Inequalities for Quantum f-Divergence of Trace Class Operators in Hilbert Spaces

INEQUALITIES FOR QUANTUM f-DIVERGENCE OF TRACE CLASS OPERATORS IN HILBERT SPACES

S.S. DRAGOMIR[1,2]

Abstract. Some inequalities for quantum f-divergence of trace class operators in Hilbert spaces are obtained. It is shown that for normalised convex functions it is nonnegative. Some upper bounds for quantum f-divergence in terms of variational and χ^2-distance are provided. Applications for some classes of divergence measures such as Umegaki and Tsallis relative entropies are also given.

1. Introduction

Let $(X,\mathcal{A})$ be a measurable space satisfying $|\mathcal{A}|>2$ and μ be a σ-finite measure on $(X,\mathcal{A})$. Let $\mathcal{P}$ be the set of all probability measures on $(X,\mathcal{A})$ which are absolutely continuous with respect to μ. For P, $Q\in\mathcal{P}$, let $p=\frac{dP}{d\mu}$ and $q=\frac{dQ}{d\mu}$ denote the *Radon-Nikodym* derivatives of P and Q with respect to μ.

Two probability measures P, $Q\in\mathcal{P}$ are said to be *orthogonal* and we denote this by $Q\perp P$ if

$$P(\{q=0\})=Q(\{p=0\})=1.$$

Let $f:[0,\infty)\to(-\infty,\infty]$ be a convex function that is continuous at 0, i.e., $f(0)=\lim_{u\downarrow 0}f(u)$.

In 1963, I. Csiszár [10] introduced the concept of f-divergence as follows.

Definition 1. *Let P, $Q\in\mathcal{P}$. Then*

$$I_f(Q,P)=\int_X p(x)\,f\left[\frac{q(x)}{p(x)}\right]d\mu(x), \tag{1.1}$$

is called the f-divergence of the probability distributions Q and P.

Remark 1. *Observe that, the integrand in the formula (1.1) is undefined when $p(x)=0$. The way to overcome this problem is to postulate for f as above that*

$$0f\left[\frac{q(x)}{0}\right]=q(x)\lim_{u\downarrow 0}\left[uf\left(\frac{1}{u}\right)\right],\ x\in X. \tag{1.2}$$

We now give some examples of f-divergences that are well-known and often used in the literature (see also [6]).

1991 *Mathematics Subject Classification.* 47A63; 47A99.

Key words and phrases. Selfadjoint bounded linear operators, Functions of operators, Trace of operators, quantum divergence measures, Umegaki and Tsallis relative entropies.

1.1. **The Class of χ^α-Divergences.** The f-divergences of this class, which is generated by the function χ^α, $\alpha \in [1,\infty)$, defined by

$$\chi^\alpha(u) = |u-1|^\alpha, \quad u \in [0,\infty)$$

have the form

$$I_f(Q,P) = \int_X p\left|\frac{q}{p}-1\right|^\alpha d\mu = \int_X p^{1-\alpha}|q-p|^\alpha d\mu. \tag{1.3}$$

From this class only the parameter $\alpha = 1$ provides a distance in the topological sense, namely the *total variation distance* $V(Q,P) = \int_X |q-p|\, d\mu$. The most prominent special case of this class is, however, *Karl Pearson's* χ^2*-divergence*

$$\chi^2(Q,P) = \int_X \frac{q^2}{p} d\mu - 1$$

that is obtained for $\alpha = 2$.

1.2. **Dichotomy Class.** From this class, generated by the function $f_\alpha : [0,\infty) \to \mathbb{R}$

$$f_\alpha(u) = \begin{cases} u - 1 - \ln u & \text{for } \alpha = 0; \\ \frac{1}{\alpha(1-\alpha)}\left[\alpha u + 1 - \alpha - u^\alpha\right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ 1 - u + u\ln u & \text{for } \alpha = 1; \end{cases}$$

only the parameter $\alpha = \frac{1}{2}$ $\left(f_{\frac{1}{2}}(u) = 2(\sqrt{u}-1)^2\right)$ provides a distance, namely, the *Hellinger distance*

$$H(Q,P) = \left[\int_X (\sqrt{q}-\sqrt{p})^2 d\mu\right]^{\frac{1}{2}}.$$

Another important divergence is the *Kullback-Leibler divergence* obtained for $\alpha = 1$,

$$KL(Q,P) = \int_X q \ln\left(\frac{q}{p}\right) d\mu.$$

1.3. **Matsushita's Divergences.** The elements of this class, which is generated by the function φ_α, $\alpha \in (0,1]$ given by

$$\varphi_\alpha(u) := |1-u^\alpha|^{\frac{1}{\alpha}}, \quad u \in [0,\infty),$$

are prototypes of metric divergences, providing the distances $\left[I_{\varphi_\alpha}(Q,P)\right]^\alpha$.

1.4. **Puri-Vincze Divergences.** This class is generated by the functions Φ_α, $\alpha \in [1,\infty)$ given by

$$\Phi_\alpha(u) := \frac{|1-u|^\alpha}{(u+1)^{\alpha-1}}, \quad u \in [0,\infty).$$

It has been shown in [27] that this class provides the distances $\left[I_{\Phi_\alpha}(Q,P)\right]^{\frac{1}{\alpha}}$.

1.5. **Divergences of Arimoto-type.** This class is generated by the functions

$$\Psi_\alpha(u) := \begin{cases} \frac{\alpha}{\alpha-1}\left[(1+u^\alpha)^{\frac{1}{\alpha}} - 2^{\frac{1}{\alpha}-1}(1+u)\right] & \text{for } \alpha \in (0,\infty)\setminus\{1\}; \\ (1+u)\ln 2 + u\ln u - (1+u)\ln(1+u) & \text{for } \alpha = 1; \\ \frac{1}{2}|1-u| & \text{for } \alpha = \infty. \end{cases}$$

It has been shown in [33] that this class provides the distances $[I_{\Psi_\alpha}(Q,P)]^{\min\left(\alpha,\frac{1}{\alpha}\right)}$ for $\alpha \in (0,\infty)$ and $\frac{1}{2}V(Q,P)$ for $\alpha = \infty$.

For f continuous convex on $[0,\infty)$ we obtain the $*$-*conjugate* function of f by

$$f^*(u) = uf\left(\frac{1}{u}\right), \quad u \in (0,\infty)$$

and

$$f^*(0) = \lim_{u\downarrow 0} f^*(u).$$

It is also known that if f is continuous convex on $[0,\infty)$ then so is f^*.

The following two theorems contain the most basic properties of f-divergences. For their proofs we refer the reader to Chapter 1 of [29] (see also [6]).

Theorem 1 (Uniqueness and Symmetry Theorem)**.** *Let f, f_1 be continuous convex on $[0,\infty)$. We have*

$$I_{f_1}(Q,P) = I_f(Q,P),$$

for all $P, Q \in \mathcal{P}$ if and only if there exists a constant $c \in \mathbb{R}$ such that

$$f_1(u) = f(u) + c(u-1),$$

for any $u \in [0,\infty)$.

Theorem 2 (Range of Values Theorem)**.** *Let $f : [0,\infty) \to \mathbb{R}$ be a continuous convex function on $[0,\infty)$.*

For any $P, Q \in \mathcal{P}$, we have the double inequality

$$f(1) \le I_f(Q,P) \le f(0) + f^*(0). \tag{1.4}$$

(i) *If $P = Q$, then the equality holds in the first part of (1.4).*

If f is strictly convex at 1*, then the equality holds in the first part of (1.4) if and only if $P = Q$;*

(ii) *If $Q \perp P$, then the equality holds in the second part of (1.4).*

If $f(0) + f^*(0) < \infty$, then equality holds in the second part of (1.4) if and only if $Q \perp P$.

The following result is a refinement of the second inequality in Theorem 2 (see [6, Theorem 3]).

Theorem 3. *Let f be a continuous convex function on $[0,\infty)$ with $f(1) = 0$ (f is normalised) and $f(0) + f^*(0) < \infty$. Then*

$$0 \le I_f(Q,P) \le \frac{1}{2}[f(0) + f^*(0)]V(Q,P) \tag{1.5}$$

for any $Q, P \in \mathcal{P}$.

For other inequalities for f-divergence see [5], [12]-[22].

Motivated by the above results, in this paper we obtain some new inequalities for quantum f-divergence of trace class operators in Hilbert spaces. It is shown that for normalised convex functions it is nonnegative. Some upper bounds for quantum f-divergence in terms of variational and χ^2-distance are provided. Applications for some classes of divergence measures such as Umegaki and Tsallis relative entropies are also given.

In what follows we recall some facts we need concerning the trace of operators and quantum f-divergence for trace class operators in infinite dimensional complex Hilbert spaces.

