MINISTRY OF EDUCATION AND TRAINING
QUY NHON UNIVERSITY

VUONG TRUNG DUNG

SOME DISTANCE FUNCTIONS IN QUANTUM INFORMATION THEORY AND RELATED PROBLEMS

DOCTORAL DISSERTATION IN MATHEMATICS

BINH DINH – 2024

MINISTRY OF EDUCATION AND TRAINING
QUY NHON UNIVERSITY

VUONG TRUNG DUNG

SOME DISTANCE FUNCTIONS IN QUANTUM INFORMATION THEORY AND RELATED PROBLEMS

Speciality: Mathematical Analysis

Speciality code: 9 46 01 02

Reviewer 1: Prof. Dang Duc Trong

Reviewer 2: Prof. Pham Tien Son

Reviewer 3: Assoc. Prof. Pham Quy Muoi

Supervisors:

1. Assoc. Prof. Dr. Le Cong Trinh
2. Assoc. Prof. Dr. Dinh Trung Hoa

BINH DINH – 2024

Declaration

This thesis was completed at the Department of Mathematics and Statistics, Quy Nhon University under the supervision of Assoc. Prof. Dr. Le Cong Trinh and Assoc. Prof. Dr. Dinh Trung Hoa. I hereby declare that the results presented in it are new and original. Most of them were published in peer-reviewed journals, others have not been published elsewhere. For using results from joint papers I have gotten permission from my co-authors.

Binh Dinh, 2024

Vuong Trung Dung

Acknowledgments

This thesis was undertaken during my years as a PhD student at the Department of Mathematics and Statistics, Quy Nhon University. Upon the completion of this thesis, I am deeply indebted to numerous individuals. On this occasion, I would like to extend my sincere appreciation to all of them.

First and foremost, I would like to express my sincerest gratitude to Assoc. Prof. Dr. Dinh Trung Hoa, who guided me into the realm of matrix analysis and taught me right from the early days. Not only that, but he also devoted a significant amount of valuable time to engage in discussions, and provided problems for me to solve. He motivated me to participate in workshops and establish connections with senior researchers in the field. He guided me to find enjoyment in solving mathematical problems and consistently nurtured my enthusiasm for my work. I can't envision having a more exceptional advisor and mentor than him.

The second person I would like to express my gratitude to is Assoc. Prof. Dr. Le Cong Trinh, who has been teaching me since my undergraduate days and also introduced me to Prof. Hoa. From the early days of sitting in lecture halls at university, Prof. Trinh has been instilling inspiration and a love for mathematics in me. It's fortunate that now I have the opportunity to be mentored by him once again. He has always provided enthusiastic support not only in my work but also in life. Without that dedicated support, it would have been difficult for me to complete this thesis.

I would like to extend a special thank you to the educators at both the Department of Mathematics and Statistic and the Department of Graduate Training at Quy Nhon University for providing the optimal environment for a postgraduate student who comes from a distant location like myself. Binh Dinh is also my hometown and the place where I have spent all my time from high school to university. The privilege and personal happiness of coming back to Quy Nhon University for advanced studies cannot be overstated.

I am grateful to the Board and Colleagues of VNU-HCM High School for the Gifted for providing me with much support to complete my PhD study. Especially, I would like to extend my heartfelt gratitude to Dr. Nguyen Thanh Hung, who has assisted to me in both material and spiritual aspects since the very first days I set foot in Saigon. He is not only a mentor and colleague but also a second father to me, who not only supported me financially and emotionally during challenging times but also constantly encouraged me to pursue a doctoral degree. Without this immense support and encouragement, I wouldn't be where I am today.

I also want to express my gratitude to Su for the wonderful time we've spent together, which has been a driving force for me to complete the PhD program and strive for even greater achievements that I have yet to attain.

Lastly, and most significantly, I would like to express my gratitude to my family. They have always been by my side throughout work, studies, and life. I want to thank my parents for giving birth to me and nurturing me to adulthood. This thesis is a gift I dedicate to them.

Binh Dinh, 2024

Vuong Trung Dung

Contents

Glossary of notation

$\mathbb{C}^n$	:	The set of all n-tuples of complex numbers
$\langle x, y\rangle$	:	The scalar product of vectors x and y
$\mathbb{M}_n$	:	The set of $n \times n$ complex matrices
$\mathcal{H}$	:	The Hilbert space
$B(\mathcal{H})$	:	The set of all bounded linear operators acting on Hilbert space $\mathcal{H}$
$\mathbb{H}_n$	:	The set of all $n \times n$ Hermitian matrices
$\mathbb{H}_n^+$	:	The set of all $n \times n$ positive semi-definite matrices
$\mathbb{P}_n$	:	The set of all $n \times n$ positive definite matrices
I, O	:	The identity and zero elements of $\mathbb{M}_n$, respectively
A^*	:	The conjugate transpose (or adjoint) of the matrix A
$\|A\|$	:	The positive semi-definite matrix $(A^*A)^{1/2}$
$\mathrm{Tr}(A)$	:	The canonical trace of matrix A
$\lambda(A)$	:	The vector of eigenvalues of matrix A in decreasing order
$s(A)$	:	The vector of singular values of matrix A in decreasing order
$Sp(A)$	:	The spectrum of matrix A
$\\|A\\|$	:	The operator norm of matrix A
$\|\|\|A\|\|\|$	:	The unitarily invariant norm of matrix A
$x \prec y$	:	x is majorized by y
$x \prec_w y$	:	x is weakly majorized by y
$A \sharp B$	:	The geometric mean of two matrices A and B

$A\sharp_t B$	:	The weighted geometric mean of two matrices A and B
$A\natural B$	:	The spectral geometric mean of two matrices A and B
$A\natural_t B$	:	The weighted spectral geometric mean of two matrices A and B
$F_t(A,B)$	:	The F-mean of two matrices A and B
$A\nabla B$	:	The arithmetic mean of two matrices A and B
$A!B$	:	The harmonic mean of two matrices A and B
$A:B$	:	The parallel sum of two matrices A and B
$\mu_p(A,B,t)$	:	The matrix p-power mean of matrices A and B

Introduction

Quantum information stands at the confluence of quantum mechanics and information theory, wielding the mathematical elegance of both realms to delve into the profound nature of information processing at the quantum level. In classical information theory, bits are the fundamental units representing 0 and 1. Quantum information theory, however, introduces the concept of qubits, the quantum counterparts to classical bits. Unlike classical bits, qubits can exist in a superposition of states, allowing them to be both 0 and 1 simultaneously. This unique property empowers quantum computers to perform certain calculations exponentially faster than classical computers.

Entanglement is a crucial phenomenon in quantum theory where two or more particles become closely connected. When particles are entangled, changing the state of one immediately affects the state of the other, no matter the distance between them. This has important implications for quantum information and computing, offering new possibilities for unique ways of handling information.

Quantum algorithms, such as Shor's algorithm for factoring large numbers and Grover's algorithm for quantum search, exemplify the power of quantum information in tackling complex computational tasks with unparalleled efficiency.

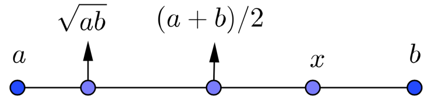
In order to treat information processing in quantum systems, it is necessary to mathematically formulate fundamental concepts such as quantum systems, states, and measurements, etc. Useful tools for researching quantum information are functional analysis and matrix theory. First, we consider the quantum system. It is described by a Hilbert space $\mathcal{H}$, which is called a *representation space*. This will be advantageous because it is not only the underlying basis of quantum mechanics but is also as helpful in introducing the special notation used for quantum mechanics. *The (pure) physical states* of the system correspond to unit vectors of the Hilbert space. This correspondence is not 1-1. When f_1 and f_2 are unit vectors, then the corresponding states are identical if $f_1 = zf_2$ for a complex number z of modulus 1 . Such z is often called *phase*. The pure physical state of the system determines a corresponding state vector up to a phase. Traditional quantum mechanics distinguishes between pure states and *mixed states*. Mixed states are described by *density matrices*. A density matrix or statistical operator is a positive matrix of trace 1 on the Hilbert space. This means that the space has a basis consisting of eigenvectors of the statistical operator and the sum of eigenvalues is 1. In quantum information theory, distance functions are used to measure the distance between two mixed states. Additionally, these distance functions can be employed to characterize the properties of a given quantum state. For instance, they can quantify the quantum entanglement between two parts of a state, representing the shortest distance between the state and the set of all separable states. These distance functions naturally extend to the set of positive semi-definite matrices, which is also the main focus of this thesis.

Nowadays, the significance of matrix theory has been widely recognized across various fields, including engineering, probability and statistics, quantum information, numerical analysis, biological and social sciences. In image processing (subdivision schemes), medical imaging (MRI), radar signal processing, statistical biology (DNA/genome), and machine learning, data from numerous experiments are stored as positive definite matrices. To work with each set of data, we need to select its representative element. In other words, we need to compute the average of the corresponding positive definite matrices. Therefore, considering global solutions of the least-squares problems for matrices is of paramount importance (refer to [2, 8, 18, 28, 67, 73] for examples).

Let $0 < a \leq x \leq b$. Consider the following least squares problem:

$$d^2(x, a) + d^2(x, b) \to \min, \quad x \in [a, b],$$

where $d := d_E(x,y) = |y-x|$, or, $d := d_R(x,y) := |\log(y) - \log(x)|$.



The arithmetic mean $(a+b)/2$ and the geometric mean $\sqrt{ab}$ are unique solutions to the above problem with respect to d_E and d_R distance, respectively. Moreover, based on the AM-GM inequality for two non-negative numbers a and b, we have a new distance as follows

$$d(a,b) = \frac{a+b}{2} - \sqrt{ab}.$$

For $A, B \in \mathbb{P}_n$, some matrix analogs of scalar distances are:

- Euclidean distance induced from Euclidean/Frobenius inner product $\langle A, B\rangle = \mathrm{Tr}(A^*B)$. The associated norm is $\|A\|_F = \langle A, A\rangle^{1/2} = (\mathrm{Tr}(A^*A))^{1/2}$.
- The Riemann distance [12] is $\delta_R(A,B) = ||\log(A^{-1}B)||_2 = \left(\sum_{i=1}^{n} \log^2 \lambda_i(A^{-1}B)\right)^{1/2}$.
- Bures-Wasserstein distance [13] in the theory of optimal transport :

$$d_b(A,B) = \left(\mathrm{Tr}(A+B) - 2\,\mathrm{Tr}\left(\left(A^{1/2}BA^{1/2}\right)^{1/2}\right)\right)^{1/2}.$$

- The Log-Determinant metric [75] in machine learning and quantum information:

$$d_l(A,B) = \log\det\frac{A+B}{2} - 2\log\det(AB).$$

- The Hellinger metric or Bhattacharya metric [73] in quantum information :

$$d_h(A,B) = \left(\mathrm{Tr}(A+B) - 2\,\mathrm{Tr}\left(A^{1/2}B^{1/2}\right)\right)^{1/2}.$$

In applications, one are sometimes 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 [11] are such distance-like functions .

Definition. A smooth function $\Phi : \mathbb{P}_n \times \mathbb{P}_n \to \mathbb{R}^+$ 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.,

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

Some divergences that have recently received a lot of attention are in [11, 14, 35, 56].

Now let us revisit the scalar mean theory which serves as a starting point for our next problem in this thesis.

A *scalar mean* of non-negative numbers is a function from $\mathbb{R}^+ \times \mathbb{R}^+$to $\mathbb{R}^+$such that:

1) $M(x, x) = x$ for every $x \in \mathbb{R}^+$.

2) $M(x, y) = M(y, x)$ for every $x, y \in \mathbb{R}^+$.

3) If $x < y$, then $x < M(x, y) < y$.

4) If $x < x_0$ and $y < y_0$, then $M(x, y) < M(x_0, y_0)$.

5) $M(x, y)$ is continuous.

6) $M(tx, ty) = tM(x, y)$ for $t, x, y \in \mathbb{R}^+$.

A two-variable function $M(x, y)$ satisfying condition 6) can be reduced to a one-variable function $f(x) := M(1, x)$. Namely, $M(x, y)$ is recovered from f as $M(x, y) = xf(x^{-1}y)$.

Notice that the function f, corresponding to M is monotone increasing on $\mathbb{R}^{+}$. And this relation forms a one-to-one correspondence between means and monotone increasing functions on $\mathbb{R}^{+}$.

The following are some desired properties of any object that is called a *"mean"* M on $\mathbb{H}_n^{+}$.

(A1). Positivity: $A, B \geqslant 0 \Rightarrow M(A, B) \geqslant 0$.

(A2). Monotonicity: $A \geqslant A', B \geqslant B' \Rightarrow M(A, B) \geqslant M(A', B')$.

(A3). Positive homogeneity: $M(kA, kB) = kM(A, B)$ for $k \in \mathbb{R}^{+}$.

(A4). Transformer inequality: $X^* M(A, B) X \leqslant M(X^* A X, X^* B X)$ for $X \in B(\mathcal{H})$.

(A5). Congruence invariance: $X^* M(A, B) X = M(X^* A X, X^* B X)$ for invertible $X \in B(\mathcal{H})$.

(A6). Concavity: $M(tA + (1-t)B, tA' + (1-t)B') \geqslant tM(A, A') + (1-t)M(B, B')$ for $t \in [0, 1]$.

(A7). Continuity from above: if $A_n \downarrow A$ and $B_n \downarrow B$, then $M(A_n, B_n) \downarrow M(A, B)$.

(A8). Betweenness: if $A \leqslant B$, then $A \leqslant M(A, B) \leqslant B$.

(A9). Fixed point property: $M(A, A) = A$.

To study matrix or operator means in general, we must first consider three classical means in mathematics: arithmetic, geometric, and harmonic means. These means are defined in the following manner, respectively,

$$A \nabla B = \frac{1}{2}(A + B),$$

$$A \sharp B = A^{1/2} \left(A^{-1/2} B A^{-1/2}\right)^{1/2} A^{1/2},$$

and

$$A ! B = 2(A^{-1} + B^{-1})^{-1}.$$

In the above definitions, if matrix A is not invertible, we replace A with $A_\epsilon = A + \epsilon I$ and then let ϵ tend to 0 (similarly for matrix B). It can be seen that the arithmetic, harmonic and geometric means share the properties (A1)-(A9) in common. In 1980, Kubo and Ando [54] developed an axiomatic theory of operator mean on $\mathbb{H}_n^+$. At first, they defined a *connection* of two matrices as follows (the term "connection" comes from the study of electrical network connections).

Definition. A connection on $\mathbb{H}_n^+$ is a binary operation σ on $\mathbb{H}_n^+$ satisfying the following axioms for all $A, A', B, B', C \in \mathbb{H}_n^+$:

(M1). Monotonicity: $A \leqslant A', B \leqslant B' \Longrightarrow A\sigma B \leqslant A'\sigma B'$.

(M2). Transformer inequality: $C(A\sigma B)C \leqslant (CAC)\sigma(CBC)$.

(M3). Joint continuity from above: if $A_n, B_n \in B(\mathcal{H})^+$satisfy $A_n \downarrow A$ and $B_n \downarrow B$, then $A_n\sigma B_n \downarrow A\sigma B$.

A *mean* is a connection with normalization condition

(M4) $I\sigma I = I$.

To each connection σ corresponds its *transpose* σ' defined by $A\sigma' B = B\sigma A$. A connection σ is *symmetric* by definition if $\sigma = \sigma'$. The *adjoint* of σ, denoted by σ^*, is defined by $A\sigma^* B = \left(A^{-1}\sigma B^{-1}\right)^{-1}$, for invertible A, B. When σ is a non-zero connection, its *dual*, in symbol $\sigma^\perp$, is defined by $\sigma^\perp = (\sigma')^* = (\sigma^*)'$.

However, Kubo-Ando theory of means still has many limitations. In applied and engineering fields, people need more classes of means that are non Kubo-Ando. For some non Kubo-Ando means we refer the interested readers to [17, 23, 25, 35, 37].

One of the famous non-Kubo-Ando means is the spectral geometric mean [37], denoted as $A\natural B$, introduced in 1997 by Fiedler and Pták . It is called the spectral geometric mean because $(A\sharp B)^2$ is similar to AB and that the eigenvalues of their spectral mean are the positive square roots of the corresponding eigenvalues of AB. In 2015, Kim and Lee [52] defined the weighted spectral mean:

$$A\natural_t B := \left(A^{-1}\sharp B\right)^t A \left(A^{-1}\sharp B\right)^t, \quad t \in [0,1].$$

In this thesis we focus on two problems:

1. **Distance function generated by operator means.** We introduce some new distance on the set of positive definite matrices in the relation to operator means, and their applications. In addition, we also study some geometric properties for means such as the in-betweenness property, and data processing inequality in quantum information.

2. **A new weighted spectral geometric mean.** We introduce a new weighted spectral geometric mean, denoted by $\mathcal{F}_t(A, B)$ and study basic properties for this quantity. We also establish a weak log-majorization relation involving $\mathcal{F}_t(A, B)$ and the Lie-Trotter formula for $\mathcal{F}_t(A, B)$.

The main tools in our research are the spectral theorem for Hermitian matrices and the theory of Kubo-Ando means. Some fundamental techniques in the theory of operator monotone functions and operator convex functions are also utilized in the dissertation. We also employ basic knowledge in matrix theory involving unitarily invariant norms, trace, etc.

The main results in this thesis are presented in the following articles:

1. Vuong T.D., Vo B.K (2020), "An inequality for quantum fidelity", *Quy Nhon Univ. J. Sci.*, 4 (3).

2. Dinh T.H., Le C.T., Vo B.K, Vuong T.D. (2021), "Weighted Hellinger distance and in betweenness property", *Math. Ine. Appls.,* 24, 157-165.

3. Dinh T.H., Le C.T., Vo B.K., Vuong T.D. (2021), "The α-z-Bures Wasserstein divergence", *Linear Algebra Appl.,* 624, 267-280.

4. Dinh T.H., Le C.T., Vuong T.D., α-z-fidelity and α-z-weighted right mean, *Submitted*.

5. Dinh T.H., Tam T.Y., Vuong T.D, On new weighted spectral geometric mean, *Submitted*.

They were presented on the seminars at the Department of Mathematics and Statistics at Quy Nhon University and at the following international workshops and conferences as follows:

1. First SIBAU-NU Workshop on Matrix Analysis and Linear Algebra, 15-17 October, 2021.

2. 20th Workshop on Optimization and Scientific Computing, April 21-23, 2022 - Ba Vi, Vietnam.

3. International Workshop on Matrix Analysis and Its Applications, June 4, 2022, Quy Nhon, Viet Nam.

4. The second international workshop on Matrix Theory and Applications, AKFA University, November, 2022.

5. International Workshop on Matrix Analysis and Its Applications, July 7-8, 2023, Quy Nhon, Viet Nam.

6. 10th Viet Nam Mathematical Congress, August 8-12, 2023, Da Nang, Viet Nam.

This thesis has introduction, three chapters, conclusion, further investigation, a list of the author's papers related to the thesis and preprints related to the topics of the thesis, and a list of references.

The introduction provides a background on the topics covered in this work and explains why they are meaningful and relevant. It also briefly summarizes the content of the thesis and highlights the main results from the main three chapters.

In the first chapter, the author collects some basic preliminaries which are used in this thesis.

In the second chapter, we introduce the weighted Hellinger distance for matrices which is an interpolating between the Euclidean distance and the Hellinger distance. In 2019, Minh [43] introduced the Alpha Procrustes distance as follows: For $\alpha > 0$, and for positive semi-definite matrices A and B,

$$d_{b,\alpha} = \frac{1}{\alpha} d_b \left(A^{2\alpha}, B^{2\alpha}\right).$$

In this chapter, by employing this approach, we define a new distance called the Weighted Hellinger distance as follows:

$$d_{h,\alpha}(A, B) = \frac{1}{\alpha} d_h \left(A^{2\alpha}, B^{2\alpha}\right)$$

and then studied its properties. In the first section of this chapter, we show that the weighted Hellinger distance, as α tends to zero, is exactly the Log-Euclidean distance (Proposition 2.1.1), that is for two positive semi-definite matrices A and B,

$$\lim_{\alpha \to 0} d_{h,\alpha}^2(A, B) = || \log(A) - \log(B)||_F^2.$$

Afterwards, in Proposition 2.1.2 we demonstrate the equivalence between the weighted Hellinger distance and the Alpha Procrustes distance, it means

$$d_{b,\alpha}(A, B) \leq d_{h,\alpha}(A, B) \leq \sqrt{2} d_{b,\alpha}(A, B).$$

We say that a matrix mean σ satisfies the *in-betweenness property* with respect to the metric d if for any pair of positive definite operators A and B,

$$d(A, A\sigma B) \leq d(A, B).$$

In the second section, we prove that the matrix power mean $\mu_p(t, A, B) = (tA^p + (1-t)B^p)^{1/p}$ satisfies the in-betweenness property in the weighted Hellinger and Alpha Procrustes distances (Theorem 2.2.1 and Theorem 2.2.2). At the end of this chapter, we prove that if σ is a symmetric mean and satisfies the in-betweenness property with respect to the Alpha Procrustes distance or the Weighted Hellinger distance, then it can only be the arithmetic mean (Theorem 2.2.3).

In chapter 3, we study a new quantum divergence so-called the α-z-Bures Wasserstein divergence. In 2015, Audenaert and Datta [7] introduced the Rényi power mean of matrices via the matrix function $P_{\alpha,z}(A, B) = \left(B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}}\right)^z$. Based on this quantity, in this chapter, the α-z-Bures Wasserstein divergence for positive semi-definite matrices A and B is defined by

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

where $Q_{\alpha,z}(A, B) = \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}}\right)^z$. Then we prove that this quantity is a quantum divergence (Theorem 3.1.1) We also solve the least square problem with respect to $\Phi(A, B)$ and showed that the solution of this problem is exactly the unique positive definite solution of the matrix equation $\sum_{i=1}^{m} w_i Q_{\alpha,z}(X, A_i) = X$ (Theorem 3.1.2). In [49], M. Jeong and co-authors investigated this solution and denoted it by $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$-called α-z-weighted right mean. In this thesis, we continue our study of this quantity and obtain some new results. An important result is an inequality for $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$, which can be considered a version of the AM-GM inequality (Theorem 3.1.3). Hwang and Kim [48] proved that for any weighted m-mean $\mathcal{G}_m$ between arithmetic mean and geometric mean, the function $\mathcal{G}_m^\omega := \mathcal{G}_m(\omega, \cdot) : \mathbb{P}^m \to \mathbb{P}$ is differentiable at $\mathbb{I} = (I, ..., I)$ with

$$D\mathcal{G}_n^\omega(\mathbb{I})(X_1, ..., X_m) = \sum_{j=1}^{m} w_j X_j.$$

Notice that the α-z-weighted right mean does not satisfy the above condition. However, we do have a similar result for $\mathcal{R}_{\alpha,z}^\omega := \mathcal{R}_{\alpha,z}(\omega, \cdot)$ (Theorem 3.1.4). The well-known Lie-Trotter formula [76] states that for $X, Y \in \mathbb{M}_n$,

$$\lim_{n\to+\infty} \Big(\exp(\frac{X}{n}) \exp(\frac{Y}{n}) \Big)^n = \exp(X+Y).$$

This formula plays an essential role in the development of Lie Theory, and frequently appears in different research fields [44, 47, 48]. In [48], J.Hwang and S.Kim introduced the multivariate Lie-Trotter mean on the convex cone $\mathbb{P}_n$ of positive definite matrices. For a positive probability vector $\omega = (w_1, ..., w_m)$ and differentiable curves $\gamma_1, ..., \gamma_m$ on $\mathbb{P}_n$ with $\gamma_i(0) = I$ ($i = 1, \cdots, m$), a weighted m-mean $\mathcal{G}_m$ (for $m \geq 2$) is the multivariate Lie-Trotter mean if

$$\lim_{s\to 0} \mathcal{G}_m(\omega, \gamma_1(s), \gamma_2(s), , ..., \gamma_m(s))^{1/s} = \exp\Big(\sum_{j=1}^{m} w_j \gamma_j'(0) \Big).$$

In the end of this section, we prove that $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ is a multivariate Lie-Trotter mean (Theorem 3.1.5) . In the second section of this chapter, we show that this divergence satisfies the data processing inequality (DPI) in quantum information (Theorem 3.2.1). The data processing inequality is an information-theoretic concept that states that the information content of a signal cannot be increased via a local physical operation. This can be expressed concisely as "post-processing cannot increase information", that is, for any completely positive trace preserving map $\mathcal{E}$ and any positive semi-definite matrices A and B,

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

Furthermore, we show that the matrix power mean $\mu(t, A, B) = ((1-t)A^p + tB^p)^{1/p}$ satisfies the in-betweenness property with respect to the α-z-Bures Wasserstein divergence (Theorem 3.2.2). Quantum fidelity is an important quantity in quantum information theory and quantum chaos theory. It is a distance measure between density matrices, which are considered as quantum states. Although it is not a metric, it has many useful properties that can be used to define a metric on the space of density matrices. In the next section, we give some properties of quantum fidelity and its extended version. An important results is we establish some variational principles for the quantum α-z-fidelity

$$f_{\alpha,z}(\rho, \sigma) := \operatorname{Tr}\left(\rho^{\alpha/2z}\sigma^{(1-\alpha)/z}\rho^{\alpha/2z}\right)^z = \operatorname{Tr}\left(\sigma^{(1-\alpha)/2z}\rho^{\alpha/z}\sigma^{(1-\alpha)/2z}\right)^z,$$

where ρ and σ are two postitive definite matrices (Theorem 3.3.4). That is, it is the extremal value of two matrix functions

$$P(X) = z \operatorname{Tr}\left(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X\right) - (z-1)\operatorname{Tr}\left(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}},$$

and

$$Q(X) = \Big(\operatorname{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X)\Big)^z . \Big(\operatorname{Tr}(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2x}})^{\frac{z}{z-1}}\Big)^{1-z}.$$

Let $U(\mathbb{H})$ be the set of $n \times n$ unitary matrices, and $\mathcal{D}_n$ the set of density matrices. For $\rho \in \mathcal{D}_n$, its unitary orbit is defined as

$$U_\rho = \{U\rho U^* : U \in U(\mathbb{H})\}.$$

In the last section we are going to obtain the maximum and minimum distance between orbits of two state ρ and σ in $\mathcal{D}_n$ via the quantum α-z-fidelity and prove that the set of these distance is a close interval in $\mathbb{R}^+$ (Theorem 2.4.2 and Theorem 3.4.3)

In chapter 4, we introduce a new weighted spectral geometric mean

$$\mathcal{F}_t(A, B) = (A^{-1}\sharp_t B)^{1/2} A^{2-2t} (A^{-1}\sharp_t B)^{1/2}, \quad t \in [0, 1],$$

where A and B are positive definite matrices. We study basic properties and inequalities for $\mathcal{F}_t(A, B)$. An important property that we obtain in this chapter is that $\mathcal{F}_t(A, B)$ satisfies the Lie-Trotter formula (Theorem 4.2.1).

