

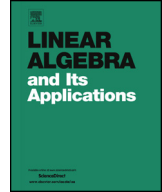


ELSEVIER

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa



The α - z -Bures Wasserstein divergence

Trung Hoa Dinh^a, Cong Trinh Le^{b,c,*}, Bich Khue Vo^d,
Trung Dung Vuong^e

^a Department of Mathematics, Troy University, Troy, AL 36082, United States

^b Division of Computational Mathematics and Engineering, Institute for Computational Science, Ton Duc Thang University, Ho Chi Minh City, Viet Nam

^c Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Viet Nam

^d Faculty of Economics and Law, University of Finance and Marketing, Ho Chi Minh City, Viet Nam

^e Department of Mathematics and Statistics, Quy Nhon University, Viet Nam



ARTICLE INFO

Article history:

Received 3 November 2020

Accepted 6 April 2021

Available online 21 April 2021

Submitted by X. Zhan

MSC:

47A63

47A56

Keywords:

Quantum divergence

α - z Bures distance

Least squares problem

Karcher mean

Matrix power mean

In-betweenness property

Data processing inequality

ABSTRACT

In this paper, we introduce the α - z -Bures Wasserstein divergence for positive semidefinite matrices A and B as

$$\Phi(A, B) = \text{Tr}((1 - \alpha)A + \alpha B) - \text{Tr}(Q_{\alpha, z}(A, B)),$$

where $Q_{\alpha, z}(A, B) = (A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}})^z$ is the matrix function in the α - z -Renyi relative entropy. We show that for $0 \leq \alpha \leq z \leq 1$, the quantity $\Phi(A, B)$ is a quantum divergence and satisfies the Data Processing Inequality in quantum information. We also solve the least squares problem with respect to the new divergence. In addition, we show that the matrix power mean $\mu(t, A, B) = ((1 - t)A^p + tB^p)^{1/p}$ satisfies

* Corresponding author at: Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Viet Nam.

E-mail addresses: thdinh@troy.edu (T.H. Dinh), leongtrinh@tdtu.edu.vn (C.T. Le), votbkhue@gmail.com (B.K. Vo), vuongtrungdung@qnu.edu.vn (T.D. Vuong).

the in-betweenness property with respect to the α - z -Bures Wasserstein divergence.

© 2021 Elsevier Inc. All rights reserved.

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.

It is well-known that in the Riemannian manifold of positive definite matrices, the weighted geometric mean $A \sharp_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$ is the unique geodesic joining A and B , where $A, B \in \mathcal{P}_n$. For $t = 1/2$, $A \sharp_{1/2} B$ is called the geometric mean of A and B . It is obvious that $A \sharp_{1/2} B$ is a matrix generalization of the geometric mean $\sqrt{ab}$ of positive numbers a and b . Let $A_1, A_2, \dots, A_m$ be positive definite matrices. In 2004, Moakher [17] and then Bhatia and Holbrook [5] studied the following least squares problem

$$\min_{X > 0} \sum_{i=1}^m \delta_2^2(X, A_i), \quad (1)$$

where $\delta_2(A, B) = \|\log(A^{-1}B)\|_2$ is the Riemannian distance between A and B . They showed that (1) has a unique solution which is called the Karcher mean of $A_1, A_2, \dots, A_m$. In literature, this mean has different names such as: Fréchet mean, Cartan mean, Riemannian center of mass. It turns out that the solution of (1) is the unique positive definite solution of the Karcher equation

$$\sum_{i=1}^n \log(X^{1/2} A_i X^{1/2}) = 0. \quad (2)$$

In [16], Lim and Palfia showed that the solution of (2) is nothing but the limit of the solution of the following matrix equation as $t \rightarrow 0$,

$$X = \sum_{i=1}^n \frac{1}{n} X \sharp_t A_i. \quad (3)$$

Let $P_{\alpha, z}(A, B) = (B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}})^z$. This is the matrix function in the α - z -Rényi relative entropy introduced by Audenaert and Datta [1] in 2015. Recently, Franco and Dumitru [12] introduced the so-called Rényi power means of matrices. More precisely, for $0 < \alpha_i \leq z_i \leq 1$ and for positive definite matrices A_i, B_i , using the approach in [16] developed by Lim and Pálfi, they showed that the following equation

$$X = \sum_{i=1}^m \omega_i P_{\alpha_i, z_i}(X, A_i) \quad (4)$$

had a unique positive definite solution, where (ω_i) is a probability vector. Notice that if we replace $P_{\alpha_i, z_i}(X, A_i)$ in (4) with the weighted geometric mean $X \#_t A_i$, the solution of the corresponding matrix equation is the weighted power mean.