2. Some Preliminary Facts

2.1. **Some Facts on Trace of Operators.** Let $(H, \langle \cdot, \cdot \rangle)$ be a complex Hilbert space and $\{e_i\}_{i \in I}$ an *orthonormal basis* of H. We say that $A \in \mathcal{B}(H)$ is a *Hilbert-Schmidt operator* if

$$\sum_{i \in I} \|Ae_i\|^2 < \infty. \tag{2.1}$$

It is well know that, if $\{e_i\}_{i \in I}$ and $\{f_j\}_{j \in J}$ are orthonormal bases for H and $A \in \mathcal{B}(H)$ then

$$\sum_{i \in I} \|Ae_i\|^2 = \sum_{j \in I} \|Af_j\|^2 = \sum_{j \in I} \|A^* f_j\|^2 \tag{2.2}$$

showing that the definition (2.1) is independent of the orthonormal basis and A is a Hilbert-Schmidt operator iff A^* is a Hilbert-Schmidt operator.

Let $\mathcal{B}_2(H)$ the set of Hilbert-Schmidt operators in $\mathcal{B}(H)$. For $A \in \mathcal{B}_2(H)$ we define

$$\|A\|_2 := \left(\sum_{i \in I} \|Ae_i\|^2 \right)^{1/2} \tag{2.3}$$

for $\{e_i\}_{i \in I}$ an orthonormal basis of H. This definition does not depend on the choice of the orthonormal basis.

Using the triangle inequality in $l^2(I)$, one checks that $\mathcal{B}_2(H)$ is a *vector space* and that $\|\cdot\|_2$ is a norm on $\mathcal{B}_2(H)$, which is usually called in the literature as the *Hilbert-Schmidt norm.*

Denote *the modulus* of an operator $A \in \mathcal{B}(H)$ by $|A| := (A^*A)^{1/2}$.

Because $\||A|x\| = \|Ax\|$ for all $x \in H$, A is Hilbert-Schmidt iff $|A|$ is Hilbert-Schmidt and $\|A\|_2 = \||A|\|_2$. From (2.2) we have that if $A \in \mathcal{B}_2(H)$, then $A^* \in \mathcal{B}_2(H)$ and $\|A\|_2 = \|A^*\|_2$.

The following theorem collects some of the most important properties of Hilbert-Schmidt operators:

Theorem 4. *We have:*

(i) $(\mathcal{B}_2(H), \|\cdot\|_2)$ *is a Hilbert space with inner product*

$$\langle A, B \rangle_2 := \sum_{i \in I} \langle Ae_i, Be_i \rangle = \sum_{i \in I} \langle B^* Ae_i, e_i \rangle \tag{2.4}$$

and the definition does not depend on the choice of the orthonormal basis $\{e_i\}_{i \in I}$*;*

(ii) We have the inequalities

$$\|A\| \leq \|A\|_2 \tag{2.5}$$

for any $A \in \mathcal{B}_2(H)$ *and*

$$\|AT\|_2, \ \ \|TA\|_2 \leq \|T\| \|A\|_2 \tag{2.6}$$

for any $A \in \mathcal{B}_2(H)$ *and* $T \in \mathcal{B}(H)$;

(iii) $\mathcal{B}_2(H)$ *is an operator ideal in* $\mathcal{B}(H)$, *i.e.*

$$\mathcal{B}(H)\mathcal{B}_2(H)\mathcal{B}(H) \subseteq \mathcal{B}_2(H);$$

(iv) $\mathcal{B}_{fin}(H)$, *the space of operators of finite rank, is a dense subspace of* $\mathcal{B}_2(H)$;

(v) $\mathcal{B}_2(H) \subseteq \mathcal{K}(H)$, *where* $\mathcal{K}(H)$ *denotes the algebra of compact operators on* H.

If $\{e_i\}_{i\in I}$ an orthonormal basis of H, we say that $A \in \mathcal{B}(H)$ is *trace class* if

$$\|A\|_1 := \sum_{i\in I} \langle |A| e_i, e_i \rangle < \infty. \tag{2.7}$$

The definition of $\|A\|_1$ does not depend on the choice of the orthonormal basis $\{e_i\}_{i\in I}$. We denote by $\mathcal{B}_1(H)$ the set of trace class operators in $\mathcal{B}(H)$.

The following proposition holds:

Proposition 1. *If* $A \in \mathcal{B}(H)$, *then the following are equivalent:*

(i) $A \in \mathcal{B}_1(H)$;

(ii) $|A|^{1/2} \in \mathcal{B}_2(H)$;

(ii) A *(or* $|A|$*) is the product of two elements of* $\mathcal{B}_2(H)$.

The following properties are also well known:

Theorem 5. *With the above notations:*

(i) We have

$$\|A\|_1 = \|A^*\|_1 \ \text{ and } \ \|A\|_2 \leq \|A\|_1 \tag{2.8}$$

for any $A \in \mathcal{B}_1(H)$;

(ii) $\mathcal{B}_1(H)$ *is an operator ideal in* $\mathcal{B}(H)$, *i.e.*

$$\mathcal{B}(H)\mathcal{B}_1(H)\mathcal{B}(H) \subseteq \mathcal{B}_1(H);$$

(iii) We have

$$\mathcal{B}_2(H)\mathcal{B}_2(H) = \mathcal{B}_1(H);$$

(iv) We have

$$\|A\|_1 = \sup\{\langle A, B\rangle_2 \mid B \in \mathcal{B}_2(H), \ \|B\| \leq 1\};$$

(v) $(\mathcal{B}_1(H), \|\cdot\|_1)$ *is a Banach space.*

(iv) We have the following isometric isomorphisms

$$\mathcal{B}_1(H) \cong K(H)^* \ \text{ and } \mathcal{B}_1(H)^* \cong \mathcal{B}(H),$$

where $K(H)^*$ *is the dual space of* $K(H)$ *and* $\mathcal{B}_1(H)^*$ *is the dual space of* $\mathcal{B}_1(H)$.

We define the *trace* of a trace class operator $A \in \mathcal{B}_1(H)$ to be

$$\operatorname{tr}(A) := \sum_{i\in I} \langle Ae_i, e_i \rangle, \tag{2.9}$$

where $\{e_i\}_{i\in I}$ an orthonormal basis of H. Note that this coincides with the usual definition of the trace if H is finite-dimensional. We observe that the series (2.9) converges absolutely and it is independent from the choice of basis.

The following result collects some properties of the trace:

Theorem 6. *We have:*

(i) If $A \in \mathcal{B}_1(H)$ *then* $A^* \in \mathcal{B}_1(H)$ *and*

$$\operatorname{tr}(A^*) = \overline{\operatorname{tr}(A)}; \tag{2.10}$$

(ii) If $A \in \mathcal{B}_1(H)$ *and* $T \in \mathcal{B}(H)$, *then* $AT,\ TA \in \mathcal{B}_1(H)$ *and*

$$\operatorname{tr}(AT) = \operatorname{tr}(TA) \text{ and } |\operatorname{tr}(AT)| \leq \|A\|_1 \|T\|; \tag{2.11}$$

(iii) $\operatorname{tr}(\cdot)$ *is a bounded linear functional on* $\mathcal{B}_1(H)$ *with* $\|\operatorname{tr}\| = 1$;

(iv) If $A,\ B \in \mathcal{B}_2(H)$ *then* $AB,\ BA \in \mathcal{B}_1(H)$ *and* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$;

(v) $\mathcal{B}_{fin}(H)$ *is a dense subspace of* $\mathcal{B}_1(H)$.

Utilising the trace notation we obviously have that

$$\langle A, B\rangle_2 = \operatorname{tr}(B^*A) = \operatorname{tr}(AB^*) \text{ and } \|A\|_2^2 = \operatorname{tr}(A^*A) = \operatorname{tr}\left(|A|^2\right)$$

for any $A,\ B \in \mathcal{B}_2(H)$.

The following Hölder's type inequality has been obtained by Ruskai in [36]

$$|\operatorname{tr}(AB)| \leq \operatorname{tr}(|AB|) \leq \left[\operatorname{tr}\left(|A|^{1/\alpha}\right)\right]^{\alpha}\left[\operatorname{tr}\left(|B|^{1/(1-\alpha)}\right)\right]^{1-\alpha} \tag{2.12}$$

where $\alpha \in (0,1)$ and $A,\ B \in \mathcal{B}(H)$ with $|A|^{1/\alpha},\ |B|^{1/(1-\alpha)} \in \mathcal{B}_1(H)$.

In particular, for $\alpha = \frac{1}{2}$ we get the Schwarz inequality

$$|\operatorname{tr}(AB)| \leq \operatorname{tr}(|AB|) \leq \left[\operatorname{tr}\left(|A|^2\right)\right]^{1/2}\left[\operatorname{tr}\left(|B|^2\right)\right]^{1/2} \tag{2.13}$$

with $A,\ B \in \mathcal{B}_2(H)$.

If $A \geq 0$ and $P \in \mathcal{B}_1(H)$ with $P \geq 0$, then

$$0 \leq \operatorname{tr}(PA) \leq \|A\| \operatorname{tr}(P). \tag{2.14}$$

Indeed, since $A \geq 0$, then $\langle Ax, x\rangle \geq 0$ for any $x \in H$. If $\{e_i\}_{i\in I}$ an orthonormal basis of H, then

$$0 \leq \left\langle AP^{1/2}e_i, P^{1/2}e_i\right\rangle \leq \|A\| \left\|P^{1/2}e_i\right\|^2 = \|A\| \langle Pe_i, e_i\rangle$$

for any $i \in I$. Summing over $i \in I$ we get

$$0 \leq \sum_{i\in I}\left\langle AP^{1/2}e_i, P^{1/2}e_i\right\rangle \leq \|A\| \sum_{i\in I}\langle Pe_i, e_i\rangle = \|A\| \operatorname{tr}(P)$$

and since

$$\sum_{i\in I}\left\langle AP^{1/2}e_i, P^{1/2}e_i\right\rangle = \sum_{i\in I}\left\langle P^{1/2}AP^{1/2}e_i, e_i\right\rangle = \operatorname{tr}\left(P^{1/2}AP^{1/2}\right) = \operatorname{tr}(PA)$$

we obtain the desired result (2.14).

