\sigma_1)) \mathrm{d}t$$

$$\leq T_R^{2\beta} C_{T_L,T_R}^f \int_{1/T_L}^{T} (S_{(t)}(\rho||\sigma) - S_{(t)}(\rho_1||\sigma_1)) \mathrm{d}\mu_f(t)$$

$$\leq T_R^{2\beta} C_{T_L,T_R}^f (S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1)) . \tag{4.45}$$

Therefore, combining (4.41), (4.42), (4.43), and (4.45) we have

$$\frac{\pi}{\sin\beta\pi} \|\sigma_1^\beta K\rho_1^{-\beta}\rho^{1/2} - \sigma^\beta K\rho^{1/2-\beta}\|_2 \leq \frac{2\|K\|}{\beta T_L^\beta} + \frac{2\|\Delta_{\sigma,\rho}\|}{(1-\beta)T_R^{1-\beta}}$$
$$+ T_R^\beta \left(C_{T,\beta}^f\right)^{1/2} (S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{1/2} .$$

Taking $T_L := T^{(1-\beta)/\beta}$ and $T_R := T$ we obtain

$$\frac{\pi}{\sin\beta\pi} \|\sigma_1^\beta K\rho_1^{-\beta}\rho^{1/2} - \sigma^\beta K\rho^{1/2-\beta}\|_2$$
$$\leq 2\left(\frac{\|K\|}{\beta} + \frac{\|\Delta_{\sigma,\rho}\|}{1-\beta}\right)\frac{1}{T^{1-\beta}} + T^\beta \left(C_{T,\beta}^f\right)^{1/2} (S_f(\rho||\sigma) - S_f(\rho_1||\sigma_1))^{1/2} .$$

$\square$

4.4 Condition for equality

Corollary 4.5 and the proof of Theorem 4.4 give a condition on the equality in the monotonicity inequality.

4.9 Corollary. *Let f be a regular function. The equality in the monotonicity inequality*

$$S_f^K(\rho||\sigma) - S_f^{K_1}(\rho_1||\sigma_1) = 0, \tag{4.46}$$

holds if and only if for all $\beta \in \mathbb{C}$ the following holds:

$$\sigma_1^\beta K\rho_1^{-\beta} = \sigma^\beta K\rho^{-\beta}. \tag{4.47}$$

Proof. If the equality in the monotonicity inequality holds, then (4.47) holds for all $\beta \in (0,1)$ following the Corollary 4.5. Since for any positive matrix X, the map $\beta \to X^\beta$ is an entire analytic function, this identity holds for all $\beta \in \mathbb{C}$.

The other way, suppose (4.47) holds for all $\beta \in \mathbb{C}$. Then from (4.41) in the proof of the Theorem 4.4, we have that

$$U(\Delta_{\sigma_1,\rho_1})^\beta (K_1 \rho_1^{1/2}) - (\Delta_{\sigma,\rho})^\beta K \rho^{1/2} = 0 \ ,$$

for all $\beta \in \mathbb{C}$. Let us use the following Taylor series expansion

$$\frac{1}{t+x} = \sum_{n=0}^\infty \frac{(-1)^n}{t^{n+1}} x^n \ .$$

Then, using the above two equalities, we obtain that for all $t \geq 0$,

$$w_t = U(t\mathbf{1} + \Delta_{\sigma_1,\rho_1})^{-1}(K_1 \rho_1^{1/2}) - (t\mathbf{1} + \Delta_{\sigma,\rho})^{-1} K \rho^{1/2} = 0 \ .$$

From (4.37) this implies that $S_{(t)}^K(\rho\|\sigma) - S_{(t)}^{K_1}(\rho_1\|\sigma_1) = 0$, and therefore, the (4.46) is satisfied, following the integral representation (4.36).

$\square$

5 Joint convexity of the quasi-relative entropy

As it was shown by Petz [26, 28] or [31, Theorem 2], the quasi relative entropy is jointly convex in ρ and σ. Here is another elegant proof of a joint convexity.

5.1 Proposition. *For an operator monotone decreasing function f, and any operator K, quasi-entropy $S_f^K(\rho\|\sigma)$ is jointly convex in $\rho, \sigma > 0$. In other words, for $\rho = \sum_j p_j \rho_j$ and $\sigma = \sum_j p_j \sigma_j$,*

$$0 \leq \sum_j p_j S_f^K(\rho_j\|\sigma_j) - S_f^K(\rho\|\sigma).$$

Moreover, the equality in the joint convexity holds if and only if

$$(\Delta_{\sigma,\rho} + tI)^{-1}(K) = (\Delta_{\sigma_j,\rho_j} + tI)^{-1}(K), \ \textit{for all } j \ \textit{and for all } t > 0.$$

Proof. Using (2.1) representation of operator monotone decreasing function, we obtain

$$
\begin{aligned}
S_f^K(\rho\|\sigma) = {} & - a\mathrm{Tr}K^*\sigma K - b\mathrm{Tr}K^*K\rho + \int_0^\infty \left\{ \mathrm{Tr}\sqrt{\rho}K^* \frac{1}{\Delta_{\sigma,\rho}+t}(K\sqrt{\rho}) - \frac{t}{t^2+1}\mathrm{Tr}K^*K\rho \right\} d\mu_f(t) \\
= {} & - a\mathrm{Tr}K^*\sigma K - b\mathrm{Tr}K^*K\rho + \int_0^\infty \left\{ \mathrm{Tr}\rho K^* \frac{1}{L_\sigma + tR_\rho}(K\rho) - \frac{t}{t^2+1}\mathrm{Tr}K^*K\rho \right\} d\mu_f(t) \\
= {} & - a\mathrm{Tr}K^*\sigma K - b\mathrm{Tr}K^*K\rho + \int_0^\infty \left\{ \mathrm{Tr}\rho K^* \frac{1}{\Delta_{\sigma,\rho}+t}(K) - \frac{t}{t^2+1}\mathrm{Tr}K^*K\rho \right\} d\mu_f(t).
\end{aligned}
$$

$$(5.1)$$

The joint convexity follows immediately from that of the map $(Y, A, B) \to \mathrm{Tr}Y^* \frac{1}{L_B + tR_A}(Y)$, which was proved in [32].

Note that equality in the joint convexity holds if and only if it holds for the first term in the integrand.

$\square$

The next theorem uses monotonicity inequality in Theorem 4.4 and provides a strengthening of the convexity inequality.