Now, notice that if we change the distance function in (1), the solution may be different, if exists. Interestingly, in applications people sometimes are interested in distance-like functions that provide distance between two data points. Such functions are not necessarily symmetric; and the triangle inequality does not need to be true. Divergences are such distance-like functions. An important example of divergences is the Bures-Wasserstein metric studied by Bhatia and coauthors [3] as follows:

$$d_b(A, B) = (Tr((A + B)/2) - Tr(A^{1/2}BA^{1/2})^{1/2})^{1/2},$$

where $Tr((A^{1/2}BA^{1/2})^{1/2})$ is the quantum fidelity of two positive definite matrices A and B . They showed that d_b^2 is a quantum divergence and solved the least squares problem with respect to the Bures-Wasserstein divergence. In another paper, these authors introduced so called the weighted Bures-Wasserstein distance as

$$d_{b,t}(A, B) = (Tr((1-t)A + tB) - Tr(F_t(A, B)))^{1/2},$$

where $F_t(A, B) = Tr(A^{\frac{1-t}{2t}}BA^{\frac{1-t}{2t}})^t$ is the sandwiched quasi-relative entropy [18,15]. They also solved the least squares problem with respect to this divergence. Mention that $(A^{1/2}BA^{1/2})^{1/2}$ and $(A^{\frac{1-t}{2t}}BA^{\frac{1-t}{2t}})^t$ are matrix generalizations of the geometric mean $\sqrt{ab}$ and the weighted geometric mean $a^{1-t}b^t$ of positive numbers a and b , respectively.

Motivated by works mentioned above, in this paper, we introduce and study different properties of the α - z -Bures Wasserstein divergence defined as

$$\Phi(A, B) = Tr((1-\alpha)A + \alpha B) - Tr(Q_{\alpha,z}(A, B)), \quad (5)$$

whenever A and B are positive definite matrices, and $Q_{\alpha,z}(A, B) = P_{\alpha,z}(B, A)$. Mention that $Q_{\alpha,z}(A, B)$ is also a parameterized matrix version of the weighted geometric mean $a^{1-\alpha}b^\alpha$.

In the next section, we show that $\Phi(A, B)$ is a quantum divergence. Using the well-known Brouwer fixed point theorem, we prove that the averaging element of m positive semidefinite matrices $A_1, A_2, \dots, A_m$ is the unique positive definite solution of the following matrix equation

$$\sum_{i=1}^m w_i Q_{\alpha,z}(X, A_i) = X.$$

In Section 3, we show that the α - z -Bures Wasserstein divergence satisfies the Data Processing Inequality in quantum information. Finally, we show that the matrix power

mean $\mu(t, A, B) = ((1-t)A^p + tB^p)^{1/p}$ satisfies the in-betweenness property in the α - z -Bures Wasserstein divergence. On the in-betweenness property we refer the readers to [2,13,8–11].

2. The α - z -Bures Wasserstein divergence and the least squares problem

Definition 2.1 ([4]). A smooth function Φ from $\mathcal{P}_n \times \mathcal{P}_n$ to the set of nonnegative real numbers is called a *quantum divergence* if

- (i) $\Phi(A, B) = 0$ if and only if $A = B$;
- (ii) The derivative $D\Phi$ with respect to the second variable vanishes on the diagonal, i.e.,

$$D\Phi(A, X)|_{X=A} = 0;$$

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

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

Recall that the function t^p ($p \in [0, 1]$) is operator monotone on $(0, \infty)$, and the following integral representation is well-known.

Lemma 1 ([7]). Let $0 < p < 1$ and A be a positive definite matrix. Then,

$$A^p = \int_0^\infty A(\lambda + A)^{-1} d\mu(\lambda) = \int_0^\infty \left(I - \lambda(\lambda + A)^{-1} \right) d\mu(\lambda) \quad (6)$$

The following integrals are elementary. So, we omit the proofs.

Lemma 2. Let $y > 0$, and $0 < p < 1$. Then we have

$$\frac{1}{\pi} \int_0^\infty \frac{\delta^{1/2}}{(\delta + y^2)^2} d\delta = \frac{y^{-1}}{2} \quad \text{and} \quad \frac{\sin p\pi}{\pi} \int_0^\infty \frac{\lambda^p d\lambda}{(\lambda + y)^2} = py^{p-1}.$$

Now we are ready to prove the first main result of this section. We show the α - z -Bures Wasserstein divergence defined in (5) is a quantum divergence.

