# On the quantum $f$-relative entropy and generalized data processing inequalities

**Naresh Sharma**

Tata Institute of Fundamental Research
Mumbai 400 005, India
Email: `nsharma@tifr.res.in`

January 9, 2019

### Abstract

We study the fundamental properties of the quantum $f$-relative entropy, where $f(\cdot)$ is an operator convex function. We give the equality conditions under various properties including monotonicity and joint convexity, and these conditions are more general than, since they hold for a class of operator convex functions, and different for $f(t) = -\ln(t)$ from, the previously known conditions. The quantum $f$-entropy is defined in terms of the quantum $f$-relative entropy and we study its properties giving the equality conditions in some cases. We then show that the $f$-generalizations of the Holevo information, the entanglement-assisted capacity, and the coherent information also satisfy the data processing inequality, and give the equality conditions for the $f$-coherent information.

## 1 Introduction

Quantum entropy is central to the study of information processing in quantum mechanical systems (see [1, 2] and references therein). The von Neumann entropy for a density matrix $\rho$, a positive semi-definite matrix ($\rho \geq 0$) with unit trace ($\text{Tr}(\rho) = 1$), is given by

$$S(\rho) = -\text{Tr}\left[\rho \ln(\rho)\right]. \tag{1}$$

Schumacher's quantum noiseless channel coding theorem gives an information-theoretic interpretation of this quantity [3]. Lieb and Ruskai showed that the von Neumann entropy satisfies, among other inequalities, the strong sub-additivity [4, 5] given by

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC}), \tag{2}$$

where $ABC$ is the composite system consisting of subsystems $A$, $B$, and $C$ with the density matrix $\rho_{ABC}$, and the density matrices of subsystem(s) is obtained by tracing out other subsystem(s). For example, $\rho_{AB} = \text{Tr}_C(\rho_{ABC})$. The equality conditions for the strong sub-additivity were given by Hayden $et\ al$ [6].

Umegaki defined the quantum relative entropy of $\rho$ to $\sigma$ as [7]

$$S(\rho||\sigma) = \text{Tr}\left\{\rho\left[\ln(\rho) - \ln(\sigma)\right]\right\}, \tag{3}$$

where $\rho$, $\sigma$ are density matrices. Lindblad proved the monotonicity of the quantum relative entropy, which is stated as

$$S(\rho_{AB}||\sigma_{AB}) \geq S(\rho_A||\sigma_A), \tag{4}$$

where $\rho_A = \text{Tr}_B(\rho_{AB})$ and $\sigma_A = \text{Tr}_B(\sigma_{AB})$ [8]. The equality conditions for the quantum relative entropy under monotonicity were given by Petz [9] and Ruskai [10], and these conditions are equivalent though not the same and are obtained using different approaches. Ibinson, Linden, and Winter later showed that monotonicity under restrictions is the only general inequality satisfied by quantum relative entropy [11]. The joint convexity property of the quantum relative entropy is stated for $0 \leq \lambda \leq 1$, and density matrices $\sigma_i$, $\rho_i$, $i = 1, 2$, as

$$S\left[\lambda\rho_1 + (1-\lambda)\rho_2||\lambda\sigma_1 + (1-\lambda)\sigma_2\right] \leq \lambda S(\rho_1||\sigma_1) + (1-\lambda)S(\rho_2||\sigma_2). \tag{5}$$

Ruskai [10] gave the equality conditions under joint convexity. Ruskai describes an elegant way of deducing strong sub-additivity and joint convexity from monotonicity in Ref. [10].

## 1.1 Operator convex functions

If $A$ is Hermitian and has a spectral decomposition given by

$$A = \sum_i \alpha_i |i\rangle\langle i|, \tag{6}$$

then the matrix valued function $f(A)$ is defined as

$$f(A) = \sum_i f(\alpha_i)|i\rangle\langle i|, \tag{7}$$

where we have implicitly assumed that the spectrum of $A$ lies in the domain of $f$.

A real valued function $f(\cdot)$ is said to be operator convex if for all Hermitian matrices $A$ and $B$, and $0 \leq \lambda \leq 1$,

$$f\left[\lambda A + (1-\lambda)B\right] \leq \lambda f(A) + (1-\lambda)f(B). \tag{8}$$

It is easy to see that if $f(t)$ is operator convex, then so is $g(t) = f(t - t_0)$ for some $t_0 \in \mathbb{R}$ assuming that $t - t_0$ is in the domain of $f(\cdot)$.

It follows from Theorem V.4.6 in Ref. [12] that for a non-affine operator convex function $f(\cdot)$ on $(-a + t_0, a + t_0)$, there exists a unique probability density function $p_X(x)$ defined on $x \in [-a, a]$ such that for $d > 0$,

$$f(t) = b + ct + d \int_{-a}^{a} \frac{(t - t_0)^2}{a^2 - (t - t_0)x} p_X(x) dx. \tag{9}$$

Choosing $t_0 = a$, $b = 1 - \ln(a)$, $c = -1/a$, $d = 1/2$, and

$$p_X(x) = \begin{cases} -\frac{2}{a^2}x, & -a \leq x \leq 0, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

we get

$$f(t) = -\ln(t), \quad t \in (0, 2a). \tag{11}$$

**Definition**: An operator convex function is said to be *diffused* if the probability density function, $p_X(x)$, in Eq. (9) is strictly positive a.e. in a subinterval of $[-a, a]$ that has $x = 0$ as the interior or the boundary point.

The function $f(t) = -\ln(t)$ clearly belongs to this class of operator convex functions since $p_X(x)$ is strictly positive $\forall\, x \in [-a, 0)$ and $x = 0$ is a boundary point of this interval.

We mention the operator Jensen's inequality that we shall use more than once in this paper.

**Theorem 1** *(Hansen and Pedersen [13]) Let $E_i$, $i = 1, ..., n$, be a set of matrices satisfying*

$$\sum_{i=1}^{n} E_i^{\dagger} E_i = \mathrm{I}. \tag{12}$$

*Then, for Hermitian matrices $\phi_i$, $i = 1, ..., n$, with bounded spectra, and an operator convex function $f(\cdot)$,*

$$f\left(\sum_{i=1}^{n} E_i^{\dagger} \phi_i E_i\right) \leq \sum_{i=1}^{n} E_i^{\dagger} f(\phi_i) E_i. \tag{13}$$

## 1.2 Quantum $f$-relative entropy

Let $A$ be a $m \times n$ matrix

$$A = [a_{ij}], \quad i = 1, ..., m, \quad j = 1, ...n, \tag{14}$$

where $a_{ij}$ is the $(i, j)$th entry of $A$. We denote conjugate, transpose, and conjugate transpose of $A$ by $A^*$, $A^{\mathsf{T}}$, and $A^{\dagger}$ respectively. One can associate a vector with matrix $A$, denoted by $\mathrm{vec}\,(A)$, whose $[n(i - 1) + j]$th entry, denoted by $\mathrm{vec}\,(A)_{n(i-1)+j}$ is given by

$$\mathrm{vec}\,(A)_{n(i-1)+j} = a_{ij}. \tag{15}$$

3

One can, of course, construct $A$ back from $\text{vec}\,(A)$. An identity that we shall frequently employ is [14]

$$\text{vec}\,(ABC) = (A \otimes C^{\mathsf{T}})\text{vec}\,(B)\,, \tag{16}$$

where $A$, $B$, $C$ are matrices with appropriate dimensions.

It is well known that many properties of the quantum relative entropy are not central to the $\ln(\cdot)$ used in its definition and a more general definition of quantum relative entropy is studied in [1, 15]. In the classical case, the $f$-generalization of the classical relative entropy was studied by Csiszár [16].

The quantum $f$-relative entropy for strictly positive $\rho$ and $\sigma$ is defined as

$$S_f(\rho||\sigma) = \text{vec}\,(\sqrt{\rho})^{\dagger}\, f\left[\sigma \otimes \left(\rho^{-1}\right)^{\mathsf{T}}\right]\text{vec}\,(\sqrt{\rho})\,, \tag{17}$$

where $f(\cdot)$ is an operator convex function. We shall implicitly assume that the domain of $f(\cdot)$ is contained in $(0, a)$ for some finite $a > 0$. Note that we don't impose the condition that $\rho$ and $\sigma$ have unit trace. Let the spectral decompositions of $\rho$ and $\sigma$ in Eq. (17) be given by

$$\rho = \sum_{i=1}^{d} p_i |i_\rho\rangle\langle i_\rho|, \tag{18}$$

$$\sigma = \sum_{j=1}^{d} q_j |j_\sigma\rangle\langle j_\sigma|, \tag{19}$$

where $d$ is the dimension of the Hilbert space that describes $\rho$ and $\sigma$. Using Eqs. (16), (17), (18), and (19), we can also write the quantum $f$-relative entropy as

$$S_f(\rho||\sigma) = \text{vec}\,(\mathrm{I})^{\dagger}\,[\mathrm{I} \otimes (\sqrt{\rho})^*]\, f\left[\sigma \otimes \left(\rho^{-1}\right)^{\mathsf{T}}\right]\left[\mathrm{I} \otimes (\sqrt{\rho})^{\mathsf{T}}\right]\text{vec}\,(\mathrm{I})\,, \tag{20}$$

$$= \sum_{i,j=1}^{d} p_i f\left(\frac{q_j}{p_i}\right)|\langle i_\rho|\, j_\sigma\rangle|^2, \tag{21}$$

$$= \sum_{j=1}^{d} \langle j_\sigma|\sqrt{\rho}f(q_j\rho^{-1})\sqrt{\rho}|j_\sigma\rangle, \tag{22}$$

$$= \sum_{i=1}^{d} p_i\,\langle i_\rho|\, f\left(\frac{\sigma}{p_i}\right)|i_\rho\rangle\,, \tag{23}$$

where I is the Identity matrix whose dimensions, if unspecified, would be apparent from the context.

