Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (2) (2022), 362 – 372
DOI : 10.15672/hujms.902879

Research Article

The matrix Heinz mean and related divergence

Trung Hoa Dinh[1], Anh Vu Le[2], Cong Trinh Le*[3], Ngoc Yen Phan[4]

[1] *Department of Mathematics and Statistics, Troy University, Troy, AL 36082, United States*
[2] *Department of Economic Mathematics, University of Economics and Law, Vietnam National University, Ho Chi Minh City, Vietnam*
[3] *Department of Mathematics and Statistics, Quy Nhon University, Vietnam*
[4] *Faculty of Fundamental Science, Van Lang University, Ho Chi Minh City, Vietnam*

Abstract

In this paper, we introduce a new quantum divergence

$$\Phi(X,Y) = \operatorname{Tr}\left[\left(\frac{1-\alpha}{\alpha} + \frac{\alpha}{1-\alpha}\right)X + 2Y - \frac{X^{1-\alpha}Y^{\alpha}}{\alpha} - \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right],$$

where $0 < \alpha < 1$. We study the least square problem with respect to this divergence. We also show that the new quantum divergence satisfies the Data Processing Inequality in quantum information theory. In addition, we show that the matrix p-power mean $\mu_p(t, A, B) = ((1-t)A^p + tB^p)^{1/p}$ satisfies the in-betweenness property with respect to the new divergence.

Mathematics Subject Classification (2020). 47A63, 47A56

Keywords. quantum divergence, Heinz mean, least squares problem, matrix power mean, in-betweenness property, data processing inequality

1. Introduction

Let $\mathbb{M}_n$ be the algebra of $n \times n$ matrices over $\mathbb{C}$ and let $\mathcal{P}_n$ denote the cone of positive definite matrices in $\mathbb{M}_n$. Denote by I the identity matrix of $\mathbb{M}_n$. For a real-valued function f and a Hermitian matrix $A \in \mathbb{M}_n$ the matrix $f(A)$ is understood by means of the functional calculus.

Definition 1.1 ([1,3])**.** A smooth function Φ from $\mathcal{P}_n \times \mathcal{P}_n$ to the set of non negative real numbers is called a *quantum divergence* if

i) $\Phi(X,Y) \geq 0, \forall X, Y \in \mathcal{P}_n$, $\Phi(X,Y) = 0$ if and only if $X = Y$;
ii) The derivative $D\Phi$ with respect to the second variable vanishes on the diagonal, i.e.,

$$D\Phi(X,Y)|_{X=Y}(B) = 0 \text{ for all Hermitian matrix } B;$$

iii) The second derivative $D^2\Phi$ is positive on the diagonal, i.e.,

$$D^2\Phi(X,Y)|_{X=Y}(B,B) \geq 0 \text{ for all Hermitian matrix } B.$$

*Corresponding Author.
Email addresses: thdinh@troy.edu (T.H. Dinh), vula@uel.edu.vn (A.V. Le), lecongtrinh@qnu.edu.vn (C.T. Le), phanngocyen@vanlanguni.vn (N.Y. Phan)
Received: 25.03.2021; Accepted: 27.09.2021

For positive definite matrices A and B, it is well-known that the geometric mean $A\sharp B$ is the midpoint of the geodesic curve $A\sharp_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$ $(t \in [0,1])$ joining A and B. In [15] J. Pitrik and D.Virosztek gave a divergence center interpretation for every symmetric Kubo-Ando mean. They also studied the weighted and multivariate versions of a large class of symmetric Kubo-Ando means. The good thing about Kubu-Ando means is that each Kubo-Ando mean σ is corresponding to an operator monotone function f_σ via the relation

$$A\sigma B = A^{1/2} f_\sigma(A^{-1/2}BA^{-1/2})A^{1/2}.$$

Using this relation, Pitrik and Virosztek constructed a corresponding quantum divergence, and showed that the given Kubo-Ando mean is the unique solution of the least squares problem with respect to this divergence. Recently, Lam and Milley [13] considered a modified quantum Hellinger divergence in which the matrix power mean $\mu_p(t, A_1, A_2, \cdots, A_n) = ((A_1^p + A_2^p + \cdots A_n^p)/n)^{1/p}$ is the unique solution of the corresponding least square problem. In another paper [12] Lam and Le introduced and studied quantum divergences with p-power means $\mu_p(t, A, B)$. Notice that one of the most important things in the least square problem is the uniqueness of the solution. It is also well-known that for any $\alpha \in [0,1]$ the function $\mathrm{Tr}\,(A^\alpha B^{1-\alpha})$ is jointly concave, hence the function

$$\mathrm{Tr}\,(H_\alpha(A,B)) = \frac{1}{2}\mathrm{Tr}\,(A^\alpha B^{1-\alpha} + A^{1-\alpha}B^\alpha) \tag{1.1}$$

is jointly concave, where $H_\alpha(A,B) = \frac{1}{2}(A^\alpha B^{1-\alpha} + A^{1-\alpha}B^\alpha)$ is the non-Kubo-Ando Heinz mean. The matrix power mean is a curve joining A and B, while the matrix Heinz mean is a curve joining the arithmetic mean $(A+B)/2$ and "the naive geometric mean" $A^{1/2}B^{1/2}$.

