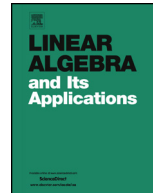Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Monotonicity versions of Epstein's Concavity Theorem and related inequalities ✩,✩✩

Eric A. Carlen [a], Haonan Zhang [b],*

[a] *Department of Mathematics, Hill Center, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA*
[b] *Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria*

A R T I C L E   I N F O

A B S T R A C T

Many trace inequalities can be expressed either as concavity/convexity theorems or as monotonicity theorems. A classic example is the joint convexity of the quantum relative entropy which is equivalent to the Data Processing Inequality. The latter says that quantum operations can never increase the relative entropy. The monotonicity versions often have many advantages, and often have direct physical application, as in the example just mentioned. Moreover, the monotonicity results are often valid for a larger class of maps than, say, quantum operations (which are completely positive). In this paper we prove several new monotonicity results, the first of which is a monotonicity theorem that has as a simple corollary a celebrated concavity theorem of Epstein. Our starting points are the monotonicity versions of the Lieb Concavity and the Lieb Convexity Theorems. We also give two new proofs of these in their general forms using interpolation. We then prove our new monotonicity theorems by several duality arguments.
© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Many important concavity and convexity theorems for matrix trace functionals have formulations as monotonicity theorems under some appropriate class of positive maps. In the following, $M_m(\mathbb{C})$ denotes the $m \times m$ complex matrices, $M_m^+(\mathbb{C})$, denotes the closed cone in $M_m(\mathbb{C})$ consisting of positive semidefinite matrices, and $M_m^{++}(\mathbb{C})$, denotes the open cone in $M_m(\mathbb{C})$ consisting of positive definite matrices. A linear map $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ is positive in case $\Phi(X) \in M_m^+(\mathbb{C})$ whenever $X \in M_n^+(\mathbb{C})$. Some more refined notions of positivity will be central here, and these are recalled in the next section. To keep the introduction brief, we assume for now that the notion of complete positivity is familiar to the reader. For each n, equip $M_n(\mathbb{C})$ with the Hilbert–Schmidt inner product, making it a Hilbert space. With respect to this Hilbert space structure, every linear $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ has an adjoint, denoted $\Phi^\dagger$, with $\Phi^\dagger : M_m(\mathbb{C}) \to M_n(\mathbb{C})$. A map $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ is *unital* in case $\Phi(\mathbb{1}) = \mathbb{1}$, and it is *trace preserving* in case $\mathrm{Tr}[\Phi(X)] = \mathrm{Tr}[X]$ for all $X \in M_n(\mathbb{C})$. It is easy to see that Φ is unital if and only if $\Phi^\dagger$ is trace preserving. Completely positive trace preserving maps are known as *quantum operations*. Any sort of physically possible manipulation of a quantum state is described by a quantum operation, hence highlighting the importance of this class of positive maps.

In a fundamental 1973 paper [20], Lieb proved a number of concavity and convexity theorems for trace functionals of matrices. The first two of these were

1.1 Theorem *(Lieb Concavity Theorem). For $s, t \geq 0$ and $s + t \leq 1$, and any fixed $K \in M_n(\mathbb{C})$, the function*

$$(X, Y) \mapsto \mathrm{Tr}[K^* Y^s K X^t] \tag{1.1}$$

is jointly concave on $M_n^+(\mathbb{C}) \times M_n^+(\mathbb{C})$.

1.2 Theorem *(Lieb Convexity Theorem). For all $s, t \geq 0$ and $s + t \leq 1$,*

$$(X, Y, K) \mapsto \mathrm{Tr}[K^* Y^{-s} K X^{-t}] \tag{1.2}$$

is jointly convex on $M_n^{++}(\mathbb{C}) \times M_n^{++}(\mathbb{C}) \times M_n(\mathbb{C})$.

In both cases, by a simple argument originally pointed out by Araki [2], the general case $s + t \leq 1$ follows from the special case $s + t = 1$; see [9, Section 2] for more information. Lindblad [21] took the further special case in which $K = \mathbb{1}$ and X and Y are both non-degenerate density matrices; i.e., elements of $M_n^{++}(\mathbb{C})$ with unit trace, to deduce that

$$D(X||Y) := \mathrm{Tr}[X(\log X - \log Y)] = \lim_{t \uparrow 1} \frac{1}{1-t}\left(1 - \mathrm{Tr}[Y^{1-t} X^t]\right) \tag{1.3}$$

is *jointly convex* in X and Y. The quantity $D(X||Y)$ is known as the *Umegaki quantum relative entropy* of X with respect to Y. Lindblad in 1975 [22] then went on to prove the monotonicity version of this result:

1.3 Theorem *(Data Processing Inequality). For all completely positive trace preserving maps $\Phi^\dagger : M_m(\mathbb{C}) \to M_n(\mathbb{C})$, and all density matrices $X, Y \in M_m^{++}(\mathbb{C})$,*

$$D(\Phi^\dagger(X)||\Phi^\dagger(Y)) \leq D(X||Y) . \tag{1.4}$$

This inequality, which roughly says that any operation one may perform on two quantum states can only make them harder to distinguish, is one of the cornerstones of quantum information theory.

It is correct to refer to the Data Processing Inequality as a *monotonicity version* of the joint convexity of the quantum relative entropy because it is easy to prove one from the other. On the one hand, Lindblad derived (1.4) from the joint convexity of $D(X||Y)$ using the structure theory for completely positive maps provided by the *Stinespring Dilation Theorem* [25]. On the other hand, one recovers the joint convexity statement by considering a very particular quantum operation, namely the partial trace. See [9] for more information, though we provide examples of this below.

Already, this example shows the merit in recasting a concavity or convexity theorem as a monotonicity theorem – *often the monotonicity has a significant physical interpretation.* There is however, another advantage of the monotonicity formulations: They often hold for a wider class of positive maps extending beyond the class of completely positive maps, and these more general results may no longer be deduced from convexity or concavity statements since the Stinespring Dilation Theorem is then not available.

The first result of this type is due to Uhlmann [26, Proposition 17] in 1977 who proved a monotonicity formulation of the Lieb Concavity Theorem. The theorem refers to the class of positive maps $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ such that for all $K \in M_n(\mathbb{C})$,

$$\Phi(K^*K) \geq \Phi(K)^*\Phi(K) . \tag{1.5}$$

Terminology is not completely standardized, but in this paper we refer to this class of maps as the class of *Schwarz maps.* Evidently such a map is positive, and as we recall in the next section, the class of Schwarz maps strictly includes the class of completely positive maps.

1.4 Theorem. *For all $0 \leq p \leq 1$, all $m, n \in \mathbb{N}$, all $A, B \in M_m^+(\mathbb{C})$ and all $K \in M_n(\mathbb{C})$, and all Schwarz maps $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$,*

$$\mathrm{Tr}[\Phi(K)^*A^p\Phi(K)B^{1-p}] \leq \mathrm{Tr}[K^*\Phi^\dagger(A)^pK\Phi^\dagger(B)^{1-p}] . \tag{1.6}$$

Assume furthermore that Φ is unital. Again, taking $K = \mathbb{1}$ and differentiation in p at $p = 1$, one obtains (1.4), but now under the more general conditions that Φ is a

unital Schwarz map. This generalization of the Data Processing Inequality was made by Uhlmann in 1977 [26], and for 40 years it remained the state of the art until Müller-Hermes and Reeb [23] proved the ultimate extension, namely that (1.4) is valid for all unital positive maps Φ.

Naturally, one may ask if (1.6) holds for all positive maps Φ. It turns out that the class of Schwarz maps is optimal, and therefore the inequality (1.6) characterizes Schwarz maps.

1.5 Theorem. *If a positive map* $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *satisfies* (1.6) *for all* $0 \leq p \leq 1$, *all* $A, B \in M_m^+(\mathbb{C})$ *and all* $K \in M_n(\mathbb{C})$, *then* Φ *is a Schwarz map. Thus, satisfaction of the inequality* (1.6) *characterizes the class of Schwarz maps.*

Proof. Note that at the endpoint case, say $p = 0$, (1.6) becomes

$$\mathrm{Tr}[\Phi(K)^*\Phi(K)B] \leq \mathrm{Tr}[K^*K\Phi^\dagger(B)] = \mathrm{Tr}[\Phi(K^*K)B] . \tag{1.7}$$

Since B is any element in $M_m^+(\mathbb{C})$, this implies the operator inequality (1.5), showing that the class of Schwarz maps is the broadest class of positive linear maps for which Theorem 1.4 is valid. $\square$

There is also a monotonicity extension of the Lieb Convexity Theorem, Theorem 1.2:

1.6 Theorem. *For all* $0 \leq t \leq 1$, *all* $m, n \in \mathbb{N}$, *all* $X, Y \in M_m^{++}(\mathbb{C})$, *all* $K \in M_m(\mathbb{C})$ *and all Schwarz maps* $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *such that* $\Phi^\dagger(\mathbb{1}) \in M_n^{++}(\mathbb{C})$,

$$\mathrm{Tr}[\Phi^\dagger(K^*)\Phi^\dagger(Y)^{t-1}\Phi^\dagger(K)\Phi^\dagger(X)^{-t}] \leq \mathrm{Tr}[K^*Y^{t-1}KX^{-t}] . \tag{1.8}$$

For unital completely positive maps, this follows from the Lieb Convexity Theorem, Theorem 1.2, and the Stinespring Dilation Theorem [25], as explained in [9, Section 3]. However, the first explicit statement appeared only in 1996, when Petz proved it directly for unital 2-positive maps, [24], a class that includes unital completely positive maps, but is strictly included in the class of unital Schwarz maps as explained below. By the very simple Lemma 1 of [18], Theorem 1.4 is equivalent to Theorem 1.6, and this together with Uhlmann's proof of Theorem 1.4 proves Theorem 1.6 with the additional condition that Φ is unital. Lieb [20] proved the equivalence of Theorem 1.1 and Theorem 1.2, but the proof of equivalence is simpler in the monotonicity formulation. Theorem 1.6 has a significant physical interpretation in terms of *monotone metrics*; see [9, Section 8] for more discussion.