5.2 Theorem. *Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, $\beta \in (0,1)$, $K \in \mathcal{B}(\mathcal{H})$. Let f be a $C^f_{T,\beta}$-regular function. Then for states $\rho = \sum_j p_j \rho_j$ and $\sigma = \sum_j p_j \sigma_j$ (with $p_j > 0$ and $\sum_j p_j = 1$), we have*

$$\frac{\pi}{\sin \beta\pi} \sum_j p_j^{1/2} \|\sigma^\beta K \rho^{-\beta} \rho_j^{1/2} - \sigma_j^\beta K \rho_j^{1/2-\beta}\|_2$$

$$\leq 2 \left(\frac{\|K\|}{\beta} + \frac{\sum_j p_j^{-1}\|\rho_j^{-1}\|}{1-\beta} \right) \frac{1}{T^{\alpha_1(\beta)}} + T^{\alpha_2(\beta)} \left(C^f_{T,\beta} \right)^{1/2} \left(\sum_j p_j S^K_f(\rho_j\|\sigma_j) - S^K_f(\rho\|\sigma) \right)^{1/2} \tag{5.2}$$

where

$$\alpha_1(\beta) = \begin{cases} \beta & \text{when } \beta \leq 1/2 \\ 1-\beta & \text{when } \beta \geq 1/2. \end{cases} \quad \text{and} \quad \alpha_2(\beta) = \begin{cases} \frac{1-\beta}{2} + \frac{\beta^2}{2(1-\beta)} & \text{when } \beta \leq 1/2 \\ \beta & \text{when } \beta \geq 1/2. \end{cases}$$

Proof. Let us form the following quantum-classical states

$$\overline{\rho} := \sum_j p_j \rho_j \otimes |j\rangle \langle j|_X,$$

$$\overline{\sigma} := \sum_j p_j \sigma_j \otimes |j\rangle \langle j|_X.$$

Then $\overline{\rho}_1 = \rho$ and $\overline{\sigma}_1 = \sigma$, i.e.

$$S^K_f(\overline{\rho}_1\|\overline{\sigma}_1) = S^K_f(\rho\|\sigma).$$

Let us use (5.1) for the expression of the quasi-relative entropy. There we see that all but one term are linear in ρ and σ. For the term with the modular operator, note that for any $A = \sum_j A_j \otimes |j\rangle \langle j|_X$ we have

$$\Delta_{\overline{\sigma},\overline{\rho}}(A) = \sum_j \Delta_{\sigma_j,\rho_j}(A_j) \otimes |j\rangle \langle j|_X. \tag{5.3}$$

Since $K = K \otimes I_X$, we may apply this to a term in (5.1), and obtain

$$\mathrm{Tr}\rho K^* \frac{1}{\Delta_{\overline{\sigma},\overline{\rho}} + t}(K) = \sum_j p_j \mathrm{Tr} \left\{ \rho K^* \frac{1}{\Delta_{\sigma_j,\rho_j} + t}(K) \right\}.$$

Therefore, from (5.1)

$$S^K_f(\overline{\rho}\|\overline{\sigma}) = \sum_j p_j S^K_f(\rho_j\|\sigma_j).$$

Note that

$$\|\Delta_{\overline{\sigma},\overline{\rho}}\| \leq \|\overline{\sigma}\|\|\overline{\rho}^{-1}\| \leq \|\overline{\rho}^{-1}\| \leq \sum_j p_j^{-1}\|\rho_j^{-1}\|.$$

Applying Theorem 4.4 will give us the right-hand side of (5.2). The left-hand side results from the following identity

$$\overline{\sigma}_1^\beta K \overline{\rho}_1^{-\beta} \overline{\rho}^{1/2} - \overline{\sigma}^\beta K \overline{\rho}^{1/2-\beta} = \sum_j p_j^{1/2} \left\{ \sigma^\beta K \rho^{-\beta} \rho_j^{1/2} - \sigma_j^\beta K \rho_j^{1/2-\beta} \right\} \otimes |j\rangle \langle j|_X .$$

Taking the Hilbert-Schmidt norm (4.7) on both sides will result in the correct left-hand side. $\square$

5.3 Corollary. *Let $\beta \in (0,1)$ and f be a $C_{T,\beta}^f$-regular function. Suppose that $C_{T,\beta}^f \le C \, T^{2c}$ for some $c, C > 0$. Let $\rho = \sum_j p_j \rho_j$ and $\sigma = \sum_j p_j \sigma_j$ (with $p_j > 0$ and $\sum_j p_j = 1$). Then there is an explicitly computable constant M depending only on the smallest non-zero eigenvalues of $\{\rho_j\}_j$, $\|K\|$, $\{p_j^{-1}\}_j$, β, C and c, such that,*

$$\sum_j p_j^{1/2} \|\sigma^\beta K \rho^{-\beta} \rho_j^{1/2} - \sigma_j^\beta K \rho_j^{1/2-\beta}\|_2 \le M \left(\sum_j p_j S_f^K(\rho_j \| \sigma_j) - S_f^K(\rho \| \sigma) \right)^{\alpha(\beta)}, \tag{5.4}$$

where

$$\alpha(\beta) = \begin{cases} \frac{\beta(1-\beta)}{1+2c(1-\beta)} & \text{when } \beta \le 1/2 \\ \frac{1}{2} \frac{1-\beta}{1+c} & \text{when } \beta \ge 1/2. \end{cases}$$

In particular, for $\beta = 1/2$,

$$\sum_j p_j^{1/2} \|\sigma^{1/2} K \rho^{-1/2} \rho_j^{1/2} - \sigma_j^{1/2} K\|_2 \le M \left(\sum_j p_j S_f^K(\rho_j \| \sigma_j) - S_f^K(\rho \| \sigma) \right)^{1/4(c+1)}. \tag{5.5}$$

5.1 Condition for equality

From Corollary 4.9 and the proof of Theorem 5.2 we obtain the condition on the equality in the joint convexity inequality.

5.4 Corollary. *An equality in the joint convexity inequality*

$$\sum_j p_j S_f^K(\rho_j \| \sigma_j) = S_f^K(\rho \| \sigma),$$

holds if and only if, for all j and all $\beta \in \mathbb{C}$

$$\sigma^\beta K \rho^{-\beta} = \sigma_j^\beta K \rho_j^{-\beta} .$$

6 Operator inequalities

In this section consider a tri-partite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Let $\rho = \rho_{ABC}$ be a state on $\mathcal{H}$. Then the strong subadditivity of quantum entropy is the following statement:

6.1 Theorem (Lieb, Ruskai '73)**.** *For ρ_{ABC} a state on $\mathcal{H}_{ABC}$, it holds that*

$$0 \leq S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) \ . \tag{6.1}$$

This theorem was proved by Lien and Ruskai [20], using Lieb's theorem that was proved in [18]. The theorem in a von Neumann algebra setting was done by Narnhofer and Thirring in [24].

Let $\sigma = \sigma_{AB} \otimes I_C$ be a state on $\mathcal{H}$. Let f be an operator monotone decreasing function. In [33] Rusaki, building on the work of Kim [17], showed that the following operator on $\mathcal{H}_C$ is positive semi-definite

$$0 \leq \mathrm{Tr}_{AB} \left([f(L_{\sigma_{AB}} R_{\rho_{ABC}}^{-1}) - f(L_{\sigma_B} R_{\rho_{BC}}^{-1})] \rho_{ABC} \right) \ . \tag{6.2}$$

In particular, taking $\sigma_{AB} = \rho_{AB}$, and $f(x) = -\log(x)$, reduces to an operator inequality of Kim's [17]

$$0 \leq \mathrm{Tr}_{AB}[\log \rho_{ABC} - \log \rho_{AB} - \log \rho_{BC} + \log \rho_B] \rho_{ABC} \ . \tag{6.3}$$

We prove the following sharpening of (6.2) inequality.