This obviously imply the fact that, if A and B are selfadjoint operators with $A \leq B$ and $P \in \mathcal{B}_1(H)$ with $P \geq 0$, then

$$\operatorname{tr}(PA) \leq \operatorname{tr}(PB). \tag{2.15}$$

Now, if A is a selfadjoint operator, then we know that

$$|\langle Ax, x\rangle| \le \langle |A|\, x, x\rangle \text{ for any } x \in H.$$

This inequality follows by Jensen's inequality for the convex function $f(t) = |t|$ defined on a closed interval containing the spectrum of A.

If $\{e_i\}_{i\in I}$ is an orthonormal basis of H, then

$$\begin{aligned} |\operatorname{tr}(PA)| &= \left|\sum_{i\in I} \left\langle AP^{1/2}e_i, P^{1/2}e_i\right\rangle\right| \le \sum_{i\in I} \left|\left\langle AP^{1/2}e_i, P^{1/2}e_i\right\rangle\right| \\ &\le \sum_{i\in I} \left\langle |A|\, P^{1/2}e_i, P^{1/2}e_i\right\rangle = \operatorname{tr}(P|A|), \end{aligned} \tag{2.16}$$

for any A a selfadjoint operator and $P \in \mathcal{B}_1^+(H) := \{P \in \mathcal{B}_1(H) \text{ with } P \ge 0\}$.

For the theory of trace functionals and their applications the reader is referred to [39].

For some classical trace inequalities see [7], [9], [32] and [43], which are continuations of the work of Bellman [3]. For related works the reader can refer to [1], [4], [7], [24], [28], [30], [31], [37] and [40].

2.2. **Quantum f-Divergence for Trace Class Operators.** On complex Hilbert space $(\mathcal{B}_2(H), \langle\cdot,\cdot\rangle_2)$, where the Hilbert-Schmidt inner product is defined by

$$\langle U, V\rangle_2 := \operatorname{tr}(V^*U),\ U,\ V \in \mathcal{B}_2(H),$$

for $A,\ B \in \mathcal{B}^+(H)$ consider the operators $\mathfrak{L}_A : \mathcal{B}_2(H) \to \mathcal{B}_2(H)$ and $\mathfrak{R}_B : \mathcal{B}_2(H) \to \mathcal{B}_2(H)$ defined by

$$\mathfrak{L}_A T := AT \text{ and } \mathfrak{R}_B T := TB.$$

We observe that they are well defined and since

$$\langle \mathfrak{L}_A T, T\rangle_2 = \langle AT, T\rangle_2 = \operatorname{tr}(T^*AT) = \operatorname{tr}\left(|T^*|^2 A\right) \ge 0$$

and

$$\langle \mathfrak{R}_B T, T\rangle_2 = \langle TB, T\rangle_2 = \operatorname{tr}(T^*TB) = \operatorname{tr}\left(|T|^2 B\right) \ge 0$$

for any $T \in \mathcal{B}_2(H)$, they are also positive in the operator order of $\mathcal{B}(\mathcal{B}_2(H))$, the Banach algebra of all bounded operators on $\mathcal{B}_2(H)$ with the norm $\|\cdot\|_2$ where $\|T\|_2 = \operatorname{tr}\left(|T|^2\right)$, $T \in \mathcal{B}_2(H)$.

Since $\operatorname{tr}\left(|X^*|^2\right) = \operatorname{tr}\left(|X|^2\right)$ for any $X \in \mathcal{B}_2(H)$, then also

$$\begin{aligned} \operatorname{tr}(T^*AT) &= \operatorname{tr}\left(T^*A^{1/2}A^{1/2}T\right) = \operatorname{tr}\left(\left(A^{1/2}T\right)^* A^{1/2}T\right) \\ &= \operatorname{tr}\left(\left|A^{1/2}T\right|^2\right) = \operatorname{tr}\left(\left|\left(A^{1/2}T\right)^*\right|^2\right) = \operatorname{tr}\left(\left|T^*A^{1/2}\right|^2\right) \end{aligned}$$

for $A \ge 0$ and $T \in \mathcal{B}_2(H)$.

We observe that $\mathfrak{L}_A$ and $\mathfrak{R}_B$ are commutative, therefore the product $\mathfrak{L}_A\mathfrak{R}_B$ is a selfadjoint positive operator in $\mathcal{B}(\mathcal{B}_2(H))$ for any positive operators $A, B \in \mathcal{B}(H)$.

For $A, B \in \mathcal{B}^+(H)$ with B invertible, we define the *Araki transform* $\mathfrak{A}_{A,B} : \mathcal{B}_2(H) \to \mathcal{B}_2(H)$ by $\mathfrak{A}_{A,B} := \mathfrak{L}_A\mathfrak{R}_{B^{-1}}$. We observe that for $T \in \mathcal{B}_2(H)$ we have $\mathfrak{A}_{A,B}T = ATB^{-1}$ and

$$\langle \mathfrak{A}_{A,B}T, T\rangle_2 = \left\langle ATB^{-1}, T\right\rangle_2 = \operatorname{tr}\left(T^*ATB^{-1}\right).$$

Observe also, by the properties of trace, that

$$\begin{aligned}\operatorname{tr}\left(T^{*}ATB^{-1}\right) &= \operatorname{tr}\left(B^{-1/2}T^{*}A^{1/2}A^{1/2}TB^{-1/2}\right)\\ &= \operatorname{tr}\left(\left(A^{1/2}TB^{-1/2}\right)^{*}\left(A^{1/2}TB^{-1/2}\right)\right) = \operatorname{tr}\left(\left|A^{1/2}TB^{-1/2}\right|^{2}\right)\end{aligned}$$

giving that

$$\left\langle \mathfrak{A}_{A,B}T,T\right\rangle_{2} = \operatorname{tr}\left(\left|A^{1/2}TB^{-1/2}\right|^{2}\right) \geq 0 \tag{2.17}$$

for any $T \in \mathcal{B}_2(H)$.

We observe that, by the definition of operator order and by (2.17) we have $r1_{\mathcal{B}_2(H)} \leq \mathfrak{A}_{A,B} \leq R1_{\mathcal{B}_2(H)}$ for some $R \geq r \geq 0$ if and only if

$$r\operatorname{tr}\left(|T|^{2}\right) \leq \operatorname{tr}\left(\left|A^{1/2}TB^{-1/2}\right|^{2}\right) \leq R\operatorname{tr}\left(|T|^{2}\right) \tag{2.18}$$

for any $T \in \mathcal{B}_2(H)$.

We also notice that a sufficient condition for (2.18) to hold is that the following inequality in the operator order of $\mathcal{B}(H)$ is satisfied

$$r|T|^{2} \leq \left|A^{1/2}TB^{-1/2}\right|^{2} \leq R|T|^{2} \tag{2.19}$$

for any $T \in \mathcal{B}_2(H)$.

Let U be a selfadjoint linear operator on a complex Hilbert space $(K;\langle\cdot,\cdot\rangle)$. The *Gelfand map* establishes a $*$-isometrically isomorphism Φ between the set $C(\operatorname{Sp}(U))$ of all *continuous functions* defined on the *spectrum* of U, denoted $\operatorname{Sp}(U)$, and the C^*-algebra $C^*(U)$ generated by U and the identity operator 1_K on K as follows:

For any $f,g \in C(\operatorname{Sp}(U))$ and any $\alpha,\beta \in \mathbb{C}$ we have

(i) $\Phi(\alpha f+\beta g) = \alpha\Phi(f)+\beta\Phi(g)$;

(ii) $\Phi(fg) = \Phi(f)\Phi(g)$ and $\Phi(\bar{f}) = \Phi(f)^{*}$;

(iii) $\|\Phi(f)\| = \|f\| := \sup_{t\in\operatorname{Sp}(U)}|f(t)|$;

(iv) $\Phi(f_0) = 1_K$ and $\Phi(f_1) = U$, where $f_0(t) = 1$ and $f_1(t) = t$, for $t \in \operatorname{Sp}(U)$.

With this notation we define

$$f(U) := \Phi(f) \quad \text{for all } f \in C(\operatorname{Sp}(U))$$

and we call it the *continuous functional calculus* for a selfadjoint operator U.

If U is a selfadjoint operator and f is a real valued continuous function on $\operatorname{Sp}(U)$, then $f(t) \geq 0$ for any $t \in \operatorname{Sp}(U)$ implies that $f(U) \geq 0$, i.e. $f(U)$ is a positive operator on K. Moreover, if both f and g are real valued functions on $\operatorname{Sp}(U)$ then the following important property holds:

$$f(t) \geq g(t) \quad \text{for any} \quad t \in \operatorname{Sp}(U) \quad \text{implies that} \quad f(U) \geq g(U) \tag{P}$$

in the operator order of $B(K)$.

