数理解析研究所講究録
第 1778 巻 2012 年 49-69

Toward the final version of Lieb-Ando concavity

東北大学・情報科学研究科　日合　文雄 (Fumio Hiai)
Graduate School of Information Sciences, Tohoku University

Introduction

The Wigner-Yanase-Dyson (WYD) skew information is an old yet new subject having a somewhat complicated history since its appearance in the paper [38] in 1963. Of the most importance among many related things are the equivalent joint concavity (also joint convexity) results due to Lieb [31] (the so-called WYDL concavity) and due to Ando [1].

The aim of these notes is two-fold. The one is to survey the history of the WYD skew information and the WYDL concavity started with [38] up to recent related developments in quantum information. The other is to pursue what is the final version of the Lieb-Ando concavity. The subject treated here has been discussed for infinite-dimensional Hilbert space operators and even in the setting of von Neumann algebras as well. But in these notes we mostly confine ourselves to the finite-dimensional setting of matrices.

Sect. 1 is for the first aim, where, besides Lieb's and Ando's work on concavity/convexity, we stress Petz' quasi-entropy and monotone metrics and also recent Carlen and Lieb's work. For the second aim, in Sect. 2 we generalize the Lieb-Ando concavity/convexity as much as possible by using Epstein's method. In Sect. 3 we treat convexity or concavity properties of the generalized quasi-entropy, that is also considered as generalizing the Lieb-Ando concavity/convexity.

It is known in joint work with J-C. Bourin that some of the concavity or the convexity results for trace functions in Sect. 2 can be extended to those for symmetric norm functions or symmetric anti-norm functions. The notion of symmetric anti-norms has recently been studied in [7, 8]. The full version of Sect. 2 together with extensions to symmetric norms or anti-norms will be presented in [9]. On the other hand, Sect. 3 is based on joint work with D. Petz and its full version will be presented in [24].

We end this section by fixing some simple notations which will be used throughout. The notations $\mathbb{M}_n$, $\mathbb{M}_n^{sa}$, $\mathbb{M}_n^+$, and $\mathbb{P}_n$ stand for the sets of all $n\times n$ complex matrices, of $n\times n$ Hermitian matrices, of $n\times n$ positive semidefinite matrices, and of $n\times n$ positive definite (i.e., invertible positive semidefinite) matrices, respectively.

1 Brief history

The purpose of this section is to briefly summarize the history started with the WYD skew information up to recent developments related to quantum information and quantum information geometry.

1.1 Wigner-Yanase-Dyson skew information

In 1963, Wigner and Yanase [38] introduced the *Wigner-Yanase skew information*

$$I_\rho^{\mathrm{WY}}(K) := -\frac{1}{2}\mathrm{Tr}\,[\rho^{1/2}, K]^2$$

for a density matrix ρ and a Hermitian matrix K, where $[X,Y] := XY - YX$, the commutator of matrices X, Y. The skew information $I_\rho^{\mathrm{WY}}(K)$ represents the amount of information which a state described by ρ contains with respect to the conserved quantity K, also measuring the non-commutativity between ρ and K. Among basic properties of the skew information shown in [38] is the convexity of $\rho \mapsto I_\rho^{\mathrm{WY}}(K)$. As mentioned in [38], Dyson extended $I_\rho^{\mathrm{WY}}(K)$ to the *WYD skew information*

$$I_\rho^{\mathrm{WYD}}(p, K) := -\frac{1}{2}\mathrm{Tr}\,[\rho^p, K]\,[\rho^{1-p}, K],$$

where $0 < p < 1$.

1.2 Lieb's concavity/convexity

The convexity of $\rho \mapsto I_\rho^{\mathrm{WYD}}(K)$ had been an open problem until Lieb [31] proved that if $p, q \geq 0$ and $p + q \leq 1$ then the trace function

$$(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+ \longmapsto \mathrm{Tr}\,X^*A^pXB^q \tag{1.1}$$

is jointly concave for any $X \in \mathbb{M}_n$. Indeed, this settles the convexity problem of the WYD skew information since

$$I_\rho^{\mathrm{WYD}}(K) := \mathrm{Tr}\,\rho K^2 - \mathrm{Tr}\,K\rho^pK\rho^{1-p}.$$

Among others, it was also proved in [31] that if $p, q \leq 0$ and $p + q \geq -1$ then $\mathrm{Tr}\,X^*A^pXB^p$ is jointly convex in $(A, B, X) \in \mathbb{P}_n \times \mathbb{P}_n \times \mathbb{M}_n$, and if $-1 \leq p, q \leq 0$ then it is jointly convex in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n$ for any $X \in \mathbb{M}_n$. These concavity and convexity properties are often called the *WYDL concavity/convexity*.

Soon after Lieb's paper, Epstein [14] used a complex function method (in particular, the integral representation of Pick or Herglotz functions) to give alternative proofs of some results in [31] and also to prove that $A \in \mathbb{M}_n^+ \mapsto \mathrm{Tr}\,(X^*A^pX)^{1/p}$ is concave for any $X \in \mathbb{M}_n$ if $0 < p \leq 1$, which was conjectured in [31].

1.3 Ando's concavity/convexity

In 1979, Ando [1] considered the map of $A, B \in \mathbb{P}_n$ into the tensor product $A^p \otimes B^q$ of the powers A^p, B^q for $p, q \in \mathbb{R}$, and proved the following joint concavity/convexity assertions: $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto A^p \otimes B^q$ is jointly concave if and only if $p, q \geq 0$ and $p + q \leq 1$, and the same is jointly convex if and only if $-1 \leq p, q \leq 0$, or $-1 \leq p \leq 0$ and $1 - p \leq q \leq 2$, or $-1 \leq q \leq 0$ and $1 - q \leq p \leq 2$. Ando's concavity is given in the positive semidefiniteness order while Lieb's is for trace functions. But they are equivalent as seen in the following way: Note that $\mathbb{M}_n$ becomes a Hilbert space when equipped with the Hilbert-Schmidt inner product $\langle X, Y\rangle_{\mathrm{HS}} := \mathrm{Tr}\, X^*Y$. One can define a faithful $*$-representation π of $\mathbb{M}_n \otimes \mathbb{M}_n$ on $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ by $\pi(A\otimes B)X := AXB^t$ for $A, B, X \in \mathbb{M}_n$ so that

$$\langle X, \pi(A^p \otimes B^q)X\rangle_{\mathrm{HS}} = \mathrm{Tr}\, X^*A^pXB^q, \qquad A, B \in \mathbb{M}_n^+,\ X \in \mathbb{M}_n,$$

from which the equivalence between Lieb's and Ando's concavity/convexity clearly follows.

1.4 Araki's WYDL concavity in von Neumann algebras and Kosaki's extension by interpolation theory

Umegaki [37] introduced the relative entropy for normal states of a semifinite von Neumann algebra, which is a non-commutative extension of the *Kullback-Leibler divergence* useful in classical information theory. In the particular case of a matrix algebra, the *relative entropy* of two density matrices ρ, σ is defined to be

$$S(\rho\|\sigma) := \mathrm{Tr}\, \rho(\log\rho - \log\sigma). \tag{1.2}$$

Later, Araki [4] extended Umegaki's relative entropy to the case for normal states (or more generally, normal positive linear functionals) of a general von Neumann algebra by using the notion of relative modular operator as follows: for normal positive linear functionals ψ, φ of a von Neumann algebra M,

$$S(\psi\|\varphi) := -\langle \xi_\psi, (\log\Delta_{\varphi,\psi})\xi_\psi\rangle,$$

where ξ_ψ is the representing vector of ψ in the standard form of M and $\Delta_{\varphi,\psi}$ is the relative modular operator of φ relative to ψ. Basic properties of relative entropy such as lower semicontinuity, joint convexity, monotonicity, and martingale convergence were established in [4]. Araki also extended the WYDL concavity to the general von Neumann algebra setting, and indeed he proved the joint convexity of relative entropy by differentiating the WYDL concavity formula at $p = 0$. In particular, in the case of a matrix algebra this is plain: Differentiating (1.1) at $p = 0$ gives

$$\lim_{p\searrow 0} \frac{1}{p}\left(\mathrm{Tr}\, X^*\rho^pX\sigma^{1-p} - \mathrm{Tr}\, X^*X\sigma\right) = \mathrm{Tr}\left(X^*(\log\rho)X\sigma - X^*X\sigma\log\sigma\right),$$

which is equal to $-S(\sigma\|\rho)$ when $X = I$. Hence the joint convexity of (1.2) follows from Lieb's concavity.

In 1986, Kosaki [26] established a convenient variational expression of relative entropy, from which all basic properties of relative entropy were shown in a simple way. Formerly in 1982, Kosaki [25] related the WYDL concavity with interpolation theory. In fact, he established a very general form of the WYDL concavity in the framework of the K-method of Peetre in interpolation theory. Then he applied it to generalize the WYDL concavity in the general von Neumann algebra setting to the form involving an operator monotone function so that

$$\langle a\xi_\psi, f(\Delta_{\varphi,\psi})a\xi_\psi\rangle \tag{1.3}$$

is jointly concave in normal positive linear functionals ψ, φ for any $a \in M$ when f is a non-negative operator monotone function on $[0, \infty)$. This reduces to Araki's WYDL concavity mentioned above when $f(x) = x^p$, $p \in (0, 1)$.