6.2 Theorem. *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, $\beta \in (0,1)$ and $\rho = \rho_{ABC}$, $\sigma = \sigma_{AB} \otimes I_C$. Let f be a $C_{T,\beta}^f$-regular function. Suppose that $C_{T,\beta}^f \leq C\,T^{2c}$ for some $c, C > 0$. Then there is an explicitly computable positive constant N depending only on the smallest non-zero eigenvalue of ρ, $\|K\|$, β, C and c, such that, the following operator inequality holds on $\mathcal{H}_C$*

$$N[\mathrm{Tr}_{AB}(P_{ABC}(\rho, \sigma) P_{ABC}^*(\rho, \sigma))]^{1/\alpha(\beta)} \leq \mathrm{Tr}_{AB} \left([f(L_{\sigma_{AB}} R_{\rho_{ABC}}^{-1}) - f(L_{\sigma_B} R_{\rho_{BC}}^{-1})] \rho_{ABC} \right) \ . \tag{6.4}$$

where

$$P_{ABC}(\rho, \sigma) := \sigma_B^\beta \rho_{BC}^{-\beta} \rho_{ABC}^{1/2} - \sigma_{AB}^\beta \rho_{ABC}^{1/2-\beta} \ ,$$

and

$$\alpha(\beta) = \begin{cases} \frac{\beta(1-\beta)}{1+2c(1-\beta)} & \text{when } \beta \leq 1/2 \\ \frac{1}{2} \frac{1-\beta}{1+c} & \text{when } \beta \geq 1/2. \end{cases}$$

The proof of this theorem is given in Section 6.1. Note that since the left-hand side of (6.4) is positive semi-definite on $\mathcal{H}_C$, this theorem implies result in [33], i.e. (6.2).

In the proof of Theorem 6.2, if one takes $\rho_{ABC} = \rho_{AB} \otimes I_C$ and σ_{ABC} (i.e. in (6.6)), we would obtain

6.3 Theorem. *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, $\beta \in (0,1)$ and $\rho = \rho_{AB} \otimes I_C$, σ_{ABC} be a state on $\mathcal{H}$. Let f be a $C_{T,\beta}^f$-regular function. Suppose that $C_{T,\beta}^f \leq C\,T^{2c}$ for some $c, C > 0$. Then there is an explicitly computable positive constant N depending only on the smallest non-zero eigenvalue of ρ, $\|K\|$, β, C and c, such that, the following operator inequality holds on $\mathcal{H}_C$*

$$N[\mathrm{Tr}_{AB}(Q_{ABC}^*(\rho_{AB}, \sigma_{ABC}) Q_{ABC}(\rho_{AB}, \sigma_{ABC}))]^{1/\alpha(\beta)} \leq \mathrm{Tr}_{AB} \left([f(L_{\sigma_{ABC}} R_{\rho_{AB}}^{-1}) - f(L_{\sigma_{BC}} R_{\rho_B}^{-1})] \rho_{AB} \right) \ ,$$

where

$$Q_{ABC}(\rho_{AB}, \sigma_{ABC}) := \sigma_{BC}^\beta \rho_B^{-\beta} \rho_{AB}^{1/2} - \sigma_{ABC}^\beta \rho_{AB}^{1/2-\beta} \ .$$

Let us take the function $\tilde{f}(x) = xf(1/x)$. By [4, Theorem V.2.9] this function is operator monotone decreasing if and only if f is. Taking $\tilde{f}$ instead of f in Theorem 6.2 and 6.3 and interchanging the roles of ρ and σ, leads to

6.4 Corollary. *With the same conditions and notations as in Theorem 6.3, we have*

$$N[\mathrm{Tr}_{AB}(P_{ABC}(\sigma,\rho)P^*_{ABC}(\sigma,\rho))]^{1/\alpha(\beta)} \le \mathrm{Tr}_{AB}\left(\rho_{AB}[f(L^{-1}_{\rho_{AB}}R_{\sigma_{ABC}}) - f(L^{-1}_{\rho_B}R_{\sigma_{BC}})]\right) .$$

6.5 Corollary. *With the same conditions and notations as in Theorem 6.2, we have*

$$N[\mathrm{Tr}_{AB}(Q^*_{ABC}(\sigma_{AB},\rho_{ABC})Q_{ABC}(\sigma_{AB},\rho_{ABC})]^{1/\alpha(\beta)} \le \mathrm{Tr}_{AB}\left(\rho_{ABC}[f(L^{-1}_{\rho_{ABC}}R_{\sigma_{AB}}) - f(L^{-1}_{\rho_{BC}}R_{\sigma_B})]\right) .$$

6.1 Proof

Proof of Theorem 6.2. We are inspired by the proof of Ruskai [33], which we provide in our case in all detail for the completeness sake.

In the monotonicity inequality Corollary 4.5, let us consider $\rho = \rho_{ABC}$, $\sigma = \sigma_{ABC}$, $\mathcal{H}_1 = \mathcal{H}_{BC}$, $\mathcal{H}_2 := \mathcal{H}_A$, and $K_{ABC} = I_A \otimes K_{BC}$. Then (4.11) is equivalent to

$$N\|\sigma^\beta_{BC}K_{BC}\rho^{-\beta}_{BC}\rho^{1/2}_{ABC} - \sigma^\beta_{ABC}K_{BC}\rho^{1/2-\beta}_{ABC}\|^{1/\alpha(\beta)}_2 \le S^{K_{BC}}_f(\rho_{ABC}\|\sigma_{ABC}) - S^{K_{BC}}_f(\rho_{BC}\|\sigma_{BC}) . \quad (6.5)$$

Let us consider the difference on right-hand side in the above inequality.

$$\begin{aligned}
D := & S^{K_{BC}}_f(\rho_{ABC}\|\sigma_{ABC}) - S^{K_{BC}}_f(\rho_{BC}\|\sigma_{BC}) \\
= & \mathrm{Tr}_{ABC}(\rho^{1/2}_{ABC}K^*_{BC}f(\Delta_{\sigma_{ABC},\rho_{ABC}})(K_{BC}\rho^{1/2}_{ABC})) - \mathrm{Tr}_{ABC}(\rho^{1/2}_{BC}K^*_{BC}f(\Delta_{\sigma_{BC},\rho_{BC}})(K_{BC}\rho^{1/2}_{BC})) \\
= & \mathrm{Tr}_{ABC}(K^*_{BC}f(\Delta_{\sigma_{ABC},\rho_{ABC}})(K_{BC}\rho_{ABC})) - \mathrm{Tr}_{ABC}(K^*_{BC}f(\Delta_{\sigma_{BC},\rho_{BC}})(K_{BC}\rho_{BC})) \quad (6.6)
\end{aligned}$$

In the first equality we used the definition of the quasi-relative entropy and the fact that there is no dependence on A in the second term. In the second equality we used the definition of the modular operator $\Delta_{A,B} = L_A R^{-1}_B$.