Let $f:[0,\infty) \to \mathbb{R}$ be a continuous function. Utilising the continuous functional calculus for the Araki selfadjoint operator $\mathfrak{A}_{Q,P} \in \mathcal{B}(\mathcal{B}_2(H))$ we can define the *quantum f-divergence* for $Q,P \in S_1(H) := \{P \in \mathcal{B}_1(H),\ P \geq 0 \text{ with } \operatorname{tr}(P) = 1\}$ and P invertible, by

$$S_f(Q,P) := \left\langle f(\mathfrak{A}_{Q,P})P^{1/2},P^{1/2}\right\rangle_2 = \operatorname{tr}\left(P^{1/2}f(\mathfrak{A}_{Q,P})P^{1/2}\right).$$

If we consider the continuous convex function $f:[0,\infty)\to\mathbb{R}$, with $f(0):=0$ and $f(t)=t\ln t$ for $t>0$ then for $Q,P\in S_1(H)$ and Q,P invertible we have

$$S_f(Q,P)=\operatorname{tr}[Q(\ln Q-\ln P)]=:U(Q,P),$$

which is the *Umegaki relative entropy.*

If we take the continuous convex function $f:[0,\infty)\to\mathbb{R}$, $f(t)=|t-1|$ for $t\geq 0$ then for $Q,P\in S_1(H)$ with P invertible we have

$$S_f(Q,P)=\operatorname{tr}(|Q-P|)=:V(Q,P),$$

where $V(Q,P)$ is the *variational distance.*

If we take $f:[0,\infty)\to\mathbb{R}$, $f(t)=t^2-1$ for $t\geq 0$ then for $Q,P\in S_1(H)$ with P invertible we have

$$S_f(Q,P)=\operatorname{tr}\left(Q^2P^{-1}\right)-1=:\chi^2(Q,P),$$

which is called the χ^2-*distance*

Let $q\in(0,1)$ and define the convex function $f_q:[0,\infty)\to\mathbb{R}$ by $f_q(t)=\frac{1-t^q}{1-q}$. Then

$$S_{f_q}(Q,P)=\frac{1-\operatorname{tr}\left(Q^qP^{1-q}\right)}{1-q},$$

which is *Tsallis relative entropy.*

If we consider the convex function $f:[0,\infty)\to\mathbb{R}$ by $f(t)=\frac{1}{2}\left(\sqrt{t}-1\right)^2$, then

$$S_f(Q,P)=1-\operatorname{tr}\left(Q^{1/2}P^{1/2}\right)=:h^2(Q,P),$$

which is known as *Hellinger discrimination.*

If we take $f:(0,\infty)\to\mathbb{R}$, $f(t)=-\ln t$ then for $Q,P\in S_1(H)$ and Q,P invertible we have

$$S_f(Q,P)=\operatorname{tr}[P(\ln P-\ln Q)]=U(P,Q).$$

The reader can obtain other particular quantum f-divergence measures by utilizing the normalized convex functions from Introduction, namely the convex functions defining the dichotomy class, Matsushita's divergences, Puri-Vincze divergences or divergences of Arimoto-type. We omit the details.

In the important case of finite dimensional space H and the generalized inverse P^{-1}, numerous properties of the quantum f-divergence, mostly in the case when f is *operator convex,* have been obtained in the recent papers [25], [26], [34], [35] and the references therein.

In what follows we obtain several inequalities for the larger class of convex functions on an interval.

3. Inequalities for f Convex and Normalized

Suppose that I is an interval of real numbers with interior $\mathring{I}$ and $f:I\to\mathbb{R}$ is a convex function on I. Then f is continuous on $\mathring{I}$ and has finite left and right derivatives at each point of $\mathring{I}$. Moreover, if $x,y\in\mathring{I}$ and $x<y$, then $f'_-(x)\leq f'_+(x)\leq f'_-(y)\leq f'_+(y)$, which shows that both f'_- and f'_+ are nondecreasing function on $\mathring{I}$. It is also known that a convex function must be differentiable except for at most countably many points.

For a convex function $f : I \to \mathbb{R}$, the subdifferential of f denoted by ∂f is the set of all functions $\varphi : I \to [-\infty, \infty]$ such that $\varphi\left(\mathring{I}\right) \subset \mathbb{R}$ and

$$f(x) \geq f(a) + (x-a)\varphi(a) \text{ for any } x, a \in I. \tag{G}$$

It is also well known that if f is convex on I, then ∂f is nonempty, f'_-, $f'_+ \in \partial f$ and if $\varphi \in \partial f$, then

$$f'_-(x) \leq \varphi(x) \leq f'_+(x) \text{ for any } x \in \mathring{I}.$$

In particular, φ is a nondecreasing function.

If f is differentiable and convex on $\mathring{I}$, then $\partial f = \{f'\}$.

We are able now to state and prove the first result concerning the quantum f-divergence for the general case of convex functions.

Theorem 7. *Let $f : [0, \infty) \to \mathbb{R}$ be a continuous convex function that is normalized, i.e. $f(1) = 0$. Then for any $Q, P \in S_1(H)$, with P invertible, we have*

$$0 \leq S_f(Q, P). \tag{3.1}$$

Moreover, if f is continuously differentiable, then also

$$S_f(Q, P) \leq S_{\ell f'}(Q, P) - S_{f'}(Q, P), \tag{3.2}$$

where the function ℓ is defined as $\ell(t) = t$, $t \in \mathbb{R}$.

Proof. Since f is convex and normalized, then by the gradient inequality (G) we have

$$f(t) \geq (t-1) f'_+(1)$$

for $t > 0$.

Applying the property (P) for the operator $\mathfrak{A}_{Q,P}$, then we have for any $T \in \mathcal{B}_2(H)$

$$\begin{aligned} \langle f(\mathfrak{A}_{Q,P}) T, T\rangle_2 &\geq f'_+(1) \left\langle \left(\mathfrak{A}_{Q,P} - 1_{\mathcal{B}_2(H)}\right) T, T\right\rangle_2 \\ &= f'_+(1) \left[\langle \mathfrak{A}_{Q,P} T, T\rangle_2 - \|T\|_2\right], \end{aligned}$$

which, in terms of trace, can be written as

$$\operatorname{tr}(T^* f(\mathfrak{A}_{Q,P}) T) \geq f'_+(1) \left[\operatorname{tr}\left(\left|Q^{1/2} T P^{-1/2}\right|^2\right) - \operatorname{tr}\left(|T|^2\right)\right] \tag{3.3}$$

for any $T \in \mathcal{B}_2(H)$.

The inequality (3.3) is of interest in itself.

Now, if we take in (3.3) $T = P^{1/2}$ where $P \in S_1(H)$, with P invertible, then we get

$$S_f(Q, P) \geq f'_+(1) [\operatorname{tr}(Q) - \operatorname{tr}(P)] = 0$$

and the inequality (3.1) is proved.

Further, if f is continuously differentiable, then by the gradient inequality we also have

$$(t-1) f'(t) \geq f(t)$$

for $t > 0$.

Applying the property (P) for the operator $\mathfrak{A}_{Q,P}$, then we have for any $T \in \mathcal{B}_2(H)$

$$\left\langle \left(\mathfrak{A}_{Q,P} - 1_{\mathcal{B}_2(H)}\right) f'(\mathfrak{A}_{Q,P}) T, T\right\rangle_2 \geq \langle f(\mathfrak{A}_{Q,P}) T, T\rangle_2,$$

namely

$$\left\langle \mathfrak{A}_{Q,P} f'\left(\mathfrak{A}_{Q,P}\right) T, T\right\rangle_2 - \left\langle f'\left(\mathfrak{A}_{Q,P}\right) T, T\right\rangle_2 \geq \left\langle f\left(\mathfrak{A}_{Q,P}\right) T, T\right\rangle_2,$$

for any $T \in \mathcal{B}_2(H)$, or in terms of trace

$$\operatorname{tr}\left(T^* \mathfrak{A}_{Q,P} f'\left(\mathfrak{A}_{Q,P}\right) T\right) - \operatorname{tr}\left(T^* f'\left(\mathfrak{A}_{Q,P}\right) T\right) \geq \operatorname{tr}\left(T^* f\left(\mathfrak{A}_{Q,P}\right) T\right), \tag{3.4}$$

for any $T \in \mathcal{B}_2(H)$.

This inequality is also of interest in itself.

If in (3.4) we take $T = P^{1/2}$, where $P \in S_1(H)$, with P invertible, then we get the desired result (3.2). □

Remark 2. *If we take in (3.2) $f : (0, \infty) \to \mathbb{R}$, $f(t) = -\ln t$ then for $Q, P \in S_1(H)$ and Q, P invertible we have*

$$0 \leq U(P, Q) \leq \chi^2(P, Q). \tag{3.5}$$

We need the following lemma that is of interest in itself.