At the end of this chapter, we compare the weak-log majorization between the $\mathcal{F}$-mean and the Wasserstein mean, which is the solution to the least square problem with respect to the Bures distance or Wasserstein distance (Theorem 4.2.3).

Chapter 1

Preliminaries

1.1 Matrix theory fundamentals

Let $\mathbb{N}$ be the set of all natural numbers. For each $n \in \mathbb{N}$, we denote by $\mathbb{M}_n$ the set of all $n \times n$ complex matrices, $\mathbb{H}_n$ is the set of all $n \times n$ Hermitian matrices, $\mathbb{H}_n^+$ is the set of $n \times n$ positive semi-definite matrices, $\mathbb{P}_n$ is the cone of positive definite matrices in $\mathbb{M}_n$, and $\mathcal{D}_n$ is the set of density matrices which are the positive definite matrices with trace equal to one. Denote by I and O the identity and zero elements of $\mathbb{M}_n$, respectively. This thesis deals with problems for matrices, which are operators in finite-dimensional Hilbert spaces $\mathcal{H}$. We will indicate if the case is infinite-dimensional.

Recall that for two vectors $x = (x_j), y = (y_j) \in \mathbb{C}^n$ the inner product $\langle x, y \rangle$ of x and y is defined as $\langle x, y \rangle \equiv \sum_j x_j \bar{y}_j$. Now let A be a matrix in $\mathbb{M}_n$, the conjugate transpose or the adjoint A^* of A is the complex conjugate of the transpose A^T. We have, $\langle Ax, y \rangle = \langle x, A^* y \rangle$.

Definition 1.1.1. A matrix $A = (a_{ij})_{i,j=1}^n \in \mathbb{M}_n$ is said to be:

(i) diagonal if $a_{ij} = 0$ when $i \neq j$.

(ii) invertible if there exists an matrix B of order $n \times n$ such that $AB = I_n$. In this situation A has a unique inverse matrix $A^{-1} \in \mathbb{M}_n$ such that $A^{-1}A = AA^{-1} = I_n$.

(iii) normal if $AA^* = A^*A$.

(iv) unitary if $AA^* = A^*A = I_n$.

(v) Hermitian if $A = A^*$.

(vi) positive semi-definite if $\langle Ax, x\rangle \geq 0$ for all $x \in \mathbb{C}^n$.

(vii) positive definite if $\langle Ax, x\rangle > 0$ for all $x \in \mathbb{C}^n \backslash \{0\}$.

Definition 1.1.2 (Löwner's Order, [86])**.** Let A and B be two Hermitian matrices of same order n. We say that $A \geq B$ if and only if $A - B$ is a positive semi-definite matrix.

Definition 1.1.3. A complex number λ is said to be an eigenvalue of a matrix A corresponding to its non-zero eigenvector x if

$$Ax = \lambda x.$$

The multiset of the eigenvalues of A is denoted by $Sp(A)$ and called the spectrum of A.

There are several conditions that characterize positive matrices. Some of them are listed in theorem below [10].

Proposition 1.1.1.

(i) *A is positive semi-definite if and only if it is Hermitian and all its eigenvalues are nonnegative. Moreover, A is positive definite if and only if it is Hermitian and all its eigenvalues are positive.*

(ii) *A is positive semi-definite if and only if it is Hermitian and all its principal minors are nonnegative. Moreover, A is positive definite if and only if it is Hermitian and all its principal minors are positive.*

(iii) *A is positive semi-definite if and only if $A = B^*B$ for some matrix B. Moreover, A is positive definite if and only if B is nonsingular.*

(iv) *A is positive semi-definite if and only if $A = T^*T$ for some upper triangular matrix T. Further, T can be chosen to have nonnegative diagonal entries. If A is positive definite, then T is unique. This is called the Cholesky decomposition of A. Moreover, A is positive definite if and only if T is nonsingular.*

(v) *A is positive semi-definite if and only if $A = B^2$ for some positive matrix B. Such a B is unique. We write $B = A^{1/2}$ and call it the (positive) square root of A. Moreover, A is positive definite if and only if B is positive definite.*

(vi) *A is positive semi-definite if and only if there exist $x_1, \ldots, x_n$ in $\mathcal{H}$ such that*

$$a_{ij} = \langle x_i, x_j \rangle .$$

A is positive definite if and only if the vectors $x_j, 1 \leq j \leq n$, are linearly independent.

Let $A \in \mathbb{M}_n$, we denote the eigenvalues of A by $\lambda_j(A)$, for $j = 1, 2, ..., n$. For a matrix $A \in \mathbb{M}_n$, the notation $\lambda(A) \equiv (\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A))$ means that $\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A)$. The *absolute value* of matrix $A \in \mathbb{M}_n$ is the square root of matrix A^*A and denoted by

$$|A| = (A^*A)^{\frac{1}{2}}.$$

We call the eigenvalues of $|A|$ by the *singular value* of A and denote as $s_j(A)$, for $j = 1, 2, ..., n$. For a matrix $A \in \mathbb{M}_n$, the notation $s(A) \equiv (s_1(A), s_2(A), \ldots, s_n(A))$ means that $s_1(A) \geq s_2(A) \geq \ldots \geq s_n(A)$.

There are some basic properties of the spectrum of a matrix.

Proposition 1.1.2. *Let $A, B \in \mathbb{M}_n$, then*

(i) *$Sp(AB) = Sp(BA)$.*

(ii) *If A is a Hermitian matrix then $Sp(A) \subset \mathbb{R}$.*

(iii) *A is a positive semi-definite (respectively positive definite) if and only if A is a Hermitian matrix and $Sp(A) \subset \mathbb{R}_{\geq 0}$ (respectively $Sp(A) \subset \mathbb{R}^+$).*

(iv) *If* $A, B \geq 0$ *then* $Sp(AB) \subset \mathbb{R}^+$.

The *trace* of a matrix $A = (a_{ij}) \in \mathbb{M}_n$, denoted by $\mathrm{Tr}(A)$, is the sum of all diagonal entries, or, we often use the sum of all eigenvalues $\lambda_i(A)$ of A, i.e.,

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i(A)$$

Related to the trace of the matrix, we recall the Araki-Lieb-Thirring trace inequality [18] used consistently throughout the thesis.

Theorem 1.1.1. *Let* A *and* B *be two positive semi-definite matrices, and let* $q > 0$, *we have*

$$\mathrm{Tr}\left[\left(B^{\frac{r}{2}} A^r B^{\frac{r}{2}}\right)^{\frac{q}{r}}\right] \leq \mathrm{Tr}\left[\left(B^{\frac{1}{2}} A B^{\frac{1}{2}}\right)^{q}\right], \text{ if } r \in (0,1],$$

and

$$\mathrm{Tr}\left[\left(B^{\frac{r}{2}} A^r B^{\frac{r}{2}}\right)^{\frac{q}{r}}\right] \geq \mathrm{Tr}\left[\left(B^{\frac{1}{2}} A B^{\frac{1}{2}}\right)^{q}\right], \text{ if } r \geq 1.$$

The *determinant* of A is denoted and defined by

$$\det(A) = \sum_{\rho \in \mathbb{S}_n} \left(\mathrm{sgn}(\rho) \prod_{i=1}^{n} a_{i\rho_i} \right) = \prod_{j=1}^{n} \lambda_j.$$

where $\mathbb{S}_n$ is the set of all permutations ρ of the set $\mathbb{S} = \{1, 2, \dots, n\}$.

Proposition 1.1.3. *Let* $A, B \in \mathbb{H}_n$ *with* $\lambda(A) = (\lambda_1, \lambda_2, \dots, \lambda_n)$ *and* $\lambda(B) = (\mu_1, \mu_2, \dots, \mu_n)$. *Then*

(i) *If* $A > 0$ *and* $B > 0$, *then* $A \geq B$ *if and only if* $B^{-1} \geq A^{-1}$.

(ii) *If* $A \geq B$, *then* $X^*AX \geq X^*BX$ *for every* $X \in \mathbb{M}_n$.

(iii) *If* $A \geq B$, *then* $\lambda_j \geq \mu_j$ *for each* $j = 1, 2, \dots, n$.

(iv) *If* $A \geq B \geq 0$, *then* $\mathrm{Tr}(A) \geq \mathrm{Tr}(B) \geq 0$.

(v) If $A \geq B \geq 0$, then $\det(A) \geq \det(B) \geq 0$.

A function $\|\cdot\| : \mathbb{M}_n \to \mathbb{R}$ is said to be a matrix norm if for all $A, B \in \mathbb{M}_n$ and $\forall \alpha \in \mathbb{C}$ we have:

(i) $\|A\| \geq 0$.

(ii) $\|A\| = 0$ if and only if $A = 0$.

(iii) $||\alpha A|| = |\alpha| \cdot ||A||$.

(iv) $\|A + B\| \leq \|A\| + \|B\|$.

In addition, a matrix norm is said to be sub-multiplicative matrix norm if

$$\|AB\| \leq \|A\| \cdot \|B\|.$$

A matrix norm is said to be a *unitarily invariant norm* if for every $A \in \mathbb{M}_n$, we have $\|UAV\| = \|A\|$ for all $U, V \in \mathbb{U}_n$ unitary matrices. It is denoted as $|||\cdot|||$.

These are some important norms over $\mathbb{M}_n$.

The operator norm of A, defined by

$$|||A|||_{op} = \sqrt{\lambda_1\left(A^*A\right)} = s_1(A).$$

The Ky Fan k-norm is the sum of all singular values, i.e.,

$$\|A\|_k = \sum_{i=1}^{k} s_i(A).$$

The Schatten p-norm is defined as

$$\|A\|_p = \left(\sum_{i=1}^{n} s_i^p(A)\right)^{1/p}.$$

When $p=2$, we have the Frobenius norm or sometimes called the Hilbert-Schmidt norm :

$$\|A\|_2 = \left(\operatorname{Tr}|A|^2\right)^{1/2} = \left(\sum_{j=1}^{n} s_j^2(A)\right)^{1/2}.$$

Let $x=(x_1,x_2,\ldots,x_n)$ and $y=(y_1,y_2,\ldots,y_n)$ be in $\mathbb{R}^n$. Let $x^{\downarrow}=\left(x_{[1]},\, x_{[2]},\ldots,x_{[n]}\right)$ denote a rearrangement of the components of x such that $x_{[1]}\geqslant x_{[2]}\geqslant\ldots\geqslant x_{[n]}$. We say that x is majorized by y, denoted by $x\prec y$, if

$$\sum_{i=1}^{k} x_{[i]} \leqslant \sum_{i=1}^{k} y_{[i]}, \quad k=1,2,\ldots,n-1, \quad \text{and} \quad \sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}.$$

We say that x is weakly majorized by y if $\sum\limits_{i=1}^{k} x_{[i]} \leqslant \sum\limits_{i=1}^{k} y_{[i]}, k=1,2,\ldots,n$, denoted by $x\prec_w y$. If $x>0$ (i.e., $x_i>0$ for $i=1,\ldots,n$) and $y>0$, we say that x is $\log$-majorized by y, denoted by $x\prec_{\log} y$, if

$$\prod_{i=1}^{k} x_{[i]} \leqslant \prod_{i=1}^{k} y_{[i]}, \quad k=1,2,\ldots,n-1, \quad \text{and} \quad \prod_{i=1}^{n} x_{[i]} = \prod_{i=1}^{n} y_{[i]}.$$

In other words, $x\prec_{\log} y$ if and only if $\log x\prec\log y$.

Matrix $P\in\mathbb{M}_n$ is called a *projection* if $P^2=P$. One says that P is a *Hermitian projection* if it is both Hermitian and a projection; P is an *orthogonal projection* if the range of P is orthogonal to its null space. The partial ordering is very simple for projections. If P and Q are projections, then the relation $P\leq Q$ means that the range of P is included in the range of Q. An equivalent algebraic formulation is $PQ=P$. The largest projection in $\mathbb{M}_n$ is the identity I and the smallest one is 0 . Therefore $0\leq P\leq I$ for any projection $P\in\mathbb{M}_n$. Assume that P and Q are projections on the same Hilbert space. Among the projections which are smaller than P and Q there is a maximal projection, denoted by $P\wedge Q$, which is the orthogonal projection onto the intersection of the ranges of P and Q.

Theorem 1.1.2. *[45] Assume that P and Q are orthogonal projections. Then*

$$P \wedge Q = \lim_{n\to\infty}(PQP)^n = \lim_{n\to\infty}(QPQ)^n.$$

1.2 Matrix function and matrix mean

Now let us recall the spectral theorem which is one of the most important tools in functional analysis and matrix theory.

Theorem 1.2.1 (Spectral decomposition, [9])**.** *Let $\lambda_1 > \lambda_2 \ldots > \lambda_k$ be eigenvalues of a Hermitian matrix A. Then*

$$A = \sum_{j=1}^{k} \lambda_j P_j,$$

where P_j is the orthogonal projection onto the subspace spanned by the eigenvectors associated to the eigenvalue λ_j.

For a real-valued function f defined on some interval $K \subset \mathbb{R}$, and for a self-adjoint matrix $A \in \mathbb{M}_n$ with spectrum in K, the matrix $f(A)$ is defined by means of the functional calculus, i.e.,

$$A = \sum_{j=1}^{k} \lambda_j P_j \quad \Longrightarrow \quad f(A) := \sum_{j=1}^{k} f(\lambda_j) P_j.$$

Or, if $A = U \operatorname{diag}(\lambda_1, \ldots, \lambda_n) U^*$ is a spectral decomposition of A (where U is some unitary), then

$$f(A) := U \operatorname{diag}(f(\lambda_1), \cdots, f(\lambda_n)) U^*.$$

We are now at the stage where we will discuss matrix/operator functions. Löwner was the first to study operator monotone functions in his seminal papers [63] in 1930. In the same time, Kraus investigated the notion operator convex function [55].

Definition 1.2.1 ([63])**.** A continuous function f defined on an interval $K (K \subset \mathbb{R})$ is said to be *operator monotone of order* n on K if for two Hermitian matrices A and B in $\mathbb{M}_n$ with spectras in K, one has

$$A \le B \quad \text{implies} \quad f(A) \le f(B).$$

If f is operator monotone of any orders then f is called *operator monotone*.

Theorem 1.2.2 (Löwner-Heinz's Inequality, [86]). *The function $f(t) = t^r$ is operator monotone on $[0, \infty)$ for $0 \le r \le 1$. More specifically, for two positive semi-definite matrices such that $A \le B$. Then*

$$A^r \le B^r, \quad 0 \le r \le 1.$$

Definition 1.2.2 ([55]). A continuous function f defined on an interval $K (K \subset \mathbb{R})$ is said to be *operator convex of order n* on K if for any Hermitian matrices A and B in $\mathbb{M}_n$ with spectra in K, and for all real numbers $0 \le \lambda \le 1$,

$$f(\lambda A + (1-\lambda)B) \le \lambda f(A) + (1-\lambda) f(B).$$

If f is operator convex of any order n then f is called *operator convex*. If $-f$ is operator convex then we call f is operator concave.

Theorem 1.2.3 ([10]). *Function $f(t) = t^r$ in $[0, \infty)$ is operator convex when $r \in [-1, 0] \cup [1, 2]$. More specifically, for any positive semi-definite matrices A, B and for any $\lambda \in [0, 1]$,*

$$(\lambda A + (1-\lambda)B)^r \le \lambda A^r + (1-\lambda) B^r.$$

Another important example is the function $f(t) = \log t$, which is operator monotone on $(0, \infty)$ and the function $g(t) = t \log t$ is operator convex. The relations between operator monotone and operator convex via the theorem below.

Theorem 1.2.4 ([9]). *Let f be a (continuous) real function on the interval $[0, \alpha)$. Then the following two conditions are equivalent:*

(i) *f is operator convex and $f(0) \le 0$.*

(ii) The function $g(t) = \dfrac{f(t)}{t}$ *is operator monotone on* $(0, \alpha)$.

Definition 1.2.3 ([10])**.** Let $f(A, B)$ be a real valued function of two matrix variables. Then, f is called *jointly concave*, if for all $0 \leq \alpha \leq 1$,

$$f(\alpha A_1 + (1-\alpha)A_2, \alpha B_1 + (1-\alpha)B_2) \geq \alpha f(A_1, B_1) + (1-\alpha) f(A_2, B_2)$$

for all A_1, A_2, B_1, B_2. If $-f$ is jointly concave, we say f is *jointly convex*.

We will review very quickly some basic concepts of the Fréchet differential calculus, with special emphasis on matrix analysis. Let X, Y be real Banach spaces, and let $\mathcal{L}(X, Y)$ be the space of bounded linear operators from X to Y. Let U be an open subset of X. A continuous map f from U to Y is said to be differentiable at a point u of U if there exists $T \in \mathcal{L}(X, Y)$ such that

$$\lim_{v \to 0} \frac{\|f(u+v) - f(u) - Tv\|}{\|v\|} = 0.$$

It is clear that if such a T exists, it is unique. If f is differentiable at u, the operator T above is called the derivative of f at u. We will use for it the notation $Df(u)$, of $\partial f(u)$. This is sometimes called the Fréchet derivative. If f is differentiable at every point of U, we say that it is differentiable on U. One can see that, if f is differentiable at u, then for every $v \in X$,

$$Df(u)(v) = \left.\frac{d}{dt}\right|_{t=0} f(u+tv).$$

This is also called the directional derivative of f at u in the direction v.

If f_1, f_2 are two differentiable maps, then $f_1 + f_2$ is differentiable and

$$D\left(f_1 + f_2\right)(u) = Df_1(u) + Df_2(u).$$

The composite of two differentiable maps f and g is differentiable and we have the chain rule

$$D(g \circ f)(u) = Dg(f(u)) \cdot Df(u).$$

One important rule of differentiation for real functions is the product rule: $(fg)' = f'g + gf'$. If f and g are two maps with values in a Banach space, their product is not defined - unless the range is an algebra as well. Still, a general product rule can be established. Let f, g be two differentiable maps from X into Y_1, Y_2, respectively. Let B be a continuous bilinear map from $Y_1 \times Y_2$ into Z. Let φ be the map from X to Z defined as $\varphi(x) = B(f(x), g(x))$. Then for all u, v in X

$$D\varphi(u)(v) = B(Df(u)(v), g(u)) + B(f(u), Dg(u)(v)).$$

This is the product rule for differentiation. A special case of this arises when $Y_1 = Y_2 = \mathcal{L}(Y)$, the algebra of bounded operators in a Banach space Y. Now $\varphi(x) = f(x)g(x)$ is the usual product of two operators. The product rule then is

$$D\varphi(u)(v) = [Df(u)(v)] \cdot g(u) + f(u) \cdot [Dg(u)(v)]$$

Higher order Fréchet derivatives can be identified with multilinear maps. Let f be a differentiable map from X to Y. At each point u, the derivative $Df(u)$ is an element of the Banach space $\mathcal{L}(X, Y)$. Thus we have a map Df from X into $\mathcal{L}(X, Y)$, defined as $Df : u \to Df(u)$. If this map is differentiable at a point u, we say that f is twice differentiable at u. The derivative of the map Df at the point u is called the second derivative of f at u. It is denoted as $D^2 f(u)$. This is an element of the space $\mathcal{L}(X, \mathcal{L}(X, Y))$. Let $\mathcal{L}_2(X, Y)$ be the space of bounded bilinear maps from $X \times X$ into Y. The elements of this space are maps f from $X \times X$ into Y that are linear in both variables, and for whom there exists a constant c such that

$$\|f(x_1, x_2)\| \leq c \|x_1\| \|x_2\|$$

for all $x_1, x_2 \in X$. The infimum of all such c is called $\|f\|$. This is a norm on the space $\mathcal{L}_2(X, Y)$, and the space is a Banach space with this norm. If φ is an element of $\mathcal{L}(X, \mathcal{L}(X, Y))$, let

$$\tilde{\varphi}(x_1, x_2) = [\varphi(x_1)](x_2) \text{ for } x_1, x_2 \in X.$$

Then $\tilde{\varphi} \in \mathcal{L}_2(X, Y)$. It is easy to see that the $\text{map}\, \varphi \to \tilde{\varphi}$ is an isometric isomorphism. Thus the second derivative of a twice differentiable map f from X to Y can be thought of as a bilinear map from $X \times X$ to Y. It is easy to see that this map is symmetric in the two variables; i.e.,

$$D^2 f(u)\,(v_1, v_2) = D^2 f(u)\,(v_2, v_1)$$

for all u, v_1, v_2. Derivatives of higher order can be defined by repeating the above procedure. The p th derivative of a map f from X to Y can be identified with a p-linear map from the space $X \times X \times \cdots \times X$ (p copies) into Y. A convenient method of calculating the p th derivative of f is provided by the formula

$$D^p f(u)\,(v_1, \ldots, v_p) = \left.\frac{\partial^p}{\partial t_1 \cdots \partial t_p}\right|_{t_1 = \cdots = t_p = 0} f\left(u + t_1 v_1 + \cdots + t_p v_p\right).$$

For the convenience of readers, let us provide some examples for the derivatives of matrices.

Example 1.2.1. In these examples $X = Y = \mathcal{L}(\mathcal{H})$.

(i) Let $f(A) = A^2$. Then

$$Df(A)(B) = AB + BA,$$

and

$$\left[D^2 f(A)\right](B_1, B_2) = B_1 B_2 + B_2 B_1.$$

(ii) Let $f(A) = A^{-1}$ for each invertible A. Then

$$Df(A)(B) = -A^{-1} B A^{-1},$$

and

$$\left[D^2 f(A)\right](B_1, B_2) = A^{-1} B_1 A^{-1} B_2 A^{-1} + A^{-1} B_2 A^{-1} B_1 A^{-1}.$$

(iii) Let $f(A) = A^{-2}$ for each invertible A. Then

$$Df(A)(B) = -A^{-1}BA^{-2} - A^{-2}BA^{-1},$$

and

$$\begin{aligned}\left[D^2 f(A)\right](B_1, B_2) =& A^{-2}B_1A^{-1}B_2A^{-1} + A^{-2}B_2A^{-1}B_1A^{-1} \\ &+ A^{-1}B_1A^{-2}B_2A^{-1} + A^{-1}B_2A^{-2}B_1A^{-1} \\ &+ A^{-1}B_1A^{-1}B_2A^{-2} + A^{-1}B_2A^{-1}B_1A^{-2}\end{aligned}$$

(iv) Let $f(A) = A^*A$. Then

$$Df(A)(B) = A^*B + B^*A,$$

and

$$D^2 f(A)(B_1, B_2) = B_1^*B_2 + B_2^*B_1.$$

In connections with electrical engineering, Anderson and Duffin [3] defined the *parallel sum* of two positive definite matrices A and B by

$$A : B = \left(A^{-1} + B^{-1}\right)^{-1}.$$

The harmonic mean is $2(A : B)$ which is the dual of the *arithmetic mean* $A\nabla B = \dfrac{A+B}{2}$. In this period time, Pusz and Woronowicz [69] introduced the *geometric mean* as

$$A\sharp B := A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2}.$$

They also proved that the geometric mean is the unique positive solution of the Riccati equation

$$XA^{-1}X = B.$$

In 2005, Moakher [65] conducted a study, and then in 2006, Bhatia and Holbrook [14] investigated the structure of the Riemannian manifold $\mathbb{H}_n^+$. They showed that the curve

$$\gamma(t) = A\sharp_t B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^t A^{1/2} \quad (t \in [0,1])$$

is the unique geodesic joining A and B, and called t-geometric mean or weighted geometric mean. The weighted harmonic and the weighted arithmetic means are defined by

$$A!_t B = \left(tA^{-1} + (1-t)B^{-1}\right)^{-1},$$

and

$$A\nabla_t B = tA + (1-t)B.$$

The well-known inequality related to these quantities is the harmonic, geometric, and arithmetic means inequality [47, 60] , that is,

$$A!_t B \leq A\sharp_t B \leq A\nabla_t B.$$

These three means are Kubo-Ando means. Let's collect the main content of the Kubo-Ando means theory in the general case [54]. For $x > 0$ and $t \geq 0$, the function $\phi(x,t) = \dfrac{x(1+t)}{x+t}$ is bounded and continuous on the extended half-line $[0,\infty]$. The Löwner theory ([9, 45]) on operator-monotone functions states that the map $m \mapsto f$, defined by

$$f(x) = \int_{[0,\infty]} \phi(x,t)dm(t) \text{ for } x > 0,$$

establishes an affine isomorphism from the class of positive Radon measures on $[0,\infty]$ onto the class of operator-monotone functions. In the representation abvove, $f(0) = \inf_x f(x) = m(\{0\})$ and $\inf_x f(x)/x = m(\{\infty\})$.

Theorem 1.2.5. *[Kubo-Ando] For each operator connection σ, there exists a unique operator* *monotone function* $f : \mathbb{R}^+ \to \mathbb{R}^+$*, satisfying*

$$f(t)I_n = I_n\sigma(tI_n), t > 0,$$

and for $A, B > 0$ *the formula*

$$A\sigma B = A^{\frac{1}{2}} f(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}}$$

holds, with the right hand side defined via functional calculus, and extended to $A, B \geq 0$ *as follows*

$$A\sigma B = \lim_{\epsilon \to 0}(A + \epsilon I_n)\sigma(B + \epsilon I_n).$$

We call f *the representing function of* σ*.*

The next theorem follows from the integral representation of matrix monotone functions and from the previous theorem.

Theorem 1.2.6. *The map,* $m \mapsto \sigma$*, defined by*

$$A\sigma B = aA + bB + \int_{(0,\infty)} \frac{1+t}{t}\{(tA) : B\} dm(t)$$

where

$$a = m(\{0\}) \text{ and } b = m(\{\infty\}),$$

establishes an affine isomorphism from the class of positive Radon measures on $[0, \infty]$ *onto the class of connections.*

If P and Q are two projections, then the explicit formulation for $P\sigma Q$ is simpler.