1.5 Petz' quasi-entropy

In [33, 34] Petz introduced the *quasi-entropy* generalizing the relative entropy and the classical f-divergence, whose definition is given in the same form as (1.3). In the matrix algebra setting (as in [34]), for any real function f on $(0, \infty)$, it is defined as

$$S_f^X(\rho\|\sigma) := \langle X, f(\mathbf{L}_\rho \mathbf{R}_\sigma^{-1}) \mathbf{R}_\sigma X \rangle_{\mathrm{HS}} = \langle X\sigma^{1/2}, f(\mathbf{L}_\rho \mathbf{R}_\sigma^{-1})(X\sigma^{1/2}) \rangle_{\mathrm{HS}}$$

for $X \in \mathbb{M}_n$ and invertible density matrices ρ, σ, where $\mathbf{L}_\rho$ and $\mathbf{R}_\sigma$ are the left and the right multiplication operators of ρ, σ, respectively. (Note that $\mathbf{L}_\rho \mathbf{R}_\sigma^{-1}$ is the relative modular operator in the matrix algebra setting.) The quasi-entropy $S_f^X(\rho\|\sigma)$ is sometimes called the *quantum f-divergence* (in particular when $X = I$). When $f(x) = x \log x$ and $X = I$, one has $S_f^I(\rho\|\sigma) = S(\rho\|\sigma)$, the relative entropy. Petz first proved the monotonicity of quasi-entropy as follows: If f is operator monotone with $f(+0) \geq 0$ and $\alpha : \mathbb{M}_n \to \mathbb{M}_m$ is a unital Schwarz map (i.e., α is a linear map such that $\alpha(I_n) = I_m$ and $\alpha(X^*X) \geq \alpha(X^*)\alpha(X)$), then

$$S_f^X(\alpha^*(\rho)\|\alpha^*(\sigma)) \leq S_f^{\alpha(X)}(\rho\|\sigma)$$

for every invertible density matrices $\rho, \sigma \in \mathbb{M}_n$ and every $X \in \mathbb{M}_n$, where α^* is the dual map of α. Next, applying the monotonicity, he proved the joint convexity of quasi-entropy as follows: If f is operator convex, then $S_f^X(\rho\|\sigma)$ is jointly convex in invertible density matrices ρ, σ for any $X \in \mathbb{M}_n$. When f is operator monotone on $(0, \infty)$, $-f$ is operator convex and so $S_f^X(\rho\|\sigma)$ is jointly concave in ρ, σ. (See [6, 21] for operator monotone/convex functions and related matters.)

Alternative proofs for the joint convexity of quasi-entropy were recently presented by Hansen [17, 18] and by Effros [13]. In a recent paper [22] the most generalized results for monotonicity and convexity of quantum f-divergences are presented though restricted to the matrix algebra setting.

1.6 Petz' monotone metrics

The set $\mathcal{D}_n$ of invertible $n \times n$ density matrices forms a smooth Riemannian manifold with tangent space $\{A \in \mathbb{M}_n^{sa} : \operatorname{Tr} A = 0\}$. A Riemannian metric γ_ρ with foot points ρ (more precisely, a sequence of smooth Riemannian metrics γ_ρ on $\mathcal{D}_n$ for all $n \in \mathbb{N}$) is called a *monotone metric* if, for any completely positive and trace-preserving map $\beta : \mathbb{M}_n \to \mathbb{M}_m$, we have

$$\gamma_{\beta(\rho)}(\beta(A), \beta(A)) \leq \gamma_\rho(A, A), \qquad \rho \in \mathcal{D}_n,\ A \in \mathbb{M}_n^{sa},\ \operatorname{Tr} A = 0.$$

In the commutative case (or when restricted on the diagonal matrices) there exists only one monotone metric (up to a scalar factor), due to Chentsov, that is the so-called Fisher-Rao metric. The situation is quite different in the non-commutative case as shown in Petz' paper [35] of 1996. That is, the monotone metrics γ_ρ with normalization $\gamma_\rho(I,I) = \operatorname{Tr}\rho^{-1}$ correspond one-to-one to the operator monotone functions f on $(0,\infty)$ with normalization $f(1)=1$ in such a way that

$$\gamma_\rho^f(A,B) := \langle A, (f(\mathbf{L}_\rho\mathbf{R}_\rho^{-1})\mathbf{R}_\rho)^{-1}B\rangle_{\mathrm{HS}}$$

for $\rho \in \mathcal{D}_n$ and $A,B \in \mathbb{M}_n^{sa}$, $\operatorname{Tr}A = \operatorname{Tr}B = 0$. (A complete characterization of monotone metrics $\gamma_D(X,Y)$ extended to $D \in \mathbb{P}_n$ and $X,Y \in \mathbb{M}_n$ was recently obtained in [29].)

Petz and Hasegawa [36] observed that the WYD skew information can be written as

$$I_\rho^{\mathrm{WYD}}(p,K) = \frac{f_p(0)}{2}\,\gamma_\rho^{f_p}(i[\rho,K], i[\rho,K]) \tag{1.4}$$

(apart from a constant factor) by using an operator monotone function defined by

$$f_p(x) := p(1-p)\,\frac{(x-1)^2}{(x^p-1)(x^{1-p}-1)}.$$

It is easy to see that a monotone metric $\gamma_\rho(A,A)$ is jointly convex in (ρ,A), so the above expression (1.4) gives an alternative proof for the convexity of $\rho \mapsto I_\rho^{\mathrm{WYD}}(K)$. Petz' monotone metrics are often called *quantum Fisher informations.*

Recently, Hansen [18] generalized the WYD skew information to that associated with a symmetric (i.e., $f(x) = xf(x^{-1})$, $x > 0$) operator monotone function f on $[0,\infty)$ with $f(1)=1$ and $f(0)>0$ as follows:

$$I_\rho^f(K) := \frac{f(0)}{2}\,\gamma_\rho^f(i[\rho,K], i[\rho,K]) = \frac{f(0)}{2}\langle[\rho,K], (f(\mathbf{L}_\rho\mathbf{R}_\rho^{-1})\mathbf{R}_\rho)^{-1}([\rho,K])\rangle_{\mathrm{HS}} \tag{1.5}$$

for $\rho \in \mathcal{D}_n$ and $K \in \mathbb{M}_n^{sa}$, and he called it the *metric adjusted skew information* corresponding to f. It was proved in [18, 10] that $\rho \mapsto I_\rho^f(K)$ is convex for any K.

Another interesting observation due to Lesniewski and Ruskai [30] is that any monotone metric γ_ρ^f corresponding to f is obtained as the Hessian of a quantum g-divergence

$$\gamma_\rho^f(A,B) = \frac{\partial^2}{\partial t\partial s} S_g^I(\rho + tA\|\rho + sB)\Big|_{t=s=0},$$

where g is an operator convex function determined by the relation

$$\frac{1}{f(x)} = \frac{g(x) + xg(x^{-1})}{(x-1)^2}.$$

For example, when $f(x) = (x-1)/\log x$ and $g(x) = x\log x$, γ_ρ^f is the famous Bogoliubov (or Kubo-Mori) metric, S_g^I is the relative entropy, and it is well known that the Bogoliubov metric is the Hessian of the relative entropy.

1.7 Carlen-Lieb's work

In two papers [11, 12] of the same title, Carlen and Lieb developed convexity/concavity properties of some trace functions related to the WYDL concavity. In the first paper [11] of 1999 they proved that the Minkowski type trace function

$$(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+ \longmapsto \mathrm{Tr}\,(A^p + B^p)^{1/p} \tag{1.6}$$

is jointly concave if $0 < p \le 1$, jointly convex if $p = 2$, and not jointly convex (also not jointly concave) if $p > 2$. Indeed, the first assertion in case of $0 < p \le 1$ is immediate from Epstein's concavity (see the last paragraph of Sect. 1.2) since

$$\mathrm{Tr}\left(\begin{bmatrix} I & 0 \\ I & 0 \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^p \begin{bmatrix} I & 0 \\ I & 0 \end{bmatrix}\right)^{1/p} = \mathrm{Tr}\,(A^p + B^p)^{1/p}.$$

The last assertion (i.e., (1.6) is not jointly convex if $p > 2$) was also shown in [2].

In [20] we treated certain trace functions such as

$$(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+ \longmapsto \mathrm{Tr}\,(\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2})^s, \tag{1.7}$$

where $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ are positive linear maps. By using Epstein's method we proved the joint concavity of (1.7) when $p, q \ge 0$, $p + q = 1$ and $1 \le s \le 1/(p + q)$. In Sect. 2 we will discuss what are the final concavity/convexity assertions for such trace functions as (1.7).