Lemma 1. *Let S be a selfadjoint operator on the Hilbert space $(K, \langle \cdot, \cdot \rangle)$ and with spectrum $\operatorname{Sp}(S) \subseteq [\gamma, \Gamma]$ for some real numbers γ, Γ. If $g : [\gamma, \Gamma] \to \mathbb{C}$ is a continuous function such that*

$$|g(t) - \lambda| \leq \rho \text{ for any } t \in [\gamma, \Gamma] \tag{3.6}$$

for some complex number $\lambda \in \mathbb{C}$ and positive number ρ, then

$$\begin{aligned} |\langle Sg(S) x, x\rangle - \langle Sx, x\rangle \langle g(S) x, x\rangle| &\leq \rho \langle |S - \langle Sx, x\rangle 1_H| x, x\rangle \\ &\leq \rho \left[\left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2\right]^{1/2} \end{aligned} \tag{3.7}$$

for any $x \in K$, $\|x\| = 1$.

Proof. We observe that

$$\langle Sg(S) x, x\rangle - \langle Sx, x\rangle \langle g(S) x, x\rangle = \langle (S - \langle Sx, x\rangle 1_H)(g(S) - \lambda 1_H) x, x\rangle \tag{3.8}$$

for any $x \in K$, $\|x\| = 1$.

For any selfadjoint operator B we have the modulus inequality

$$|\langle Bx, x\rangle| \leq \langle |B| x, x\rangle \text{ for any } x \in K, \|x\| = 1. \tag{3.9}$$

Also, utilizing the continuous functional calculus we have for each fixed $x \in K$, $\|x\| = 1$

$$\begin{aligned} |(S - \langle Sx, x\rangle 1_H)(g(S) - \lambda 1_H)| &= |S - \langle Sx, x\rangle 1_H| \, |g(S) - \lambda 1_H| \\ &\leq \rho \, |S - \langle Sx, x\rangle 1_H|, \end{aligned}$$

which implies that

$$\langle |(S - \langle Sx, x\rangle 1_H)(g(S) - \lambda 1_H)| x, x\rangle \leq \rho \langle |S - \langle Sx, x\rangle 1_H| x, x\rangle \tag{3.10}$$

for any $x \in K$, $\|x\| = 1$.

Therefore, by taking the modulus in (3.8) and utilizing (3.9) and (3.10) we get

$$\begin{aligned} &|\langle Sg(S) x, x\rangle - \langle Sx, x\rangle \langle g(S) x, x\rangle| \\ &= |\langle (S - \langle Sx, x\rangle 1_H)(g(S) - \lambda 1_H) x, x\rangle| \\ &\leq \langle |(S - \langle Sx, x\rangle 1_H)(g(S) - \lambda 1_H)| x, x\rangle \\ &\leq \rho \langle |S - \langle Sx, x\rangle 1_H| x, x\rangle \end{aligned} \tag{3.11}$$

for any $x \in K, \|x\| = 1$, which proves the first inequality in (3.7).

Using Schwarz inequality we also have

$$\begin{aligned}\langle |S - \langle Sx, x\rangle 1_H| \, x, x\rangle &\le \left\langle (S - \langle Sx, x\rangle 1_H)^2 x, x\right\rangle^{1/2} \\ &= \left[\left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2\right]^{1/2}\end{aligned}$$

for any $x \in K, \|x\| = 1$, and the lemma is proved. □

Corollary 1. *With the assumption of Lemma 1, we have*

$$\begin{aligned}(3.12)\qquad 0 &\le \left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2 \le \frac{1}{2}(\Gamma - \gamma)\langle |S - \langle Sx, x\rangle 1_H| \, x, x\rangle \\ &\le \frac{1}{2}(\Gamma - \gamma)\left[\left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2\right]^{1/2} \le \frac{1}{4}(\Gamma - \gamma)^2,\end{aligned}$$

for any $x \in K$, $\|x\| = 1$.

Proof. If we take in Lemma 1 $g(t) = t$, $\lambda = \frac{1}{2}(\Gamma + \gamma)$ and $\rho = \frac{1}{2}(\Gamma - \gamma)$, then we get

$$\begin{aligned}(3.13)\qquad 0 &\le \left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2 \le \frac{1}{2}(\Gamma - \gamma)\langle |S - \langle Sx, x\rangle 1_H| \, x, x\rangle \\ &\le \frac{1}{2}(\Gamma - \gamma)\left[\left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2\right]^{1/2}\end{aligned}$$

for any $x \in K$, $\|x\| = 1$.

From the first and last terms in (3.13) we have

$$\left[\left\langle S^2 x, x\right\rangle - \langle Sx, x\rangle^2\right]^{1/2} \le \frac{1}{2}(\Gamma - \gamma),$$

which proves the rest of (3.12). □

We can prove the following result that provides simpler upper bounds for the quantum f-divergence when the operators P and Q satisfy the condition (2.18).

Theorem 8. *Let $f : [0, \infty) \to \mathbb{R}$ be a continuous convex function that is normalized. If $Q, P \in S_1(H)$, with P invertible, and there exists $R \ge 1 \ge r \ge 0$ such that*

$$(3.14)\qquad r \operatorname{tr}\left(|T|^2\right) \le \operatorname{tr}\left(\left|Q^{1/2} T P^{-1/2}\right|^2\right) \le R \operatorname{tr}\left(|T|^2\right)$$

for any $T \in \mathcal{B}_2(H)$, then

$$\begin{aligned}(3.15)\qquad 0 \le S_f(Q, P) &\le \frac{1}{2}\left[f'_-(R) - f'_+(r)\right] V(Q, P) \\ &\le \frac{1}{2}\left[f'_-(R) - f'_+(r)\right] \chi(Q, P) \\ &\le \frac{1}{4}(R - r)\left[f'_-(R) - f'_+(r)\right].\end{aligned}$$

Proof. Without loosing the generality, we prove the inequality in the case that f is continuously differentiable on $(0, \infty)$.

Since f' is monotonic nondecreasing on $[r, R]$ we have that

$$f'(r) \le f'(t) \le f'(R) \text{ for any } t \in [r, R],$$

which implies that

$$\left|f'(t)-\frac{f'(R)+f'(r)}{2}\right|\le\frac{1}{2}\left[f'(R)-f'(r)\right]$$

for any $t\in[r,R]$.

Applying Lemma 1 and Corollary 1 in the Hilbert space $(\mathcal{B}_2(H),\langle\cdot,\cdot\rangle_2)$ and for the selfadjoint operator $\mathfrak{A}_{Q,P}$ we have

$$\begin{aligned}
&\left|\langle\mathfrak{A}_{Q,P}f'(\mathfrak{A}_{Q,P})T,T\rangle_2-\langle\mathfrak{A}_{Q,P}T,T\rangle_2\langle f'(\mathfrak{A}_{Q,P})T,T\rangle_2\right|\\
&\le\frac{1}{2}\left[f'(R)-f'(r)\right]\left\langle\left|\mathfrak{A}_{Q,P}-\langle\mathfrak{A}_{Q,P}T,T\rangle_2 1_{\mathcal{B}_2(H)}\right|T,T\right\rangle_2\\
&\le\frac{1}{2}\left[f'(R)-f'(r)\right]\left[\left\langle\mathfrak{A}_{Q,P}^2T,T\right\rangle_2-\langle\mathfrak{A}_{Q,P}T,T\rangle_2^2\right]^{1/2}\\
&\le\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right]
\end{aligned}$$

for any $T\in\mathcal{B}_2(H)$, $\|T\|_2=1$, which is an inequality of interest in itself as well.

If in this inequality we take $T=P^{1/2}$, $P\in S_1(H)$, with P invertible, then we get

$$\begin{aligned}
&\left|\left\langle\mathfrak{A}_{Q,P}f'(\mathfrak{A}_{Q,P})P^{1/2},P^{1/2}\right\rangle_2-\left\langle f'(\mathfrak{A}_{Q,P})P^{1/2},P^{1/2}\right\rangle_2\right|\\
&\le\frac{1}{2}\left[f'(R)-f'(r)\right]\left\langle\left|\mathfrak{A}_{Q,P}-\left\langle\mathfrak{A}_{Q,P}P^{1/2},P^{1/2}\right\rangle_2 1_{\mathcal{B}_2(H)}\right|P^{1/2},P^{1/2}\right\rangle_2\\
&\le\frac{1}{2}\left[f'(R)-f'(r)\right]\left[\left\langle\mathfrak{A}_{Q,P}^2P^{1/2},P^{1/2}\right\rangle_2-\left\langle\mathfrak{A}_{Q,P}P^{1/2},P^{1/2}\right\rangle_2^2\right]^{1/2}\\
&\le\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right],
\end{aligned}$$

which can be written as

$$\begin{aligned}
\left|S_{\ell f'}(Q,P)-S_{f'}(Q,P)\right|&\le\frac{1}{2}\left[f'_-(R)-f'_+(r)\right]V(Q,P)\\
&\le\frac{1}{2}\left[f'_-(R)-f'_+(r)\right]\chi(Q,P)\\
&\le\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right].
\end{aligned}$$

Making use of Theorem 7 we deduce the desired result (3.15). □

Remark 3. *If we take in (3.15) $f(t)=t^2-1$, then we get*

$$\begin{aligned}
0&\le\chi^2(Q,P)\le\frac{1}{2}(R-r)V(Q,P)\le\frac{1}{2}(R-r)\chi(Q,P)\\
&\le\frac{1}{4}(R-r)^2
\end{aligned}\tag{3.16}$$

for $Q,P\in S_1(H)$, with P invertible and satisfying the condition (3.14).