## 1.3 Overview

We follow the "vec" notation throughout this paper as was used in our definition of the quantum $f$-relative entropy in Eq. (17) as opposed to the linear super-operators in [1, 17, 15].

We shall see that this notation along with the operator Jensen's inequality in Theorem 1 gives alternate and more accessible proofs of many inequalities and equality conditions.

We note here that the quantum $f$-relative entropy defined in Eq. (17) is a special case of quantum quasi relative entropy defined by Petz [1, 17]. However, we shall see that we arrive at the equality conditions under certain properties for a class of operator convex functions, which are more general than, and different in case of $f(t) = -\ln(t)$ from, those given by Petz [9] and Ruskai [10].

We note here that since we define the quantum $f$-relative entropy for strictly positive matrices, in some cases in this paper, it shall put implicit restrictions. For example, when we deal with the quantum $f$-relative entropy after processing, i.e., $S_f[\mathcal{E}(\rho)||\mathcal{E}(\sigma)]$, where $\mathcal{E}(\cdot)$ is a quantum operation, we shall implicitly assume that $\mathcal{E}(\rho)$ and $\mathcal{E}(\sigma)$ are strictly positive, which puts restrictions on the choices of $\rho$, $\sigma$, and $\mathcal{E}(\cdot)$. A way out could have been to extend the definition of the quantum $f$-relative entropy for positive semi-definite matrices. This could be accomplished by defining the terms of the form $f(0)$, $0 \times f(0/0)$, and $0 \times f(a/0)$, $a > 0$. But we refrain from doing that in this paper since we deal with a class of operator convex functions and leave the extension of the definition to the time when a specific choice of the function $f(\cdot)$ is made in Eq. (17), which we won't do in this paper.

We define the quantum $f$-entropy in terms of the quantum $f$-relative entropy and study some of its properties and give the equality conditions for some cases. We also show the $f$-generalizations of some well-known quantum information-theoretic quantities also satisfy the data processing inequalities as is the case for $f(t) = -\ln(t)$.

## 2  Properties of the quantum $f$-relative entropy

We now list some useful properties of the quantum $f$-relative entropy.

**Lemma 1** *For strictly positive $\rho$ and $\sigma$, the following properties hold:*

*(i) The quantum $f$-relative entropy is invariant under Unitary transformation, i.e.,*

$$S_f\left(U\rho U^\dagger||U\sigma U^\dagger\right) = S_f(\rho||\sigma), \tag{24}$$

*where $U^\dagger U = \mathrm{I}$.*

*(ii) For any strictly positive $\kappa$,*

$$S_f(\rho \otimes \kappa||\sigma \otimes \kappa) = S_f(\rho||\sigma). \tag{25}$$

*(iii) For any scalar $c > 0$,*

$$\frac{1}{c}S_f(c\rho||c\sigma) = S_f(\rho||\sigma). \tag{26}$$

*Proof*  These properties follow easily from Eq. (21) and we omit the proof.  ∎

## 2.1 Monotonicity

Petz [17], Nielsen and Petz [18] provide an elegant proof the monotonicity of the quantum $f$-relative entropy. We restate their proof in the "vec" notation.

**Lemma 2** *(Petz [17], Nielsen and Petz [18]) Let $\rho_{AB}$ and $\sigma_{AB}$ be two strictly positive matrices in the composite system consisting of systems $A$ and $B$, and let $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ and $\sigma_A = \mathrm{Tr}_B(\sigma_{AB})$. Then*

$$S_f(\rho_{AB}||\sigma_{AB}) \geq S_f(\rho_A||\sigma_A). \tag{27}$$

*Proof* Let us assume that there exists a matrix $V$ such that

$$
\begin{align}
V \mathrm{vec}\left(\sqrt{\rho_A}\right) &= \mathrm{vec}\left(\sqrt{\rho_{AB}}\right), \tag{28}\\
V^\dagger V &= \mathrm{I}, \tag{29}\\
V^\dagger \left[\sigma_{AB} \otimes \left(\rho_{AB}^{-1}\right)^\mathsf{T}\right] V &= \sigma_A \otimes \left(\rho_A^{-1}\right)^\mathsf{T}. \tag{30}
\end{align}
$$

To show that such a $V$ does exist, let us consider a linear super-operator $\mathcal{U}(\cdot)$ such that

$$\mathcal{U}(X) = \left(X\rho_A^{-1/2} \otimes \mathrm{I}\right)\sqrt{\rho_{AB}}. \tag{31}$$

Its adjoint is given by

$$\mathcal{U}^\dagger(Y) = \mathrm{Tr}_B\left[Y\sqrt{\rho_{AB}}\left(\rho_A^{-1/2} \otimes \mathrm{I}\right)\right]. \tag{32}$$

That this is indeed the adjoint is evident from

$$\langle \mathcal{U}^\dagger(Y), X\rangle = \langle Y, \mathcal{U}(X)\rangle, \tag{33}$$

where $\langle E, F\rangle = \mathrm{Tr}(E^\dagger F)$ is the Hilbert-Schmidt inner product. Let us associate a matrix $V$ with $\mathcal{U}$ such that

$$\mathrm{vec}\left[\mathcal{U}(X)\right] = V\mathrm{vec}\left(X\right) \tag{34}$$

and hence,

$$\mathrm{vec}\left[\mathcal{U}^\dagger(Y)\right] = V^\dagger\mathrm{vec}\left(Y\right). \tag{35}$$

Note that since

$$
\begin{align}
\mathcal{U}\left(\sqrt{\rho_A}\right) &= \sqrt{\rho_{AB}}, \tag{36}\\
\mathcal{U}^\dagger\left[\mathcal{U}(X)\right] &= X, \tag{37}\\
\mathcal{U}^\dagger\left[\sigma_{AB}\mathcal{U}(X)\rho_{AB}^{-1}\right] &= \sigma_A X\rho_A^{-1}, \tag{38}
\end{align}
$$

6

Eqs. (28), (29) and (30) must hold. We now have

$$
\begin{aligned}
S_f(\rho_{AB}||\sigma_{AB}) &= \text{vec}\left(\sqrt{\rho_{AB}}\right)^\dagger f\left[\sigma_{AB} \otimes \left(\rho_{AB}^{-1}\right)^\mathsf{T}\right] \text{vec}\left(\sqrt{\rho_{AB}}\right) && (39) \\
&= \text{vec}\left(\sqrt{\rho_A}\right)^\dagger V^\dagger f\left[\sigma_{AB} \otimes \left(\rho_{AB}^{-1}\right)^\mathsf{T}\right] V \text{vec}\left(\sqrt{\rho_A}\right) && (40) \\
&\geq \text{vec}\left(\sqrt{\rho_A}\right)^\dagger f\left\{V^\dagger \left[\sigma_{AB} \otimes \left(\rho_{AB}^{-1}\right)^\mathsf{T}\right] V\right\} \text{vec}\left(\sqrt{\rho_A}\right) && (41) \\
&= \text{vec}\left(\sqrt{\rho_A}\right)^\dagger f\left[\sigma_A \otimes \left(\rho_A^{-1}\right)^\mathsf{T}\right] \text{vec}\left(\sqrt{\rho_A}\right) && (42) \\
&= S_f(\rho_A||\sigma_A), && (43)
\end{aligned}
$$

where Eq. (41) follows from Eq. (40) by using the operator Jensen's inequality in Eq. (13) with $n = 1$ and $E_1 = V$. ∎

We now give the conditions for the equality in Eq. (27) for a class of operator convex functions.

**Lemma 3** *For a non-affine and diffused operator convex function $f(\cdot)$, positive $\rho_{AB}$, $\sigma_{AB}$, $\rho_A = \text{Tr}_B(\rho_{AB})$ and $\sigma_A = \text{Tr}_B(\sigma_{AB})$, the equality in*

$$
S_f\left(\rho_{AB}||\sigma_{AB}\right) \geq S_f\left(\rho_A||\sigma_A\right) \tag{44}
$$

*holds if and only if*

$$
\text{Tr}\left(\sigma_{AB}^{\imath t}\rho_{AB}^{-\imath t+1}\right) = \text{Tr}\left(\sigma_A^{\imath t}\rho_A^{-\imath t+1}\right), \quad \forall\, t \in \mathbb{C}, \tag{45}
$$

*where $\imath = \sqrt{-1}$.*

*Proof* Let

$$
\gamma_{AB} = \sigma_{AB} \otimes \left(\rho_{AB}^{-1}\right)^\mathsf{T}, \tag{46}
$$

$$
\gamma_A = \sigma_A \otimes \left(\rho_A^{-1}\right)^\mathsf{T}. \tag{47}
$$

Since

$$
S_f\left(\rho_{AB}||\sigma_{AB}\right) = \text{vec}\left(\sqrt{\rho_{AB}}\right)^\dagger f(\gamma_{AB})\text{vec}\left(\sqrt{\rho_{AB}}\right), \tag{48}
$$

$$
S_f(\rho_A||\sigma_A) = \text{vec}\left(\sqrt{\rho_A}\right)^\dagger f(\gamma_A)\text{vec}\left(\sqrt{\rho_A}\right), \tag{49}
$$

hence, using Eq. (9), we get

$$
S_f\left(\rho_{AB}||\sigma_{AB}\right) - S_f(\rho_A||\sigma_A) = d\int_{(-a,a)} r(x)p_X(x)dx, \tag{50}
$$

where $d > 0$ since $f(\cdot)$ is non-affine, and

$$
\begin{aligned}
r(x) &= \text{vec}\left(\sqrt{\rho_{AB}}\right)^\dagger \left(\gamma_{AB} - t_0\mathrm{I}\right)^2 \left[a^2\mathrm{I} - (\gamma_{AB} - t_0\mathrm{I})x\right]^{-1} \text{vec}\left(\sqrt{\rho_{AB}}\right) - && (51) \\
&\quad \text{vec}\left(\sqrt{\rho_A}\right)^\dagger \left(\gamma_A - t_0\mathrm{I}\right)^2 \left[a^2\mathrm{I} - (\gamma_A - t_0\mathrm{I})x\right]^{-1} \text{vec}\left(\sqrt{\rho_A}\right). && (52)
\end{aligned}
$$

7

Since $f(\cdot)$ is diffused, let us assume that $p_X(x)$ in Eq. (9) is strictly positive a.e. over $\acute{K}$, a subinterval of $[-a, a]$ with the point $x = 0$ as the interior or the boundary point. Consider $K$, a subinterval of $\acute{K}$, with the point $x = 0$ as the interior or the boundary point, and that does not contain the points at which $a^2 I - (\gamma_{AB} - t_0 I)x$ or $a^2 I - (\gamma_A - t_0 I)x$ become singular. Such a $K$ is always possible since there are finite values of $x$ for which the above two matrices become singular and these matrices are non-singular in the neighborhood of $x = 0$.