Theorem 1.2.7. *If* σ *is a mean, then for every pair of projections* P *and* Q

$$P\sigma Q = a(P - P \wedge Q) + b(Q - P \wedge Q) + P \wedge Q,$$

where

$$a = 1\sigma 0 \quad \textit{and} \quad b = \lim_{x\to\infty} (1\sigma x)/x.$$

An immediate consequence of the above theorem is the following relation for projections P and Q

$$P!Q = P \wedge Q \quad \text{and} \quad P\#Q = P \wedge Q.$$

Let f be the representing function of σ. Since $xf(x^{-1})$ is the representing function of the transpose σ', then σ is symmetric if and only if $f(x) = xf\left(x^{-1}\right)$. The next theorem gives the representation for a symmetric connection.

Theorem 1.2.8. *The map, $n \mapsto \sigma$, defined by*

$$A\sigma B = \frac{c}{2}(A+B) + \int_{(0,1]} \frac{1+t}{2t}\{(tA) : B + A : (tB)\}dn(t)$$

where $c = n(\{0\})$, establishes an affine isomorphism from the class of positive Radon measures on the unit interval $[0,1]$ onto the class of symmetric connections.

Chapter 2

Weighted Hellinger distance

In recent years, many researchers have paid attention to different distance functions on the set $\mathbb{P}_n$ of positive definite matrices. Along with the traditional Riemannian metric $d_R(A, B) = \left(\sum_{i=1}^{n} \log^2 \lambda_i(A^{-1}B)\right)^{1/2}$ (where $\lambda_i(A^{-1}B)$ are eigenvalues of the matrix $A^{-1/2}BA^{-1/2}$), there are other important functions. Two of them are the Bures-Wasserstein distance [13], which are adapted from the theory of optimal transport :

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

and the Hellinger metric or Bhattacharya metric [11] in quantum information :

$$d_h(A, B) = \left(\operatorname{Tr}(A + B) - 2\operatorname{Tr}(A^{1/2}B^{1/2})\right)^{1/2}.$$

Notice that the metric d_h is the same as the Euclidean distance between $A^{1/2}$ and $B^{1/2}$, i.e., $\|A^{1/2} - B^{1/2}\|_F$.

Recently, Minh [43] introduced the Alpha Procrustes distance as follows: For $\alpha > 0$ and for two positive semi-definite matrices A and B,

$$d_{b,\alpha} = \frac{1}{\alpha} d_b(A^{2\alpha}, B^{2\alpha}).$$

He showed that the Alpha Procrustes distances are the Riemannian distances corresponding to a family of Riemannian metrics on the manifold of positive definite matrices, which encompass both the Log-Euclidean and Wasserstein Riemannian metrics. Since the Alpha Procrustes distances are defined based on the Bures-Wasserstein distance, we also call them the *weighted Bures-Wasserstein distances*. In that flow, in this chapter we can define the *weighted Hellinger metric* for two positive semi-definite matrices as follows:

$$d_{h,\alpha}(A,B) = \frac{1}{\alpha} d_h(A^{2\alpha}, B^{2\alpha}),$$

then investigate its properties within this framework.

The results of this chapter are taken from [32].

2.1 Weighted Hellinger distance

Definition 2.1.1. For two positive semi-definite matrices A and B and for $\alpha > 0$, the weighted Hellinger distance between A and B is defined as

$$d_{h,\alpha}(A,B) = \frac{1}{\alpha} d_h(A^{2\alpha}, B^{2\alpha}) = \frac{1}{\alpha}(\operatorname{Tr}(A^{2\alpha} + B^{2\alpha}) - 2\operatorname{Tr}(A^{\alpha}B^{\alpha}))^{\frac{1}{2}}. \tag{2.1.1}$$

It turns out that $d_{h,\alpha}(A,B)$ is an interpolating metric between the Log-Euclidean and the Hellinger metrics. We start by showing that the limit of the weighted Hellinger distance as α tends to 0 is the Log-Euclidean distance. We also show that the weighted Bures-Wasserstein and weighted Hellinger distances are equivalent (Proposition 2.1.2).

Proposition 2.1.1. *For two positive semi-definite matrices A and B,*

$$\lim_{\alpha \to 0} d^2_{h,\alpha}(A,B) = ||\log(A) - \log(B)||_F^2.$$

Proof. We rewrite the expression of $d_{h,\alpha}(A,B)$ as

$$
\begin{aligned}
d^2_{h,\alpha}(A,B) &= \frac{1}{\alpha^2} d^2_h(A^{2\alpha}B^{2\alpha}) \\
&= \frac{1}{\alpha^2}\Big[\operatorname{Tr}\Big(A^{2\alpha}+B^{2\alpha}-2A^{\alpha}B^{\alpha}\Big)\Big] \\
&= \frac{||A^{\alpha}-I||_F^2}{\alpha^2}+\frac{||B^{\alpha}-I||_F^2}{\alpha^2}-\frac{2}{\alpha^2}\operatorname{Tr}\Big(A^{\alpha}B^{\alpha}-A^{\alpha}-B^{\alpha}+I\Big).
\end{aligned}
$$

We have

$$
\lim_{\alpha\to 0}\frac{||A^{\alpha}-I||_F^2}{\alpha^2}=||\log A||_F^2, \quad \lim_{\alpha\to 0}\frac{||B^{\alpha}-I||_F^2}{\alpha^2}=||\log B||_F^2.
$$

Since

$$
A^{\alpha}=\exp(\alpha\log A)=I+\alpha\log A+\frac{\alpha^2}{2!}(\log A)^2+\cdots,
$$

$$
B^{\alpha}=\exp(\alpha\log B)=I+\alpha\log B+\frac{\alpha^2}{2!}(\log B)^2+\cdots,
$$

we have

$$
A^{\alpha}B^{\alpha}=I+\alpha(\log A+\log B)+\frac{\alpha^2}{2}\Big((\log A)^2+(\log B)^2+2\log A.\log B\Big)+\cdots.
$$

Therefore,

$$
A^{\alpha}B^{\alpha}-A^{\alpha}-B^{\alpha}+I=\alpha^2\log A\cdot\log B+\cdots.
$$

Consequently,

$$
\begin{aligned}
d^2_{h,\alpha}(A,B) &= \frac{||A^{\alpha}-I||_F^2}{\alpha^2}+\frac{||B^{\alpha}-I||_F^2}{\alpha^2}-2\operatorname{Tr}(\log A.\log B) \\
&= \frac{||A^{\alpha}-I||_F^2}{\alpha^2}+\frac{||B^{\alpha}-I||_F^2}{\alpha^2}-2\Big\langle\log A,\log B\Big\rangle_F.
\end{aligned}
$$

Tending α to zero, we obtain

$$
d^2_{h,\alpha}(A,B)=||\log A||_F^2+||\log B||_B^2-2\Big\langle\log A,\log B\Big\rangle_F=||\log A-\log B||_F^2.
$$

This completes the proof. $\square$

It is interesting to note that the weighted Bures-Wasserstein and weighted Hellinger distances are equivalent.

Proposition 2.1.2. *For two positive semi-definite matrices A and B,*

$$d_{b,\alpha}(A,B) \leq d_{h,\alpha}(A,B) \leq \sqrt{2}d_{b,\alpha}(A,B).$$

Proof. According the Araki-Lieb-Thirring inequality [43] , we have

$$\mathrm{Tr}(A^{1/2}BA^{1/2})^r \geq \mathrm{Tr}(A^rB^r), \quad |r| \leq 1.$$

Replace A with $A^{2\alpha}$, B with $B^{2\alpha}$ and r with $\frac{1}{2}$ we obtain the following

$$\mathrm{Tr}(A^\alpha B^{2\alpha}A^\alpha)^{1/2} \geq \mathrm{Tr}(A^\alpha B^\alpha).$$

Thus,

$$\frac{1}{\alpha^2}\,\mathrm{Tr}\left(A^{2\alpha} + B^{2\alpha} - 2(A^\alpha B^{2\alpha}A^\alpha)^{1/2}\right) \leq \frac{1}{\alpha^2}\,\mathrm{Tr}\left(A^{2\alpha} + B^{2\alpha} - 2A^\alpha B^\alpha\right).$$

In other words,

$$d_{b,\alpha}(A,B) \leq d_{h,\alpha}(A,B).$$

With $\rho, \sigma \in \mathcal{D}_n$, we have

$$d_h^2(\rho,\sigma) = 2 - 2\ \mathrm{Tr}(\rho^{1/2}\sigma^{1/2}) \leq 4 - 4\ \mathrm{Tr}((\rho^{1/2}\sigma\rho^{1/2})^{1/2}) = 2d_b^2(\rho,\sigma),$$

or,

$$2\,\mathrm{Tr}((\rho^{1/2}\sigma\rho^{1/2})^{1/2}) \leq 1 + \mathrm{Tr}(\rho^{1/2}\sigma^{1/2}).$$

In the above inequality replace ρ with $\frac{A^{2\alpha}}{\mathrm{Tr}(A^{2\alpha})}$ and σ with $\frac{B^{2\alpha}}{\mathrm{Tr}(B^{2\alpha})}$ we have

$$
\begin{aligned}
2\,\mathrm{Tr}\left[(A^{\alpha}B^{2\alpha}A^{\alpha})^{1/2}\right] &\leq \mathrm{Tr}(A^{2\alpha})^{1/2}\,\mathrm{Tr}(B^{2\alpha})^{1/2} + \mathrm{Tr}(A^{\alpha}B^{\alpha}) \\
&\leq \frac{1}{2}\,\mathrm{Tr}(A^{2\alpha} + B^{2\alpha}) + \mathrm{Tr}(A^{\alpha}B^{\alpha}).
\end{aligned}
$$

It follows that

$$4\,\mathrm{Tr}[(A^{\alpha}B^{2\alpha}A^{\alpha})^{1/2}] \leq \mathrm{Tr}(A^{2\alpha} + B^{2\alpha}) + 2\,\mathrm{Tr}(A^{\alpha}B^{\alpha}).$$

The above inequality is equivalent to

$$2[\mathrm{Tr}(A^{2\alpha} + B^{2\alpha} - 2\,\mathrm{Tr}(A^{\alpha}B^{2\alpha}A^{\alpha})^{1/2}] \geq \mathrm{Tr}(A^{2\alpha} + B^{2\alpha} - 2A^{\alpha}B^{\alpha}),$$

or,

$$d^2_{h,\alpha}(A, B) \leq 2d^2_{b,\alpha}(A, B).$$

Consequently,

$$d_{h,\alpha}(A, B) \leq \sqrt{2}d_{b,\alpha}(A, B).$$

□

2.2 In-betweenness property

In 2016, Audenaert [5] introduced the in-betweenness property of matrix means . We say that a matrix mean σ satisfies the *in-betweenness property* with respect to the metric d if for any pair of positive definite operators A and B,

$$d(A, A\sigma B) \leq d(A, B).$$

In [34], the authors introduced and studied the in-sphere property of matrix means. Dinh, Franco and Dumitru also published several papers [26, 28] on geometric properties of the matrix power mean $\mu_p(t; A, B) := (tA^p + (1-t)B^p)^{1/p}$ with respect to different distance functions. They also considered the case of the matrix power mean in the sense of Kubo-Ando [54] which is defined as

$$P_p(t, A, B) = A^{1/2}\left(tI + (1-t)(A^{-1/2}BA^{-1/2})^p\right)^{1/p} A^{1/2}.$$

In this section, we focus our study on the in-betweenness properties of the matrix power means with respect to the weighted Bures-Wasserstein and weighted Hellinger distances. As a consequence of the equivalence, using the operator convexity and concavity of the power functions, we show that the matrix power mean satisfies the in-betweenness property with respect to $d_{h,\alpha}$ (Theorem 2.2.1) and $d_{b,\alpha}$ (Theorem 2.2.2). We also show that among symmetric means, the arithmetic mean is the only one that satisfies the in-betweenness property in the weighted Bures-Wasserstein and weighted Hellinger distances.

Now we are ready to show that the matrix power means $\mu_p(t; A, B)$ satisfy the in-betweenness property in $d_{h,\alpha}$ and $d_{b,\alpha}$.

Theorem 2.2.1. *Let* $0 < p/2 \leq \alpha \leq p$ *and* $0 \leq t \leq 1$. *Then*

$$d_{h,\alpha}(A, \mu_p(t; A, B)) \leq d_{h,\alpha}(A, B),$$

for all $A, B \in \mathbb{H}_n^+$.

Proof. We have

$$d^2_{h,\alpha}(A, \mu_p(t; A, B)) = \frac{1}{\alpha^2}\operatorname{Tr}\Big(A^{2\alpha} + \mu_p^{2\alpha} - 2A^{\alpha}\mu_p^{\alpha}(t; A, B)\Big),$$

and

$$d^2_{h,\alpha}(A, B) = \frac{1}{\alpha^2}\operatorname{Tr}\Big(A^{2\alpha} + B^{2\alpha} - 2A^{\alpha}B^{\alpha}\Big).$$

Therefore, the above result follows if

$$\operatorname{Tr}\Big(\mu_p^{2\alpha}(t; A, B) - 2A^{\alpha}\mu_p^{\alpha}(t; A, B)\Big) \leq \operatorname{Tr}\Big(B^{2\alpha} - 2A^{\alpha}B^{\alpha}\Big).$$

By the operator convexity of the map $x \mapsto x^{2\alpha/p}$, when $\frac{p}{2} \leq \alpha \leq p$,

$$\mu_p^{2\alpha}(t; A, B) = \Big(tA^p + (1-t)B^p\Big)^{2\alpha/p} \leq tA^{2\alpha} + (1-t)B^{2\alpha}.$$

Thus, the desired result follows if

$$\operatorname{Tr}\Big[t\Big(A^{2\alpha} - B^{2\alpha}\Big) - 2A^\alpha \mu_p^\alpha(t; A, B)\Big] \leq -2\operatorname{Tr}(A^\alpha B^\alpha).$$

By the operator concavity of the map $x \mapsto x^{\alpha/p}$, when $\frac{p}{2} \leq \alpha \leq p$,

$$\mu_p^{\alpha}(t; A, B) = \Big(tA^p + (1-t)B^p\Big)^{\alpha/p} \geq tA^{\alpha} + (1-t)B^{\alpha}.$$

Therefore, the distance monotonicity follows if

$$\operatorname{Tr}\Big[t(A^{2\alpha} - B^{2\alpha}) - 2A^\alpha\Big(tA^\alpha + (1-t)B^\alpha\Big)\Big] \leq -2\operatorname{Tr}(A^\alpha B^\alpha),$$

or

$$t\operatorname{Tr}\Big(A^{2\alpha} + B^{2\alpha} - 2A^\alpha B^\alpha\Big) \geq 0,$$

which is from AM-GM inequality. $\square$

Theorem 2.2.2. *Let* $0 < p/2 \leq \alpha \leq p$ *and* $1/2 \leq t \leq 1$. *Then,*

$$d_{b,\alpha}(A, \mu_p(t; A, B)) \leq d_{b,\alpha}(A, B),$$

for all $A, B \in \mathbb{H}_n^+$.

Proof. Firstly, we show that for any positive semi-definite matrices A and B, for $p/2 \leq \alpha \leq p$ and $1/2 \leq t \leq 1$,

$$d_{b,\alpha}(A, \mu_p(t; A, B)) \leq d_{h,\alpha}(A, \mu_p(t; A, B)) \leq \sqrt{1-t}\, d_{h,\alpha}(A, B).$$

By the Araki-Lieb-Thirring inequality, we have

$$\operatorname{Tr}\left(A^\alpha B^{2\alpha} A^\alpha\right)^{1/2} \geq \operatorname{Tr}\left(A^\alpha B^\alpha\right).$$

Therefore,

$$\begin{aligned} d_{b,\alpha}^2(A, \mu_p(t; A, B)) &= \frac{1}{\alpha^2} d_b(A^{2\alpha}, \mu_p^{2\alpha}(t; A, B)) \\ &= \frac{1}{\alpha^2} \operatorname{Tr}\left(A^{2\alpha} + \mu_p^{2\alpha}(t; A, B) - 2(A^\alpha \mu_p^{2\alpha}(t; A, B) A^\alpha)^{1/2}\right) \\ &\leq \frac{1}{\alpha^2} \operatorname{Tr}\left(A^{2\alpha} + \mu_p^{2\alpha}(t; A, B) - 2A^\alpha \mu_p^{\alpha}(t; A, B)\right). \end{aligned}$$

By the operator convexity of the function $x \mapsto x^{2\alpha/p}$ and the operator concavity of the function $x \mapsto x^{\alpha/p}$, we obtain

$$\begin{aligned} d_{b,\alpha}^2(A, \mu_p(t; A, B)) &\leq \frac{1}{\alpha^2} \operatorname{Tr}\left[A^{2\alpha} + tA^{2\alpha} + (1-t)B^{2\alpha} - 2A^\alpha\left(tA^\alpha + (1-t)B^\alpha\right)\right] \\ &= \frac{1-t}{\alpha^2} \operatorname{Tr}\left(A^{2\alpha} + B^{2\alpha} - 2A^\alpha B^\alpha\right) \\ &= (1-t) d_{h,\alpha}^2(A, B). \end{aligned}$$

From here, applying the square root function to both sides with $t \in [\frac{1}{2}, 1]$, we have

$$d_{b,\alpha}(A, \mu_p(t; A, B)) \leq \sqrt{1-t}\, d_{h,\alpha}(A, B) \leq \frac{1}{\sqrt{2}} d_{h,\alpha}(A, B) \leq d_{b,\alpha}(A, B).$$

This completes the proof. $\square$

In [28, Theorem 2] the authors proved that the matrix Kubo-Ando power mean $P_p(t, A, B)$ satisfies the in-betweenness property which follows from the fact that the function $g(t) = \operatorname{Tr}(A^{1/2} P_p(t; A, B)^{1/2})$ is concave. Note that $P_t(A, B) \neq P_t(B, A)$, i.e., P_t is not symmetric. However, for the symmetric means we may have the following result whose proof is adapted from [22].

Theorem 2.2.3. *Let σ be a symmetric mean and assume that one of the following inequalities* *holds for any pair of positive definite matrices A and B:*

$$d_{h,\alpha}(A, A\sigma B) \leq d_{h,\alpha}(A, B) \tag{2.2.2}$$

or

$$d_{b,\alpha}(A, A\sigma B) \leq d_{b,\alpha}(A, B). \tag{2.2.3}$$

Then σ is the arithmetic mean.

Proof. By Theorem 1.2.6 and 1.2.8, the symmetric operator mean σ is represented as follows:

$$A\sigma B = \frac{\delta}{2}(A+B) + \int_{(0,\infty)} \frac{\lambda+1}{\lambda}\{(\lambda A) : B + A : (\lambda B)\} d\mu(\lambda), \tag{2.2.4}$$

where $A, B \geq 0$, $\lambda \geq 0$ and μ is a positive measure on $(0, \infty)$ with $\delta + \mu((0, \infty)) = 1$, and the parallel sum $A : B$ is given by $A : B = (A^{-1} + B^{-1})^{-1}$, where A and B are invertible.

For two orthogonal projections P, Q acting on a Hilbert space H, let us denote by $P \wedge Q$ their infimum which is the orthogonal projection on the subspace $P(H) \cap Q(H)$. If $P \wedge Q = 0$, then by Theorem 1.2.7,

$$(\lambda P) : Q = P : (\lambda Q) = \frac{\lambda}{\lambda+1} P \wedge Q.$$

Consequently, from (2.2.4) we get

$$P\sigma Q = \frac{\delta}{2}(P+Q).$$

Let us consider the following orthogonal projections

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q_\theta = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}.$$

Notice that $Q_\theta \to P$ as $\theta \to 0$ and $Q_\theta \wedge P = 0$. From the projections above, it is easy to see that the inequality (2.2.2) becomes

$$d_{h,\alpha}(P, \delta(P+Q_\theta)/2) \le d_{h,\alpha}(P, Q_\theta).$$

Since this is true for all $\theta > 0$, we can take a limit as $\theta \to 0^+$ to obtain

$$d_{h,\alpha}(P, \delta P) \le d_{h,\alpha}(P, P)$$

whose equality occurs if and only if $\delta = 1$. This shows that $\mu = 0$ and σ is the arithmetic mean.

The statement for $d_{h,\alpha}$ can be proved similarly. □

In this chapter, we introduce a new distance called the weighted Hellinger distance and investigate its properties. This distance is constructed based on Minh's approach when he constructed the weighted Bures distance. The weighted Bures distance is an extended version with one parameter of the Bures distance. In the next chapter, we introduce a new quantum divergence called the α-z-Bures Wasserstein divergence, which is considered as an extension with two parameters of the Bures distance.

Chapter 3

The α-z-Bures Wasserstein divergence

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 $A, B \in \mathbb{P}_n$. 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 $A_1, A_2, \cdots, A_m$ be positive definite matrices. In 2004, Moakher [65] and then Bhatia and Holbrook [14] studied the following least squares problem

$$\min_{X>0} \sum_{i=1}^{m} \delta_2^2(X, A_i), \tag{3.0.1}$$

where $\delta_2(A, B) = ||\log(A^{-1}B)||_2$ is the Riemannian distance between A and B. They showed that (3.0.1) has a unique solution which is called the Karcher mean of $A_1, A_2, \cdots, A_m$. In literature, this mean has different names such as: Fréchet mean, Cartan mean, Riemannian center of mass. It turns out that the solution of (3.0.1) is the unique positive definite solution of the Karcher equation

$$\sum_{i=1}^{m} \log(X^{1/2}A_iX^{1/2}) = 0. \tag{3.0.2}$$

In [60], Lim and Palfia showed that the solution of (3.0.2) is nothing but the limit of the solution of the following matrix equation as $t \to 0$,

$$X = \sum_{i=1}^{m} \frac{1}{m} X \sharp_t A_i. \tag{3.0.3}$$

Recently, Franco and Dumitru [38] 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 [60] developed by Lim and Pálfia, they showed that the following equation

$$X = \sum_{i=1}^{m} \omega_i P_{\alpha_i, z_i}(X, A_i) \tag{3.0.4}$$

had a unique positive definite solution, where (ω_i) is a probability vector (it means, $\omega_i \geq 0$ and $\omega_1 + \omega_2 + ... + \omega_m = 1$) and $P_{\alpha,z}(A,B) = (B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}})^z$-the matrix function in the α-z-Rényi relative entropy introduced by Audenaert and Datta [7] in 2015. Notice that if we replace $P_{\alpha_i,z_i}(X, A_i)$ in (3.0.4) with the weighted geometric mean $X \sharp_t A_i$, the solution of the corresponding matrix equation is the weighted power mean.

Now, notice that if we change the distance function in (3.0.1), 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 [13] as follows:

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

where $\mathrm{Tr}((A^{1/2}BA^{1/2})^{1/2})$ 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 [14], these authors introduced so called the weighted Bures-Wasserstein distance as

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

where $F_t(A,B) = \mathrm{Tr}(A^{\frac{1-t}{2t}}BA^{\frac{1-t}{2t}})^t$ is the sandwiched quasi-relative entropy [59, 79]. 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 the first section of this chapter, we introduce and study different properties of the α-z-Bures Wasserstein divergence defined as

$$\Phi(A,B) = \mathrm{Tr}((1-\alpha)A+\alpha B) - \mathrm{Tr}(Q_{\alpha,z}(A,B)), \tag{3.0.5}$$

whenever A and B are positive definite matrices, and $Q_{\alpha,z}(A,B) = P_{\alpha,z}(B,A)$. Note 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 semi-definite matrices $A_1, A_2, \cdots, A_m$ is the unique positive definite solution of the following matrix equation

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

which is called the α-z-weighted right mean. We also establish some properties for this quantity.

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 [5, 26, 28, 39, 34].

In 1976 [78], A. Uhlmann introduced the concept of fidelity which is one of the most important concepts in quantum information. Besides the distance functions like Bures distance, Hellinger distance or relative entropy etc., people also use quantum fidelity to measure the distance between two quantum states. In addition, quantum fidelity can be used to characterize the property of a given quantum state, for instance, to quantify the quantum entanglement between two parts of a state, which is the shortest distance between the state and the set of all separable states. In the next section, in relation to quantum fidelity, we provide a refinement for Fuchs-van de Graaf inequality [80, 84] and a inequality for the parameterized version of quantum fidelity. Recall that the quantity we use to define the α-z-Bures Wasserstein divergence was first introduced in 2015 by Audenaert and Datta [7]. They referred to it as α-z-fidelity and denoted by

$$f_{\alpha,z}(\rho,\sigma) := \operatorname{Tr}\left(\rho^{\alpha/2z}\sigma^{(1-\alpha)/z}\rho^{\alpha/2z}\right)^{z} = \operatorname{Tr}\left(\sigma^{(1-\alpha)/2z}\rho^{\alpha/z}\sigma^{(1-\alpha)/2z}\right)^{z}. \tag{3.0.6}$$

This is the matrix function in the α-z-Rényi relative entropy which is a generalization of sandwiched Rényi relative entropy [66] and Rényi relative entropy [70]. In the past few years, many mathematicians and theoretical physicists paid a lot of attention to it [7, 20, 50, 83]. The above quantity is considered an extension involving two parameters of quantum fidelity. We show that (Theorem 3.3.4) for $0 < \alpha < z < 1$, the quantum α-z-fidelity $f_{\alpha,z}(\rho,\sigma)$ is the minimum of the following functions

$$P(X) = z\operatorname{Tr}\left(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X\right) - (z-1)\operatorname{Tr}\left(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}}$$

and

$$Q(X) = \Big(\operatorname{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X)\Big)^{z}.\Big(\operatorname{Tr}(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2x}})^{\frac{z}{z-1}}\Big)^{1-z}.$$

In the fourth section, we use quantum α-z-fidelity to measure the distance between two quantum orbits and show that the set of all distances forms a closed interval in $\mathbb{R}^{+}$. (Theorems 3.4.2 and 3.4.3).

The results of this chapter are taken from [30, 31, 32, 77].

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

The first main result of this section is that we show the α-z-Bures Wasserstein divergence, defined in (3.0.5), is a quantum divergence.