If we take in (3.15) $f(t)=t\ln t$, then we get the inequality

$$\begin{aligned}
0&\le U(Q,P)\le\frac{1}{2}\ln\left(\frac{R}{r}\right)V(Q,P)\le\frac{1}{2}\ln\left(\frac{R}{r}\right)\chi(Q,P)\\
&\le\frac{1}{4}(R-r)\ln\left(\frac{R}{r}\right)
\end{aligned}\tag{3.17}$$

provided that $Q,P \in S_1(H)$, *with* P, Q *invertible and satisfying the condition (3.14).*

With the same conditions and if we take $f(t) = -\ln t$, *then*

$$0 \le U(P,Q) \le \frac{R-r}{2rR} V(Q,P) \le \frac{R-r}{2rR}\chi(Q,P) \le \frac{(R-r)^2}{4rR}. \tag{3.18}$$

If we take in (3.15) $f(t) = f_q(t) = \frac{1-t^q}{1-q}$, *then we get*

$$\begin{aligned} 0 \le S_{f_q}(Q,P) &\le \frac{q}{2(1-q)}\left(\frac{R^{1-q}-r^{1-q}}{R^{1-q}r^{1-q}}\right) V(Q,P) \\ &\le \frac{q}{2(1-q)}\left(\frac{R^{1-q}-r^{1-q}}{R^{1-q}r^{1-q}}\right)\chi(Q,P) \\ &\le \frac{q}{4(1-q)}\left(\frac{R^{1-q}-r^{1-q}}{R^{1-q}r^{1-q}}\right)(R-r) \end{aligned} \tag{3.19}$$

provided that $Q,P \in S_1(H)$, *with* P, Q *invertible and satisfying the condition (3.14).*

4. Other Reverse Inequalities

Utilising different techniques we can obtain other upper bounds for the quantum f-divergence as follows. Applications for Umegaki relative entropy and χ^2-divergence are also provided.

Theorem 9. *Let* $f:[0,\infty) \to \mathbb{R}$ *be a continuous convex function that is normalized. If* $Q,P \in S_1(H)$, *with* P *invertible, and there exists* $R \ge 1 \ge r \ge 0$ *such that the condition (3.14) is satisfied, then*

$$0 \le S_f(Q,P) \le \frac{(R-1)f(r) + (1-r)f(R)}{R-r}. \tag{4.1}$$

Proof. By the convexity of f we have

$$f(t) = f\left(\frac{(R-t)r + (t-r)R}{R-r}\right) \le \frac{(R-t)f(r) + (t-r)f(R)}{R-r}$$

for any $t \in [r,R]$.

This inequality implies the following inequality in the operator order of $\mathcal{B}(\mathcal{B}_2(H))$

$$f(\mathfrak{A}_{Q,P}) \le \frac{\left(R1_{\mathcal{B}_2(H)} - \mathfrak{A}_{Q,P}\right)f(r) + \left(\mathfrak{A}_{Q,P} - r1_{\mathcal{B}_2(H)}\right)f(R)}{R-r},$$

which can be written as

$$\begin{aligned} &\langle f(\mathfrak{A}_{Q,P})T,T\rangle_2 \\ &\le \frac{f(r)}{R-r}\left\langle \left(R1_{\mathcal{B}_2(H)} - \mathfrak{A}_{Q,P}\right)T,T\right\rangle_2 + \frac{f(R)}{R-r}\left\langle \left(\mathfrak{A}_{Q,P} - r1_{\mathcal{B}_2(H)}\right)T,T\right\rangle_2 \end{aligned} \tag{4.2}$$

for any $T \in \mathcal{B}_2(H)$.

This inequality is of interest in itself.

Now, if we take in (4.2) $T = P^{1/2}$, $P \in S_1(H)$, then we get the desired result (4.2). □

Remark 4. *If we take in (4.1) $f(t) = t^2 - 1$, then we get*

$$0 \le \chi^2(Q,P) \le (R-1)(1-r)\frac{R+r+2}{R-r} \tag{4.3}$$

for $Q, P \in S_1(H)$, with P invertible and satisfying the condition (3.14).

If we take in (4.1) $f(t) = t\ln t$, then we get the inequality

$$0 \le U(Q,P) \le \ln\left[r^{\frac{(R-1)r}{R-r}}R^{\frac{R(1-r)}{R-r}}\right] \tag{4.4}$$

provided that $Q, P \in S_1(H)$, with P, Q invertible and satisfying the condition (3.14).

If we take in (4.1) $f(t) = -\ln t$, then we get the inequality

$$0 \le U(P,Q) \le \ln\left[r^{\frac{1-R}{R-r}}R^{\frac{r-1}{R-r}}\right] \tag{4.5}$$

for $Q, P \in S_1(H)$, with P, Q invertible and satisfying the condition (3.14).

We also have:

Theorem 10. *Let $f : [0,\infty) \to \mathbb{R}$ be a continuous convex function that is normalized. If $Q, P \in S_1(H)$, with P invertible, and there exists $R > 1 > r \ge 0$ such that the condition (3.14) is satisfied, then*

$$\begin{aligned} 0 \le S_f(Q,P) &\le \frac{(R-1)(1-r)}{R-r}\Psi_f(1;r,R) \\ &\le \frac{(R-1)(1-r)}{R-r}\sup_{t\in(r,R)}\Psi_f(t;r,R) \\ &\le (R-1)(1-r)\frac{f'_-(R)-f'_+(r)}{R-r} \\ &\le \frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right] \end{aligned} \tag{4.6}$$

where $\Psi_f(\cdot;r,R) : (r,R) \to \mathbb{R}$ is defined by

$$\Psi_f(t;r,R) = \frac{f(R)-f(t)}{R-t} - \frac{f(t)-f(r)}{t-r}. \tag{4.7}$$

We also have

$$\begin{aligned} 0 \le S_f(Q,P) &\le \frac{(R-1)(1-r)}{R-r}\Psi_f(1;r,R) \\ &\le \frac{1}{4}(R-r)\Psi_f(1;r,R) \\ &\le \frac{1}{4}(R-r)\sup_{t\in(r,R)}\Psi_f(t;r,R) \\ &\le \frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right]. \end{aligned} \tag{4.8}$$

Proof. By denoting

$$\Delta_f(t;r,R) := \frac{(t-r)f(R)+(R-t)f(r)}{R-r} - f(t), \quad t \in [r,R]$$

we have

$$\begin{aligned}(4.9)\qquad \Delta_f(t;r,R) &= \frac{(t-r)f(R)+(R-t)f(r)-(R-r)f(t)}{R-r}\\ &= \frac{(t-r)f(R)+(R-t)f(r)-(T-t+t-r)f(t)}{R-r}\\ &= \frac{(t-r)[f(R)-f(t)]-(R-t)[f(t)-f(r)]}{M-m}\\ &= \frac{(R-t)(t-r)}{R-r}\Psi_f(t;r,R)\end{aligned}$$

for any $t\in(r,R)$.

From the proof of Theorem 9 we have

$$\begin{aligned}(4.10)\qquad &\langle f(\mathfrak{A}_{Q,P})T,T\rangle_2\\ &\le \frac{f(r)}{R-r}\left\langle \left(R1_{\mathcal{B}_2(H)}-\mathfrak{A}_{Q,P}\right)T,T\right\rangle_2+\frac{f(R)}{R-r}\left\langle\left(\mathfrak{A}_{Q,P}-r1_{\mathcal{B}_2(H)}\right)T,T\right\rangle_2\\ &= \frac{\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2-r\right)f(R)+\left(R-\langle\mathfrak{A}_{Q,P}T,T\rangle_2\right)f(r)}{R-r}\end{aligned}$$

for any $T\in\mathcal{B}_2(H)$, $\|T\|_2=1$.

This implies that

$$\begin{aligned}(4.11)\quad 0&\le \langle f(\mathfrak{A}_{Q,P})T,T\rangle_2-f\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2\right)\\ &\le \frac{\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2-r\right)f(R)+\left(R-\langle\mathfrak{A}_{Q,P}T,T\rangle_2\right)f(r)}{R-r}-f\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2\right)\\ &=\Delta_f\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2;r,R\right)\\ &=\frac{\left(R-\langle\mathfrak{A}_{Q,P}T,T\rangle_2\right)\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2-r\right)}{R-r}\Psi_f\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2;r,R\right)\end{aligned}$$

for any $T\in\mathcal{B}_2(H)$, $\|T\|_2=1$.