Now note that the function $g(t) = (t - t_0)^2 / [a^2 - (t - t_0)x]$ is operator convex. To see this, first note that $1/t$, $t \in (0, \infty)$ is operator convex (see Corollary V.2.6 in Ref. [12] and Ref. [18]). It follows that $h(t) = 1/(a^2 t - x)$ is also operator convex in $t \in (x/a^2, \infty)$. Using Lemma 5 in Sub-section 2.2, it follows that $t\, h(1/t) = t^2/(a^2 - tx)$, $t \in (0, a^2/x)$, is also operator convex, which implies that $g(t)$ is also operator convex in the neighborhood of $x = 0$. Hence, using monotonicity, $r(x) \geq 0$. For the equality to hold,

$$0 = \int_{(-a,a)} r(x) p_X(x) dx \geq \int_K r(x) p_X(x) dx \geq 0. \tag{53}$$

Hence, $r(x) p_X(x) = 0$ a.e. over $K$. Since $p_X(x) > 0$ a.e. over $K$, hence, $r(x) = 0$ a.e. over $K$. But since $r(x)$ is continuous in $K$, hence, $r(x) = 0$ over $K$. Hence, the coefficients of the Taylor series expansion of $r(x)$ around $x = 0$ must be zero. (If $x = 0$ is the boundary point of $K$, then we shall take the Taylor series expansion of $r(x)$ around $x = \epsilon \in K$, where $|\epsilon|$ is arbitrarily small and then in the limit of $\epsilon$ approaching zero, we shall arrive at the same conclusions as below.) Equating the Taylor series coefficients to be zero (and taking appropriate limits if $x = 0$ is the boundary point of $K$), we have

$$\mathrm{vec}\left(\sqrt{\rho_{AB}}\right)^\dagger (\gamma_{AB} - t_0 I)^n \mathrm{vec}\left(\sqrt{\rho_{AB}}\right) = \mathrm{vec}\left(\sqrt{\rho_A}\right)^\dagger (\gamma_A - t_0 I)^n \mathrm{vec}\left(\sqrt{\rho_A}\right), \quad \forall\, n \geq 2. \tag{54}$$

The above equation is trivially true for $n = 0, 1$, and it follows that the above equation is equivalent to

$$\mathrm{vec}\left(\sqrt{\rho_{AB}}\right)^\dagger \gamma_{AB}^n \mathrm{vec}\left(\sqrt{\rho_{AB}}\right) = \mathrm{vec}\left(\sqrt{\rho_A}\right)^\dagger \gamma_A^n \mathrm{vec}\left(\sqrt{\rho_A}\right), \quad \forall\, n \geq 0, \tag{55}$$

$$\mathrm{Tr}\left(\sigma_{AB}^n \rho_{AB}^{-n+1}\right) = \mathrm{Tr}\left(\sigma_A^n \rho_A^{-n+1}\right), \quad \forall\, n \geq 0. \tag{56}$$

Let the spectral decompositions of the matrices in the above equation be given by $\rho_{AB} = \sum_k p_k |k_{AB}^\rho\rangle\langle k_{AB}^\rho|$, $\sigma_{AB} = \sum_k q_k |k_{AB}^\sigma\rangle\langle k_{AB}^\sigma|$, $\rho_A = \sum_k r_k |k_A^\rho\rangle\langle k_A^\rho|$, and $\sigma_A = \sum_k s_k |k_A^\sigma\rangle\langle k_A^\sigma|$. Substituting in the above equation, we get

$$\sum_{k,j} \left(\frac{q_j}{p_k}\right)^n p_k \, |\langle j_{AB}^\sigma| k_{AB}^\rho\rangle|^2 = \sum_{k,j} \left(\frac{s_j}{r_k}\right)^n r_k \, |\langle j_A^\sigma| k_A^\rho\rangle|^2, \quad \forall\, n \geq 0. \tag{57}$$

Now consider the terms in the LHS such that $q_j/p_k = \max_{k,j} q_j/p_k$ and in the RHS such that $s_j/r_k = \max_{k,j} s_j/r_k$. It is clear that for large $n$, these two set of terms dominate all other terms, and since LHS = RHS, $\max_{k,j} q_j/p_k = \max_{k,j} s_j/r_k$, and the sum of their coefficients in the LHS must be the same as their sum in the RHS. Subtracting these two sets from both

sides and arguing similarly for the maximum in the pruned summations and continuing till no term is left, it follows that Eq. (56) amounts to Eq. (45). ■

We note that the conditions for the equality in Eq. (27) for $f(t) = -\ln(t)$ were given by Petz [9] as

$$\sigma_{AB}^{it}\rho_{AB}^{-it} = \sigma_A^{it}\rho_A^{-it} \otimes \mathrm{I}. \tag{58}$$

Since $f(t) = -\ln(t)$ is a non-affine and diffused operator convex function, Eq. (45) should be satified if Eq. (58) is true, which, indeed, is the case. Ruskai [10] gave the following conditions for the equality in Eq. (27) for $f(t) = -\ln(t)$ as

$$\ln(\sigma_{AB}) - \ln(\rho_{AB}) = [\ln(\sigma_A) - \ln(\rho_A)] \otimes \mathrm{I}. \tag{59}$$

Ruskai showed that Eq. (59) can be obtained from Eq. (58) by taking the derivative of both sides of Eq. (58) w.r.t. $t$ at $t = 0$ [10].

**Corollary 1** *For $f(t) = -\ln(t)$, the necessary and sufficient conditions for the equality in Eq. (44) are given by Eq. (45).*

The proof is along the same lines as that of Lemma 3. It is interesting to note that following Ruskai's approach [10], by taking the derivative of both sides of Eq. (45) w.r.t. $t$ at $t = 0$, we obtain $S_{(-\ln)}(\rho_{AB}||\sigma_{AB}) = S_{(-\ln)}(\rho_A||\sigma_A)$.

We now consider the following special case, which is applicable to a variety of cases for both the quantum $f$-relative entropy and the quantum $f$-entropy.

**Corollary 2** *For a non-affine and diffused operator convex function $f(\cdot)$, strictly positive $\rho_{AB}$, $\sigma_B$, and $\rho_A = \mathrm{Tr}_B(\rho_{AB})$, the equality in*

$$S_f\left(\rho_{AB}||\mathrm{I} \otimes \sigma_B\right) \geq S_f\left[\rho_A||\mathrm{Tr}(\sigma_B)\mathrm{I}\right] \tag{60}$$

*holds if and only if $\rho_{AB} = \rho_A \otimes \sigma_B/\mathrm{Tr}(\sigma_B)$.*

*Proof* The inequality in Eq. (60) is true, of course, because of monotonicity. Using Eq. (56), the equality conditions are

$$\mathrm{Tr}\left[(\mathrm{I} \otimes \phi_B^n)\,\rho_{AB}^{-n+1}\right] = \mathrm{Tr}\left(\rho_A^{-n+1}\right), \quad \forall\, n \geq 0, \tag{61}$$

where $\phi_B = \sigma_B/\mathrm{Tr}(\sigma_B)$. Let the spectral decompositions of the matrices in the above equation be given by $\rho_A = \sum_{j=1}^{d_\rho} \lambda_j|j_A\rangle\langle j_A|$, $\phi_B = \sum_{i=1}^{d_\sigma} \beta_i|i_B\rangle\langle i_B|$, $\rho_{AB} = \sum_{k=1}^{d_\rho d_\sigma} \alpha_k|k_{AB}\rangle\langle k_{AB}|$, where $d_\rho$, $d_\sigma$ are the dimensions of the Hilbert spaces describing $\rho_A$, $\sigma_B$ respectively. Then Eq. (61) can be restated as

$$\sum_{j=1}^{d_\rho}\sum_{i=1}^{d_\sigma}\sum_{k=1}^{d_\rho d_\sigma} \left(\frac{\beta_i}{\alpha_k}\right)^n \alpha_k\,|\langle j_A|\,\langle i_B|\,k_{AB}\rangle|^2 = \sum_{j=1}^{d_\rho} \lambda_j^{-n}\lambda_j, \quad \forall\, n \geq 0, \tag{62}$$

$$\sum_{j=1}^{d_\rho}\sum_{i=1}^{d_\sigma}\sum_{k=1}^{d_\rho d_\sigma} \left[\left(\frac{\beta_i}{\alpha_k}\right)^n \alpha_k - \left(\frac{1}{\lambda_j}\right)^n \beta_i\lambda_j\right] |\langle j_A|\,\langle i_B|\,k_{AB}\rangle|^2 = 0, \quad \forall\, n \geq 0. \tag{63}$$