Bekjan [5] treated joint concavity/convexity of trace functions complementing (1.6) and proved that when $0 < p \le 1$, $\mathrm{Tr}\,(A^{-p}+B^{-p})^{-1/p}$ is jointly concave in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n$, and $\mathrm{Tr}\,(A^{-p}+B^{-p})^{1/p}$ and $\mathrm{Tr}\,(A^p+B^p)^{-1/p}$ are jointly convex in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n$.

In the second paper [12] of 2008, Carlen and Lieb proved that the function

$$A \in \mathbb{M}_n^+ \longmapsto \mathrm{Tr}\,(X^*A^pX)^s \tag{1.8}$$

is, for any $X \in \mathbb{M}_n$, convex if $1 \le p \le 2$ and $s \ge 1/p$, concave if $0 \le p \le 1$ and $1 \le s \le 1/p$, and neither convex nor concave if $p > 2$. From this they further proved that the function

$$(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+ \longmapsto \|(A^p + B^p)^{1/p}\|_q = \{\mathrm{Tr}\,(A^p + B^p)^{q/p}\}^{1/q}$$

is jointly convex if $1 \le p \le 2$ and $q \ge 1$, jointly concave if $0 \le p \le q \le 1$, and neither convex nor concave if $p > 2$ and $0 < q \ne p$. By letting $q = 1$ this affirmatively settles the conjecture that $\mathrm{Tr}\,(A^p+B^p)^{1/p}$ is jointly convex in $(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+$ if $1 \le p \le 2$.

1.8 More relations to quantum information

In quantum information we need to treat many different quantities (certain kinds of distances) to distinguish given states of a quantum system (see e.g., [22]). Needless to say, the relative entropy is the most significant quantity to discriminate two states; it typically appears as the asymptotic error exponent in quantum Stein's lemma of quantum hypothesis testing. In other types of quantum hypothesis testing, the function

$\psi(s) := \log \operatorname{Tr} \rho^s \sigma^{1-s}$ of $s \in \mathbb{R}$ (particularly, $s \in [0,1]$) for density matrices ρ, σ shows up. Obviously, $\operatorname{Tr} \rho^s \sigma^{1-s}$ is a special case of (1.1) with $X = I$. Another similar quantity widely used in quantum information is the fidelity (due to Uhlmann) given for density matrices ρ, σ by

$$F(\rho, \sigma) := \operatorname{Tr} |\rho^{1/2}\sigma^{1/2}| = \operatorname{Tr} (\rho^{1/2}\sigma\rho^{1/2})^{1/2},$$

which is a particular case of (1.7). The monotonicity and the WYDL type convexity/concavity properties of such distance-like quantities seem essential in quantum information. We are tempted to generalize the fidelity to the function

$$F_{p,\alpha}(\rho, \sigma) := \operatorname{Tr} (\rho^{p\alpha}\sigma^{2p(1-\alpha)}\rho^{p\alpha})^{1/2p} \tag{1.9}$$

with two parameters $p > 0$ and $\alpha \in (0, 1)$, while we do not know if this generalization is meaningful in quantum information.

Moreover, the notion of monotone metrics (or quantum Fisher informations generalizing the WYD skew information, see Sect. 1.6) has been playing a crucial role in quantum information geometry. For example, it has successfully been used in the recently developed topic of inequalities related to uncertainty principle, first given in terms of the WYD skew information ([32, 27, 39]) and then generalized in terms of a general quantum Fisher information (e.g., [16, 15]).

2 The first approach with Epstein's method

In this section, following up the paper [20], we aim to generalize the Lieb-Ando concavity/convexity (see Sect. 1.2) as much as possible by using Epstein's method. Our first target is the joint concavity/convexity of the trace functions

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \operatorname{Tr} (\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2})^s, \tag{2.1}$$

$$A \in \mathbb{P}_n \longmapsto \operatorname{Tr} (\Phi(A^p))^s, \tag{2.2}$$

where p, q, s may be negative but $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ are strictly positive linear maps. Here, Φ being strictly positive means that Φ is positive and $\Phi(I_n)$ is in $\mathbb{P}_k$. Such restrictions to strictly positive ones are harmless since it is automatic to extend the result by continuity to general positive semidefinite A, B and to general positive linear maps Φ, Ψ as far as p, q, s are non-negative. Note that (2.1) reduces to (1.1) when $\Phi(\cdot) = X^* \cdot X$, $\Psi = \mathrm{id}$ and $s = 1$, and (2.2) reduces to (1.8) when $\Phi(\cdot) = X^* \cdot X$. So (2.1) and (2.2) seem to have a sufficiently general form in generalizing the Lieb-Ando concavity/convexity. Our final goal is to fix the necessary and sufficient condition on p, q, s for which (2.1) is jointly concave (or convex) for any Φ, Ψ, and also the condition on p, s for which (2.2) is concave (or convex) for any Φ.

In the following we always assume that $(p, q) \neq (0, 0)$ and $s \neq 0$ for (2.1) and that $p \neq 0$ and $s \neq 0$ for (2.2); otherwise (2.1) or (2.2) is constant. The next proposition is concerned with the necessity parts of the joint concavity of (2.1) and of the concavity of (2.2).

Proposition 2.1. (1) *Assume that $p, s \neq 0$. If $A \in \mathbb{P}_2 \mapsto \operatorname{Tr} (X^*A^pX)^s$ is concave for any invertible $X \in \mathbb{M}_2$, then either $0 < p \leq 1$ and $0 < s \leq 1/p$, or $-1 \leq p < 0$ and $1/p \leq s < 0$.*

(2) *Assume that* $(p,q) \neq (0,0)$ *and* $s \neq 0$. *If* $(A,B) \in \mathbb{P}_2 \times \mathbb{P}_2 \mapsto \mathrm{Tr}\,(A^{p/2}B^qA^{p/2})^s$ *is jointly concave, then either* $0 \le p,q \le 1$ *and* $0 < s \le 1/(p+q)$, *or* $-1 \le p,q \le 0$ *and* $1/(p+q) \le s < 0$.

Proof. (1) First assume that $s > 0$. By assumption, x^{ps} is concave in $x > 0$ so that $0 < ps \le 1$. For every $a,b,\varepsilon > 0$ let $A := \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and $X_\varepsilon := \begin{bmatrix} 1 & 0 \\ 1 & \varepsilon \end{bmatrix}$; then $\mathrm{Tr}\,(X_\varepsilon^* A^p X_\varepsilon)^s \to (a^p + b^p)^s$ as $\varepsilon \searrow 0$. So $(a^p+b^p)^s$ is concave in $a,b>0$. Since

$$\frac{d^2}{dx^2}(x^p+b)^s = psx^{p-2}(x^p+b)^{s-2}\{(ps-1)x^p + (p-1)b\},$$

we must have $(ps-1)x^p + (p-1)b \le 0$ for all $x,b>0$, which gives $p \le 1$. When $s<0$, the result follows from the above case since $\mathrm{Tr}\,(X^*A^pX)^s = \mathrm{Tr}\,(X^{-1}A^{-p}(X^{-1})^*)^{-s}$.

(2) As in the proof of (1) it suffices to assume that $s > 0$. By assumption, $x^{(p+q)s}$ is concave in $x > 0$ so that $(p+q)s \le 1$. The assumption also implies that $A \in \mathbb{P}_2 \mapsto \mathrm{Tr}\,(X^*A^pX)^s$ is concave for every invertible $X \in \mathbb{M}_2$, as readily seen by taking the polar decomposition of X. Hence (1) implies that $0 \le p \le 1$. Similarly, $0 \le q \le 1$. $\square$

For the sufficiency part of the joint concavity of (2.1) we extend the result in [20] (mentioned in Sect. 1.7) as follows:

Theorem 2.2. *If either* $0 \le p,q \le 1$ *and* $1/2 \le s \le 1/(p+q)$, *or* $-1 \le p,q \le 0$ *and* $1/(p+q) \le s \le -1/2$, *then the function* (2.1) *is jointly concave for every strictly positive linear maps* $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ *and* $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ *of every* n,m,k.

Sketch of proof. Let us prove the case where $0 \le p,q \le 1$ and $1/2 \le s \le 1/(p+q)$ (the proof is easily modified for the other case). The proof is a slight improvement of that of [20, Theorem 2.3], so we will only give a sketch. First, we note that the assertion in the case $1/2 \le s < 1/(p+q)$ follows from that in the case $s = 1/(p+q)$. Indeed, since $1/2 \le s < 1/(p+q)$, one can choose $p' \in [p,1]$ and $q' \in [q,1]$ such that $s = 1/(p'+q')$. (The assumption $s \ge 1/2$ is used only here.) Then, since x^α $(x \ge 0)$ with $0 < \alpha \le 1$ is operator concave as well as operator monotone, it is not difficult to show the assertion for p,q,s follows from that for p',q',s.