Since

$$\begin{aligned}(4.12)\quad \Psi_f\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2;r,R\right) &\le \sup_{t\in(r,R)}\Psi_f(t;r,R)\\ &=\sup_{t\in(r,R)}\left[\frac{f(R)-f(t)}{R-t}-\frac{f(t)-f(r)}{t-r}\right]\\ &\le \sup_{t\in(r,R)}\left[\frac{f(R)-f(t)}{R-t}\right]+\sup_{t\in(r,R)}\left[-\frac{f(t)-f(r)}{t-r}\right]\\ &=\sup_{t\in(r,R)}\left[\frac{f(R)-f(t)}{R-t}\right]-\inf_{t\in(r,R)}\left[\frac{f(t)-f(r)}{t-r}\right]\\ &=f'_-(R)-f'_+(r),\end{aligned}$$

and, obviously

$$(4.13)\qquad \frac{1}{R-r}\left(R-\langle\mathfrak{A}_{Q,P}T,T\rangle_2\right)\left(\langle\mathfrak{A}_{Q,P}T,T\rangle_2-r\right)\le\frac{1}{4}(R-r),$$

then by (4.11)-(4.13) we have

$$
\begin{aligned}
0 &\le \langle f(\mathfrak{A}_{Q,P})T,T\rangle_2 - f\left(\langle \mathfrak{A}_{Q,P}T,T\rangle_2\right) \\
&\le \frac{\left(R-\langle \mathfrak{A}_{Q,P}T,T\rangle_2\right)\left(\langle \mathfrak{A}_{Q,P}T,T\rangle_2 - r\right)}{R-r}\Psi_f\left(\langle \mathfrak{A}_{Q,P}T,T\rangle_2; r,R\right) \\
&\le \frac{\left(R-\langle \mathfrak{A}_{Q,P}T,T\rangle_2\right)\left(\langle \mathfrak{A}_{Q,P}T,T\rangle_2 - r\right)}{R-r}\sup_{t\in(r,R)}\Psi_f(t;r,R) \\
&\le \left(R-\langle \mathfrak{A}_{Q,P}T,T\rangle_2\right)\left(\langle \mathfrak{A}_{Q,P}T,T\rangle_2 - r\right)\frac{f'_-(R)-f'_+(r)}{R-r} \\
&\le \frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right]
\end{aligned} \tag{4.14}
$$

for any $T\in\mathcal{B}_2(H)$, $\|T\|_2=1$.

This inequality is of interest in itself.

Now, if we take in (4.14) $T=P^{1/2}$, then we get the desired result (4.6).

The inequality (4.8) is obvious from (4.6). □

Remark 5. *If we consider the convex normalized function $f(t)=t^2-1$, then*

$$
\Psi_f(t;r,R)=\frac{R^2-t^2}{R-t}-\frac{t^2-r^2}{t-r}=R-r,\ \ t\in(r,R)
$$

and we get from (4.6) the simple inequality

$$
0\le \chi^2(Q,P)\le (R-1)(1-r) \tag{4.15}
$$

for $Q,P\in S_1(H)$, with P invertible and satisfying the condition (3.14), which is better than (4.3).

If we take the convex normalized function $f(t)=t^{-1}-1$, then we have

$$
\Psi_f(t;r,R)=\frac{R^{-1}-t^{-1}}{R-t}-\frac{t^{-1}-r^{-1}}{t-r}=\frac{R-r}{rRt},\ \ t\in[r,R].
$$

Also

$$
S_f(Q,P)=\chi^2(P,Q).
$$

Using (4.6) we get

$$
0\le \chi^2(P,Q)\le \frac{(R-1)(1-r)}{Rr} \tag{4.16}
$$

for $Q,P\in S_1(H)$, with Q invertible and satisfying the condition (3.14).

If we consider the convex function $f(t)=-\ln t$ defined on $[r,R]\subset(0,\infty)$, then

$$
\begin{aligned}
\Psi_f(t;r,R) &= \frac{-\ln R+\ln t}{R-t}-\frac{-\ln t+\ln r}{t-r} \\
&= \frac{(R-r)\ln t-(R-t)\ln r-(t-r)\ln R}{(M-t)(t-m)} \\
&= \ln\left(\frac{t^{R-r}}{r^{R-t}M^{t-r}}\right)^{\frac{1}{(R-t)(t-r)}},\ \ t\in(r,R).
\end{aligned}
$$

Then by (4.6) we have

$$
0\le U(P,Q)\le \ln\left[r^{\frac{1-R}{R-r}}R^{\frac{r-1}{R-r}}\right]\le \frac{(R-1)(1-r)}{rR} \tag{4.17}
$$

for $Q,P\in S_1(H)$, with P, Q invertible and satisfying the condition (3.14).

If we consider the convex function $f(t) = t\ln t$ *defined on* $[r,R] \subset (0,\infty)$, *then*

$$\Psi_f(t;r,R) = \frac{R\ln R - t\ln t}{R-t} - \frac{t\ln t - r\ln r}{t-r},\ t \in (r,R),$$

which gives that

$$\Psi_f(1;r,R) = \frac{R\ln R}{R-1} - \frac{r\ln r}{1-r}.$$

Using (4.6) we get

$$\begin{aligned} 0 &\le U(Q,P) \le \ln\left[R^{\frac{(1-r)R}{R-r}} r^{\frac{(1-R)r}{R-r}}\right] \\ &\le (R-1)(1-r)\ln\left[\left(\frac{R}{r}\right)^{\frac{1}{R-r}}\right] \end{aligned} \tag{4.18}$$

for $Q,P \in S_1(H)$, *with* P, Q *invertible and satisfying the condition (3.14).*

We also have:

Theorem 11. *Let* $f:[0,\infty)\to\mathbb{R}$ *be a continuous convex function that is normalized. If* $Q,P\in S_1(H)$, *with* P *invertible, and there exists* $R>1>r\ge 0$ *such that the condition (3.14) is satisfied, then*

$$0 \le S_f(Q,P) \le 2\left[\frac{f(r)+f(R)}{2} - f\left(\frac{r+R}{2}\right)\right]. \tag{4.19}$$

Proof. We recall the following result (see for instance [11]) that provides a refinement and a reverse for the weighted Jensen's discrete inequality:

$$\begin{aligned} & n\min_{i\in\{1,\dots,n\}}\{p_i\}\left[\frac{1}{n}\sum_{i=1}^n f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^n x_i\right)\right] \\ & \le \frac{1}{P_n}\sum_{i=1}^n p_i f(x_i) - f\left(\frac{1}{P_n}\sum_{i=1}^n p_i x_i\right) \\ & \le n\max_{i\in\{1,\dots,n\}}\{p_i\}\left[\frac{1}{n}\sum_{i=1}^n f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^n x_i\right)\right], \end{aligned} \tag{4.20}$$

where $f:C\to\mathbb{R}$ is a convex function defined on the convex subset C of the linear space X, $\{x_i\}_{i\in\{1,\dots,n\}}\subset C$ are vectors and $\{p_i\}_{i\in\{1,\dots,n\}}$ are nonnegative numbers with $P_n := \sum_{i=1}^n p_i > 0$.

For $n=2$ we deduce from (3.6) that

$$\begin{aligned} & 2\min\{s,1-s\}\left[\frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right)\right] \\ & \le sf(x) + (1-s)f(y) - f(sx+(1-s)y) \\ & \le 2\max\{s,1-s\}\left[\frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right)\right] \end{aligned} \tag{4.21}$$

for any $x,y\in C$ and $s\in[0,1]$.

Now, if we use the second inequality in (4.21) for $x = r$, $y = R$, $s = \frac{R-t}{R-r}$ with $t \in [r, R]$, then we have

$$\begin{aligned} &\frac{(R-t) f(r) + (t-r) f(R)}{R-r} - f(t) \\ &\leq 2 \max\left\{\frac{R-t}{R-r}, \frac{t-r}{R-r}\right\} \left[\frac{f(r) + f(R)}{2} - f\left(\frac{r+R}{2}\right)\right] \\ &= \left[1 + \frac{2}{R-r}\left|t - \frac{r+R}{2}\right|\right] \left[\frac{f(r) + f(R)}{2} - f\left(\frac{r+R}{2}\right)\right] \end{aligned} \tag{4.22}$$

for any $t \in [r, R]$.

This implies in the operator order of $\mathcal{B}(\mathcal{B}_2(H))$

$$\begin{aligned} &\frac{\left(R 1_{\mathcal{B}_2(H)} - \mathfrak{A}_{Q,P}\right) f(r) + \left(\mathfrak{A}_{Q,P} - r 1_{\mathcal{B}_2(H)}\right) f(R)}{R-r} - f(\mathfrak{A}_{Q,P}) \\ &\leq \left[\frac{f(r) + f(R)}{2} - f\left(\frac{r+R}{2}\right)\right] \\ &\times \left[1_{\mathcal{B}_2(H)} + \frac{2}{R-r}\left|\mathfrak{A}_{Q,P} - \frac{r+R}{2} 1_{\mathcal{B}_2(H)}\right|\right] \end{aligned}$$

which implies that

$$\begin{aligned} 0 &\leq \langle f(\mathfrak{A}_{Q,P}) T, T\rangle_2 - f\left(\langle \mathfrak{A}_{Q,P} T, T\rangle_2\right) \\ &\leq \frac{\left(\langle \mathfrak{A}_{Q,P} T, T\rangle_2 - r\right) f(R) + \left(R - \langle \mathfrak{A}_{Q,P} T, T\rangle_2\right) f(r)}{R-r} - f\left(\langle \mathfrak{A}_{Q,P} T, T\rangle_2\right) \\ &\leq \left[\frac{f(r) + f(R)}{2} - f\left(\frac{r+R}{2}\right)\right] \\ &\times \left[1 + \frac{2}{R-r}\left\langle \left|\mathfrak{A}_{Q,P} - \frac{r+R}{2} 1_{\mathcal{B}_2(H)}\right| T, T\right\rangle_2\right] \\ &\leq 2\left[\frac{f(r) + f(R)}{2} - f\left(\frac{r+R}{2}\right)\right] \end{aligned} \tag{4.23}$$

for any $T \in \mathcal{B}_2(H)$, $\|T\|_2 = 1$.