9

Let there be $M$ distinct eigenvalues of $\rho_A$ denoted by $\lambda^{(m)}$, $m = 1, ..., M$. We follow similar reasoning as in Lemma 3 to claim that we have $M$ disjoint sets $\mathcal{Q}_m$, $m = 1, ..., M$, such that $\forall\, (j, i, k) \in \mathcal{Q}_m$,

$$\frac{\alpha_k}{\beta_i} = \lambda_j = \lambda^{(m)}. \tag{64}$$

For completeness, we shall also define $\mathcal{Q}_0 = \bigcap_{m=1}^{M} \mathcal{Q}_m^c$, where $\mathcal{Q}^c$ denotes the complement of $\mathcal{Q}$. It follows that for $(j, i, k) \in \mathcal{Q}_0$, $|\langle j_A| \langle i_B| k_{AB}\rangle| = 0$. We now have

$$\mathrm{Tr}\left[\rho_{AB}\left(\rho_A \otimes \phi_B\right)\right] = \sum_{j=1}^{d_\rho}\sum_{i=1}^{d_\sigma}\sum_{k=1}^{d_\rho d_\sigma} \alpha_k \lambda_j \beta_i \left|\langle j_A| \langle i_B| k_{AB}\rangle\right|^2 \tag{65}$$

$$= \sum_{m=1}^{M} \sum_{(j,i,k)\in\mathcal{Q}_m} \alpha_k \lambda_j \beta_i \left|\langle j_A| \langle i_B| k_{AB}\rangle\right|^2 \tag{66}$$

$$= \sum_{m=1}^{M} \sum_{(j,i,k)\in\mathcal{Q}_m} \alpha_k^2 \left|\langle j_A| \langle i_B| k_{AB}\rangle\right|^2 \tag{67}$$

$$= \sum_{j=1}^{d_\rho}\sum_{i=1}^{d_\sigma}\sum_{k=1}^{d_\rho d_\sigma} \alpha_k^2 \left|\langle j_A| \langle i_B| k_{AB}\rangle\right|^2 \tag{68}$$

$$= \mathrm{Tr}\left(\rho_{AB}^2\right). \tag{69}$$

Similarly, one can show that

$$\mathrm{Tr}\left(\rho_A^2 \otimes \phi_B^2\right) = \mathrm{Tr}\left(\rho_{AB}^2\right). \tag{70}$$

Using the above two equations, it now follows that

$$\|\rho_{AB} - \rho_A \otimes \phi_B\|_\mathrm{F}^2 = \mathrm{Tr}\left(\rho_{AB}^2\right) + \mathrm{Tr}\left(\rho_A^2 \otimes \phi_B^2\right) - 2\mathrm{Tr}\left[\rho_{AB}\left(\rho_A \otimes \phi_B\right)\right] \tag{71}$$

$$= 0, \tag{72}$$

or $\rho_{AB} = \rho_A \otimes \sigma_B/\mathrm{Tr}(\sigma_B)$. ∎

**Corollary 3** *The inequality*

$$S_f\left(\rho_{AB}||\mathrm{I} \otimes \rho_B\right) \leq S_f\left(\rho_{ABC}||\mathrm{I} \otimes \rho_{BC}\right) \tag{73}$$

*holds and the equality conditions are given by*

$$\mathrm{Tr}\left[\left(\mathrm{I} \otimes \rho_B^{it}\right)\rho_{AB}^{-it+1}\right] = \mathrm{Tr}\left[\left(\mathrm{I} \otimes \rho_{BC}^{it}\right)\rho_{ABC}^{-it+1}\right], \quad \forall\, t \in \mathbb{C}. \tag{74}$$

Note that the inequality follows immediately from the monotonicity and the equality conditions from Eq. (45). Ruskai [10] showed that the above inequality for $f(t) = -\ln(t)$ is just a restatement of the strong sub-additivity in Eq. (2).

It follows that for strictly positive $\rho$ and $\sigma$, and a non-affine and diffused operator convex function $f(\cdot)$, that

$$S_f(\rho||\sigma) \geq f\left[\frac{\text{Tr}(\sigma)}{\text{Tr}(\rho)}\right]\text{Tr}(\rho) \tag{75}$$

with equality if and only if $\rho/\text{Tr}(\rho) = \sigma/\text{Tr}(\sigma)$. To see this, substitute $\rho_{AB} = \rho$, $\rho_A = \text{Tr}(\rho)$, and $\sigma_B = \sigma$ in Eq. (60).

However, Petz showed the same result for a non-affine operator convex function with no requirement that the function has to be diffused [17]. We provide an alternate derivation of his result.

**Lemma 4** *(Petz [17]) For strictly positive $\rho$ and $\sigma$, and a non-affine operator convex function $f(\cdot)$, the following holds*

$$S_f(\rho||\sigma) \geq f\left[\frac{\text{Tr}(\sigma)}{\text{Tr}(\rho)}\right]\text{Tr}(\rho) \tag{76}$$

*with equality if and only if $\rho/\text{Tr}(\rho) = \sigma/\text{Tr}(\sigma)$.*

*Proof* We first note that since $f(\cdot)$ is a non-affine operator convex function, hence, it is strictly convex. Secondly, note that for a strictly convex function $f(\cdot)$, $a_k > 0$, $k = 1, ..., d$,

$$\sum_{k=1}^{d} a_k f\left(\frac{b_k}{a_k}\right) \geq \left(\sum_{k=1}^{d} a_k\right) f\left(\frac{\sum_{k=1}^{d} b_k}{\sum_{k=1}^{d} a_k}\right), \tag{77}$$

with equality if and only if $b_k/a_k$ is constant $\forall\, k = 1, ..., d$. Let the spectral decompositions of $\rho$ and $\sigma$ be given by Eqs. (18) and (19) respectively. Using Eqs. (21) and (77), we get

$$\begin{aligned}
S_f(\rho||\sigma) &= \sum_{i,j}^{d} p_i|\langle i_\sigma|\, j_\rho\rangle|^2 f\left(\frac{q_j}{p_i}\right) \tag{78}\\
&\geq \sum_{j=1}^{d}\left(\sum_{i=1}^{d} p_i|\langle i_\sigma|\, j_\rho\rangle|^2\right) f\left(\frac{\sum_{i=1}^{d} q_j|\langle i_\sigma|\, j_\rho\rangle|^2}{\sum_{i=1}^{d} p_i|\langle i_\sigma|\, j_\rho\rangle|^2}\right) \tag{79}\\
&\geq \left(\sum_{i,j=1}^{d} p_i|\langle i_\sigma|\, j_\rho\rangle|^2\right) f\left(\frac{\sum_{i,j=1}^{d} q_j|\langle i_\sigma|\, j_\rho\rangle|^2}{\sum_{i,j=1}^{d} p_i|\langle i_\sigma|\, j_\rho\rangle|^2}\right) \tag{80}\\
&= f\left[\frac{\text{Tr}(\sigma)}{\text{Tr}(\rho)}\right]\text{Tr}(\rho). \tag{81}
\end{aligned}$$

The conditions for the equality are $q_j|\langle i_\sigma|\, j_\rho\rangle|^2 = c_1 p_i|\langle i_\sigma|\, j_\rho\rangle|^2$, $\forall\, i = 1, ..., d$, and $q_j = c_2 \sum_{i=1}^{d} p_i|\langle i_\sigma|\, j_\rho\rangle|^2$, $\forall\, j = 1, ..., d$, where $c_1$, $c_2$ are positive constants. It now follows that $c_1 = c_2 = \text{Tr}(\sigma)/\text{Tr}(\rho)$, $\text{Tr}(\rho\sigma) = \text{Tr}(\sigma^2)/c_1$, and $\text{Tr}(\rho^2) = \text{Tr}(\sigma^2)/c_1^2$. Let $||\kappa||_\text{F} = \sqrt{\text{Tr}\left(\kappa^\dagger \kappa\right)}$ denote the Frobenius norm of $\kappa$. Then

$$\left\|\frac{\rho}{\text{Tr}(\rho)} - \frac{\sigma}{\text{Tr}(\sigma)}\right\|_\text{F}^2 = \frac{\text{Tr}(\rho^2)}{[\text{Tr}(\rho)]^2} + \frac{\text{Tr}(\sigma^2)}{[\text{Tr}(\sigma)]^2} - 2\frac{\text{Tr}(\rho\sigma)}{\text{Tr}(\rho)\text{Tr}(\sigma)} = 0. \tag{82}$$

11

QED. ∎

The Klein's inequality is a special case of the above Lemma for $f(t) = -\ln(t)$.

## 2.2 Convexity

We now examine the convexity properties of the quantum $f$-relative entropy .

**Lemma 5** *Let*

$$g(x) \stackrel{\Delta}{=} \sqrt{x} f \left( \frac{1}{x} \right) \sqrt{x} = x f \left( \frac{1}{x} \right), \quad x \in (0, \infty). \tag{83}$$

*Then if $f(\cdot)$ is operator convex, so is $g(\cdot)$.*

*Proof* The proof is not much different from that of Theorem 2.2 in Ref. [19] though without using the linear super-operators. Note first from Lemma 5.1.5 in Ref. [12] that if $A \leq B$, then for any matrix $X$ with appropriate dimensions, $X^\dagger A X \leq X^\dagger B X$.