In the following we may assume that $0 < p,q \le 1$ and $s = 1/(p+q)$. Set $\gamma := p+q \in (0,2]$, so $s = 1/\gamma$. As in [20] we will use the following notations:

$$\mathbb{C}^+ := \{z \in \mathbb{C} : \mathrm{Im}\, z > 0\},$$

$$\mathcal{I}_n^+ := \{X \in \mathbb{M}_n : \mathrm{Im}\, X > 0\}, \quad \mathcal{I}_n^- := \{X \in \mathbb{M}_n : \mathrm{Im}\, X < 0\},$$

and

$$\Gamma_{\gamma\pi} := \{re^{i\theta} : r > 0,\ 0 < \theta < \gamma\pi\}.$$

Note that, for each $\alpha > 0$, the function x^α $(x > 0)$ has the analytic continuation z^α in $\mathbb{C} \setminus [0,\infty)$ (in particular, in $\mathbb{C}^+$) defined by

$$z^\alpha := r^\alpha e^{i\alpha\theta} \quad \text{for} \quad z = re^{i\theta} \ (r > 0,\ 0 < \theta < 2\pi).$$

To obtain the joint concavity result, it suffices to prove that if $A, H \in \mathbb{M}_n$ and $B, K \in \mathbb{M}_m$ are such that A, B are positive definite and H, K are Hermitian, then

$$\frac{d^2}{dx^2}\mathrm{Tr}\left\{\Phi((A+xH)^p)^{1/2}\Psi((B+xK)^q)\Phi((A+xH)^p)^{1/2}\right\}^s \le 0$$

for every sufficiently small $x > 0$. For $z \in \mathbb{C}$ set $X(z) := zA + H$ and $Y(z) := zB + K$. For any $z \in \mathbb{C}^+$, since $X(z) \in \mathcal{I}_n^+$, $Y(z) \in \mathcal{I}_m^+$ and $p, q \in (0, 1]$, one can define $X(z)^p$ and $Y(z)^q$ by analytic functional calculus by [20, Lemma 1.1]. Since [20, Lemma 1.2] implies that

$$\mathrm{Im}\,\Phi(X(z)^p) = \Phi(\mathrm{Im}\,X(z)^p) > 0, \quad \mathrm{Im}\,\Psi(Y(z)^q) = \Psi(\mathrm{Im}\,Y(z)^q) > 0,$$

one has $\Phi(X(z)^p), \Psi(Y(z)^q) \in \mathcal{I}_k^+$ and hence $\Phi(X(z)^p)^{1/2} \in \mathcal{I}_k^+$ is also well defined. Now, define

$$F(z) := \Phi(X(z)^p)^{1/2}\Psi(Y(z)^q)\Phi(X(z)^p)^{1/2}, \qquad z \in \mathbb{C}^+,$$

which is analytic in $\mathbb{C}^+$. Let $\sigma(F(z))$ denote the set of eigenvalues of $F(z)$. As in the proof of [20, Theorem 2.3] we can show the following properties:

(a) When $z = re^{i\theta}$ with a fixed $0 < \theta < \pi$, $\sigma(F(z)) \subset \Gamma_{\gamma\pi}$ for sufficiently large $r > 0$.

(b) $\sigma(F(z)) \cap [0, \infty) = \emptyset$ for all $z \in \mathbb{C}^+$.

(c) $\sigma(F(z)) \cap \{re^{i\gamma\pi} : r \ge 0\} = \emptyset$ for all $z \in \mathbb{C}^+$.

Furthermore, we obtain

$$F(z) := z^\gamma \Phi((A + z^{-1}H)^p)^{1/2}\Psi((B + z^{-1}K)^q)\Phi((A + z^{-1}H)^p)^{1/2}, \quad z \in \mathbb{C}^+. \tag{2.3}$$

The above properties (a)–(c) imply that

$$\sigma(F(z)) \subset \Gamma_{\gamma\pi} \quad \text{if} \quad z \in \mathbb{C}^+. \tag{2.4}$$

In fact, if (2.4) fails to hold for some $z_0 = r_0e^{i\theta_0} \in \mathbb{C}^+$, then according to (a) and the continuity of the eigenvalues of $F(z)$ we must have $\sigma(F(z)) \cup \partial\Gamma_{\gamma\pi} \neq \emptyset$ for some $z \in \{re^{\theta_0} : r > r_0\}$, which says that (b) or (c) must be violated. Hence (2.4) follows, so one can define $F(z)^s$ for $z \in \mathbb{C}^+$ by applying the analytic functional calculus by z^s on $\Gamma_{\gamma\pi}$ to $F(z)$. Since $\gamma s = 1$ by assumption, note that z^s maps $\Gamma_{\gamma\pi}$ into $\mathbb{C}^+$. Thus, $F(z)^s$ is an analytic function so that $\sigma(F(z)^s) \subset \mathbb{C}^+$ for all $z \in \mathbb{C}^+$ (see [20, Sect. 1]). In view of (2.3), we can choose an $R > 0$ so that $F(z)^s$ in $\mathbb{C}^+$ is continuously extended to $\mathbb{C}^+ \cup (R, \infty)$ with

$$F(x)^s = x\left\{\Phi((A + x^{-1}H)^p)^{1/2}\Psi((B + x^{-1}K)^q)\Phi((A + x^{-1}H)^p)^{1/2}\right\}^s, \quad x \in (R, \infty).$$

Since $\mathrm{Tr}\,(F(x)^s) \in \mathbb{R}$ for all $x \in (R, \infty)$, by the reflection principle we obtain a Pick function φ on $\mathbb{C} \setminus (-\infty, R]$ such that $\varphi(x) = \mathrm{Tr}\,(F(x)^s)$ for all $x \in (R, \infty)$. For every $x \in (0, R^{-1})$ we have

$$x\varphi(x^{-1}) = \mathrm{Tr}\left\{\Phi((A + xH)^p)^{1/2}\Psi((B + xK)^q)\Phi((A + xH)^p)^{1/2}\right\}^s.$$

It thus remains to show that

$$\frac{d^2}{dx^2}(x\varphi(x^{-1})) \le 0, \qquad x \in (0, R^{-1}). \tag{2.5}$$

According to Nevanlinna's theorem for Pick functions (see, e.g., [21, Theorem 2.6.2]), φ admits an integral expression

$$\varphi(z) = a + bz + \int_{-\infty}^{\infty} \frac{1+tz}{t-z}\, d\nu(t),$$

where $a \in \mathbb{R}$, $b \ge 0$, and ν is a finite measure on $\mathbb{R}$. Since φ is analytically continued across the interval (R, ∞), the measure ν is supported in $(-\infty, R]$. Therefore,

$$x\varphi(x^{-1}) = ax + b + \int_{-\infty}^{R} \frac{x(x+t)}{xt-1}\, d\nu(t), \qquad x \in (0, R^{-1}).$$

Compute

$$\frac{d}{dx}\left(\frac{x(x+t)}{xt-1}\right) = \frac{x^2t - 2x - t}{(xt-1)^2}, \qquad \frac{d^2}{dx^2}\left(\frac{x(x+t)}{xt-1}\right) = \frac{2(t^2+1)}{(xt-1)^3} \le 0$$

for all $x \in (0, R^{-1})$, and hence (2.5) follows. $\square$

Remark 2.3. The difference between a necessary condition in Proposition 2.1 (2) and a sufficient condition in Theorem 2.2 is rather small: $0 < s < 1/2$ for $0 \le p, q \le 1$, or $-1/2 < s < 0$ for $-1 \le p, q \le 0$. Although we cannot at the moment fix the joint concavity of (2.1) in these remaining cases, the following remark is worth noting: Assume that $0 < p, q \le 1$ and $0 < s \le 1$. For every positive linear maps $\Phi : \mathbb{M}_n \to \mathbb{M}_k$, $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ and for every $A_1, A_2 \in \mathbb{M}_n^+$, $B_1, B_2 \in \mathbb{M}_m^+$ one has

$$\begin{aligned}
&\left\{\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2} \Psi\left(\left(\frac{B_1+B_2}{2}\right)^q\right) \Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2}\right\}^s \\
&\quad \ge \left\{\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2} \left(\frac{\Psi(B_1^q)+\Psi(B_2^q)}{2}\right) \Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2}\right\}^s \\
&\quad \simeq \left\{\left(\frac{\Psi(B_1^q)+\Psi(B_2^q)}{2}\right)^{1/2} \Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right) \left(\frac{\Psi(B_1^q)+\Psi(B_2^q)}{2}\right)^{1/2}\right\}^s \\
&\quad \ge \left\{\left(\frac{\Psi(B_1^q)+\Psi(B_2^q)}{2}\right)^{1/2} \left(\frac{\Phi(A_1^p)+\Phi(A_2^p)}{2}\right) \left(\frac{\Psi(B_1^q)+\Psi(B_2^q)}{2}\right)^{1/2}\right\}^s \\
&\quad \simeq \left\{\left(\frac{\Phi(A_1^q)+\Phi(A_2^q)}{2}\right)^{1/2} \left(\frac{\Psi(B_1^p)+\Psi(B_2^p)}{2}\right) \left(\frac{\Phi(A_1^q)+\Phi(A_2^q)}{2}\right)^{1/2}\right\}^s,
\end{aligned}$$

where $\simeq$ means unitary conjugation. Hence, to settle the case $0 < s < 1/2$ (and $0 \le p, q \le 1$) of Theorem 2.2, we need to prove that $(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+ \mapsto \mathrm{Tr}\,(A^{1/2}BA^{1/2})^s$ is jointly concave if $0 < s < 1/2$.