Theorem 3. Let $0 \leq \alpha \leq z \leq 1$. Then the quantity

$$\Phi(X, Y) = \text{Tr}((1 - \alpha)X + \alpha Y) - \text{Tr}(Q_{\alpha, z}(X, Y)) \quad (X, Y > 0)$$

is a divergence.

Proof. For $z = \alpha$, the theorem was obtained in [5]. The case $0 \leq \alpha \leq z = 1$ was proved in a recent paper by Nguyen Lam and Phi Long Le in [14]. We need to consider the case $0 < \alpha < z < 1$.

Let $f(X, Y) = \text{Tr}((1 - \alpha)X + \alpha Y)$ and $g(X, Y) = \text{Tr}(Q_{\alpha, z}(X, Y))$. Since $z/\alpha > 1$, we have

$$\begin{aligned} \text{Tr}\left((X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}})^z\right) &= \text{Tr}\left((X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}})^{\frac{z}{\alpha}\alpha}\right) \\ &\leq \text{Tr}\left((X^{\frac{1-\alpha}{2\alpha}} Y X^{\frac{1-\alpha}{2\alpha}})^{\alpha}\right) \quad (\text{the Araki-Lieb-Thirring inequality}) \\ &\leq \text{Tr}((1 - \alpha)X + \alpha Y) \quad ([5, \text{Theorem 11}]). \end{aligned}$$

The equality occurs if and only if $X = Y$. So, $\Phi(X, Y)$ satisfies the first property in Definition 2.1.

Next, we need to verify that $D\Phi(A, X)|_{X=A} = 0$. We have

$$\frac{\partial f(X, Y)}{\partial Y}(B) = \text{Tr}(\alpha B). \quad (7)$$

Since $0 < z < 1$, we have

$$\frac{\partial Y^z}{\partial Y}(B) = \int_0^\infty \delta^z (\delta + Y)^{-1} B (\delta + Y)^{-1} \frac{\sin(z\pi)}{\pi} d\delta. \quad (8)$$

If we put $\varphi(X, Y) = X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}}$, then according to (8) and the Chain Rule,

$$\frac{\partial \varphi^z}{\partial Y}(B) = \int_0^\infty \delta^z (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + \varphi)^{-1} \frac{\sin(z\pi)}{\pi} d\delta.$$

Therefore,

$$\begin{aligned} \frac{\partial g(X, Y)}{\partial Y}(B) &= \text{Tr} \left[\int_0^\infty \delta^z (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + \varphi)^{-1} \frac{\sin(z\pi)}{\pi} d\delta \right] \\ &= \text{Tr} \left[\int_0^\infty \frac{\delta^z}{(\delta + \varphi)^2} \frac{\sin(z\pi)}{\pi} d\delta \left(\frac{\partial \varphi}{\partial Y}(B) \right) \right] \\ &= \text{Tr} \left[z \varphi^{z-1} \frac{\partial \varphi}{\partial Y}(B) \right] \quad (\text{By Lemma 1}) \\ &= \text{Tr} \left[z \left(X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}} \right)^{z-1} \frac{\partial \varphi}{\partial Y}(B) \right]. \end{aligned}$$

On the other hand,

$$\begin{aligned}\frac{\partial \varphi}{\partial Y}(B) &= X^{\frac{1-\alpha}{2z}} \frac{\partial Y^{\frac{\alpha}{z}}}{\partial Y}(B) X^{\frac{1-\alpha}{2z}} \\ &= X^{\frac{1-\alpha}{2z}} \left(\int_0^\infty \lambda^{\frac{\alpha}{z}} (\lambda + Y)^{-1} B (\lambda + Y)^{-1} \frac{\sin(\frac{\alpha}{z}\pi)}{\pi} d\lambda \right) X^{\frac{1-\alpha}{2z}}.\end{aligned}$$