For $0 \leq \lambda \leq 1$, choose $n = 2$, strictly positive $A$, $B$, $C = \lambda A + (1 - \lambda)B$, $E_1 = \sqrt{\lambda A C^{-1}}$, $E_2 = \sqrt{(1 - \lambda) B C^{-1}}$, $\phi_1 = A^{-1}$, and $\phi_2 = B^{-1}$, and substitute in Eq. (13) to get

$$f \left( C^{-1} \right) \leq \sqrt{C^{-1}} \left[ \lambda \sqrt{A} f \left( A^{-1} \right) \sqrt{A} + (1 - \lambda) \sqrt{B} f \left( B^{-1} \right) \sqrt{B} \right] \sqrt{C^{-1}}. \tag{84}$$

The results follows by pre-multiplying and post-multiplying both sides by $\sqrt{C}$ and noting that

$$g(X) = \sqrt{X} f \left( X^{-1} \right) \sqrt{X}. \tag{85}$$

QED. ∎

The operator convexity of $g(t) = t \ln(t)$ follows from that $f(t) = -\ln(t)$ by using the above result. It is easy to see that from Eq. (17) that the quantum $f$-relative entropy is convex in the second argument since

$$
\begin{aligned}
S_f \left[ \rho || \lambda \sigma_1 + (1 - \lambda)\sigma_2 \right] &= \text{vec} \left( \sqrt{\rho} \right)^\dagger f \left\{ [\lambda \sigma_1 + (1 - \lambda)\sigma_2] \otimes \left( \rho^{-1} \right)^{\mathsf{T}} \right\} \text{vec} \left( \sqrt{\rho} \right) \quad (86) \\
&\leq \lambda \text{vec} \left( \sqrt{\rho} \right)^\dagger f \left[ \sigma_1 \otimes (\rho^{-1})^{\mathsf{T}} \right] \text{vec} \left( \sqrt{\rho} \right) + \\
&\quad (1 - \lambda) \text{vec} \left( \sqrt{\rho} \right)^\dagger f \left[ \sigma_2 \otimes \left( \rho^{-1} \right)^{\mathsf{T}} \right] \text{vec} \left( \sqrt{\rho} \right) \quad (87) \\
&= \lambda S_f(\rho || \sigma_1) + (1 - \lambda) S_f(\rho || \sigma_2). \quad (88)
\end{aligned}
$$

It is easy to check from Eq. (21) that

$$S_f(\rho || \sigma) = S_g(\sigma || \rho), \tag{89}$$

where $g(\cdot)$ is as defined in Eq. (83). It now follows from Eq. (88) that the quantum $f$-relative entropy is convex in its first argument as well since

$$
\begin{align}
S_f\left[\lambda\rho_1 + (1-\lambda)\rho_2||\sigma\right] &= S_g\left[\sigma||\lambda\rho_1 + (1-\lambda)\rho_2\right] \tag{90}\\
&\leq \lambda S_g(\sigma||\rho_1) + (1-\lambda)S_g(\sigma||\rho_2) \tag{91}\\
&= \lambda S_f(\rho_1||\sigma) + (1-\lambda)S_f(\rho_2||\sigma). \tag{92}
\end{align}
$$

We now show that the quantum $f$-relative entropy is jointly convex in its arguments which is a stronger result than the convexity of any one of its arguments. Petz proved the joint convexity of quantum quasi relative entropy [17]. We provide an alternate proof that is more accessible. Furthermore, we give the equality conditions for a class of operator convex functions.

**Lemma 6** *For $0 < \lambda < 1$, strictly positive $\rho_1$, $\rho_2$, $\sigma_1$, $\sigma_2$, and $f(\cdot)$ operator convex,*

$$
S_f\left(\rho_\lambda||\sigma_\lambda\right) \leq \lambda S_f(\rho_1||\sigma_1) + (1-\lambda)S_f(\rho_2||\sigma_2), \tag{93}
$$

*where $\rho_\lambda = \lambda\rho_1 + (1-\lambda)\rho_2$ and $\sigma_\lambda = \lambda\sigma_1 + (1-\lambda)\sigma_2$. The equality holds for a non-affine and diffused operator convex function $f(\cdot)$ if and only if*

$$
\mathrm{Tr}\left(\sigma_\lambda^{\imath t}\rho_\lambda^{-\imath t+1}\right) = \lambda\mathrm{Tr}\left(\sigma_1^{\imath t}\rho_1^{-\imath t+1}\right) + (1-\lambda)\mathrm{Tr}\left(\sigma_2^{\imath t}\rho_2^{-\imath t+1}\right), \quad \forall\, t \in \mathbb{C}. \tag{94}
$$

*Proof* Choose

$$
\begin{align}
E_1 &= \sqrt{\lambda}\left[\mathrm{I} \otimes (\sqrt{\rho_1})^\mathsf{T}\right]\left[\mathrm{I} \otimes \left(\rho_\lambda^{-1/2}\right)^\mathsf{T}\right], \tag{95}\\
E_2 &= \sqrt{1-\lambda}\left[\mathrm{I} \otimes [\sqrt{\rho_2})^\mathsf{T}\right]\left[\mathrm{I} \otimes \left(\rho_\lambda^{-1/2}\right)^\mathsf{T}\right], \tag{96}\\
\phi_1 &= \sigma_1 \otimes \left(\rho_1^{-1}\right)^\mathsf{T}, \tag{97}\\
\phi_2 &= \sigma_2 \otimes \left(\rho_2^{-1}\right)^\mathsf{T}. \tag{98}
\end{align}
$$

It is easy to check that $E_1^\dagger E_1 + E_2^\dagger E_2 = \mathrm{I}$ and $E_1^\dagger \phi_1 E_1 + E_2^\dagger \phi_2 E_2 = \sigma_\lambda \otimes \left(\rho_\lambda^{-1}\right)^\mathsf{T}$. Using Eq. (13), and pre-multiplying both sides by $X^\dagger = \left[\mathrm{I} \otimes (\sqrt{\rho_\lambda})^*\right]$ and post-multiplying both sides by $X = \left[\mathrm{I} \otimes (\sqrt{\rho_\lambda})^\mathsf{T}\right]$, we get

$$
\begin{align}
X^\dagger f\left[\sigma_\lambda \otimes \left(\rho_\lambda^{-1}\right)^\mathsf{T}\right] X &\leq \lambda\left[\mathrm{I} \otimes (\sqrt{\rho_1})^*\right] f\left[\sigma_1 \otimes \left(\rho_1^{-1}\right)^\mathsf{T}\right]\left[\mathrm{I} \otimes (\sqrt{\rho_1})^\mathsf{T}\right]\\
&\quad + (1-\lambda)\left[\mathrm{I} \otimes (\sqrt{\rho_2})^*\right] f\left[\sigma_2 \otimes \left(\rho_2^{-1}\right)^\mathsf{T}\right]\left[\mathrm{I} \otimes (\sqrt{\rho_2})^\mathsf{T}\right]. \tag{99}
\end{align}
$$

Pre-multiplying both sides by $\mathrm{vec}\,(\mathrm{I})^\dagger$, post-multiplying by $\mathrm{vec}\,(\mathrm{I})$, and using Eq. (21), we get

$$
S_f(\rho_\lambda||\sigma_\lambda) \leq \lambda S(\rho_1||\lambda_1) + (1-\lambda)S(\rho_2||\lambda_2). \tag{100}
$$

13

To prove the equality conditions, we follow the analysis in Lemma 3 to reduce the equality conditions to

$$\text{Tr}\left(\sigma_\lambda^n \rho_\lambda^{-n+1}\right) = \lambda \text{Tr}\left(\sigma_1^n \rho_1^{-n+1}\right) + (1-\lambda)\text{Tr}\left(\sigma_2^n \rho_2^{-n+1}\right), \quad \forall\, n \geq 0, \tag{101}$$

which, using the reasoning in Lemma 3, can be shown to be equivalent to Eq. (94). ∎

For $f(t) = -\ln(t)$, Ruskai [10] gave the equality conditions for Eq. (93) as

$$\ln(\sigma_\lambda) - \ln(\rho_\lambda) = \ln(\sigma_i) - \ln(\rho_i), \quad i = 1, 2. \tag{102}$$

**Corollary 4** *The quantum $f$-relative entropy is sub-additive, i.e., for strictly positive $\rho_i$, $\sigma_i$, $i = 1, 2$,*

$$S_f\left(\rho_1 + \rho_2 || \sigma_1 + \sigma_2\right) \leq S_f\left(\rho_1 || \sigma_1\right) + S_f\left(\rho_2 || \sigma_2\right) \tag{103}$$

*and the equality holds if and only if*

$$\text{Tr}\left[(\sigma_1 + \sigma_2)^{\imath t}(\rho_1 + \rho_2)^{-\imath t + 1}\right] = \text{Tr}\left(\sigma_1^{\imath t}\rho_1^{-\imath t + 1}\right) + \text{Tr}\left(\sigma_2^{\imath t}\rho_2^{-\imath t + 1}\right), \quad \forall\, t \in \mathbb{C}. \tag{104}$$

*Proof* Joint convexity implies the sub-additivity of the quantum $f$-relative entropy since

$$S_f\left(\rho_1 + \rho_2 || \sigma_1 + \sigma_2\right) = S_f\left(\frac{2\rho_1}{2} + \frac{2\rho_2}{2} \middle|\middle| \frac{2\sigma_1}{2} + \frac{2\sigma_2}{2}\right) \tag{105}$$

$$\leq \frac{1}{2}S_f\left(2\rho_1 || 2\sigma_1\right) + \frac{1}{2}S_f\left(2\rho_2 || 2\sigma_2\right) \tag{106}$$

$$= S_f\left(\rho_1 || \sigma_1\right) + S_f\left(\rho_2 || \sigma_2\right). \tag{107}$$

The equality conditions follow from Eq. (94). ∎

**Lemma 7** *Joint convexity of the quantum $f$-relative entropy implies monotonicity, and for a completely positive trace-preserving (CPTP) quantum operation $\mathcal{E}(\cdot)$,*

$$S_f\left[\mathcal{E}(\rho) || \mathcal{E}(\sigma)\right] \leq S_f(\rho || \sigma), \tag{108}$$

*and the equality holds if and only if*

$$\text{Tr}\left\{[\mathcal{E}(\sigma)]^{\imath t}[\mathcal{E}(\rho)]^{-\imath t + 1}\right\} = \text{Tr}\left(\sigma^{\imath t}\rho^{-\imath t + 1}\right), \quad \forall\, t \in \mathbb{C}. \tag{109}$$

We omit the proof.