Recall that for $0 < p < 1$, and let μ be the measure on $(0, \infty)$ defined by

$$d\mu(\lambda) = \frac{\sin p\pi}{\pi}\lambda^{p-1}d\lambda,$$

then for $x > 0$, we have

$$x^p = \frac{\sin p\pi}{\pi}\int_0^\infty \frac{x}{\lambda + x}\lambda^{p-1}d\lambda.$$

The next lemma follows from the Spectral Theorem and the above representation.

Lemma 3.1.1 ([9]). *Let $0 < p < 1$ and A be a positive definite matrix. Then,*

$$A^p = \int_0^\infty A(\lambda + A)^{-1}d\mu(\lambda) = \int_0^\infty \Big(I - \lambda(\lambda + A)^{-1}\Big)d\mu(\lambda), \tag{3.1.7}$$

where $d\mu(\lambda) = \dfrac{\sin(p\pi)}{\pi}\lambda^{p-1}d\lambda$.

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

Lemma 3.1.2. *Let $y > 0$, and $0 < p < 1$. Then we have*

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

Theorem 3.1.1. *Let $0 \le \alpha \le z \le 1$. Then the quantity*

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

is a divergence.

Proof. For $z = \alpha$, the theorem was obtained in [14]. The case $0 \leq \alpha \leq z = 1$ was proved in a recent paper by Nguyen and Le in [56]. We need to consider the case $0 < \alpha < z < 1$.

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

$$\begin{aligned}\mathrm{Tr}((X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}})^z) &= \mathrm{Tr}((X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}})^{\frac{z}{\alpha}\alpha}) \\ &\leq \mathrm{Tr}((X^{\frac{1-\alpha}{2\alpha}} Y X^{\frac{1-\alpha}{2\alpha}})^{\alpha}) \quad \text{(the Araki-Lieb-Thirring inequality)} \\ &\leq \mathrm{Tr}((1-\alpha)X + \alpha Y) \quad \text{([14, Theorem 11])}.\end{aligned}$$

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

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

$$\frac{\partial f(X,Y)}{\partial Y}(B) = \mathrm{Tr}(\alpha B). \tag{3.1.8}$$

Since $0 < z < 1$, we have

$$\frac{\partial Y^z}{\partial Y}(B) = \int_0^\infty \delta^z(\delta + Y)^{-1} B (\delta + Y)^{-1} \frac{\sin(z\pi)}{\pi} d\delta. \tag{3.1.9}$$

If we put $\varphi(X,Y) = X^{\frac{1-\alpha}{2z}} Y^{\frac{\alpha}{z}} X^{\frac{1-\alpha}{2z}}$, then according to (3.1.9) 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,

$$\begin{aligned}\frac{\partial g(X,Y)}{\partial Y}(B) &= \mathrm{Tr}\Big[\int_0^\infty \delta^z(\delta+\varphi)^{-1}\frac{\partial\varphi}{\partial Y}(B)(\delta+\varphi)^{-1}\frac{\sin(z\pi)}{\pi}d\delta\Big]\\ &= \mathrm{Tr}\Big[\int_0^\infty \frac{\delta^z}{(\delta+\varphi)^2}\frac{\sin(z\pi)}{\pi}d\delta\left(\frac{\partial\varphi}{\partial Y}(B)\right)\Big]\\ &= \mathrm{Tr}\left[z\varphi^{z-1}\frac{\partial\varphi}{\partial Y}(B)\right]\quad \text{(By Lemma3.3.2)}\\ &= \mathrm{Tr}\left[z\Big(X^{\frac{1-\alpha}{2z}}Y^{\frac{\alpha}{z}}X^{\frac{1-\alpha}{2z}}\Big)^{z-1}\frac{\partial\varphi}{\partial Y}(B)\right].\end{aligned}$$

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 $Y=X$,

$$\begin{aligned}\frac{\partial g(X,Y)}{\partial Y}(B)\Big|_{Y=X} &= \mathrm{Tr}\left(zY^{\frac{z-1}{z}}Y^{\frac{1-\alpha}{2z}}\left(\int_0^\infty \lambda^{\frac{\alpha}{z}}(\lambda+Y)^{-1}B(\lambda+Y)^{-1}\frac{\sin(\frac{\alpha}{z}\pi)}{\pi}d\lambda\right)Y^{\frac{1-\alpha}{2z}}\right)\\ &= \mathrm{Tr}\left(zY^{\frac{z-\alpha}{z}}\int_0^\infty \lambda^{\frac{\alpha}{z}}(\lambda+Y)^{-1}B(\lambda+Y)^{-1}\frac{\sin(\frac{\alpha}{z}\pi)}{\pi}d\lambda\right)\\ &= \mathrm{Tr}\left(zY^{\frac{z-\alpha}{z}}\int_0^\infty \frac{\lambda^{\frac{\alpha}{z}}}{(\lambda+Y)^2}\frac{\sin(\frac{\alpha}{z}\pi)}{\pi}d\lambda B\right)\\ &= \mathrm{Tr}\left(zY^{\frac{z-\alpha}{z}}\frac{\alpha}{z}Y^{\frac{\alpha}{z}-1}B\right)\quad \text{(by Lemma 3.3.2)}\\ &= \mathrm{Tr}(\alpha B).\end{aligned}$$

Therefore, on account of (3.1.8) from the last identity we have

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

Property (ii) in definition of quantum divergence is fulfilled.

Finally, we check that for every Hermitian matrix B,

$$\frac{\partial^2\Phi(X,Y)}{\partial Y^2}\Big|_{Y=X}(B,B) \geq 0. \tag{3.1.10}$$

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

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

By the product rule, we have

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

Put

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

and

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

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

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

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 (3.1.10) is true. Indeed,

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

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

$$\left.\frac{\partial^2 \varphi}{\partial Y^2}\right|_{Y=X}(B, B) = -\frac{2\sin(\frac{\alpha}{z}\pi)}{\pi}\operatorname{Tr}\left(Y^{\frac{1-\alpha}{z}}\int_0^\infty \lambda^{\frac{\alpha}{z}}(\lambda+Y)^{-1}B(\lambda+Y)^{-1}B(\lambda+Y)^{-1}d\lambda\right). \tag{3.1.11}$$

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

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

Consequently, the integral (3.1.11) 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 $\mathbb{P}_n$. □

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

$$\min_{X>0} \sum_{i=1}^{m} \omega_i \Phi(X, A_i). \tag{3.1.12}$$

This problem was solved by Bhatia, Lim and Jain in [14] when $z = \alpha$.

Theorem 3.1.2. *For $0 \le \alpha \le z \le 1$, 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

$$\sum_{i=1}^{m} w_i Q_{\alpha,z}(X, A_i) = X. \tag{3.1.13}$$

Proof. We have

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

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

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

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

$$\alpha I = \int_0^\infty (\lambda + X)^{-1} C (\lambda + X)^{-1} \frac{\lambda^{\frac{\alpha}{z}} \sin(\frac{\alpha}{z}\pi)}{\pi} d\lambda. \tag{3.1.14}$$

Now, let us choose an orthonormal basis in which the matrix X is diagonal, i.e., $X = diag(x_1, x_2, \cdots, x_n)$

and let $C = (c_{ij})$ be the representation of C in this basis. From the equation (3.1.14) 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,

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

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

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

Thus, X satisfies the equation (3.1.13).

Finally, we show that the equation (3.1.13) has a unique solution. Firstly, notice that the function $F(X)$ is strictly convex [14]. Therefore, if the equation (3.1.13) 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 $aI \le A_i \le bI$, for all $1 \le i \le m$. We have

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

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. $\square$

In [49], Jeong et al denoted this solution as $\mathcal{R}_{\alpha,z}(\omega;\mathbb{A})$ and referred to it as the α-z-weighted right mean. They have identified several favorable properties for this quantity. For the convenience of the readers, let's recall some notations.

Let Δ_m be the set of all positive probability vectors in $\mathbb{R}^m$ convexly spanned by the unit coordinate vectors. Let $\mathbb{A} = (A_1, \ldots, A_m) \in \mathbb{P}_n^m, \omega = (w_1, \ldots, w_m) \in \Delta_m, \sigma \in S^m$ be a permutation on m-letters, $p \in \mathbb{R}$, and $M \in \mathrm{GL}_n$, the set of $n \times n$ invertible matrices. Denote

$$\begin{aligned}
\omega_\sigma &:= (w_{\sigma_1}, \ldots, w_{\sigma_m}), \\
\mathbb{A}_\sigma &:= (A_{\sigma_1}, \ldots, A_{\sigma_m}), \\
\mathbb{A}^p &:= (A_1^p, \ldots, A_m^p), \\
M\mathbb{A}M^* &:= (MA_1M^*, \ldots, MA_mM^*),
\end{aligned}$$

and

$$\hat{\omega} := \frac{1}{1-w_m}\left(w_1, \ldots, w_{m-1}\right) \in \Delta_{m-1},$$
$$\omega^{(k)} := \frac{1}{k}(\underbrace{w_1, \ldots, w_1}_{k}, \ldots, \underbrace{w_m, \ldots, w_m}_{k}) \in \Delta_{mk},$$
$$\mathbb{A}^{(k)} := (\underbrace{A_1, \ldots, A_1}_{k}, \ldots, \underbrace{A_m, \ldots, A_m}_{k}) \in \mathbb{P}_n^{mk}.$$

For the completeness, we recall some properties that were obtained in [49].

Proposition 3.1.1. *The weighted right mean $\mathcal{R}_{\alpha,z}$ satisfies the following:*

(i) $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A}) = \left(\sum_{j=1}^{m} w_j A_j^{1-\alpha}\right)^{\frac{1}{1-\alpha}}$ *if A_j are commuting.*

(ii) $\mathcal{R}_{\alpha,z}(\omega, c\mathbb{A}) = c\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ *for any $c > 0$.*

(iii) $\mathcal{R}_{\alpha,z}\left(\omega_\pi, \mathbb{A}_\pi\right) = \mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ *for any permutation π on $\{1, \ldots, m\}$.*

(iv) $\mathcal{R}_{\alpha,z}\left(\omega^{(k)}, \mathbb{A}^{(k)}\right) = \mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ *for any natural number k.*

(v) $\mathcal{R}_{\alpha,z}\left(\omega, U\mathbb{A}U^*\right) = U\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})U^*$ *for any unitary matrix U.*

(vi) $\det \mathcal{R}_{\alpha,z}(\omega, \mathbb{A}) \geq \prod_{j=1}^{m} \left(\det A_j\right)^{w_j}$*, and equality holds if and only if $A_1 = \cdots = A_m$.*

(vii) $X = \mathcal{R}_{\alpha,z}\left(\omega, A_1, \ldots, A_{m-1}, X\right)$ *implies that* $X = \mathcal{R}_{\alpha,z}\left(\hat{\omega}, A_1, \ldots, A_{m-1}\right)$.

(viii) $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A}) = \mathcal{R}_{\alpha,z}\left(\sum_{j=1}^{k} w_j, w_{k+1}, \ldots, w_m; A_1, A_{k+1}, \ldots, A_m\right)$ *if $A_1 = \cdots = A_k$ for $1 \leq k < m$.*

(ix) For $\frac{1}{2} \leq z \leq 1$,

$$\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})^{\frac{1-\alpha}{z}} \leq \mathcal{A}\left(\omega, \mathbb{A}^{\frac{1-\alpha}{z}}\right).$$

(x) Let $0 < \alpha \leq z < 1$. If $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A}) \leq I$, then

$$\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})^{1-\frac{\alpha}{z}} \geq \mathcal{A}\left(\omega, \mathbb{A}^{1-\alpha}\right).$$

If $\mathcal{R}_{\alpha,z}(\omega,\mathbb{A}) \geq I$, then the reverse inequality holds.

(xi) Let $0 < \alpha \leq z < 1$. If $\mathcal{R}_{\alpha,z}(\omega,\mathbb{A}) \geq I$ then

$$\mathcal{R}_{\alpha,z}(\omega,\mathbb{A})^{1-\frac{\alpha}{z}} \leq P_z\left(\omega, \mathbb{A}^{\frac{1-\alpha}{z}}\right).$$

If $\mathcal{R}_{\alpha,z}(\omega,\mathbb{A}) \leq I$, then the reverse inequality holds.

The matrix norm $|||\cdot|||$ on $\mathbb{M}_n$ is said to be unitarily invariant if $|||UAV||| = |||A|||$ for any matrix $A \in \mathbb{M}_n$ and unitary matrices U, V. In [49, Remark 3.6] the authors showed the following

$$|||\mathcal{A}\left(\omega, \mathbb{A}^{1-\alpha}\right)||| \leq |||\mathcal{R}_{\alpha,z}(\omega,\mathbb{A})||| \quad \text{and} \quad |||P_z\left(\omega, \mathbb{A}^{\frac{1-\alpha}{z}}\right)||| \leq |||\mathcal{R}_{\alpha,z}(\omega,\mathbb{A})|||.$$

We establish an upper bound as follows.

Proposition 3.1.2. *For $0 \leq \alpha \leq z \leq 1, \alpha \neq 1$, we have*

$$\|\mathcal{R}_{\alpha,z}(\omega,\mathbb{A})\| \leq \Big(\sum_{j=1}^{m} w_j \|A_j\|^{1-\alpha}\Big)^{\frac{1}{1-\alpha}}.$$

Proof. Let $X = \mathcal{R}_{\alpha,z}(\omega;\mathbb{A})$. By the triangle inequality, the sub-multiplicativity for the operator norm, and the fact that $\|A^t\| = \|A\|^t$ for any positive definite A and $t \geq 0$, from equation (3.1.13) we get

$$\begin{aligned}
\|\mathcal{R}_{\alpha,z}(\omega,\mathbb{A})\| = \|X\| &= \Big\|\sum_{j=1}^{m} w_j \Big(X^{\frac{\alpha}{2z}} A_j^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}}\Big)^z\Big\| \\
&\leq \sum_{j=1}^{m} w_j \Big\|X^{\frac{\alpha}{2z}} A_j^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}}\Big\|^z \\
&\leq \sum_{j=1}^{m} w_j \|X\|^{\alpha}.\|A_j\|^{1-\alpha}.
\end{aligned}$$

Therefore,

$$\|X\|^{1-\alpha} \leq \sum_{j=1}^{m} w_j \|A_j\|^{1-\alpha}.$$

Consequently,

$$\|\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})\| = \|X\| \leq \Big(\sum_{j=1}^{m} w_j \|A_j\|^{1-\alpha}\Big)^{\frac{1}{1-\alpha}}.$$

□

We derive a version of the AM-GM inequality for the α-z-weighted right mean. However, we need the following lemma ([61]).

Lemma 3.1.3. *Let $T > 0$. The following inequalites hold:*

1. $\lambda T + I - \lambda I \geq T^{\lambda}$ *for* $\lambda \in [0, 1]$.
2. $\lambda T + I - \lambda I \leq T^{\lambda}$ *for* $\lambda > 1$.
3. $\lambda T + I - \lambda I \leq T^{\lambda}$ *for* $\lambda < 0$.
4. $\lambda T + I - \lambda I \geq T^{\lambda} \geq (\lambda T^{-1} + I - \lambda I)^{-1}$ *for* $\lambda \in [0, 1]$.

Theorem 3.1.3. *Let $0 \leq \alpha \leq z \leq 1, \alpha \neq 1, z \neq 0$. Let $\mathbb{A} = (A_1, ..., A_m)$ be a m-tuple of positive definite matrices, and $\omega = (w_1, ..., w_m)$ a probability vector. We have*

$$\frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} w_j A_j^{-\frac{1-\alpha}{z}} \leq \mathcal{R}_{\alpha,z}(\omega, \mathbb{A}) \leq \Big(\frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} w_j A_j^{\frac{1-\alpha}{z}}\Big)^{-1}.$$

The second inequality holds when $(1 + z - \alpha)I - z \sum_{j=1}^{m} w_j A_j^{\frac{1-\alpha}{z}}$ is invertible.

Proof. Recall that $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ is the unique solution of the following equation

$$X = \sum_{j=1}^{m} w_j X^{\frac{\alpha}{2z}} \Big(A_j^{\frac{1-\alpha}{z}} \sharp_{1-z} X^{-\frac{\alpha}{z}}\Big) X^{\frac{\alpha}{2z}}.$$

Multiplying both sides of the above inequality from the left and from the right by $X^{-\frac{\alpha}{2z}}$ we get

$$X^{\frac{z-\alpha}{z}} = \sum_{j=1}^{m} w_j \Big(A_j^{\frac{1-\alpha}{z}} \sharp_{1-z} X^{-\frac{\alpha}{z}} \Big).$$

By the AM-GM inequality,

$$\begin{aligned} X^{\frac{z-\alpha}{z}} &\leq \sum_{j=1}^{m} w_j \Big(z A_j^{\frac{1-\alpha}{z}} + (1-z) X^{-\frac{\alpha}{z}} \Big) \\ &= \sum_{j=1}^{m} w_j z A_j^{\frac{1-\alpha}{z}} + (1-z) X^{-\frac{\alpha}{z}}. \end{aligned}$$

Therefore,

$$X^{1-\frac{\alpha}{z}} - (1-z) X^{-\frac{\alpha}{z}} \leq \sum_{j=1}^{m} w_j z A_j^{\frac{1-\alpha}{z}}. \tag{3.1.15}$$

Let $\varphi(X) = X^{1+t} - (1-z)X^t$, where $t = -\dfrac{\alpha}{z} \in [-1, 0]$. By Lemma 3.1.3, we have

$$\begin{aligned} \varphi(X) &= X^{-\frac{1}{2}} X^{2+t} X^{-\frac{1}{2}} - (1-z) X^{-\frac{1}{2}} X^{1+t} X^{-\frac{1}{2}} \\ &\geq X^{-\frac{1}{2}} \Big((2+t)X - (1+t)I \Big) X^{-\frac{1}{2}} - (1-z) X^{-\frac{1}{2}} \Big((1+t)X - tI \Big) X^{-\frac{1}{2}} \\ &= (2+t)I - (1+t)X^{-1} - (1-z)(1+t)I + (1-z)tX^{-1} \\ &= (\alpha - 1)X^{-1} + (1 + z - \alpha)I. \end{aligned}$$

Therefore, from (3.1.15) we get

$$(\alpha - 1)X^{-1} + (1 + z - \alpha)I \leq \sum_{j=1}^{m} w_j z A^{\frac{1-\alpha}{z}},$$

or,

$$X^{-1} \geq \frac{1 + z - \alpha}{1 - \alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} A_j^{\frac{1-\alpha}{z}}.$$

Consequently,

$$X \leq \Big(\frac{1+\alpha-z}{1-\alpha}I - \frac{z}{1-\alpha}\sum_{j=1}^{m} w_j A_j^{\frac{1-\alpha}{z}}\Big)^{-1}.$$

Now let us prove the first inequality in Theorem. By the harmonic mean-geometric mean inequality we have

$$X^{\frac{z-\alpha}{z}} = \sum_{j=1}^{m} w_j\Big(A_j^{\frac{1-\alpha}{z}}\sharp_{1-z}X^{-\frac{\alpha}{z}}\Big) \geq \sum_{j=1}^{m} w_j\Big(zA_j^{-\frac{1-\alpha}{z}} + (1-z)X^{\frac{\alpha}{z}}\Big)^{-1}.$$

Since the map $x \mapsto x^{-1}$ is convex, from the last inequality we get

$$\begin{aligned} X^{\frac{\alpha-z}{z}} &\leq \Big[\sum_{j=1}^{m} w_j\Big(zA_j^{-\frac{1-\alpha}{z}} + (1-z)X^{\frac{\alpha}{z}}\Big)^{-1}\Big]^{-1} \\ &\leq \sum_{j=1}^{m} w_j\Big(zA_j^{-\frac{1-\alpha}{z}} + (1-z)X^{\frac{\alpha}{z}}\Big) \\ &= \sum_{j=1}^{m} w_j z A_j^{-\frac{1-\alpha}{z}} + (1-z)X^{\frac{\alpha}{z}}. \end{aligned}$$

Consequently,

$$X^{\frac{\alpha}{z}-1} - (1-z)X^{\frac{\alpha}{z}} \leq \sum_{j=1}^{m} w_j z A_j^{-\frac{1-\alpha}{z}}. \tag{3.1.16}$$

Let $\Psi(X) = X^{p-1} - (1-z)X^p$, where $p = \dfrac{\alpha}{z} \in [0,1]$. Using Lemma 3.1.3 again we obtain

$$\begin{aligned} \Psi(X) &\geq (p-1)X + (2-p)I - (1-z)\Big(pX + (1-p)I\Big) \\ &= \Big(p-1-(1-z)p\Big)X + \Big(2-p-(1-z)(1-p)\Big)I \\ &= (\alpha-1)X + (1+z-\alpha)I. \end{aligned}$$

From (3.1.16) we get

$$(\alpha-1)X + (1+z-\alpha)I \leq \sum_{j=1}^{m} w_j z A_j^{-\frac{1-\alpha}{z}},$$

or,

$$X \geq \frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} A_j^{-\frac{1-\alpha}{z}}.$$

$\square$

The matrix power mean $P_t(\omega, \mathbb{A})$ for $t \in (0, 1]$ was introduced by Lim and Palfia [60] as the unique solution $X \in \mathbb{P}_n$ of the following equation

$$X = \sum_{j=1}^{m} w_j X \#_t A_j. \tag{3.1.17}$$

Note that for $t \in [-1, 0)$ we define $P_t(\omega, \mathbb{A}) = P_{-t}\left(\omega, \mathbb{A}^{-1}\right)^{-1}$. Especially,

$$P_1(\omega, \mathbb{A}) = \sum_{j=1}^{m} w_j A_j = \mathcal{A}(\omega, \mathbb{A}),$$

$$P_{-1}(\omega, \mathbb{A}) = \left[\sum_{j=1}^{m} w_j A_j^{-1}\right]^{-1} = \mathcal{H}(\omega, \mathbb{A})$$

are the weighted arithmetic and harmonic means, respectively.

Let $\omega = (w_1, w_2, ..., w_m)$ be a probability vector, for $m \geq 2$, a weighted m-mean $\mathcal{G}_m$ defined on $\mathbb{P}_n^m$ is an idempotent map $\mathcal{G}_m(\omega, \cdot) : \mathbb{P}_n^m \longrightarrow \mathbb{P}_n$, that is, $\mathcal{G}_m(\omega, X, X, ..., X) = X$, for all $X \in \mathbb{P}_n$. Let $\mathcal{A}_m := \mathcal{A}_m(\omega, \mathbb{A}) = \sum_{j=1}^{m} w_j A_j$ and $\mathcal{H}_m := \mathcal{H}_m(\omega, \mathbb{A}) = \left(\sum_{j=1}^{m} w_j A_j^{-1}\right)^{-1}$ be the arithmetic mean and the harmonic mean, respectively. In [48], Hwang and Kim proved that for any $\mathcal{G}_m$ between $\mathcal{A}_m$ and $\mathcal{H}_m$, i.e.,

$$\mathcal{H}_m \leq \mathcal{G}_m \leq \mathcal{A}_m, \tag{3.1.18}$$

the function $\mathcal{G}_m^{\omega} := \mathcal{G}_m(\omega, \cdot) : \mathbb{P}^m \to \mathbb{P}$ is differentiable at $\mathbb{I} = (I, ..., I)$ with

$$D\mathcal{G}_n^{\omega}(\mathbb{I})(X_1, ..., X_m) = \sum_{j=1}^{m} w_j X_j.$$

Notice that the α-z-weighted right mean does not satisfy the inequality (3.1.18). However we have the following result.