This is an inequality of interest in itself.

If we take in (4.23) $T = P^{1/2}$, $P \in S_1(H)$, then we get the desired result (4.19). □

Remark 6. *If we take $f(t) = t^2 - 1$ in (4.19), then we get*

$$0 \leq \chi^2(Q, P) \leq \frac{1}{2}(R-r)^2$$

for $Q, P \in S_1(H)$, with P invertible and satisfying the condition (3.14), which is not as good as (4.15).

If we take in (4.19) $f(t) = t^{-1} - 1$, then we have

$$0 \leq \chi^2(P, Q) \leq \frac{(R-r)^2}{rR(r+R)} \tag{4.24}$$

for $Q, P \in S_1(H)$, with P invertible and satisfying the condition (3.14).

If we take in (4.19) $f(t) = -\ln t$, *then we have*

$$0 \leq U(P,Q) \leq \ln\left(\frac{(R+r)^2}{4rR}\right) \tag{4.25}$$

for $Q, P \in S_1(H)$, *with* P *invertible and satisfying the condition (3.14).*

From (3.18) we have the following absolute upper bound

$$0 \leq U(P,Q) \leq \frac{(R-r)^2}{4rR} \tag{4.26}$$

for $Q, P \in S_1(H)$, *with* P *invertible and satisfying the condition (3.14).*

Utilising the elementary inequality $\ln x \leq x - 1$, $x > 0$, *we have that*

$$\ln\left(\frac{(R+r)^2}{4rR}\right) \leq \frac{(R-r)^2}{4rR},$$

which shows that (4.25) is better than (4.26).

References

[1] T. Ando, Matrix Young inequalities, *Oper. Theory Adv. Appl.* **75** (1995), 33–38.

[2] G. de Barra, *Measure Theory and Integration*, Ellis Horwood Ltd., 1981.

[3] R. Bellman, Some inequalities for positive definite matrices, in: E.F. Beckenbach (Ed.), *General Inequalities 2*, Proceedings of the 2nd International Conference on General Inequalities, Birkhäuser, Basel, 1980, pp. 89–90.

[4] E. V. Belmega, M. Jungers and S. Lasaulce, A generalization of a trace inequality for positive definite matrices. *Aust. J. Math. Anal. Appl.* **7** (2010), no. 2, Art. 26, 5 pp.

[5] P. Cerone and S. S. Dragomir, Approximation of the integral mean divergence and f-divergence via mean results. *Math. Comput. Modelling* **42** (2005), no. 1-2, 207–219.

[6] P. Cerone, S. S. Dragomir and F. Österreicher, Bounds on extended f-divergences for a variety of classes, Preprint, *RGMIA Res. Rep. Coll.* **6**(2003), No.1, Article 5. [ONLINE: http://rgmia.vu.edu.au/v6n1.html]. *Kybernetika* (Prague) **40** (2004), no. 6, 745–756.

[7] D. Chang, A matrix trace inequality for products of Hermitian matrices, *J. Math. Anal. Appl.* **237** (1999) 721–725.

[8] L. Chen and C. Wong, Inequalities for singular values and traces, *Linear Algebra Appl.* **171** (1992), 109–120.

[9] I. D. Coop, On matrix trace inequalities and related topics for products of Hermitian matrix, *J. Math. Anal. Appl.* **188** (1994) 999–1001.

[10] I. Csiszár, Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. (German) *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8** (1963) 85–108.

[11] S. S. Dragomir, Bounds for the normalized Jensen functional, *Bull. Austral. Math. Soc.* **74**(3)(2006), 471-476.

[12] S. S. Dragomir, Some inequalities for (m, M)-convex mappings and applications for the Csiszár Φ-divergence in information theory. *Math. J. Ibaraki Univ.* **33** (2001), 35–50.

[13] S. S. Dragomir, Some inequalities for two Csiszár divergences and applications. *Mat. Bilten* No. **25** (2001), 73–90.

[14] S. S. Dragomir, An upper bound for the Csiszár f-divergence in terms of the variational distance and applications. *Panamer. Math. J.* **12** (2002), no. 4, 43–54.

[15] S. S. Dragomir, Upper and lower bounds for Csiszár f-divergence in terms of Hellinger discrimination and applications. *Nonlinear Anal. Forum* **7** (2002), no. 1, 1–13

[16] S. S. Dragomir, Bounds for f-divergences under likelihood ratio constraints. *Appl. Math.* **48** (2003), no. 3, 205–223.

[17] S. S. Dragomir, New inequalities for Csiszár divergence and applications. *Acta Math. Vietnam.* **28** (2003), no. 2, 123–134.

[18] S. S. Dragomir, A generalized f-divergence for probability vectors and applications. *Panamer. Math. J.* **13** (2003), no. 4, 61–69.

[19] S. S. Dragomir, Some inequalities for the Csiszár φ-divergence when φ is an L-Lipschitzian function and applications. *Ital. J. Pure Appl. Math.* No. **15** (2004), 57–76.
[20] S. S. Dragomir, A converse inequality for the Csiszár Φ-divergence. *Tamsui Oxf. J. Math. Sci.* **20** (2004), no. 1, 35–53.
[21] S. S. Dragomir, Some general divergence measures for probability distributions. *Acta Math. Hungar.* **109** (2005), no. 4, 331–345.
[22] S. S. Dragomir, A refinement of Jensen's inequality with applications for f-divergence measures. *Taiwanese J. Math.* **14** (2010), no. 1, 153–164.
[23] S. S. Dragomir, A generalization of f-divergence measure to convex functions defined on linear spaces. *Commun. Math. Anal.* **15** (2013), no. 2, 1–14.
[24] S. Furuichi and M. Lin, Refinements of the trace inequality of Belmega, Lasaulce and Debbah. *Aust. J. Math. Anal. Appl.* **7** (2010), no. 2, Art. 23, 4 pp.
[25] F. Hiai, Fumio and D. Petz, From quasi-entropy to various quantum information quantities. *Publ. Res. Inst. Math. Sci.* **48** (2012), no. 3, 525–542.
[26] F. Hiai, M. Mosonyi, D. Petz and C. Bény, Quantum f-divergences and error correction. *Rev. Math. Phys.* **23** (2011), no. 7, 691–747.
[27] P. Kafka, F. Österreicher and I. Vincze, On powers of f-divergence defining a distance, *Studia Sci. Math. Hungar.,* **26** (1991), 415-422.
[28] H. D. Lee, On some matrix inequalities, *Korean J. Math.* **16** (2008), No. 4, pp. 565-571.
[29] F. Liese and I. Vajda, *Convex Statistical Distances,* Teubuer – Texte zur Mathematik, Band **95**, Leipzig, 1987.
[30] L. Liu, A trace class operator inequality, *J. Math. Anal. Appl.* 328 (2007) 1484–1486.
[31] S. Manjegani, Hölder and Young inequalities for the trace of operators, *Positivity* **11** (2007), 239–250.
[32] H. Neudecker, A matrix trace inequality, *J. Math. Anal. Appl.* **166** (1992) 302–303.
[33] F. Österreicher and I. Vajda, A new class of metric divergences on probability spaces and its applicability in statistics. *Ann. Inst. Statist. Math.* **55** (2003), no. 3, 639–653.
[34] D. Petz, From quasi-entropy. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **55** (2012), 81–92.
[35] D. Petz, From f-divergence to quantum quasi-entropies and their use. *Entropy* **12** (2010), no. 3, 304–325.
[36] M. B. Ruskai, Inequalities for traces on von Neumann algebras, *Commun. Math. Phys.* **26**(1972), 280—289.
[37] K. Shebrawi and H. Albadawi, Operator norm inequalities of Minkowski type, *J. Inequal. Pure Appl. Math.* **9**(1) (2008), 1–10, article 26.
[38] K. Shebrawi and H. Albadawi, Trace inequalities for matrices, *Bull. Aust. Math. Soc.* **87** (2013), 139–148.
[39] B. Simon, *Trace Ideals and Their Applications,* Cambridge University Press, Cambridge, 1979.
[40] Z. Ulukök and R. Türkmen, On some matrix trace inequalities. *J. Inequal. Appl.* **2010**, Art. ID 201486, 8 pp.
[41] X. Yang, A matrix trace inequality, *J. Math. Anal. Appl.* **250** (2000) 372–374.
[42] X. M. Yang, X. Q. Yang and K. L. Teo, A matrix trace inequality, *J. Math. Anal. Appl.* **263** (2001), 327–331.
[43] Y. Yang, A matrix trace inequality, *J. Math. Anal. Appl.* **133** (1988) 573–574.

[1]Mathematics, School of Engineering & Science, Victoria University, PO Box 14428, Melbourne City, MC 8001, Australia.

E-mail address: `sever.dragomir@vu.edu.au`

URL: `http://rgmia.org/dragomir`

[2]School of Computational & Applied Mathematics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa