Mathematical
Inequalities
& Applications
Volume 1, Number 1 (1998), 1–30

HÖLDER TYPE INEQUALITIES FOR MATRICES

TSUYOSHI ANDO AND FUMIO HIAI

(*communicated by R. Mathias*)

Abstract. We discuss Hölder type inequalites involving $(A^p + B^p)^{1/p}$ for positive semi-definite matrices A, B. Matrix or trace inequalities of Hölder type as well as weak majorizations of similar type are obtained. Also we give counter-examples for expected Hölder type inequalities.

0. Introduction

When $1 < p, q < \infty$ and $1/p + 1/q = 1$, the simplest form of numerical Hölder inequality is written as

$$(|a|^p + |b|^p)^{1/p}(|c|^q + |d|^q)^{1/q} \geqslant |ac + bd| \tag{0.1}$$

for $a, b, c, d \in \mathbb{C}$. Moreover, we have the variational expression

$$(|a|^p + |b|^p)^{1/p} = \max\{|ac + bd| : |c|^q + |d|^q = 1\}. \tag{0.2}$$

Another formulation of Hölder inequality is

$$(a_1 + a_2)^{1/p}(b_1 + b_2)^{1/q} \geqslant a_1^{1/p}b_1^{1/q} + a_2^{1/p}b_2^{1/q} \quad \text{for } a_j, b_j \geqslant 0, \tag{0.3}$$

which means joint concavity of the function $(a, b) \mapsto a^{1/p}b^{1/q}$ in $a, b \geqslant 0$.

A well-known Hölder inequality for matrices (or operators) is that for the Schatten p-norms and is not viewed as a genuine matrix inequality. Although many kinds of matrix inequalilties, trace inequalities, and (weak) majorization relations are known so far as summarized in [3], we have very few ones of Hölder type. The aim of this paper is to make the first step to develop Hölder type inequalities for matrices themselves. Namely, we want to get inequalities involving $(|A|^p + |B|^p)^{1/p}$ or $(A^p + B^p)^{1/p}$ for (positive semi-definite) matrices A, B.

A direct generalization of (0.1) or (0.2) to matrices may be formulated as follows: If $1/p + 1/q = 1$ and $|C^*|^q + |D^*|^q = I$, then

$$(|A|^p + |B|^p)^{1/p} \geqslant |CA + DB| \tag{0.4}$$

Mathematics subject classification (1991): Primary 15A45, 15A42; Secondary 15A60, 47A63.

Key words and phrases: Positive semi-definite matrices, Hölder inequality, trace inequalities, weak majorization, Oppenheim inequality.

Paper MIA-01-01

holds in the order of positive semi-definiteness. As is well known to experts, the case $p = 2$ (or a matrix Cauchy-Schwarz inequality) is shown in this form by the method of 2×2 block matrices (see the beginning of Sec. 1). However, the matrix inequality (0.4) turns out to be false for any $1 < p < \infty$ except $p = 2$ (see Proposition 3.1), so that some obstacles must be encountered in our study. Nevertheless, we can obtain, related to (0.1) and (0.2), several (trace) inequalities and weak majorizations of Hölder type for matrices.

In connection with (0.3) it is well known [12] (also [1]) that the function $(A,B) \mapsto \mathrm{Tr}(A^{1/p}B^{1/q})$ is jointly concave in $A,B \geqslant 0$ whenever $1 < p,q < \infty$ and $1/p+1/q = 1$. On the other hand, even when we consider $B^{1/2q}A^{1/p}B^{1/2q}$ instead of $A^{1/p}B^{1/q}$, it is impossible to generalize (0.3) to a matrix inequality. In fact, $B \mapsto B^{1/2q}A^{1/p}B^{1/2q}$ is not operator concave for any $q > 0$ (see Proposition 4.1). So our concern is to obtain weak majorization relations for eigenvalue products generalizing joint concavity (0.3).

The paper is organized as follows. In Sec. 1, we prove a matrix Hölder inequality

$$(|A|^p + |B|^p)^{2/p} \geqslant |\alpha^{1/r}CA + (1-\alpha)^{1/r}DB|^2$$

if $0 \leqslant \alpha \leqslant 1$ and $|C^*|^q + |D^*|^q \leqslant I$, where $2 \leqslant p,q < \infty$, $1 < r < \infty$, and $1/p + 1/q = 1 - 1/r$. This may be unsatisfactory in the points that an additional parameter α is contained and that the case $1 < p < 2$ is excluded. But there must be some weakening for a matrix Hölder inequality because the best form (0.4) is false.

In Sec. 2, we discuss Hölder type inequalities for matrices under trace. For $A,B \geqslant 0$ a main trace inequality presented is

$$\mathrm{Tr}(A^p + B^p)^{1/p} \geqslant \mathrm{Tr}(CA + DB)$$

if $C,D \geqslant 0$ and $C^q + D^q \leqslant I$, where $1 < p,q < \infty$ and $1/p + 1/q = 1$. As a corollary we have

$$\mathrm{Tr}(A^p + B^p)^{1/p} \geqslant \sum_{i=1}^{n} (a_{ii}^p + b_{ii}^p)^{1/p},$$

and this can be strengthened to a form of weak majorization. Furthermore, we compare $\mathrm{Tr}(A^p + B^p)^{1/p}$ with some variational expressions such as

$$\max\{\mathrm{Tr}\,|CA + DB| : |C^*|^q + |D^*|^q \leqslant I\},$$
$$\max\{\mathrm{Tr}(CA + DB) : C,D \geqslant 0,\ C^q + D^q \leqslant I\}.$$