When $Y = X$,

$$\begin{aligned}\left. \frac{\partial g(X, Y)}{\partial Y}(B) \right|_{Y=X} &= \text{Tr} \left(z Y^{\frac{z-1}{z}} Y^{\frac{1-\alpha}{2z}} \left(\int_0^\infty \lambda^{\frac{\alpha}{z}} (\lambda + Y)^{-1} B (\lambda + Y)^{-1} \frac{\sin(\frac{\alpha}{z}\pi)}{\pi} d\lambda \right) Y^{\frac{1-\alpha}{2z}} \right) \\ &= \text{Tr} \left(z Y^{\frac{z-\alpha}{z}} \int_0^\infty \lambda^{\frac{\alpha}{z}} (\lambda + Y)^{-1} B (\lambda + Y)^{-1} \frac{\sin(\frac{\alpha}{z}\pi)}{\pi} d\lambda \right) \\ &= \text{Tr} \left(z Y^{\frac{z-\alpha}{z}} \int_0^\infty \frac{\lambda^{\frac{\alpha}{z}}}{(\lambda + Y)^2} \frac{\sin(\frac{\alpha}{z}\pi)}{\pi} d\lambda B \right) \\ &= \text{Tr} \left(z Y^{\frac{z-\alpha}{z}} \frac{\alpha}{z} Y^{\frac{\alpha}{z}-1} B \right) \quad (\text{by Lemma 1}) \\ &= \text{Tr}(\alpha B).\end{aligned}$$

Therefore, on account of (7) from the last identity we have

$$\left. \frac{\partial \Phi(X, Y)}{\partial Y}(B) \right|_{Y=X} = \left. \frac{\partial f(X, Y)}{\partial Y}(B) \right|_{Y=X} - \left. \frac{\partial g(X, Y)}{\partial Y}(B) \right|_{Y=X} = 0.$$

Property (ii) in Definition 2.1 is fulfilled.

Finally, we check that for every Hermitian matrix B ,

$$\left. \frac{\partial^2 \Phi(X, Y)}{\partial Y^2} \right|_{Y=X} (B, B) \geq 0. \quad (9)$$

Since $\frac{\partial^2 f}{\partial Y^2}(B_1, B_2) = 0$, we only need to consider the second derivative of $g(X, Y)$ in Y . Let us recall

$$\frac{\partial g(X, Y)}{\partial Y}(B) = \text{Tr} \left[\int_0^\infty \delta^z (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + \varphi)^{-1} \frac{\sin(z\pi)}{\pi} d\delta \right].$$

By the product rule, we have

$$\begin{aligned} \frac{\partial^2 g}{\partial Y^2}(B_1, B_2) &= -\frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_2) (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_1) (\delta + \varphi)^{-1} \right] d\delta \\ &\quad + \frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + \varphi)^{-1} \frac{\partial^2 \varphi}{\partial Y^2}(B_1, B_2) (\delta + \varphi)^{-1} \right] d\delta \\ &\quad - \frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_1) (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_2) (\delta + \varphi)^{-1} \right] d\delta. \end{aligned}$$

Put

$$\begin{aligned} I_1 &= \frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_2) (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_1) (\delta + \varphi)^{-1} \right] d\delta, \\ I_3 &= \frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_1) (\delta + \varphi)^{-1} \frac{\partial \varphi}{\partial Y}(B_2) (\delta + \varphi)^{-1} \right] d\delta, \end{aligned}$$

and

$$I_2 = \frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + \varphi)^{-1} \frac{\partial^2 \varphi}{\partial Y^2}(B_1, B_2) (\delta + \varphi)^{-1} \right] d\delta.$$

When $Y = X$ and $B_1 = B_2 = B$, we have

$$I_1 = I_3 = \frac{\sin(z\pi)}{\pi} \int_0^\infty \delta^z \text{Tr} \left[(\delta + Y^{\frac{1}{z}})^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + Y^{\frac{1}{z}})^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + Y^{\frac{1}{z}})^{-1} \right] d\delta.$$

Since $(\delta + Y^{\frac{1}{z}})^{-1} \geq 0$, it is obvious that

$$(\delta + Y^{\frac{1}{z}})^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + Y^{\frac{1}{z}})^{-1} \frac{\partial \varphi}{\partial Y}(B) (\delta + Y^{\frac{1}{z}})^{-1} \geq 0.$$

Consequently, the integrals I_1 and I_3 are nonnegative. If we show that the integral I_2 is nonpositive, then we (9) is true. Indeed,

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial Y^2}(B_1, B_2) &= -\frac{\sin(\frac{\alpha}{z}\pi)}{\pi} \left(X^{\frac{1-\alpha}{2z}} \int_0^\infty \left(\lambda^{\frac{\alpha}{z}} (\lambda + Y)^{-1} B_2 (\lambda + Y)^{-1} B_1 (\lambda + Y)^{-1} \right. \right. \\ &\quad \left. \left. + \lambda^{\frac{\alpha}{z}} (\lambda + Y)^{-1} B_1 (\lambda + Y)^{-1} B_2 (\lambda + Y)^{-1} \right) d\lambda \right) X^{\frac{1-\alpha}{2z}}. \end{aligned}$$