Motivated by works mentioned above, in this paper we define and study the following quantity

$$\Phi(X,Y) = \mathrm{Tr}\left[\left(\frac{1-\alpha}{\alpha} + \frac{\alpha}{1-\alpha}\right)X + 2Y - \frac{X^{1-\alpha}Y^\alpha}{\alpha} - \frac{X^\alpha Y^{1-\alpha}}{1-\alpha}\right],$$

where $0 < \alpha < 1$. We show that the quantity $\Phi(A,B)$ is a quantum divergence (Theorem 2.1). Then, we use the joint concavity of the function (1.1) to show that the least square problem

$$\min_{X>0} \sum_{i=1}^m \omega_i \Phi\,(A_i, X)$$

has a unique solution which is the unique solution of the following matrix equation

$$\sum_{i=1}^m \omega_i H_\alpha\,(X, A_i) = X,$$

where H_α is the matrix Heinz mean (Theorem 2.2). In order to show the usefulness of the new quantum divergence, we prove that the new quantum divergence satisfies the Data Processing Inequality (Theorem 3.1). Finally, we also show that the matrix power mean $\mu_p(t, A, B)$ satisfies the in-betweenness property (Theorem 3.2). On the in-betweenness property we refer the readers to [2, 4–9, 12].

2. A new quantum divergence and the least squares problem

Theorem 2.1. *For any $0 < \alpha < 1$, the quantity $\Phi(X,Y)$ is a quantum divergence.*

Proof. Firstly, we show that $\Phi(X,Y) \geq 0$. Indeed, according to the AGM inequality [11] we have

$$\mathrm{Tr}\left(\frac{X^{1-\alpha}Y^\alpha}{\alpha}\right) = \frac{1}{\alpha}\mathrm{Tr}\,(X^{1-\alpha}Y^\alpha) \leq \frac{1}{\alpha}\mathrm{Tr}\,((1-\alpha)X + \alpha Y),$$

and

$$\mathrm{Tr}\left(\frac{X^\alpha Y^{1-\alpha}}{1-\alpha}\right) = \frac{1}{1-\alpha}\mathrm{Tr}\,(X^\alpha Y^{1-\alpha}) \le \frac{1}{1-\alpha}\mathrm{Tr}\,(\alpha X + (1-\alpha)Y).$$

From here, it follows that

$$\mathrm{Tr}\left(\frac{X^{1-\alpha}Y^\alpha}{\alpha} + \frac{X^\alpha Y^{1-\alpha}}{1-\alpha}\right) \le \mathrm{Tr}\,\left[\left(\frac{1-\alpha}{\alpha} + \frac{\alpha}{1-\alpha}\right)X + 2Y\right]. \tag{2.1}$$

The equality happens if and only if $X = Y$, so $\Phi(X,Y)$ satisfies the first property in the definition of divergence.

We now prove that for every Hermitian matrix B, $\frac{\partial \Phi}{\partial Y}(X,Y)|_{X=Y}(B) = 0$. Recall [11, Chapter 3] that $\frac{d}{dt}|_{t=0}\mathrm{Tr}\,(f(A+tB)) = \mathrm{Tr}\,(f'(A)B)$. Using this fact, we have

$$\frac{\partial \Phi}{\partial Y}(X,Y)(B) = \mathrm{Tr}\,\left[2B - \frac{X^{1-\alpha}D(Y^\alpha)(B)}{\alpha} - \frac{X^\alpha D(Y^{1-\alpha})(B)}{1-\alpha}\right].$$

When $X = Y$,

$$\frac{\partial \Phi}{\partial Y}(X,Y)|_{X=Y}(B) = \mathrm{Tr}\,\left[2B - \frac{X^{1-\alpha}D(X^\alpha)(B)}{\alpha} - \frac{X^\alpha D(X^{1-\alpha})(B)}{1-\alpha}\right].$$

Differentiating both sides of the identity $\mathrm{Tr}\,\left((X^\alpha)^{\frac{1}{\alpha}}\right) = \mathrm{Tr}\,(X)$, we obtain

$$\frac{d}{dt}|_{t=0}\mathrm{Tr}\,\left((X^\alpha)^{\frac{1}{\alpha}}(B)\right) = \mathrm{Tr}\,(B), \quad \text{or} \quad \mathrm{Tr}\,\left(\frac{1}{\alpha}(X^\alpha)^{\frac{1}{\alpha}-1}D(X^\alpha)(B)\right) = \mathrm{Tr}\,(B).$$

Equivalently,

$$\mathrm{Tr}\,\left(\frac{X^{1-\alpha}D(X^\alpha)(B)}{\alpha}\right) = \mathrm{Tr}\,(B). \tag{2.2}$$

On the other hand, letting $t = 1-\alpha$, on account of (2.2) we get

$$\mathrm{Tr}\,\left(\frac{X^\alpha D(X^{1-\alpha})(B)}{1-\alpha}\right) = \mathrm{Tr}\,\left(\frac{X^{1-t}D(X^t)(B)}{t}\right) = \mathrm{Tr}\,(B). \tag{2.3}$$

Therefore, $\frac{\partial \Phi}{\partial Y}(X,Y)|_{X=Y}(B) = \mathrm{Tr}\,(2B) - \mathrm{Tr}\,(2B) = 0$.

Finally, we need to check that $\frac{\partial^2 \Phi}{\partial Y^2}(X,Y)|_{X=Y}(B,B) \ge 0$ for every Hermitian matrix B. We have

$$\frac{\partial^2 \Phi}{\partial Y^2}(X,Y)(B,B) = \mathrm{Tr}\,\left[-\frac{1}{\alpha}X^{1-\alpha}D^2(Y^\alpha)(B,B) - \frac{1}{1-\alpha}X^\alpha D^2(Y^{1-\alpha})(B,B)\right].$$

When $X = Y$,

$$\frac{\partial^2 \Phi}{\partial Y^2}(X,Y)|_{X=Y}(B,B) = \mathrm{Tr}\,\left[-\frac{1}{\alpha}X^{1-\alpha}D^2(X^\alpha)(B,B) - \frac{1}{1-\alpha}X^\alpha D^2(X^{1-\alpha})(B,B)\right].$$

Now, differentiating both sides of (2.2) and (2.3) we obtain

$$\frac{1}{\alpha}\mathrm{Tr}\,\left[D(X^{1-\alpha})(B)D(X^\alpha)(B) + X^{1-\alpha}D^2(X^\alpha)(B,B)\right] = 0,$$

and

$$\frac{1}{1-\alpha}\mathrm{Tr}\,\left[D(X^\alpha)(B)D(X^{1-\alpha})(B) + X^\alpha D^2(X^{1-\alpha})(B,B)\right] = 0.$$

Therefore,

$$\begin{aligned}\frac{\partial^2\Phi}{\partial Y^2}(X,Y)|_{X=Y}(B,B) &= \frac{1}{\alpha}\mathrm{Tr}\left[D(X^{1-\alpha})(B)D(X^{\alpha})(B)\right]\\ &+ \frac{1}{1-\alpha}\mathrm{Tr}\left[D(X^{\alpha})(B)D(X^{1-\alpha})(B)\right]\\ &= \left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right)\mathrm{Tr}\left[D(X^{1-\alpha})(B)D(X^{\alpha})(B)\right].\end{aligned}$$

Now, assume that $X = U^*D_XU$, where $D_X = \mathrm{diag}(\lambda_i)$ and U is a unitary matrix. Then

$$\begin{aligned}D(X^{1-\alpha})(B) = \frac{d}{dt}|_{t=0}(X+tB)^{1-\alpha} &= \frac{d}{dt}|_{t=0}(U^*D_XU + tU^*UBU^*U)^{1-\alpha}\\ &= U^*\frac{d}{dt}|_{t=0}(D_X + tUBU^*)^{1-\alpha}U\\ &= U^*D(X^{1-\alpha})(UBU^*)|_{X=D_X}U.\end{aligned}$$

Similarly,

$$D(X^{\alpha})(B) = U^*D(X^{\alpha})(UBU^*)|_{X=D_X}.$$

Consequently,

$$\begin{aligned}\mathrm{Tr}\left[D(X^{1-\alpha})(B)D(X^{\alpha})(B)\right] &= \mathrm{Tr}\left[U^*D(X^{1-\alpha})(UBU^*)|_{X=D_X}UU^*D(X^{\alpha})(UBU^*)|_{X=D_X}U\right]\\ &= \mathrm{Tr}\left[U^*D(X^{1-\alpha})(UBU^*)|_{X=D_X}D(X^{\alpha})(UBU^*)|_{X=D_X}U\right]\\ &= \mathrm{Tr}\left[D(X^{1-\alpha})(UBU^*)|_{X=D_X}D(X^{\alpha})(UBU^*)|_{X=D_X}\right].\end{aligned}$$

Therefore, we can assume that X is diagonal. Suppose that $X = diag(\lambda_i)$. In this case, on account of the Mean Value Theorem we get

$$[D(X^{\alpha})(B)]_{ij} = \begin{cases}\dfrac{\lambda_i^{\alpha}-\lambda_j^{\alpha}}{\lambda_i-\lambda_j}B_{ij}, \lambda_i \neq \lambda_j\\ \alpha\lambda_i^{\alpha-1}B_{ij}, \lambda_i = \lambda_j\end{cases} = \alpha s_{ij}^{\alpha-1}B_{ij},$$

where s_{ij} between λ_i and λ_j. Similarly,

$$\left[D(X^{1-\alpha})(B)\right]_{ij} = \begin{cases}\dfrac{\lambda_i^{1-\alpha}-\lambda_j^{1-\alpha}}{\lambda_i-\lambda_j}B_{ij}, \lambda_i \neq \lambda_j\\ (1-\alpha)\lambda_i^{-\alpha}B_{ij}, \lambda_i = \lambda_j\end{cases} = (1-\alpha)t_{ij}^{-\alpha}B_{ij},$$

where t_{ij} is between λ_i and λ_j. Hence,

$$\begin{aligned}\mathrm{Tr}\left[D(X^{1-\alpha})(B)D(X^{\alpha})(B)\right] &= \sum_{i,k}\left[D(X^{1-\alpha})(B)\right]_{ik}[D(X^{\alpha})(B)]_{ki}\\ &= \sum_{i,k}(1-\alpha)t_{ik}^{-\alpha}\alpha s_{ki}^{\alpha-1}B_{ik}B_{ki}\\ &= \sum_{i,k}(1-\alpha)t_{ik}^{-\alpha}\alpha s_{ki}^{\alpha-1}\left|B_{ik}\right|^2 \geq 0.\end{aligned}$$

Therefore,

$$\mathrm{Tr}\left[D(X^{1-\alpha})(B)D(X^{\alpha})(B)\right]\left(\frac{1}{\alpha}+\frac{1}{1-\alpha}\right) \geq 0.$$

We can conclude that $\dfrac{\partial^2\Phi}{\partial Y^2}(X,Y)|_{X=Y}(B,B) \geq 0$ for every Hermitian matrix B. □

Now, let $A_1, A_2, \cdots A_m$ be positive definite matrices and $\omega = (\omega_1, \omega_2, \cdots \omega_n)$ be a probability vector. Consider the least squares problem

$$\min_{X>0} \sum_{i=1}^{m} \omega_i \Phi\left(A_i, X\right). \tag{2.4}$$

Theorem 2.2. *For $0 < \alpha < 1$, the function*

$$F(X) = \sum_{i=1}^{m} \omega_i \Phi\left(A_i, X\right)$$

attains its minimum at X_0 which is the unique positive definite solution of the following matrix equation

$$\sum_{i=1}^{m} \omega_i H_\alpha\left(X, A_i\right) = X, \tag{2.5}$$

where $H_\alpha\left(X, A_i\right)$ is the non-Kubo-Ando matrix Heinz mean.

Proof. We have

$$\Phi(A_i, X) = \operatorname{Tr}\left[\left(\frac{1-\alpha}{\alpha} + \frac{\alpha}{1-\alpha}\right) A_i + 2X - \frac{A_i^{1-\alpha} X^\alpha}{\alpha} - \frac{A_i^\alpha X^{1-\alpha}}{1-\alpha}\right],$$

and

$$\frac{\partial F(X)}{\partial X}(B) = \operatorname{Tr}\left[2B - \sum_{i=1}^{m} \omega_i \frac{A_i^\alpha}{1-\alpha} D(X^{1-\alpha})(B) - \sum_{i=1}^{m} \omega_i \frac{A_i^{1-\alpha}}{\alpha} D(X^\alpha)(B)\right].$$

Recall that

$$D(X^{1-\alpha})(B) = \int_0^\infty (\lambda + X)^{-1} B (\lambda + X)^{-1} \lambda^{1-\alpha} d\lambda \frac{\sin((1-\alpha)\pi)}{\pi},$$

and

$$D(X^\alpha)(B) = \int_0^\infty (\lambda + X)^{-1} B (\lambda + X)^{-1} \lambda^{\alpha} d\lambda \frac{\sin(\alpha\pi)}{\pi}.$$

Then we have

$$\begin{aligned}
\frac{\partial F(X)}{\partial X}(B) &= \operatorname{Tr}\left[2B - \sum_{i=1}^{m} \omega_i \frac{A_i^\alpha}{1-\alpha} \int_0^\infty (\lambda + X)^{-1} B (\lambda + X)^{-1} \lambda^{1-\alpha} d\lambda \frac{\sin((1-\alpha)\pi)}{\pi}\right] \\
&\quad - Tr\left[\sum_{i=1}^{m} \omega_i \frac{A_i^{1-\alpha}}{\alpha} \int_0^\infty (\lambda + X)^{-1} B (\lambda + X)^{-1} \lambda^{\alpha} d\lambda \frac{\sin(\alpha\pi)}{\pi}\right] \\
&= \operatorname{Tr}\left[2I - \int_0^\infty (\lambda + X)^{-1} \sum_{i=1}^{m} \omega_i \frac{A_i^\alpha}{1-\alpha} (\lambda + X)^{-1} \lambda^{1-\alpha} d\lambda \frac{\sin((1-\alpha)\pi)}{\pi} (B)\right] \\
&\quad - \operatorname{Tr}\left[\int_0^\infty (\lambda + X)^{-1} \sum_{i=1}^{m} \omega_i \frac{A_i^{1-\alpha}}{\alpha} (\lambda + X)^{-1} \lambda^{\alpha} d\lambda \frac{\sin(\alpha\pi)}{\pi} (B)\right].
\end{aligned}$$

The condition $\dfrac{\partial F(X)}{\partial X}(B) = 0$ for every $B \geq 0$ implies that

$$2I = \int_0^\infty (\lambda + X)^{-1} \sum_{i=1}^m \omega_i \frac{A_i^\alpha}{1-\alpha} (\lambda + X)^{-1} \lambda^{1-\alpha} d\lambda \frac{\sin((1-\alpha)\pi)}{\pi}$$
$$+ \int_0^\infty (\lambda + X)^{-1} \sum_{i=1}^m \omega_i \frac{A_i^{1-\alpha}}{\alpha} (\lambda + X)^{-1} \lambda^{\alpha} d\lambda \frac{\sin(\alpha\pi)}{\pi}. \quad (2.6)$$

Now, we choose a basis in which the matrix $Y_1 = \sum_{i=1}^m \omega_i \frac{A_i^\alpha}{1-\alpha} = diag(y_{1ii})$ is diagonal. Let Y_2 is the representation of $\sum_{i=1}^m \omega_i \frac{A_i^{1-\alpha}}{\alpha}$ in the new basis. Now, we are going to show that there exists a diagonal matrix $X = diag(x_i)$ $(x_i > 0)$ satisfying the equation (2.6). We have

$$2I = \int_0^\infty (\lambda + diag(x_i))^{-1} Y_1 (\lambda + diag(x_i))^{-1} \lambda^{1-\alpha} d\lambda \frac{\sin((1-\alpha)\pi)}{\pi}$$
$$+ \int_0^\infty (\lambda + diag(x_i))^{-1} Y_2 (\lambda + diag(x_i))^{-1} \lambda^{\alpha} d\lambda \frac{\sin(\alpha\pi)}{\pi} \lambda^{1-\alpha} d\lambda,$$

which implies that

$$\int_0^\infty \frac{y_{1ij}}{(\lambda + x_i)(\lambda + x_j)} \frac{\sin((1-\alpha)\pi)}{\pi} \lambda^{1-\alpha} d\lambda + \int_0^\infty \frac{y_{2ij}}{(\lambda + x_i)(\lambda + x_j)} \frac{\sin(\alpha\pi)}{\pi} \lambda^{\alpha} d\lambda = 2\delta_{ij}. \quad (2.7)$$

Mention that for any $\alpha \in (0,1)$ and for any $x_j > x_i > 0$,

$$\int_0^\infty \frac{\lambda^\alpha}{(\lambda + x_i)(\lambda + x_j)} d\lambda \geq \int_0^\infty \frac{\lambda^\alpha}{(\lambda + x_j)^2} d\lambda > 0.$$

From (2.7), for $i \neq j$,

$$\int_0^\infty \frac{y_{2ij}}{(\lambda + x_i)(\lambda + x_j)} \frac{\sin(\alpha\pi)}{\pi} \lambda^\alpha d\lambda = 0.$$

Consequently, $y_{2ij} = 0$ for $i \neq j$, hence, the matrix Y_2 is diagonal in the new basis, and

$$\int_0^\infty \frac{y_{1ii}}{(\lambda + x_i)^2} \frac{\sin((1-\alpha)\pi)}{\pi} \lambda^{1-\alpha} d\lambda + \int_0^\infty \frac{y_{2ii}}{(\lambda + x_i)^2} \frac{\sin(\alpha\pi)}{\pi} \lambda^\alpha d\lambda = 2. \quad (2.8)$$

Differentiating both sides of the following identities

$$x^{1-\alpha} = \cos\left(\frac{(1-\alpha)\pi}{2}\right) + \frac{\sin(1-\alpha)\pi}{\pi} \int_0^\infty \left(\frac{\lambda}{\lambda^2 + 1} - \frac{1}{\lambda + x}\right) \lambda^{1-\alpha} d\lambda,$$

$$x^\alpha = \cos\left(\frac{\alpha\pi}{2}\right) + \frac{\sin(\alpha\pi)}{\pi} \int_0^\infty \left(\frac{\lambda}{\lambda^2 + 1} - \frac{1}{\lambda + x}\right) \lambda^\alpha d\lambda,$$

we get

$$\frac{\sin(1-\alpha)\pi}{\pi} \int_0^\infty \frac{1}{(\lambda + x_i)^2} \lambda^{1-\alpha} d\lambda = (1-\alpha) x_i^{-\alpha},$$

and

$$\frac{\sin(\alpha\pi)}{\pi}\int_0^\infty \frac{1}{(\lambda+x_i)^2}\lambda^\alpha d\lambda = \alpha x_i^{\alpha-1}.$$

From (2.8) we have

$$(1-\alpha)y_{1ii}x_i^{-\alpha} + \alpha y_{2ii}x_i^{\alpha-1} = 2. \tag{2.9}$$

If we show that the equation (2.9) has solution for any $i = 1, 2, \cdots, n$, then from here we obtain the matrix $X > 0$ satisfying

$$\sum_{i=1}^m \omega_i(X^{-\alpha}A_i^\alpha + X^{\alpha-1}A_i^{1-\alpha}) = 2I.$$

And then, multiplying both sides of the last identity from the left by $X/2$, we get

$$X = \sum_{i=1}^m \omega_i \frac{X^{1-\alpha}A_i^\alpha + X^\alpha A_i^{1-\alpha}}{2} = \sum_{i=1}^m \omega_i H_\alpha(X, A_i) \tag{2.10}$$

Indeed, equation (2.9) is equivalent to the following

$$c_i x_i^{1-\alpha} + d_i x_i^\alpha = 2x_i$$

or

$$x_i^{1-\alpha}(c_i - x_i^\alpha) + x_i^\alpha\left(d_i - x_i^{1-\alpha}\right) = 0,$$

where $(1-\alpha)y_{1ii} = c_i > 0$ and $\alpha y_{2ii} = d_i > 0$. Let $f(x_i) = x_i^{1-\alpha}(c_i - x_i^\alpha) + x_i^\alpha\left(d_i - x_i^{1-\alpha}\right)$. We have $f\left(c_i^{\frac{1}{\alpha}}\right) = c_i\left(d_i - c_i^{\frac{1-\alpha}{\alpha}}\right)$, and $f\left(d_i^{\frac{1}{1-\alpha}}\right) = d_i\left(c_i - d_i^{\frac{\alpha}{1-\alpha}}\right)$. Therefore,

$$f\left(c^{\frac{1}{\alpha}}\right)f\left(d^{\frac{1}{1-\alpha}}\right) = \left(cd - c^{\frac{1}{\alpha}}\right)\left(cd - d^{\frac{1}{1-\alpha}}\right).$$

If $d^\alpha = c^{1-\alpha}$, then $f\left(c^{\frac{1}{\alpha}}\right)f\left(d^{\frac{1}{1-\alpha}}\right) = 0$. That means, the equation $f(x_i) = 0$ have a positive solution. In the case $d^\alpha > c^{1-\alpha}$, we have $cd - c^{\frac{1}{\alpha}} > 0$ and $cd - d^{\frac{1}{1-\alpha}} < 0$. Consequently, $f\left(c^{\frac{1}{\alpha}}\right)f\left(d^{\frac{1}{1-\alpha}}\right) < 0$. By the Intermediate Value Theorem, the function f has a zero in $(c^{1-\alpha}, d^\alpha)$. Similarly, one can show that the function f also has a zero in $(d^\alpha, c^{1-\alpha})$ if $d^\alpha < c^{1-\alpha}$. Therefore, we finish the first part of the theorem that the function $F(X)$ attains it minimum at a solution of the equation (2.10).

Now, if we show that the function $F(X)$ is strictly convex and has a critical point, then the function attains global minimum at that point. In that case, the the equation (2.10) has a unique solution. In order to show the convexity of $F(X)$, it is suffice to show the convexity of $\Phi(A_i, X)$. For $0 < \alpha < 1$, the function $X \mapsto X^\alpha$ is matrix concave. Then for any $t \in [0, 1]$ and $X_1, X_2 \geq 0$,

$$(tX_1 + (1-t)X_2)^\alpha \geq tX_1^\alpha + (1-t)X_2^\alpha.$$

Consequently,

$$\mathrm{Tr}\,(A_i^{1-\alpha}(tX_1 + (1-t)X_2)^\alpha) \geq t\mathrm{Tr}\,(A_i^{1-\alpha}X_1^\alpha) + (1-t)\mathrm{Tr}\,(A_I^{1-\alpha}X_2^\alpha).$$

Therefore,

$$\mathrm{Tr}\left(\frac{A_i^{1-\alpha}(tX_1 + (1-t)Y_2)^\alpha}{\alpha}\right) \geq t\mathrm{Tr}\left(\frac{A_i^{1-\alpha}X_1^\alpha}{\alpha}\right) + (1-t)\mathrm{Tr}\left(\frac{A_i^{1-\alpha}X_2^\alpha}{\alpha}\right).$$

Similarly,

$$\mathrm{Tr}\left(\frac{A_i^\alpha(tX_1 + (1-t)X_2)^{1-\alpha}}{1-\alpha}\right) \geq t\mathrm{Tr}\left(\frac{A_i^\alpha X_1^{1-\alpha}}{1-\alpha}\right) + (1-t)\mathrm{Tr}\left(\frac{A_i^\alpha X_2^{1-\alpha}}{1-\alpha}\right).$$

Combining the last two inequalities we obtain the convexity of $\Phi(A_i, X)$.

Now, we show that $F(X)$ is strictly convex. Suppose that

$$\mathrm{Tr}\left(A_i^{1-\alpha}(tX_1+(1-t)X_2)^{\alpha}\right) = t\mathrm{Tr}\,(A_i^{1-\alpha}X_1^{\alpha}) + (1-t)\mathrm{Tr}\,(A_i^{1-\alpha}X_2^{\alpha}). \tag{2.11}$$

Since

$$A_i^{\frac{1-\alpha}{2}}(tX_1+(1-t)X_2)^{\alpha}A_i^{\frac{1-\alpha}{2}} - tA_i^{\frac{1-\alpha}{2}}X_1^{\alpha}A_i^{\frac{1-\alpha}{2}} - A_i^{\frac{1-\alpha}{2}}(1-t)X_2^{\alpha}A_i^{\frac{1-\alpha}{2}} \le 0,$$

from (2.11) it follows that

$$A_i^{\frac{1-\alpha}{2}}(tX_1+(1-t)X_2)A_i^{\frac{1-\alpha}{2}} = tA_i^{\frac{1-\alpha}{2}}X_1^{\alpha}A_i^{\frac{1-\alpha}{2}} + A_i^{\frac{1-\alpha}{2}}(1-t)X_2^{\alpha}A_i^{\frac{1-\alpha}{2}}$$

and

$$(tX_1+(1-t)X_2)^{\alpha} = tX_1^{\alpha} + (1-t)X_2^{\alpha}. \tag{2.12}$$

For $\alpha \in (0,1)$, the function $f(x) = x^{\alpha}$ is strictly operator concave. Therefore, from (2.12) it follows that $X_1 = X_2$. Thus, the function $\Phi(A_i, X)$ is strictly convex, and hence, $F(X)$ is strictly convex. □

Remark 2.3. The existence of the solution of (2.10) can be obtained using Brouwer's fixed point theorem. Indeed, let a and b be positive numbers such that $aI \le A_i \le bI$, for all $1 \le i \le m$. It is obvious that for $aI \le X \le bI$,

$$aI \le \sum_{i=1}^{m} \omega_i \frac{X^{1-\alpha}A_i^{\alpha} + X^{\alpha}A_i^{1-\alpha}}{2} \le bI.$$

In other words, $G(X)$ is a self-map on the compact and convex $\mathcal{K}$, where

$$\mathcal{K} = \{X \in \mathcal{P}_n : aI \le X \le bI\}.$$

According to Brouwer's fixed point theorem, $G(X)$ has a fixed point.

3. Data processing inequality and In-betweenness property

Recall that the data processing inequality with respect to a quantum divergence Φ means that for any completely positive trace preserving map $\mathcal{E}$ and for any positive semidefinite matrices A and B,

$$\Phi(\mathcal{E}(A), \mathcal{E}(B)) \le \Phi(A, B).$$

Recall that (see, for examples, [14, Theorem 5.16]) if a map $\Phi(A,B)$ is jointly convex, unitarily invariant and invariant under tensor product, then Φ is monotone with respect to all completely positive trace-preserving map.

By the definition, the map $\Phi(A,B)$ is jointly convex because $\mathrm{Tr}\,(A^{\alpha}B^{1-\alpha})$ is jointly concave for any $\alpha \in (0,1)$ [10]. Therefore, from the following theorem it implies that the new quantum divergence satisfies the Data Processing Inequality.

Theorem 3.1. *$\Phi(X,Y)$ is invariant under all unitary matrix U and invariant under tensoring with another density matrix τ, $\mathrm{Tr}\,(\tau) = 1$.*

Proof. For an arbitrary unitary matrix U, we have

$$\begin{aligned}
&\Phi(U^*XU, U^*YU) \\
&= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha} + \frac{1-\alpha}{\alpha}\right)U^*XU + 2U^*YU\right] \\
&\quad - \mathrm{Tr}\left[\frac{(U^*XU)^{1-\alpha}(U^*YU)^{\alpha}}{\alpha} + \frac{(U^*XU)^{\alpha}(U^*YU)^{1-\alpha}}{1-\alpha}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)U^*XU+2U^*YU\right] \\
&\quad - \mathrm{Tr}\left[\frac{U^*X^{1-\alpha}UU^*Y^{\alpha}U}{\alpha}-\frac{U^*X^{\alpha}UU^*Y^{1-\alpha}U}{1-\alpha}\right] \\
&= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)U^*XU+2U^*YU-\frac{U^*X^{1-\alpha}Y^{\alpha}U}{\alpha}-\frac{U^*X^{\alpha}Y^{1-\alpha}U}{1-\alpha}\right] \\
&= \mathrm{Tr}\left[U^*\left(\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X+2Y-\frac{X^{1-\alpha}Y^{\alpha}}{\alpha}-\frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right)U\right] \\
&= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X+2Y-\frac{X^{1-\alpha}Y^{\alpha}}{\alpha}-\frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right] \\
&= \Phi(X,Y).
\end{aligned}
$$

Now, suppose that τ is an arbitrary density matrix. We have

$$
\begin{aligned}
\Phi(X\otimes\tau, Y\otimes\tau) &= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X\otimes\tau+2Y\otimes\tau\right] \\
&\quad - \mathrm{Tr}\left[\frac{(X\otimes\tau)^{1-\alpha}(Y\otimes\tau)^{\alpha}}{\alpha}+\frac{(X\otimes\tau)^{\alpha}(Y\otimes\tau)^{1-\alpha}}{1-\alpha}\right] \\
&= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X\otimes\tau+2Y\otimes\tau\right] \\
&\quad - \mathrm{Tr}\left[\frac{(X^{1-\alpha}\otimes\tau^{1-\alpha})(Y^{\alpha}\otimes\tau^{\alpha})}{\alpha}-\frac{(X^{\alpha}\otimes\tau^{\alpha})(Y^{1-\alpha}\otimes\tau^{1-\alpha})}{1-\alpha}\right] \\
&= \mathrm{Tr}\left[\left((\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X\otimes\tau+2Y\otimes\tau\right] \\
&\quad - \mathrm{Tr}\left(\frac{X^{1-\alpha}Y^{\alpha}}{\alpha}\otimes\tau\right)-\mathrm{Tr}\left(\frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\otimes\tau\right) \\
&= \mathrm{Tr}\left[\left(\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X+2Y-\frac{X^{1-\alpha}Y^{\alpha}}{\alpha}-\frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right)\otimes\tau\right] \\
&= \mathrm{Tr}\left[\left(\frac{\alpha}{1-\alpha}+\frac{1-\alpha}{\alpha}\right)X+2Y-\frac{X^{1-\alpha}Y^{\alpha}}{\alpha}-\frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right]\mathrm{Tr}\,(\tau) \\
&= \Phi(X,Y).
\end{aligned}
$$

□

The in-betweenness was introduced by Audenaert in [2]. He showed that the matrix power mean $\mu_p(t;X,Y)=(tX^p+(1-t)Y^p)^{\frac{1}{p}}$ satisfies the in-betweenness property. This property was investigated by the first author and co-authors in [4–9]. To finish this paper, we show that the matrix power mean $\mu_p(t;X,Y)$ also satisfies the in-betweenness property with respect to the new divergence in the previous section.

Theorem 3.2. *Let $X,Y\in\mathcal{P}_n$, and α and p such that $0<\max\{\alpha,1-\alpha\}\le p\le 1$. Then*

$$\Phi(X,\mu_p)\le\Phi(X,Y), \tag{3.1}$$

where $\mu_p:=\mu_p(t;X,Y)=(tX^p+(1-t)Y^p)^{\frac{1}{p}}$.

Proof. The inequality (3.1) is equivalent to the following

$$\operatorname{Tr}\left[2\mu_p - \frac{X^{1-\alpha}\mu_p^{\alpha}}{\alpha} - \frac{X^{\alpha}\mu_p^{1-\alpha}}{1-\alpha}\right] \le \operatorname{Tr}\left[2Y - \frac{X^{1-\alpha}Y^{\alpha}}{\alpha} - \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right],$$

or,

$$\operatorname{Tr}\left[2\mu_p + \frac{X^{1-\alpha}Y^{\alpha}}{\alpha} + \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right] \le \operatorname{Tr}\left[2Y + \frac{X^{1-\alpha}\mu_p^{\alpha}}{\alpha} + \frac{X^{\alpha}\mu_p^{1-\alpha}}{1-\alpha}\right]. \tag{3.2}$$

By the operator convexity of the map $x \mapsto x^{\frac{1}{p}}$ when $\frac{1}{2} \le p \le 1$, we have

$$\mu_p = (tX^p + (1-t)Y^p)^{\frac{1}{p}} \le tX + (1-t)Y.$$

Therefore,

$$\operatorname{Tr}(2\mu_p) \le \operatorname{Tr}(2tX + 2(1-t)Y).$$

Consequently, it is enough to prove

$$\operatorname{Tr}\left[2tX + 2(1-t)Y + \frac{X^{1-\alpha}Y^{\alpha}}{\alpha} + \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right] \le \operatorname{Tr}\left[2Y + \frac{X^{1-\alpha}\mu_p^{\alpha}}{\alpha} + \frac{X^{\alpha}\mu_p^{1-\alpha}}{1-\alpha}\right].$$

