Generalizing Lieb's Concavity Theorem via Operator Interpolation

De Huang[1]

Applied and Computational Mathematics, California Institute of Technology, Pasadena, CA 91125, USA

Abstract

We introduce the notion of k-trace and use interpolation of operators to prove the joint concavity of the function $(A,B)\mapsto \mathrm{Tr}_k\big[(B^{\frac{qs}{2}}K^*A^{ps}KB^{\frac{qs}{2}})^{\frac{1}{s}}\big]^{\frac{1}{k}}$, which generalizes Lieb's concavity theorem from trace to a class of homogeneous functions $\mathrm{Tr}_k[\cdot]^{\frac{1}{k}}$. Here $\mathrm{Tr}_k[A]$ denotes the k_{th} elementary symmetric polynomial of the eigenvalues of A. This result gives an alternative proof for the concavity of $A\mapsto \mathrm{Tr}_k\big[\exp(H+\log A)\big]^{\frac{1}{k}}$ that was obtained and used in a recent work to derive expectation estimates and tail bounds on partial spectral sums of random matrices.

Keywords: trace inequalities, concave/convex matrix functions, interpolation of operators.
2010 MSC: 47A57, 47A63, 15A42, 15A16, 15A75

1. Introduction

A fundamental result in the study of trace inequalities is the joint concavity of the function

$$(A,B)\ \longmapsto\ \mathrm{Tr}[K^*A^pKB^q] \tag{1}$$

on $\mathbf{H}_n^+\times\mathbf{H}_m^+$, for any $K\in\mathbb{C}^{n\times m}$, $p,q\in(0,1], p+q\le 1$, known as Lieb's Concavity Theorem [1]. Here $\mathbf{H}_n^+$ is the convex cone of all $n\times n$ Hermitian, positive semidefinite matrices. This theorem answered affirmatively to an important conjecture by Wigner, Yanase and Dyson [2] in information theory. It also led to Lieb's three-matrix extension of the Golden-Thompson inequality, which was then used by Lieb and Ruskai [3] to prove the strong subadditivity of quantum entropy.

In this paper, we generalize Lieb's concavity theorem from trace to a class of homogeneous matrix functions. In particular, we prove that the function

$$(A,B)\ \longmapsto\ \mathrm{Tr}_k\big[(B^{\frac{qs}{2}}K^*A^{ps}KB^{\frac{qs}{2}})^{\frac{1}{s}}\big]^{\frac{1}{k}} \tag{2}$$

is jointly concave on $\mathbf{H}_n^+\times\mathbf{H}_m^+$, for any $K\in\mathbb{C}^{n\times m}$, $s,p,q\in(0,1], p+q\le 1$. $\mathrm{Tr}_k[A]$ denotes the k_{th} elementary symmetric polynomial of the eigenvalues of A, specially $\mathrm{Tr}_1[A]=\mathrm{Tr}[A]$ and $\mathrm{Tr}_n[A]=\det[A]$ if $A\in\mathbb{C}^{n\times n}$. The map $A\mapsto\mathrm{Tr}_k[A]^{\frac{1}{k}}$ is homogeneous of order 1 and concave on $\mathbf{H}_n^+$. When $k=1, s=1$, function (2) reduces to function (1). The motivation of deriving our generalized Lieb's concavity theorem is to provide an alternative proof for the concavity of the map

$$A\ \longmapsto\ \mathrm{Tr}_k\big[\exp(H+\log A)\big]^{\frac{1}{k}} \tag{3}$$

on $\mathbf{H}_n^{++}$ (positive definite), for any Hermitian matrix H of the same size, due to a recent work by Huang [4], and hence completing the theory behind it. For $k=1$, the concavity of

$$A\ \longmapsto\ \mathrm{Tr}\big[\exp(H+\log A)\big], \tag{4}$$

[1]E-mail address: dhuang@caltech.edu.

also due to Lieb [1], is an equivalence of Lieb's concavity theorem. Tropp [5],[6] made use of the concavity of (4) to establish his master bounds on the largest (or smallest) eigenvalue of a sum of random matrices. Huang [4] introduced the notion of k-trace functions Tr_k to provide estimates on partial spectral sums of Hermitian matrices, and then used the concavity of (3) to generalize Tropp's master bounds from the largest (or smallest) eigenvalue to the sum of the the k largest (or smallest) eigenvalues. Huang's proof of the concavity of (3) in [4] was an imitation of Lieb's original arguments using matrix derivatives, which, however, failed to extend to the more complicated function (2). We then looked for a more profound approach to see a bigger picture.

Since Lieb's original establishment of his concavity theorem, alternative proofs have been developed from different aspects of matrix theories, including matrix tensors (Ando [7], Carlen [8], Nikoufar et al. [9]), the theory of Herglotz functions (Epstein [10]), and interpolation theories (Uhlmann [11], Kosaki [12]). The tensor approaches prove the theorem elegantly by translating the concavity of (1) to the operator concavity of the map $(A, B) \mapsto A^p \otimes B^q$, but have difficulties in generalizing to our k-trace case due to the nonlinearity of Tr_k. We, therefore, turned to the more generalizable methods of operator interpolation based essentially on Hölder's inequality that applies to k-trace as well. Originating from the Hadamard three-lines theorem [13], interpolation of operators has been a powerful tool in operator and functional analysis, with variant versions including the Riesz-Thorin interpolation theorem [14], Stein's interpolation of holomorphic operators [15], Peetre's K-method [16] and many others. In particular, we found Stein's complex interpolation technique most compatible and easiest to use in the k-trace setting. Our use of interpolation technique was inspired by a recent work of Sutter et al. [17], in which they applied Stein's interpolation to derive a multivariate extension of the Golden-Thompson inequality.

outline

The rest of the paper is organized as follows. Section 2 is devoted to introductions of general notations and the notion of k-trace. We will present in Section 3 our main theorems and their proofs, and briefly review previous proofs of Lieb's original concavity theorem. We also introduce the operator interpolation technique that we will use in Section 3. Two other results on k-trace will be given in Section 4. Some details of background knowledge are discussed in the appendices.

2. Notations and K-trace

2.1. General conventions

For any positive integers n, m, we write $\mathbb{C}^n$ for the n-dimensional complex vector spaces equipped with the standard l_2 inner products, and $\mathbb{C}^{n\times m}$ for the space of all complex matrices of size $n \times m$. Let $\mathbf{H}_n$ be the space of all $n \times n$ Hermitian matrices, $\mathbf{H}_n^+$ be the convex cone of all $n \times n$ Hermitian, positive semi-definite matrices, and $\mathbf{H}_n^{++}$ be the convex cone of all $n \times n$ Hermitian, positive definite matrices. We write $\mathbf{0}$ for square zero matrices of suitable size according to the context, and I_n for the identity matrix of size $n \times n$.

Following the notation in [4], we define the k-trace of a matrix $A \in \mathbb{C}^{n\times n}$ to be

$$\mathrm{Tr}_k[A] = \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}, \qquad 1 \leq k \leq n, \tag{5}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are all eigenvalues of A. In particular, $\mathrm{Tr}_1[A] = \mathrm{Tr}[A]$ is the normal trace of A, and $\mathrm{Tr}_n[A] = \det[A]$ is the determinant of A. If we write $A_{(i_1\cdots i_k, i_1\cdots i_k)}$ for the $k \times k$ principal submatrix of A corresponding to the indices $i_1, i_2, \cdots, i_k$, then an equivalent definition of the k-trace of A is given by

$$\mathrm{Tr}_k[A] = \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} \det[A_{(i_1\cdots i_k, i_1\cdots i_k)}], \qquad 1 \leq k \leq n. \tag{6}$$

Using the second definition (6), one can check that for any $1 \leq k \leq n$, the k-trace enjoys the cyclic invariance like the normal trace and the determinant. That is for any $A, B \in \mathbb{C}^n$, $\mathrm{Tr}_k[AB] = \mathrm{Tr}_k[BA]$.

For any function $f : \mathbb{R} \to \mathbb{R}$, the extension of f to a function from $\mathbf{H}_n$ to $\mathbf{H}_n$ is given by

$$f(A) = \sum_{i=1}^{n} f(\lambda_i) u_i u_i^*, \quad A \in \mathbf{H}_n,$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A, and $u_1, u_2, \cdots, u_n \in \mathbb{C}^n$ are the corresponding normalized eigenvectors. A function f is said to be operator monotone increasing (or decreasing) if $A \succeq B$ implies $f(A) \succeq f(B)$ (or $f(A) \preceq f(B)$); f is said to be operator convex (or concave) on some set S, if

$$\tau f(A) + (1-\tau) f(B) \succeq f(\tau A + (1-\tau)B) \text{ (or } \preceq f(\tau A + (1-\tau)B)),$$

for any $A, B \in S$ and any $\tau \in [0,1]$. For example, the function $A \mapsto A^r$ is both operator monotone increasing and operator concave on $\mathbf{H}_n^+$ for $r \in [0,1]$ (the Löwner-Heinz theorem [18], [19], [20], see also [8]). One can find more details and properties of matrix functions in [8, 21]. For any $A \in \mathbb{C}^{n\times m}$, we denote by $\|A\|_p$ the standard Schatten p-norm,

$$\|A\|_p = \mathrm{Tr}[|A|^p]^{\frac{1}{p}}, \tag{7}$$

where $|A| = (A^*A)^{\frac{1}{2}}$. In particular, we write $\|A\| = \|A\|_\infty =$ the largest singular value of A.

2.2. k-trace

Huang [4] introduced the notion of k-trace to provide bounds on the sum of the k largest (or smallest) eigenvalues of a matrix $A \in \mathbf{H}_n$,

$$\sum_{i=1}^{k} \lambda_i(A) \le \log \mathrm{Tr}_k\big[\exp(A)\big] \le \sum_{i=1}^{k} \lambda_i(A) + \log \binom{n}{k},$$
$$\sum_{i=1}^{k} \lambda_{n-i+1}(A) \ge -\log \mathrm{Tr}_k\big[\exp(-A)\big] \ge \sum_{i=1}^{k} \lambda_{n-i+1}(A) - \log \binom{n}{k},$$

where $\lambda_i(A)$ denotes the i_{th} largest eigenvalue of A. Huang then used these estimates and the concavity of $A \mapsto \mathrm{Tr}_k[\exp(H + \log A)]^{\frac{1}{k}}$ to derive expectation estimates and tail bounds for the sum of the k largest (or smallest) eigenvalues of a class of random matrices. Apart from this particular application, the k-trace is of theoretical interest by itself, as it has many interpretations corresponding to different aspects of matrix theories. Writing $D(A^{(1)}, A^{(2)}, \cdots, A^{(n)})$ the mixed discriminant of any n matrices $A^{(1)}, A^{(2)}, \cdots, A^{(n)} \in \mathbb{C}^{n\times n}$, we then have the identity

$$\mathrm{Tr}_k[A] = \binom{n}{k} D(\underbrace{A, \cdots, A}_{k}, \underbrace{I_n, \cdots, I_n}_{n-k}),$$

which as well connects the k-trace to the notion of the k_{th} intrinsic volume in convex geometry. Also, if we consider the k_{th} exterior algebra $\wedge^k(\mathbb{C}^n)$, we can then interpret the k-trace of A as

$$\mathrm{Tr}_k[A] = \mathrm{Tr}_{\mathcal{L}(\wedge^k(\mathbb{C}^n))}\big[\mathcal{M}_0^{(k)}(A)\big],$$

where $\mathrm{Tr}_{\mathcal{L}(\wedge^k(\mathbb{C}^n))}$ is the normal trace on the operator space $\mathcal{L}(\wedge^k(\mathbb{C}^n))$, and $\mathcal{M}_0^{(k)}(A) \in \mathcal{L}(\wedge^k(\mathbb{C}^n))$ is defined as $\mathcal{M}_0^{(k)}(A)(v_1 \wedge v_2 \wedge \cdots \wedge v_k) = Av_1 \wedge Av_2 \wedge \cdots \wedge Av_k$, for any $v_1 \wedge v_2 \wedge \cdots \wedge v_k \in \wedge^k(\mathbb{C}^n)$. More discussions on these two interpretations and how they can be used to study the k-trace will be presented in Appendix A and Appendix B.

Throughout the paper, we will be using the following properties of the k-trace.

Proposition 2.1. *For any positive integers n, k, $1 \le k \le n$, the k-trace function $\mathrm{Tr}_k[\cdot]$ satisfies the following:*

(i) Cyclicity: $\mathrm{Tr}_k[AB] = \mathrm{Tr}_k[BA]$, $A, B \in \mathbb{C}^{n\times n}$.

(ii) Homogeneity: $\mathrm{Tr}_k[\alpha A] = \alpha^k \mathrm{Tr}_k[A]$, $A \in \mathbb{C}^{n\times n}, \alpha \in \mathbb{C}$.

(iii) Monotonicity: For any $A, B \in \mathbf{H}_n^+$, $\mathrm{Tr}_k[A] \geq \mathrm{Tr}_k[B]$, *if* $A \succeq B$; $\mathrm{Tr}_k[A] > \mathrm{Tr}_k[B]$, *if* $A \succ B$. *In particular,* $\mathrm{Tr}_k[A] \geq 0, A \in \mathbf{H}_n^+$.

(iv) Concavity: The function $A \mapsto (\mathrm{Tr}_k[A])^{\frac{1}{k}}$ *is concave on* $\mathbf{H}_n^+$.

(v) Hölder's Inequality: $\mathrm{Tr}_k[AB] \leq \mathrm{Tr}_k[|A|^p]^{\frac{1}{p}} \mathrm{Tr}_k[|B|^q]^{\frac{1}{q}}$, *for any* $p, q \in [0, +\infty], \frac{1}{p} + \frac{1}{q} = 1$, *and any* $A, B \in \mathbb{C}^{n\times n}$.

(vi) Consistency: For any $\tilde{n}$, $k \leq \tilde{n} \leq n$, *and any* $A \in \mathbb{C}^{\tilde{n}\times\tilde{n}}$,

$$\mathrm{Tr}_k\left[\begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}_{n\times n}\right] = \mathrm{Tr}_k[A].$$

Proof. (i), (ii), (iii) and (vi) can be easily verified by definitions (5) and (6). (iv) is a result of the general Brunn-Minkowski theorem (Corollary A.3) in Appendix A. (v) is a direct result of expression (B.5) in Appendix B. In fact, since the normal trace enjoys the Hölder's inequality, we have

$$\mathrm{Tr}_k[AB] = \mathrm{Tr}[\mathcal{M}_0^{(k)}(A)\mathcal{M}_0^{(k)}(B)] \leq \mathrm{Tr}[|\mathcal{M}_0^{(k)}(A)|^p]^{\frac{1}{p}} \mathrm{Tr}[|\mathcal{M}_0^{(k)}(B)|^q]^{\frac{1}{q}} = \mathrm{Tr}_k[|A|^p]^{\frac{1}{p}} \mathrm{Tr}_k[|B|^q]^{\frac{1}{q}}.$$

We have used multiple properties of the operator $\mathcal{M}_0^{(k)}(A)$ introduced in Appendix B. □

3. Generalizing Lieb's Conacvity Theorem

3.1. Main Theorems

In what follows, we will fix the integer k and always write

$$\phi(A) = (\mathrm{Tr}_k[A])^{\frac{1}{k}}$$

for simplicity. Note that the function ϕ also satisfies (i) cyclicity, (iii) monotonicity, (v) Hölder's inequality and (vi) consistency as in Proposition 2.1. But now the map $A \mapsto \phi(A)$ is homogeneous of order 1 and is concave on $\mathbf{H}_n^+$. Abusing notation, we will also refer the function ϕ as the k-trace. Our main results of this paper are the following.

Lemma 3.1. *For any* $s, r \in (0, 1]$ *and any* $K \in \mathbb{C}^{n\times n}$, *the function*

$$A \longmapsto \phi\big((K^* A^{rs} K)^{\frac{1}{s}}\big) \tag{8}$$

is concave on $\mathbf{H}_n^+$.

Theorem 3.2 (Generalized Lieb's Concavity Theorem)**.** *For any* $s, p, q \in (0, 1], p + q \leq 1$, *and any* $K \in \mathbb{C}^{n\times m}$, *the function*

$$(A, B) \longmapsto \phi\big((B^{\frac{qs}{2}} K^* A^{ps} K B^{\frac{qs}{2}})^{\frac{1}{s}}\big) \tag{9}$$

is jointly concave on $\mathbf{H}_n^+ \times \mathbf{H}_m^+$.

Theorem 3.3. *For any* $H \in \mathbf{H}_n$ *and any* $\{p_j\}_{j=1}^m \subset (0, 1]$ *such that* $\sum_{j=1}^m p_j \leq 1$, *the function*

$$(A^{(1)}, A^{(2)}, \ldots, A^{(m)}) \longmapsto \phi\Big(\exp\big(H + \sum_{j=1}^m p_j \log A^{(j)}\big)\Big) \tag{10}$$

is jointly concave on $(\mathbf{H}_n^{++})^{\times m}$. *In particular,* $A \mapsto \phi\big(\exp(H + \log A)\big)$ *is concave on* $\mathbf{H}_n^{++}$.

Lemma 3.1 is a k-trace extension of the concave part of Lemma 2.8 in [22]. The latter is a direct consequence of the original Lieb's concavity theorem. However, we will first apply the technique of operator interpolation to prove Lemma 3.1 independently, and then use it to derive the other results. Theorem 3.2 is our generalized Lieb's concavity theorem, which not only extends the original Lieb's concavity to a k-trace version, but also strengthens its form by adding a power s. We will perform operator interpolation with respect to p to derive Theorem 3.2 from Lemma 3.1, hence inheriting the power s.

Theorem 3.3 is a generalization of both Corollary 6.1 in [1] (from trace to k-trace) and Theorem 2.1 in [4] (from univariate to multivariate). Lieb [1] proved the original trace version by checking the non-positiveness of the second order directional derivatives (or Hessians). Huang [4] imitated Lieb's derivative arguments and proved the concavity of $A \mapsto \phi\big(\exp(H + \log A)\big)$, which he then used to derive concentrations of partial spectral sums of random matrices. We will first prove Theorem 3.3 for $m = 1$ by applying the Lie product formula to Lemma 3.1 (taking $s \to 0$), and hence providing an alternative proof of the concavity of $A \mapsto \phi\big(\exp(H + \log A)\big)$. We then improve the result from $m = 1$ to $m \geq 1$ using a k-trace version of the Araki-Lieb-Thirring inequality (Lemma 3.6). All proofs of our main results will be presented in Section 3.3.

Apart from the above theorems, we will also prove (i) a k-trace version of the multivariate extension of the Golden-Thompson inequality, and (ii) the monotonicity preserving and concavity preserving properties of k-trace. Section 4 will be devoted to these two results.

3.2. Operator Interpolation

Our main tool is Stein's interpolation of linear operators [15], that was developed from Hirschman's stronger version of the Hadamard three-line theorem [23]. This technique was recently adopted by Sutter et al. [17] to establish a multivariate extension of the Golden-Thompson inequality, which inspired our use of interpolation in proving the generalized Lieb's concavity theorem. We will follow the notations in [17]. For any $\theta \in (0, 1)$, we define a density $\beta_\theta(t)$ on $\mathbb{R}$ by

$$\beta_\theta(t) = \frac{\sin(\pi\theta)}{2\theta\big(\cosh(\pi t) + \cos(\pi\theta)\big)}, \quad t \in \mathbb{R}. \tag{11}$$

Specially, we define

$$\beta_0(t) = \lim_{\theta \searrow 0} \beta_\theta(t) = \frac{\pi}{2(\cosh(\pi t) + 1)}, \quad \text{and} \quad \beta_1(t) = \lim_{\theta \nearrow 1} \beta_\theta(t) = \delta(t).$$

$\beta_\theta(t)$ is a density since $\beta_\theta(t) \geq 0, t \in \mathbb{R}$ and $\int_{-\infty}^{+\infty} \beta_\theta(t)dt = 1$. We will always use $\mathcal{S}$ to denote a vertical strip on the complex plane $\mathbb{C}$:

$$\mathcal{S} = \{z \in \mathbb{C} : 0 \leq \mathrm{Re}(z) \leq 1\}. \tag{12}$$

Theorem 3.4 (Stein-Hirschman)**.** *Let $G(z)$ be a map from $\mathcal{S}$ to bounded linear operators on a separable Hilbert space that that is holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary. Let $p_0, p_1 \in [1, +\infty], \theta \in [0, 1]$, and define p_θ by*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Then if $\|G(z)\|_{p_{\mathrm{Re}(z)}}$ is uniformly bounded on $\mathcal{S}$, the following inequality holds:

$$\log \|G(\theta)\|_{p_\theta} \leq \int_{-\infty}^{+\infty} dt \Big(\beta_{1-\theta}(t) \log \|G(it)\|_{p_0}^{1-\theta} + \beta_\theta(t) \log \|G(1+it)\|_{p_1}^{\theta}\Big). \tag{13}$$

A k-trace analog of the above theorem, that is more convenient for our use, is as follows,

Lemma 3.5. *Let $G(z) : \mathcal{S} \to \mathbb{C}^{n\times n}$ be holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary. Let $p_0, p_1 \in [1, +\infty]$, $\theta \in [0, 1]$, and define p_θ by*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Then if $\|G(z)\|$ is uniformly bounded on $\mathcal{S}$, the following inequality holds:

$$\begin{aligned} &\log\left[\phi(|G(\theta)|^{p_\theta})^{\frac{1}{p_\theta}}\right] \\ &\leq \int_{-\infty}^{+\infty} dt\Big(\beta_{1-\theta}(t)\log\left[\phi\big(|G(it)|^{p_0}\big)^{\frac{1-\theta}{p_0}}\right] + \beta_\theta(t)\log\left[\phi\big(|G(1+it)|^{p_1}\big)^{\frac{\theta}{p_1}}\right]\Big). \end{aligned} \tag{14}$$

More discussions on operator interpolation and the proof of Lemma 3.5 will be given in Appendix C. For $p_0, p_1 \in [1, +\infty)$, we can rewrite inequality (14) as

$$\begin{aligned} &\log\phi(|G(\theta)|^{p_\theta}) \\ &\leq \int_{-\infty}^{+\infty} dt\Big(\frac{(1-\theta)p_\theta}{p_0}\beta_{1-\theta}(t)\log\phi\big(|G(it)|^{p_0}\big) + \frac{\theta p_\theta}{p_1}\beta_\theta(t)\log\phi\big(|G(1+it)|^{p_1}\big)\Big) \end{aligned} \tag{15}$$

Notice that

$$\int_{-\infty}^{+\infty}\Big(\frac{(1-\theta)p_\theta}{p_0}\beta_{1-\theta} + \frac{\theta p_\theta}{p_1}\beta_\theta(t)\Big)dt = 1.$$

Then using Jensen's inequality on the concavity of logarithm, we can immediately conclude from (15) that for $p_0, p_1 \in [1, +\infty)$,

$$\phi(|G(\theta)|^{p_\theta}) \leq \int_{-\infty}^{+\infty} dt\Big(\frac{(1-\theta)p_\theta}{p_0}\beta_{1-\theta}(t)\phi\big(|G(it)|^{p_0}\big) + \frac{\theta p_\theta}{p_1}\beta_\theta(t)\phi\big(|G(1+it)|^{p_1}\big)\Big), \tag{16}$$

under the same setting as in Lemma 3.5.

3.3. Proof of Main Results

The key of applying Lemma 3.5 is to choose some proper holomorphic function $G(z)$ and then interpolating on some power in $[0, 1]$. In particular, we will interpolation on s to prove Lemma 3.1, and then on p to prove Theorem 3.2. Our choices of the holomorphic functions $G(z)$ in the following proofs are inspired by Lieb's constructions in [1] for the use of maximum modulus principle.

Proof of Lemma 3.1. We need to show that, for any $A, B \in \mathbf{H}_n^+$ and any $\tau \in [0, 1]$,

$$\tau\phi\big((K^*A^{rs}K)^{\frac{1}{s}}\big) + (1-\tau)\phi\big((K^*B^{rs}K)^{\frac{1}{s}}\big) \leq \phi\big((K^*C^{rs}K)^{\frac{1}{s}}\big),$$

where $C = \tau A + (1-\tau)B$. We may assume that $A, B \in \mathbf{H}_n^{++}$ and K is invertible. Once this is done, the general result for $A, B \in \mathbf{H}_n^+$ and $K \in \mathbb{C}^{n\times n}$ can be obtained by continuity. Let $M = C^{\frac{rs}{2}}K$, and let $M = Q|M|$ be the polar decomposition of M for some unitary matrix Q. Since C, K are both invertible, $|M| \in \mathbf{H}_n^{++}$. We then define two functions from $\mathcal{S}$ to $\mathbb{C}^{n\times n}$:

$$G_A(z) = A^{\frac{rz}{2}}C^{-\frac{rz}{2}}Q|M|^{\frac{z}{s}}, \quad G_B(z) = B^{\frac{rz}{2}}C^{-\frac{rz}{2}}Q|M|^{\frac{z}{s}}, \quad z \in \mathcal{S},$$

where $\mathcal{S}$ is given by (12). In what follows we will use X for A or B. We then have

$$\begin{aligned} \phi\big((K^*X^{rs}K)^{\frac{1}{s}}\big) &= \phi\big((M^*C^{-\frac{rs}{2}}X^{rs}C^{-\frac{rs}{2}}M)^{\frac{1}{s}}\big) \\ &= \phi\big((|M|Q^*C^{-\frac{rs}{2}}X^{\frac{rs}{2}}X^{\frac{rs}{2}}C^{-\frac{rs}{2}}Q|M|)^{\frac{1}{s}}\big) \\ &= \phi\big(|G_X(s)|^{\frac{2}{s}}\big). \end{aligned}$$

Since A, B, C, M are now fixed matrices in $\mathbf{H}_n^{++}$, $G_A(z)$ and $G_B(z)$ are apparently holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary. Also, it is easy to check that $\|G_A(z)\|$ and $\|G_B(z)\|$ are uniformly bounded on $\mathcal{S}$, since $\mathrm{Re}(z) \in [0, 1]$. Therefore we can use inequality (16) with $\theta = s, p_\theta = \frac{2}{s}$ to obtain

$$\phi(|G_X(s)|^{\frac{2}{s}}) \leq \int_{-\infty}^{+\infty} dt\Big(\frac{2(1-s)}{sp_0}\beta_{1-s}(t)\phi\big(|G_X(it)|^{p_0}\big) + \frac{2}{p_1}\beta_s(t)\phi\big(|G_X(1+it)|^{p_1}\big)\Big).$$

We still need to choose some $p_0, p_1 \geq 1$ satisfying $\frac{1-s}{p_0} + \frac{s}{p_1} = \frac{1}{p_s} = \frac{s}{2}$ to proceed. Note that $G_X(it) = X^{\frac{irt}{2}} C^{-\frac{irt}{2}} Q|M|^{\frac{it}{s}}$ are now unitary matrices for all $t \in \mathbb{R}$ since $X, C, |M| \in \mathbf{H}_n^{++}$, and thus $|G_X(it)|^{p_0} = I_n$ for all p_0. Therefore we can take $p_0 \to +\infty, p_1 = 2$ to obtain

$$\phi(|G_X(s)|^{\frac{2}{s}}) \leq \int_{-\infty}^{+\infty} dt \beta_s(t) \phi\big(|G_X(1+it)|^2\big).$$

Further, for each $t \in \mathbb{R}$, we have

$$\begin{aligned}
&\phi\big(|G_X(1+it)|^2\big) \\
&= \phi\big(G_X(1+it)^* G_X(1+it)\big) \\
&= \phi\big(|M|^{\frac{(1-it)}{s}} Q^* C^{-\frac{r(1-it)}{2}} X^r C^{-\frac{r(1+it)}{2}} Q|M|^{\frac{(1+it)}{s}}\big) \\
&= \phi\big(|M|^{\frac{1}{s}} Q^* C^{-\frac{r(1-it)}{2}} X^r C^{-\frac{r(1+it)}{2}} Q|M|^{\frac{1}{s}}\big)
\end{aligned}$$

where we have used the cyclicity of ϕ. Therefore we have

$$\begin{aligned}
&\tau\phi\big(|G_A(1+it)|^2\big) + (1-\tau)\phi\big(|G_B(1+it)|^2\big) \\
&= \tau\phi\big(|M|^{\frac{1}{s}} Q^* C^{-\frac{r(1-it)}{2}} A^r C^{-\frac{r(1+it)}{2}} Q|M|^{\frac{1}{s}}\big) \\
&\quad + (1-\tau)\phi\big(|M|^{\frac{1}{s}} Q^* C^{-\frac{r(1-it)}{2}} B^r C^{-\frac{r(1+it)}{2}} Q|M|^{\frac{1}{s}}\big) \\
&\leq \phi\big(|M|^{\frac{1}{s}} Q^* C^{-\frac{r(1-it)}{2}} (\tau A^r + (1-\tau)B^r) C^{-\frac{r(1+it)}{2}} Q|M|^{\frac{1}{s}}\big) \\
&\leq \phi\big(|M|^{\frac{1}{s}} Q^* C^{-\frac{r(1-it)}{2}} C^r C^{-\frac{r(1+it)}{2}} Q|M|^{\frac{1}{s}}\big) \\
&= \phi\big(|M|^{\frac{2}{s}}\big) \\
&= \phi\big((M^*M)^{\frac{1}{s}}\big).
\end{aligned}$$

The first inequality above is due to the concavity of ϕ, the second inequality is due to (i) that ϕ is monotone increasing on $\mathbf{H}_n^+$ and (ii) that $X \mapsto X^r$ is operator concave on $\mathbf{H}_n^+$ for $r \in (0,1]$. Finally, since $\phi\big((M^*M)^{\frac{1}{s}}\big)$ is independent of t, and $\beta_s(t)$ is a density on $\mathbb{R}$, we obtain that

$$\begin{aligned}
&\tau\phi\big((K^*A^{rs}K)^{\frac{1}{s}}\big) + (1-\tau)\phi\big((K^*B^{rs}K)^{\frac{1}{s}}\big) \\
&= \tau\phi\big(|G_A(s)|^{\frac{2}{s}}\big) + (1-\tau)\phi\big(|G_B(s)|^{\frac{2}{s}}\big) \\
&\leq \phi\big((M^*M)^{\frac{1}{s}}\big) \\
&= \phi\big((K^*C^{rs}K)^{\frac{1}{s}}\big).
\end{aligned}$$

So we have proved the concavity of (8) on $\mathbf{H}_n^+$. □

Proof of Theorem 3.2. We first claim that, to prove the joint concavity of (9), we only need to prove the concavity of

$$A \longmapsto \phi\big((A^{\frac{qs}{2}} K^* A^{ps} K A^{\frac{qs}{2}})^{\frac{1}{s}}\big) \tag{17}$$

on $\mathbf{H}_n^+$ for any $K \in \mathbb{C}^{n\times n}$. In fact, if we define

$$\widehat{A} = \begin{pmatrix} B & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} \in \mathcal{H}_{m+n}^+, \quad \widehat{K} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ K & \mathbf{0} \end{pmatrix} \in \mathbb{C}^{(m+n)\times(m+n)},$$

we will have

$$(\widehat{A}^{\frac{qs}{2}} \widehat{K}^* \widehat{A}^{ps} \widehat{K} \widehat{A}^{\frac{qs}{2}})^{\frac{1}{s}} = \begin{pmatrix} (B^{\frac{qs}{2}} K^* A^{ps} K B^{\frac{qs}{2}})^{\frac{1}{s}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

and thus by consistency of ϕ,

$$\phi\big((\widehat{A}^{\frac{qs}{2}} \widehat{K}^* \widehat{A}^{ps} \widehat{K} \widehat{A}^{\frac{qs}{2}})^{\frac{1}{s}}\big) = \phi\big((B^{\frac{qs}{2}} K^* A^{ps} K B^{\frac{qs}{2}})^{\frac{1}{s}}\big).$$

Then the concavity of $\widehat{A} \mapsto \phi\big((\widehat{A}^{\frac{qs}{2}}\widehat{K}^*\widehat{A}^{ps}\widehat{K}\widehat{A}^{\frac{qs}{2}})^{\frac{1}{s}}\big)$ immediately implies the joint concavity of $(A,B) \mapsto \phi\big((B^{\frac{qs}{2}}K^*A^{ps}KB^{\frac{qs}{2}})^{\frac{1}{s}}\big)$. So the claim is justified. Now to prove the concavity of (17), we need to show that, for any $A,B \in \mathbf{H}_n^+$ and any $\tau \in [0,1]$,

$$\tau\phi\big((A^{\frac{qs}{2}}K^*A^{ps}KA^{\frac{qs}{2}})^{\frac{1}{s}}\big) + (1-\tau)\phi\big((B^{\frac{qs}{2}}K^*B^{ps}KB^{\frac{qs}{2}})^{\frac{1}{s}}\big) \le \phi\big((C^{\frac{qs}{2}}K^*C^{ps}KC^{\frac{qs}{2}})^{\frac{1}{s}}\big),$$

where $C = \tau A + (1-\tau)B$. Again, we may assume that $A,B \in \mathbf{H}_n^{++}$. Once this is done, the general result for $A,B \in \mathbf{H}_n^+$ can be obtained by continuity. Let $M = C^{\frac{ps}{2}}KC^{\frac{qs}{2}}$. We then define two functions from $\mathcal{S}$ to $\mathbb{C}^{n\times n}$:

$$G_A(z) = A^{\frac{rsz}{2}}C^{-\frac{rsz}{2}}MC^{-\frac{rs(1-z)}{2}}A^{\frac{rs(1-z)}{2}}, \quad z \in \mathcal{S},$$
$$G_B(z) = B^{\frac{rsz}{2}}C^{-\frac{rsz}{2}}MC^{-\frac{rs(1-z)}{2}}B^{\frac{rs(1-z)}{2}}, \quad z \in \mathcal{S},$$

where $\mathcal{S}$ is given by (12), and $r = p+q \in (0,1]$. In what follows we may use X for A or B. We then have

$$\begin{aligned}\phi\big((X^{\frac{qs}{2}}K^*X^{ps}KX^{\frac{qs}{2}})^{\frac{1}{s}}\big) &= \phi\big((X^{\frac{qs}{2}}C^{-\frac{qs}{2}}M^*C^{-\frac{ps}{2}}X^{\frac{ps}{2}}X^{\frac{ps}{2}}C^{-\frac{ps}{2}}MC^{-\frac{qs}{2}}X^{\frac{qs}{2}})^{\frac{1}{s}}\big)\\ &= \phi\big(|G_X(\frac{p}{r})|^{\frac{2}{s}}\big).\end{aligned}$$

Since A,B,C,M are now fixed matrices in $\mathbf{H}_n^{++}$, $G_A(z)$ and $G_B(z)$ are apparently holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary. Also, it is easy to check that $\|G_A(z)\|$ and $\|G_B(z)\|$ are uniformly bounded on $\mathcal{S}$, since $\mathrm{Re}(z) \in [0,1]$. Therefore we can use inequality (16) with $\theta = \frac{p}{r}, p_0 = p_1 = p_\theta = \frac{2}{s}$ to obtain

$$\phi(|G_X(\frac{p}{r})|^{\frac{2}{s}}) \le \int_{-\infty}^{+\infty} dt\Big(\frac{q}{r}\beta_{1-\frac{p}{r}}(t)\phi\big(|G_X(it)|^{\frac{2}{s}}\big) + \frac{p}{r}\beta_{\frac{p}{r}}(t)\phi\big(|G_X(1+it)|^{\frac{2}{s}}\big)\Big).$$

Further, for each $t \in \mathbb{R}$, we have

$$\begin{aligned}&\phi\big(|G_X(1+it)|^{\frac{2}{s}}\big)\\ &= \phi\Big(\big(G_X(1+it)^*G_X(1+it)\big)^{\frac{1}{s}}\Big)\\ &= \phi\Big(\big(X^{\frac{irst}{2}}C^{-\frac{irst}{2}}M^*C^{-\frac{rs(1-it)}{2}}X^{rs}C^{-\frac{rs(1+it)}{2}}MC^{\frac{irst}{2}}X^{-\frac{irst}{2}}\big)^{\frac{1}{s}}\Big)\\ &= \phi\Big(\big(M^*C^{-\frac{rs(1-it)}{2}}X^{rs}C^{-\frac{rs(1+it)}{2}}M\big)^{\frac{1}{s}}\Big),\end{aligned}$$

where we have used (i) that X^{it} is unitary for any $X \in \mathbf{H}_n^{++}, t \in \mathbb{R}$, (ii) that $f(U^*XU) = U^*f(X)U$ for any $X \in \mathbf{H}_n$, any unitary $U \in \mathbb{C}^{n\times n}$ and any function f, and (iii) the cyclicity of ϕ. Now since $r,s \in (0,1]$, we can use Lemma 3.1 to obtain

$$\begin{aligned}&\tau\phi\big(|G_A(1+it)|^{\frac{2}{s}}\big) + (1-\tau)\phi\big(|G_B(1+it)|^{\frac{2}{s}}\big)\\ &= \tau\phi\Big(\big(M^*C^{-\frac{rs(1-it)}{2}}A^{rs}C^{-\frac{rs(1+it)}{2}}M\big)^{\frac{1}{s}}\Big)\\ &\quad + (1-\tau)\phi\Big(\big(M^*C^{-\frac{rs(1-it)}{2}}B^{rs}C^{-\frac{rs(1+it)}{2}}M\big)^{\frac{1}{s}}\Big)\\ &\le \phi\Big(\big(M^*C^{-\frac{rs(1-it)}{2}}(\tau A+(1-\tau)B)^{rs}C^{-\frac{rs(1+it)}{2}}M\big)^{\frac{1}{s}}\Big)\\ &= \phi\Big(\big(M^*C^{-\frac{rs(1-it)}{2}}C^{rs}C^{-\frac{rs(1+it)}{2}}M\big)^{\frac{1}{s}}\Big)\\ &= \phi\big((M^*M)^{\frac{1}{s}}\big).\end{aligned}$$

Similarly, we have that for each $t \in \mathbb{R}$,

$$\begin{aligned}\phi\big(|G_X(it)|^{\frac{2}{s}}\big) &= \phi\Big(\big(X^{\frac{rs(1+it)}{2}}C^{-\frac{rs(1+it)}{2}}M^*MC^{-\frac{rs(1-it)}{2}}X^{\frac{rs(1-it)}{2}}\big)^{\frac{1}{s}}\Big)\\ &= \phi\Big(\big(MC^{-\frac{rs(1-it)}{2}}X^{rs}C^{-\frac{rs(1+it)}{2}}M^*\big)^{\frac{1}{s}}\Big).\end{aligned}$$

We have used the fact that $\phi\big(f(X^*X)\big) = \phi\big(f(XX^*)\big)$ for any $X \in \mathbb{C}^{n\times n}$ and any function f, since ϕ is only a function of eigenvalues and the spectrums of $f(X^*X)$ and $f(XX^*)$ are the same. Then again using Lemma 3.1 we obtain that

$$\tau\phi\big(|G_A(it)|^{\frac{2}{s}}\big) + (1-\tau)\phi\big(|G_B(it)|^{\frac{2}{s}}\big) \leq \phi\big((MM^*)^{\frac{1}{s}}\big) = \phi\big((M^*M)^{\frac{1}{s}}\big).$$

Finally we have

$$\begin{aligned}
&\tau\phi\big((A^{\frac{qs}{2}}K^*A^{ps}KA^{\frac{qs}{2}})^{\frac{1}{s}}\big) + (1-\tau)\phi\big((B^{\frac{qs}{2}}K^*B^{ps}KB^{\frac{qs}{2}})^{\frac{1}{s}}\big)\\
&= \tau\phi\big(|G_A(\frac{p}{r})|^{\frac{2}{s}}\big) + (1-\tau)\phi\big(|G_B(\frac{p}{r})|^{\frac{2}{s}}\big)\\
&\leq \int_{-\infty}^{+\infty} dt\Big\{\frac{q}{r}\beta_{1-\frac{p}{r}}(t)\Big(\tau\phi\big(|G_A(it)|^{\frac{2}{s}}\big) + (1-\tau)\phi\big(|G_B(it)|^{\frac{2}{s}}\big)\Big)\\
&\qquad + \frac{p}{r}\beta_{\frac{p}{r}}(t)\Big(\tau\phi\big(|G_A(1+it)|^{\frac{2}{s}}\big) + (1-\tau)\phi\big(|G_B(1+it)|^{\frac{2}{s}}\big)\Big)\Big\}\\
&\leq \phi\big((M^*M)^{\frac{1}{s}}\big)\int_{-\infty}^{+\infty}\Big(\frac{q}{r}\beta_{1-\frac{p}{r}}(t) + \frac{p}{r}\beta_{\frac{p}{r}}(t)\Big)dt\\
&= \phi\big((C^{\frac{qs}{2}}K^*C^{ps}KC^{\frac{qs}{2}})^{\frac{1}{s}}\big).
\end{aligned}$$

So we have proved the concavity of (17), and thus the joint concavity of (9). □

Proof of Theorem 3.3 (Part I). We first prove the theorem for $m = 1$. Let $r = p_1 \in (0, 1]$, and $K^{(N)} = (K^{(N)})^* = \exp\big(\frac{1}{2N}H\big), N \geq 1$. Then using the Lie product formula

$$\lim_{N\to+\infty}\Big(\exp\big(\frac{1}{2N}Y\big)\exp\big(\frac{1}{N}X\big)\exp\big(\frac{1}{2N}Y\big)\Big)^N = \exp(X+Y), \quad X, Y \in \mathbf{H}_n,$$

we have

$$\begin{aligned}
\lim_{N\to+\infty}\phi\Big(\big((K^{(N)})^*A^{\frac{r}{N}}K^{(N)}\big)^N\Big) &= \lim_{N\to+\infty}\phi\left(\Big(\exp\big(\frac{1}{2N}H\big)\exp\big(\frac{r}{N}\log A\big)\exp\big(\frac{1}{2N}H\big)\Big)^N\right)\\
&= \phi\big(\exp(H + r\log A)\big).
\end{aligned}$$

By Theorem 3.2, for each $N \geq 1$, $\phi\Big(\big((K^{(N)})^*A^{\frac{r}{N}}K^{(N)}\big)^N\Big)$ is concave in A, thus the limit function $\phi\big(\exp(H + r\log A)\big)$ is also concave in A. □

To go from $m = 1$ to $m > 1$ in Theorem 3.3, we need to use the convexity of the map $A \mapsto \phi(\exp(A))$, which we will prove via the following lemmas. They are the k-trace extensions of the Araki-Lieb-Thirring inequality [24], the Golden-Thompson inequality and a variant of the Peierls-Bogoliubov inequality (see Theorem 2.12 in [8]).

Lemma 3.6 (k-trace Araki-Lieb-Thirring Inequality)**.** *For any $A, B \in \mathbf{H}_n^+$, the function*

$$t \mapsto \mathrm{Tr}_k\big[(B^{\frac{t}{2}}A^tB^{\frac{t}{2}})^{\frac{1}{t}}\big]$$

is monotone increasing on $(0, +\infty)$, that is

$$\mathrm{Tr}_k\big[(B^{\frac{t}{2}}A^tB^{\frac{t}{2}})^{\frac{1}{t}}\big] \leq \mathrm{Tr}_k\big[(B^{\frac{s}{2}}A^sB^{\frac{s}{2}})^{\frac{1}{s}}\big], \quad 0 < t \leq s. \tag{18}$$

Proof. Using the definition and properties of the operator $\mathcal{M}_0^{(k)}$ in Appendix B, we have that

$$\begin{aligned}
\mathrm{Tr}_k\big[(B^{\frac{t}{2}}A^tB^{\frac{t}{2}})^{\frac{1}{t}}\big] &= \mathrm{Tr}\big[\mathcal{M}_0^{(k)}\big((B^{\frac{t}{2}}A^tB^{\frac{t}{2}})^{\frac{1}{t}}\big)\big]\\
&= \mathrm{Tr}\big[\big((\mathcal{M}_0^{(k)}(B))^{\frac{t}{2}}(\mathcal{M}_0^{(k)}(A))^t(\mathcal{M}_0^{(k)}(B))^{\frac{t}{2}}\big)^{\frac{1}{t}}\big].
\end{aligned}$$

Since $A, B \in \mathbf{H}_n^+$, $\mathcal{M}_0^{(k)}(A)$ and $\mathcal{M}_0^{(k)}(B)$ are both Hermitian and positive semi-definite. Then inequality (18) follows immediately from the original Araki-Lieb-Thirring inequality [24] for normal trace. □

Lemma 3.7 (k-trace Golden-Thompson Inequality)**.** *For any $A, B \in \mathbf{H}_n$,*

$$\mathrm{Tr}_k\big[\exp(A+B)\big] \le \mathrm{Tr}_k\big[\exp(A)\exp(B)\big], \tag{19}$$

with equality holds if and only if $AB = BA$.

Proof. We here only prove the inequality. The condition for equality will be justified in an alternative proof of this lemma in Appendix B. For any $A, B \in \mathbf{H}_n^+$, we have

$$\begin{aligned}\mathrm{Tr}_k\big[\exp(A+B)\big] &= \lim_{m\to+\infty} \mathrm{Tr}_k\Big[\Big(\exp\big(\frac{1}{2m}B\big)\exp\big(\frac{1}{m}A\big)\exp\big(\frac{1}{2m}B\big)\Big)^m\Big] \\ &\le \mathrm{Tr}_k\big[\exp\big(\frac{1}{2}B\big)\exp\big(A\big)\exp\big(\frac{1}{2}B\big)\big] \\ &= \mathrm{Tr}_k\big[\exp\big(A\big)\exp\big(B\big)\big].\end{aligned}$$

The first equality above is the Lie product formula, and the inequality is due to Lemma 3.6. □

Lemma 3.8 (k-trace Peierls-Bogoliubov Inequality)**.** *The function*

$$A \longmapsto \log \mathrm{Tr}_k\big[\exp(A)\big] \tag{20}$$

is convex on $\mathbf{H}_n$.

Proof. For any $A, B \in \mathbf{H}_n^+$, $\tau \in (0,1)$, by Lemma 3.7 we have

$$\begin{aligned}\mathrm{Tr}_k\big[\exp(\tau A + (1-\tau)B)\big] &\le \mathrm{Tr}_k\big[\exp\big(\tau A\big)\exp\big((1-\tau)B\big)\big] \\ &\le \mathrm{Tr}_k\big[\exp(A)\big]^{\tau}\mathrm{Tr}_k\big[\exp(B)\big]^{1-\tau}.\end{aligned}$$

The second inequality above is Hölder's. Therefore

$$\log \mathrm{Tr}_k\big[\exp(\tau A + (1-\tau)B)\big] \le \tau \log \mathrm{Tr}_k\big[\exp(A)\big] + (1-\tau)\log \mathrm{Tr}_k\big[\exp(B)\big].$$

□

We remark that Lemma 3.8 can also be proved using the operator interpolation in Lemma 3.5. Lemma 3.8 immediately implies that $A \mapsto \log\phi\big(\exp(A)\big) = \frac{1}{k}\log \mathrm{Tr}_k\big[\exp(A)\big]$ is convex, and thus $A \mapsto \phi\big(\exp(A)\big)$ is convex. This will help us prove improve from $m = 1$ to $m \ge 1$ in Theorem 3.3.

Proof of Theorem 3.3 (Part II). Given any $\{A^{(j)}\}_{j=1}^m, \{B^{(j)}\}_{j=1}^m \subset \mathbf{H}_n^{++}$, and any $\tau \in [0,1]$, let $C^{(j)} = \tau A^{(j)} + (1-\tau)B^{(j)}, 1 \le j \le m$. Since the map $X \mapsto \phi(\exp(X))$ is convex on $\mathbf{H}_n$, the map $X \mapsto \phi(\exp(L+X))$ is also convex on $\mathbf{H}_n$ for arbitrary $L \in \mathbf{H}_n$. Now define

$$L = H + \sum_{j=1}^m p_j \log C^{(j)}, \quad r = \sum_{j=1}^m p_j \le 1.$$

We then have that

$$\begin{aligned}\phi\big(\exp(H + \sum_{j=1}^m p_j \log X^{(j)})\big) &= \phi\Big(\exp\big(H + r\sum_{j=1}^m \frac{p_j}{r}(\log X^{(j)} - \log C^{(j)}) + \sum_{j=1}^m p_j \log C^{(j)}\big)\Big) \\ &= \phi\Big(\exp\big(L + r\sum_{j=1}^m \frac{p_j}{r}(\log X^{(j)} - \log C^{(j)})\big)\Big) \\ &\le \sum_{j=1}^m \frac{p_j}{r}\phi\big(\exp(L + r\log X^{(j)} - r\log C^{(j)})\big), \quad X^{(j)} = A^{(j)}, B^{(j)}.\end{aligned}$$

For each j, by the concavity of (10) for $m = 1$, we have

$$\begin{aligned}
&\tau\phi\big(\exp(L + r\log A^{(j)} - r\log C^{(j)})\big) + (1-\tau)\phi\big(\exp(L + r\log B^{(j)} - r\log C^{(j)})\big)\\
&\leq \phi\big(\exp(L + r\log(\tau A^{(j)} + (1-\tau)B^{(j)}) - r\log C^{(j)})\big)\\
&= \phi\big(\exp(L)\big).
\end{aligned}$$

Therefore we obtain that

$$\begin{aligned}
&\tau\phi\big(\exp(H + \sum_{j=1}^{m} p_j \log A^{(j)})\big) + (1-\tau)\phi\big(\exp(H + \sum_{j=1}^{m} p_j \log B^{(j)})\big)\\
&\leq \sum_{j=1}^{m} \frac{p_j}{r}\phi\big(\exp(L)\big)\\
&= \phi\big(\exp(H + \sum_{j=1}^{m} p_j \log C^{(j)})\big),
\end{aligned}$$

that is, (10) is jointly concave on $(\mathbf{H}_n^{++})^{\times m}$ for all $m \geq 1$. □

3.4. Some Corollaries

The following corollary follows from standard arguments on homogeneous, concave functions.

Corollary 3.9. *For any $s, p, q \in (0, 1], p + q \leq 1$, and any $K \in \mathbb{C}^{n\times m}$, the functions*

$$(A, B) \longmapsto \phi\big((B^{\frac{qs}{2}} K^* A^{ps} K B^{\frac{qs}{2}})^{\frac{1}{s}}\big)^{\frac{1}{p+q}} \tag{21}$$

is jointly concave on $\mathcal{H}_n^{+} \times \mathbf{H}_m^{+}$. For any $H \in \mathbf{H}_n$ and any $\{p_j\}_{j=1}^{m} \subset (0, 1]$ such that $\sum_{j=1}^{m} p_j \leq 1$, the function

$$(A^{(1)}, A^{(2)}, \dots, A^{(m)}) \longmapsto \phi\big(\exp\big(H + \sum_{j=1}^{m} p_j \log A^{(j)}\big)\big)^{\frac{1}{\sum_{j=1}^{m} p_j}}, \tag{22}$$

is jointly concave on $(\mathcal{H}_n^{++})^{\times m}$.

Proof. Consider any matrix function $F : \mathbf{H}_n^{+} \to [0, +\infty)$(or $\mathbf{H}_n^{++} \to [0, +\infty)$) that is positively homogeneous of order 1, i.e. $F(\lambda A) = \lambda F(A), \forall \lambda \geq 0$. By Lemma D.1, we have that F is concave $\Longleftrightarrow$ F^s is concave for some $s \in (0, 1]$.

One can easily check that the functions (21) and (22) are positively homogeneous of order 1. Then this corollary follows from Theorem 3.2, Theorem 3.3 and Lemma D.1. □

The following corollary is an analog of the concave part of Lemma 3.1 in [22].

Corollary 3.10. *For any $s, r \in (0, 1]$ and any $\{K^{(j)}\}_{j=1}^{m} \subset \mathbb{C}^{n\times n}$, the function*

$$(A^{(1)}, A^{(2)}, \dots, A^{(m)}) \longmapsto \phi\Big(\Big(\sum_{j=1}^{m} (K^{(j)})^* (A^{(j)})^{rs} K^{(j)}\Big)^{\frac{1}{s}}\Big) \tag{23}$$

is jointly concave on $(\mathbf{H}_n^{+})^{\times m}$.

Proof. Define

$$\widehat{A} = \begin{pmatrix} A^{(1)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A^{(1)} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & A^{(m)} \end{pmatrix} \in \mathcal{H}_{mn}^{+}, \quad \widehat{K} = \begin{pmatrix} K^{(1)} & \mathbf{0} & \dots & \mathbf{0} \\ K^{(2)} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ K^{(m)} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix} \in \mathbb{C}^{2n\times 2n}.$$

Then we have

$$(\widehat{K}^*\widehat{A}^{rs}\widehat{K})^{\frac{1}{s}} = \begin{pmatrix} \left(\sum_{j=1}^m (K^{(j)})^*(A^{(j)})^{rs}K^{(j)}\right)^{\frac{1}{s}} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} \end{pmatrix},$$

and thus

$$\phi\big((\widehat{K}^*\widehat{A}^{rs}\widehat{K})^{\frac{1}{s}}\big) = \phi\Big(\Big(\sum_{j=1}^m (K^{(j)})^*(A^{(j)})^{rs}K^{(j)}\Big)^{\frac{1}{s}}\Big).$$

By Theorem 3.2, the left hand side above is concave in $\widehat{A}$, therefore the right hand side is jointly concave in $(A^{(1)}, A^{(2)}, \dots, A^{(m)})$. □

3.5. Revisiting previous proofs in trace case

As mentioned in the introduction, many alternative proofs of Lieb's concavity theorem have been found since its original establishment by Lieb in 1973. An proof using matrix tensors was given by Ando [7] in 1979 (see also Carlen [8]). Ando interpreted $\mathrm{Tr}[K^*A^pKB^q]$ as an inner product on the tensor space $\mathbb{C}^n \otimes \mathbb{C}^m$ and translated the Lieb's concavity theorem to the statement that the map $(A, B) \mapsto A^p \otimes B^q$ is operator concave. Ando then proved the latter using the integral representation of A^p. Here $\otimes$ is the Kronecker product. Later, Nikoufar et al. [9] provided a simpler proof for the concavity of $(A, B) \mapsto A^p \otimes B^q$ using the concept of matrix perspectives [25]. We summarize the ideas of their proofs as follows. For simplicity, we assume that $p+q = 1$. The result for $p+q = r < 1$ can be further obtained by using the fact that $A \mapsto A^r$ is operator monotone increasing and operator concave for $r \in (0, 1]$. For $p \in (0, 1]$, the map $A \mapsto (A \otimes I_m)^p = A^p \otimes I_m$ from $\mathbf{H}_n^+$ to $\mathbf{H}_{nm}^+$ is operator concave, and thus its perspective from $\mathbf{H}_n^+ \times \mathbf{H}_m^+$ to $\mathbf{H}_{nm}^+$,

$$(A, B) \mapsto (I_n \otimes B)^{\frac{1}{2}}\big((I_n \otimes B)^{-\frac{1}{2}}(A \otimes I_m)(I_n \otimes B)^{-\frac{1}{2}}\big)^p(I_n \otimes B)^{\frac{1}{2}} = A^p \otimes B^{1-p},$$

is jointly operator concave in (A, B). The simplified expression above results from the fact that $A \otimes I_m$ commutes with $I_n \otimes B$. For any $K \in \mathbb{C}^{n\times m}$, we have the identity (a variant of Ando's interpretation)

$$\mathrm{Tr}[K^*A^pKB^{1-p}] = \Big\langle \sum_{j=1}^m (Ke_j^{(m)}) \otimes e_j^{(m)}, A^p \otimes (B^T)^{1-p} \sum_{j=1}^m (Ke_j^{(m)}) \otimes e_j^{(m)} \Big\rangle_{\mathbb{C}^n \otimes \mathbb{C}^m}, \tag{24}$$

where B^T is the transpose of B, and $e_j^{(m)} = (0, \dots, \overset{j_{\text{th}}}{1}, \dots, 0) \in \mathbb{C}^m$. Note that since $B \in \mathbf{H}_n^+$, B^T is also in $\mathbf{H}_n^+$. Since $B \mapsto B^T$ is linear, the joint operator concavity of $(A, B) \mapsto A^p \otimes B^{1-p}$ then implies the joint concavity of $(A, B) \mapsto \mathrm{Tr}[K^*A^pKB^{1-p}]$.

As an application, Carlen and Lieb [22] applied the Lieb's concavity theorem to prove the concavity of $A \mapsto \mathrm{Tr}[(K^*A^{rs}K)^{\frac{1}{s}}]$ for $s, r \in (0, 1]$ (they used a slightly different but equivalent expression). They performed a variational argument based essentially on the Hölder's inequality for trace. We here provide a simplified proof that captures the main spirit. For any $A, B \in \mathbf{H}_n^+$, $K \in \mathbb{C}^{n\times n}$, let

$$X = (K^*A^{rs}K)^{\frac{1}{s}}, \quad Y = (K^*B^{rs}K)^{\frac{1}{s}}.$$

Then for any $\tau \in [0, 1]$, note that $rs + (1 - s) \le 1$, we have

$$\begin{aligned} \tau\mathrm{Tr}[X] + (1-\tau)\mathrm{Tr}[Y] &= \tau\mathrm{Tr}\big[K^*A^{rs}KX^{1-s}\big] + (1-\tau)\mathrm{Tr}\big[K^*B^{rs}KY^{1-s}\big] \\ &\le \mathrm{Tr}\big[K^*(\tau A + (1-\tau)B)^{rs}K(\tau X + (1-\tau)Y)^{1-s}\big] \\ &\le \mathrm{Tr}\big[\big(K^*(\tau A + (1-\tau)B)^{rs}K\big)^{\frac{1}{s}}\big]^s\mathrm{Tr}\big[\tau X + (1-\tau)Y\big]^{1-s} \\ &= \mathrm{Tr}\big[\big(K^*(\tau A + (1-\tau)B)^{rs}K\big)^{\frac{1}{s}}\big]^s\big(\tau\mathrm{Tr}[X] + (1-\tau)\mathrm{Tr}[Y]\big)^{1-s}, \end{aligned} \tag{25}$$

where the first inequality is due to Lieb's concavity theorem with $p = rs, q = 1-s, p+q = rs+1-s \le 1$, and the second inequality is Hölder's inequality for trace. The above then simplifies to

$$\tau\mathrm{Tr}\big[(K^*A^{rs}K)^{\frac{1}{s}}\big] + (1-\tau)\mathrm{Tr}[(K^*B^{rs}K)^{\frac{1}{s}}\big] \le \mathrm{Tr}\big[\big(K^*(\tau A + (1-\tau)B)^{rs}K\big)^{\frac{1}{s}}\big],$$

which concludes the concavity of $A \mapsto \mathrm{Tr}[(K^*A^{rs}K)^{\frac{1}{s}}]$.

These arguments using matrix tensors and variational forms were also adopted by Tropp [6] to provide an alternative proof of the concavity of $A \mapsto \mathrm{Tr}[\exp(H + \log A)]$. Tropp's proof is based on his variational formula for trace,

$$\mathrm{Tr}[M] = \sup_{T \in \mathbf{H}_n^{++}} \mathrm{Tr}\big[T \log M - T \log T + T\big], \quad M \in \mathbf{H}_n^{++}, \tag{26}$$

which relies on the non-negativeness of the matrix relative entropy

$$D(T; M) = \mathrm{Tr}\big[T(\log T - \log M) - (T - M)\big], \quad T, M \in \mathbf{H}_n^{++}.$$

The non-negativeness of $D(T; M)$ is a classical result of Klein's inequality (see Petz [26], Carlen [8] or Tropp [6]). Tropp substituted $M = \exp(H + \log A)$ in (26) to obtain

$$\mathrm{Tr}[\exp(H + \log A)] = \sup_{T \in \mathbf{H}_n^{++}} \big(\mathrm{Tr}[TH] + \mathrm{Tr}[A] - D(T; A)\big).$$

The concavity of $A \mapsto \mathrm{Tr}[\exp(H + \log A)]$ then follows from this variational expression, the joint convexity of $D(T; A)$ in (T, A), and the fact that $g(x) = \sup_{y \in \Omega} f(x, y)$ is concave in x if $f(x, y)$ is jointly concave in y and Ω is convex. The joint convexity of the relative entropy $D(T; A)$ was first due to Lindblad [27]. One can also see Ando [7], Carlen [8] and Tropp [6] for alternative proofs.

However elegant, the above approaches are found hardly generalizable to k-trace, as they more or less rely on the linearity of the normal trace. For example, the Ando's identity (24), the last step in (25) and Tropp's variational formula (26). Our k-trace function ϕ for $k > 1$ is at best sub-additive since it is concave and homogeneous of order 1.

We therefore look back to Lieb's original proof of his concavity theorem in [1]. Roughly speaking, Lieb made use of the maximum modulus principle to concentrate the powers of A^p, B^q in $\mathrm{Tr}[K^*A^pKB^q]$ to the only power $p+q$ on A or B, and then proceeded with the operator concavity of $X \mapsto X^{p+q}$ for $p + q \in (0, 1]$. This technique, relying on the holomorphicity of X^z ($X \in \mathbf{H}_n^+$) as a function of z, shed some light on the use of complex interpolation theories. In the same year, Epstein [10] pushed forward the use of holomorphicity with the theory of Herglotz functions, and provided an unified way of proving the concavity of $A \mapsto \mathrm{Tr}[K^*A^pKA^q]$, $A \mapsto \mathrm{Tr}[(K^*A^sK)^{\frac{1}{s}}]$ and $A \mapsto \mathrm{Tr}[\exp(H + \log A)]$. Later, Uhlmann [11](1977) applied interpolation theories explicitly to again prove Lieb's concavity theorem, by interpreting $\mathrm{Tr}[K^*A^pKB^{1-p}]$ as an interpolation between $\mathrm{Tr}[K^*AK]$ and $\mathrm{Tr}[K^*KB]$. Uhlmann's quadratic interpolation of seminorms extended the relevant works of Lieb to positive linear forms of arbitrary *-algebras. Kosaki [12](1982) further explored the idea of quadratic interpolation of seminorms and captured Lieb's concavity theorem in the frame of general interpolation theories. All these previous works have suggested us to adopt complex interpolation theories for extending Lieb's concavity to our k-trace version. In particular, we found the operator interpolation technique (Theorem 3.4) developed even earlier by Stein [15](1956) more compatible and flexible to our problem. As we have seen in the proof of Lemma 3.1 and Theorem 3.2, we can derive variant interpolation inequalities systematically by choosing $G(z)$ properly in inequality (13).

4. Other Results on *k*-trace

4.1. Multivariate Golden-Thompson Inequality

Sutter et al. [17] recently applied the operator interpolation in Theorem 3.4 to derive a multivariate extension of the Golden-Thompson (GT) inequality, which covers the original GT inequality and its three-matrix extension by Lieb [1].

Following the ideas in [17], we may also use Lemma 3.5 to further extend the multivariate GT inequality to a k-trace form. In what follows, we write $\prod_{j=1}^m X^{(j)}$ for the matrix multiplication in the index order, i.e. $\prod_{j=1}^m X^{(j)} = X^{(1)}X^{(2)} \cdots X^{(m)}$. We first present an analog of Theorem 3.2 in [17].

Lemma 4.1. *For any $A^{(1)}, A^{(2)}, \ldots, A^{(m)} \in \mathbf{H}_n^+$, $p \in [1, +\infty)$, $r \in (0, 1]$, the following inequality holds:*

$$\log \phi\Big(\Big|\prod_{j=1}^m (A^{(j)})^r\Big|^{\frac{p}{r}}\Big) \le \int_{-\infty}^{+\infty} dt \beta_r(t) \log \phi\Big(\Big|\prod_{j=1}^m (A^{(j)})^{1+it}\Big|^p\Big). \tag{27}$$

Proof. Define

$$G(z)=\prod_{j=1}^{m}(A^{(j)})^{z},\quad z\in\mathcal{S},$$

where $\mathcal{S}$ is defined as in Lemma 3.5. One can check that $G(z)$ is holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary, and $\|G(z)\|$ is uniform bounded on $\mathcal{S}$. We may first assume that each $A^{(j)}\in\mathbf{H}_n^{++}$ so that $(A^{(j)})^{it}, t\in\mathbb{R}$ is unitary. The result for $A^{(j)}\in\mathbf{H}_n^{+}$ can be obtained by continuity. Thus $G(it)$ is unitary for all $t\in\mathbb{R}$, and $|G(it)|^{p_0}=I_n$ for all p_0. Thus we can apply inequality (15) with $\theta=r, p_0\to+\infty, p_1=p, p_\theta=\frac{p}{r}$ to obtain

$$\log\phi\big(|G(r)|^{\frac{p}{r}}\big)\le\int_{-\infty}^{+\infty}dt\beta_r(t)\log\phi\big(|G(1+it)|^{p}\big),$$

which is exactly (27). □

Using a multivariate version of the Lie product formula, we immediately obtain the following from Lemma 4.1.

Theorem 4.2 (Multivariate Golden-Thompson Inequality for k-trace)**.** *For any* $A^{(1)},A^{(2)},\dots,A^{(m)}\in\mathbf{H}_n$*, the following inequality holds:*

$$\log\phi\Big(\Big(\exp\big(\sum_{j=1}^{m}A^{(j)}\big)\Big)^{p}\Big)\le\int_{-\infty}^{+\infty}dt\beta_0(t)\log\phi\Big(\Big|\prod_{j=1}^{m}\exp\big((1+it)A^{(j)}\big)\Big|^{p}\Big).\tag{28}$$

Proof. We only need to replace $A^{(j)}$ in inequality (27) by $\exp(A^{(j)})$, and take $r\to0$. Since each $\|\exp((1+it)A^{(j)})\|=\|\exp(A^{(j)})\exp(itA^{(j)})\|=\|\exp(A^{(j)})\|$ is uniformly bounded for all $t\in\mathbb{R}$, the right hand side of (27) then becomes the right hand side of (28). By a multivariate Lie product formula (see e.g. [17])

$$\lim_{r\searrow0}\Big(\exp(rX^{(1)})\exp(rX^{(2)})\cdots\exp(rX^{(m)})\Big)^{\frac{1}{r}}=\exp\big(\sum_{i=1}^{m}X^{(j)}\big),$$

the left hand side of (27) then becomes

$$\begin{aligned}\lim_{r\searrow0}\log\phi\Big(\Big|\prod_{j=1}^{m}\exp\big(rA^{(j)}\big)\Big|^{\frac{p}{r}}\Big)&=\lim_{r\searrow0}\log\phi\Big(\Big(\prod_{j=1}^{m}\exp\big(rA^{(m-j+1)}\big)\prod_{j=1}^{m}\exp\big(rA^{(j)}\big)\Big)^{\frac{p}{2r}}\Big)\\&=\log\phi\Big(\Big(\exp\big(\sum_{j=1}^{m}2A^{(j)}\big)\Big)^{\frac{p}{2}}\Big)\\&=\log\phi\Big(\Big(\exp\big(\sum_{j=1}^{m}A^{(j)}\big)\Big)^{p}\Big).\end{aligned}$$

□

If we choose $m=2,p=2$ in Theorem 4.2 and replace $A^{(j)}$ by $\frac{1}{2}A^{(j)}$, the right hand side of inequality (28) is independent of t by cyclicity of ϕ. We then recover the k-trace GT inequality

$$\phi\big(\exp(A^{(1)}+A^{(2)})\big)\le\phi\big(\exp(A^{(1)})\exp(A^{(2)})\big),$$

that we have obtained in Lemma 3.7. If we choose $m=3,p=2$ in Theorem 4.2 and again replace $A^{(j)}$ by $\frac{1}{2}A^{(j)}$, we have

$$\begin{aligned}&\log\phi\big(\exp(A^{(1)}+A^{(2)}+A^{(3)})\big)\\\le&\int_{-\infty}^{+\infty}dt\beta_0(t)\log\phi\Big(\exp(A^{(1)})\exp\big(\frac{1+it}{2}A^{(2)}\big)\exp(A^{(3)})\exp\big(\frac{1-it}{2}A^{(2)}\big)\Big)\\\le&\log\phi\Big(\int_{-\infty}^{+\infty}dt\beta_0(t)\exp(A^{(1)})\exp\big(\frac{1+it}{2}A^{(2)}\big)\exp(A^{(3)})\exp\big(\frac{1-it}{2}A^{(2)}\big)\Big)\end{aligned}$$

The second inequality above is due to concavity of logarithm and ϕ. If we define

$$\mathcal{T}_A[B] = \int_0^{+\infty} dt(A+tI_n)^{-1}B(A+tI_n)^{-1}, \quad A,B \in \mathbf{H}_n^{++},$$

and use Lemma 3.4 in [17] that

$$\int_0^{+\infty} dt(A^{-1}+tI_n)^{-1}B(A^{-1}+tI_n)^{-1} = \int_{-\infty}^{+\infty} dt\beta_0(t)A^{\frac{1+it}{2}}BA^{\frac{1-it}{2}}, \quad A,B \in \mathbf{H}_n^{++},$$

we then further obtain

$$\phi\big(\exp(A^{(1)}+A^{(2)}+A^{(3)})\big) \le \phi\big(\exp(A^{(1)})\mathcal{T}_{\exp(-A^{(2)})}[\exp(A^{(3)})]\big).$$

This can be seen as a k-trace generalization of Lieb's [1] three-matrix extension of the GT inequality that $\mathrm{Tr}[\exp(A+B+C)] \le \mathrm{Tr}[\exp(A)\mathcal{T}_{\exp(-B)}[\exp(C)]]$.

4.2. Monotonicity Preserving and Concavity Preserving

If a scalar function f is monotone, its extension to Hermitian matrices is not necessarily operator monotone. For instance, let $f(x) = x^3$, and

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix},$$

then f is monotone increasing, and $A \preceq B$. But neither $A^3 \preceq B^3$ nor $A^3 \succeq B^3$ is true. However, a composition with trace will preserve the monotonicity. That is, $\mathrm{Tr}[f(A)]$ is monotone increasing (or decreasing) in A with respect to Löwner partial order, if f is monotone increasing (or decreasing). Likewise, if f is concave (or convex), its extension to Hermitian matrices is not necessarily operator concave (or convex), but $A \mapsto \mathrm{Tr}[f(A)]$ is still concave (or convex). One can see Theorem 2.10 in [8]. This means that, some partial information like trace may preserve monotonicity and concavity. In fact, we will show that for any integer k the partial information $\phi(\cdot) = \mathrm{Tr}_k[\cdot]^{\frac{1}{k}}$ also preserves monotonicity and concavity of scalar functions. But we need to restrict to f that only takes values in $[0,+\infty)$. We need the following lemma for proving concavity preserving.

Lemma 4.3. *For any $A \in \mathbf{H}_n^+$, let $\mathrm{diag}(A)$ denote the diagonal part of A, then*

$$\phi(A) \le \phi(\mathrm{diag}(A)).$$

Proof. Let D be a $n \times n$ diagonal matrix, whose diagonal entries follow independent Rademacher distributions, i.e.

$$D_{ii} = \begin{cases} 1, & \text{with prob. } 0.5, \\ -1, & \text{with prob. } 0.5. \end{cases}$$

Since $D^2 \equiv I_n$, we have $\phi(A) = \phi(DAD)$. Also notice that $\mathbb{E}[DAD] = \mathrm{diag}(A)$, since $\mathbb{E}[D_{ii}D_{jj}] = \delta_{ij}$. Then by concavity of ϕ, we have

$$\phi(A) = \mathbb{E}\phi(DAD) \le \phi(\mathbb{E}[DAD]) = \phi(\mathrm{diag}(A)).$$

□

Lemma 4.3 can be proved, instead, using the concept of Majorization. Let $\boldsymbol{a} = \{a_i\}_i^n$ and $\boldsymbol{b} = \{b_i\}_i^n$ be two sequences, both in descending order, i.e. $a_1 \ge a_2 \ge \cdots \ge a_n$, $b_1 \ge b_2 \ge \cdots \ge b_n$. We say $\boldsymbol{b}$ majorizes $\boldsymbol{a}$, denoted by $\boldsymbol{b} \succeq \boldsymbol{a}$, if

$$\sum_{i=1}^k b_i \ge \sum_{i=1}^k a_i, \quad 1 \le k \le n,$$

and the equality holds for $k = n$. It is not hard to show that, if $\boldsymbol{b} \succeq \boldsymbol{a}$, and $\boldsymbol{a}, \boldsymbol{b} \subset [0, +\infty)$, then $\prod_{i=1}^d b_i \leq \prod_{i=1}^d a_i$. Now for any $A \in \mathbf{H}_n^+$, let $\boldsymbol{\lambda} = \{\lambda_i\}_{i=1}^n$ and $\boldsymbol{a} = \{a_i\}_i^n$ be the eigenvalues and the diagonal entries of A respectively, both in descending order. Then since

$$\sum_{i=1}^k \lambda_i = \max_{\substack{\{v_i\}_{i=1}^k \subset \mathbb{C}^n \\ v_i^* v_j = \delta_{ij}}} \sum_{i=1}^k v_i^* A v_i \geq \sum_{i=1}^k e_i^* A e_i = \sum_{i=1}^k a_i, \quad 1 \leq k < n$$

and $\sum_{i=1}^n \lambda_i = \mathrm{Tr}[A] = \sum_{i=1}^n a_i$, we have $\boldsymbol{\lambda} \succeq \boldsymbol{a}$. Therefore we know that

$$\det[A] = \prod_{i=1}^n \lambda_i \leq \prod_{i=1}^n a_i = \det[\mathrm{diag}(A)].$$

Then using the equivalent definition (6) of k-trace, we have that, for any $1 \leq k \leq n$,

$$\begin{aligned} \mathrm{Tr}_k[A] &= \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \det[A_{(i_1 \cdots i_k, i_1 \cdots i_k)}] \\ &\leq \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} A_{i_1 i_1} A_{i_2 i_2} \cdots A_{i_k i_k}, \\ &= \mathrm{Tr}_k[\mathrm{diag}(A)]. \end{aligned}$$

Theorem 4.4. *Given a function $f : \mathbb{R} \to [0, +\infty)$, if f is monotone increasing (or decreasing) as a scalar function, then $\phi(f(\cdot))$ is monotone increasing (or decreasing) on $\mathbf{H}_n^+$, in the sense that $\phi(f(A)) \geq \phi(f(B))$(or $\phi(f(A)) \leq \phi(f(B))$) if $A, B \in \mathbf{H}_n^+, A \succeq B$; if f is concave as a scalar function, then $\phi(f(\cdot))$ is concave on $\mathbf{H}_n^+$.*

Proof. We first prove the monotonicity preserving property of ϕ. For any matrix $A \in \mathbf{H}_n$, we denote by $\lambda_i(A)$ the i_{th} largest eigenvalue of A. For any $A, B \in \mathbf{H}_n^+$, if $A \succeq B$, then $\lambda_i(A) \geq \lambda_i(B), 1 \leq i \leq n$. Therefore if f is monotone increasing, we immediately have

$$\lambda_i(f(A)) = f(\lambda_i(A)) \geq f(\lambda_i(B)) = \lambda_i(f(B)), \quad 1 \leq i \leq n,$$

and thus $\phi(f(A)) \geq \phi(f(B))$ by definition ϕ. Similarly, if f is monotone decreasing, we have

$$\lambda_i(f(A)) = f(\lambda_{n-i+1}(A)) \leq f(\lambda_{n-i+1}(B)) = \lambda_i(f(B)), \quad 1 \leq i \leq n,$$

and thus $\phi(f(A)) \leq \phi(f(B))$.

Next we prove the concavity preserving property of ϕ. Given any $A, B \in \mathbf{H}_n^+$, and any $\tau \in [0, 1]$, we define $C = \tau A + (1 - \tau)B$. Let $U = [u_1, u_2, \cdots, u_n]$ be a unitary matrix such that the columns are all eigenvectors of C, then

$$U^* f(C) U = f(U^* C U) = f(\mathrm{diag}(U^* C U)).$$

If f is concave, then

$$f(u_i^* C u_i) = f(\tau u_i^* A u_i + (1 - \tau) u_i^* B u_i) \geq \tau f(u_i^* A u_i) + (1 - \tau) f(u_i^* B u_i), \quad 1 \leq i \leq n,$$

and thus $f(\mathrm{diag}(U^* C U)) \succeq \tau f(\mathrm{diag}(U^* A U)) + (1 - \tau) f(\mathrm{diag}(U^* B U))$. Further, for any unit vector $u \in \mathbb{C}^n$, we have

$$\begin{aligned} f(u^* A u) &= f\Big(\sum_{i=1}^n \lambda_i(A) u^* v_i v_i^* u\Big) = f\Big(\sum_{i=1}^n \lambda_j(A) |v_i^* u|^2\Big) \\ &\geq \sum_{i=1}^n |v_i^* u|^2 f\big(\lambda_i(A)\big) = u^* \Big(\sum_{i=1}^n f\big(\lambda_i(A)\big) v_i v_i^*\Big) u = u^* f(A) u, \end{aligned}$$

where $v_1, v_2, \cdots, v_n$ are all eigenvectors of A, and we have used that $\sum_{i=1}^n |v_i^* u|^2 = \|u\|_2^2 = 1$. Then we have $f(\mathrm{diag}(U^*AU)) \succeq \mathrm{diag}(U^* f(A) U)$. Similar, we have $f(\mathrm{diag}(U^*BU)) \succeq \mathrm{diag}(U^* f(B) U)$. Finally, we can compute

$$\begin{aligned}
\phi\big(f(C)\big) &= \phi\big(U^* f(C) U\big) \\
&= \phi\big(f(\mathrm{diag}(U^*CU))\big) \\
&\geq \phi\big(\tau f(\mathrm{diag}(U^*AU)) + (1-\tau) f(\mathrm{diag}(U^*BU))\big) \\
&\geq \tau\phi\big(f(\mathrm{diag}(U^*AU))\big) + (1-\tau)\phi\big(f(\mathrm{diag}(U^*BU))\big) \\
&\geq \tau\phi\big(\mathrm{diag}(f(U^*AU))\big) + (1-\tau)\phi\big(\mathrm{diag}(f(U^*BU))\big) \\
&\geq \tau\phi\big(f(U^*AU)\big) + (1-\tau)\phi\big(f(U^*BU)\big) \\
&= \tau\phi\big(U^* f(A) U\big) + (1-\tau)\phi\big(U^* f(B) U\big) \\
&= \tau\phi\big(f(A)\big) + (1-\tau)\phi\big(f(B)\big).
\end{aligned}$$

We have used Lemma 4.3 for the last inequality above. Therefore $\phi(f(\cdot))$ is concave on $\mathbf{H}_n^+$. □

Appendix

A. Mixed Discriminant

The mixed discriminant $D(A^{(1)}, A^{(2)}, \cdots, A^{(n)})$ of n matrices $A^{(1)}, A^{(2)}, \cdots, A^{(n)} \in \mathbb{C}^{n\times n}$ is defined as

$$D(A^{(1)}, A^{(2)}, \cdots, A^{(n)}) = \frac{1}{n!}\sum_{\sigma\in S_n} \det \begin{bmatrix} A_{11}^{(\sigma(1))} & A_{12}^{(\sigma(2))} & \cdots & A_{1n}^{(\sigma(n))} \\ A_{21}^{(\sigma(1))} & A_{22}^{(\sigma(2))} & \cdots & A_{1n}^{(\sigma(n))} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}^{(\sigma(1))} & A_{n2}^{(\sigma(2))} & \cdots & A_{nn}^{(\sigma(n))} \end{bmatrix}, \tag{A.1}$$

where S_n denotes the symmetric group of order n. We here list some basic facts about mixed discriminants. For more properties of mixed discriminants, one may refer to [28, 29].

- Symmetry: $D(A^{(1)}, A^{(2)}, \cdots, A^{(n)})$ is symmetric in $A^{(1)}, A^{(2)}, \cdots, A^{(n)}$, i.e.
$$D(A^{(1)}, A^{(2)}, \cdots, A^{(n)}) = D(A^{\sigma(1)}, A^{\sigma(2)}, \cdots, A^{\sigma(n)}), \quad \sigma \in S_n.$$
- Multilinearity: for any $\alpha, \beta \in \mathbb{R}$,
$$D(\alpha A + \beta B, A^{(2)}, \cdots, A^{(n)}) = \alpha D(A, A^{(2)}, \cdots, A^{(n)}) + \beta D(B, A^{(2)}, \cdots, A^{(n)}).$$
- Positiveness [28]: If $A^{(1)}, A^{(2)}, \cdots, A^{(n)} \in \mathbf{H}_n^+$, then $D(A^{(1)}, A^{(2)}, \cdots, A^{(n)}) \geq 0$; if $A^{(1)}, A^{(2)}, \cdots, A^{(n)} \in \mathbf{H}_n^{++}$, then $D(A^{(1)}, A^{(2)}, \cdots, A^{(n)}) > 0$.

The relation between the mixed discriminant and Tr_k is obvious. If we calculate the mixed discriminant for k copies of $A \in \mathbb{C}^{n\times n}$ and $n-k$ copies of I_n, we can find that

$$D(\underbrace{A, \cdots, A}_{k}, \underbrace{I_n, \cdots, I_n}_{n-k}) = \binom{n}{k}^{-1} \sum_{1\leq i_1 < i_2 < \cdots < i_k \leq n} \det[A_{(i_1\cdots i_k, i_1\cdots i_k)}] = \binom{n}{k}^{-1} \mathrm{Tr}_k[A]. \tag{A.2}$$

This is why the mixed discriminant plays an important role in the proof of our main theorem. In particular, we will need the following inequality on mixed discriminant by Alexandrov [30].

Theorem A.1. ***(Alexandrov-Fenchel Inequality for Mixed Discriminants)*** *For any $B \in \mathbf{H}_n$ and any $A, \underbrace{A^{(3)}, \cdots, A^{(n)}}_{n-2} \in \mathbf{H}_n^{++}$, we have*

$$D(A, B, A^{(3)}, \cdots, A^{(n)})^2 \geq D(A, A, A^{(3)}, \cdots, A^{(n)}) D(B, B, A^{(3)}, \cdots, A^{(n)}), \tag{A.3}$$

with equality if and only if $B = \lambda A$ for some $\lambda \in \mathbb{R}$.

This theorem originally applied to real symmetric matrices when established in 1937. A proof of its extension to Hermitian matrices can be found in [31]. By continuity, inequality (A.3) can extend to the case that $A, A^{(3)}, \cdots, A^{(n)} \in \mathbf{H}_n^+$, but the necessity of the condition for equality is no longer valid.

Repeatedly applying the Alexandrov-Fenchel inequality (A.3) grants us the following corollary.

Corollary A.2. *For any $0 \le l \le k \le n$, and any $A, B, \underbrace{A^{(k+1)}, \cdots, A^{(n)}}_{n-k} \in \mathbf{H}_n^+$, we have*

$$\begin{aligned} &D(\underbrace{A, \cdots, A}_{l}, \underbrace{B, \cdots, B}_{k-l}, \underbrace{A^{(k+1)}, \cdots, A^{(n)}}_{n-k})^k \\ \ge\ &D(\underbrace{A, \cdots, A}_{k}, \underbrace{A^{(k+1)}, \cdots, A^{(n)}}_{n-k})^l \cdot D(\underbrace{B, \cdots, B}_{k}, \underbrace{A^{(k+1)}, \cdots, A^{(n)}}_{n-k})^{k-l}. \end{aligned} \tag{A.4}$$

A direct result of Corollary A.2 is the following general Brunn-Minkowski theorem for mixed discriminants.

Corollary A.3. ***(General Brunn-Minkowski Theorem for Mixed Discriminants)*** *For any $1 \le k \le n$, and any fixed $\underbrace{A^{(k+1)}, \cdots, A^{(n)}}_{n-k} \in \mathbf{H}_n^+$, the function*

$$\begin{aligned} \mathbf{H}_n^+ &\longrightarrow \mathbb{R} \\ A &\longmapsto D(\underbrace{A, \cdots, A}_{k}, \underbrace{A^{(k+1)}, \cdots, A^{(n)}}_{n-k})^{\frac{1}{k}} \end{aligned} \tag{A.5}$$

is concave.

If we choose $A^{(k+1)}, \cdots, A^{(n)}$ to be $n-k$ copies of I_n, we can immediately conclude from Corollary A.3 that the function $A \mapsto \big(\mathrm{Tr}_k[A]\big)^{\frac{1}{k}}$ is concave on $\mathbf{H}_n^+$.

B. Exterior Algebra

We give a brief review of exterior algebras on the vector space $\mathbb{C}^n$. For more details, one may refer to [32, 33]. For the convenience of our use, the notations in our paper might be different from those in other materials. For any $1 \le k \le n$, let $\wedge^k(\mathbb{C}^n)$ denote the vector space of the k_{th} exterior algebra of $\mathbb{C}^n$, equipped with the inner product

$$\begin{aligned} \langle\cdot,\cdot\rangle_{\wedge^k}: \quad \wedge^k(\mathbb{C}^n) \times \wedge^k(\mathbb{C}^n) &\longrightarrow \mathbb{C} \\ \langle u_1 \wedge \cdots \wedge u_k, v_1 \wedge \cdots \wedge v_k\rangle_{\wedge^k} &= \det \begin{bmatrix} \langle u_1, v_1\rangle & \langle u_1, v_2\rangle & \cdots & \langle u_1, v_k\rangle \\ \langle u_2, v_1\rangle & \langle u_2, v_2\rangle & \cdots & \langle u_2, v_k\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_k, v_1\rangle & \langle u_k, v_2\rangle & \cdots & \langle u_k, v_k\rangle \end{bmatrix}, \end{aligned}$$

where $\langle u, v\rangle = u^* v$ is the standard l_2 inner product on $\mathbb{C}^n$.

Let $\mathcal{L}(\wedge^k(\mathbb{C}^n))$ denote the space of all linear operators from $\wedge^k(\mathbb{C}^n)$ to itself. For any matrices $A^{(1)}, A^{(2)}, \cdots, A^{(k)} \in \mathbb{C}^{n\times n}$, we can define an element in $\mathcal{L}(\wedge^k(\mathbb{C}^n))$:

$$\begin{aligned} \mathcal{M}^{(k)}(A^{(1)}, A^{(2)}, \cdots, A^{(k)}): \quad \wedge^k(\mathbb{C}^n) &\longrightarrow \wedge^k(\mathbb{C}^n) \\ v_1 \wedge v_2 \wedge \cdots \wedge v_k &\longmapsto \sum_{\sigma\in S_k} A^{(\sigma(1))}v_1 \wedge A^{(\sigma(2))}v_2 \wedge \cdots \wedge A^{(\sigma(k))}v_k, \end{aligned} \tag{B.1}$$

where S_k is the symmetric group of order k. Apparently, the map $\mathcal{M}^{(k)}(A^{(1)}, A^{(2)}, \cdots, A^{(k)})$ is symmetric in $A^{(1)}, A^{(2)}, \cdots, A^{(k)}$, and linear in each single $A^{(i)}$. For simplicity, we will use the following notations for any matrix $A, B \in \mathbb{C}^{n\times n}$:

$$\mathcal{M}_0^{(k)}(A) = \frac{1}{k!}\mathcal{M}^{(k)}(A, \cdots, A), \tag{B.2a}$$

$$\mathcal{M}_1^{(k)}(A; B) = \frac{1}{(k-1)!}\mathcal{M}^{(k)}(A, B, \cdots, B). \tag{B.2b}$$

Obviously the identity operator in $\mathcal{L}(\wedge^k(\mathbb{C}^n))$ is $\mathcal{M}_0(I_n)$. Note that due to the skew-symmetric property of exterior algebra, the spectrum of $\mathcal{M}_0^{(k)}(A)$ is $\{\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}\}_{1\le i_1<i_2<\cdots<i_k\le n}$, where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A. It is easy to verify that $\mathcal{M}_0^{(k)}$ has the following properties:

- Invertibility: If $A \in \mathbb{C}^{n\times n}$ is invertible, then $(\mathcal{M}_0^{(k)}(A))^{-1} = \mathcal{M}_0^{(k)}(A^{-1})$.
- Adjoint: For any $A \in \mathbb{C}^{n\times n}$, $(\mathcal{M}_0^{(k)}(A))^* = \mathcal{M}_0^{(k)}(A^*)$, with respect to the inner product $\langle\cdot,\cdot\rangle_{\wedge^k}$.
- Power: For any $A \in \mathbf{H}_n$ and any $t \in \mathbb{C}$, $\big(\mathcal{M}_0^{(k)}(A)\big)^t = \mathcal{M}_0^{(k)}(A^t)$.
- Positiveness: If $A \in \mathbf{H}_n$, then $\mathcal{M}_0^{(k)}(A)$ is Hermitian; if $A \in \mathbf{H}_n^+$, then $\mathcal{M}_0^{(k)}(A) \succeq \mathbf{0}$; if $A \in \mathbf{H}_n^{++}$, then $\mathcal{M}_0^{(k)}(A) \succ \mathbf{0}$.
- Product: For any $A, B \in \mathbb{C}^{n\times n}$, $\mathcal{M}_0^{(k)}(AB) = \mathcal{M}_0^{(k)}(A)\mathcal{M}_0^{(k)}(B)$.

Using these properties, one can check that

$$|\mathcal{M}_0^{(k)}(A)| = \big((\mathcal{M}_0^{(k)}(A))^*\mathcal{M}_0^{(k)}(A)\big)^{\frac{1}{2}} = \mathcal{M}_0^{(k)}\big((A^*A)^{\frac{1}{2}}\big) = \mathcal{M}_0^{(k)}(|A|).$$

Next we consider the natural basis of $\wedge^k(\mathbb{C}^n)$,

$$\{e_{i_1}\wedge e_{i_2}\wedge\cdots\wedge e_{i_k}\}_{1\le i_1<i_2<\cdots<i_k\le n},$$

which is orthogonal under the inner product $\langle\cdot,\cdot\rangle_{\wedge^k}$. Then the trace function on $\mathcal{L}(\wedge^k(\mathbb{C}^n))$ is defined as

$$\begin{aligned} \mathrm{Tr}: \quad \mathcal{L}(\wedge^k(\mathbb{C}^n)) &\longrightarrow \mathbb{C} \\ \mathrm{Tr}\big[\mathcal{F}\big] &= \sum_{1\le i_1<i_2<\cdots<i_k\le n} \langle e_{i_1}\wedge e_{i_2}\wedge\cdots\wedge e_{i_k}, \mathcal{F}(e_{i_1}\wedge e_{i_2}\wedge\cdots\wedge e_{i_k})\rangle_{\wedge^k}. \end{aligned} \tag{B.3}$$

It is not hard to check that this trace function is also invariant under cyclic permutation, i.e. $\mathrm{Tr}\big[\mathcal{F}\mathcal{G}\big] = \mathrm{Tr}\big[\mathcal{G}\mathcal{F}\big]$ for any $\mathcal{F}, \mathcal{G} \in \mathcal{L}(\wedge^k(\mathbb{C}^n))$. Then for any $A^{(1)}, \cdots, A^{(k)} \in \mathbb{C}^{n\times n}$, the trace $\mathrm{Tr}[\mathcal{M}^{(k)}(A^{(1)}, \cdots, A^{(k)})]$ coincides with the definition of the mixed discriminant, as one can check that

$$\begin{aligned} \mathrm{Tr}\big[\mathcal{M}^{(k)}(A^{(1)}, \cdots, A^{(k)})\big] &= \sum_{\sigma\in S_k}\sum_{1\le i_1<\cdots<i_k\le n} \langle e_{i_1}\wedge\cdots\wedge e_{i_k}, A^{(\sigma(1))}e_{i_1}\wedge\cdots\wedge A^{(\sigma(k))}e_{i_k}\rangle_{\wedge^k} \\ &= \frac{n!}{(n-k)!} D(A^{(1)}, \cdots, A^{(k)}, \underbrace{I_n, \cdots, I_n}_{n-k}). \end{aligned} \tag{B.4}$$

From this observation, we can now express the k-trace of a matrix $A \in \mathbb{C}^{n\times n}$ as

$$\mathrm{Tr}_k[A] = \mathrm{Tr}\big[\mathcal{M}_0^{(k)}(A)\big]. \tag{B.5}$$

For those who are familiar with exterior algebra, it is clear that the spectrum of $\mathcal{M}_0^{(k)}$ is just $\{\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}\}_{1\le i_1<i_2<\cdots<i_k\le n}$, where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A. So in this way it is more convenient to see that $\mathrm{Tr}\big[\mathcal{M}_0^{(k)}(A)\big] = \mathrm{sum}(\text{spectrum of } \mathcal{M}_0^{(k)}(A)) = \sum_{1\le i_1<\cdots<i_k\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k} = \mathrm{Tr}_k[A]$.

We here provide an alternative of Lemma 3.7 using the following lemma.

Lemma B.1. *For any $A \in \mathbf{H}_n$, we have*

$$\mathcal{M}_0^{(k)}\big(\exp(A)\big) = \exp\big(\mathcal{M}_1^{(k)}(A; I_n)\big).$$

Proof. We need to show that for any $v_1 \wedge v_2 \wedge \cdots \wedge v_k \in \wedge^k(\mathbb{C}^n)$,

$$\mathcal{M}_0^{(k)}\big(\exp(A)\big)(v_1 \wedge v_2 \wedge \cdots \wedge v_k) = \exp\big(\mathcal{M}_1^{(k)}(A; I_n)\big)(v_1 \wedge v_2 \wedge \cdots \wedge v_k). \tag{B.6}$$

We use Taylor expansion of e^x to expand

$$\mathcal{M}_0^{(k)}\big(\exp(A)\big) = \mathcal{M}_0^{(k)}\Big(\sum_{j=0}^{+\infty}\frac{1}{j!}A^j\Big), \quad \exp\big(\mathcal{M}_1^{(k)}(A; I_n)\big) = \sum_{j=0}^{+\infty}\frac{1}{j!}\big(\mathcal{M}_1^{(k)}(A; I_n)\big)^j.$$

Then for any integers $j_1, j_2, \ldots, j_k \geq 0$, the coefficient of the term $A^{j_1}v_1 \wedge A^{j_2}v_2 \wedge \cdots \wedge A^{j_k}v_k$ in the left hand side of (B.6) is

$$\frac{1}{j_1!j_2!\cdots j_k!},$$

and the coefficient of the same term in the right hand side of (B.6) is also

$$\frac{1}{J!}\binom{J}{j_1}\binom{J-j_1}{j_2}\cdots\binom{J-j_1-j_2-\cdots-j_{k-1}}{j_k} = \frac{1}{j_1!j_2!\cdots j_k!} \quad (J = j_1 + j_2 + \cdots + j_k).$$

□

An alternative proof of Lemma 3.7. Using Lemma B.1 and the original GT inequality for normal trace, we have

$$\begin{aligned}
\mathrm{Tr}_k[\exp(A+B)] &= \mathrm{Tr}\big[\mathcal{M}_0^{(k)}\big(\exp(A+B)\big)\big] \\
&= \mathrm{Tr}\big[\exp\big(\mathcal{M}_1^{(k)}(A+B; I_n)\big)\big] \\
&= \mathrm{Tr}\big[\exp\big(\mathcal{M}_1^{(k)}(A; I_n) + \mathcal{M}_1^{(k)}(B; I_n)\big)\big] \\
&\leq \mathrm{Tr}\big[\exp\big(\mathcal{M}_1^{(k)}(A; I_n)\big)\exp\big(\mathcal{M}_1^{(k)}(B; I_n)\big)\big] \\
&= \mathrm{Tr}\big[\mathcal{M}_0^{(k)}\big(\exp(A)\big)\mathcal{M}_0^{(k)}\big(\exp(B)\big)\big] \\
&= \mathrm{Tr}_k\big[\exp(A)\exp(B)\big],
\end{aligned}$$

where we have used that $\mathcal{M}_1^{(k)}(X; I_n)$ is linear in X. As shown by Petz [26], in the original GT inequality, the equality $\mathrm{Tr}[\exp(A+B)] = \mathrm{Tr}[\exp(A)\exp(B)]$ holds for $A, B \in \mathbf{H}_n$ if and only if $AB = BA$. Therefore, according to our calculation above, the equality $\mathrm{Tr}_k[\exp(A+B)] = \mathrm{Tr}_k[\exp(A)\exp(B)]$ holds if and only if

$$\mathcal{M}_1^{(k)}(A; I_n)\mathcal{M}_1^{(k)}(B; I_n) = \mathcal{M}_1^{(k)}(B; I_n)\mathcal{M}_1^{(k)}(A; I_n). \tag{B.7}$$

However, one can check by definition that (B.7) is true if and only if $AB = BA$. □

C. Complex Interpolation

The idea of complex interpolation originates from an important result in harmonic analysis, the Hadamard three-lines theorem [13], that if $f(z)$ is uniformly bounded on $\mathcal{S} = \{z \in \mathbb{C} : 0 \leq \mathrm{Re}(z) \leq 1\}$, holomorphic in the interior and continuous on the boundary, then $g(x) = \log\sup_y|f(x+iy)|$ is a convex function on $[0,1]$. Hirschman [23] imporved this theorem to the following.

Theorem C.1 (Hirschman)**.** *Let $f(z)$ be uniformly bounded on $\mathcal{S}$, holomorphic in the interior and continuous on the boundary. Then for $\theta \in (0,1)$, we have*

$$\log|f(\theta)| \leq \int_{-\infty}^{+\infty} dt\big(\beta_{1-\theta}(t)\log|f(it)|^{1-\theta} + \beta_\theta(t)\log|f(1+it)|^\theta\big).$$

Moreover, the assumption that $f(z)$ is uniformly bounded can be relaxed to

$$\log|f(z)| \leq Ce^{a\mathrm{Im}(z)}, \quad \forall z \in \mathcal{S}, \quad \textit{for some constants } C < +\infty,\ a < \pi.$$

Stein [15] further generalized this complex interpolation theory to interpolation of linear operators as described in Theorem 3.4. Our key lemma, Lemma 3.5, is then a natural extension of Theorem 3.4 from Schatten norm to k-trace. We provide a proof of Lemma 3.5 as follows.