Consider a special case when $\sigma_{ABC} = \sigma_{AB} \otimes I_C$. Then

$$D = \mathrm{Tr}_{ABC}(K^*_{BC}f(\Delta_{\sigma_{AB},\rho_{ABC}})(K_{BC}\rho_{ABC})) - \mathrm{Tr}_{ABC}(K^*_{BC}f(\Delta_{\sigma_B,\rho_{BC}})(K_{BC}\rho_{BC})) . \quad (6.7)$$

Choose $K_{BC} = I_B \otimes K_C$. Then $K = I_{AB} \otimes K_C$ commutes with $\sigma_{ABC} = \sigma_{AB} \otimes I_C$. Therefore,

$$D = \mathrm{Tr}_{ABC}(K^*_C K_C f(\Delta_{\sigma_{AB},\rho_{ABC}})(\rho_{ABC})) - \mathrm{Tr}_{ABC}(K^*_C K_C f(\Delta_{\sigma_B,\rho_{BC}})(\rho_{BC})) .$$

Furthermore, take $K_C = |\phi\rangle\langle\phi|_C$ to be a projector onto a vector $|\phi\rangle$. Then

$$D = \langle\phi|_C \mathrm{Tr}_{AB}\left(f(\Delta_{\sigma_{AB},\rho_{ABC}})(\rho_{ABC}) - f(\Delta_{\sigma_B,\rho_{BC}})(\rho_{BC})\right)|\phi\rangle_C . \quad (6.8)$$

Bringing this expression back into (6.5), we obtain

$$\begin{aligned}
N\|\sigma^\beta_B |\phi\rangle\langle\phi|_C \rho^{-\beta}_{BC}\rho^{1/2}_{ABC} - & \sigma^\beta_{AB}|\phi\rangle\langle\phi|_C \rho^{1/2-\beta}_{ABC}\|^{1/\alpha(\beta)}_2 \le \\
& \langle\phi|_C \mathrm{Tr}_{AB}\left([f(\Delta_{\sigma_{AB},\rho_{ABC}}) - f(\Delta_{\sigma_B,\rho_{BC}})](\rho_{ABC})\right)|\phi\rangle_C . \quad (6.9)
\end{aligned}$$

Denoting

$$P_{ABC}(\rho, \sigma) := \sigma_B^\beta \rho_{BC}^{-\beta} \rho_{ABC}^{1/2} - \sigma_{AB}^\beta \rho_{ABC}^{1/2-\beta} ,$$

we calculate the norm on the left-hand side by definition (4.7)

$$\|\sigma_B^\beta |\phi\rangle \langle\phi|_C \rho_{BC}^{-\beta} \rho_{ABC}^{1/2} - \sigma_{AB}^\beta |\phi\rangle \langle\phi|_C \rho_{ABC}^{1/2-\beta}\|_2 = \| |\phi\rangle \langle\phi|_C [\sigma_B^\beta \rho_{BC}^{-\beta} \rho_{ABC}^{1/2} - \sigma_{AB}^\beta \rho_{ABC}^{1/2-\beta}]\|_2 \quad (6.10)$$

$$= \text{Tr}_{ABC}(|\phi\rangle \langle\phi|_C P_{ABC}(\rho, \sigma) P_{ABC}^*(\rho, \sigma)) \quad (6.11)$$

$$= \langle\phi|_C [\text{Tr}_{AB}(P_{ABC}(\rho, \sigma) P_{ABC}^*(\rho, \sigma))] |\phi\rangle_C . \quad (6.12)$$

Therefore, we have for all $|\phi\rangle_C$

$$\langle\phi|_C [\text{Tr}_{AB}(P_{ABC}(\rho, \sigma) P_{ABC}^*(\rho, \sigma))] |\phi\rangle_C \leq \langle\phi|_C \text{Tr}_{AB} ([f(\Delta_{\sigma_{AB}, \rho_{ABC}}) - f(\Delta_{\sigma_B, \rho_{BC}})](\rho_{ABC})) |\phi\rangle_C .$$

Since the above inequality holds for all $|\phi\rangle_C$ in $\mathcal{H}_C$, we have the following operator inequality

$$N[\text{Tr}_{AB}(P_{ABC}(\rho, \sigma) P_{ABC}^*(\rho, \sigma))]^{1/\alpha(\beta)} \leq \text{Tr}_{AB} ([f(L_{\sigma_{AB}} R_{\rho_{ABC}}^{-1}) - f(L_{\sigma_B} R_{\rho_{BC}}^{-1})](\rho_{ABC})) .$$

$\square$

6.2   Condition for equality

From the proof above it is evident that in Theorems 6.2 and 6.3 and Corollaries 6.4 and 6.5 the right-hand side is zero if and only if for all $\beta \in (0, 1)$, $P_{ABC}(\rho, \sigma) = 0$ or $Q_{ABC}(\rho, \sigma) = 0$. Let us take an example

6.6 Proposition. *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, $\beta \in (0, 1)$ and $\rho = \rho_{ABC}$, $\sigma = \sigma_{AB} \otimes I_C$. Let f be a $C_{T,\beta}^f$-regular function. The equality*

$$\text{Tr}_{AB}([f(L_{\sigma_{AB}} R_{\rho_{ABC}}^{-1}) \rho_{ABC}) = \text{Tr}_B(f(L_{\sigma_B} R_{\rho_{BC}}^{-1})] \rho_{BC}) \quad (6.13)$$

holds if and only if

$$\sigma_B^\beta \rho_{BC}^{-\beta} = \sigma_{AB}^\beta \rho_{ABC}^{-\beta}, \quad \text{for all } \beta \in (0, 1) . \quad (6.14)$$

Moreover, (6.13) is equivalent to the Petz's recovery condition

$$\mathscr{R}_{\rho_{ABC}}(\sigma_B) = \sigma_{AB} . \quad (6.15)$$

Proof. From Theorem 6.2, it is clear that if (6.13) is satisfied, the following holds

$$0 = \text{Tr}_{AB}(P_{ABC} P_{ABC}^*).$$

Then for all $|\phi\rangle_C$ in $\mathcal{H}_C$,

$$0 = \text{Tr}(P_{ABC} P_{ABC}^* |\phi\rangle_C \langle\phi|_C) .$$

The operator inside trace is positive semi-definite, since it can written as $A^* A$ for $A = |\phi\rangle_C \langle\phi|_C P_{ABC}$. Therefore, the operator itself is zero, i.e. for all $|\phi\rangle_C$

$$0 = |\phi\rangle_C \langle\phi|_C P_{ABC} .$$

Since this holds for all $|\phi\rangle_C$, the operator P_{ABC} is zero, leading to the required (6.14). Taking $\beta = 1/2$, equality (6.14) leads to (6.15).