Equivalently,

$$\operatorname{Tr}\left[2t(X-Y) + \frac{X^{1-\alpha}Y^{\alpha}}{\alpha} + \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right] \le \operatorname{Tr}\left[\frac{X^{1-\alpha}\mu_p^{\alpha}}{\alpha} + \frac{X^{\alpha}\mu_p^{1-\alpha}}{1-\alpha}\right]. \tag{3.3}$$

Indeed, by the operator concavity of the map $x \mapsto x^{\frac{\alpha}{p}}$, when $0 < \alpha < p \le 1$ and $x \mapsto x^{\frac{1-\alpha}{p}}$, when $0 < 1-\alpha < p \le 1$, we have

$$\mu_p^{\alpha} = (tX^p + (1-t)Y^p)^{\frac{\alpha}{p}} \ge tX^{\alpha} + (1-t)Y^{\alpha}$$

and

$$\frac{1}{\alpha}\operatorname{Tr}(X^{1-\alpha}\mu_p^{\alpha}) \ge \frac{1}{\alpha}\operatorname{Tr}\left[tX + (1-t)X^{1-\alpha}Y^{\alpha}\right].$$

Similarly,

$$\mu_p^{1-\alpha} = (tX^p + (1-t)Y^p)^{\frac{1-\alpha}{p}} \ge tX^{1-\alpha} + (1-t)Y^{1-\alpha}$$

and

$$\frac{1}{1-\alpha}\operatorname{Tr}(X^{\alpha}\mu_p^{1-\alpha}) \ge \frac{1}{1-\alpha}\operatorname{Tr}\left[tX + (1-t)X^{\alpha}Y^{1-\alpha}\right].$$

Consequently,

$$\operatorname{Tr}\left[t\left(\frac{1}{\alpha} + \frac{1}{1-\alpha}\right)X + (1-t)\left(\frac{X^{1-\alpha}Y^{\alpha}}{\alpha} + \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right)\right] \le \operatorname{Tr}\left[\frac{X^{1-\alpha}\mu_p^{\alpha}}{\alpha} + \frac{X^{\alpha}\mu_p^{1-\alpha}}{1-\alpha}\right].$$

Therefore, the inequality (3.3) follows if

$$\begin{aligned}&\operatorname{Tr}\left[2t(X-Y) + \frac{X^{1-\alpha}Y^{\alpha}}{\alpha} + \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right] \\ &\le \operatorname{Tr}\left[t\left(\frac{1}{\alpha} + \frac{1}{1-\alpha}\right)X + (1-t)\left(\frac{X^{1-\alpha}Y^{\alpha}}{\alpha} + \frac{X^{\alpha}Y^{1-\alpha}}{1-\alpha}\right)\right],\end{aligned}$$

which is equivalent to the inequality (2.1).

□

Acknowledgement. Research of the first author is funded by Troy University under the grant 2021/2022. Research of the second author is funded by Viet Nam National University Ho Chi Minh City (VNU-HCM), under the grant B2022-34-05. Research of the third author is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under the grant number 101.01-2020.12.

References

[1] S. Amari, *Information Geometry and Its Applications*. Springer, 2016.

[2] K.M.R. Audenaert, *In-betweenness, a geometrical monotonicity property for operator means*, Linear Algebra Appl. **438**, 1769-1778, 2013.

[3] R. Bhatia, S. Gaubert and T. Jain, *Matrix versions of Hellinger distance*, Lett. Math. Phys. **109**, 2779-2781, 2019.

[4] T.H. Dinh, R. Dumitru and J.A. Franco, *On the monotonicity of weighted power means for matrices*, Linear Algebra Appl. **527**, 128-140, 2017.

[5] T.H. Dinh, R. Dumitru and J.A. Franco, *Some geometric properties of matrix means in different distance functions*, Positivity, **24**, 1419-1434, 2020.

[6] T.H. Dinh, B.K. Vo and T.Y. Tam, *In-sphere property and reverse inequalities for matrix means*, Electron. J. Linear Algebra, **35** (1), 35-41, 2019.

[7] T.H. Dinh, C.T. Le, B.K. Vo and T.D. Vuong, *Weighted Hellinger Distance and In-betweenness property*, Math. Inequal. Appl. **24** (1), 157-165, 2021.

[8] T.H. Dinh, C.T. Le, B.K. Vo and T.D. Vuong, *The α-z-Bures Wesserstein divergence*, Linear Algebra Appl. **624**, 267-280, 2021.

[9] F. Franco and R. Dumitru, *Generalized Hellinger Metrics and Audenaert's In-Betweenness*, Linear Algebra Appl. **585**, 191-198, 2020.

[10] F. Frank and E. Lieb, *Monotonicity of a relative Renyi entropy*, J. Math. Phys. **54**, 122201, 2013.

[11] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Application*, Springer, 2014.

[12] N. Lam and P.L. Le, *Quantum divergences with p-power means*, Linear Algebra Appl. **609**, 289-307, 2021.

[13] N. Lam and R. Milley, *Some notes on quantum Hellinger divergences with Heinz means*, Electron. J. Linear Algebra, **36**, 704-722, 2020.

[14] M.M. Wolf, *Quantum channels and operations: Guided tour*, Lecture Notes Available at http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf, 2012.

[15] J. Pitrik and D. Virosztek, *A divergence center interpretation of general symmetric Kubo-Ando means, and related weighted multivariate operator means*, Linear Algebra Appl. **609**, 203-217, 2021.