# 3 Quantum $f$-entropy

We now define quantum $f$-entropy for strictly positive $\rho$, denoted by $S_f(\rho)$, in terms of the quantum $f$-relative entropy as

$$S_f(\rho) \triangleq -S_f(\rho||\mathrm{I}) = -\mathrm{Tr}\left[\rho f\left(\rho^{-1}\right)\right] = -\sum_{i=1}^d p_i f\left(\frac{1}{p_i}\right), \tag{110}$$

where the spectral decomposition of $\rho$ is given by

$$\rho = \sum_{i=1}^d p_i |i_\rho\rangle\langle i_\rho|, \tag{111}$$

and $d$ is the dimension of the Hilbert space that describes $\rho$. For a density matrix $\rho$, and $f(t) = -\ln(t)$, quantum $f$-entropy coincides with the von-Neumann entropy [1, 2].

**Lemma 8** *The following holds:*

(i) *For strictly positive $\rho$ with dimension $d$,*

$$S_f(\rho) \leq -\mathrm{Tr}(\rho) f\left[\frac{d}{\mathrm{Tr}(\rho)}\right]. \tag{112}$$

*For a non-affine operator convex function $f(\cdot)$, the equality holds if and only if $\rho = \mathrm{Tr}(\rho)\mathrm{I}/d$.*

(ii) *Let the joint state in system $AB$ be a pure state. Then $S(A) = S(B)$.*

(iii) *Projective measurements increase quantum $f$-entropy, and for a non-affine and diffused operator convex function $f(\cdot)$, the equality holds if and only if the projective measurement leaves the state unchanged.*

*Proof*

(i) follows by using Eq. (76).

(ii) Let the joint state in system $AB$ be a pure state denoted by $|\phi_{AB}\rangle$, and let its Schmidt decomposition be given by

$$|\phi_{AB}\rangle = \sum_k \sqrt{\lambda_k}|k_A\rangle|k_B\rangle \tag{113}$$

where $\{\lambda_k\}$ is a probability vector, $\{|k_A\rangle\}$ and $\{|k_B\rangle\}$ are orthonormal states in $A$ and $B$ respectively. Then since

$$\rho_A = \mathrm{Tr}_B\left(|\phi_{AB}\rangle\langle\phi_{AB}|\right) \;=\; \sum_k \lambda_k|k_A\rangle\langle k_A|, \tag{114}$$

$$\rho_B = \mathrm{Tr}_A\left(|\phi_{AB}\rangle\langle\phi_{AB}|\right) \;=\; \sum_k \lambda_k|k_B\rangle\langle k_B|, \tag{115}$$

it follows that

$$S_f\left(\rho_A\right) = S_f\left(\rho_B\right). \tag{116}$$

(iii) Let $\{P_i\}$ be a complete set of projectors and $\sum_i P_i = \mathrm{I}$. Projective measurements increase quantum $f$-entropy since it follows using Eq. (108) that

$$S_f(\rho) \;=\; -S_f(\rho||\mathrm{I}) \tag{117}$$

$$\geq \;-S_f\left(\sum_i P_i\rho P_i\Big|\Big|\mathrm{I}\right) \tag{118}$$

$$= \;S_f\left(\sum_i P_i\rho P_i\right). \tag{119}$$

To prove the equality condition, we use Eq. (108) to get

$$\mathrm{Tr}\left(\rho^{-n+1}\right) = \mathrm{Tr}\left[\left(\sum_i P_i\rho P_i\right)^{-n+1}\right], \quad \forall\, n \geq 0, \tag{120}$$

or the eigenvalues of $\rho$ and $\sum_i P_i\rho P_i$ are the same including multiplicities. In particular,

$$\mathrm{Tr}(\rho^2) = \mathrm{Tr}\left[\left(\sum_i P_i\rho P_i\right)^2\right]. \tag{121}$$

Then it is easy to show that

$$\left\|\rho - \sum_i P_i\rho P_i\right\|_{\mathrm{F}}^2 = \mathrm{Tr}(\rho^2) - \mathrm{Tr}\left[\left(\sum_i P_i\rho P_i\right)^2\right] = 0, \tag{122}$$

which proves the result. ■

**Lemma 9** *Let $\{\rho_i\}$ be a set of strictly positive matrices with unit trace and described by the same Hilbert space, and let their spectral decompositions be given by*

$$\rho_i = \sum_j q_{ij}|i,j\rangle\langle i,j| \tag{123}$$

*Then for any probability vector $\{p_i\}$,*

$$\sum_i p_i S_f(\rho_i) \leq S_f\left(\sum_i p_i\rho_i\right) \leq -\sum_{ij} p_i q_{ij} f\left(\frac{1}{p_i q_{ij}}\right). \tag{124}$$

*For a non-affine and diffused operator convex $f(\cdot)$, the equality in the first inequality holds if and only if the $\rho_i$'s with $p_i > 0$ are identical, and the equality in the second inequality holds if and only $\rho_i$'s have support on orthogonal subspaces.*

16

*Proof* Consider a joint state in system $AB$ as

$$\rho_{AB} = \sum_i p_i \rho_i \otimes |i_B\rangle\langle i_B|, \tag{125}$$

where $\rho_i$ lie in the system $A$ (our system of interest), and $\{|i_B\rangle\}$ is any orthonormal basis in system $B$ (ancilla). Then for

$$\rho_B = \mathrm{Tr}_A(\rho_{AB}) = \sum_i p_i |i_B\rangle\langle i_B|, \tag{126}$$

we have

$$S_f\left(\rho_{AB}||\mathrm{I}\otimes\rho_B\right) = S_f\left(\sum_{i,j} p_i q_{ij}|i,j\rangle\langle i,j| \otimes |i_B\rangle\langle i_B| \,\Big|\Big|\, \sum_{i,j} p_i |i,j\rangle\langle i,j| \otimes |i_B\rangle\langle i_B|\right) \tag{127}$$

$$= \sum_{i,j} f\left(\frac{1}{q_{ij}}\right) p_i q_{ij} \tag{128}$$

$$= -\sum_i p_i S_f(\rho_i). \tag{129}$$

Note that

$$\rho_A = \mathrm{Tr}_B\left(\rho_{AB}\right) = \sum_i p_i \rho_i. \tag{130}$$

It now follows that

$$S_f\left(\sum_i p_i \rho_i\right) = -S_f\left(\sum_i p_i \rho_i \Big|\Big| \mathrm{I}\right) \tag{131}$$

$$= -S_f\left(\rho_A||\mathrm{I}\right) \tag{132}$$

$$\geq -S_f\left(\rho_{AB}||\mathrm{I}\otimes\rho_B\right) \tag{133}$$

$$= \sum_i p_i S_f(\rho_i). \tag{134}$$

Using Eq. (60), the equality holds if and only if $\rho_{AB} = \rho_A \otimes \rho_B$, or

$$\sum_i p_i \rho_i \otimes |i_B\rangle\langle i_B| = \sum_i p_i \rho_i \otimes \sum_k p_k |k_B\rangle\langle k_B|, \tag{135}$$

or the $\rho_i$ with $p_i > 0$ are identical.

To prove the upper bound, let us first assume that $\rho_i$'s are all pure and $\rho_i = |\psi_i\rangle\langle\psi_i|$. We attach an ancilla $B$ to our system $A$ such that

$$|\phi_{AB}\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle|i_B\rangle, \tag{136}$$

17

where $\{|i_B\rangle\}$ is an orthonormal basis in $B$. It easily follows that

$$\rho_A \;=\; \mathrm{Tr}_B\left(|\phi_{AB}\rangle\langle\phi_{AB}|\right) = \sum_i p_i \rho_i, \tag{137}$$

$$\rho_B \;=\; \mathrm{Tr}_A\left(|\phi_{AB}\rangle\langle\phi_{AB}|\right) = \sum_i \sqrt{p_i p_j}\,\langle\psi_j|\,\psi_i\rangle\,|i_B\rangle\langle j_B|. \tag{138}$$

Define projectors $P_i = |i_B\rangle\langle i_B|$ and

$$\rho_{\acute{B}} = \sum_i P_i \rho_B P_i = \sum_i p_i |i_B\rangle\langle i_B|. \tag{139}$$

Hence using Lemma 8, we have

$$-\sum_i p_i f\left(\frac{1}{p_i}\right) = S_f\left(\rho_{\acute{B}}\right) \;\geq\; S_f\left(\rho_B\right) = S_f\left(\rho_A\right) \tag{140}$$

$$= \; S_f\left(\sum_i p_i \rho_i\right) = S_f\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right). \tag{141}$$

Using Lemma 8, the equality holds if and only if $\rho_B = \rho_{\acute{B}}$ or $\langle\psi_j|\,\psi_i\rangle = \delta_{i,j}$. This proves the Lemma when $\rho_i$'s are pure. For mixed $\rho_i$'s, we use the above result to have

$$S_f\left(\sum_{i,j} p_i q_{i,j} |i,j\rangle\langle i,j|\right) \leq -\sum_{i,j} p_i q_{i,j} f\left(\frac{1}{p_i q_{ij}}\right), \tag{142}$$

with equality if and only if $\rho_i$ have support on orthogonal subspaces. ∎

# 4    Generalized Data Processing Inequalities

In this section, we show that the $f$-generalizations of well-known quantum information theoretic quantities also satisfy the data processing inequalities [2, 20] as they do for $f(t) = -\ln(t)$.

## 4.1    Holevo information

Consider a state in the composite system consisting of $A$ and $B$ given by

$$\rho_{AB} = \sum_i p_i \rho_i \otimes |i_B\rangle\langle i_B|, \tag{143}$$

where $\{|i_B\rangle\}$ is an orthonormal basis in $B$ and $\rho_i$'s are strictly positive with unit trace.. Let $\mathcal{E}(\cdot)$ denote the CPTP quantum operation acting on $A$. Then $f$-Holevo $\chi_f(\mathcal{E})$ quantity is defined as

$$\chi_f(\mathcal{E}) = \max_{\{p_i,\rho_i\}} S_f\left[(\mathcal{E}\otimes\mathcal{I}_B)\,\rho_{AB}||\mathcal{E}(\rho_A)\otimes\rho_B\right], \tag{144}$$

where $\rho_{AB}$ is given by Eq. (143), $\rho_A = \text{Tr}_B(\rho_{AB})$, and $\rho_B = \text{Tr}_A(\rho_{AB})$. For $f(\cdot) = -\ln(\cdot)$, this quantity is the product state capacity for the quantum channel $\mathcal{E}(\cdot)$ for transmitting classical information as proved by Holevo, Schumacher, and Westmoreland (HSW theorem) [21, 22].