Theorem 3.1.4. *Let* $\omega = (w_1, ..., w_m)$ *be a probability vector and let* $\mathcal{R}^{\omega}_{\alpha,z} := \mathcal{R}_{\alpha,z}(\omega, \cdot) : \mathbb{P}_n^m \longrightarrow \mathbb{P}_n$*. Then* $\mathcal{R}^{\omega}_{\alpha,z}$ *is differentiable at* $\mathbb{I} = (I, ..., I)$*, and*

$$D\mathcal{R}^{\omega}_{\alpha,z}(\mathbb{I})(X_1, ..., X_m) = \sum_{j=1}^{m} w_j X_j.$$

Proof. Let $X_1, X_2, ..., X_m \in \mathbb{H}_n$. If $X_1 = X_2 = ... = X_m = 0$ then the result is obvious. Without loss of generality, we assume that there exists $i \in \{1, 2, ..., m\}$ such that $X_i \neq 0$. Let $\tau = \max\{\mathrm{spr}(X_i) : i = 1, 2, ..., m\} > 0$, where $\mathrm{spr}(X)$ is the spectral radius of X. Consider the following functions

$$f(t) = \frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} w_j (I + tX_j)^{-\frac{1-\alpha}{z}},$$

and

$$g(t) = \Big(\frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} w_j (I + tX_j)^{\frac{1-\alpha}{z}}\Big)^{-1}.$$

Since

$$\begin{aligned} \lambda(I + tX_j) &= 1 + t\lambda(X_j) \\ &\geq 1 - |t||\lambda(X_j)| \\ &\geq 1 - \rho|t| \\ &> 0, \end{aligned}$$

where $\lambda(X)$ is some eigenvalue of X, we have, $I + tX_j \in \mathbb{P}_n$, for all $t \in (-\frac{1}{\tau}, \frac{1}{\tau})$. Therefore, $f(t)$ and $g(t)$ are well-defined on $(-\frac{1}{\tau}, \frac{1}{\tau})$, and $f(0) = g(0) = I$. We have

$$\frac{d}{dt}(I + tX_j)^{-\frac{1-\alpha}{z}} = -(I + tX_j)^{-\frac{1-\alpha}{z}} \frac{d}{dt}(I + tX_j)^{\frac{1-\alpha}{z}} (I + tX_j)^{-\frac{1-\alpha}{z}}.$$

At $t = 0$, we have

$$\frac{d}{dt}(I + tX_j)^{-\frac{1-\alpha}{z}}\bigg|_{t=0} = -\frac{1-\alpha}{z}X_j.$$

Consequently,

$$\frac{d}{dt}f(t)\bigg|_{t=0} = \sum_{j=1}^{m} w_j X_j.$$

Similarly,

$$\frac{d}{dt}g(t) = -h(t)^{-1}\frac{d}{dt}h(t)h(t)^{-1},$$

where $h(t) = g(t)^{-1}$. Since $h(0) = I$ and

$$\frac{d}{dt}h(t)\bigg|_{t=0} = -\sum_{j=1}^{m} w_j X_j,$$

we obtain

$$\frac{d}{dt}g(t)\bigg|_{t=0} = \sum_{j=1}^{m} w_j X_j.$$

By Theorem 3.1.3, we have

$$f(t) - I \leq \mathcal{R}^{\omega}_{\alpha,z}(\omega, I + tX_1, ..., I + tX_m) - I \leq g(t) - I.$$

Since $\mathcal{R}^{\omega}_{\alpha,z}(\omega, I, ..., I) = I$, for any sufficiently small $t > 0$, we have

$$\frac{f(t) - f(0)}{t} \leq \frac{\mathcal{R}^{\omega}_{\alpha,z}(\omega, I + tX_1, ..., I + tX_m) - \mathcal{R}^{\omega}_{\alpha,z}(\omega, I, ..., I)}{t} \leq \frac{g(t) - g(0)}{t}.$$

Let $t \to 0^+$, from the above inequality, we obtain

$$\lim_{t \to 0^+} \frac{\mathcal{R}^{\omega}_{\alpha,z}(\omega, I + tX_1, ..., I + tX_m) - \mathcal{R}^{\omega}_{\alpha,z}(\omega, I, ..., I)}{t} = \sum_{j=1}^{m} w_j X_j.$$

Similarly, for $t < 0$ we have

$$\lim_{t\to 0^-} \frac{\mathcal{R}^{\omega}_{\alpha,z}(\omega, I + tX_1, ..., I + tX_m) - \mathcal{R}^{\omega}_{\alpha,z}(\omega, I, ..., I)}{t} = \sum_{j=1}^{m} w_j X_j.$$

Thus, $\mathcal{R}^{\omega}_{\alpha,z}$ is differentiable at $\mathbb{I} = (I, ..., I)$ and

$$D\mathcal{R}^{\omega}_{\alpha,z}(\mathbb{I})(X_1, ..., X_m) = \sum_{j=1}^{m} w_j X_j.$$

□

The well-known Lie-Trotter formula [76] states that for $X, Y \in \mathbb{M}_n$,

$$\lim_{n\to+\infty} \Big(\exp(\frac{X}{n}) \exp(\frac{Y}{n}) \Big)^n = \exp(X + Y).$$

This formula plays an essential role in the development of Lie Theory, and frequently appears in different research fields [44, 47, 48]. In [48], J.Hwang and S.Kim introduced the multivariate Lie-Trotter mean on the convex cone $\mathbb{P}_n$ of positive definite matrices. For a positive probability vector $\omega = (w_1, ..., w_m)$ and differentiable curves $\gamma_1, ..., \gamma_m$ on $\mathbb{P}_n$ with $\gamma_i(0) = I$ ($i = 1, \cdots, m$), a weighted m-mean $\mathcal{G}_m$ (for $m \geq 2$) is the multivariate Lie-Trotter mean if

$$\lim_{s\to 0} \mathcal{G}_m(\omega, \gamma_1(s), \gamma_2(s), , ..., \gamma_m(s))^{1/s} = \exp\Big(\sum_{j=1}^{m} w_j \gamma_j'(0) \Big).$$

In the following theorem, we show that the mean $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ is also a multivariate Lie-Trotter mean.

Theorem 3.1.5. *The $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$ is the multivariate Lie-Trotter mean, that means, for any probability vector $\omega = (w_1, w_2, ., w_m)$, we have*

$$\lim_{s\to 0} \mathcal{R}_{\alpha,z}(\omega, \gamma_1(s), ..., \gamma_m(s))^{1/s} = \exp\Big(\sum_{j=1}^{m} w_j \gamma_j'(0) \Big),$$

where for $\varepsilon > 0, \gamma_j : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{P}_n$ *are differentiable curves with* $\gamma_j(0) = I$*, for all* $j = 1, 2, ..., m$.

Proof. For each $j = 1, 2, ..., m$, since γ_j is a continuous map with $\gamma_j(0) = I$, there exists $\delta_j > 0$ such that $\gamma_j(s) \in B_r(I) = \left\{ A \in \mathbb{H}_n : ||A - I|| < r = 1 - \left(\frac{z}{1+z-\alpha}\right)^{\frac{z}{1-\alpha}} > 0 \right\}, \forall s \in (-\delta_j, \delta_j)$. Therefore,

$$\left|\lambda_i\Big(\gamma_j(s)\Big) - 1\right| = \left|\lambda_i\Big(\gamma_j(s) - I\Big)\right| \leq ||\gamma_j(s) - I|| \leq r,$$

where $\lambda_i(A)$ is the i-th eigenvalue of $A \in \mathbb{H}_n$ in the decreasing order. Therefore,

$$\lambda_i\big(\gamma_j(s)\big) \geq 1 - r = \left(\frac{z}{1+z-\alpha}\right)^{\frac{z}{1-\alpha}}.$$

From here it implies that

$$\gamma_j(s) \geq \left(\frac{z}{1+z-\alpha}\right)^{\frac{z}{1-\alpha}} I, \quad \text{or,} \quad \gamma_j(s)^{-\frac{1-\alpha}{z}} \leq \frac{1+z-\alpha}{z} I.$$

Consequently,

$$z \sum_{j=1}^{m} w_j \gamma_j(s)^{-\frac{1-\alpha}{z}} \leq (1+z-\alpha) I.$$

By Theorem 3.1.3, we have

$$\frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} w_j \gamma_j^{-\frac{1-\alpha}{z}}(s) \leq \mathcal{R}_{\alpha,z}(\omega, \mathbb{A}) \leq \left(\frac{1+z-\alpha}{1-\alpha} I - \frac{z}{1-\alpha} \sum_{j=1}^{m} w_j \gamma_j^{\frac{1-\alpha}{z}}(s) \right)^{-1}.$$

According to the operator monotonicity of the logarithmic function, from the last inequalities for $s > 0$ we get

$$\begin{aligned}\frac{1}{s}\log\left(\frac{1+z-\alpha}{1-\alpha}I - \frac{z}{1-\alpha}\sum_{j=1}^{m} w_j\gamma_j^{-\frac{1-\alpha}{z}}(s)\right) &\leq \frac{1}{s}\log\mathcal{R}_{\alpha,z}(\omega,\gamma_1(s),...,\gamma_m(s))\\ &\leq \frac{1}{s}\log\left(\frac{1+z-\alpha}{1-\alpha}I - \frac{z}{1-\alpha}\sum_{j=1}^{m} w_j\gamma_j^{\frac{1-\alpha}{z}}(s)\right)^{-1}.\end{aligned}$$

Using the L'Hopital rule, we obtain

$$\lim_{s\to 0^+}\log\mathcal{R}_{\alpha,z}(\omega,\gamma_1(s),...,\gamma_m(s))^{\frac{1}{s}} = \sum_{j=1}^{m}\gamma_j'(0).$$

Similarly, for $s < 0$ we also have

$$\lim_{s\to 0^-}\log\mathcal{R}_{\alpha,z}(\omega,\gamma_1(s),...,\gamma_m(s))^{\frac{1}{s}} = \sum_{j=1}^{m}\gamma_j'(0).$$

Thus,

$$\lim_{s\to 0}\mathcal{R}_{\alpha,z}(\omega,\gamma_1(s),...,\gamma_m(s))^{1/s} = \exp\Big(\sum_{j=1}^{m} w_j\gamma_j'(0)\Big).$$

□

3.2 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 semi-definite matrices A and B,

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

It is worth noting that (see, for examples, [81, 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 [19], the trace function $\Theta_{p,q,s}(X, Y) = \operatorname{Tr}\left[\left(X^{\frac{q}{2}} Y^p X^{\frac{q}{2}}\right)^s\right]$ is jointly concave if only if $0 \leq p, q \leq 1$ and $0 \leq s \leq \dfrac{1}{p+q}$. For $q = \dfrac{1-\alpha}{z}$, $p = \dfrac{\alpha}{z}$, and $s = z$, we have that the function $\operatorname{Tr}\left(Q_{\alpha,z}(X, Y)\right)$ 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 3.2.1. $\Phi(X, Y)$ *is invariant under all unitary matrix* U *and invariant under tensoring with another density matrix* τ.

Proof. For an arbitrary unitary U, we have

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

Also for an arbitrary density matrix τ, we have

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

□

To finish this section, 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 divergcence.

Theorem 3.2.2. *Let $A, B\in\mathbb{P}_n, 0<\alpha\le z\le 1, 1/2\le p\le 1$ and $\alpha\le zp$. Then for any positive definite matrices A and B,*

$$\Phi(A,\mu_p)\le\Phi(A,B). \tag{3.2.19}$$

Proof. We have

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

and

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

Therefore, the inequality (3.2.19) is equivalent to the following

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

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,

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

On the other hand, from the conditions $0 < \alpha \leq z$ and $\alpha \leq zp$ it implies that $0 \leq \dfrac{\alpha}{zp} \leq 1$. By the operator concavity of the map $x \mapsto x^{\frac{\alpha}{zp}}$, we have

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

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

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

Thus, the desired result follows if

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

Or, equivalently,

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

which was proved in Theorem 3.1.1. Thus, the matrix power mean μ_p satisfies the in-betweenness property in the α-z-Bures Wasserstein divergence. □

3.3 Quantum fidelity and its parameterized versions

Quantum fidelity is an important quantity in quantum information theory and quantum chaos theory. It is a distance measure between density matrices which are considered as quantum states. Although it is not a metric, it has many useful properties that can be used to define a metric on the space of density matrices. In this section, we present an inequality for quantum fidelity, along with several of its extended forms and properties. Firstly, let we recall the definition of quantum fidelity [64, 80].

Definition 3.3.1. Let $A, B \in \mathbb{P}_n$ be positive semi-definite matrices. The *fidelity* between two elements A and B is defined as

$$F(A,B) = ||\sqrt{A}\sqrt{B}||_1, \tag{3.3.20}$$

where $||.||_1$ is Schatten 1-norm (trace norm),

$$||A||_1 = \mathrm{Tr}\,|A| = \mathrm{Tr}\,\sqrt{AA^*}.$$

Alternatively, the trace norm of an operator (or a matrix) A can be expressed as the sum of its singular values, $||A||_1 = \sum\limits_{i=1}^{n} s_i(A)$.

In quantum theory, quantum fidelity is defined for density matrices, and it can be generalized to the set of positive semi-definite matrices. By (3.3.20), we have

$$F(A,B) = \mathrm{Tr}\left(A^{1/2}BA^{1/2}\right)^{1/2}.$$

When $A, B \in \mathcal{D}_n$, quantum fidelity have several important properties [64, 78, 80], which can be proved in the sense of unital C^*-algebras

(1) Bounds: $0 \leq F(A,B) \leq 1$. Furthermore $F(A,B) = 1$ iff $A = B$, while $F(A,B) = 0$ iff $supp(A) \perp supp(B)$.

(2) Symmetry: $F(A,B) = F(B,A)$.

(3) Unitary Invariance: $F(A,B) = F(UAU^*, UBU^*)$, for any unitary matrix U.

(4) Concavity: $F(A, tB + (1-t)C) \geq tF(A,B) + (1-t)F(A,C)$, for $t \in [0,1]$ and $A, B, C \in \mathcal{D}_n$.

(5) Multiplicativity: $F(A \otimes B, C \otimes D) = F(A,C) \cdot F(B,D)$, for A, B, C,and $D \in \mathcal{D}_n$.

(6) Joint concavity: $F(tA + (1-t)B, tc + (1-t)D) \geq tF(A,C) + (1-t)F(B,D)$, for $t \in [0,1]$ and A, B, C, and $D \in \mathcal{D}_n$.

One of the most important inequalities of quantum fidelity is Fuchs de Graaf's inequality [80, 84].

Theorem 3.3.1. *(Fuchs-van de Graaf's inequality) For $A, B \in \mathcal{D}_n$, we have*

$$1 - \frac{1}{2}||A-B||_1 \leq F(A,B) \leq \sqrt{1 - \frac{1}{4}||A-B||_1^2}. \tag{3.3.21}$$

Equivalently,

$$2 - 2F(A,B) \leq ||A-B||_1 \leq 2\sqrt{1 - F(A,B)^2}. \tag{3.3.22}$$

The above inequality provides an upper bound and lower bound of quantum fidelity. It is also a tight relationship between different distances between A and B. The proof of the right inequality of (3.3.22) is based on Uhlmann's theorem [68] while the proof of the left inequality of (3.3.22) based on the following result [80].

Lemma 3.3.1. *Let $A, B \in \mathbb{P}_n$ be positive semi-definite matrices. It holds that*

$$||A-B||_1 \geq ||\sqrt{A} - \sqrt{B}||_2^2,$$

where $||.||_2$ is the Shcatten 2-norm,

$$||A||_2 = \Big(\sum_{i=1}^{n} s_i^2(A)\Big)^{1/2}.$$

It is worth mentioning that it is difficult to improve the Fuchs-van de Graaf inequality. In [84], the authors established a lower bound for $F(A, B)$ as follows.

Let $\lambda_0 = \lambda_{\max}(B^{-1/2}AB^{-1/2})$, where $\lambda_{\max}(X)$ is used to denote the maximum eigenvalue of the matrix X. Then

$$F(A, B) \geq 1 - \frac{1}{2}\frac{\sqrt{\lambda_0}}{\sqrt{\lambda_0}+1}||A - B||_1.$$

Now we establish an estimate for the trace-norm of the difference for two density matrices A and B and the fidelity of A and the convex combination $tA + (1-t)B, t \in [0, 1]$ of A and B. Before presenting this result, let's recall the following well-known inequality [4, 73, 74]

$$d_b(A, B) \leq d_1^{1/2}(A, B),$$

where $A, B \in \mathbb{P}_n$. This inequality was first proved in C^*-algebra setting by Araki in [4]. However, we can prove this inequality by another way as follows.

By Lemma (3.3.1), we have

$$\begin{aligned}
||A - B||_1 \geq ||\sqrt{A} - \sqrt{B}||_2^2 &= \mathrm{Tr}(\sqrt{A} - \sqrt{B})^2 \\
&= \mathrm{Tr}(A + B - 2\sqrt{A}\sqrt{B}) \\
&\geq \mathrm{Tr}\, A + \mathrm{Tr}\, B - 2F(A, B) \\
&= d_b^2(A, B),
\end{aligned}$$

where the last inequality follows from the fact that

$$F(A, B) = \mathrm{Tr}(A^{1/2}BA^{1/2})^{1/2} \geq \mathrm{Tr}(A^{1/2}B^{1/2}),$$

which is the consequence of the famous Araki-Lieb-Thirring inequality [43].

Theorem 3.3.2. *Let* $A, B \in \mathcal{D}_n$ *and* $t \in [0,1]$. *Then*

$$\sqrt{F(A, tAa + (1-t)B} \geq 1 - \frac{1}{4}(1-\sqrt{t})||A-B||_1^{1/2}.$$

Proof. Firstly, let us recall the Jensen inequality for trace. Let f be a continuous and concave function on an interval J and m be a natural number. Then for self-adjoint matrices $X_1, \cdots, X_m$ with spectra in J,

$$\mathrm{Tr}\left(f\Big(\sum_{i=1}^{m} A_i^* X_i A_i\Big)\right) \geq \mathrm{Tr}\left(\sum_{i=1}^{m} A_i^* f(X_i) A_i\right),$$

where $A_1, \cdots, A_m$ satisfy $\sum\limits_{i=1}^{m} A_i^* A_i = I$.

We have

$$\begin{aligned} F(A, tA + (1-t)B) &= \mathrm{Tr}[A^{1/2}(tA + (1-t)B)A^{1/2}]^{1/2} \\ &= \mathrm{Tr}[tA^2 + (1-t)A^{1/2}BA^{1/2}]^{1/2} \\ &\geq \mathrm{Tr}[tA + (1-t)(A^{1/2}BA^{1/2})^{1/2}] \\ &= t + (1-t)F(A, B), \end{aligned}$$

where the inequality is valid according to the concavity of the function $x \mapsto x^{1/2}$ and Jensen's trace inequality.

From $d_b(A, B) \leq d_1^{1/2}(A, B) = ||A-B||_1^{1/2}$, we have

$$\begin{aligned} & 1 - \frac{1}{4}(1-\sqrt{t})||A-B||_1^{1/2} \\ \leq\ & 1 - \frac{1}{4}(1-\sqrt{t})d_b(A, B) \\ =\ & 1 - \frac{1}{4}(1-\sqrt{t})\sqrt{(2 - 2\,\mathrm{Tr}(A^{1/2}BA^{1/2})^{1/2})} \\ =\ & 1 - (1-\sqrt{t})\sqrt{(1 - F(A, B))}. \end{aligned}$$

Thus, it is necessary to prove

$$\sqrt{t+(1-t)F(A,B)} \geq 1-(1-\sqrt{t})\sqrt{(1-F(A,B))}.$$

Since $0 \leq t \leq 1$, and $0 \leq F(A,B) \leq 1$, squaring both sides of this inequality, we have

$$\begin{aligned}
& t+F-tF \geq 1-2(1-\sqrt{t})\sqrt{1-F}+(1-\sqrt{t})^2(1-F) \\
\Leftrightarrow\ & (t-1)(1-F)+2(1-\sqrt{t})\sqrt{1-F}-(1-\sqrt{t})^2(1-F) \geq 0 \\
\Leftrightarrow\ & (1-F)[(t-1)-(1-\sqrt{t})^2]+2(1-\sqrt{t})\sqrt{1-F} \geq 0 \\
\Leftrightarrow\ & 2(1-\sqrt{t})[\sqrt{1-F}-(1-F)] \geq 0.
\end{aligned}$$

In the above transformations, F is used to denote for $F(A,B)$. The last inequality is evident because $0 \leq \sqrt{t} \leq 1$, and $0 \leq 1-F(A,B) \leq 1$. □

Remark 3.3.1. For $t=0$, with $||A-B||_1 \leq 16$ and from the theorem we have

$$F(A,B) \geq (1-\frac{1}{4}||A-B||_1^{1/2})^2.$$

Let's compare the value $(1-\frac{1}{4}||A-B||_1^{1/2})^2$ and the value $1-\dfrac{1}{2}||A-B||_1$ on the left-hand-side part in the Fuchs-van de Graaf inequality. By a simple computation, if $||A-B||_1 \geq 64/81$ then we have

$$F(A,B) \geq (1-\frac{1}{4}||A-B||_1^{1/2})^2 \geq 1-\frac{1}{2}||A-B||_1.$$

Indeed, from the last inequality we have

$$\begin{aligned}
& (1-\frac{1}{4}||A-B||_1^{1/2})^2 \geq 1-\frac{1}{2}||A-B||_1 \\
\Leftrightarrow\ & 1-\frac{1}{2}||A-B||_1^{1/2}+\frac{1}{16}||A-B||_1 \geq 1-\frac{1}{2}||A-B||_1 \\
\Leftrightarrow\ & \frac{||A-B||_1^{1/2}}{2}\Big(\frac{9}{8}||A-B||_1^{1/2}-1\Big) \geq 0,
\end{aligned}$$

which is equivalent to that

$$||A-B||_1 \geq 64/81.$$

Therefore, the above result is a refinement of the Fuchs-van de Graaf inequality for a big set of quantum states A and B.

Next, in relation to the matrix power mean, we prove an inequality for quantity-called a *parameterized version of quantum fidelity*, which was introduced by Bhatia, Jain, and Lim [14].

Definition 3.3.2. Let $A, B \in \mathbb{P}_n$, a parameterized version of fidelity defined as

$$F_\alpha(A,B) = \operatorname{Tr}\left(A^{\frac{1-\alpha}{2\alpha}} B A^{\frac{1-\alpha}{2\alpha}}\right)^\alpha, \alpha \in (0,\infty).$$

Proposition 3.3.1. *Let $A, B \in \mathcal{D}_n$, $p \geq 1$ and $0 \leq t \leq 1, 0 < \alpha < 1$. Then*

$$F_\alpha(A, \mu_p(t;A,B)) \geq F_\alpha(A,B)$$

and

$$F_\alpha(A, P_p(t;A,B)) \geq F_\alpha(A,B).$$

Proof. Let $p=1$. Notice that the function x^α $(0<\alpha<1)$ is operator concave, and that $0 \leq F_\alpha(A,B) \leq Tr\Big(tA+(1-t)B\Big) = t+1-t = 1$ [14, Theorem 11]. We have

$$\begin{aligned}
F_\alpha(A,\mu_1(t;A,B)) &= \operatorname{Tr}\left(A^{\frac{1-\alpha}{2\alpha}}\Big(tA+(1-t)B\Big)A^{\frac{1-\alpha}{2\alpha}}\right)^\alpha \\
&= \operatorname{Tr}\left(tA^{\frac{1}{\alpha}}+(1-t)A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\right)^\alpha \\
&\geq \operatorname{Tr}\left(tA+(1-t)\Big(A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\Big)^\alpha\right) \\
&= t+(1-t)F_\alpha(A,B) \\
&\geq F_\alpha(A,B).
\end{aligned}$$

Now, let us consider the case where $p>1$. In this case, the function $x \mapsto x^{1/p}$ is operator concave, hence

$$\mu_p(t; A, B) = \Big(tA^p + (t-1)B^p\Big)^{1/p} \geq tA + (1-t)B = \mu_1(t; A, B).$$

This implies

$$F_\alpha(A, \mu_p(t; A, B)) \geq F_\alpha(A, \mu_1(t; A, B)),$$

from which the result for $\mu_p(t; A, B)$ follows.

The proof for $P_p(t; A, B)$ is similar to $P_1(t; A, B) = \mu_1(t; A, B)$ and

$$\Big(tI + (1-t)(A^{-1/2}BA^{-1/2})^p\Big)^{1/p} \geq tI + (1-t)(A^{-1/2}BA^{-1/2}),$$

which implies

$$F_\alpha(A, P_p(t; A, B)) \geq F_\alpha(A, P_1(t; A, B)).$$

□

A parameterized version with two parameters was introduced by Audenaert and Datta, called α-z-fidelity [7] and defined by

$$f_{\alpha,z}(\rho, \sigma) := \mathrm{Tr}\left(\rho^{\alpha/2z}\sigma^{(1-\alpha)/z}\rho^{\alpha/2z}\right)^z = \mathrm{Tr}\left(\sigma^{(1-\alpha)/2z}\rho^{\alpha/z}\sigma^{(1-\alpha)/2z}\right)^z. \tag{3.3.23}$$

We list some basic properties of the α-z-fidelity as follows [7, 83].

Proposition 3.3.2. *Let ρ and σ be density matrices. Then,*

1. $f_{\alpha,z}(\rho, \sigma) \leq 1$ *for* $0 < \alpha < 1$ *and* $z > 0$.

2. $f_{\alpha,z}(\rho, \sigma) \geq 1$ *for* $\alpha > 1$ *and* $z > 0$.

3. *For* $\alpha \in (0, 1) \cup (1, +\infty)$ *and* $z > 0$, $f_{\alpha,z}(\rho, \sigma) = 1$ *if and only if* $\rho = \sigma$.

Proposition 3.3.3. *Let ρ and σ be density matrices such that supp $\rho \subseteq$ supp σ, and Λ be a completely positive trace preserving map (a quantum channel).*

1. *If* $\alpha \in (0,1]$ *and* $z \geq \max\{\alpha, 1-\alpha\}$*, then*

$$f_{\alpha,z}(\Lambda(\rho), \Lambda(\sigma)) \geq f_{\alpha,z}(\rho, \sigma).$$

2. *If* $\alpha \in [1,2]$ *and* $z \in \left\{1, \frac{\alpha}{2}\right\}$, *or* $\alpha \geq 1$ *and* $z = \alpha$*, then*

$$f_{\alpha,z}(\Lambda(\rho), \Lambda(\sigma)) \leq f_{\alpha,z}(\rho, \sigma).$$

For $z = \alpha$, Bhatia and coauthors [14] established some variational formulas for $f_{\alpha,\alpha}(\rho,\sigma)$ via extreme values of the following matrix functions.

Theorem 3.3.3. *Let* ρ *and* σ *be positive definite matrices and let* $0 < \alpha < 1$*. Then*

1. $f_{\alpha,\alpha}(\rho,\sigma) = \min\limits_{X>0} \mathrm{Tr}[(1-\alpha)\left(\sigma^{\frac{\alpha-1}{2\alpha}} X \sigma^{\frac{\alpha-1}{2\alpha}}\right)^{\frac{\alpha}{\alpha-1}} + \alpha X\rho]$.
2. $f_{\alpha,\alpha}(\rho,\sigma) = \min\limits_{X>0} \mathrm{Tr}[(\sigma^{\frac{\alpha-1}{2\alpha}} X \sigma^{\frac{\alpha-1}{2\alpha}})^{\frac{\alpha}{\alpha-1}}]^{1-\alpha}[\mathrm{Tr}(X\rho)]^{\alpha}$.
3. $f_{\alpha,\alpha}(\rho,\sigma) = \min\limits_{X>0} \mathrm{Tr}[\alpha\sigma^{\frac{1-\alpha}{\alpha}} X + (1-\alpha)(\rho^{-\frac{1}{2}} X \rho^{-\frac{1}{2}})^{\frac{\alpha}{\alpha-1}}]$.
4. $f_{\alpha,\alpha}(\rho,\sigma) = \min\limits_{X>0} [\mathrm{Tr}\,\sigma^{\frac{1-\alpha}{\alpha}} X]^{\alpha}[\mathrm{Tr}((\rho^{-\frac{1}{2}} X \rho^{-\frac{1}{2}})^{\frac{\alpha}{\alpha-1}}]^{1-\alpha}$.

Recently, S.Chehade [20] used the classical matrix inequalities to prove that for $\alpha > 1$ and $z > 1$,

$$f_{\alpha,z}(\rho,\sigma) = \max_{X>0} P(X),$$

where

$$P(X) = z\,\mathrm{Tr}\left(\sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} X\right) - (z-1)\,\mathrm{Tr}\left(\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}}.$$

In this section, we show that $f_{\alpha,z}(\rho,\sigma)$ is also the minimum of $P(X)$ when $0 < \alpha < z < 1$. In addition, $f_{\alpha,z}(\rho,\sigma) = \min\limits_{X>0} Q(X)$, where

$$Q(X) = \left(\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} X)\right)^{z} . \left(\mathrm{Tr}(\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2x}})^{\frac{z}{z-1}}\right)^{1-z}.$$

In order to prove the main result of this section we need the following lemmas that can be found in [14].

Lemma 3.3.2. *Let f be a smooth function on $\mathbb{R}^+$ and let $\hat{f}$ be the function on $\mathbb{P}_n$ defined as $\hat{f}(X) = \operatorname{Tr} f(X)$. Then for all $X \in \mathbb{P}_n$ and $Y \in \mathbb{H}_n$.*

$$D\hat{f}(X)(Y) = \operatorname{Tr}\left(f'(X)Y\right).$$

Lemma 3.3.3. *The function $f(X) = \operatorname{Tr} X^t$ on the set of positive definite matrices is strictly concave if $0 < t < 1$ and strictly convex if $t \in (-\infty, 0) \cup (1, \infty)$.*

Theorem 3.3.4. *Let ρ, σ be positive definite matrices and $0 < \alpha < z < 1$. We have*

(i) $f_{\alpha,z}(\rho,\sigma) = \min_{X>0} P(X)$.

(ii) $f_{\alpha,z}(\rho,\sigma) = \min_{X>0} Q(X)$.

Furthermore, the minimum is achieved at $X_0 = \sigma^{\frac{1-z}{2z}}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^{z-1}\sigma^{\frac{1-z}{2z}}$.

Proof. (i). Let $k(X) = \operatorname{Tr} X^{\frac{z}{z-1}}$ and $h(X) = \operatorname{Tr}(\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}})^{\frac{z}{z-1}}$. By Lemma 3.3.2, we have $Dk(X)(Y) = \dfrac{z}{z-1}\operatorname{Tr}(X^{\frac{1}{z-1}}Y)$. By the chain rule, we have

$$\begin{aligned}
Dh(X)(Y) &= Dk(\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}})(\sigma^{\frac{z-1}{2z}} Y \sigma^{\frac{z-1}{2z}}) \\
&= \frac{z}{z-1}\operatorname{Tr}\left((\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}})^{\frac{1}{z-1}}(\sigma^{\frac{z-1}{2z}} Y \sigma^{\frac{z-1}{2z}})\right) \\
&= \frac{z}{z-1}\operatorname{Tr}\left(\sigma^{\frac{z-1}{2z}}(\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}})^{\frac{1}{z-1}}\sigma^{\frac{z-1}{2z}} Y\right) \\
&= \frac{z}{z-1}\operatorname{Tr}\left(\sigma^{\frac{z-1}{2z}}(\sigma^{\frac{1-z}{2z}} X^{-1} \sigma^{\frac{1-z}{2z}})^{\frac{1}{1-z}}\sigma^{\frac{z-1}{2z}} Y\right).
\end{aligned}$$

Consequently,

$$DP(X)(Y) = z\operatorname{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}Y) - z\operatorname{Tr}\left(\sigma^{\frac{z-1}{2z}}(\sigma^{\frac{1-z}{2z}} X^{-1}\sigma^{\frac{1-z}{2z}})^{\frac{1}{1-z}}\sigma^{\frac{z-1}{2z}} Y\right),$$

and hence, $DP(X)(Y) = 0$ for all $Y \in \mathbb{H}_n$ if only if

$$\sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} = \sigma^{\frac{z-1}{2z}} (\sigma^{\frac{1-z}{2z}} X^{-1} \sigma^{\frac{1-z}{2z}})^{\frac{1}{1-z}} \sigma^{\frac{z-1}{2z}}.$$

Multiplying both sides of the above identity from the left and from the right by $\sigma^{\frac{1-z}{2z}}$ we get

$$\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} = (\sigma^{\frac{1-z}{2z}} X^{-1} \sigma^{\frac{1-z}{2z}})^{\frac{1}{1-z}} = (\sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}})^{\frac{1}{z-1}}.$$

From here we get

$$(\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} = \sigma^{\frac{z-1}{2z}} X \sigma^{\frac{z-1}{2z}},$$

or,

$$X = \sigma^{\frac{1-z}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} \sigma^{\frac{1-z}{2z}}.$$

According to Lemma 3.3.3, the function $P(X)$ is convex. Therefore, $P(X)$ attains a minimum at

$$X_0 = \sigma^{\frac{1-z}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} \sigma^{\frac{1-z}{2z}}.$$

Now we have

$$\begin{aligned}
\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} X_0) &= \mathrm{Tr}\left(\sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} \sigma^{\frac{1-z}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} \sigma^{\frac{1-z}{2z}}\right) \\
&= \mathrm{Tr}\left(\rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} \sigma^{\frac{1-\alpha}{2z}}\right) \\
&= \mathrm{Tr}\left(\rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} (\sigma^{\frac{\alpha-1}{2z}} \rho^{-\frac{\alpha}{z}} \sigma^{\frac{\alpha-1}{2z}})^{1-z} \sigma^{\frac{1-\alpha}{2z}}\right) \\
&= \mathrm{Tr}\left(\rho^{\frac{\alpha}{z}} (\sigma^{\frac{1-\alpha}{z}} \sharp_{1-z} \rho^{-\frac{\alpha}{z}})\right) \\
&= \mathrm{Tr}\left(\rho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{z}} \rho^{\frac{\alpha}{2z}} \sharp_{1-z} I\right) \\
&= \mathrm{Tr}(\rho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{z}} \rho^{\frac{\alpha}{2z}})^{z} \\
&= \mathrm{Tr}(\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Tr}\left((\sigma^{\frac{z-1}{2z}}X_0\sigma^{\frac{z-1}{2z}})^{\frac{z}{z-1}}\right) &= \mathrm{Tr}\left(\sigma^{\frac{z-1}{2z}}\sigma^{\frac{1-z}{2z}}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^{z-1}\sigma^{\frac{1-z}{2z}}\sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}} \\
&= \mathrm{Tr}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^z.
\end{aligned}$$

Therefore,

$$\begin{aligned}
P(X_0) &= z\,\mathrm{Tr}\left(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X_0\right) - (z-1)\,\mathrm{Tr}\left(\sigma^{\frac{z-1}{2z}}X_0\sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}} \\
&= z\,\mathrm{Tr}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^z - (z-1)\,\mathrm{Tr}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^z \\
&= \mathrm{Tr}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^z \\
&= f_{\alpha,z}(\rho,\sigma).
\end{aligned}$$

(ii). Now, we prove $f_{\alpha,z}(\rho,\sigma) = \min_{X>0} Q(X)$. The Hölder inequality states that for positive numbers p, q, r with $\dfrac{1}{p} + \dfrac{1}{q} = \dfrac{1}{r}$, we have

$$\mathrm{Tr}\,|ST|^r \leq (\mathrm{Tr}\,|S|^p)^{\frac{r}{p}}(\mathrm{Tr}\,|T|^q)^{\frac{r}{q}}, \tag{3.3.24}$$

where $|X| = (X^*X)^{\frac{1}{2}}$ be the absolute value of X. Applying (3.3.24) for $S = \rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}}, T = X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}}$ and $r = 2z, p = 2$ and $q = \frac{2z}{1-z}$, we obtain

$$\begin{aligned}
\mathrm{Tr}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^z &= \mathrm{Tr}\left(\sigma^{\frac{1-z}{2z}}X^{-\frac{1}{2}}X^{\frac{1}{2}}\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{2z}}\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}}X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}}\right)^z \\
&= \mathrm{Tr}\left((X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}})^*(\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}})^*(\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}})(X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}})\right)^z \\
&= \mathrm{Tr}\left|(\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}})(X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}})\right|^{2z} \\
&\leq \left(\mathrm{Tr}\,|\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}}|^2\right)^{\frac{2z}{2}}\left(\mathrm{Tr}\,|X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}}|^{\frac{2z}{1-z}}\right)^{1-z} \\
&= \left(\mathrm{Tr}(X^{\frac{1}{2}}\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{2z}}\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}X^{\frac{1}{2}})\right)^z\left(\mathrm{Tr}(\sigma^{\frac{1-z}{2z}}X^{-\frac{1}{2}}X^{-\frac{1}{2}}\sigma^{\frac{1-z}{2z}})^{\frac{z}{1-z}}\right)^{1-z} \\
&= \left(\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{z}}X)\right)^z\left(\mathrm{Tr}(\sigma^{\frac{1-z}{2z}}X^{-1}\sigma^{\frac{1-z}{2z}})^{\frac{z}{1-z}}\right)^{1-z} \\
&= \left(\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{z}}X)\right)^z\left(\mathrm{Tr}(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2z}})^{\frac{z}{z-1}}\right)^{1-z}.
\end{aligned}$$

From the proof of (i), functions $\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{z}}X)$ and $\mathrm{Tr}(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2z}})^{\frac{z}{z-1}}$ attain minimum at $X = X_0 = \sigma^{\frac{1-z}{2z}}(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}})^{z-1}\sigma^{\frac{1-z}{2z}}$. We also have

$$\Big(\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X_0)\Big)^z = f_{\alpha,z}(\rho,\sigma),$$

and

$$\left[\mathrm{Tr}\left(\sigma^{\frac{z-1}{2z}}X_0\sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}}\right]^{1-z} = f_{\alpha,z}(\rho,\sigma)^{1-z}.$$

Thus, (ii) is proven. $\square$

3.4 The α-z-fidelity between unitary orbits

In [85], Zhang and then in [82], Yan and their co-authors used the fidelity and the α-fidelity to determine the distance between two quantum orbits. In this section, adapting their techniques we use quantum α-z-fidelity to measure the distance between two quantum orbits. We also show that the set of these distances is a close interval in $\mathbb{R}^+$.

Definition 3.4.1 (([85])). Let $U(\mathbb{H})$ be the set of $n\times n$ unitary matrices, and $\mathcal{D}_n$ the set of density matrices. For $\rho \in \mathcal{D}_n$, its unitary orbit is defined as

$$U_\rho = \{U\rho U^* : U \in U(\mathbb{H})\}.$$

In this section we are going to obtain the maximum and minimum distance between orbits of two state ρ and σ in $\mathcal{D}_n$ via the quantum α-z-fidelity. Note that $f_{\alpha,z}(V\rho V^*, W\sigma W^*) = f_{\alpha,z}(\rho, U\sigma U^*)$, where $U = V^*W$. Thus, the problems are reduced to computing $\max\limits_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*)$ and $\min\limits_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*)$.

For the proof of the folowing theorem, let us recall the Golden Thomson inequality [9].

Theorem 3.4.1. *For two Hermitian operators $A, B \in L(H)$, we have*

$$\mathrm{Tr}[\exp(A+B)] \le \mathrm{Tr}[\exp(A)\exp(B)].$$

Furthermore, the equality holds if and only if A and B commute.

Theorem 3.4.2. *Let ρ and $\sigma \in \mathcal{D}_n$, the α-z-fidelity $f_{\alpha,z}(\rho,\sigma) = \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^z$ between the unitary orbits U_ρ and U_σ satifies*

$$\max_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) = \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\downarrow}(\sigma)^{1-\alpha},$$

and

$$\min_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) = \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\uparrow}(\sigma)^{1-\alpha},$$

where $\lambda(\rho) = (\lambda_1, \ldots, \lambda_n)$ are the eigenvalues of ρ and $\lambda^{\downarrow}(\rho)$ (resp. $\lambda^{\uparrow}(\rho)\big)$ is a rearrangement of $\lambda(\rho)$ in decreasing order (resp. increasing order).

Proof. For two Hermitian matrices A and B, there exist unitary matrices U and V such that $A = U\Lambda(A)U^*$ and $B = V\Lambda(B)V^*$, where $\Lambda(A)$ and $\Lambda(B)$ are denote the diagonal matrices whose entries are the eigenvalues of A and B, respectively [9]. In addition, for a density matrix ρ, there is an orthonormal basis $\{|i\rangle : i = 1, 2, ..., n\}$ such that $\rho = \sum_{i=1}^{n} \lambda_i(\rho)|i\rangle\langle i|$ [71]. Therefore, the eigenvectors of two density matrices can be connected via a unitary matrix, we can assume that ρ and σ have the following spectral decompositions

$$\rho = \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)|i\rangle\langle i| \text{ and } \sigma = \sum_{i=1}^{n} \lambda_i^{\downarrow}(\sigma)W_0|i\rangle\langle i|W_0^*,$$

where $\lambda_i^{\downarrow}(p), \lambda_i^{\downarrow}(\sigma) > 0$ for all $i = 1, 2, ..., n$ and W_0 is a unitary matrix.

We know that for any two $n\times n$ Hermitian matrices A and B, there exist two unitary matrices U_1 and U_2 such that [72]

$$\exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right) = \exp\left(U_1AU_1^* + U_2BU_2^*\right).$$

Let $A = \dfrac{1-\alpha}{z} U \ln \sigma U^*$ and $B = \dfrac{\alpha}{z} \ln \rho$, we have

$$\begin{aligned}
\exp\left(\frac{A}{2}\right)\exp(B)\exp\left(\frac{A}{2}\right) &= U\sigma^{\frac{1-\alpha}{2z}}U^*\rho^{\frac{\alpha}{z}}U\sigma^{\frac{1-\alpha}{2z}}U^* \\
&= \exp\Big(\frac{1-\alpha}{z}U_1 U \ln \sigma U_1^* U^* + \frac{\alpha}{z}U_2 \ln \sigma U_2^*\Big).
\end{aligned}$$

Therefore, for $0 < \alpha < 1$, we have

$$\begin{aligned}
f_{\alpha,z}(\rho,\sigma) &= \operatorname{Tr}\Big(U\sigma^{\frac{1-\alpha}{2z}}U^*\rho^{\frac{\alpha}{z}}U\sigma^{\frac{1-\alpha}{2z}}U^*\Big)^z = \operatorname{Tr}\left(\exp\Big(\frac{1-\alpha}{z}U_1 U \ln \sigma U^* U_1^* + \frac{\alpha}{z}U_2 \ln \sigma U_2^*\Big)\right)^z \\
&= \operatorname{Tr}\left(\exp\Big(\alpha U_2 \ln \sigma U_2^* + (1-\alpha)U_1 U \ln \sigma U^* U_1^*\Big)\right) \\
&= \operatorname{Tr}\Big(\exp(\alpha \ln \rho + (1-\alpha)\tilde{U}\ln \sigma \tilde{U}^*)\Big),
\end{aligned}$$

where $\tilde{U} = U_2^* U_1 U$. By the Golden-Thomson inequality we have

$$\operatorname{Tr}\Big(\exp(\alpha \ln \rho + (1-\alpha)\tilde{U}\ln \sigma \tilde{U}^*)\Big) \leq \operatorname{Tr}\Big(\rho^\alpha \tilde{U}\sigma^{1-\alpha}\tilde{U}^*\Big).$$

Using the Araki-Lieb-Thirring inequality we have

$$\operatorname{Tr}\Big(\rho^\alpha \tilde{U}\sigma^{1-\alpha}\tilde{U}^*\Big) = \operatorname{Tr}\Big(\tilde{U}\sigma^{\frac{1-\alpha}{2}}\tilde{U}^*\rho^\alpha\tilde{U}\sigma^{\frac{1-\alpha}{2}}\tilde{U}^*\Big) \leq \operatorname{Tr}\Big(\sigma^{\frac{1-\alpha}{2z}}\tilde{U}^*\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\tilde{U}^*\Big)^z = f_{\alpha,z}(\rho,\sigma\tilde{U}^*).$$

Since the unitary group $U(\mathbb{H})$ is compact, there exists some unitary $U_0 \in U(\mathbb{H})$ such that the maximum is attained, that is

$$\begin{aligned}
\max_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) = f_{\alpha,z}(\sigma, U_0\sigma U_0^*) &= \operatorname{Tr}\Big(\exp(\alpha \ln \rho + (1-\alpha)\tilde{U}_0 \ln \sigma \tilde{U}_0^*)\Big) \\
&= \operatorname{Tr}\Big(\rho^\alpha \tilde{U}\sigma^{1-\alpha}\tilde{U}^*\Big).
\end{aligned}$$

By the condition of equality of the Golden-Thomson's inequality we must have $[\rho^\alpha, \tilde{U}_0\sigma^{1-\alpha}\tilde{U}_0^*] = 0$. Since $[\rho^\alpha, W_0^*\sigma^{1-\alpha}W_0] = 0$, it follows that $\tilde{U}_0 = W^*$. As a result, if $[\rho, W_0^*\sigma W_0] = 0$ and

$0 < \alpha < z < 1$, then

$$\max_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) = f_{\alpha,z}(\rho, W_0^*\sigma W_0) = \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\downarrow}(\sigma)^{1-\alpha}.$$

On the other hand, we have

$$\begin{aligned} f_{\alpha,z}(\rho, U\sigma U^*) \geq \operatorname{Tr}\left(\rho^{\alpha}U\sigma^{1-\alpha}U^*\right) &\geq \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho^{\alpha})\lambda_i^{\uparrow}(U\sigma^{1-\alpha}U^*) \\ &= \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\uparrow}(\sigma)^{1-\alpha}, \end{aligned}$$

where the first inequality holds by Araki-Lieb-Thirring's inequality and the second inequality holds because for any two positive definite matrices A and B we have [9]

$$\sum_{i=1}^{n} \lambda_i^{\downarrow}(A).\lambda_i^{\uparrow}(B) \leqslant \operatorname{Tr}(AB) \leqslant \sum_{i=1}^{n} \lambda_i^{\downarrow}(A).\lambda_i^{\downarrow}(B).$$

Therefore,

$$\min_{U\in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) \geq \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\uparrow}(\sigma)^{1-\alpha}.$$

The equality occurs when $UW_0|i\rangle = |n-i+1\rangle$. $\square$

Theorem 3.4.3. *For* $0 \leq \alpha \leq z \leq 1$,

$$\{f_{\alpha,z}(\rho, U\sigma U^*) : U \in U(\mathbb{H})\} = \Big[\sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\uparrow}(\sigma)^{1-\alpha}, \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha}\lambda_i^{\downarrow}(\sigma)^{1-\alpha}\Big]. \tag{3.4.25}$$

Proof. Recall that the Stone theorem [21] states that any unitary matrix U can be parameterized as $U = \exp(tK)$ for some skew Hermitian matrix K and $t \in \mathbb{R}$. Note that

$$\operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{z}}\right)^{z} = \operatorname{Tr}\left(\rho^{\frac{\alpha}{2z}}\sigma^{\frac{1-\alpha}{z}}\rho^{\frac{\alpha}{2z}}\right)^{z}.$$

By Theorem 3.4.2, there exists $t \in \mathbb{R}$ such that function

$$g(t) = f_{\alpha,z}(\rho, U_t\sigma U_t^*) = \operatorname{Tr}\left(\rho^{\frac{\alpha}{2z}} U_t \sigma^{\frac{1-\alpha}{z}} U_t^* \rho^{\frac{\alpha}{2z}}\right)^z$$

achieves maximum and minimum. In order to prove (3.4.25) we need to verify that the function $g(t)$ is continuous in t.

The function $U_t = \exp(tK)$ is continuous in t with respect to the Schatten 1-norm, this means that for all $\varepsilon > 0$, there exists $\delta > 0$ such that $\|U_{t+\delta} - U_t\|_1 < \frac{\varepsilon}{2}$. Let $\theta_t = \rho^{\frac{\alpha}{2z}} U_t \sigma^{\frac{1-\alpha}{z}} U_t^* \rho^{\frac{\alpha}{2z}}$, then we have

$$\begin{aligned}
\|\theta_{t+\delta} - \theta_t\|_1 &= \|\rho^{\frac{\alpha}{2z}} U_{t+\delta} \sigma^{\frac{1-\alpha}{z}} U_{t+\delta}^* \rho^{\frac{\alpha}{2z}} - \rho^{\frac{\alpha}{2z}} U_t \sigma^{\frac{1-\alpha}{z}} U_t^* \rho^{\frac{\alpha}{2z}}\|_1 \\
&= \|\rho^{\frac{\alpha}{2z}} [(U_{t+\delta} - U_t)\sigma^{\frac{1-\alpha}{z}} U_{t+\delta}^* + U_t \sigma^{\frac{1-\alpha}{z}} (U_{t+\delta}^* - U_t^*)] \rho^{\frac{\alpha}{2z}}\|_1 \\
&\leq \|\rho^{\frac{\alpha}{2z}}\|_\infty^2 . \|(U_{t+\delta} - U_t)\sigma^{\frac{1-\alpha}{z}} U_{t+\delta}^* + U_t \sigma^{\frac{1-\alpha}{z}} (U_{t+\delta}^* - U_t^*)\|_1 \\
&\leq \|\rho^{\frac{\alpha}{2z}}\|_\infty^2 . (\|U_{t+\delta} - U_t\|_1 . \|\sigma^{\frac{1-\alpha}{z}} U_{t+\delta}^*\|_\infty + \|U_t \sigma^{\frac{1-\alpha}{z}}\|_\infty . \|U_{t+\delta}^* - U_t^*\|_1) \\
&\leq \|\rho^{\frac{1-\alpha}{2z}}\|_\infty^2 . \|\sigma^{\frac{1-\alpha}{z}}\|_\infty . (\|U_{t+\delta} - U_t)\|_1 + \|U_{t+\delta}^* - U_t^*\|_1) \\
&\leq \varepsilon,
\end{aligned}$$

where the first and the second inequalities hold since for any matrices A, B and C of the same order and for any $1 \leq p \leq \infty$, we have [80]

$$\|ABC\|_p \leq \|A\|_\infty \|B\|_p \|C\|_\infty.$$

According to the integral representation of operator monotone function a^z (see, for example, [31]), for $0 < z < 1$ and $a > 0$ we have

$$\theta_t^z = \frac{\sin(z\pi)}{\pi} \int_0^\infty x^z \Big(\frac{1}{x} - \frac{1}{x + \theta_t}\Big) dx.$$

Therefore,

$$\begin{aligned}
\|\theta^z_{t+\delta} - \theta^z_t\|_1 &= \Big\|\frac{\sin(z\pi)}{\pi}\int_0^\infty x^z\Big((x+\theta_t)^{-1} - (x+\theta_{t+\delta})^{-1}\Big)dx\Big\|_1 \\
&= \Big\|\frac{\sin(z\pi)}{\pi}\int_0^\infty x^z(x+\theta_t)^{-1}(\theta_{t+\delta}-\theta_t)(x+\theta_{t+\delta})^{-1}dx\Big\|_1 \\
&\leq \frac{1}{\pi}\int_0^\infty x^z\|(x+\theta_t)^{-1}\|_\infty.\|\theta_{t+\delta}-\theta_t\|_1.\|(x+\theta_{t+\delta})^{-1}\|_\infty dx.
\end{aligned}$$

Note that for the positive definite matrix A we have $\|A\|_\infty = \lambda_1^\downarrow(A)$. Then, for $x > 0$ we have

$$\|(x+\theta_t)^{-1}\|_\infty = (x+\lambda_n^\downarrow(\theta_t))^{-1},$$

and

$$\|(x+\theta_{t+\delta})^{-1}\|_\infty = (x+\lambda_n^\downarrow(\theta_{t+\delta}))^{-1}.$$

Consequently,

$$\|(x+\theta_t)^{-1}\|_\infty.\|(x+\theta_{t+\delta})^{-1}\|_\infty \leq (x+c)^{-2},$$

where $c = \min\{\lambda_n^\downarrow(\theta_t), \lambda_n^\downarrow(\theta_{t+\delta})\}$. Therefore,

$$\begin{aligned}
\|\theta^z_{t+\delta} - \theta^z_t\|_1 &\leq \frac{1}{\pi}\|\theta_{t+\delta}-\theta_t\|_1\int_0^\infty \frac{x^z}{(x+c)^2}dx \\
&\leq \frac{\varepsilon}{\pi}\int_c^\infty \frac{(y-c)^z}{y^2}dy \ \ \text{(where } y = x+c\text{)} \\
&\leq \frac{\varepsilon}{\pi}\int_c^\infty \frac{dy}{y^{2-z}} \\
&= k\varepsilon,
\end{aligned}$$

where $k = \dfrac{1}{\pi}\displaystyle\int_c^\infty y^{z-2}dy < \infty$. Finally, we have

$$|g(t+\delta) - g(t)| = |\operatorname{Tr}(\theta^z_{t+\delta}) - \operatorname{Tr}(\theta^z_t)| \leq \|\theta^z_{t+\delta} - \theta^z_t\|_1 \leq k\varepsilon.$$

Finally, we have

$$|g(t+\delta)-g(t)| = |\operatorname{Tr}(\theta^z_{t+\delta}) - \operatorname{Tr}(\theta^z_t)| \leq \|\theta^z_{t+\delta} - \theta^z_t\|_1 \leq k\varepsilon.$$

This means, the function $g(t)$ is continuous in t. By the Intermediate Value Theorem, we have (3.4.25). □

In this chapter, we introduce a new quantum divergence called the α-z-Bures Wasserstein distance, which is an extension with two parameters of the Bures distance. Then we investigate its properties. In particular, we solve the least square problem with respect to this divergence and study its solution. In the next chapter, we introduce a new weighted spectral geometric mean denoted by $\mathcal{F}_t(A,B)$ and study the properties of this quantity. Additionally, we provide some comparisons between $\mathcal{F}_t(A,B)$ and $A \diamond_t B$, which is the solution to the least square problem with respect to the Bures Wasserstein distance.

Chapter 4

A new weighted spectral geometric mean

It is well-known [10] that the geometric mean $A\sharp B$ is the midpoint of the geodesic

$$A\sharp_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}, \quad t \in [0,1],$$

joining A and B under the Riemannian metric $\delta_R(A,B) = \|\log(A^{-1/2}BA^{-1/2})\|_F$, where $\|\cdot\|_F$ denotes the Frobenius norm [11].

The spectral geometric mean of $A, B \in \mathbb{P}_n$ was introduced by Fiedler and Pták in 1997 [37], and one of its formulations is

$$A\natural B := (A^{-1}\sharp B)^{1/2} A (A^{-1}\sharp B)^{1/2}. \tag{4.0.1}$$

It is called *the spectral geometric mean* because $(A\natural B)^2$ is similar to AB and that the eigenvalues of their spectral mean are the positive square roots of the corresponding eigenvalues of AB [37, Theorem 3.2].

Kim and Lee [52] defined the weighted spectral mean:

$$A\natural_t B := \left(A^{-1}\sharp B\right)^t A \left(A^{-1}\sharp B\right)^t, \quad t \in [0,1]. \tag{4.0.2}$$

It is obvious that $A\natural_t B$ is a curve joining A and B. They studied the relative operator entropy related to the spectral geometric mean and several properties similar to those of the relative entropy of Tsallis operator defined via the matrix geometric mean. Recently, Gan, Liu, and Tam [41] and Gan and Tam [40] studied $A\natural_t B$ and obtained some nice properties.

Note that in (4.0.2) the geometric mean $A^{-1}\sharp B$ is a main component of the weighted spectral mean $A\natural_t B$ while the middle term is A, independent of t.

Following that sequence of events, in this chapter we define a new weighted mean, called $\mathcal{F}$-mean.

The results of this chapter are taken from [33].

4.1 A new weighted spectral geometric mean and its basic properties

Definition 4.1.1. Let $A, B \in \mathbb{P}_n$. Define

$$\mathcal{F}_t(A,B) := (A^{-1}\sharp_t B)^{1/2} A^{2-2t} (A^{-1}\sharp_t B)^{1/2}, \quad t \in [0,1]. \tag{4.1.3}$$

It is obvious that $\mathcal{F}_0(A,B) = A$ and $\mathcal{F}_1(A,B) = B$, and hence $\mathcal{F}_t(A,B)$ is a curve joining A and B. For $t = \frac{1}{2}$, $\mathcal{F}_{\frac{1}{2}}(A,B)$ is the spectral geometric mean (4.0.1). We call $\mathcal{F}_t(A,B)$ weighted $\mathcal{F}$-mean and it is different from (4.0.2).

From the Riccati equation, it is obvious that $A\sharp X = B$ if and only if $X = BA^{-1}B$. Therefore, $\mathcal{F}_t(A,B)$ is the unique positive definite solution X to

$$A^{2(t-1)}\sharp X = (A^{-1}\sharp_t B)^{1/2}.$$

Let's recall some known properties of the weighted geometric mean [57] .

Lemma 4.1.1. *Let $A, B, C, D \in \mathbb{P}_n$ and $t \in [0,1]$. We have*

1. $A\sharp_t B = A^{1-t}B^t$ *if A and B commute.*

2. $(aA)\sharp_t(bB) = a^{1-t}b^t(A\sharp_t B)$ *for* $a, b > 0$.

3. $A\sharp_t B = B\sharp_{1-t}A$.

4. $(A\sharp_t B)^{-1} = A^{-1}\sharp_t B^{-1}$.

5. $U^*(A\sharp_t B)U = (U^*AU)\sharp_t(U^*BU)$ *for any* $U \in U(n)$.

6. *(Löwner-Heinz)* $A\sharp_t B \leq C\sharp_t D$ *if* $A \leq C$, $B \leq D$.

7. $(\lambda A + (1-\lambda)B)\sharp_t(\lambda C + (1-\lambda)D) \geq \lambda(A\sharp_t C) + (1-\lambda)(B\sharp_t D)$, *for* $\lambda \in [0, 1]$.

8. $((1-t)A^{-1} + tB^{-1})^{-1} \leq A\sharp_t B \leq (1-t)A + tB$.

The following proposition lists some basic properties of $\mathcal{F}_t(A, B)$. Some properties are similar to those of weighted geometric mean [57], and are not hard to prove. Proofs are presented here for the sake of completeness.

Proposition 4.1.1. *Let* $A, B \in \mathbb{P}_n$. *The following properties hold for all* $t \in [0, 1]$.

1. $\mathcal{F}_t(A, B) = A^{1-t}B^t$ *if* A *and* B *commute.*

2. $\mathcal{F}_t(aA, bB) = a^{1-t}b^t\mathcal{F}_t(A, B)$ *for* $a, b > 0$.

3. $U^*\mathcal{F}_t(A, B)U = \mathcal{F}_t(U^*AU, U^*BU)$ *for* $U \in U(n)$.

4. $\mathcal{F}_t^{-1}(A, B) = \mathcal{F}_t(A^{-1}, B^{-1})$.

5. $\det \mathcal{F}_t(A, B) = (\det A)^{1-t}(\det B)^t$.

6. $2((1-t)A + tB^{-1})^{-1/2} - A^{2(t-1)} \leq \mathcal{F}_t(A, B) \leq [2((1-t)A^{-1} + tB)^{-1/2} - A^{-2(t-1)}]^{-1}$.

Proof. (1) Since A and B commute, so are A^{-1} and B. Thus $A^{-1}\sharp_t B = (A^{-1})^{1-t}B^t$ and we have

$$\mathcal{F}_t(A, B) = (A^{-1}\sharp_t B)^{1/2}A^{2-2t}(A^{-1}\sharp_t B)^{1/2} = (A^{-1+t}B^t)^{1/2}A^{2-2t}(A^{-1+t}B^t)^{1/2} = A^{1-t}B^t.$$

(2) For any $a, b > 0$, we have $(aA)\sharp_t(bB) = a^{1-t}b^t(A\sharp_t B)$. Consequently,

$$\begin{aligned}
\mathcal{F}_t(aA, bB) &= \Big((aA)^{-1}\sharp_t(bB)\Big)^{1/2}(aA)^{2-2t}\Big((aA)^{-1}\sharp_t(bB)\Big)^{1/2} \\
&= a^{1-t}b^t(A^{-1}\sharp_t B)^{1/2}A^{2-2t}(A^{-1}\sharp_t B)^{1/2} \\
&= a^{1-t}b^t\mathcal{F}_t(A, B).
\end{aligned}$$

(3) Note that $U^*(A\sharp_t B)^{1/2}U = (U^*(A\sharp_t B)U)^{1/2}$ and $U^*A^{2-2t}U = (U^*AU)^{2-2t}$ for any $U \in U(n)$. Then

$$\begin{aligned}
U^*\mathcal{F}_t(A, B)U &= U^*(A\sharp_t B)^{1/2}A^{2-2t}(A\sharp_t B)^{1/2}U \\
&= U^*(A\sharp_t B)^{1/2}UU^*A^{2-2t}UU^*(A\sharp_t B)^{1/2}U \\
&= ((U^*\left(A\sharp_t B\right)U)^{1/2}\left(U^*AU\right)^{2-2t}\left(U^*\left(A\sharp_t B\right)U\right)^{1/2} \\
&= \mathcal{F}_t(U^*AU, U^*BU),
\end{aligned}$$

where the last equality follows from $U^*(A\sharp_t B)U = (U^*AU)\sharp_t(U^*BU)$.

(4) Applying $(A\sharp_t B)^{-1} = A^{-1}\sharp_t B^{-1}$, we obtain

$$\begin{aligned}
\mathcal{F}_t(A, B)^{-1} &= \left[\left(A^{-1}\sharp_t B\right)^{1/2}A^{2-2t}\left(A^{-1}\sharp_t B\right)^{1/2}\right]^{-1} \\
&= \left(A^{-1}\sharp_t B\right)^{-1/2}A^{2t-2}\left(A^{-1}\sharp_t B\right)^{-1/2} \\
&= \left(A\sharp_t B^{-1}\right)^{1/2}A^{2t-2}\left(A\sharp_t B^{-1}\right)^{1/2} \\
&= \mathcal{F}_t(A^{-1}, B^{-1}).
\end{aligned}$$

(5) Since $\det(AB) = \det A \,\det B$, we obtain

$$\det\mathcal{F}_t(A, B) = \det\left(A^{-1}\sharp_t B\right)\det\left(A^{2-2t}\right) = (\det A)^{t-1}(\det B)^t(\det A)^{2-2t} = (\det A)^{1-t}(\det B)^t.$$

(6) Let $X = \mathcal{F}_t(A, B)$. By the Arithmetic-Geometric-Harmonic inequality and the operator

monotonicity of the function $X \mapsto X^t$ when $t \in [0,1]$, we have

$$\left(\frac{A^{-2(t-1)} + X^{-1}}{2}\right)^{-1} \leq A^{2(t-1)} \sharp X = (A^{-1} \sharp_t B)^{1/2} \leq \Big((1-t)A^{-1} + tB\Big)^{1/2}. \tag{4.1.4}$$

Then we have

$$\frac{A^{-2(t-1)} + X^{-1}}{2} \geq \Big((1-t)A^{-1} + tB\Big)^{-1/2}.$$

Hence

$$X^{-1} \geq 2\Big((1-t)A^{-1} + tB\Big)^{-1/2} - A^{-2(t-1)}.$$

Consequently,

$$X \leq \left[2\Big((1-t)A^{-1} + tB\Big)^{-1/2} - A^{-2(t-1)}\right]^{-1}.$$

Since $\mathcal{F}_t(A,B) = (\mathcal{F}_t(A^{-1}, B^{-1}))^{-1}$, we obtain the first inequality.

Using the second inequality in (4.1.4) and similar arguments, one can prove the second inequality. $\square$

Remark 4.1.1. An analog of Lemma 4.1.1(3) for $\mathcal{F}_t(A,B)$ is not true, i.e., the equality $\mathcal{F}_t(A,B) = \mathcal{F}_{1-t}(B,A)$ does not hold. Indeed, from the last identity we have

$$(A^{-1} \sharp_t B)^{1/2} A^{2-2t} (A^{-1} \sharp_t B)^{1/2} = (A^{-1} \sharp_t B)^{-1/2} B^{2t} (A^{-1} \sharp_t B)^{-1/2},$$

or equivalently,

$$B^{2t} = (A^{-1} \sharp_t B) A^{2-2t} (A^{-1} \sharp_t B).$$

According to the Riccati equation, it implies that

$$A^{-1} \sharp_t B = B^{2t} \sharp A^{2t-2}$$

which is not true.

4.2 The Lie-Trotter formula and weak log-majorization

Let $B(\mathcal{H})$ be the Banach space of bounded operators on Hilbert space $\mathcal{H}$ and $P(\mathcal{H})$ be the open convex cone of positive definite operators. A straightforward outcome of calculus applied to mappings on operators and operator-valued functions is the possibility to expand the classical Lie-Trotter formula in the following manner.

Proposition 4.2.1 ([1]). *For any differentiable curve $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{P}_n$ with $\gamma(0) = I$,*

$$e^{\gamma'(0)} = \lim_{t\to 0} \gamma^{1/t}(t) = \lim_{n\to\infty} \gamma^n(1/n).$$

Indeed, the exponential function $e : B(\mathscr{H}) \to P(\mathscr{H})$ and the logarithm function $\log : P(\mathscr{H}) \to B(\mathscr{H})$ are both well-defined and diffeomorphic. The derivative of the exponential function at the origin $0 \in B(\mathscr{H})$ is the identity map on $B(\mathscr{H})$. Consequently, the derivative of its inverse function $\log$ at the identity operator $I \in P(\mathscr{H})$ is the identity map on $B(\mathscr{H})$. Therefore

$$\begin{aligned} \gamma'(0) = (\log \circ \gamma)'(0) &= \lim_{t\to 0} \frac{\log(\gamma(t)) - \log(\gamma(0))}{t} = \lim_{t\to 0} \frac{\log(\gamma(t))}{t} \\ &= \lim_{t\to 0} \log\left(\gamma(t)^{1/t}\right) = \lim_{n\to\infty} \log\left(\gamma(1/n)^n\right) \end{aligned}.$$

Notice that for $X, Y \in \mathbb{H}_n$ and $\alpha \in [0, 1]$, the following curves are smooth and pass through the identity matrix I at $t = 0$:

$$\begin{aligned} \gamma_1(t) &= e^{t(1-\alpha)X/2} e^{t\alpha Y} e^{t(1-\alpha)X/2}, \\ \gamma_2(t) &= (1-\alpha)e^{tX} + \alpha e^{tY}, \\ \gamma_3(t) &= ((1-\alpha)e^{-tX} + \alpha e^{-tY})^{-1}, \\ \gamma_4(t) &= e^{tX} \sharp_\alpha e^{tY}, \\ \gamma_5(t) &= e^{tX} \natural_\alpha e^{tY}. \end{aligned}$$

Applying Proposition 4.2.1 one obtains the following Lie-Trotter formulas:

$$\begin{aligned}
e^{(1-\alpha)X+\alpha Y} &= \lim_{n\to\infty}(e^{t(1-\alpha)X/2n}e^{t\alpha Y/n}e^{t(1-\alpha)X/2n})^n \\
&= \lim_{n\to\infty}((1-\alpha)e^{tX/n}+\alpha e^{tY/n})^n \\
&= \lim_{n\to\infty}((1-\alpha)e^{-tX/n}+\alpha e^{-tY/n})^{-n} \\
&= \lim_{n\to\infty}(e^{tX/n}\sharp_\alpha e^{tY/n})^n \\
&= \lim_{n\to\infty}(e^{tX/n}\natural_\alpha e^{tY/n})^n.
\end{aligned}$$

In next theorem, we show the Lie-Trotter formula for $\mathcal{F}_t$, namely,

$$\lim_{p\to 0}\mathcal{F}_t^{1/p}(e^{pA},e^{pB}) = e^{(1-t)A+tB},$$

when $A, B \in \mathbb{H}_n$ and $t \in [0,1]$.

Theorem 4.2.1. *Let* $A, B \in \mathbb{H}_n$ *and* $t \in [0,1]$. *Then*

$$\lim_{p\to 0}\mathcal{F}_t^{1/p}(e^{pA},e^{pB}) = e^{(1-t)A+tB}.$$

Proof. Since $\mathcal{F}_t^{-1}(A,B) = \mathcal{F}_t(A^{-1},B^{-1})$ we have

$$\lim_{p\to 0^-}\mathcal{F}_t^{-1/p}\left(e^{pA},e^{pB}\right) = \lim_{p\to 0^-}\mathcal{F}_t^{-1/p}\left(e^{-pA},e^{-pB}\right) = \lim_{p\to 0^+}\mathcal{F}_t^{1/p}\left(e^{pA},e^{pB}\right).$$

So we only need to prove

$$\lim_{p\to 0^+}\mathcal{F}_t(e^{pA},e^{pB})^{1/p} = e^{(1-t)A+tB}.$$

For $p \in (0,1)$ we may express $p = \frac{1}{m+s}$, where $m \in \mathbb{N}$, and $s \in (0,1)$. Set

$$X(p) := \mathcal{F}_t(e^{pA},e^{pB}), \qquad Y(p) := e^{p[(1-t)A+tB]}.$$

We have

$$\begin{aligned}
& \|\mathcal{F}_t(e^{pA}, e^{pB})^{1/p} - e^{(1-t)A+tB}\| \\
= \ & \|X(p)^{1/p} - Y(p)^{1/p}\| \\
\leq \ & \|X(p)^{1/p} - X(p)^m\| + \|X(p)^m - Y(p)^m\| + \|Y(p)^m - Y(p)^{1/p}\|. \qquad (4.2.5)
\end{aligned}$$

By [62, Theorem 1.1],

$$e^{pA}\sharp_t e^{pB} \prec_{\log} e^{p[(1-t)A+tB]}$$

so we have

$$\|e^{pA}\sharp_t e^{pB}\| \leq \|Y(p)\| \leq e^{p[(1-t)\|A\|+t\|B\|]}.$$

Therefore,

$$\begin{aligned}
\|X(p)\| &= \| \left(e^{-pA}\sharp_t e^{pB}\right)^{\frac{1}{2}} e^{p(2-2t)A}(e^{-pA}\sharp_t e^{pB})^{\frac{1}{2}}\| \\
&\leq \|e^{-pA}\sharp_t e^{pB}\|^{\frac{1}{2}}\|e^{p(2-2t)A}\| \, \|e^{-pA}\sharp_t e^{pB}\|^{\frac{1}{2}} \\
&\leq e^{\frac{p}{2}[(1-t)\|A\|+t\|B\|]} e^{p(2-2t)\|A\|} e^{\frac{p}{2}[(1-t)\|A\|+t\|B\|]} \\
&= e^{p[(3-3t)\|A\|+2t\|B\|]}.
\end{aligned}$$

As $pm \leq 1$, we have $\|X(p)\|^m \leq e^{pm[(3-3t)\|A\|+2t\|B\|]} \leq e^{(3-3t)\|A\|+2t\|B\|} < \infty$. Consequently, the first term in (4.2.5)

$$\|X(p)^{1/p} - X(p)^m\| = \|X(p)^{m+s} - X(p)^m\| \leq \|X(p)\|^m\|X(p)^s - I\| \to 0 \quad \text{as} \quad p \to 0^+,$$

since $X(p) \to I$ as $p \to 0^+$ by (4.1.3) and $s \in (0,1)$. Similarly, the third term in (4.2.5)

$$\|Y(p)^m - Y(p)^{1/p}\| = \|Y(p)^m - Y(p)^{m+s}\| \leq \|Y(p)\|^m\|I - Y(p)^s\| \to 0 \quad \text{as} \quad p \to 0^+$$

Now the second term in (4.2.5)

$$\|X(p)^m - Y(p)^m\| = \|\sum_{j=0}^{m-1} X(p)^{m-1-j}(X(p)-Y(p))Y(p)^j\| \leq mM^{m-1}\|X(p)-Y(p)\|,$$

where $M := \max\{\|X(p)\|, \|Y(p)\|\}$. As $p(m-1) \leq 1$, we have

$$\begin{aligned} M^{m-1} &\leq \max\left\{e^{p(m-1)[(3-3t)\|A\|+2t\|B\|}, e^{p(m-1)[(1-t)\|A\|+t\|B\|]}\right\} \\ &\leq \max\left\{e^{(3-3t)\|A\|+2t\|B\|}, e^{(1-t)\|A\|+t\|B\|}\right\} \\ &< \infty. \end{aligned}$$

Using the power series expansion of the matrix exponential $e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$, we have

$$\begin{aligned} e^{-pA}\sharp_t e^{pB} &= e^{\frac{-pA}{2}}\left(e^{\frac{pA}{2}}e^{pB}e^{\frac{pA}{2}}\right)^t e^{\frac{-pA}{2}} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{-pA}{2}\right)^k\left[\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{pA}{2}\right)^k\sum_{k=0}^{\infty}\frac{(pB)^k}{k!}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{pA}{2}\right)^k\right]\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{-pA}{2}\right)^k \\ &= \left(I - \frac{pA}{2} + o(p)\right)\left[\left(I + \frac{pA}{2} + o(p)\right)(I + pB + o(p))\left(I + \frac{pA}{2} + o(p)\right)\right]^t \\ &\cdot\left(I - \frac{pA}{2} + o(p)\right) \\ &= \left(I - \frac{pA}{2} + o(p)\right)[I + p(A+B) + o(p)]^t\left(I - \frac{pA}{2} + o(p)\right) \\ &= I + p[-(1-t)A + tB] + o(p) \end{aligned}$$

and

$$e^{p(2-2t)A} = \sum_{k=0}^{\infty}\frac{1}{k!}(p(2-2t)A)^k = I + p(2-2t)A + o(p).$$

Hence

$$
\begin{aligned}
X(p) &= \left(e^{-pA}\sharp_t e^{pA}\right)^{\frac{1}{2}} e^{p(2-2t)A} \left(e^{-pA}\sharp_t e^{pB}\right)^{\frac{1}{2}} \\
&= [I + p(-(1-t)A + tB) + o(p)]^{\frac{1}{2}} [I + p(2-2t)A + o(p)] [I + p(-(1-t)A + tB)] \frac{1}{2} \\
&= \left[I + \frac{p}{2}(-(1-t)A + tB) + o(p)\right] [I + p(2-2t)A + o(p)] \left[I + \frac{p}{2}(-(1-t)A + tB) + o(p)\right] \\
&= I + p((1-t)A + tB) + o(p).
\end{aligned}
$$

As $Y(p) := e^{p[(1-t)A+tB]} = I + p((1-t)A + tB) + o(p)$, we have $\|X(p) - Y(p)\| \leq cp^2$ for some constant c. Then

$$
\|X(p)^m - Y(p)^m\| \leq mM^{m-1}cp^2 \leq \frac{m}{(m+s)} M^{m-1}cp \to 0 \quad \text{as} \quad p \to 0^+,
$$

since M^{m-1} is bounded. Thus all three terms in (4.2.5) converge to 0 as $p \to 0^+$ and hence the proof is completed. $\square$

Theorem 4.2.2. *Let $(\alpha_1, ..., \alpha_{m-1}) \in \mathbb{R}^{m-1}$, and $X_1, X_2, ..., X_m \in \mathbb{H}_n$. The curve*

$$
\gamma(t) := \mathcal{F}_{\alpha_{m-1}}\left(e^{tX_m}, \mathcal{F}_{\alpha_m-2}\left(e^{tX_{m-1}}, \mathcal{F}_{\alpha_m-3}(...\mathcal{F}_{\alpha_1}(e^{tX_2}, e^{tX_1})...)\right)\right)
$$

is a differentiable curve with $\gamma(0) = I$ and

$$
\gamma'(0) = \sum_{k=1}^{m} \prod_{i=k}^{m} \alpha_i \left(1 - \alpha_{k-1}\right) X_k,
$$

where $\alpha_0 = 0$ and $\alpha_m = 1$. In particular, if $\alpha_k = \dfrac{k}{k+1}$, for $k = 1, 2, ..., m-1$ then $\gamma'(0) = \dfrac{1}{m}\sum_{k=1}^{m} X_k$.

Proof. Let

$$\begin{aligned}\beta(t) := \mathcal{F}_{\alpha_1}\Big(e^{tX_2}, e^{tX_1}\Big) &= \Big(e^{-tX_2}\sharp_{\alpha_1}e^{tX_1}\Big)^{\frac{1}{2}} e^{t(2-2\alpha_1)X_2}\Big(e^{-tX_2}\sharp_{\alpha_1}e^{tX_1}\Big)^{\frac{1}{2}} \\ &= \varphi(t)^{\frac{1}{2}e^{t(2-2\alpha_1)X_2}\varphi(t)^{\frac{1}{2}}},\end{aligned}$$

where $\varphi(t) = e^{-tX_2}\sharp_{\alpha_1}e^{tX_1} = e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1}e^{-\frac{tX_2}{2}}$. We have

$$\begin{aligned}&\frac{d}{dt}\varphi(t) \\ =\ & -\frac{X_2}{2}e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1}e^{-\frac{tX_2}{2}} - e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1}e^{-\frac{tX_2}{2}}\frac{X_2}{2} \\ & +\alpha_1 e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1-1}\frac{d}{dt}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)e^{-\frac{tX_2}{2}} \\ =\ & -\frac{X_2}{2}e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1}e^{-\frac{tX_2}{2}} - e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1}e^{-\frac{tX_2}{2}}\frac{X_2}{2} \\ & +\alpha_1 e^{-\frac{tX_2}{2}}\Big(e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\Big)^{\alpha_1-1}\Big(\frac{X_2}{2}e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}} + e^{\frac{tX_2}{2}}e^{tX_1}e^{\frac{tX_2}{2}}\frac{X_2}{2} + e^{\frac{tX_2}{2}}e^{tX_1}X_1e^{\frac{tX_2}{2}}\Big)e^{-\frac{tX_2}{2}}.\end{aligned}$$

Therefore

$$\frac{d}{dt}\varphi(t)\Big|_{t=0} = -X_2 + \alpha_1(X_2+X_1) = (\alpha_1-1)X_2 + \alpha_1X_1.$$

On the other hand,

$$\begin{aligned}\frac{d}{dt}\beta(t) &= \frac{1}{2}\varphi(t)^{-\frac{1}{2}}\frac{d}{dt}\varphi(t)e^{t(2-2\alpha_1)X_2}\varphi(t)^{\frac{1}{2}} + \frac{1}{2}\varphi(t)^{\frac{1}{2}}e^{t(2-2\alpha_1)X_2}\varphi(t)^{-\frac{1}{2}}\frac{d}{dt}\varphi(t) \\ &\quad +(2-2\alpha_1)\varphi(t)^{\frac{1}{2}}e^{t(2-2\alpha_1)X_2}X_2\varphi(t)^{\frac{1}{2}}.\end{aligned}$$

Thus,

$$\frac{d}{dt}\beta(t)\Big|_{t=0} = (\alpha_1-1)X_2 + \alpha_1X_1 + (2-2\alpha_1)X_2 = (1-\alpha_1)X_2 + \alpha_1X_1.$$

Set

$$\xi(t) := \mathcal{F}_{\alpha_m}\Big(e^{tX_{m+1}}, \gamma(t)\Big) = L(t)^{\frac{1}{2}}e^{t(2-2\alpha_m)X_{m+1}}L(t)^{\frac{1}{2}},$$

where $L(t) = e^{-tX_{m+1}} \sharp_{\alpha_m} \gamma(t)$. Since

$$\begin{aligned} \frac{d}{dt}\xi(t) &= \frac{1}{2}L(t)^{-\frac{1}{2}}\frac{d}{dt}L(t)e^{t(2-2\alpha_m)X_{m+1}}L(t)^{\frac{1}{2}} + \frac{1}{2}L(t)^{\frac{1}{2}}e^{t(2-2\alpha_m)X_{m+1}}L(t)^{-\frac{1}{2}}\frac{d}{dt}L(t) \\ &\quad +(2-2\alpha_m)L(t)^{\frac{1}{2}}e^{t(2-2\alpha_m)X_{m+1}}X_{m+1}L(t)^{\frac{1}{2}}, \end{aligned}$$

by the previous argument, we have

$$\left.\frac{d}{dt}L(t)\right|_{t=0} = -X_{m+1} + \alpha_m(X_{m+1} + \gamma'(0)).$$

Therefore,

$$\left.\frac{d}{dt}\xi(t)\right|_{t=0} = (1-\alpha_m)X_{m+1} + \sum_{k=1}^{m}\prod_{i=k}^{m}\alpha_i\left(1-\alpha_{k-1}\right)X_k = \sum_{k=1}^{m+1}\prod_{i=k}^{m+1}\alpha_i\left(1-\alpha_{k-1}\right)X_k,$$

where $\alpha_0 = 0$ and $\alpha_{m+1} = 1$. $\square$

The Wasserstein distance or Bures distance of $A, B \in \mathbb{P}_m$ is the Riemannian metric given by [13, 42]

$$d_b(A,B) = \left[\operatorname{tr}\left(\frac{A+B}{2}\right) - \operatorname{tr}\left(A^{1/2}BA^{1/2}\right)^{1/2}\right]^{1/2}.$$

The Wasserstein mean of $A_1, A_2, ..., A_m$ belonging to $\mathbb{P}_n$ is the solution to the least squares mean problem for the Wasserstein distance, which is defined as follows:

$$\Omega\left(\omega; A_1, \ldots, A_m\right) = \underset{X\in\mathbb{P}_n}{\arg\min}\sum_{j=1}^{m} w_j d_b^2\left(X, A_j\right),$$

Here, ω is a positive probability vector represented as $\omega = (w_1, \ldots, w_m)$. Specifically, when there are two distributions (i.e., when $n = 2$), the Wasserstein mean of A and B with respect to $\omega = (1-t, t)$ where $t \in [0,1]$ is exactly

$$A \diamond_t B = A^{-1/2}\left[(1-t)A + t\left(A^{1/2}BA^{1/2}\right)^{1/2}\right]^2 A^{-1/2}.$$

Now we compare the weak log-majorization between the $\mathcal{F}$-mean and the Wasserstein mean.