When $Y = X$ and $B_1 = B_2 = B$, we have

$$\left. \frac{\partial^2 \varphi}{\partial Y^2} \right|_{Y=X} (B, B) = -\frac{2 \sin(\frac{\alpha}{z} \pi)}{\pi} \text{Tr} \left(Y^{\frac{1-\alpha}{z}} \int_0^\infty \lambda^{\frac{\alpha}{z}} (\lambda + Y)^{-1} B (\lambda + Y)^{-1} B (\lambda + Y)^{-1} d\lambda \right). \quad (10)$$

Since $(\lambda + Y)^{-1} \geq 0$, we have

$$(\lambda + Y)^{-1} B (\lambda + Y)^{-1} B (\lambda + Y)^{-1} \geq 0.$$

Consequently, the integral (10) is nonpositive. Therefore, $I_2 \leq 0$ and

$$\left. \frac{\partial^2 \Phi(X, Y)}{\partial Y^2} (B, B) \right|_{Y=X} = - \left. \frac{\partial^2 g(X, Y)}{\partial Y^2} (B, B) \right|_{Y=X} \geq 0.$$

Thus, $\Phi(A, B)$ is a quantum divergence on $\mathcal{P}_n$. $\square$

Now, let us consider the least squares problem with respect to the new quantum divergence. Let $A_1, A_2, \dots, A_m$ be positive definite matrices and $\omega = (\omega_1, \omega_2, \dots, \omega_m)$ be a probability vector. Consider the least squares problem as shown below:

$$\min_{X > 0} \sum_{i=1}^m \omega_i \Phi(X, A_i). \quad (11)$$

This problem was solved by Bhatia, Lim and Jain in [5] when $z = \alpha$.

Theorem 4. For $0 \leq \alpha \leq z \leq 1$, the function

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

attains minimum at X_0 , where X_0 is the unique positive definite solution of the following matrix equation

$$\sum_{i=1}^m \omega_i Q_{\alpha, z}(X, A_i) = X. \quad (12)$$

Proof. We have

$$\frac{\partial \Phi(A_i, X)}{\partial X} (B) = \text{Tr} \left(\alpha B - z \varphi^{z-1} \frac{\partial \varphi}{\partial X} (B) \right),$$

where $\varphi = \varphi(A_i, X) = A_i^{\frac{1-\alpha}{2z}} X^{\frac{\alpha}{z}} A_i^{\frac{1-\alpha}{2z}}$. Consequently,

$$\begin{aligned}
\frac{\partial F(X)}{\partial X}(B) &= \text{Tr} \left[\alpha B - \sum_{i=1}^m w_i z \varphi^{z-1} \frac{\partial \varphi}{\partial X}(B) \right] \\
&= \text{Tr} \left[\alpha B - \sum_{i=1}^m w_i z A_i^{\frac{1-\alpha}{2z}} \varphi^{z-1} A_i^{\frac{1-\alpha}{2z}} \int_0^\infty (\lambda + X)^{-1} B (\lambda + X)^{-1} \frac{\lambda^{\frac{\alpha}{z}} \sin(\frac{\alpha}{z} \pi)}{\pi} d\lambda \right] \\
&= \text{Tr} \left[\alpha B - B \int_0^\infty (\lambda + X)^{-1} C (\lambda + X)^{-1} \frac{\lambda^{\frac{\alpha}{z}} \sin(\frac{\alpha}{z} \pi)}{\pi} d\lambda \right] \\
&= \text{Tr} \left[B \left(\alpha I - \int_0^\infty (\lambda + X)^{-1} C (\lambda + X)^{-1} \frac{\lambda^{\frac{\alpha}{z}} \sin(\frac{\alpha}{z} \pi)}{\pi} d\lambda \right) \right],
\end{aligned}$$

where $C = \sum_{i=1}^m w_i z A_i^{\frac{1-\alpha}{2z}} \varphi^{z-1} A_i^{\frac{1-\alpha}{2z}}$. Therefore, the only critical point of $F(X)$ is the solution of the equation

$$\alpha I = \int_0^\infty (\lambda + X)^{-1} C (\lambda + X)^{-1} \frac{\lambda^{\frac{\alpha}{z}} \sin(\frac{\alpha}{z} \pi)}{\pi} d\lambda. \quad (13)$$