We note here that $\chi_f(\mathcal{E})$ is independent of the choice of $\{|i_B\rangle\}$. To see this, let $\{U|i_B\rangle\}$ be the new orthonormal basis chosen for the system $B$ associated with $\acute{\chi}_f(\mathcal{E})$. Then for $V = \mathrm{I} \otimes U$, we have

$$\acute{\chi}_f(\mathcal{E}) = \max_{\{p_i,\rho_i\}} S_f \left\{ V \left[ \sum_i p_i \mathcal{E}(\rho_i) \otimes |i_B\rangle\langle i_B| \right] V^\dagger \middle\| \right. \tag{145}$$

$$\left. V \left[ \sum_i p_i \mathcal{E}(\rho_i) \otimes \sum_j p_j |j_B\rangle\langle j_B| \right] V^\dagger \right\} \tag{146}$$

$$= \chi_f(\mathcal{E}). \tag{147}$$

Let $\mathcal{E}_1(\cdot)$ and $\mathcal{E}_2(\cdot)$ be two CPTP quantum operations, then using Eq. (108), we have

$$\chi_f(\mathcal{E}_1) = \max_{\{p_i,\rho_i\}} S_f \left[ (\mathcal{E}_1 \otimes \mathcal{I}_B) \rho_{AB} \| (\mathcal{E}_1 \otimes \mathcal{I}_B)(\rho_A \otimes \rho_B) \right] \tag{148}$$

$$\geq \max_{\{p_i,\rho_i\}} S_f \left[ (\mathcal{E}_2 \otimes \mathcal{I}_B)(\mathcal{E}_1 \otimes \mathcal{I}_B) \rho_{AB} \| (\mathcal{E}_2 \otimes \mathcal{I}_B)(\mathcal{E}_1 \otimes \mathcal{I}_B)(\rho_A \otimes \rho_B) \right] \tag{149}$$

$$= \chi_f(\mathcal{E}_2 \circ \mathcal{E}_1). \tag{150}$$

$$\chi_f(\mathcal{E}_2 \circ \mathcal{E}_1) = \max_{\{p_i,\rho_i\}} S_f \left\{ (\mathcal{E}_2 \otimes \mathcal{I}_B) \left[ \sum_i p_i \mathcal{E}_1(\rho_i) \otimes |i_B\rangle\langle i_B| \right] \middle\| \right.$$

$$\left. (\mathcal{E}_2 \otimes \mathcal{I}_B) \left[ \sum_i p_i \mathcal{E}_1(\rho_i) \otimes \rho_B \right] \right\} \tag{151}$$

$$= \max_{\{p_i,\sigma_i:\sigma_i=\mathcal{E}_1(\rho_i)\}} S_f \left\{ (\mathcal{E}_2 \otimes \mathcal{I}_B) \left[ \sum_i p_i \sigma_i \otimes |i_B\rangle\langle i_B| \right] \middle\| \right.$$

$$\left. (\mathcal{E}_2 \otimes \mathcal{I}_B) \left[ \sum_i p_i \sigma_i \otimes \rho_B \right] \right\} \tag{152}$$

$$\leq \chi_f(\mathcal{E}_2). \tag{153}$$

Hence,

$$\chi_f(\mathcal{E}_2 \circ \mathcal{E}_1) \leq \min\{\chi_f(\mathcal{E}_1), \chi_f(\mathcal{E}_2)\}, \tag{154}$$

which is the data processing inequality.

## 4.2   Entanglement-assisted capacity

Bennett *et al* gave an expression for the capacity known as the entanglement-assisted classical capacity if the sender and receiver have a shared quantum entanglement [23].

Let $Q$ be the system of interest and the purification of a state in $Q$ is given in the joint system $RQ$. Then the $f$-generalization of the entanglement-assisted channel capacity is defined as

$$C_{E,f}(\mathcal{E}) = \max_{\rho_Q} S_f \left[ (\mathcal{E} \otimes \mathcal{I}_R) (|\psi_{QR}\rangle\langle\psi_{QR}|) \,\middle|\middle|\, (\mathcal{E} \otimes \mathcal{I}_R) (\rho_Q \otimes \rho_R) \right], \tag{155}$$

where $|\psi_{QR}\rangle$ is a purification of the density matrix $\rho_Q$ and $\rho_R = \mathrm{Tr}_Q (|\psi_{QR}\rangle\langle\psi_{QR}|)$.

$$
\begin{aligned}
C_{E,f}(\mathcal{E}_1) &= \max_{\rho_Q} S_f \left[ (\mathcal{E}_1 \otimes \mathcal{I}_R) (|\psi_{QR}\rangle\langle\psi_{QR}|) \,\middle|\middle|\, (\mathcal{E}_1 \otimes \mathcal{I}_R) (\rho_Q \otimes \rho_R) \right] & (156) \\
&\geq \max_{\rho_Q} S_f \left[ (\mathcal{E}_2 \otimes \mathcal{I}_R) (\mathcal{E}_1 \otimes \mathcal{I}_R) (|\psi_{QR}\rangle\langle\psi_{QR}|) \,\middle|\middle| \right. & (157) \\
&\qquad\qquad \left. (\mathcal{E}_2 \otimes \mathcal{I}_R) (\mathcal{E}_1 \otimes \mathcal{I}_R) (\rho_Q \otimes \rho_R) \right] & (158) \\
&= C_{E,f}(\mathcal{E}_2 \circ \mathcal{E}_1). & (159)
\end{aligned}
$$

Let us introduce an ancilla $E_1$ with a Unitary operation $V_1$ over the composite system $QRE_1$ to mock up the quantum operation $\mathcal{E}_1 \otimes \mathcal{I}_R$, i.e.,

$$(\mathcal{E}_1 \otimes \mathcal{I}_R) (\rho_{QR}) = \mathrm{Tr}_{E_1} \left[ V_1 (\rho_{QR} \otimes |0_{E_1}\rangle\langle 0_{E_1}|) V_1^\dagger \right], \tag{160}$$

where $|0_{E_1}\rangle$ is the initial state of the ancilla. Let $|\psi_{QRE_1}\rangle = |\psi_{QR}\rangle|0_{E_1}\rangle$ and $\rho_{E_1} = |0_{E_1}\rangle\langle 0_{E_1}|$. Then

$$
\begin{aligned}
C_{E,f}(\mathcal{E}_2 \circ \mathcal{E}_1) &= \max_{\rho_Q} S_f \Big\{ (\mathcal{E}_2 \otimes \mathcal{I}_R) \Big[ \mathrm{Tr}_{E_1} \left( V_1 |\psi_{QRE_1}\rangle\langle\psi_{QRE_1}| V_1^\dagger \right) \Big] \,\Big|\Big| \\
&\qquad \mathcal{E}_2 \Big[ \mathrm{Tr}_{RE_1} \left( V_1 |\psi_{QRE_1}\rangle\langle\psi_{QRE_1}| V_1^\dagger \right) \Big] \otimes \rho_R \Big\} & (161) \\
&= \max_{\rho_Q} S_f \Big( \mathrm{Tr}_{E_1} \Big[ (\mathcal{E}_2 \otimes \mathcal{I}_R \otimes \mathcal{I}_{E_1}) \left( V_1 |\psi_{QRE_1}\rangle\langle\psi_{QRE_1}| V_1^\dagger \right) \Big] \,\Big|\Big| \\
&\qquad \mathrm{Tr}_{E_1} \Big\{ \mathcal{E}_2 \Big[ \mathrm{Tr}_{RE_1} \left( V_1 |\psi_{QRE_1}\rangle\langle\psi_{QRE_1}| V_1^\dagger \right) \Big] \otimes \rho_R \otimes \rho_{E_1} \Big\} \Big) & (162) \\
&\leq \max_{\rho_Q} S_f \Big\{ (\mathcal{E}_2 \otimes \mathcal{I}_R \otimes \mathcal{I}_{E_1}) \left( V_1 |\psi_{QRE_1}\rangle\langle\psi_{QRE_1}| V_1^\dagger \right) \,\Big|\Big| \\
&\qquad \mathcal{E}_2 \Big[ \mathrm{Tr}_{RE_1} \left( V_1 |\psi_{QRE_1}\rangle\langle\psi_{QRE_1}| V_1^\dagger \right) \Big] \otimes \rho_R \otimes \rho_{E_1} \Big\} & (163) \\
&= \max_{\mathcal{E}_1(\rho_Q)} S \Big\{ (\mathcal{E}_2 \otimes \mathcal{I}_R \otimes \mathcal{I}_{E_1}) \left( |\acute{\psi}_{QRE_1}\rangle\langle\acute{\psi}_{QRE_1}| \right) \,\Big|\Big| \\
&\qquad \mathcal{E}_2 \Big[ \mathrm{Tr}_{RE_1} \left( |\acute{\psi}_{QRE_1}\rangle\langle\acute{\psi}_{QRE_1}| \right) \Big] \otimes \rho_R \otimes \rho_{E_1} \Big\} & (164) \\
&\leq C_{E,f}(\mathcal{E}_2). & (165)
\end{aligned}
$$

Hence,

$$C_{E,f}(\mathcal{E}_2 \circ \mathcal{E}_1) \leq \min \left\{ C_{E,f}(\mathcal{E}_1), C_{E,f}(\mathcal{E}_2) \right\}, \tag{166}$$

which is the data processing inequality.