Corollary 2.4. *Assume that $p > 0$ and $\alpha \in (0,1)$. The function (1.9) is jointly concave in $n \times n$ density matrices ρ, σ of every n if and only only if $p \le \min\{1/2\alpha, 1/2(1-\alpha)\}$.*

The sufficiency part of the corollary follows from Theorem 2.2. For the necessity part, it suffices to prove that if $p, q, s > 0$ and $\mathrm{Tr}\,(\rho^{p/2}\sigma^q\rho^{p/2})^s$ is jointly concave in 4×4 density matrices ρ, σ, then $p, q \le 1$ (and $s \le 1/(p+q)$). The proof is a modification of Carlen and Lieb's argument in [12]; the details are omitted here.

For the sufficiency part of the concavity of (2.2) we have

Theorem 2.5. *If either $0 < p \le 1$ and $0 < s \le 1/p$, or $-1 \le p < 0$ and $1/p \le s \le -1/2$, then the function (2.2) is concave for every strictly positive linear map $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ of every n, k.*

Proof. When $0 < p \le 1$ and $0 < s \le 1$, one has the matrix inequality

$$\Phi\left(\left(\frac{A+B}{2}\right)^p\right)^s \ge \left(\frac{\Phi(A^p)+\Phi(B^p)}{2}\right)^s \ge \frac{\Phi(A^p)^s+\Phi(B^p)^s}{2} \tag{2.6}$$

for every $A, B \in \mathbb{M}_n^+$. Next, assume that $0 < p \le 1$ and $1 \le s \le 1/p$. Then one can easily reduce the proof to the case $s = 1/p$. Then the result is a special case of [20, Theorem 2.1] (or follows from Theorem 2.2 by taking $\Psi = \mathrm{id}$ and $q = 0$). The case where $-1 \le p < 0$ and $1/p \le s \le -1/2$ also follows from Theorem 2.2 by taking Ψ and q as above. $\square$

Remark 2.6. The gap between a necessary condition in Proposition 2.1 (1) and a sufficient condition in Theorem 2.5 is the case $-1/2 < s < 0$ for $-1 \le p < 0$. Note that this case cannot be treated in a way similar to (2.6).

We now turn to the (joint) convexity property of (2.1) and (2.2). Concerning the necessity part we have

Proposition 2.7. (1) *Assume that $p, s \ne 0$. If $A \in \mathbb{P}_4 \mapsto \mathrm{Tr}\,(X^*A^pX)^s$ is convex for every $X \in \mathbb{P}_4$, then $-1 \le p < 0$ and $s > 0$, or $0 < p \le 1$ and $s < 0$, or $1 \le p \le 2$ and $s \ge 1/p$, or $-2 \le p \le -1$ and $s \le 1/p$.*

(2) *Assume that $(p,q) \ne (0,0)$ and $s \ne 0$. If $(A,B) \in \mathbb{P}_4 \times \mathbb{P}_4 \mapsto \mathrm{Tr}\,(A^{p/2}B^qA^{p/2})^s$ is jointly convex, then one has the following restrictions:*

- $-1 \le p, q \le 0$, $s > 0$,
- $0 \le p, q \le 1$, $s < 0$,
- $-1 \le p \le 0$, $1 \le q \le 2$, $p+q > 0$, $s \ge 1/(p+q)$,
- $0 \le p \le 1$, $-2 \le q \le -1$, $p+q < 0$, $s \le 1/(p+q)$,
- $1 \le p \le 2$, $-1 \le q \le 0$, $p+q > 0$, $s \ge 1/(p+q)$,
- $-2 \le p \le -1$, $0 \le q \le 1$, $p+q < 0$, $s \le 1/(p+q)$.

Proof. As in the proof of Proposition 2.1 it suffices to assume that $s > 0$.

(1) Let $X_\varepsilon := \begin{bmatrix} I & 0 \\ I & \varepsilon I \end{bmatrix} \in \mathbb{M}_4$ for $\varepsilon > 0$. For any $A, B \in \mathbb{P}_2$, since

$$\mathrm{Tr}\left(X_\varepsilon^* \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^p X_\varepsilon\right)^s \longrightarrow \mathrm{Tr}\,(A^p + B^p)^s \quad \text{as } \varepsilon \searrow 0,$$

the assumption implies that $(A, B) \in \mathbb{P}_2 \times \mathbb{P}_2 \mapsto \mathrm{Tr}\,(A^p + B^p)^s$ is jointly convex, so $\varphi_t(A) := \mathrm{Tr}\,(tA^p + B)^s$ is convex in $A \in \mathbb{P}_2$ for any $t > 0$ and $B \in \mathbb{P}_2$. Since

$$\frac{d}{dt}\varphi_t(A)\Big|_{t=0} = s\mathrm{Tr}\, B^{s-1}A^p,$$

we notice that

$$\varphi_t(A) = \mathrm{Tr}\, B^s + st\mathrm{Tr}\, B^{s-1}A^p + O(t^2) \quad \text{as } t \searrow 0.$$

Therefore, for $A_1, A_2 \in \mathbb{P}_2$ we have

$$\begin{aligned} 0 &\geq \varphi_t\left(\frac{A_1 + A_2}{2}\right) - \frac{\varphi_r(A_1) + \varphi_r(A_2)}{2} \\ &= st\left\{\mathrm{Tr}\, B^{s-1}\left(\frac{A_1 + A_2}{2}\right)^p - \mathrm{Tr}\, B^{s-1}\left(\frac{A_1^p + A_2^p}{2}\right)\right\} + O(t^2) \quad \text{as } t \searrow 0 \end{aligned}$$

so that

$$\mathrm{Tr}\, B^{s-1}\left(\frac{A_1 + A_2}{2}\right) \leq \mathrm{Tr}\, B^{s-1}\left(\frac{A_1^p + A_2^p}{2}\right).$$

When $s \neq 1$, this means that x^p $(x > 0)$ is matrix convex of order 2, which is also clear from the assumption itself when $s = 1$. Hence by [19, Proposition 3.1] we must have $-1 \leq p \leq 0$ or $1 \leq p \leq 2$. When $1 \leq p \leq 2$, $ps \geq 1$ since x^{ps} is convex.

(2) Since the assumption here implies that of (1), it follows that either $-1 \leq p < 0$, or $1 \leq p \leq 2$ and $s \geq 1/p$. Similarly, either $-1 \leq q < 0$, or $1 \leq q \leq 2$ and $s \geq 1/q$. Since $x^{ps}y^{qs}$ is jointly convex in $x, y > 0$, a simple computation yields

$$pq\{(p + q)s - 1\} \leq 0. \tag{2.7}$$

Hence the case where both $ps \geq 1$ and $qs \geq 1$ occur is impossible, so the following three cases are possible (when $s > 0$):

- $-1 \leq p, q \leq 0$,
- $-1 \leq p \leq 0$, $1 \leq q \leq 2$,
- $1 \leq p \leq 2$, $-1 \leq q \leq 0$.

For the second case, if $p = 0$ then $s \geq 1/q = 1/(p + q)$, and if $p < 0$ then (2.7) gives $p + q > 0$ and $s \geq 1/(p + q)$. The third case is similar. When $s < 0$, the additional three cases appear correspondingly. $\square$

The next theorem is concerned with the sufficiency part of the joint convexity of (2.1). We omit the proof that is essentially similar to that of Theorem 2.2. We should say that there is quite a big gap between a necessary condition in Proposition 2.7 (2) and a sufficient condition in the next theorem. It seems that Epstein's method cannot work to prove the joint convexity of (2.1) unless $p, q \in [0, 1]$ or $p, q \in [-1, 0]$.

Theorem 2.8. *If either $-1 \le p, q \le 0$ and $1/2 \le s \le -1/(p+q)$, or $0 \le p, q \le 1$ and $-1/(p+q) \le s \le -1/2$, then the function* (2.1) *is jointly convex for every strictly positive linear maps $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ of every n, m, k.*

Concerning the sufficiency part of the convexity of (2.2) we have

Theorem 2.9. *If $-1 \le p < 0$ and $s \ge 1/2$, or $0 < p \le 1$ and $s < 0$, or $1 \le p \le 2$ and $s \ge 1$, then the function* (2.2) *is convex for any strictly positive linear map $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ of every n, k.*

It is plain to check the conclusion of Theorem 2.9 if $-1 \le p < 0$ and $s \ge 1$, or $0 < p \le 1$ and $s < 0$, or $1 \le p \le 2$ and $s \ge 1$. The conclusion when $-1 \le p < 0$ and $1/2 \le s \le -1/p$ (≥ 1) follows by taking $\Psi = \mathrm{id}$ and $q = 0$ in the first case of Theorem 2.8. As mentioned in Sect. 1.7 it was proved in [12] that $A \in \mathbb{P}_n \mapsto \mathrm{Tr}\,(X^* A^p X)^s$ (a special form of (2.2)) is convex for any $X \in \mathbb{M}_n$ if $1 \le p \le 2$ and $s \ge 1/p$; hence so is this for any invertible $X \in \mathbb{M}_n$ also if $-2 \le p \le -1$ and $s \le 1/p$.

Our second target in this section is the joint concavity/convexity of the trace function

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \mathrm{Tr}\,(\Phi(A^p)\,\sigma\,\Phi(B^q))^s \tag{2.8}$$

involving an operator mean σ in the sense of Kubo and Ando [28], where p, q, s and Φ, Ψ are as in (2.1). Assume that $(p, q) \ne (0, 0)$ and $s \ne 0$ as before. The next theorem extends [20, Theorem 4.3]. The proof can be done by improving that in [20]; the details are omitted here.

Theorem 2.10. *Assume that σ is any operator mean and $0 \le p, q \le 1$. If $0 < s \le 1/\max\{p, q\}$, then the function* (2.8) *is jointly concave for every positive linear maps $\Phi : \mathbb{M}_n \to \mathbb{M}_k$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_k$ of every n, m, k. If $-1/\max\{p, q\} \le s < 0$, then* (2.8) *is jointly convex for every strictly positive Φ, Ψ as above.*

In particular, when $\Phi = \Psi = \mathrm{id}$, since

$$(A^p\,\sigma\,B^q)^s = (A^{-p}\,\sigma^*\,B^{-q})^{-s}$$

with the operator mean $A\,\sigma^*\,B := (A^{-1}\,\sigma\,B^{-1})^{-1}$, the adjoint of σ, we have

Corollary 2.11. *Assume that σ is any operator mean and $0 \le p, q \le 1$. If $0 < s \le 1/\max\{p, q\}$, then $\mathrm{Tr}\,(A^p\,\sigma\,B^q)^s$ and $\mathrm{Tr}\,(A^{-p}\,\sigma\,B^{-q})^{-s}$ are jointly concave in $A, B \in \mathbb{P}_n$, and $\mathrm{Tr}\,(A^p\,\sigma\,B^q)^{-s}$ and $\mathrm{Tr}\,(A^{-p}\,\sigma\,B^{-q})^s$ are jointly convex in $A, B \in \mathbb{P}_n$.*

Furthermore, when σ is the arithmetic mean, we have slight extensions of some results in [11] and [5] (see Sect. 1.7)

Corollary 2.12. *Assume that* $0 \le p, q \le 1$. *If* $0 < s \le 1/\max\{p, q\}$, *then* $\mathrm{Tr}\,(A^p + B^q)^s$ *and* $\mathrm{Tr}\,(A^{-p} + B^{-q})^{-s}$ *are jointly concave in* $A, B \in \mathbb{P}_n$, *and* $\mathrm{Tr}\,(A^p + B^q)^{-s}$ *and* $\mathrm{Tr}\,(A^{-p} + B^{-q})^s$ *are jointly convex in* $A, B \in \mathbb{P}_n$.

Assume that $p, q > 0$ and $s \ne 0$. If $(x^p + y^q)^s$ is concave in $x, y > 0$, then we must have $p, q \le 1$ and $0 < s \le 1/\max\{p, q\}$ as in the proof of Proposition 2.1 (1). So the joint concavity part of Theorem 2.10 is best possible.

3 The second approach with Petz' quasi-entropies

In this section we treat a one-parameter generalization of Petz' quasi-entropy and consider its convexity/concavity properties of several types. We then clarify the relation among those properties, thus generalizing the Lieb-Ando concavity/convexity.

Let f be a continuous positive function on $(0, \infty)$ and θ an arbitrary real number. For $A, B \in \mathbb{P}_n$ define an operator $\mathbf{J}^f_{A,B}$ on $\mathbb{M}_n$ by $\mathbf{J}^f_{A,B} := f(\mathbf{L}_A \mathbf{R}_B^{-1})\mathbf{R}_B$, that is,

$$\mathbf{J}^f_{A,B} X = f(\mathbf{L}_A \mathbf{R}_B^{-1})\mathbf{R}_B X = \sum_{i=1}^{k} \sum_{j=1}^{l} f(\alpha_i \beta_j^{-1}) \beta_j P_i X Q_j, \qquad X \in \mathbb{M}_n,$$

where $A = \sum_{i=1}^k \alpha_i P_i$ and $B = \sum_{j=1}^l \beta_j Q_j$ are the spectral decompositions of A and B. It is immediate to see that $\mathbf{J}^f_{A,B}$ is a positive linear operator on the Hilbert-Schmidt Hilbert space $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ for any $A, B \in \mathbb{P}_n$. In particular, $\mathbf{J}^f_{A,A}$ is denoted by $\mathbf{J}^f_A$ for short. We define a three-variable function $I^\theta_f(A, B, X)$ on $\mathbb{P}_n \times \mathbb{P}_n \times \mathbb{M}_n$ by

$$I^\theta_f(A, B, X) := \langle X, (\mathbf{J}^f_{A,B})^{-\theta} X\rangle_{\mathrm{HS}}, \qquad A, B \in \mathbb{P}_n,\ X \in \mathbb{M}_n. \tag{3.1}$$

With the spectral decompositions of A, B as above, $I^\theta_f(A, B, X)$ is more explicitly written as

$$I^\theta_f(A, B, X) = \sum_{i=1}^{k} \sum_{j=1}^{l} (f(\alpha_i \beta_j^{-1})\beta_j)^{-\theta} \mathrm{Tr}\, X^* P_i X Q_j.$$

When $\theta = 0$, $I^\theta_f(A, B, X)$ is reduced to the trivial function $\langle X, X\rangle_{\mathrm{HS}}$, so in this section we always assume that $\theta \ne 0$.

The function I^θ_f has been discussed by several authors in its special cases from different viewpoints. In particular, when $f(x) = x^\alpha$ $(x > 0)$ with $\alpha \in \mathbb{R}$, one has

$$I^\theta_f(A, B, X) = \mathrm{Tr}\, X^* A^{-\alpha\theta} X B^{-(1-\alpha)\theta}.$$

For any $p, q \in \mathbb{R}$ with $p + q \ne 0$ there are unique $\alpha, \theta \in \mathbb{R}$ such that $-\alpha\theta = p$ and $-(1-\alpha)\theta = q$, so the function I^θ_f covers the trace functions (1.1). The quasi-entropy $S^X_f(A, B)$ in Sect. 1.5 is written in the form of $I^\theta_f(A, B, X)$ with $\theta = -1$. When f is an operator monotone function on $[0, \infty)$ and $\theta = 1$,

$$I^\theta_f(A, A, X) = \langle X, (\mathbf{J}^f_A)^{-1} X\rangle_{\mathrm{HS}} = \mathrm{Tr}\, X^* ((\mathbf{J}^f_A)^{-1} X)$$

is the monotone metric on the manifold $\mathbb{P}_n$ (or rather restricted on $\mathcal{D}_n$) in Sect. 1.6. The minus sign of $-\theta$ in definition (3.1) is adjusted to the expression of monotone metrics. The Riemannian metrics on $\mathbb{P}_n$ recently discussed in [23] are written as $\langle H, (\mathbf{J}_A^f)^{-\theta}K\rangle$ for $A \in \mathbb{P}_n$ (foot points) and $H, K \in \mathbb{M}_n^{sa}$ (tangent vectors) when $M(x,y) := f(xy^{-1})y$ is a symmetric homogeneous mean. This metric is written in the form $I_f^\theta(A, A, H)$ when $K = H$. Furthermore, the WYD skew information and Hansen's generalization (or metric adjusted skew informations) are rephrased by $I_f^\theta(\rho, \rho, i[\rho, K])$ (up to a constant factor) for $\rho \in \mathcal{D}_n$ and $K \in \mathbb{M}_n^{sa}$ with $\theta = -1$ and an appropriate choice of f (see Sect. 1.6).

Our goal is to determine f and θ for which $I_f^\theta(A, B, X)$ is jointly convex or concave in $(A, B, X) \in \mathbb{P}_n \times \mathbb{P}_n \times \mathbb{M}_n$ or in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n$ for any $X \in \mathbb{M}_n$. For this purpose, given a continuous function f (always assumed to be positive, i.e., $f(x) > 0$ for $x > 0$) and $\theta \in \mathbb{R} \setminus \{0\}$, we consider the following convexity/concavity properties (where each property means that the condition holds for every n):

(i) $(A, B, X) \in \mathbb{P}_n \times \mathbb{P}_n \times \mathbb{M}_n \mapsto I_f^\theta(A, B, X)$ is jointly convex,

(ii) $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto \log I_f^\theta(A, B, X)$ is jointly convex for any $X \in \mathbb{M}_n$,

(iii) $\theta > 0$ and $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto I_f^\theta(A, B, X)$ is jointly convex for any $X \in \mathbb{M}_n$,

(iv) $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto I_f^\theta(A, B, X)$ is jointly convex for any $X \in \mathbb{M}_n$,

(v) $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto I_f^{-\theta}(A, B, X)$ is jointly concave for any $X \in \mathbb{M}_n$,

(vi) $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto \log I_f^{-\theta}(A, B, X)$ is jointly concave for any $X \in \mathbb{M}_n$.

For each of the above properties, we also consider the property reduced to $A = B$, that is,

(i′) $(A, X) \in \mathbb{P}_n \times \mathbb{M}_n \longmapsto I_f^\theta(A, A, X)$ is jointly convex,

(ii′) $A \in \mathbb{P}_n \longmapsto \log I_f^\theta(A, A, X)$ is convex for any $X \in \mathbb{M}_n$,

and similarly for (iii′)-(vi′).

Furthermore, we raise the the following intrinsic conditions for f and θ:

(vii) f is operator monotone on $(0, \infty)$ and $\theta \in (0, 1]$,

(viii) f is operator monotone on $(0, \infty)$ and $\theta \in (0, 2]$.

We then have

Theorem 3.1. *Concerning the above properties the following hold:*

(a) *Each of* (i)–(vi) *is equivalent to the corresponding condition with prime.*

(b) (vii) $\Leftrightarrow$ (i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii) $\Rightarrow$ (viii).

(c) (vii) $\Leftrightarrow$ (v) $\Rightarrow$ (vi) $\Rightarrow$ (viii).

(d) (iii) $\Rightarrow$ (iv) $\Rightarrow \theta \in [-2,-1] \cup (0,2]$.

It is remarkable that all the convexity/concavity conditions (i)–(vi) except (iv) sit between (vii) and (viii), and the difference between the last two is only the range ($(0,1]$ or $(0,2]$) of θ.

The proof of (a) is an easy application of the 2×2 block matrix trick. For each $A, B \in \mathbb{P}_n$ set

$$\tilde{A} := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathbb{P}_{2n}.$$

For any $\tilde{X} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \in \mathbb{M}_{2n}$ $(= \mathbb{M}_n \otimes \mathbb{M}_2)$, since

$$\mathbf{L}_{\tilde{A}}\tilde{X} = \tilde{A}\tilde{X} = \begin{bmatrix} AX_{11} & AX_{12} \\ BX_{21} & BX_{22} \end{bmatrix}, \quad \mathbf{R}_{\tilde{A}}\tilde{X} = \tilde{X}\tilde{A} = \begin{bmatrix} X_{11}A & X_{12}B \\ X_{21}A & X_{22}B \end{bmatrix},$$

one can write

$$\mathbf{L}_{\tilde{A}} = \mathbf{L}_A \oplus \mathbf{L}_A \oplus \mathbf{L}_B \oplus \mathbf{L}_B, \quad \mathbf{R}_{\tilde{A}} = \mathbf{R}_A \oplus \mathbf{R}_B \oplus \mathbf{R}_A \oplus \mathbf{R}_B$$

under the identification of the Hilbert-Schmidt Hilbert space $\mathbb{M}_{2n}$ with the direct sum $\mathbb{M}_n \oplus \mathbb{M}_n \oplus \mathbb{M}_n \oplus \mathbb{M}_n$ by the isomorphism $X \mapsto X_{11} \oplus X_{12} \oplus X_{21} \oplus X_{22}$. Hence we have

$$\mathbf{J}^f_{\tilde{A}} = \mathbf{J}^f_A \oplus \mathbf{J}^f_{A,B} \oplus \mathbf{J}^f_{B,A} \oplus \mathbf{J}^f_B$$

so that

$$I^\theta_f\left(\tilde{A}, \tilde{A}, \begin{bmatrix} 0 & X \\ 0 & 0 \end{bmatrix}\right) = \left\langle \begin{bmatrix} 0 & X \\ 0 & 0 \end{bmatrix}, (\mathbf{J}^f_{\tilde{A}})^{-\theta} \begin{bmatrix} 0 & X \\ 0 & 0 \end{bmatrix} \right\rangle_{\mathrm{HS}} = I^\theta_f(A, B, X),$$

from which each case in (a) follows immediately.

The results (b)–(d) are the main part of the theorem. Their proofs are rather long based on [3, Theorems 3.1 and 3.7], which are omitted here. But we present some necessary statements from [3] as a lemma below. Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$, and $B(\mathcal{H})^{++}$ be the set of all positive and invertible bounded operators on $\mathcal{H}$. Let f be a continuous positive function on $(0,\infty)$, and $f(A)$ be defined for $A \in B(\mathcal{H})^{++}$ via functional calculus as usual.

Lemma 3.2. *The following* (a1)–(a4) *are equivalent:*

(a1) *f is operator monotone decreasing on* $(0,\infty)$;

(a2) $(A, \xi) \in B(\mathcal{H})^{++} \times \mathcal{H} \mapsto \langle \xi, f(A)\xi \rangle$ *is jointly convex;*

(a3) $A \in B(\mathcal{H})^{++} \mapsto \log\langle \xi, f(A)\xi \rangle$ *is convex for every* $\xi \in \mathcal{H}$;

(a4) *f is operator convex on* $(0,\infty)$ *and the numerical function* $f(x)$ *is non-increasing on* $(0,\infty)$.

Also, the following (b1) *and* (b2) *are equivalent:*

(b1) *f is operator monotone (or equivalently, operator concave) on* $(0,\infty)$*;*

(b2) $A \in B(\mathcal{H})^{++} \mapsto \log\langle \xi, f(A)\xi\rangle$ *is concave for every* $\xi \in \mathcal{H}$.

Note that log-convexity is stronger than convexity for continuous positive functions while log-concavity is weaker than concavity. The log-convexity condition (a3) characterizes the operator monotone decreasingness of f that is a stronger version of operator convexity. On the other hand, the log-concavity condition (b2) is equivalent to the operator concavity of f.

Remark 3.3. Let $f(x) := \sqrt{x}$ $(x > 0)$. By [31, Corollary 8.1 (2)] the function

$$\log I_f^\theta(A,B,X) = \log \operatorname{Tr} X^* A^{-\theta/2} X B^{-\theta/2}$$

is jointly convex in $(A,B) \in \mathbb{P}_n \times \mathbb{P}_n$ for any $\theta \in (0,2]$. Hence (ii) ($\Leftrightarrow$ (iii)) does not imply (vii), and the restriction $\theta \in (0,2]$ from (ii) is best possible.

Remark 3.4. Let $f(x) := x^\alpha$ $(x > 0)$, where $\alpha \in \mathbb{R}$. Recall (see Sect. 1.3) that the function

$$I_f^\theta(A,B,X) = \operatorname{Tr} X^* A^{-\alpha\theta} X B^{-(1-\alpha)\theta}$$

is jointly convex in $(A,B) \in \mathbb{P}_n \times \mathbb{P}_n$ if and only if $(A,B) \in \mathbb{P}_n \times \mathbb{P}_n \mapsto A^{-\alpha\theta} \otimes B^{-(1-\alpha)\theta}$ is jointly convex. This joint convexity holds if and only if one of the following cases is satisfied ([1]):

- $0 \le \alpha \le 1$ and $0 < \theta \le \min\{\frac{1}{\alpha}, \frac{1}{1-\alpha}\}$,
- $1 \le \alpha \le 2$ and $-\min\{\frac{2}{\alpha}, \frac{1}{\alpha-1}\} \le \theta \le -1$,
- $-1 \le \alpha \le 0$ and $-\min\{\frac{2}{1-\alpha}, \frac{1}{-\alpha}\} \le \theta \le -1$,

where $1/0 := +\infty$ by convention. In particular, (iv) is satisfied when $\alpha = 1/2$ and $\theta \in (0,2]$ or when $\alpha = 1$ and $\theta \in [-2,-1]$. Hence the restriction $\theta \in [-2,-1] \cup (0,2]$ from (iv) is best possible. Also, note that (viii) does not imply (iv).

Next, we consider the *generalized skew information* defined as

$$J_A^f(X) := I_f^\theta(A,A,i[A,X]) = \langle i[A,X], (\mathbf{J}_A^f)^{-1}(i[A,X])\rangle_{\mathrm{HS}}. \quad A \in \mathbb{P}_n,\ X \in \mathbb{M}_n.$$

This is the same as (1.5) up to a constant factor that is irrelevant to our aim. As mentioned in Sect. 1.6, Hansen [18] (also [10]) showed that $J_\rho^f(K)$ is convex in a density matrix ρ for any $K \in \mathbb{M}_n^{sa}$ if f is a symmetric operator monotone function. The following two theorems are slight refinements of this, showing that the converse direction is also valid.

Theorem 3.5. *Define the harmonic symmetrization of f by*

$$f^{\mathrm{sym}}(x) := \left(\frac{f(x)^{-1} + \tilde f(x)^{-1}}{2}\right)^{-1} \quad \textit{where} \quad \tilde f(x) := f(x^{-1})x, \quad x > 0.$$

Then the following conditions are equivalent:

(1) *for every* n, $A \in \mathbb{P}_n \mapsto J_A^f(K)$ *is convex for any* $K \in \mathbb{M}_n^{sa}$;

(2) *the function* $(x-1)^2/f_{\mathrm{sym}}(x)$ *is operator convex on* $(0,\infty)$;

(3) f_{sym} *is operator monotone on* $(0,\infty)$.