Proof of Lemma 3.5. For any $X \in \mathbb{C}^{n\times n}$ and $p \in [1, +\infty)$, we have that

$$\mathcal{M}_0^{(k)}\big(|X|^p\big) = \mathcal{M}_0^{(k)}\big((X^*X)^{\frac{p}{2}}\big) = \big(\mathcal{M}_0^{(k)}(X^*X)\big)^{\frac{p}{2}} = \big((\mathcal{M}_0^{(k)}(X))^*\mathcal{M}_0^{(k)}(X)\big)^{\frac{p}{2}} = \big|\mathcal{M}_0^{(k)}(X)\big|^p,$$

and thus

$$\mathrm{Tr}_k\big[|X|^p\big]^{\frac{1}{p}} = \mathrm{Tr}\big[\mathcal{M}_0^{(k)}\big(|X|^p\big)\big]^{\frac{1}{p}} = \mathrm{Tr}\big[\big|\mathcal{M}_0^{(k)}(X)\big|^p\big]^{\frac{1}{p}} = \big\|\mathcal{M}_0^{(k)}(X)\big\|_p.$$

The above equality also holds for $p \to +\infty$ since we are dealing with finite dimensional operators. If $G(z)$ is holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary, then so is $\mathcal{M}_0^{(k)}\big(G(z)\big)$. And if $\|G(z)\|$ is uniformly bounded on $\mathcal{S}$, then $\|\mathcal{M}_0^{(k)}\big(G(z)\big)\|_{p_{\mathrm{Re}(z)}}$ is also uniformly bounded on $\mathcal{S}$, since all norms are equivalent for finite dimensional operators. Therefore we can use Theorem 3.4 with $G(z)$ replaced by $\mathcal{M}_0^{(k)}\big(G(z)\big)$ to get

$$\begin{aligned}&\log\big(\mathrm{Tr}_k\big[|G(\theta)|^{p_\theta}\big]^{\frac{1}{p_\theta}}\big)\\ &\le \int_{-\infty}^{+\infty} dt\Big(\beta_{1-\theta}(t)\log\big(\mathrm{Tr}_k\big[|G(it)|^{p_0}\big]^{\frac{1-\theta}{p_0}}\big) + \beta_\theta(t)\log\big(\mathrm{Tr}_k\big[|G(1+it)|^{p_1}\big]^{\frac{\theta}{p_1}}\big)\Big).\end{aligned}$$

We then multiply both sides by with $\frac{1}{k}$ to obtain (14). □

We remark that the operator interpolation inequality in Theorem 3.4 is interestingly powerful and user-friendly for proving matrix inequalities, as we have seen in the proofs of Lemma 3.1 and Theorem 3.2. In fact this technique can also prove many fundamental results in matrix theories. We here show one more example to see how we can use the trick of interpolation to prove that the map $A \mapsto A^r$ is operator concave on $\mathbf{H}_n^+$ for $r \in (0, 1]$. Note that, in general, this result is proved by using an integral expression for A^r (e.g. see [8]).

We need to show that for any $A, B \in \mathbf{H}_n^+$, $\tau \in [0, 1]$,

$$\tau A^r + (1-\tau)B^r \le C^r, \tag{C.1}$$

where $C = \tau A + (1-\tau)B$. We may assume that C is invertible. The case when C is not invertible can be handled by continuity. Notice that $X \preceq I_n$ if and only if $\mathrm{Tr}[K^*XK] \le \mathrm{Tr}[K^*K]$ for all $K \in \mathbb{C}^{n\times n}$. (C.1) is then equivalent to the statement that

$$\tau\mathrm{Tr}[K^*C^{-\frac{r}{2}}A^rC^{-\frac{r}{2}}K] + (1-\tau)\mathrm{Tr}[K^*C^{-\frac{r}{2}}B^rC^{-\frac{r}{2}}K] \le \mathrm{Tr}[K^*K], \quad \forall K \in \mathbb{C}^{n\times n}.$$

Now we fix K and define

$$G_X(z) = X^{\frac{z}{2}}C^{-\frac{z}{2}}K, \quad X = A, B,$$

so we have

$$\mathrm{Tr}[K^*C^{-\frac{r}{2}}X^rC^{-\frac{r}{2}}K] = \|G_X(r)\|_2^2.$$

$G_X(z)$ is holomorphic in the interior of $\mathcal{S}$ and continuous on the boundary, and $\|G_X(z)\|$ is uniformly bounded on $\mathcal{S}$. We then use inequality (13) in Theorem 3.4 with $\theta = r, p_\theta = p_0 = p_1 = 2$ to obtain

$$\|G_X(r)\|_2^2 \le \int_{-\infty}^{+\infty} dt\Big((1-r)\beta_{1-r}(t)\|G_X(it)\|_2^2 + r\beta_r(t)\|G_X(1+it)\|_2^2\Big).$$

We have again used Jensen's inequality to get rid of the logarithms. For each $t \in \mathbb{R}$, we have that

$$\|G_X(it)\|_2^2 = \mathrm{Tr}[K^*C^{\frac{it}{2}}X^{-\frac{it}{2}}X^{\frac{it}{2}}C^{-\frac{it}{2}}K] = \mathrm{Tr}[K^*K],$$

and

$$\|G_X(1+it)\|_2^2 = \mathrm{Tr}[K^*C^{-\frac{1-it}{2}}X^{\frac{1-it}{2}}X^{\frac{1+it}{2}}C^{-\frac{1+it}{2}}K] = \mathrm{Tr}[K^*C^{-\frac{1-it}{2}}XC^{-\frac{1+it}{2}}K].$$

We then have

$$\tau\|G_A(1+it)\|_2^2 + (1-\tau)\|G_B(1+it)\|_2^2 = \mathrm{Tr}[K^*C^{-\frac{1-it}{2}}(\tau A + (1-\tau)B)C^{-\frac{1+it}{2}}K] = \mathrm{Tr}[K^*K].$$

Finally we obtain

$$\tau\|G_A(r)\|_2^2 + (1-\tau)\|G_B(r)\|_2^2 \le \mathrm{Tr}[K^*K].$$

D. Homogeneous Convex/Concave Functions

Lemma D.1. *Let $\mathcal{C}$ be a convex cone in some linear space, i.e $\mathcal{C} = conv(\mathcal{C})$ and $C = \lambda\mathcal{C}$ for any $\lambda > 0$. Let function $f : \mathcal{C} \to [0, +\infty)$ be positively homogeneous of order 1, i.e. $f(\lambda x) = \lambda f(x)$, for any $x \in \mathcal{C}$ and $\lambda > 0$. Then for any $s \in (0, 1)$, $f(x)$ is concave if and only if $f(x)^s$ is concave; for any $s \in (1, +\infty)$, $f(x)$ is convex if and only if $f(x)^s$ is convex.*

In general this lemma is proved via an argument of level sets. Here we provide a more direct proof.

Proof. One direction is trivial. If $f(x)$ is concave, then $f(x)^s$ is concave for $s \in (0, 1)$, since $(\cdot)^s$ is concave and monotone increasing. Conversely, if $f(x)^s$ is concave for some $s \in (0, 1)$, then $f(\tau x + (1-\tau)y)^s \geq \tau f(x)^s + (1-\tau)f(y)^s$, for any $x, y \in \mathcal{C}, \tau \in [0, 1]$. Now given any fixed $x, y \in \mathcal{C}, \tau \in [0, 1]$, we need to show that $\tau f(x) + (1-\tau)f(y) \leq f(\tau x + (1-\tau)y)$. If $f(x) = f(y) = 0$, then we are done. Otherwise, we may assume that $f(x) > 0$, and define $M = \tau f(x) + (1-\tau)(f(y) + \epsilon)$ for some $\epsilon > 0$(this ϵ is not necessary if $f(y) > 0$). We then have

$$
\begin{aligned}
f(\tau x + (1-\tau)y) &= f\left(\frac{\tau f(x)}{M}\frac{Mx}{f(x)} + \frac{(1-\tau)(f(y)+\epsilon)}{M}\frac{My}{f(y)+\epsilon}\right)^{s\cdot\frac{1}{s}} \\
&\geq \left(\frac{\tau f(x)}{M} f\left(\frac{Mx}{f(x)}\right)^s + \frac{(1-\tau)(f(y)+\epsilon)}{M} f\left(\frac{My}{f(y)+\epsilon}\right)^s\right)^{\frac{1}{s}} \\
&= \left(\frac{\tau f(x)}{M}M^s + \frac{(1-\tau)(f(y)+\epsilon)^{1-s}f(y)^s}{M}M^s\right)^{\frac{1}{s}} \\
&= M\left(\frac{\tau f(x) + (1-\tau)(f(y)+\epsilon)^{1-s}f(y)^s}{\tau f(x) + (1-\tau)(f(y)+\epsilon)}\right)^{\frac{1}{s}}.
\end{aligned}
$$

We then take $\epsilon \to 0$ to obtain $f(\tau x + (1-\tau)y) \geq \tau f(x) + (1-\tau)f(y)$. Therefore $f(x)$ is concave. The convexity part can be proved similarly. □

Acknowledgment

The research was in part supported by the NSF Grant DMS-1613861. The author would like to thank Thomas Y. Hou for his wholehearted mentoring and supporting.

References

References

[1] E. H. Lieb, Convex trace functions and the wigner-yanase-dyson conjecture, Advances in Mathematics 11 (3) (1973) 267 – 288.

[2] E. P. Wigner, M. M. Yanase, Information contents of distributions, Proceedings of the National Academy of Sciences of the United States of America 49 (6) (1963) 910.

[3] E. H. Lieb, M. B. Ruskai, Proof of the strong subadditivity of quantummechanical entropy, Journal of Mathematical Physics 14 (12) (1973) 1938–1941.

[4] D. Huang, A generalized Lieb's theorem and its applications to spectrum estimates for a sum of random matrices, arXiv e-prints (2018) arXiv:1808.05550`arXiv:1808.05550`.

[5] J. A. Tropp, User-friendly tail bounds for sums of random matrices, Foundations of Computational Mathematics 12 (4) (2012) 389–434.

[6] J. A. Tropp, An introduction to matrix concentration inequalities, Foundations and Trends in Machine Learning 8 (1-2) (2015) 1–230.

[7] T. Ando, Concavity of certain maps on positive definite matrices and applications to hadamard products, Linear Algebra and its Applications 26 (1979) 203 – 241.

[8] E. A. Carlen, Trace inequalities and quantum entropy: an introductory course, Entropy and the quantum 529 (2010) 73–140.

[9] I. Nikoufar, A. Ebadian, M. E. Gordji, The simplest proof of lieb concavity theorem, Advances in Mathematics 248 (2013) 531 – 533.

[10] H. Epstein, Remarks on two theorems of e. lieb, Communications in Mathematical Physics 31 (4) (1973) 317–325.

[11] A. Uhlmann, Relative entropy and the wigner-yanase-dyson-lieb concavity in an interpolation theory, Communications in Mathematical Physics 54 (1) (1977) 21–32.

[12] H. Kosaki, Interpolation theory and the wigner-yanase-dyson-lieb concavity, Communications in Mathematical Physics 87 (3) (1982) 315–329.

[13] J. Hadamard, et al., Théorème sur les séries entières, Acta Mathematica 22 (1899) 55–63.

[14] M. Riesz, et al., Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires, Acta mathematica 49 (3-4) (1926) 465–497.

[15] E. M. Stein, Interpolation of linear operators, Transactions of the American Mathematical Society 83 (2) (1956) 482–492.

[16] J. Peetre, A theory of interpolation spaces, Notes, Universidade de Brasilia.

[17] D. Sutter, M. Berta, M. Tomamichel, Multivariate trace inequalities, Communications in Mathematical Physics 352 (1) (2017) 37–58.

[18] K. Löwner, Über monotone matrixfunktionen, Mathematische Zeitschrift 38 (1) (1934) 177–216.

[19] E. Heinz, Beiträge zur störungstheorie der spektralzerleung, Mathematische Annalen 123 (1) (1951) 415–438.

[20] F. Kraus, Über konvexe matrixfunktionen, Mathematische Zeitschrift 41 (1) (1936) 18–42.

[21] A. Vershynina, E. A. Carlen, E. H. Lieb, Matrix and Operator Trace Inequalities, Scholarpedia 8 (4) (2013) 30919, revision #132324.

[22] E. A. Carlen, E. H. Lieb, A minkowski type trace inequality and strong subadditivity of quantum entropy ii: Convexity and concavity, Letters in Mathematical Physics 83 (2) (2008) 107–126.

[23] I. I. Hirschman, A convexity theorem for certain groups of transformations, Journal d'Analyse Mathématique 2 (2) (1952) 209–218.

[24] H. Araki, On an inequality of lieb and thirring, Letters in Mathematical Physics 19 (2) (1990) 167–170.

[25] A. Ebadian, I. Nikoufar, M. E. Gordji, Perspectives of matrix convex functions, Proceedings of the National Academy of Sciences 108 (18) (2011) 7313–7314.

[26] D. Petz, A survey of certain trace inequalities, Banach Center Publications 30 (1) (1994) 287–298.

[27] G. Lindblad, Entropy, information and quantum measurements, Communications in Mathematical Physics 33 (4) (1973) 305–322.

[28] R. Bapat, Mixed discriminants of positive semidefinite matrices, Linear Algebra and its Applications 126 (1989) 107–124.

[29] A. A. Panov, On some properties of mixed discriminants, Sbornik: Mathematics 56 (2) (1987) 279–293.

[30] A. Aleksandrov, On the theory of mixed volumes of convex bodies, iv, mixed discriminants and mixed volumes, Mat. Sb.(NS) 3 (1) (1938) 938.

[31] P. Li, The Alexandrov-Fenchel type inequalities, revisited, arXiv e-prints (2017) arXiv:1710.00520`arXiv:1710.00520`.

[32] R. Bishop, S. Goldberg, Tensor Analysis on Manifolds, Dover Books on Mathematics, Dover Publications, 1980.

[33] J. J. Rotman, Advanced modern algebra; 2nd ed., Graduate studies in mathematics, American Mathematical Society, Providence, RI, 2010.