## 4.3 Coherent Information

Let $Q$ be our system of interest with density matrix $\rho$ that is purified in the composite system $RQ$. We consider two CPTP quantum operations $\mathcal{E}_1(\cdot)$ and $\mathcal{E}_2(\cdot)$ that are mocked up by introducing ancillae $E_1$ and $E_2$ respectively. The system after the operation $\mathcal{E}_1(\cdot)$ is denoted by $R'Q'E_1'E_2{}'$ and after the operation $\mathcal{E}_2(\cdot)$ by $R''Q''E_1''E_2{}''$.

The $f$-generalization of the coherent information is defined as

$$I_f(\rho, \mathcal{E}_1) = -S_f\left(\rho_{R'E_1'}\Big|\Big|\mathrm{I} \otimes \rho_{E_1'}\right), \tag{167}$$

where $\rho_{R'E_1'} = \mathrm{Tr}_{Q'E_2'}(\rho_{R'Q'E_1'E_2'})$ and $\rho_{E_1'} = \mathrm{Tr}_{R'Q'E_2'}(\rho_{R'Q'E_1'E_2'})$. For $f(t) = -\ln(t)$, it was shown by Shor that coherent information is related to the quantum channel capacity [24].

**Lemma 10** *For CPTP quantum operations $\mathcal{E}_1(\cdot)$ and $\mathcal{E}_2(\cdot)$, and strictly positive density matrix $\rho$,*

$$S_f(\rho) \geq I_f(\rho, \mathcal{E}_1) \geq I_f(\rho, \mathcal{E}_2 \circ \mathcal{E}_1). \tag{168}$$

*For a non-affine and diffused operator convex function $f(\cdot)$, the equality holds in the first inequality if and only if there exists a CPTP quantum operation $\mathcal{E}_2(\cdot)$ such that $F(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) = 1$.*

*Proof* We first prove the inequalities. Since $\mathcal{E}_1(\cdot)$ does not affect $R$, hence, $\rho_R = \rho_{R'}$. Similar reasoning for $\mathcal{E}_2(\cdot)$ yields $\rho_{R'E_1'} = \rho_{R''E_1''}$. We have using monotonicity

$$\begin{align}
S_f(\rho) &= -S_f(\rho_R||\mathrm{I}) \tag{169}\\
&= -S_f\left(\rho_{R'}||\mathrm{I}\right) \tag{170}\\
&\geq -S_f\left(\rho_{R'E_1'}||\mathrm{I} \otimes \rho_{E_1'}\right) \tag{171}\\
&= I_f(\rho, \mathcal{E}_1) \tag{172}
\end{align}$$

and

$$\begin{align}
I_f(\rho, \mathcal{E}_1) &= -S_f\left(\rho_{R'E_1'}||\mathrm{I} \otimes \rho_{E_1'}\right) \tag{173}\\
&= -S_f\left(\rho_{R''E_1''}||\mathrm{I} \otimes \rho_{E_1''}\right) \tag{174}\\
&\geq -S_f\left(\rho_{R''E_1''E_2''}||\mathrm{I} \otimes \rho_{E_1''E_2''}\right) \tag{175}\\
&= I_f(\rho, \mathcal{E}_2 \circ \mathcal{E}_1), \tag{176}
\end{align}$$

where Eqs. (171) and (175) follow using Eq. (60).

To prove the equality condition, let us first assume that $F(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) = 1$. This implies that

$$\langle\psi_{RQ}|\rho_{R''Q''}|\psi_{RQ}\rangle = 1, \tag{177}$$

and since $\rho_{R''Q''}$ is a density matrix, hence, it follows that $\langle \psi_{RQ} | \rho_{R''Q''} | \psi_{RQ} \rangle \leq 1$ with equality if and only if $| \psi_{RQ} \rangle$ is an eigenvector of $\rho_{R''Q''}$ with eigenvalue 1, or

$$\rho_{R''Q''} = | \psi_{RQ} \rangle \langle \psi_{RQ} |. \tag{178}$$

Since $\rho_{R''Q''}$ is in a pure state, hence, it follows that

$$\rho_{R''Q''E_1''E_2''} = | \psi_{RQ} \rangle \langle \psi_{RQ} | \otimes \rho_{E_1''E_2''} \tag{179}$$

and by tracing out $Q''$, we get

$$\rho_{R''E_1''E_2''} = \rho_R \otimes \rho_{E_1''E_2''}. \tag{180}$$

Hence,

$$I_f(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) = -S_f \left( \rho_{R''E_1''E_2''} || \mathrm{I} \otimes \rho_{E_1''E_2''} \right) \tag{181}$$

$$= -S_f \left( \rho_R \otimes \rho_{E_1''E_2''} || \mathrm{I} \otimes \rho_{E_1''E_2''} \right) \tag{182}$$

$$= -S_f(\rho_R || \mathrm{I}) \tag{183}$$

$$= S_f(\rho). \tag{184}$$

Result follows by using Eq. (168) and Eq. (184) to get

$$S_f(\rho) \geq I_f(\rho, \mathcal{E}_1) \geq I_f(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) = S_f(\rho). \tag{185}$$

To prove the statement in the other direction, let us assume that

$$I_f(\rho, \mathcal{E}_1) = S_f(\rho). \tag{186}$$

This implies that

$$S_f \left( \rho_{R'E_1'} || \mathrm{I} \otimes \rho_{E_1'} \right) = S_f \left( \rho_{R'} || \mathrm{I} \right). \tag{187}$$

Using Eq. (60), it follows that the above equality is true if and only if

$$\rho_{R'E_1'} = \rho_{R'} \otimes \rho_{E_1'}, \tag{188}$$

which is the same condition as the one for $f(t) = -\ln(t)$ in Ref. [25], following which, we can construct a recovery operation $\mathcal{E}_2(\cdot)$ such that $\rho_{R''Q''} = \rho_{RQ}$ or $F(\rho, \mathcal{E}_2 \circ \mathcal{E}_1) = 1$. ∎

# 5  Conclusions and Acknowledgements

In conclusion, we have studied the fundamental properties of the quantum $f$-relative entropy and the quantum $f$-entropy . We give the equality conditions under some properties for a class of operator convex functions. These conditions are more general than the previously known conditions and also apply to the case of $f(t) = -\ln(t)$. We define the quantum $f$-entropy in terms of the quantum $f$-relative entropy and study its properties giving the equality conditions in some cases. We also show that the $f$-generalizations of many well-known information-theoretic quantities also satisfy the data processing inequality and for the case of $f$-coherent information, we give the equality conditions.

# References

[1] M. Ohya and D. Petz, *Quantum Entropy and its use*, 1st ed. Berlin: Springer-Verlag, 1993.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[3] B. Schumacher, "Quantum coding," *Phys. Rev. A*, vol. 51, pp. 2738–2747, 1995.

[4] E. H. Lieb and M. B. Ruskai, "A fundamental property of quantum mechanical entropy," *Phys. Rev. Lett.*, vol. 30, pp. 434–436, 1973.

[5] ——, "Proof of the strong subadditivity of quantum mechanical entropy," *J. Math. Phys.*, vol. 14, pp. 1938–1941, 1973.

[6] P. Hayden, R. Jozsa, D. Petz, and A. Winter, "Structure of states which satisfy strong subadditivity of quantum entropy with equality," *Commun. Math. Phys.*, vol. 246, pp. 359–374, Feb. 2004.

[7] R. Umegaki, *Ködai Math. Sem. Rep.*, vol. 14, pp. 59–85, 1962.

[8] G. Lindblad, "Completely positive maps and entropy inequalities," *Commun. Math. Phys.*, vol. 40, pp. 147–151, June 1975.

[9] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von Neumann algebra," *Commun. Math. Phys.*, vol. 105, pp. 123–131, 1986.

[10] M. B. Ruskai, "Inequalities for quantum entropy: A review with conditions for equality," *J. Math. Phys.*, vol. 43, pp. 4358–4375, Sept. 2002.

[11] B. Ibinson, N. Linden, and A. Winter, "All inequalities for the relative entropy," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2006.

[12] R. Bhatia, *Matrix Analysis*. Springer-Verlag, 1997.

[13] F. Hansen and G. Pedersen, "Jensen's inequality for operators and Löwner's theorem," *Math. Ann.*, vol. 258, pp. 229–241, Sept. 1982.

[14] J. Watrous, "Theory of quantum information," *Online: http://www.cs.uwaterloo.ca/~watrous/lecture-notes.html*.

[15] A. Jencova and M. B. Ruskai, "A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality," *arXiv:0903.2895v2*.

[16] I. Csiszár, "A class of measures of informativity of observation channels," *Period. Math. Hung.*, vol. 2, pp. 191–213, Mar. 1970.

[17] D. Petz, "Quasi-entropies for finite quantum systems," *Rep. Math. Phys.*, vol. 23, pp. 57–65, Feb. 1986.

[18] M. A. Nielsen and D. Petz, "A simple proof of the strong subadditivity inequality," *Quantum Information and Computation*, vol. 5, pp. 507–513, Sept. 2005.

[19] E. G. Effros, "A matrix convexity approach to some celebrated quantum inequalities," in *Proc. Nat. Acad. Sciences USA*, R. V. Kadison, Ed., vol. 106, Jan. 2009, pp. 1006–1008.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[21] A. S. Holevo, "The capacity of a quantum channel with general signal states," *IEEE Trans. Inf. Theory*, vol. 22, pp. 269–273, Jan. 1998.

[22] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy classical channels," *Phys. Rev. A*, vol. 56, pp. 131–138, July 1997.

[23] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, pp. 2637–2655, Oct. 2002.

[24] P. W. Shor, "The quantum channel capacity and coherent information," in *MSRI Workshop on Quantum Computation*, http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/, 2002.

[25] B. Schumacher and M. A. Nielsen, "Quantum data processing and error correction," *Phys. Rev. A*, vol. 54, pp. 2629–2635, Oct. 1996.