Theorem 3.6. *The following conditions are equivalent:*

(1′) *for every* n, $A \in \mathbb{P}_n \mapsto J_A^f(X)$ *is convex for any* $X \in \mathbb{M}_n$;

(2′) *the function* $(x-1)^2/f(x)$ *is operator convex on* $(0,\infty)$;

(3′) f *is operator monotone on* $(0,\infty)$.

To prove the theorems, we define

$$h(x) := \frac{(x-1)^2}{f(x)}, \quad \tilde{h}(x) := \frac{(x-1)^2}{\tilde{f}(x)}, \quad h_{\mathrm{sym}}(x) = \frac{(x-1)^2}{f^{\mathrm{sym}}(x)}, \qquad x > 0.$$

It is clear that $\tilde{h}(x) = h(x^{-1})x$ and

$$h_{\mathrm{sym}}(x) := \frac{h(x) + \tilde{h}(x)}{2}, \qquad x > 0.$$

As mentioned in [18, 10] it is easy to see that, for every $A \in \mathbb{P}_n$ and $X \in \mathbb{M}_n$,

$$J_A^f(X) = \langle X, \mathbf{J}_A^h X\rangle_{\mathrm{HS}}, \quad J_A^{\tilde{f}}(X) = \langle X, \mathbf{J}_A^{\tilde{h}} X\rangle_{\mathrm{HS}}, \quad J_A^{f^{\mathrm{sym}}}(X) = \langle X, \mathbf{J}_A^{h_{\mathrm{sym}}} X\rangle_{\mathrm{HS}}. \quad (3.2)$$

Furthermore, we have

$$J_A^f(X) = J_A^{\tilde{f}}(X^*).$$

Now the proof can be done from the above equations together with the next lemma. Here we omit the proof of the lemma as well as the details of the proofs of the theorems.

Lemma 3.7. *For a positive function* f *on* $(0,\infty)$ *the following conditions are equivalent:*

(c1) f *is operator monotone on* $(0,\infty)$;

(c2) $(x-1)/f(x)$ *is operator monotone on* $(0,\infty)$;

(c3) $(x-1)f(x)$ *is operator convex on* $(0,\infty)$;

(c4) $(x-1)^2/f(x)$ *is operator convex on* $(0,\infty)$.

Remark 3.8. The equivalence between (c1) and (c4) is crucial; other conditions (c2) and (c3) are stated for the completeness of statements. In [10], (c1) $\Rightarrow$ (c4) was indeed proved and it was remarked that the converse is not true. For example, $f(x) := (x-1)^2$ clearly satisfies (c4) while it is not operator monotone. But this function is not strictly positive on $(0,\infty)$, which we excluded in the lemma.

Although it is obvious from (3.2) that the function $J_A^f(X)$ is convex in X, the joint convexity of $J_A^f(K)$ in (A,K) is impossible as follows:

Proposition 3.9. *For any continuous positive function* f *on* $(0,\infty)$, *the function* $I_A^f(K)$ *is not jointly convex in* $(A,K) \in \mathbb{P}_n \times \mathbb{M}_n^{sa}$ *for some* $n \in \mathbb{N}$.

References

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard Products, *Linear Algebra Appl.* **26** (1979), 203–241.

[2] T. Ando and F. Hiai, Hölder type inequalities for matrices, *Math. Ineq. Appl.* **1** (1998), 1–30.

[3] T. Ando and F. Hiai, Operator log-convex functions and operator means, *Math. Ann.* **350** (2011), 611–630.

[4] H. Araki, Relative entropy of states of von Neumann algebras I, II, *Publ. Res. Inst. Math. Sci.* **11** (1976), 809–833; **13** (1977), 173–192.

[5] T. N. Bekjan, On joint convexity of trace functions, *Linear Algebra Appl.* **390** (2004), 321–327.

[6] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1996.

[7] J.-C. Bourin and F. Hiai, Norm and anti-norm inequalities for positive semi-definite matrices, *Internat. J. Math.* **22** (2011), 1121–1138.

[8] J.-C. Bourin and F. Hiai, Jensen and Minkowski type inequalities for operator means, Preprint, 2011, arXiv:1106.2213.

[9] J.-C. Bourin and F. Hiai, in preparation.

[10] L. Cai and F. Hansen, Metric adjusted skew information: Convexity and restricted forms of superadditivity, *Lett. Math. Phys.* **93** (2010), 1–13.

[11] E. A. Carlen and E. H. Lieb, A Minkowski type trace inequality and strong subadditivity of quantum entropy, *Advances in the Mathematical Sciences, Amer. Math. Soc. Transl. Ser. 2* **189** (1999), 59–68.

[12] E. A. Carlen and E. H. Lieb, A Minkowski type trace inequality and strong subadditivity of quantum entropy II: Convexity and concavity, *Lett. Math. Phys.* **83** (2008), 107–126.

[13] E. G. Effros, A matrix convexity approach to some celebrated quantum inequalities, *Proc. Nat. Acad. Sci. USA* **106** (2009), 1006–1008.

[14] H. Epstein, Remarks on two theorems of E. Lieb, *Comm. Math. Phys.* **31** (1973), 317–325.

[15] P. Gibilisco, F. Hiai and D. Petz, Quantum covariance, quantum Fisher information and the uncertainty principle, *IEEE Trans. Inform. Theory* **55** (2009), 439–443.

[16] P. Gibilisco, D. Imparato and T. Isola, Uncertainty principle and quantum Fisher information. II, *J. Math. Phys.* **48** (2007), 072109.

[17] F. Hansen, Extensions of Lieb's concavity theorem, *J. Stat. Phys.* **124**, (2006), 87–101.

[18] F. Hansen, Metric adjusted skew information, *Proc. Nat. Acad. Sci. USA* **105** (2008), 9909–9916.

[19] F. Hansen and J. Tomiyama, Differential analysis of matrix convex functions II, *J. Inequal. Pure Appl. Math.* **10** (2009), Article 32, 5 pp.

[20] F. Hiai, Concavity of certain matrix trace functions, *Taiwanese J. Math.* **5** (2001), 535–554.

[21] F. Hiai, Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization (GSIS selected lectures), *Interdisciplinary Information Sciences* **16** (2010), 139–248.

[22] F. Hiai, M. Mosonyi, D. Petz and C. Bény, Quantum f-divergences and error correction, *Rev. Math. Phys.* **23** (2011), 691–747.

[23] F. Hiai and D. Petz, Riemannian metrics on positive definite matrices related to means, *Linear Algebra Appl.* **430** (2009), 3105–3130.

[24] F. Hiai and D. Petz, in preparation.

[25] H. Kosaki, Interpolation theory and the Wigner-Yanase-Dyson-Lieb concavity, *Comm. Math. Phys.* **87** (1982), 315–329.

[26] H. Kosaki, Relative entropy of states: a variational expression, *J. Operator Theory* **16** (1986), 335–348.

[27] H. Kosaki, Matrix trace inequality related to uncertainty principle, *Internat. J. Math.* **16** (2005), 629–645.

[28] F. Kubo and T. Ando, Means of positive linear operator, *Math. Ann.* **246** (1980), 205–224.

[29] W. Kumagai, A characterization of extended monotone metrics, *Linear Algebra Appl.* **434** (2011), 224–231.

[30] A. Lesniewski and M. B. Ruskai, Monotone Riemannian metrics and relative entropy on noncommutative probability spaces, *J. Math. Phys.* **40** (1999), 5702–5724.

[31] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Advances in Math.* **11** (1973), 267–288.

[32] S. Luo and Z. Zhang, An information characterization of Schrödinger's uncertainty principle, *J. Stat. Phys.* **114** (2004), 1557–1576.

[33] D. Petz, Quasi-entropies for states of a von Neumann algebra, *Publ. Res. Inst. Math. Sci.* **21** (1985), 787—800.

[34] D. Petz, Quasi-entropies for finite quantum systems, *Rep. Math. Phys.* **23** (1986), 57–65.

[35] D. Petz, Monotone metrics on matrix spaces, *Linear Algebra Appl.* **244** (1996), 81–96.

[36] D. Petz and H. Hasegawa, On the Riemannian metric of α-entropies of density matrices, *Lett. Math. Phys.* **38** (1996), 221–225.

[37] H. Umegaki, Conditional expectation in an operator algebra, IV (entropy and information), *Kōdai Math. Sem. Rep.* **14** (1962), 59–85.

[38] E. P. Wigner and M. M. Yanase, Information contents of distributions, *Proc. Nat. Acad. Sci. USA* **49** (1963), 910–918.

[39] K. Yanagi, S. Furuichi and K. Kuriyama, A generalized skew information and uncertainty relation, *IEEE Trans. Inform. Theory* **51** (2005), 4401–4404.