On the other side, if (6.14) is satisfied, it means that in particular, Petz's recovery map recovers both ρ_{BC} and σ_B perfectly, i.e. (6.15) holds. Therefore, by the equivalence of the recovery of both states and the saturation of the monotonicity inequality, we have that D in (6.6) is zero, i.e. (6.13) is satisfied. $\qquad\square$

7 Logarithmic function

Let us take $f(x) = -\log(x)$ and $K = I$. We may explicitly calculate the power $\alpha(\beta)$. From Example 2.7 we have that $d\mu(x) = dx$. Therefore, in Corollary 5.3 and Theorem 6.2 the constants are: $c = 0$, $C = 1$, and

$$\alpha(\beta) = \begin{cases} \beta(1-\beta) & \text{when } \beta \le 1/2 \\ (1-\beta)/2 & \text{when } \beta \ge 1/2. \end{cases}$$

Moreover, from [9, Corollary 5.1]

$$N = \begin{cases} \left(\frac{\pi(1-2\beta+2\beta^2)\beta}{\sin\beta\pi}\right)^{\frac{1}{\beta(1-\beta)}} \left(\|K\| + \frac{\beta}{1-\beta}D\right)^{-\frac{1-2\beta+2\beta^2}{\beta(1-\beta)}} 2^{-\frac{1-2\beta+2\beta^2}{\beta(1-\beta)}} \left(\frac{1-2\beta+2\beta^2}{2(1-\beta)}\right)^{-2} & \text{when } \beta \le 1/2 \\ \left(\frac{\pi\beta(1-\beta)}{\sin\beta\pi}\right)^{\frac{1}{1-\beta}} \left(\frac{1-\beta}{\beta}\|K\| + D\right)^{-\frac{2\beta}{1-\beta}} 2^{-\frac{2\beta}{1-\beta}}\beta^{-2} & \text{when } \beta \ge 1/2. \end{cases}$$

where $D = \sum_j p_j^{-1}\|\rho_j^{-1}\|$ in Corollary 7.1 below, $D = \|\Delta_{\rho_A \otimes \rho_{BC}, \rho_{ABC}}\|$ in Corollary 7.2, and $D = \|\Delta_{\sigma,\rho}\|$ in Corollary 7.3.

Taking $K = I$ in Corollary 5.3, we obtain the following sharpening of the joint convexity of quantum relative entropy.

7.1 Corollary (Joint convexity). *Let $\rho = \sum_j p_j\rho_j$ and $\sigma = \sum_j p_j\sigma_j$ (with $p_j > 0$ and $\sum_j p_j = 1$). With $\alpha(\beta)$ and N defined above, we have*

$$N\left(\sum_j p_j^{1/2}\|\sigma^\beta\rho^{-\beta}\rho_j^{1/2} - \sigma_j^\beta\rho_j^{1/2-\beta}\|_2\right)^{1/\alpha(\beta)} \le \sum_j p_j S(\rho_j\|\sigma_j) - S(\rho\|\sigma) . \tag{7.1}$$

In particular, for $\beta = 1/2$,

$$N\left(\sum_j p_j^{1/2}\|\sigma^{1/2}\rho^{-1/2}\rho_j^{1/2} - \sigma_j^{1/2}\|_2\right)^4 \le \sum_j p_j S(\rho_j\|\sigma_j) - S(\rho\|\sigma) . \tag{7.2}$$

From the last inequality, we have that if the joint convexity is saturated, i.e.

$$\sum_j p_j S(\rho_j\|\sigma_j) = S(\rho\|\sigma),$$

then for all j and all $\beta \in (0,1)$

$$\sigma^\beta\rho^{-\beta} = \sigma_j^\beta\rho_j^{-\beta} .$$

From the sharpening of the monotonicity inequality, Corollary 4.6, or the previous result (4.11), we obtain the sharpening of the strong subadditivity when taking $\sigma_{ABC} = \rho_{AB} \otimes \rho_C$ and tracing out system A.

7.2 Corollary (Strong subadditivity). *For $\rho := \rho_{ABC}$ a state on $\mathcal{H}_{ABC}$ and $\beta \in (0,1)$, it holds that*

$$N \| \rho_B^{\beta} \otimes \rho_C^{\beta} \rho_{BC}^{-\beta} \rho^{1/2} - \rho_{AB}^{\beta} \otimes \rho_C^{\beta} \rho^{1/2-\beta} \|_2^{1/\alpha(\beta)} \leq S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) .$$

In particular, for $\beta = 1/2$ the excplicit bound involving Petz's recovery map holds

$$\left(\frac{\pi}{8} \right)^4 \| \rho^{-1} \|^{-2} \| \mathcal{R}_\rho(\rho_B \otimes \rho_C) - \rho_{AB} \otimes \rho_C \|_1^4 \leq S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) .$$

Moreover, it is clear that the equality in the strong subadditivity inequality holds if and only if Petz's map $\mathcal{R}_\rho$ recovers state $\rho_{AB} \otimes \rho_C$ perfectly.

7.3 Corollary (Operator strong-subadditivity). *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ and $\beta \in (0,1)$. Let ρ_{ABC} be a state on $\mathcal{H}$, and $\sigma_{ABC} = \sigma_{AB} \otimes I_C$ with σ_{AB} being a state on $\mathcal{H}_A \otimes \mathcal{H}_B$. With $\alpha(\beta)$ and N defined above, the following results hold.*

1. *Theorem 6.2 leads to*

   $$N \left[\mathrm{Tr}_{AB}(P(\rho_{ABC}, \sigma_{AB}) P^*(\rho_{ABC}, \sigma_{AB})) \right]^{1/\alpha(\beta)} \leq \mathrm{Tr}_{AB}[\log \rho_{ABC} - \log \sigma_{AB} - \log \rho_{BC} + \log \sigma_B] \rho_{ABC} .$$

2. *When $\mathcal{H}_B$ is one-dimensional, this becomes*

   $$N \left[\mathrm{Tr}_A(P^*(\rho_{AC}, \sigma_A) P(\rho_{AC}, \sigma_A)) \right]^{1/\alpha(\beta)} \leq \mathrm{Tr}_A[\log \rho_{AC} - \log \sigma_A - \log \rho_C] \rho_{AC} .$$

   Note that $P(\rho_{AC}, \sigma_A) = \rho_C^{-\beta} \rho_{AC}^{1/2} - \sigma_A^{\beta} \rho_{AC}^{1/2-\beta}$.

3. *And when $\sigma_A = \frac{1}{d_A} I_A$ is a maximally mixed state,*

   $$N \left[\mathrm{Tr}_A(P(\rho_{AC}, \sigma_A) P^*(\rho_{AC}, \sigma_A)) \right]^{1/\alpha(\beta)} \leq \mathrm{Tr}_A[(\log \rho_{AC}) \rho_{AC} - (\log \rho_C) \rho_C] + \log(d_A) \rho_C ,$$

   with $P(\rho_{AC}, \sigma_A) = \rho_C^{-\beta} \rho_{AC}^{1/2} - \frac{1}{d_A^{\beta}} \rho_{AC}^{1/2-\beta}$.

4. *Taking $\sigma_{AB} = \rho_{AB}$ in part 1, we obtain*

   $$N \left[\mathrm{Tr}_{AB}(P(\rho_{ABC}, \rho_{AB}) P^*(\rho_{ABC}, \rho_{AB})) \right]^{1/\alpha(\beta)} \leq \mathrm{Tr}_{AB}[\log \rho_{ABC} - \log \rho_{AB} - \log \rho_{BC} + \log \rho_B] \rho_{ABC} , \tag{7.3}$$

   where

   $$P(\rho_{ABC}, \rho_{AB}) = \rho_B^{\beta} \rho_{BC}^{-\beta} \rho_{ABC}^{1/2} - \rho_{AB}^{\beta} \rho_{ABC}^{1/2-\beta} .$$

5. *In particular, taking $\mathcal{H}_B$ to be one-dimensional in the last inequality,*

$$N\left[\text{Tr}_A(P(\rho_A, \rho_{AC})P^*(\rho_A, \rho_{AC}))\right]^{1/\alpha(\beta)} \leq \text{Tr}_A[\log \rho_{AC} - \log \rho_A - \log \rho_C]\rho_{AC} \ , \qquad (7.4)$$

where

$$P(\rho_{AC}, \rho_A) = \rho_C^{-\beta}\rho_{AC}^{1/2} - \rho_A^\beta\rho_{AC}^{1/2-\beta} \ .$$

6. *Theorem 6.3 leads to*

$$N\left[\text{Tr}_{AB}(Q^*(\sigma_{AB}, \rho_{ABC})Q(\sigma_{AB}, \rho_{ABC}))\right]^{1/\alpha(\beta)} \leq \text{Tr}_{AB}\rho_{AB}[-\log \rho_{ABC} + \log \sigma_{AB} + \log \rho_{BC} - \log \sigma_B] \ .$$
$$(7.5)$$

7. *When $\mathcal{H}_B$ is one-dimensional in the last inequality, we obtain*

$$N\left[\text{Tr}_A(Q^*(\sigma_A, \rho_{AC})Q(\sigma_A, \rho_{AC}))\right]^{1/\alpha(\beta)} \leq \text{Tr}_A\rho_A[-\log \rho_{AC} + \log \sigma_A] + \log \rho_C \ .$$

8. *Taking $\rho_{AB} = \sigma_{AB}$ in part 6, we obtain a stronger version of Kim's inequality (6.3)*

$$N\left[\text{Tr}_{AB}(Q^*(\rho_{AB}, \rho_{ABC})Q(\rho_{AB}, \rho_{ABC}))\right]^{1/\alpha(\beta)} \leq \text{Tr}_{AB}\rho_{AB}[-\log \rho_{ABC} + \log \rho_{AB} + \log \rho_{BC} - \log \rho_B] \ .$$

9. *When $\mathcal{H}_B$ is one-dimensional in the last inequality, we obtain*

$$N\left[\text{Tr}_A(Q^*(\rho_A, \rho_{AC})Q(\rho_A, \rho_{AC}))\right]^{1/\alpha(\beta)} \leq \text{Tr}_A\rho_A[-\log \rho_{AC} + \log \rho_A] + \log \rho_C \ .$$

8 Wigner-Yanase-Dyson-type inequalities

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

$$S_{f_p}^K(\rho||\sigma) = \frac{1}{p(1-p)}\text{Tr}(K\rho K^* - K^*\sigma^p K\rho^{1-p}) \ .$$

From Proposition 3.7 we obtain the lower bound on the quasi-relative entropy in terms of the trace distance.

8.1 Corollary (Pinsker inequality)**.** *For a unitary U, $p \in (-1, 2)$ and $p \neq 0, 1$ and states ρ and σ, we have*

$$\frac{1}{p(1-p)}[1 - \text{Tr}(U^*\sigma^p U\rho^{1-p})] \geq \frac{1}{2}||\rho - U^*\sigma U||_1^2 \ .$$

Note that in the above inequality, as well in all below ones, it is important to leave the factor $1/p(1-p)$ in place, as it changed sign at $p = 0, 1$.

For the power function f_p we may explicitly calculate the power $\alpha(\beta)$. Then from Example 2.8 we have $d\mu_f(x) = \frac{\sin(p\pi)}{p(1-p)} x^p \, dx$, for $p \in (0,1)$. Therefore, in Corollary 6.5, we have

$$c = \begin{cases} p/2 & \text{when } \beta \leq 1/2 \\ p(1-\beta)/(2\beta) & \text{when } \beta \geq 1/2. \end{cases},$$

and therefore,

$$\alpha(\beta) = \begin{cases} \frac{\beta(1-\beta)}{1+p(1-\beta)} & \text{when } \beta \leq 1/2 \\ \frac{\beta(1-\beta)}{2\beta+p(1-\beta)} & \text{when } \beta \geq 1/2. \end{cases}$$

Moreover, from [9, Corollary 5.2] for $\beta \leq 1/2$, the constant N is defined as

$$N = (\|K\| + \frac{\beta}{1-\beta}\|\Delta_{\sigma,\rho}\|)^{-\frac{p(1-\beta)+1-2\beta+2\beta^2}{\beta(1-\beta)}} 2^{-\frac{p(1-\beta)+1-2\beta+2\beta^2}{\beta(1-\beta)}} \tag{8.1}$$

$$\cdot \frac{\sin(p\pi)}{\pi} \left(\frac{\pi\beta(p(1-\beta)+1-2\beta+2\beta^2)}{(1+p(1-\beta))\sin\beta\pi} \right)^{\frac{1+p(1-\beta)}{\beta(1-\beta)}} \left(\frac{p(1-\beta)+1-2\beta+2\beta^2}{2(1-\beta)} \right)^{-2},$$

and for $\beta \geq 1/2$,

$$N = \left(\frac{1-\beta}{\beta}\|K\| + \|\Delta_{\sigma,\rho}\| \right)^{-\frac{2\beta^2+p(1-\beta)}{\beta(1-\beta)}} 2^{-\frac{2\beta^2+p(1-\beta)}{\beta(1-\beta)}} \tag{8.2}$$

$$\cdot \frac{\sin p\pi}{\pi} \left(\frac{\pi(1-\beta)(2\beta^2+p(1-\beta))}{(2\beta+p(1-\beta))\sin\beta\pi} \right)^{\frac{2\beta+p(1-\beta)}{\beta(1-\beta)}} \left(\frac{2\beta^2+p(1-\beta)}{2\beta} \right)^{-2}.$$

Recall, from part 5 of Example 3.2, we have that the quasi-relative entropy is the Wigner-Yanase-Dyson p-skew information $I_p(\rho, K)$ for $K^* = K$, i.e.

$$S_{f_p}^K(\rho\|\rho) = -\frac{1}{2p(1-p)}\text{Tr}[K, \rho^p][K, \rho^{1-p}] = -\frac{1}{p(1-p)}I_p(\rho, K) .$$

It was conjectured by Wigner and Yanase in [44] that p-skew information $I_p(\rho, K)$ is concave as a function of a density matrix ρ for a fixed $p \in (0,1)$. A more general expression

$$S_{f_p}^K(\rho\|\sigma) = \frac{1}{p(1-p)}\text{Tr}(K\rho K^* - K^*\sigma^p K\rho^{1-p}),$$

shows that the concavity of WYD information follows from the joint concavity of the term $\text{Tr}(K^*\sigma^p K\rho^{1-p})$, since the first term in the above expression is linear in ρ. The concavity of this term was shown by Lieb [18] for powers of ρ and σ that sum up to a number no greater than one.

Using Corollary 5.3, we have the following strengthening of the joint concavity.

8.2 Corollary (Joint concavity of WYD information). *Let $p \in (-1, 2)$ and $p \neq 0, 1$, and $\rho = \sum_j p_j \rho_j$ and $\sigma = \sum_j p_j \sigma_j$ (with $p_j > 0$ and $\sum_j p_j = 1$). With $\alpha(\beta)$ and M defined above, we have*

$$N \left(\sum_j p_j^{1/2} \| \sigma^\beta K \rho^{-\beta} \rho_j^{1/2} - \sigma_j^\beta K \rho_j^{1/2-\beta} \|_2 \right)^{1/\alpha(\beta)} \leq \frac{1}{p(1-p)} \left(\mathrm{Tr} K^* \sigma^p K \rho^{1-p} - \sum_j p_j \mathrm{Tr} K^* \sigma_j^p K \rho_j^{1-p} \right) \quad (8.3)$$

In particular, for $\beta = 1/2$,

$$N \left(\sum_j p_j^{1/2} \| \sigma^{1/2} K \rho^{-1/2} \rho_j^{1/2} - \sigma_j^{1/2} K \|_2 \right)^4 \leq \frac{1}{p(1-p)} \left(\mathrm{Tr} K^* \sigma^p K \rho^{1-p} - \sum_j p_j \mathrm{Tr} K^* \sigma_j^p K \rho_j^{1-p} \right) .$$
$$(8.4)$$

From Theorem 6.3 we might obtain some interesting sharpening of the operator version of Wigner-Yanase-Dyson inequalities. In [33] Ruskai showed that taking f_p in the operator version of SSA, with ρ_{ABC} and σ_{AB} leads to

$$0 \leq \frac{1}{p(1-p)} [-\mathrm{Tr}_{AB} \rho_{ABC}^{1-p} \sigma_{AB}^p + \mathrm{Tr}_B \rho_{BC}^{1-p} \sigma_B^p] .$$

From Corollary 6.5 we obtain the error term for the above difference:

8.3 Corollary (Operator version of WYD inequality). *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, $\beta \in (0, 1)$, $p \in (-1, 2)$, $p \neq 0, 1$ and $\rho = \rho_{ABC}$, $\sigma = \sigma_{AB} \otimes I_C$. For constants defined above, the following operator inequality holds*

$$N \left[\mathrm{Tr}_{AB} (Q_{ABC}^*(\sigma_{AB}, \rho_{ABC}) Q_{ABC}(\sigma_{AB}, \rho_{ABC})) \right]^{1/\alpha(\beta)} \leq \frac{1}{p(1-p)} \left[-\mathrm{Tr}_{AB} \rho_{ABC}^{1-p} \sigma_{AB}^p + \mathrm{Tr}_B \rho_{BC}^{1-p} \sigma_B^p \right],$$
$$(8.5)$$

where

$$Q_{ABC}(\sigma_{AB}, \rho_{ABC}) := \rho_{BC}^\beta \sigma_B^{-\beta} \sigma_{AB}^{1/2} - \rho_{ABC}^\beta \sigma_{AB}^{1/2-\beta} .$$

In particular, when $\mathcal{H}_B$ is one-dimensional, we have

$$N \left[\mathrm{Tr}_A (Q_{AC}^*(\sigma_A, \rho_{AC}) Q_{AC}(\sigma_A, \rho_{AC})) \right]^{1/\alpha(\beta)} \leq \frac{1}{p(1-p)} \left[-\mathrm{Tr}_A \rho_{AC}^{1-p} \sigma_A^p + \rho_C^{1-p} \right] .$$

8.1 Cauchy-Schwartz inequality

In particular, taking $p = 2$ in f_p defined above, gives the following right-hand side in Corollary 8.3

$$\mathrm{Tr}_{AB} (\sigma_{AB}^2 \rho_{ABC}^{-1} - \sigma_B^2 \rho_{BC}^{-1}) = \mathrm{Tr}_{AB} (\sigma_{AB} \rho_{ABC}^{-1} \sigma_{AB}) - \mathrm{Tr}_B (\sigma_B \rho_{BC}^{-1} \sigma_B) .$$

In [21] Lieb and Ruskai showed the positivity of the following operator

$$\mathrm{Tr}_A X_{AC}^* Q_{AC}^{-1} X_{AC} \geq X_C^* Q_C^{-1} X_C,$$

for any X_{AC} and positive semi-definite Q_{AC} with $\ker Q_{AC} \subseteq \ker X_{AC}^*$.

From Corollary 8.3 we obtain the following sharpening of the Cauchy-Schwarz inequality:

8.4 Corollary. *Let ρ_{ABC} and $\sigma = \sigma_{AB} \otimes I_C$ be two states on the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Let $\beta \in (0,1)$. Then*

$$N\left[\text{Tr}_{AB}(Q^*_{ABC}(\sigma_{AB}, \rho_{ABC})Q_{ABC}(\sigma_{AB}, \rho_{ABC})\right]^{1/\alpha(\beta)} \leq \text{Tr}_{AB}(\sigma_{AB}\rho^{-1}_{ABC}\sigma_{AB}) - \text{Tr}_B(\sigma_B\rho^{-1}_{BC}\sigma_B) \ ,$$

where, recall, $Q_{ABC}(\sigma_{AB}, \rho_{ABC}) := \rho^{\beta}_{BC}\sigma^{-\beta}_B\sigma^{1/2}_{AB} - \rho^{\beta}_{ABC}\sigma^{1/2-\beta}_{AB}$.

Moreover, the equality

$$\text{Tr}_{AB}(\sigma_{AB}\rho^{-1}_{ABC}\sigma_{AB}) = \text{Tr}_B(\sigma_B\rho^{-1}_{BC}\sigma_B)$$

holds if and only if

$$\rho^{\beta}_B\sigma^{-\beta}_{BC} = \rho^{\beta}_{AB}\sigma^{-\beta}_{ABC}, \quad \text{for all } \beta \in (0,1) \ .$$

In particular, for $\beta = 1/2$, it is equivalent to the Petz's recovery condition

$$\mathscr{R}_{\sigma_{AB}}(\rho_{BC}) = \rho_{ABC} \ .$$

The reasoning for the equality condition if is similar to the one in Proposition 6.6.