Now, let us choose an orthonormal basis in which the matrix X is diagonal, i.e., $X = \text{diag}(x_1, x_2, \dots, x_n)$ and let $C = (c_{ij})$ be the representation of C in this basis. From the equation (13) we have

$$\alpha \delta_{ij} = \int_0^\infty \frac{c_{ij}}{(\lambda + x_i)(\lambda + x_j)} \frac{\lambda^{\frac{\alpha}{z}} \sin(\frac{\alpha}{z} \pi)}{\pi} d\lambda.$$

From here, it implies that C is diagonal, and,

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

Therefore,

$$C = \sum_{i=1}^m w_i z A_i^{\frac{1-\alpha}{2z}} \varphi^{z-1} A_i^{\frac{1-\alpha}{2z}} = z X^{1-\frac{\alpha}{z}}.$$

Multiplying both sides of the last identity from the left and from the right by $X^{\frac{\alpha}{2z}}$, we get

$$X = \sum_{i=1}^m w_i X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{2z}} \left(A_i^{\frac{1-\alpha}{2z}} X^{\frac{\alpha}{z}} A_i^{\frac{1-\alpha}{2z}} \right)^{z-1} A_i^{\frac{1-\alpha}{2z}} X^{\frac{\alpha}{2z}}$$

$$\begin{aligned}
 &= \sum_{i=1}^m w_i X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{2z}} \left(A_i^{\frac{\alpha-1}{2z}} X^{-\frac{\alpha}{z}} A_i^{\frac{\alpha-1}{2z}} \right)^{1-z} A_i^{\frac{1-\alpha}{2z}} X^{\frac{\alpha}{2z}} \\
 &= \sum_{i=1}^m w_i X^{\frac{\alpha}{2z}} \left(A_i^{\frac{1-\alpha}{z}} \#_{1-z} X^{-\frac{\alpha}{z}} \right) X^{\frac{\alpha}{2z}} \\
 &= \sum_{i=1}^m w_i \left(X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \#_{1-z} I \right) \\
 &= \sum_{i=1}^m w_i \left(X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \right)^z.
 \end{aligned}$$

Thus, X satisfies the equation (12).

Finally, we show that the equation (12) has a unique solution. Firstly, notice that the function $F(X)$ is strictly convex [5]. Therefore, if the equation (12) has a solution, then it is unique. So, if we show that the function $F(X)$ has a fixed point, we finish the proof. Indeed, let a and b be positive numbers such that $aI \leq A_i \leq bI$, for all $1 \leq i \leq m$. We have

$$X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \geq X^{\frac{\alpha}{2z}} a^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \geq a^{\frac{1}{z}} I.$$

By the operator monotone of the map $x \mapsto x^z$, when $z \in [0, 1]$ we have

$$\left(X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \right)^z \geq aI.$$

Similarly, $\left(X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \right)^z \leq bI$. Therefore,

$$aI \leq F(X) = \sum_{i=1}^m \omega_i \left(X^{\frac{\alpha}{2z}} A_i^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}} \right)^z \leq bI.$$

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

$$\mathcal{K} = \{X \in \mathbb{P} : aI \leq X \leq bI\}.$$

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

3. Data processing inequality and in-betweenness property

In this section we show that the α - z -Bures Wasserstein divergence satisfies the Data Processing Inequality in quantum information theory. We also show that the matrix power mean satisfies the in-betweenness property in this divergence.

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

$$\Psi(\mathcal{E}(A), \mathcal{E}(B)) \leq \Psi(A, B).$$

Recall that (see, for examples, [19, Theorem 5.16]) if a map $\Psi(A, B)$ is jointly convex, unitarily invariant and invariant under tensor product, then Ψ is monotone with respect to all completely positive trace-preserving maps.

By the definition, the map $\Phi(A, B)$ is jointly convex. Indeed, according to [6], the trace function $\Theta_{p,q,s}(X, Y) = \text{Tr} \left[\left(X^{\frac{\alpha}{2}} Y^p X^{\frac{\alpha}{2}} \right)^s \right]$ is jointly concave if and only if $0 \leq p, q \leq 1$ and $0 \leq s \leq \frac{1}{p+q}$. For $q = \frac{1-\alpha}{z}$, $p = \frac{\alpha}{z}$, and $s = z$, we have that the function $\text{Tr} \left(Q_{\alpha,z}(X, Y) \right)$ is jointly concave. Hence, $\Phi(X, Y)$ is jointly convex. Therefore, from the following theorem it implies that the α - z -Bures Wasserstein divergence satisfies the Data Processing Inequality.

Theorem 5. $\Phi(X, Y)$ is invariant under all unitary matrix U and invariant under tensoring with another density matrix τ .

Proof. For an arbitrary unitary U , we have

$$\begin{aligned} & \Phi(U^* X U, U^* Y U) \\ &= \text{Tr} \left[(1-\alpha) U^* X U + \alpha U^* Y U - \left((U^* X U)^{\frac{1-\alpha}{2z}} (U^* Y U)^{\frac{\alpha}{z}} (U^* X U)^{\frac{1-\alpha}{2z}} \right)^z \right] \\ &= \text{Tr} \left[(1-\alpha) U^* X U + \alpha U^* Y U - \left(U^* X^{\frac{1-\alpha}{2z}} U U^* Y^{\frac{\alpha}{z}} U U^* X^{\frac{1-\alpha}{2z}} U \right)^z \right] \\ &= \text{Tr} \left[(1-\alpha) U^* X U + \alpha U^* Y U - \left(U^* X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}} U \right)^z \right] \\ &= \text{Tr} \left[(1-\alpha) U^* X U + \alpha U^* Y U - U^* \left(X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}} \right)^z U \right] \\ &= \text{Tr} \left[U^* \left((1-\alpha) X + \alpha Y - \left(X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}} \right)^z \right) U \right] \\ &= \text{Tr} \left((1-\alpha) X + \alpha Y - \left(X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}} \right)^z \right) \\ &= \Phi(X, Y) \end{aligned}$$

Also for an arbitrary density matrix τ , we have

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

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

To finish this paper, we show that the matrix power mean $\mu_p(t; A, B) = \left(tA^p + (1-t)B^p \right)^{\frac{1}{p}}$ satisfies the in-betweenness property with respect to the α - z -Bures Wasserstein divergence. We refer the interested readers to other papers for related results on in-betweenness [2,13,8–11].

Theorem 6. Let $A, B \in \mathcal{P}_n$, $0 < \alpha \leq z \leq 1$, $1/2 \leq p \leq 1$ and $\alpha \leq zp$. Then for any positive definite matrices A and B ,

$$\Phi(A, \mu_p) \leq \Phi(A, B). \quad (14)$$

Proof. We have

$$\Phi(A, \mu_p) = \text{Tr} \left[(1-\alpha)A + \alpha \mu_p - \left(A^{\frac{1-\alpha}{z}} \mu_p^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{z}} \right)^z \right],$$

and

$$\Phi(A, B) = \text{Tr} \left[(1-\alpha)A + \alpha B - \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right)^z \right].$$

Therefore, the inequality (14) is equivalent to the following

$$\begin{aligned}
&\text{Tr} \left[\alpha \left(tA^p + (1-t)B^p \right)^{\frac{1}{p}} - \left(A^{\frac{1-\alpha}{2z}} (tA^p + (1-t)B^p)^{\frac{\alpha}{zp}} A^{\frac{1-\alpha}{2z}} \right)^z \right] \\
&\leq \text{Tr} \left[\alpha B - \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right)^z \right].
\end{aligned}$$

Since $1 \leq \frac{1}{p} \leq 2$ when $\frac{1}{2} \leq p \leq 1$, the map $x \mapsto x^{\frac{1}{p}}$ is operator convex. Therefore,

$$\left(tA^p + (1-t)B^p \right)^{\frac{1}{p}} \leq tA + (1-t)B.$$

On the other hand, from the conditions $0 < \alpha \leq z$ and $\alpha \leq zp$ it implies that $0 \leq \frac{\alpha}{zp} \leq 1$.

By the operator concavity of the map $x \mapsto x^{\frac{\alpha}{zp}}$, we have

$$A^{\frac{1-\alpha}{2z}} \mu_p^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} = A^{\frac{1-\alpha}{2z}} \left(tA^p + (1-t)B^p \right)^{\frac{\alpha}{zp}} A^{\frac{1-\alpha}{2z}}$$

$$\begin{aligned}
&\geq A^{\frac{1-\alpha}{2z}} \left(tA^{\frac{\alpha}{z}} + (1-t)B^{\frac{\alpha}{z}} \right) A^{\frac{1-\alpha}{z}} \\
&= tA^{\frac{1}{z}} + (1-t)A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}}.
\end{aligned}$$

For $0 < z \leq 1$ the function t^z is operator concave on $(0, \infty)$. Then we have

$$\begin{aligned}
\left[A^{\frac{1-\alpha}{2z}} \left(tA^p + (1-t)B^p \right)^{\frac{\alpha}{zp}} A^{\frac{1-\alpha}{2z}} \right]^z &\geq \left[tA^{\frac{1}{z}} + (1-t)A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right]^z \\
&\geq tA + (1-t) \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right)^z.
\end{aligned}$$

Thus, the desired result follows if

$$Tr \left[\alpha tA + \alpha B - \alpha tB - tA - (1-t) \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right)^z \right] \leq Tr \left[\alpha B - \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right)^z \right].$$

Or, equivalently,

$$Tr \left[(1-\alpha)A + \alpha B - \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}} \right)^z \right] \geq 0$$

which was proved in Theorem 3. Thus, the matrix power mean μ_p satisfies the in-betweenness property in the α - z -Bures Wasserstein divergence. $\square$

Declaration of competing interest

There is no competing interest.

Acknowledgements

The authors would like to express their sincere gratitude to the anonymous referees for their useful comments and suggestions. The fourth author is partially supported by the Vietnam Ministry of Education and Training under grant number B2021-DQN-01.

References

- [1] K.M.R. Audenaert, N. Datta, α - z -Renyi relative entropies, *J. Math. Phys.* 56 (2) (2015) 022202.
- [2] K.M.R. Audenaert, In-betweenness, a geometrical monotonicity property for operator means, *Linear Algebra Appl.* 438 (2013) 1769–1778.
- [3] R. Bhatia, T. Jain, Y. Lim, On the Bures-Wasserstein distance between positive definite matrices, *Expo. Math.* 37 (2019) 165–191.
- [4] R. Bhatia, S. Gaubert, T. Jain, Matrix versions of the Hellinger distance, *Lett. Math. Phys.* 109 (8) (2019) 1777–1804.
- [5] R. Bhatia, T. Jain, Y. Lim, Strong convexity of sandwiched entropies and related optimization problems, *Rev. Math. Phys.* 30 (9) (2018) 1850014.
- [6] E.A. Carlen, R.L. Frank, E.H. Lieb, Some operator and trace function convexity theorems, *Linear Algebra Appl.* 490 (2016) 174–185.
- [7] E.H. Carlen, Trace inequalities and quantum entropy: an introductory course, in: *Entropy and the Quantum*, in: *Contemp. Math.*, vol. 529, 2010, pp. 73–140.

- [8] T.H. Dinh, R. Dumitru, J.A. Franco, Some geometric properties of matrix means in different distance functions, *Positivity* 24 (2020) 1419–1434.
- [9] T.H. Dinh, R. Dumitru, J.A. Franco, On the monotonicity of weighted power means for matrices, *Linear Algebra Appl.* 527 (2017) 128–140.
- [10] T.H. Dinh, B.K. Vo, T.Y. Tam, In-sphere property and reverse inequalities for matrix means, *Electron. J. Linear Algebra* 35 (1) (2019) 35–41.
- [11] T.H. Dinh, C.T. Le, B.K. Vo, T.D. Vuong, Weighted Hellinger distance and in-betweenness property, *Math. Inequal. Appl.* 24 (1) (2021) 157–165.
- [12] J. Franco, R. Dumitru, The Renyi power means of matrices, *Linear Algebra Appl.* 607 (2020) 45–57.
- [13] F. Franco, R. Dumitru, Generalized Hellinger metrics and Audenaert’s in-betweenness, *Linear Algebra Appl.* 585 (2020) 191–198.
- [14] N. Lam, P.L. Le, Quantum divergences with p -power means, *Linear Algebra Appl.* 609 (2021) 289–307.
- [15] E.H. Lieb, M.B. Ruskai, Proof of the strong subadditivity of quantum-mechanical entropy, With an appendix by B. Simon, *J. Math. Phys.* 14 (1973) 1938–1941.
- [16] Y. Lim, M. Palfia, Matrix power means and the Karcher mean, *J. Funct. Anal.* 262 (4) (2012) 1498–1514.
- [17] M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices, *SIAM J. Matrix Anal. Appl.* 26 (3) (2005) 735–747.
- [18] M. Wilde, A. Winter, D. Yang, Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Renyi relative entropy, *Commun. Math. Phys.* 331 (2014) 593–622.
- [19] M.M. Wolf, Quantum channels & operations: guided tour, Lecture notes available at <http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf>, July 2012.