Again, the class of Schwarz maps is optimal, and therefore the inequality (1.8) characterizes Schwarz maps:

1.7 Theorem. *If a positive map* $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ *satisfies* (1.8) *for all* $0 \leq t \leq 1$, *all* $X, Y \in M_m^+(\mathbb{C})$ *and all* $K \in M_n(\mathbb{C})$, *then* Φ *is a Schwarz map. Thus, satisfaction of the inequality* (1.8) *characterizes the class of Schwarz maps.*

Proof. In the endpoint $t = 0$ case, (1.8) becomes

$$\text{Tr}\left[\Phi^\dagger(K^*)\Phi^\dagger(Y)^{-1}\Phi^\dagger(K)\right] \leq \text{Tr}[K^*Y^{-1}K] . \tag{1.9}$$

By a recent result [11, Theorem 4], the inequality (1.9) characterizes Schwarz maps. $\square$

1.8 Remark. That Φ is Schwarz is equivalent to (1.9) was proved recently in [11]. In a preliminary version of this paper, we also obtained this characterization, in the class of unital maps, using results in an earlier version of [11] and the work of Hiai and Petz [18]. Although the argument made use of an additional hypothesis, namely unitality, the argument may still be of interest, and so we sketch it here.

The method of Hiai and Petz in [18] uses the integral representation for operator monotone increasing functions f:

$$f(x) := \beta + \gamma x + \int\limits_{(0,\infty)} \frac{x}{\lambda + x}(1 + \lambda)\mathrm{d}\mu(\lambda) ,$$

where μ is a finite Borel measure on $(0, \infty)$. The key observation underlying our characterization of unital Schwarz maps is that in the integral representation of

$$x^r = \int\limits_{(0,\infty)} \frac{x}{\lambda + x}(1 + \lambda)\mathrm{d}\mu(\lambda) \quad \text{where} \quad \mathrm{d}\mu(\lambda) = \frac{\sin(\pi r)}{\pi}\frac{\lambda^{r-1}}{1 + \lambda}\mathrm{d}\lambda ,$$

with $0 < r < 1$, one does not have the affine term, i.e. $\beta = \gamma = 0$. In the argument of Hiai and Petz, these terms were treated using the Schwarz inequality. The remaining integral term only required a *tracial inequality* equivalent to (1.9) when Φ is unital. This tracial inequality was shown in an earlier version [11] to be valid whenever Φ is a unital Schwarz map. On the other hand, we observed that when the tracial inequality holds, the method of Hiai and Petz yields (1.6) without assuming that Φ is a Schwarz map since the affine terms are not present. But then by Theorem 1.5, Φ is a Schwarz map.

A final advantage of the monotonicity formulations is that they are often easier to directly prove than the corresponding convexity or concavity theorems, which one then readily recovers with a simple partial trace argument as we demonstrate below. Indeed, a very simple proof of Theorems 1.4 and 1.6 was recently given in [11], building on previous work of Hiai and Petz [18], extending their argument to the wider generality of Schwarz maps. Later in the paper, we give a new proof of Theorem 1.4 that is modeled on Lieb's original proof of Theorem 1.1 by interpolation. The interpolation proof of the monotonicity result, which is different from Uhlmann's interpolation argument [26], seems especially simple to us. We also use interpolation to give another proof of Theorem 1.6.

The main theme of this paper is to obtain new monotonicity theorems from Theorems 1.4 and 1.6, among others, by a duality technique involving Legendre transforms

or variations upon them. This is a technique that both authors have successfully applied to prove new convexity and concavity results. One of us recently used this approach to prove a monotonicity version of Lieb's theorem [20, Theorem 6] stating that for all self-adjoint $H \in M_n(\mathbb{C})$,

$$X \mapsto \mathrm{Tr}[\exp(H + \log X)] \tag{1.10}$$

is concave on $M_n^{++}(\mathbb{C})$; see [8].

Our first new result is a monotonicity version of a theorem of Epstein [15]:

1.9 Theorem. *Let $0 < p \leq 1$, and let $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ be a unital Schwarz map. Then for any $A \in M_n^{++}(\mathbb{C})$ and any $B \in M_m(\mathbb{C})$,*

$$\mathrm{Tr}[(\Phi(B)^* A^p \Phi(B))^{1/p}] \leq \mathrm{Tr}[(B^* \Phi^\dagger(A)^p B)^{1/p}] . \tag{1.11}$$

Below we give two proofs of this theorem, each of which lends itself to different generalizations. First consider the special case in which $n = 2m$ and

$$\Phi(X) = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \quad \text{so that} \quad \Phi^\dagger\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \right) = A + D . \tag{1.12}$$

Clearly Φ is unital completely positive and hence Schwarz, as explained below. Now consider $A_1, A_2 \in M_m^+(\mathbb{C})$, $B \in M_m(\mathbb{C})$ and define $A := \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$. Then

$$\begin{aligned} \mathrm{Tr}[(\Phi(B)^* A^p \Phi(B))^{1/p}] &= \mathrm{Tr} \begin{bmatrix} (B^* A_1^p B)^{1/p} & 0 \\ 0 & (B^* A_2^p B)^{1/p} \end{bmatrix} \\ &= \mathrm{Tr}[(B^* A_1^p B)^{1/p}] + \mathrm{Tr}[(B^* A_2^p B)^{1/p}] , \end{aligned}$$

and

$$\mathrm{Tr}[(B^* \Phi^\dagger(A)^p B)^{1/p}] = \mathrm{Tr}[(B^* (A_1 + A_2)^p B)^{1/p}] .$$

Thus, by homogeneity,

$$\frac{1}{2} \mathrm{Tr}[(B^* A_1^p B)^{1/p}] + \frac{1}{2} \mathrm{Tr}[(B^* A_2^p B)^{1/p}] \leq \mathrm{Tr} \left[\left(B^* \left(\frac{A_1 + A_2}{2} \right)^p B \right)^{1/p} \right] .$$

Together with the obvious continuity, this proves:

1.10 Corollary *(Epstein's Theorem [15]). For any $B \in M_n(\mathbb{C})$ and all $0 \leq p \leq 1$, the map*

$$A \mapsto \mathrm{Tr}[(B^*A^pB)^{1/p}] \tag{1.13}$$

is concave on $M_n^+(\mathbb{C})$.

1.11 Remark. In the endpoint case $p = 1$, (1.11) reduces to $\mathrm{Tr}[\Phi(B)\Phi(B)^*A] \leq \mathrm{Tr}[BB^*\Phi^\dagger(A)]$, and again since $A \in M_n^+$ is arbitrary, Φ satisfies (1.5). Hence the validity of (1.11) for $0 < p \leq 1$ is characteristic of the class of Schwarz maps, which is the broadest class of unital positive linear maps for which Theorem 1.9 is valid.

Acknowledgment

2. Positivity and duality

In this section, we recall some different notions of positivity, and some duality lemmas we shall use in later sections.

Consider a linear map $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$. We identify $M_n(\mathbb{C}) \otimes M_2(\mathbb{C})$ with the 2×2 block matrices $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ and then

$$\Phi \otimes \mathrm{Id} \left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \right) = \begin{bmatrix} \Phi(A) & \Phi(B) \\ \Phi(C) & \Phi(D) \end{bmatrix} . \tag{2.1}$$

The map Φ is defined to be 2-*positive* in case the matrix on the right in (2.1) belongs to $M_{2m}^+(\mathbb{C})$ whenever $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in M_{2n}^+(\mathbb{C})$. The notion of k-*positivity* is defined in the analogous way for all positive integers $k \geq 2$, and the map Φ is *completely positive* in case it is k-positive for all k. By a theorem of Choi [13], which generalizes an earlier result of Kadison [19], whenever Φ is a unital 2-positive map, for all $X \in M_n(\mathbb{C})$,

$$\Phi(X^*X) \geq \Phi(X)^*\Phi(X) . \tag{2.2}$$

A *Schwarz map* is a linear map satisfying (2.2). Evidently, every Schwarz map is positive. Choi also showed [14] that while every 2-positive map is a Schwarz map, the converse is not true: There are Schwarz maps that are not 2-positive. The notion of a *generalized Schwarz map* may be found in [11]: The map $\Phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$, not necessarily unital, is a *generalized Schwarz map* in case

$$\begin{bmatrix} \Phi(\mathbb{1}) & \Phi(X) \\ \Phi(X)^* & \Phi(X^*X) \end{bmatrix} \geq 0 . \tag{2.3}$$

Using the well-known fact that for $A \in M_n^+(\mathbb{C})$, with A^+ denoting its Moore–Penrose generalized inverse,

$$\begin{bmatrix} A & Z \\ Z^* & B \end{bmatrix} \geq 0 \iff B \geq Z^* A^+ Z , \tag{2.4}$$

one sees, as noted in [11] that every 2-positive map is a generalized Schwarz map, and that a unital map is Schwarz if and only if it is generalized Schwarz. Thus we have the following hierarchy of classes of positive maps:

$$\text{Completely Positive} \subset 2 - \text{Positive} \subset \text{Generalized Schwarz} \subset \text{Positive} . \tag{2.5}$$

We shall also use some duality lemmas. The first of these depends on the reverse Hölder inequality for traces that we briefly recall for completeness; see [12] and [17, Appendix] for a more general result without the positivity condition that meets our needs here. Let $X, Y \in M_n^{++}(\mathbb{C})$. We take the trace of XY using a basis in which $X_{i,j} = \delta_{i,j} \lambda_j$ and $Y_{j,i} = \sum_k U_{j,k} \mu_k \overline{U_{i,k}}$. Then

$$\text{Tr}[XY] = \sum_{j,k} \lambda_j |U_{j,k}|^2 \mu_k .$$

The matrix S with $S_{j,k} := |U_{j,k}|^2$ is doubly stochastic, and hence by Birkhoff's Theorem [6], S is a convex combination of permutation matrices $P^{(\pi)}$, $\pi \in \mathcal{S}_n$ with $\mathcal{S}_n$ the symmetric group on n letters. Hence $\text{Tr}[XY]$ is a convex combination of sums of the form $\sum_{j=1}^n \lambda_j \mu_{\pi(j)}$, $\pi \in \mathcal{S}_n$. To each of these we may apply the classical reverse Hölder inequality to conclude that for $0 < r < 1$ or $-\infty < r < 0$,

$$\sum_{j=1}^n \lambda_j \mu_{\pi(j)} \geq \left(\sum_{j=1}^n \lambda_j^r \right)^{1/r} \left(\sum_{j=1}^n \mu_{\pi(j)}^{r/(r-1)} \right)^{(r-1)/r} = (\text{Tr}[X^r])^{1/r} \left(\text{Tr}[Y^{r/(r-1)}] \right)^{(r-1)/r} .$$

We conclude that for $X, Y \in M_n^{++}(\mathbb{C})$: If $0 < r < 1$ or $-\infty < r < 0$, then

$$\text{Tr}[XY] \geq (\text{Tr}[X^r])^{1/r} \left(\text{Tr}[Y^{r/(r-1)}] \right)^{(r-1)/r} . \tag{2.6}$$

2.1 Lemma. *Let $X \in M_n^+(\mathbb{C})$. Then for $1 < r < \infty$,*

$$(\text{Tr}[X^r])^{1/r} = \max \left\{ \text{Tr}[XY] \ : \ Y \in M_n^+(\mathbb{C}) , \quad \text{Tr}[Y^{r/(r-1)}] = 1 \right\} , \tag{2.7}$$

while for $0 < r < 1$ or $-\infty < r < 0$, and $X \in M_n^{++}(\mathbb{C})$,

$$(\text{Tr}[X^r])^{1/r} = \min \left\{ \text{Tr}[XY] \ : \ Y \in M_n^{++}(\mathbb{C}) , \quad \text{Tr}[Y^{r/(r-1)}] = 1 \right\} . \tag{2.8}$$

Proof. First, consider $1 < r < \infty$. For $Y := (\text{Tr}[X^r])^{(1-r)/r} X^{r-1}$, we have

$$\text{Tr}[XY] = (\text{Tr}[X^r])^{1/r} \qquad \text{and} \qquad \text{Tr}[Y^{r/(r-1)}] = 1 .$$

Then by Hölder's inequality for traces [7, Section 7.1],

$$\text{Tr}[XY] \leq (\text{Tr}[X^r])^{1/r} \left(\text{Tr}[Y^{r/(r-1)}] \right)^{(r-1)/r} ,$$

and this proves (2.7).

Next, suppose $X \in M_n^{++}(\mathbb{C})$, and $0 < r < 1$ or $-\infty < r < 0$. Again with $Y := (\text{Tr}[X^r])^{(1-r)/r} X^{r-1}$, we have

$$\text{Tr}[XY] = (\text{Tr}[X^r])^{1/r} \qquad \text{and} \qquad \text{Tr}[Y^{r/(r-1)}] = 1 . \qquad (2.9)$$

Now the reverse Hölder inequality (2.6) together with the example in (2.9) proves (2.8). $\square$

3. Proofs of Theorem 1.9 and related inequalities

3.1. Proofs of Theorem 1.9

First proof of Theorem 1.9. When $p = 1$, the inequality follows directly from the assumption that Φ is Schwarz. Now let $0 < p < 1$ and put $r := 1/p > 1$, so that

$$(\text{Tr}[(\Phi(B)^* A^p \Phi(B))^{1/p}])^p = (\text{Tr}[(\Phi(B)^* A^p \Phi(B))^r])^{1/r} .$$

Then by (2.7)

$$
\begin{aligned}
&(\text{Tr}[(\Phi(B)^* A^p \Phi(B))^r])^{1/r} \\
&= \max \left\{ \text{Tr}[\Phi(B)^* A^p \Phi(B) Y] \ : \ Y \in M_n^+(\mathbb{C}) , \quad \text{Tr}[Y^{r/(r-1)}] = 1 \right\} \\
&= \max \left\{ \text{Tr}[\Phi(B)^* A^p \Phi(B) Y] \ : \ Y \in M_n^+(\mathbb{C}) , \quad \text{Tr}[Y^{1/(1-p)}] = 1 \right\} \\
&= \max \left\{ \text{Tr}[\Phi(B)^* A^p \Phi(B) Y^{1-p}] \ : \ Y \in M_n^+(\mathbb{C}) , \quad \text{Tr}[Y] = 1 \right\} \qquad (3.1)
\end{aligned}
$$

By Theorem 1.4, $\text{Tr}[\Phi(B)^* A^p \Phi(B) Y^{1-p}] \leq \text{Tr}[(B^* \Phi^\dagger (A)^p B \Phi^\dagger (Y)^{1-p}]$. Therefore,

$$
\begin{aligned}
&(\text{Tr}[(\Phi(B)^* A^p \Phi(B))^{1/p}])^p \\
&\leq \max \left\{ \text{Tr} \left[B^* \Phi^\dagger (A)^p B \Phi^\dagger (Y)^{1-p} \right] \ : \ Y \in M_n^+(\mathbb{C}) , \quad \text{Tr}[Y] = 1 \right\} \\
&\leq \max \left\{ \text{Tr} \left[B^* \Phi^\dagger (A)^p B Z^{1-p} \right] \ : \ Z \in M_m^+(\mathbb{C}) , \quad \text{Tr}[Z] = 1 \right\} \\
&= \max \left\{ \text{Tr} \left[B^* \Phi^\dagger (A)^p B Z \right] \ : \ Z \in M_m^+(\mathbb{C}) , \quad \text{Tr}[Z^{1/(1-p)}] = 1 \right\} \\
&= (\text{Tr}[(B^* \Phi^\dagger (A)^p B)^{1/p}])^p , \qquad (3.2)
\end{aligned}
$$

where in the second inequality we have used the fact that $\Phi^\dagger$ is positive trace preserving and the definition of max, and in the last line we have used (2.7) once again. Now take the p-th root of both sides. $\square$

We now give a second proof of Theorem 1.9 that uses a different but related variational formula. The starting point is the generalized Hölder's inequality for traces. For $0 < p < \infty$, and $X \in M_n(\mathbb{C})$, define $\|X\|_p := (\mathrm{Tr}[(X^*X)^{p/2}])^{1/p}$. For $p \geq 1$, this is the Schatten p-norm, and otherwise it is a quasi-norm. Then one may prove using, for example, the theory of majorization, that for $X, Y \in M_n(\mathbb{C})$ and for $p_i > 0, i = 0, 1, 2$ such that $1/p_0 = 1/p_1 + 1/p_2$ [4, Exercise IV.2.7]

$$\|XY\|_{p_0} \leq \|X\|_{p_1}\|Y\|_{p_2} \ . \tag{3.3}$$

As proved in [27], this gives us more variational formulae that are very useful in proving the concavity/convexity theorems. In fact, for any $X, Y, Z \in M_n(\mathbb{C})$ with Y invertible, (3.3) yields (here we use the notation $|A| := (A^*A)^{1/2}$)

$$\mathrm{Tr}[|XZ|^{p_0}] \leq \|XY\|_{p_1}^{p_0}\|Y^{-1}Z\|_{p_2}^{p_0} \leq \frac{p_0}{p_1}\mathrm{Tr}\left[|XY|^{p_1}\right] + \frac{p_0}{p_2}\mathrm{Tr}\left[|Y^{-1}Z|^{p_2}\right] \ ,$$

or equivalently

$$\mathrm{Tr}\left[|XY|^{p_1}\right] \geq \frac{p_1}{p_0}\mathrm{Tr}\left[|XZ|^{p_0}\right] - \frac{p_1}{p_2}\mathrm{Tr}\left[|Y^{-1}Z|^{p_2}\right] \ .$$

One can find $Z \in M_n(\mathbb{C})$ to saturate the inequality so that [27, Theorem 3.3]

$$\begin{aligned}
\mathrm{Tr}\left[|XY|^{p_1}\right] &= \max_{Z \in M_n(\mathbb{C})} \left\{\frac{p_1}{p_0}\mathrm{Tr}\left[|XZ|^{p_0}\right] - \frac{p_1}{p_2}\mathrm{Tr}\left[|Y^{-1}Z|^{p_2}\right]\right\} \\
&= \max_{H \in M_n^+(\mathbb{C})} \left\{\frac{p_1}{p_0}\mathrm{Tr}\left[(XHX^*)^{p_0/2}\right] - \frac{p_1}{p_2}\mathrm{Tr}\left[(|Y^{-1}HY^{*-1})^{p_2/2}\right]\right\} \ .
\end{aligned}$$

Now for any $0 < p < 1$, take $(X, Y) = (B^*, A^{p/2})$ with $A \in M_n^{++}(\mathbb{C})$ and $(p_0, p_1, p_2) = (2, 2/p, 2/(1-p))$:

$$\mathrm{Tr}\left[(B^*A^pB)^{1/p}\right] = \max_{H \in M_n^+(\mathbb{C})} \left\{\frac{1}{p}\mathrm{Tr}[B^*HB] - \frac{1-p}{p}\mathrm{Tr}\left[\left(A^{-p/2}HA^{-p/2}\right)^{1/(1-p)}\right]\right\} \ . \tag{3.4}$$

Second proof of Theorem 1.9. We apply (3.4) in conjunction with the monotonicity of *sandwiched α-Rényi relative entropy* for $\alpha > 1$; see Müller-Hermes and Reeb [23, Theorem 2] and Beigi [3]. Recall that the monotonicity result of Müller-Hermes and Reeb [23, Theorem 2] states: For $\alpha > 1$,

$$\mathrm{Tr}\left[\left(\Lambda(\sigma)^{(1-\alpha)/(2\alpha)}\Lambda(\rho)\Lambda(\sigma)^{(1-\alpha)/(2\alpha)}\right)^\alpha\right] \leq \mathrm{Tr}\left[\left(\sigma^{(1-\alpha)/(2\alpha)}\rho\sigma^{(1-\alpha)/(2\alpha)}\right)^\alpha\right] \ , \tag{3.5}$$

for any positive trace preserving map $\Lambda : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ and any $\rho, \sigma \in M_n^+(\mathbb{C})$.

If Φ is unital Schwarz, then

$$\mathrm{Tr}\left[\left(\Phi(B)^*A^p\Phi(B)\right)^{1/p}\right]$$

$$= \max_{H \in M_n^+(\mathbb{C})} \left\{ \frac{1}{p}\mathrm{Tr}\left[\Phi(B)^*H\Phi(B)\right] - \frac{1-p}{p}\mathrm{Tr}\left[\left(A^{-p/2}HA^{-p/2}\right)^{1/(1-p)}\right]\right\}$$

$$\leq \max_{H \in M_n^+(\mathbb{C})} \left\{ \frac{1}{p}\mathrm{Tr}\left[\Phi(B)^*H\Phi(B)\right] - \frac{1-p}{p}\mathrm{Tr}\left[\left(\Phi^\dagger(A)^{-p/2}\Phi^\dagger(H)\Phi^\dagger(A)^{-p/2}\right)^{1/(1-p)}\right]\right\}$$

$$\leq \max_{H \in M_n^+(\mathbb{C})} \left\{ \frac{1}{p}\mathrm{Tr}\left[B^*\Phi^\dagger(H)B\right] - \frac{1-p}{p}\mathrm{Tr}\left[\left(\Phi^\dagger(A)^{-p/2}\Phi^\dagger(H)\Phi^\dagger(A)^{-p/2}\right)^{1/(1-p)}\right]\right\}$$

$$\leq \max_{H \in M_m^+(\mathbb{C})} \left\{ \frac{1}{p}\mathrm{Tr}\left[B^*HB\right] - \frac{1-p}{p}\mathrm{Tr}\left[\left(\Phi^\dagger(A)^{-p/2}H\Phi^\dagger(A)^{-p/2}\right)^{1/(1-p)}\right]\right\}$$

$$= \mathrm{Tr}\left[\left(B^*\Phi^\dagger(A)^pB\right)^{1/p}\right] \ ,$$

where we have used (3.4) in the two equalities, the monotonicity of sandwiched α-Rényi relative entropy (3.5) with $\alpha = 1/(1-p) > 1$ in the first inequality, the fact that Φ is Schwarz in the second inequality, and the definition of max in the third inequality. $\qquad\square$

3.2. Generalizations of Theorem 1.9

In Lieb Concavity Theorem, one does not require the exponents to be t and $1 - t$. Theorem 1.1 holds, as stated, when $0 < s, t$ and $s + t \leq 1$. There is a corresponding version of the monotonicity version, but for $s + t < 1$, we require Φ to be *semiunital*.

3.1 Definition. A linear map $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ is *semiunital* in case

$$\Phi^\dagger(\mathbb{1}) = \frac{n}{m}\mathbb{1} \ . \tag{3.6}$$

It is *sesquiunital* in case it is unital and semiunital.

For any unital map $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$, $\Phi^\dagger$ is trace preserving, so that $n = \mathrm{Tr}[\Phi^\dagger(\mathbb{1})]$ and hence if $\Phi^\dagger(\mathbb{1}) = c\mathbb{1}$ for some c, then necessarily $c = n/m$. As a fundamental example, consider the unital map Φ defined in (1.12). In this case $n = 2m$ so that $c := n/m = 2$ no matter how large m may be. Note that (1.12) is precisely the map that was used to recover Epstein's Concavity Theorem from Theorem 1.9. In the same way, we shall be able to recover the full statement of Lieb Concavity Theorem from the extension of Theorem 1.4 that we next state and prove:

3.2 Theorem. *Let* $0 < s, t$, $s + t \leq 1$, *and let* $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ *be a semiunital Schwarz map. Then for all* $X, Y \in M_n^+(\mathbb{C})$ *and all* $K \in M_m(\mathbb{C})$,

$$\mathrm{Tr}[\Phi(K)^* X^s \Phi(K) Y^t] \leq \left(\frac{n}{m}\right)^{1-(s+t)} \mathrm{Tr}[K^* \Phi^\dagger(X)^s K \Phi^\dagger(Y)^t] \,. \tag{3.7}$$

Proof. In case $s + t = 1$, the theorem is already proved under the weaker assumption that Φ is a Schwarz map. Therefore suppose that $t < 1 - s$ and define $0 < r < 1$ so that $t = r(1 - s)$. Define $Z = Y^r$. Then by the definitions and Theorem 1.4,

$$\mathrm{Tr}[\Phi(K)^* X^s \Phi(K) Y^t] = \mathrm{Tr}[\Phi(K)^* X^s \Phi(K) Z^{1-s}] \leq \mathrm{Tr}[K^* \Phi^\dagger(X)^s K \Phi^\dagger(Z)^{1-s}] \,.$$

Therefore,

$$\frac{m}{n} \mathrm{Tr}[\Phi(K)^* X^s \Phi(K) Y^t] \leq \mathrm{Tr}\left[K^* \left(\frac{m}{n}\Phi^\dagger(X)\right)^s K \left(\frac{m}{n}\Phi^\dagger(Y^r)\right)^{1-s}\right] \,.$$

Since $\frac{m}{n}\Phi^\dagger$ is unital, an operator Jensen type inequality for unital positive maps [13, Theorem 2.1] (see also [5, Proposition 2.7.1]) yields $\frac{m}{n}\Phi^\dagger(Y^r) \leq \left(\frac{m}{n}\Phi^\dagger(Y)\right)^r$. Then by the operator monotonicity of $A \mapsto A^{1-s}, 0 < s < 1$:

$$\left(\frac{m}{n}\Phi^\dagger(Y^r)\right)^{1-s} \leq \left(\frac{m}{n}\Phi^\dagger(Y)\right)^t \,.$$

This finishes the proof. $\square$

Now applying the method used in our first proof of Theorem 1.9, we prove:

3.3 Theorem. *Let $0 < s \leq 1$ and $s \leq r \leq 1$. Let Φ be any sesquiunital Schwarz map from $M_m(\mathbb{C})$ to $M_n(\mathbb{C})$ for $m, n \in \mathbb{N}$. Then for all $A \in M_n^+(\mathbb{C})$ and all $B \in M_m(\mathbb{C})$,*

$$\mathrm{Tr}[(\Phi(B)^* A^s \Phi(B))^{r/s}] \leq \left(\frac{n}{m}\right)^{1-r} \mathrm{Tr}[(B^* \Phi^\dagger(A)^s B)^{r/s}] \,. \tag{3.8}$$

Before proving the theorem, we first extract the concavity theorem that is its corollary. This concavity result was proved in [10, Theorem 1.1] and [16, Theorem 4.1].

3.4 Corollary. *For all $B \in M_m(\mathbb{C})$, $0 < s < 1$ and $0 < r \leq 1$, the map $A \mapsto \mathrm{Tr}[(B^* A^s B)^{r/s}]$ is concave on $M_m^+(\mathbb{C})$.*

Proof. First suppose that $s \leq r \leq 1$ so that we may apply Theorem 3.3. As in the discussion below the statement of Theorem 1.9, consider $A_1, A_2 \in M_m^+(\mathbb{C})$, $B \in M_m(\mathbb{C})$ and define $A := \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$. Define Φ as in (1.12), and note that Φ is Schwarz and sesquiunital. One computes

$$\mathrm{Tr}[(\Phi(B)^* A^s \Phi(B))^{r/s}] = \mathrm{Tr}[(B^* A_1^s B)^{r/s}] + \mathrm{Tr}[(B^* A_2^s B)^{r/s}] \,,$$

and

$$\text{Tr}[(B^*\Phi^\dagger(A)^s B)^{r/s}] = \text{Tr}[(B^*(A_1 + A_2)^s B)^{r/s}] .$$

In this case, $n = 2m$ so that $(n/m)^r = 2^r$. Therefore, by Theorem 3.3

$$\frac{1}{2}\left(\text{Tr}[(B^* A_1^s B)^{r/s}] + \text{Tr}[(B^* A_2^s B)^{r/s}]\right) = \frac{1}{2}\text{Tr}[(\Phi(B)^* A^s \Phi(B))^{r/s}]$$

$$\leq \left(\frac{1}{2}\right)^r \text{Tr}[(B^*\Phi^\dagger(A)^s B)^{r/s}]$$

$$= \text{Tr}\left[\left(B^*\left(\frac{A_1 + A_2}{2}\right)^s B\right)^{r/s}\right] .$$

As before, with the obvious continuity, midpoint concavity implies concavity.

Next, suppose that $0 < r < s$. (This part is elementary, and does not depend on Theorem 1.4.) Define $p := r/s$, and note that $0 < p < 1$. Then $(\text{Tr}[(B^* A^s B)^{r/s}])^{s/r} = (\text{Tr}[(B^* A^s B)^p])^{1/p} = \|B^* A^s B\|_p$. By (2.8),

$$\|B^* A^s B\|_p = \min\{\text{Tr}[B^* A^s BY] : Y \in M_n^{++}(\mathbb{C}) , \ \text{Tr}[Y^{p/(p-1)}] = 1\} ,$$

and then since $A \mapsto A^s$ is concave, so is $A \mapsto (\text{Tr}[(B^* A^s B)^{r/s}])^{s/r}$. Define $f(A) := (\text{Tr}[(B^* A^s B)^{r/s}])^{s/r}$ and $g(A) := \text{Tr}[(B^* A^s B)^{r/s}]$ so that $g(A) = (f(A))^p$. Since f is concave, and $h(x) := x^p$ is concave and monotone increasing on $[0, \infty)$, g is also concave. $\square$

3.5 Remark. Note that we have recovered a dimension independent concavity theorem from a monotonicity theorem that may at first appear to be dimension dependent. But of course, the dimensions m and n enter the inequality provided by the monotonicity theorem only through the ratio n/m, and this can be finite or constant in meaningful examples even as the dimensions tend to infinity.

The same argument also yields the full Lieb Concavity Theorem from its monotonicity version Theorem 3.2.

Proof of Theorem 3.3. We proceed exactly as in the first proof of Theorem 1.9. By continuity we may assume that $0 < s < 1$ and $s < r < 1$, so $0 < t < 1$ and $s + t < 1$ for $t := (r - s)/r$. By (2.7), we have

$$(\text{Tr}[(\Phi(B)^* A^s \Phi(B))^{r/s}])^{s/r} = \max\{\text{Tr}[\Phi(B)^* A^s \Phi(B) Y^{(r-s)/r}] : \text{Tr}[Y] = 1 \} .$$

By Theorem 3.2, for all $Y \in M_n^+(\mathbb{C})$ with $\text{Tr}[Y] = 1$,

$$\text{Tr}[\Phi(B)^* A^s \Phi(B) Y^{(r-s)/r}] \leq \left(\frac{n}{m}\right)^{s(1-r)/r} \text{Tr}[B^*\Phi^\dagger(A)^s B\Phi^\dagger(Y)^{(r-s)/r}] .$$

Therefore, since $\Phi^\dagger$ is trace preserving and using (2.7), we have

$$(\mathrm{Tr}[(\Phi(B)^*A^s\Phi(B))^{r/s}])^{s/r} \leq \left(\frac{n}{m}\right)^{s(1-r)/r} \max\{\mathrm{Tr}[B^*\Phi^\dagger(A)^sBZ^{(r-s)/r}] \ : \ \mathrm{Tr}[Z]=1\}$$

$$= \left(\frac{n}{m}\right)^{s(1-r)/r} \left(\mathrm{Tr}[(B^*\Phi^\dagger(A)^sB)^{r/s}]\right)^{s/r} \ .$$

Hence

$$(\mathrm{Tr}[(\Phi(B)^*A^s\Phi(B))^{r/s}])^{s/r} \leq \left(\frac{n}{m}\right)^{s(1-r)/r} \left(\mathrm{Tr}[(B^*\Phi^\dagger(A)^sB)^{r/s}]\right)^{s/r} \ ,$$

which is the same as (3.8). □

4. Duality, monotonicity and the Lieb Convexity Theorem

The same ideas that we have developed concerning the monotonicity version of the Lieb Concavity Theorem may be developed with respect to the Lieb Convexity Theorem, Theorem 1.2, as we now explain. We begin with the analog of Theorem 1.6. Afterwards, we will deduce the analog of Theorem 3.3, and deduce its convexity analog.

4.1 Theorem. *Let $0 < p < 1$, and let $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ be a unital Schwarz map. Then for any $A \in M_n^{++}(\mathbb{C})$ and any invertible $B \in M_m(\mathbb{C})$,*

$$\mathrm{Tr}[(B^*A^{-p}B)^{1/(2-p)}] \geq \mathrm{Tr}[(\Phi^\dagger(B)^*\Phi^\dagger(A)^{-p}\Phi^\dagger(B))^{1/(2-p)}] \ . \tag{4.1}$$

Proof. Define $r := 1/(2-p)$ and note that $1/2 < r < 1$. Then by (2.8),

$$(\mathrm{Tr}[(B^*A^{-p}B)^{1/(2-p)}])^{1/r}$$
$$= \left(\mathrm{Tr}[(B^*A^{-p}B)^r]\right)^{1/r}$$
$$= \min\left\{\mathrm{Tr}[B^*A^{-p}BY] \ : \ Y \in M_n^{++}(\mathbb{C}) \ , \quad \mathrm{Tr}[Y^{r/(r-1)}]=1\right\}$$
$$= \min\left\{\mathrm{Tr}[B^*A^{-p}BY^{1-1/r}] \ : \ Y \in M_n^{++}(\mathbb{C}) \ , \quad \mathrm{Tr}[Y]=1\right\}$$
$$= \min\left\{\mathrm{Tr}[B^*A^{-p}BY^{p-1}] \ : \ Y \in M_n^{++}(\mathbb{C}) \ , \quad \mathrm{Tr}[Y]=1\right\}$$
$$\geq \min\left\{\mathrm{Tr}[\Phi^\dagger(B)^*\Phi^\dagger(A)^{-p}\Phi^\dagger(B)\Phi^\dagger(Y)^{p-1}] \ : \ Y \in M_n^{++}(\mathbb{C}) \ , \quad \mathrm{Tr}[Y]=1\right\}$$
$$\geq \min\left\{\mathrm{Tr}[\Phi^\dagger(B)^*\Phi^\dagger(A)^{-p}\Phi^\dagger(B)Z^{p-1}] \ : \ Z \in M_m^{++}(\mathbb{C}) \ , \quad \mathrm{Tr}[Z]=1\right\}$$
$$= \min\left\{\mathrm{Tr}[\Phi^\dagger(B)^*\Phi^\dagger(A)^{-p}\Phi^\dagger(B)Z] \ : \ Z \in M_m^{++}(\mathbb{C}) \ , \quad \mathrm{Tr}[Z^{r/(r-1)}]=1\right\}$$
$$= (\mathrm{Tr}[(\Phi^\dagger(B)^*\Phi^\dagger(A)^{-p}\Phi^\dagger(B))^{1/(2-p)}])^{1/r} \ .$$

The first inequality is (1.8), and the second results from taking the minimum over a larger set. Now raise both sides to the r-th power. □

4.2 Remark. Taking the limit $p \uparrow 1$ in (4.1), we obtain that for all $A \in M_n^{++}(\mathbb{C})$, all invertible $B \in M_n(\mathbb{C})$ and all unital Schwarz maps $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$,

$$\mathrm{Tr}[B^* A^{-1} B] \geq \mathrm{Tr}[\Phi^\dagger(B)^* \Phi^\dagger(A)^{-1} \Phi^\dagger(B))] , \qquad (4.2)$$

a special case of a result that was recently proved in [11]. Actually, they proved that (4.2) holds whenever Φ is Schwarz (and vice versa), which we have already noted is an end-point case of Theorem 1.6. Thus, one way to prove this special case (without assuming Φ to be unital), is to take the interpolation proof of Theorem 1.6 given in Section 5, or Uhlmann's proof of Theorem 1.4 and [18, Lemma 1] to prove Theorem 1.6 (assuming Φ to be unital). However, the simple and direct proof of a more general result in [11] is by far the most direct route to (4.2).

Notice that if we choose $A = \mathbb{1}$ in (4.2), then

$$\|\Phi^\dagger(\mathbb{1})\| \mathrm{Tr}[B^* B] \geq \mathrm{Tr}[\Phi^\dagger(B)^* \Phi^\dagger(B))] . \qquad (4.3)$$

We now prove the extension of Theorem 1.6 corresponding to the full range of homogeneities in Theorem 1.2.

4.3 Theorem. *For all $s, t > 0$ with $s + t \leq 1$, all $m, n \in \mathbb{N}$, all $X, Y \in M_n^{++}(\mathbb{C})$ and all $K \in M_n(\mathbb{C})$ and all semiunital Schwarz maps $\Phi : M_m(\mathbb{C}) \to M_n(\mathbb{C})$ such that $\Phi^\dagger(\mathbb{1}) \in M_m^{++}(\mathbb{C})$,*

$$\mathrm{Tr}[\Phi^\dagger(K^*) \Phi^\dagger(Y)^{-s} \Phi^\dagger(K) \Phi^\dagger(X)^{-t}] \leq \left(\frac{n}{m}\right)^{1-(s+t)} \mathrm{Tr}[K^* Y^{-s} K X^{-t}] . \qquad (4.4)$$

Proof. We proceed as in the proof of Theorem 3.2. As in that case, we may assume $s < 1 - t$, and we define $r := s/(1-t)$ so that $0 < r < 1$. Since Φ is semiunital, $\frac{m}{n} \Phi^\dagger$ is unital and positive. Again, by [13, Theorem 2.1] or [5, Proposition 2.7.1], since $A \mapsto A^r$ is operator concave,

$$\frac{m}{n} \Phi^\dagger(Y^r) \leq \left(\frac{m}{n} \Phi^\dagger(Y)\right)^r ,$$

and hence by the fact that $A \mapsto A^{t-1}$ is operator decreasing,

$$\left(\frac{m}{n} \Phi^\dagger(Y^r)\right)^{t-1} \geq \left(\frac{m}{n} \Phi^\dagger(Y)\right)^{-s} . \qquad (4.5)$$

By Theorem 1.6 and then (4.5)

$$\frac{m}{n} \mathrm{Tr}[K^* Y^{-s} K X^{-t}] = \frac{m}{n} \mathrm{Tr}[K^* (Y^r)^{t-1} K X^{-t}]$$

$$\geq \frac{m}{n} \mathrm{Tr}[\Phi^\dagger(K)^* \Phi^\dagger(Y^r)^{t-1} \Phi^\dagger(K) \Phi^\dagger(X)^{-t}]$$

$$= \mathrm{Tr}\left[\left(\frac{m}{n}\Phi^\dagger(K)\right)^* \left(\frac{m}{n}\Phi^\dagger(Y^r)\right)^{t-1} \left(\frac{m}{n}\Phi^\dagger(K)\right) \left(\frac{m}{n}\Phi^\dagger(X)\right)^{-t}\right]$$

$$\geq \mathrm{Tr}\left[\left(\frac{m}{n}\Phi^\dagger(K)\right)^* \left(\frac{m}{n}\Phi^\dagger(Y)\right)^{-s} \left(\frac{m}{n}\Phi^\dagger(K)\right) \left(\frac{m}{n}\Phi^\dagger(X)\right)^{-t}\right] ,$$

and this is the same as (4.4). □

4.4 Remark. It is easy to see that taking Φ as in (1.12), one recovers Theorem 1.2 from Theorem 4.3.

We now combine duality with Theorem 4.3 to prove:

4.5 Theorem. *Let $0 \leq s < 1$. Let Φ be any sesquiunital Schwarz map from $M_m(\mathbb{C})$ to $M_n(\mathbb{C})$ for $m, n \in \mathbb{N}$. Then for all $A \in M_n^{++}(\mathbb{C})$ and all invertible $B \in M_n(\mathbb{C})$, and all*

$$\frac{1}{2-s} \leq q < 1 , \tag{4.6}$$

we have

$$\left(\frac{n}{m}\right)^{(2-s)q-1} \mathrm{Tr}[(B^*A^{-s}B)^q] \geq \mathrm{Tr}[(\Phi^\dagger(B)^*\Phi^\dagger(A)^{-s}\Phi^\dagger(B))^q] . \tag{4.7}$$

Proof. Let $0 < q < 1$. Then by (2.8),

$$(\mathrm{Tr}[(B^*A^{-s}B)^q])^{1/q} = \max\{\mathrm{Tr}[B^*A^{-s}BY] \;:\; Y \in M_n^{++}(\mathbb{C}) , \; \mathrm{Tr}[Y^{q/(q-1)}] = 1 \}$$
$$= \max\{\mathrm{Tr}[B^*A^{-s}BY^{(q-1)/q}] \;:\; Y \in M_n^{++}(\mathbb{C}) , \; \mathrm{Tr}[Y] = 1 \} .$$

In order to apply Theorem 4.3, we require $s + (1-q)/q \leq 1$, and together with $q < 1$, this yields the requirement that (4.6) be satisfied. Define $\delta := 2 - s - 1/q$. Then with (4.6) satisfied, Theorem 4.3 and the previous calculation yield

$$\left(\frac{n}{m}\right)^\delta (\mathrm{Tr}[(B^*A^{-s}B)^q])^{1/q}$$
$$= \left(\frac{n}{m}\right)^\delta \max\{\mathrm{Tr}[B^*A^{-s}BY^{(q-1)/q}] \;:\; Y \in M_n^{++}(\mathbb{C}) , \; \mathrm{Tr}[Y] = 1 \}$$
$$\geq \max\{\mathrm{Tr}[\Phi^\dagger(B)^*\Phi^\dagger(A)^{-s}\Phi^\dagger(B)\Phi^\dagger(Y)^{(q-1)/q}] \;:\; Y \in M_n^{++}(\mathbb{C}) , \; \mathrm{Tr}[Y] = 1 \}$$
$$\geq \max\{\mathrm{Tr}[\Phi^\dagger(B)^*\Phi^\dagger(A)^{-s}\Phi^\dagger(B)Z^{(q-1)/q}] \;:\; Z \in M_n^{++}(\mathbb{C}) , \; \mathrm{Tr}[Z] = 1 \}$$
$$= (\mathrm{Tr}[(\Phi^\dagger(B)^*\Phi^\dagger(A)^{-s}\Phi^\dagger(B))^q])^{1/q} .$$

Taking the q-th power of both sides yields (4.7). □

4.6 Remark. Note that for q saturating the lower bound in (4.6), (4.7) reduces to (4.1), so that Theorem 4.5 generalizes Theorem 4.1.

Now take $n = 2m$, $A_1, A_2 \in M_m^{++}(\mathbb{C})$ and invertible $B_1, B_2 \in M_m(\mathbb{C})$, and let

$$A := \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \quad \text{and} \quad B := \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}.$$

Let Φ be given by (1.12). Then

$$\frac{1}{2}\mathrm{Tr}[(B^*A^{-s}B)^q] = \frac{1}{2}\left(\mathrm{Tr}[(B_1^*A_1^{-s}B_1)^q] + \mathrm{Tr}[(B_2^*A_2^{-s}B_2)^q]\right),$$

and

$$\left(\frac{1}{2}\right)^{(2-s)q}\mathrm{Tr}[(\Phi^\dagger(B)^*\Phi^\dagger(A)^{-s}\Phi^\dagger(B))^q]$$

$$= \left(\frac{1}{2}\right)^{(2-s)q}\mathrm{Tr}[((B_1+B_2)^*(A_1+A_2)^{-s}(B_1+B_2))^q]$$

$$= \mathrm{Tr}\left[\left(\left(\frac{B_1+B_2}{2}\right)^*\left(\frac{A_1+A_2}{2}\right)^{-s}\left(\frac{B_1+B_2}{2}\right)\right)^q\right].$$

By Theorem 4.5 this proves:

4.7 Corollary. *Let $0 \le s < 1$ and let $1/(2-s) \le q < 1$. Then the map*

$$(A, B) \mapsto \mathrm{Tr}[(B^*A^{-s}B)^q] \tag{4.8}$$

is jointly convex on $M_n^{++}(\mathbb{C}) \times M_n(\mathbb{C})$.

A further corollary recover a result first proved in [10]:

4.8 Corollary. *Let $1 < p \le 2$, and let $r \ge 1$. Then for all $B \in M_n(\mathbb{C})$, the map*

$$A \mapsto \mathrm{Tr}[(B^*A^pB)^{r/p}] \tag{4.9}$$

is convex on $M_n^{++}(\mathbb{C})$.

Proof. First suppose that $1 \le r < p$. As in the proof of the Ando Convexity Theorem [1] as a direct consequence of the Lieb Convexity Theorem given in [9], replace B by AB in (4.8), so that $\mathrm{Tr}[(B^*A^{-s}B)^q]$ becomes $\mathrm{Tr}[(B^*A^{2-s}B)^q]$. Define $p := 2 - s$, and note that $1 < p \le 2$. Then $1/(2-s) \le q < 1$ becomes $1/p \le q < 1$. Finally, replacing q by r/p yields (4.9) in this case.

The case $r \ge p$ is elementary and does not depend on Theorem 1.6. As in [10, Lemma 2.1] (see also second part of the proof of Corollary 3.4), one first shows by duality that with $q = r/p$, $f(A) := \|B^*A^pB\|_q$ is convex. Then $g(A) := \mathrm{Tr}[(B^*A^pB)^q] = (f(A))^q$ is convex since $q \ge 1$. $\square$

5. Interpolation proof of Theorems 1.4 and 1.6

We give a simple proof of the monotonicity version of the Lieb Concavity Theorem for Schwarz maps using Lieb's original interpolation argument [20] adapted to this setting. This gives a third proof of this result and a further illustration of how direct proof of monotonicity can be both simpler and more general than direct proofs of concavity/convexity.

Interpolation Proof of Theorem 1.4. We don't need to reduce it to the case when $A = B$ that requires 2-positivity of Φ. By approximation, we may assume that all the matrices A, B, K and their images under $\Phi^\dagger$ are invertible. Denote $M := \Phi^\dagger(A)^{p/2} K \Phi^\dagger(B)^{(1-p)/2}$. Then

$$\mathrm{Tr}[K^* \Phi^\dagger(A)^p K \Phi^\dagger(B)^{1-p}] = \mathrm{Tr}[M^*M] . \tag{5.1}$$

Denote by S the strip $\{z \in \mathbb{C} : 0 < \Re z < 1\}$ and $\overline{S}$ its closure. Consider the following function

$$f(z) := \mathrm{Tr}\left[\Phi\left(\Phi^\dagger(B)^{-\frac{1-z}{2}} M^* \Phi^\dagger(A)^{-\frac{z}{2}}\right) A^z \Phi\left(\Phi^\dagger(A)^{-\frac{z}{2}} M \Phi^\dagger(B)^{-\frac{1-z}{2}}\right) B^{1-z}\right]$$

on $\overline{S}$. Then it is analytic on S and continuous on $\overline{S}$. It is also uniformly bounded on $\overline{S}$. In fact, for any $z = x + iy \in \overline{S}$ with $x \in [0,1]$ and $y \in \mathbb{R}$, by Hölder's inequality

$$
\begin{aligned}
|f(z)| &\leq \|\Phi(Y(z)) A^z \Phi(X(z)) B^{1-z}\|_1 \\
&\leq \|\Phi(X(z))\|_2 \|\Phi(Y(z))\|_2 \|A^z\| \|B^{1-z}\| \\
&\leq \|\Phi(X(z))\|_2 \|\Phi(Y(z))\|_2 \|A\|^x \|B\|^{1-x},
\end{aligned}
$$

where $X(z) := \Phi^\dagger(A)^{-z/2} M \Phi^\dagger(B)^{(z-1)/2}$ and

$$Y(z) := X(\overline{z})^* = \Phi^\dagger(B)^{(z-1)/2} M^* \Phi^\dagger(A)^{-z/2}.$$

Since Φ is a Schwarz map, we have

$$\|\Phi(X(z))\|_2^2 = \mathrm{Tr}\left[\Phi(X(z))^* \Phi(X(z))\right] \leq \mathrm{Tr}\left[\Phi(X(z)^* X(z))\right] \leq \|\Phi^\dagger(\mathbb{1})\| \mathrm{Tr}\left[X(z)^* X(z)\right],$$

where the last inequality follows from Hölder's inequality. By definition,

$$\mathrm{Tr}\left[X(z)^* X(z)\right] = \mathrm{Tr}\left[M^* \Phi^\dagger(A)^{-x} M \Phi^\dagger(B)^{x-1}\right] .$$

Similarly, we have

$$\|\Phi(Y(z))\|_2^2 = \|\Phi(X(\overline{z}))^*\|_2^2 \leq \|\Phi^\dagger(\mathbb{1})\| \mathrm{Tr}\left[M^* \Phi^\dagger(A)^{-x} M \Phi^\dagger(B)^{x-1}\right] .$$

Therefore we have proved

$$\sup_{z \in \overline{S}} |f(z)| \leq \|\Phi^\dagger(\mathbb{1})\| \sup_{x \in [0,1]} \|A\|^x \|B\|^{1-x} \mathrm{Tr}\left[M^* \Phi^\dagger(A)^{-x} M \Phi^\dagger(B)^{x-1}\right] < \infty \ .$$

So f is uniformly bounded on $\overline{S}$. Note that by definition

$$f(p) = \mathrm{Tr}[\Phi(K)^* A^p \Phi(K) B^{1-p}] \ .$$

By the three-lines lemma, it remains to check that

$$\sup_{y \in \mathbb{R}} |f(k + iy)| \leq \mathrm{Tr}[M^* M], \quad \text{for} \quad k = 0, 1 \ . \tag{5.2}$$

By definition,

$$\begin{aligned} f(iy) =& \mathrm{Tr}\left[\Phi\left(\Phi^\dagger(B)^{-\frac{1-iy}{2}} M^* \Phi^\dagger(A)^{-\frac{iy}{2}}\right) A^{iy} \Phi\left(\Phi^\dagger(A)^{-\frac{iy}{2}} M \Phi^\dagger(B)^{-\frac{1-iy}{2}}\right) B^{1-iy}\right] \\ =& \mathrm{Tr}\left[\Phi^\dagger(B)^{-\frac{1-iy}{2}} M^* \Phi^\dagger(A)^{-\frac{iy}{2}} \Phi^\dagger(X)\right] \\ =& \mathrm{Tr}\left[\Phi^\dagger(B)^{\frac{iy}{2}} M^* \Phi^\dagger(A)^{-\frac{iy}{2}} \cdot \Phi^\dagger(X) \Phi^\dagger(B)^{-\frac{1}{2}}\right] \end{aligned}$$

where $X = A^{iy} \Phi\left(\Phi^\dagger(A)^{-iy/2} M \Phi^\dagger(B)^{(iy-1)/2}\right) B^{1-iy}$. By Cauchy–Schwarz inequality, we have

$$|f(iy)|^2 \leq \mathrm{Tr}[M^* M] \mathrm{Tr}[\Phi^\dagger(X) \Phi^\dagger(B)^{-1} \Phi^\dagger(X^*)] \ .$$

Since Φ is Schwarz, by (4.2) or [11, Theorem 4]

$$\begin{aligned} \mathrm{Tr}[\Phi^\dagger(X) \Phi^\dagger(B)^{-1} \Phi^\dagger(X^*)] &\leq \mathrm{Tr}[X B^{-1} X^*] \\ &= \mathrm{Tr}\left[B\Phi\left(\Phi^\dagger(B)^{-\frac{1+iy}{2}} M^* \Phi^\dagger(A)^{\frac{iy}{2}}\right) \Phi\left(\Phi^\dagger(A)^{-\frac{iy}{2}} M \Phi^\dagger(B)^{-\frac{1-iy}{2}}\right)\right] . \end{aligned}$$

Using again that Φ is Schwarz, this is bounded from above by

$$\mathrm{Tr}\left[B\Phi\left(\Phi^\dagger(B)^{-\frac{1+iy}{2}} M^* M \Phi^\dagger(B)^{-\frac{1-iy}{2}}\right)\right] = \mathrm{Tr}[M^* M] \ .$$

So we have proved (5.2) for $k = 0$. The proof of (5.2) for $k = 1$ is similar. □

Using a similar argument, we can also give an interpolation proof of the monotonicity version of Lieb Convexity Theorem, Theorem 1.6.

Interpolation Proof of Theorem 1.6. Again, we may assume that all the matrices X, Y, K and their images under $\Phi^\dagger$ are invertible. Denote $M := Y^{(t-1)/2} K X^{-t/2}$. Then

$$\mathrm{Tr}[K^*Y^{t-1}KX^{-t}] = \mathrm{Tr}[M^*M] \ .$$

Let S be the open strip as above. Consider the following function

$$g(z) := \mathrm{Tr}\left[\Phi^\dagger\left(X^{\frac{z}{2}}M^*Y^{\frac{1-z}{2}}\right)\Phi^\dagger(Y)^{z-1}\Phi^\dagger\left(Y^{\frac{1-z}{2}}MX^{\frac{z}{2}}\right)\Phi^\dagger(X)^{-z}\right]$$

on $\overline{S}$. Then it is analytic on S and continuous on $\overline{S}$. It is uniformly bounded on $\overline{S}$, which we now argue as above. For any $z = x + iy \in \overline{S}$ with $x \in [0,1]$ and $y \in \mathbb{R}$, by Hölder's inequality

$$\begin{aligned}
|g(z)| &\leq \|\Phi^\dagger(B(z))\Phi^\dagger(Y)^{z-1}\Phi^\dagger(A(z))\Phi^\dagger(X)^{-z}\|_1 \\
&\leq \|\Phi^\dagger(A(z))\|_2\|\Phi^\dagger(B(z))\|_2\|\Phi^\dagger(Y)^{z-1}\|\|\Phi^\dagger(X)^{-z}\| \ ,
\end{aligned}$$

where $A(z) := Y^{(1-z)/2}MX^{z/2}$ and $B(z) := A(\overline{z})^* = X^{z/2}M^*Y^{(1-z)/2}$. Note that $x - 1, -x \leq 0$, we get

$$\|\Phi^\dagger(Y)^{z-1}\|\|\Phi^\dagger(X)^{-z}\| = \|\Phi^\dagger(Y)^{-1}\|^{1-x}\|\Phi^\dagger(X)^{-1}\|^x \ .$$

Since Φ is Schwarz, by (4.3)

$$\begin{aligned}
\|\Phi^\dagger(A(z))\|_2^2 = \mathrm{Tr}\left[\Phi^\dagger(A(z))^*\Phi^\dagger(A(z))\right] &\leq \|\Phi^\dagger(\mathbb{1})\|\mathrm{Tr}\left[A(z)^*A(z)\right] \\
&= \|\Phi^\dagger(\mathbb{1})\|\mathrm{Tr}\left[M^*Y^{1-x}MX^x\right] \ .
\end{aligned}$$

Similarly,

$$\|\Phi^\dagger(B(z))\|_2^2 = \|\Phi^\dagger(A(\overline{z}))^*\|_2^2 \leq \|\Phi^\dagger(\mathbb{1})\|\mathrm{Tr}\left[M^*Y^{1-x}MX^x\right] \ .$$

So we have shown that

$$\sup_{z\in\overline{S}}|g(z)| \leq \|\Phi^\dagger(\mathbb{1})\|\sup_{x\in[0,1]}\|\Phi^\dagger(Y)^{-1}\|^{1-x}\|\Phi^\dagger(X)^{-1}\|^x\mathrm{Tr}\left[M^*Y^{1-x}MX^x\right] < \infty \ .$$

Hence g is uniformly bounded on $\overline{S}$. Note that by definition

$$g(t) = \mathrm{Tr}[\Phi^\dagger(K)^*\Phi^\dagger(Y)^{t-1}\Phi^\dagger(K)\Phi^\dagger(X)^{-t}] \ .$$

So by the three-lines lemma, it suffices to show that

$$\sup_{y\in\mathbb{R}}|g(k+iy)| \leq \mathrm{Tr}[M^*M], \quad \text{for} \quad k = 0,1. \tag{5.3}$$

When $k = 0$, we have

$$g(iy) = \text{Tr}\left[\Phi^\dagger \left(X^{\frac{iy}{2}} M^* Y^{\frac{1-iy}{2}}\right) \Phi^\dagger (Y)^{iy-1} \Phi^\dagger \left(Y^{\frac{1-iy}{2}} M X^{\frac{iy}{2}}\right) \Phi^\dagger (X)^{-iy}\right]$$

$$= \text{Tr}\left[X^{\frac{iy}{2}} M^* Y^{\frac{1-iy}{2}} \Phi(A)\right]$$

$$= \text{Tr}\left[X^{\frac{iy}{2}} M^* Y^{-\frac{iy}{2}} \cdot Y^{\frac{1}{2}} \Phi(A)\right] \ ,$$

where $A := \Phi^\dagger (Y)^{iy-1} \Phi^\dagger \left(Y^{(1-iy)/2} M X^{iy/2}\right) \Phi^\dagger (X)^{-iy}$. By Cauchy–Schwarz inequality,

$$|g(iy)|^2 \leq \text{Tr}[M^* M] \text{Tr}[\Phi(A)^* Y \Phi(A)] \ .$$

Since Φ is a Schwarz map, we have

$$\text{Tr}[\Phi(A)^* Y \Phi(A)] \leq \text{Tr}\left[A^* \Phi^\dagger (Y) A\right] = \text{Tr}\left[\Phi^\dagger (B)^* \Phi^\dagger (Y)^{-1} \Phi^\dagger (B)\right] \ ,$$

where $B := Y^{(1-iy)/2} M X^{iy/2}$. Again, since Φ is Schwarz, by (4.2) or [11, Theorem 4] that

$$\text{Tr}\left[\Phi^\dagger (B)^* \Phi^\dagger (Y)^{-1} \Phi^\dagger (B)\right] \leq \text{Tr}\left[B^* Y^{-1} B\right] = \text{Tr}[M^* M] \ .$$

So we have proved (5.3) for $k = 0$. The proof for $k = 1$ is similar. $\quad\square$

Declaration of competing interest

None.

Data availability

No data was used for the research described in the article.

References

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

[2] H. Araki, Inequalities in von Neumann algebras, in: Les rencontres physiciens-mathématiciens de Strasbourg, vol. 22, 1975, pp. 1–25.

[3] S. Beigi, Sandwiched Rényi divergence satisfies data processing inequality, J. Math. Phys. 54 (12) (2013) 122202, 11.

[4] R. Bhatia, Matrix Analysis, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, New York, 1997.

[5] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2009.

[6] G. Birkhoff, Tres observaciones sobre el algebra lineal, Rev. Univ. Nac. Tucumán, Mat. Fis. Teor. 5 (1946) 147–151.

[7] E.A. Carlen, Trace inequalities and quantum entropy: an introductory course, in: Entropy and the Quantum, in: Contemp. Math., vol. 529, Amer. Math. Soc., Providence, RI, 2010, pp. 73–140.

[8] E.A. Carlen, A monotonicity version of a concavity theorem of Lieb, e-prints, arXiv:2203.02789, March 2022.

[9] E.A. Carlen, On some convexity and monotonicity inequalities of Elliott Lieb, e-prints, arXiv: 2202.03591, February 2022.

[10] E.A. Carlen, E.H. Lieb, A Minkowski-type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity, Lett. Math. Phys. 83 (2008) 107–126.
[11] E.A. Carlen, A. Müller-Hermes, Characterizing schwarz maps by tracial inequalities, e-prints, arXiv: 2203.03433, 2022.
[12] V. Chayes, Reverse Hölder, Minkowski, and Hanner inequalities for matrices, e-prints, arXiv:2103. 09915, 2021.
[13] M.D. Choi, A Schwarz inequality for positive linear maps on C^* algebras, Ill. J. Math. 18 (1974) 565–574.
[14] M.D. Choi, Some assorted inequalities for positive linear maps on C^* algebras, J. Oper. Theory 4 (1980) 271–285.
[15] H. Epstein, Remarks on two theorems of E. Lieb, Commun. Math. Phys. 31 (1973) 317–325.
[16] F. Hiai, Concavity of certain matrix trace and norm functions, Linear Algebra Appl. 439 (5) (2013) 1568–1589.
[17] F. Hiai, M. Mosonyi, D. Petz, C. Bény, Quantum f-divergences and error correction, Rev. Math. Phys. 23 (07) (2011) 691–747.
[18] F. Hiai, D. Petz, From quasi-entropy to various quantum information quantities, Publ. Res. Inst. Math. Sci. 48 (3) (2012) 525–542.
[19] R.V. Kadison, A generalized Schwarz inequality and algebraic invariants for operator algebras, Ann. Math. 58 (1952) 494–503.
[20] E.H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, Adv. Math. 11 (1973) 267–288.
[21] G. Lindblad, Expectations and entropy inequalities for finite quantum systems, Commun. Math. Phys. 39 (1974) 111–119.
[22] G. Lindblad, Completely positive maps and entropy inequalities, Commun. Math. Phys. 40 (2) (1975) 147–151.
[23] A. Müller-Hermes, D. Reeb, Monotonicity of the quantum relative entropy under positive maps, Ann. Henri Poincaré 18 (5) (2017) 1777–1788.
[24] D. Petz, Monotone metrics on matrix spaces, Linear Algebra Appl. 244 (1996) 81–96.
[25] W.F. Stinespring, Positive functions on C^*-algebras, Proc. Am. Math. Soc. 6 (1955) 211–216.
[26] A. Uhlmann, Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory, Commun. Math. Phys. 54 (1) (1977) 21–32.
[27] H. Zhang, From Wigner-Yanase-Dyson conjecture to Carlen-Frank-Lieb conjecture, Adv. Math. 365 (2020) 107053, 18.