References

[1] H. Araki, "Relative Entropy of States of von Neumann Algebras", Publ. RIMS Kyoto Univ. 9: 809-833, (1976)

[2] H. Araki, E. H. Lieb, "Entropy Inequalities", Commun. Math. Phys. 18: 160-170, (1970)

[3] M. Berta, M. Tomamichel, "The fidelity of recovery is multiplicative", IEEE Trans. Inf. Theory 62.4: 1758-1763, (2016)

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

[5] F. G. S. L. Brandão, A. W. Harrow, J. Oppenheim, S. Strelchuk, "Quantum conditional mutual information, reconstructed states, and state redistribution", Phys. Rev. Lett. 115.5: 050501, (2015)

[6] E. A. Carlen, E.H. Lieb, "Bounds for Entanglement via an Extension of Strong Subadditivity of Entropy", Lett. Math. Phys. 101: 1-11, (2012)

[7] E. A. Carlen, E. H. Lieb, "Remainder Terms for Some Quantum Entropy Inequalities", J. Math. Phys. 55: 042201, (2014)

[8] E. Carlen, A. Vershynina, "Recovery map stability for the Data Processing Inequality", preprint, arxiv:1710.02409

[9] E. Carlen, A. Vershynina, "Recovery and the Data Processing Inequality for quasi-entropies", IEEE Trans. Info. Th. 64(10): 6929-6938, (2018)

[10] I. Csiszár, "Information type measure of difference of probability distributions and indirect observations", Studia Sci. Math. Hungar. 2: 299-318, (1967)

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

[12] O. Fawzi, R. Renner, "Quantum conditional mutual information and approximate Markov chains", Commun. Math. Phys. 340.2: 575-611, (2015)

[13] P. Hayden, R. Jozsa, D. Petz, A. Winter, "Structure of states which satisfy strong subadditivity of quantum entropy with equality", Comm. Math. Phys. 246.2: 359-374, (2004)

[14] F. Hiai, M. Mosonyi, "Different quantum f-divergences and the reversibility of quantum operations," Rev. Math. Phys. 29.07: 1750023, (2017)

[15] 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: 1099-1121, (2010)

[16] M. Junge, R. Renner, D. Sutter, M. Wilde, A. Winter, "Universal recovery from a decrease of quantum relative entropy", Ann.Henri Poincare 19.10: 2955-2978, (2015)

[17] I. H. Kim, "Operator extension of strong subadditivity of entropy", J. Math. Phys. 53.12: 122204, (2012)

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

[19] E. H. Lieb, "Some Convexity and Subadditivity Properties of Entropy", Bull. AMS 81.1: 1-13, (1975)

[20] E. H. Lieb, M. B. Ruskai, "Proof of the strong subadditivity of quantum-mechanical entropy", J. Math. Phys. 14.12: 1938-1941, (1973)

[21] E. H. Lieb, M.B. Ruskai, "Some Operator Inequalities of the Schwarz Type," Adv. Math. 12: 269-273, (1974)

[22] G. Lindblad, "Expectations and Entropy Inequalities for Finite Quantum Systems", Commun. Math. Phys. 39: 111-119, (1974)

[23] G. Lindblad, "Completely Positive Maps and Entropy Inequalities", Commun. Math. Phys. 40.2: 147-151, (1975)

[24] H. Narnhofer, W. Thirring, "From Relative Entropy to Entropy", Fizika 17: 257-265, (1985)

[25] M. Nielsen, D. Petz, "A simple proof of the strong subadditivity inequality", Q, Info. Comp., 5.6: 507-513, (2005)

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

[27] D. Petz, "Quasi-entropies for states of a von Neumann algebra", Publ. RIMS. Kyoto Univ. 21: 781-800, (1985)

[28] D. Petz, "Quasi-Entropies for Finite Quantum Systems", Rep. Math. Phys. 23: 57-65, (1986)

[29] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von Neumann algebra", Comm. Math. Phys. 105.1: 123-131, (1986).

[30] D. Petz, "Sufficiency of channels over von Neumann algebras", Quart. J. Math. Oxford Ser. (2), 39.153: 97-108, (1988)

[31] D. Petz, "From quasi-entropy", arXiv:1009.2679 (2010)

[32] M. B. Ruskai, "Inequalities for Quantum Entropy: A Review with Conditions for Equality", J. Math. Phys. 43: 4358, (2002); erratum 46: 019901, (2005)

[33] M. B. Ruskai, "Remarks on Kim's Strong Subadditivity Matrix Inequality: Extensions and Equality Conditions", J. Math. Phys. 54.10: 102202, (2013)

[34] N. Sharma, "Equality condition for quantum quasi-entropies under monotonicity and joint-convexity", Communications (NCC), 2014 Twentieth National Conference on. IEEE: 1-6, (2014)

[35] W. F. Stinespring, "Positive Functions on C*-algebras", Proc. American Math. Soc., **6**, 211-216 (1955)

[36] D. Sutter, O. Fawzi, R. Renner, "Universal recovery map for approximate Markov chains", Proc. R. Soc. Lond. A Math. Phys. Eng. Sci. 472.2186: 20150623, (2016)

[37] D. Sutter, M. Tomamichel, A.W. Harrow, "Strengthened monotonicity of relative entropy via pinched Petz recovery map", IEEE Trans. Inf. Theory 62.5: 2907-2913, (2016)

[38] A. Ulhmann, "Relative Entropy and the Wigner-Yanase-Dyson-Lieb Concavity in an Interpolation Theory", Comm. Math. Phys, 54: 21-32, (1977)

[39] H. Umegaki, "Conditional Expectation in an Operator Algebra. IV. Entropy and Information", Kodai Math. Sem. Rep. 14: 59-85, (1962)

[40] I. Vajda, "Note on discrimination information and variation", IEEE Trans. Inf. Theory, 16.6: 771-773, (1970)

[41] D. Virosztek, "Quantum entropies, relative entropies, and related preserver problems", Ph.D. Thesis, (2016)

[42] J. Watrous, "The theory of quantum information", Cambridge University Press, 2018

[43] E. P. Wigner, M. M. Yanase, "Information Content of Distributions", Proc. Natl. Acad. Sci. USA 49: 910-918, (1963)

[44] E. P. Wigner, M. M. Yanase, "On the Positive Semi-Definite Nature of a Certain Matrix Expression", Can. J. Math. 16: 397-406, (1964)

[45] M. Wilde, "Recoverability in quantum information theory", Proc. R. Soc. A. 471: 20150338, The Royal Society, (2015)

[46] L. Zhang, "A Strengthened Monotonicity Inequality of Quantum Relative Entropy: A Unifying Approach Via Rényi Relative Entropy", Lett. Math. Phys. 106.4: 557-573, (2016)