Theorem 4.2.3. *Let $A, B \in \mathbb{P}_n$ and $t \in [0,1]$.*

(i) If $0 \le t \le \frac{1}{2}$ then

$$\mathcal{F}_t(A,B) \prec_{w\log} A \diamond_t B;$$

(ii) If $\frac{1}{2} \le t \le 1$ then

$$\mathcal{F}_{1-t}(B,A) \prec_{w\log} A \diamond_t B.$$

Proof. By using the technique of the k-th antisymmetric tensor power, we only need to prove that $A \diamond_t B \le I$ implies $\mathcal{F}_t(A,B) \le I$. This is equivalent to proving that $\lambda_1(A \diamond_t B) \le 1$, which in turn implies $\lambda_1(\mathcal{F}_t(A,B)) \le 1$, when $0 \le t \le \frac{1}{2}$. The same reasoning applies to the second inequality.

(i) Let $0 \le t \le \frac{1}{2}$, set $C = A^{-1}\sharp_t B = A^{-1/2}(A^{1/2}BA^{1/2})^t A^{-1/2}$. Consequently,

$$(A^{1/2}BA^{1/2})^{1/2} = (A^{1/2}CA^{1/2})^{1/2t}.$$

Assuming that $A \diamond_t B \le I$. This is equivalent to

$$\left[(1-t)A + t\left(A^{1/2}BA^{1/2}\right)^{1/2}\right]^2 \le A,$$

since map $x \mapsto x^{1/2}$ is operator monotone, then we have

$$\begin{aligned} & (1-t)A + t(A^{1/2}BA^{1/2})^{1/2} \le A^{1/2} \\ \Leftrightarrow \quad & (1-t)A + t(A^{1/2}CA^{1/2})^{1/2t} \le A^{1/2}. \end{aligned}$$

This leads to

$$(A^{1/2}CA^{1/2})^{1/2t} \le \left(1 - \frac{1}{t}\right)A + \frac{1}{t}A^{1/2}.$$

Since $2t \in [0,1]$, then we have

$$A^{1/2}CA^{1/2} \leq \Big((1-\frac{1}{t})A + \frac{1}{t}A^{1/2}\Big)^{2t}.$$

Thus,

$$C \leq A^{-1/2}\Big((1-\frac{1}{t})A + \frac{1}{t}A^{1/2}\Big)^{2t}A^{-1/2}.$$

Now,

$$\begin{aligned}
\lambda_1(\mathcal{F}_t(A,B)) &= \lambda_1(C^{1/2}A^{2-2t}C^{1/2}) \\
&= \lambda_1(A^{1-t}CA^{1-t}) \\
&\leq \lambda_1(A^{1/2-t}((1-\frac{1}{t})A + \frac{1}{t}A^{1/2})^{2t}A^{1/2-t}) \\
&= \lambda_1\Big(((1-\frac{1}{t})A + \frac{1}{t}A^{1/2})^{2t}A^{1-2t}\Big).
\end{aligned}$$

Since $A > 0$, there exists a unitary matrix U and a diagonal matrix $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ such that $A = UDU^*$. Therefore

$$\begin{aligned}
\Big((1-\frac{1}{t})A + \frac{1}{t}A^{1/2}\Big)^{2t}A^{1-2t} &= U\Big((1-\frac{1}{t})D + \frac{1}{t}D^{1/2}\Big)^{2t}D^{1-2t}U^* \\
&= UEU^*,
\end{aligned}$$

where $E = \mathrm{diag}((1-\frac{1}{t})\lambda_1 + \frac{1}{t}\lambda_1^{1/2})^{2t}\lambda_1^{1-2t}, ..., ((1-\frac{1}{t})\lambda_n + \frac{1}{t}\lambda_1^{1/2})^{2t}\lambda_n^{1-2t})$.

Now we prove

$$\Big((1-\frac{1}{t})x + \frac{1}{t}x^{1/2}\Big)^{2t}x^{1-2t} = \Big((1-\frac{1}{t})x^{1/2t} + \frac{1}{t}x^{(1-t)/2t}\Big)^{2t} \leq 1,$$

where $x > 0$ and $0 < t \leq \dfrac{1}{2}$. This is equivalent to

$$f(x) := \Big(\frac{t-1}{t}x^{1/2t} + \frac{1}{t}x^{(1-t)/2t}\Big) \leq 1,$$

where $x > 0$ and $0 < t \leq \dfrac{1}{2}$. We have

$$\begin{aligned} f'(x) &= \frac{t-1}{2t^2}x^{(1-2t)/2t} + \frac{1-t}{2t^2}x^{(1-3t)/2t} \\ &= \frac{t-1}{2t^2}x^{(1-3t)/2t}(x^{1/2}-1). \end{aligned}$$

Thus, $f'(x) = 0$ if only if $x = 1$. Hence, $f(x)$ attains its maximum at $f(1) = 1$, and $f(x) \leq 1$, for all $0 \leq x \leq \frac{1}{2}$ and $x > 0$. Therefore, $\lambda_1(\mathcal{F}_t(A, B)) \leq 1$, which implies $F_t(A, B) \leq I$.

(ii) Let $\frac{1}{2} \leq t \leq 1$, set $C = B^{-1}\sharp_{1-t}A = B^{-1/2}(B^{1/2}AB^{1/2})^{1-t}B^{-1/2}$. Consequently, $B^{1/2}CB^{1/2} = (B^{1/2}AB^{1/2})^{1-t}$. This implies $(B^{1/2}CB^{1/2})^{1/(2-2t)} = (B^{1/2}AB^{1/2})^{1/2}$. Recall that,

$$A \diamond_t B = B \diamond_{1-t} A = B^{-1/2}(tB + (1-t)(B^{1/2}AB^{1/2})^{1/2})^2B^{-1/2}.$$

If $A \diamond_t B \leq I$, then we have

$$tB + (1-t)(B^{1/2}AB^{1/2})^{1/2} \leq B^{1/2}.$$

This is equivalent to

$$(B^{1/2}CB^{1/2})^{1/(2-2t)} \leq \frac{t}{t-1}B + \frac{1}{1-t}B^{1/2},$$

since the map $x \mapsto x^{2-2t}$ is operator monotone when $\frac{1}{2} \leq t \leq 1$. Then we have

$$B^{1/2}CB^{1/2} \leq \Big(\frac{t}{t-1}B + \frac{1}{1-t}B^{1/2}\Big)^{2-2t}.$$

Hence,

$$C \leq B^{-1/2}\Big(\frac{t}{t-1}B + \frac{1}{1-t}B^{1/2}\Big)^{2-2t}B^{-1/2}.$$

Now,

$$\begin{aligned}
\lambda_1(\mathcal{F}_{1-t}(B,A)) &= \lambda_1(C^{1/2}B^{2t}C^{1/2}) \\
&= \lambda_1(B^tCB^t) \\
&\leq \lambda_1\Big(B^{t-1/2}(\frac{t}{t-1}B + \frac{1}{1-t}B^{1/2})^{2-2t}B^{t-1/2}\Big) \\
&= \lambda_1\Big((\frac{t}{t-1}B + \frac{1}{1-t}B^{1/2})^{2-2t}B^{2t-1}\Big).
\end{aligned}$$

Since $B > 0$, there exists a unitary matrix U and a diagonal matrix $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ such that $B = UDU^*$. Therefore,

$$\begin{aligned}
\Big(\frac{t}{t-1}B + \frac{1}{1-t}B^{1/2}\Big)^{2-2t}B^{2t-1} &= U\Big(\frac{t}{t-1}D + \frac{1}{1-t}D^{1/2}\Big)^{2-2t}D^{2t-1}U^* \\
&= UEU^*,
\end{aligned}$$

where $E = \mathrm{diag}\Big((\frac{t}{t-1}\lambda_1 + \frac{1}{1-t}\lambda_1^{1/2})^{2-2t}\lambda_1^{2t-1}, ..., (\frac{t}{t-1}\lambda_n + \frac{1}{1-t}\lambda_n^{1/2})^{2-2t}\lambda_n^{2t-1}\Big)$. Now we prove

$$\Big(\frac{t}{t-1}x + \frac{1}{1-t}x^{1/2})^{2-2t} = (\frac{t}{t-1}x^{1(2-2t)} + \frac{1}{1-t}x^{t/(2-2t)}\Big)^{2-2t} \leq 1,$$

where $\frac{1}{2} \leq t \leq 1$ and $x > 0$. Let

$$f(x) = \frac{t}{t-1}x^{1(2-2t)} + \frac{1}{1-t}x^{t/(2-2t)},$$

where $\frac{1}{2} \leq t \leq 1$ and $x > 0$. We have

$$\begin{aligned} f'(x) &= \frac{t}{(t-1)(2-2t)} x^{(2t-1)/(2-2t)} + \frac{t}{(1-t)(2-2t)} x^{(3t-2)/2-2t} \\ &= \frac{t}{(t-1)(2-2t)} x^{(2t-1)/(2-2t)} (1 - x^{-1/2}) = 0. \end{aligned}$$

Thus, $f'(x) = 0$ if only if $x = 1$. Hence, $f(x)$ attains its maximum at $f(1) = 1$ and $f(x) \leq 1$, for all $\frac{1}{2} \leq x \leq 1$ and $x > 0$. Therefore, $\lambda_1(\mathcal{F}_{1-t}(B, A)) \leq 1$, that is $\mathcal{F}_{1-t}(B, A) \leq I$.

$\square$

In this chapter, we introduce a new spectral geometric mean, called the $\mathcal{F}$-mean. Besides providing some basic properties of this quantity, we prove that the $\mathcal{F}$-mean satisfies the Lie-Trotter formula, and then we compare it with the solution of the least square problem with respect to the Bures distance.

Conclusions

This thesis obtained the following main results:

1. We introduce a new Weighted Hellinger distance, denoted as $d_{h,\alpha}(A,B)$, and prove that it acts as an interpolating metric between the Log-Euclidean and Hellinger metrics. Additionally, we establish the equivalence between the weighted Bures-Wasserstein and weighted Hellinger distances. Moreover, we demonstrate that both distances satisfy the in-betweenness property. Moreover, we also show that among symmetric means, the arithmetic mean is the only one that satisfies the in-betweenness property in the weighted Bures-Wasserstein and weighted Hellinger distances.

2. We construct a new quantum divergence called the α-z-Bures-Wasserstein divergence and demonstrate that this divergence satisfies the in-betweenness property and the data processing inequality in quantum information theory. Furthermore, we solve the least squares problem with respect to this divergence and establish that the solution to this problem corresponds exactly to the unique positive solution of the matrix equation

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

where $Q_{\alpha,z}(A,B) = \left(A^{\frac{1-\alpha}{2z}} B^{\frac{\alpha}{z}} A^{\frac{1-\alpha}{2z}}\right)^z$ and $0 < \alpha \leq z \leq 1$. Afterwards, we proceed to study the properties of the solution to this problem and achieve several significant results. In addition, we provide an inequality for quantum fidelity and its parameterized versions. Then, we utilize α-z-fidelity to measure the distance between two quantum orbits.

3. We introduce a new weighted geometric mean, called the $\mathcal{F}$-mean. We establish some properties for the $\mathcal{F}$-mean and prove that it satisfies the Lie-Trotter formula, Furthermore, we provide a comparison in weak-log majorization between the $\mathcal{F}$-mean and the Wasserstein mean.

Further investigation

In the future, we intend to continue the investigation in the following directions:

- Construct some new distance function based on non-Kubo-Ando means.
- Construct a new distance function between two matrices with different dimensions.
- For $X, Y > 0$ and $0 < t < 1$, verify whether the two quantities

$$\Phi_1(X,Y) = \mathrm{Tr}((1-t)X + tY) - \mathrm{Tr}\left(X \natural_t Y\right)$$

  and

$$\Phi_2(X,Y) = \mathrm{Tr}((1-t)X + tY) - \mathrm{Tr}\left(\mathcal{F}_t(X,Y)\right)$$

  are divergences and simultaneously solve related problems.
- Quantity $\mathcal{F}_t(X,Y)$ is new; therefore, we need to establish new properties for this quantity while also comparing it with the previously known means.

List of Author's related to the thesis

1. Vuong T.D., Vo B.K (2020), "An inequality for quantum fidelity", *Quy Nhon Univ. J. Sci.*, 4 (3).

2. Dinh T.H., Le C.T., Vo B.K, Vuong T.D. (2021), "Weighted Hellinger distance and in betweenness property", *Math. Ine. Appls.,* 24, 157-165.

3. Dinh T.H., Le C.T., Vo B.K., Vuong T.D. (2021), "The α-z-Bures Wasserstein divergence", *Linear Algebra Appl.,* 624, 267-280.

4. Dinh T.H., Le C.T., Vuong T.D., α-z-fidelity and α-z-weighted right mean, *Submitted.*

5. Dinh T.H., Tam T.Y., Vuong T.D, On new weighted spectral geometric mean, *Submitted.*

Bibliography

[1] Ahn E., Kim S., Lim Y. (2007), "An extended Lie-Trotter formula and its applications", *Linear Algebra Appl.*, 427, 190-196.

[2] Absil P.A., Mahony R., Sepulchresea. R (2007), *Optimization Algorithms on Matrix Manifolds*, Princeton.

[3] Anderson W., Duffin W. (1969), "Series and parallel addition of matrices", *J. Math. Anal. Appl.,* 26, 576-594.

[4] Araki H. (1971), "A remark on Bures distance function for normal states", *RIMS Kyoto Uni.,* 6, 477-482.

[5] Audenaert K.M.R. (2013), "In-betweenness-a geometrical monotonicity property for operator means", *Linear Algebra Appl.,* 438 (4), 1769-1778.

[6] Audenaert K.M.R. (2017), "A singular value inequality for Heinz means", *Linear Algebra Appl.*, 442, 279-283.

[7] Audenaert K.M.R., Datta N. (2015), "α-z-Rényi relative entropies", *J. Math. Phys.,* 56.

[8] Audenaert K.M.R., Calsamiglia J., Masanes L., Munoz-Tapia R., Acin A., Bagan E., Verstraete D. (2007), "Discriminating States: The Quantum Chernoff Bound", *Phys. Rev. Lett.*, 98.

[9] Bhatia R. (1997), *Matrix analysis*, Graduate Texts in Mathematics, Springer-Verlag, New York.

[10] Bhatia R. (2007), *Positive definite matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NewYork.

[11] Bhatia R., Gaubert S., Jain T. (2019), “Matrix version of Hellinger distance”, *L. Math. Phys.,* 109, 1777-1804.

[12] Bhatia R., Holbrook J. (2006), “Riemannian geometry and matrix geometric means”, *Linear Algebra Appl.*, 413 (2), 594–618.

[13] Bhatia R., Jain T., Lim Y. (2019), “On the Bures-Wasserstein distance between positive definite matrices”, *Expo. Math.,* 37, 165–191.

[14] Bhatia R., Jain T., Lim Y. (2018), “Strong convexity of sandwiched entropies and related optimization problems”, *Rev. Math. Phys.,* 30.

[15] Bhatia R., Jain T., Lim Y. (1019), “Inequalities for the Wasserstein mean of positive definite matrices”, *Linear Algerba Appl.*, 576 (1), 108-123.

[16] Bhatia R., Lim Y., Yamazaki T. (2016), “Some norm inequalities for matrix means”, *Linear Algebra Appl.*, 501, 112-122.

[17] Bhagwat K.V., Subramanian R. (1978), “Inequalities between means of positive operators”, *Math. Proc. Camb. Phil. Soc.*, 83 (5), 393-401.

[18] Carlen, E. (2010), “Trace inequalities and quantum entropy: An introductory course”, *Entropy Quantum*, 529 (73), 73-140.

[19] Carlen E., Frank R.L., Lieb E.H. (2016), “Some operator and trace function convexity theorems”, *Linear Algebra Appl.*, 490, 174-185.

[20] Chehade S. (2020), “Saturating the Data Processing Inequality for Rényi relative entropy”, https://arxiv.org/pdf/2006.07726.pdf.

[21] Conway J.B. (1985), *A Course in Functional Analysis Graduate Texts in Mathematics*, 96, Springer, Berlin.

[22] Dinh T.H. (2015), "On characterization of operator monotone functions", *Linear Algebra Appl.*, 487, 260-267.

[23] Dinh T.H. (2017), "Some inequalities for the matrix Heron mean", *Linear Algebra Appl.*, 528, 321-330.

[24] Dinh T.H., Ahsani S., Tam T.Y. (2016), "Geometry and inequalities of geometric mean", *Czechoslovak Math. J.*, 66 (3), 777–792.

[25] Dinh T.H., Dumitru R., Franco J. (2022), "On the matrix Heron means and Rényi divergences", *Linear and Multilinear Alg.,* 70 (8), 1442-1450.

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

[27] Dinh T.H., Dumitru R., Franco J. (2019), "Non-Linear Interpolation of the Harmonic-Geometric-Arithmetic Matrix Means", *Loba. J. Math .*, 40, 101-105.

[28] Dinh T.H., Dumitru R., Franco J. (2020), "Some geometric properties of matrix means with respect to different metrics", *Positivity*, 24, 1419-1434.

[29] Dinh T.H., Du T.H.B., Nguyen D.A.N., Vuong T.D. (2023), "On new quantum divergences", *Linear and Multilinear Alg.*

[30] Dinh T.H., Le C.T, Vuong T.D. (2023), "α-z-fidelity and α-z-weighted right mean", *Submitted*.

[31] Dinh T.H., Le, B.K. Vo., Vuong T.D. (2021), "The α-z-Bures Wasserstein divergence", *Linear Algebra Appl.,* 624, 267-280.

[32] Dinh T.H., Le C.T., Vo B.K., Vuong T.D. (2021), "Weighted Hellinger distance and in betweenness property", *Math. Ine. Appls.,* 24, 157-165.

[33] Dinh T.H., Tam T.Y., Vuong T.D. (2023), "On new weighted spectral geometric mean", *Submitted*.

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

[35] Dinh T.H., Le A.V., Le C.T., Phan N.Y. (2022), "The matrix Heinz mean and related divergence", *Hacet. J. Math. Stat.*, 51, 362-372.

[36] Dumitru R., Franco F. (2022), "Geodesic in-betweenness for means of several matrices", *Linear Algebra Appl.*, 631, 77-92.

[37] Fiedler M., Ptak V. (1997), "A new positive definite geometric mean of two positive definite matrices", *Linear Algebra Appl.*, 251, 1-20.

[38] Franco J., Dumitru R. (2020), "The Rényi power means of matrices", *Linear Algebra Appl.*, 607, 45-57.

[39] Franco J., Dumitru R. (2020), "Generalized Hellinger metrics and Audenaert's in-betweenness", *Linear Algebra Appl.*, 585, 191-198.

[40] Gan L., Tam T.Y. (2022), "Inequalities and limits of weighted spectral geometric mean", *Linear and Multilinear Alg.,*.

[41] Gan L., Liu X., Tam T.Y. (2021), "On two geometric means and sum of adjoint orbits", *Linear Algebra Appl.*, 631,156-173.

[42] Gang L., Kim S. (2023), "Weak log-majorization between the geometric and Wasserstein means", *J. Math. Anal. Appl.*, 530 (2).

[43] Ha Q.M. (2019), "A Unified Formulation for the Bures-Wasserstein and Log-Euclidean/Log-Hilbert-Schmidt Distances between Positive Definite Operators", *Geometric Science of Information*, 475-483, Springer.

[44] Hiai F., Petz D. (1993), "The Golden-Thompson trace inequality is complemented", *Linear Algebra Appl.*, 181, 153-185.

[45] Hiai F., Petz D. (2014), *Introduction to Matrix Analysis and Applications*, Universitext, Springer Cham.

[46] Hwang J., Jeong M., Kim S. (2021), "Right Rényi mean and tensor product", *J. Appl. Math. Informatics.*, 39, 751-760.

[47] Hwang J., Kim S. (2018), "Bounds for the Wasserstein mean with applications to the Lie-Trotter mean", *J. Math. Anal. Appl.*, 475, 1744-1753.

[48] Hwang J., Kim K. (2017), "Lie-Trotter means of positive definite operators", *Linear Algebra Appl.*, 531, 268-280.

[49] Jeong M., Hwang J., Kim K. (2023), "Right mean for the α-z-Bures Wasserstein quantum divergence", *Act. Math, Sci,* 43, 2320-2332.

[50] Huang H., Wu Z., Zhu C., Fei S. (2021), "Quantifying the quantumness of ensembles via generalized relative Rényi entropy", *Int. J. Theor. Phys.*, 60, 2368-2379.

[51] Kim S. (2021), "Operator inequalities and gyrolines of the weighted geometric means", *Math. Ine. Appls.,* 2, 491-514.

[52] Kim S., Lee H. (2015), "Relative operator entropy related with the spectral geometric mean", *Anal.Math.Phys.*, 5, 233-240.

[53] Kim H., Lim Y. (2007), "An extended matrix exponential formula", *J. Math. Inequal.*, 1 (3), 443-447.

[54] Kubo F., Ando T. (1980), "Means of positive linear operators", *Math. Ann.*, 246 (3), 205-224.

[55] Kraus K. (1936), "Überkonvexe Matrixfunktionen", *Math. Ann.,* 246 (3), 205-224.

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

[57] Lee H., Lim Y. (2007), "Metric and spectral geometric means on sysmmetric cones", *Kyungpook Math. J.*, 47, 133-150.

[58] Li L., Molnar L., Wang L. (2021), "On preservers related to the spectral geometric mean", *Linear Algebra Appl.*, 610, 647-672.

[59] Lieb E., Ruskai M. (1973), "Proof of the strong subadditivity of quantum-mechanical entropy", *J. Math. Phys.*, 14, 1938-1941.

[60] Lim Y., Palfia P. (2012), "Matrix power means and the Karcher mean", *J. Funct. Anal.*, 262, 1498–1514.

[61] Lin C. (2000), "Bounds of operator functions and Furuta inequalities", *Publ. RIMS*, Kyoto Univ., 36, 483-490.

[62] Liao M., Liu X., Tam T.Y. (2014), "A geometric mean for symmetric spaces of noncompact type", *J. Lie Theory*, 24, 725-736.

[63] Loewner C. (1934), "Uber monotone matrix function", *Math. Z.*, 38, 177-216.

[64] Miszczak J.A., Puchala Z., Horodecki P., Uhlmanx A., Zyczkowski K. (2009), "Sub-and super-fidelity as bounds for quantum fidelity", *Quantum Information and Computation,* 12, 103-130.

[65] Moakher M. (2005), "A differential geometric approach to the geometric mean of symmetric positive-definite matrices", *SIAM J. Matrix Anal. Appl.*, 26, 735-747 .

[66] Müller M., Lennert. (2013), *Quantum Relative Rényi Entropies*, Master thesis, ETH Zurich..

[67] Nielsen F., Barbaresco F. (2019), *Geometric Science of Information*, Lecture Notes in Computer Science, Springer, Cham.

[68] Nielsen M.A., Chuang L.L. (2000), *Quantum Computation and Quantum Information,* Cambridge University Press, Cambridge.

[69] Pusz W., Woronowicz S.L. (1975), "Functional calculus for sesquilinear forms and the purification map", *Rep. Math. Phys.*, 5, 159-170.

[70] Rényi A. (1961), "On Measures of Information and Entropy", *In Proc. Symp. Math., Stat. Probab.*, Berkeley, University of California Press, 547-561.

[71] Scherer W. (2019), *Mathematics of quantum computing*, Springer.

[72] So W. (2004), "The high road to an exponential formula", *Linear Algebra Appl.*, 379, 69-75.

[73] Spehner S., Illuminati F., Orszag M., Roga W. (2017), "Geometric measures of quantum correlations with Bures and Hellinger distances", *Lectures on general quantum correlations and their applications*, 105-157.

[74] Spehner D., Orszag M. (2013), "Geometric quantum discord with Bures distance", *New journal of physics,* 15.

[75] Sra S. (2012), "A new metric on the manifold of kernel matrices with application to matrix geometric means", *In advances in Neural Information Processing Systems*, 1, 144-152.

[76] Trotter H. (1959), "On the product of semigroups of operators", *P.Am Math Soc.*, 10, 545-551.

[77] Vuong T.D, Vo B.K (2020), "An inequality for quantum fidelity", *Quy Nhon Univ. J. Sci*, 4 (3).

[78] Uhlmann A. (1976), "The transition probability in the state space of a $*$-algebra", *Rep. Math. Phys.*, 9, 273-279.

[79] Wilde W., Winter A., Yang D. (2014), "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy", *Commun. Math. Phys.*, 331, 593-622.

[80] Watrous J. (2008), *Theory of Quantum Information, Lecture notes from Fall 2008, Institute for Quantum Computing*, University of Waterloo, Canada.

[81] Wolf M.M. (2012), *Quantum channels & operations: guided tour, Lecture notes*, available at http://wwwm5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf.

[82] Yan X., Yin Z., Li L. (2020), "Quantum α-fidelity of unitary orbits", *Quantum infor. Pro.*, 307.

[83] Zhu X.X., Jin Z.X., Fei S.M. (2019), "Quantifying quantum coherence based on the generalized relative Rényi entrropy", *Quantum Infor. Pro.*, 18.

[84] Zhang, Bu K., Wu J. (2016), "A lower bound on the fidelity between two states in terms of their trace-distance and max-relative entropy", *Linear and Multilinear Alg.,* 64, 801-806.

[85] Zhang L., Fei S.M. (2014), "Quantum fidelity and relative entropy between unitary orbits", *J. Phys. A: Math. Theor.*, 47.

[86] Zhan X. (2002), *Matrix inequalities*, volume 1790 of Lecture Notes in Mathematics, Springer-Verlag, Berlin.

Index