In Sec. 3, we take as A, B two non-commuting orthogonal projections of rank one to give counter-examples for expected matrix or trace inequalities of Hölder type. In particular, all possible candidates of variational expressions of $\mathrm{Tr}(A^p + B^p)^{1/p}$ fail to hold true for any $2 < p < \infty$. This tells us that an attempt to prove joint concavity of $(A,B) \mapsto \mathrm{Tr}(A^p + B^p)^{1/p}$ via variational expression seems hopeless. Indeed, we show that $\mathrm{Tr}(A^p + B^p)^{1/p}$ is not jointly concave in $A,B \geqslant 0$ for any $2 < p < \infty$, while this is the case for $p = 2$ because of variational expression. The case $1 < p < 2$ is a challenging open problem.

In Sec. 4, we obtain a Hölder type inequality given under taking eigenvalue products $\prod_{i=1}^{k} \lambda_{n-i+1}(\cdot)$ where $\lambda_1(A),\dots,\lambda_n(A)$ are the eigenvalues of A in decreasing order. This is a kind of weak majorization viewed as an extension of (0.3) and also a generalization of the Oppenheim inequality [14]. For the proof the technique of antisymmetric tensor powers and the concavity property given in [12] and [1] are useful. In Sec. 5 and 6, we further give similar weak majorizations involving α-power mean and Hadamard product.

Acknowledgments. We thank Professor D. Petz who suggested us the problem on joint convexity of $\mathrm{Tr}(A^p + B^p)^{1/p}$ when $p > 1$. Our study was motivated by this problem. We are grateful to Professor M. B. Ruskai who informed us that E. A. Carlen and E. H. Lieb studied the same problem. But we do not know the details.

1. Matrix inequalities of Hölder type

In this paper we always consider $n \times n$ complex matrices, which are denoted by A, B, C, etc. We use the usual notations: the identity matrix I, the adjont A^*, the absolute value $|A| = (A^*A)^{1/2}$, the operator norm or the spectral norm $\|A\|$, the trace $\mathrm{Tr}\,A$, the determinant $\det A$, and so on. Matrix inequalities (between Hermitian matrices) mean those with respect to the order of positive semi-definiteness; in particular, $A \geqslant 0$ means that A is positive semi-definite.

Let f be a real continuous function on an interval J. Recall that f is said to be operator monotone if $A \geqslant B$ implies $f(A) \geqslant f(B)$ for any Hermitian matrices (of any size) A, B whose eigenvalues are in J. Also, f is said to be operator convex if $f(\alpha A + (1-\alpha)B) \leqslant \alpha f(A) + (1-\alpha)f(B)$ for every Hermitian A, B with eigenvalues in J, and operator concave if inequality sign is reversed.

As is well known, the trick of 2×2 block matrices is quite useful when one discusses matrix inequalities. This is based on the fact that if A, B, C are matrices and $A, B \geqslant 0$, then $\begin{bmatrix} A & C^* \\ C & B \end{bmatrix} \geqslant 0$ if and only if $C = B^{1/2}WA^{1/2}$ for some W with $\|W\| \leqslant 1$. This says in particular that $\begin{bmatrix} A & C^* \\ C & I \end{bmatrix} \geqslant 0$ if and only if $A \geqslant C^*C$. The following is a typical example of 2×2 arguments. For any matrices A, B, C, D we have

$$\begin{bmatrix} A^*A + B^*B & A^*C^* + B^*D^* \\ CA + DB & CC^* + DD^* \end{bmatrix} = \begin{bmatrix} A & C^* \\ B & D^* \end{bmatrix}^* \begin{bmatrix} A & C^* \\ B & D^* \end{bmatrix} \geqslant 0,$$

so that if $CC^* + DD^* = I$ (or more weakly $CC^* + DD^* \leqslant I$), then

$$A^*A + B^*B \geqslant |CA + DB|^2 \tag{1.1}$$

and hence by operator monotonity of $t^{1/2}$

$$(A^*A + B^*B)^{1/2} \geqslant |CA + DB|. \tag{1.2}$$

Furthermore, when $C = (A^*A + B^*B)^{-1/2}A^*$ and $D = (A^*A + B^*B)^{-1/2}B^*$ (with a slight modification if $A^*A + B^*B$ is not invertible), we obtain $CC^* + DD^* = I$ and equality occurs in (1.1). Inequality (1.1) or (1.2) under $CC^* + DD^* = I$ is regarded as a matrix Cauchy-Schwarz inequality.

One may expect the generalization of (1.1) or (1.2) to a matrix Hölder inequality. Namely, if $1 < p,q < \infty$, $1/p + 1/q = 1$, and $|C^*|^q + |D^*|^q = I$, then does the inequality

$$(|A|^p + |B|^p)^{1/p} \geqslant |CA + DB| \tag{1.3}$$

hold true? Unfortunately, this is false for every $1 < p < \infty$ except $p = 2$ as we shall explicitly see in Sec. 3. So one has to weaken (1.3) to get some matrix inequality of Hölder type. The aim of this section is to establish an inequality of this type.

The next simple lemma will be sometimes useful in this paper.

LEMMA 1.1. *If $1 < p < \infty$ and $0 \leqslant \alpha \leqslant 1$, then for every $A, B \geqslant 0$*

$$(A^p + B^p)^{1/p} \geqslant \alpha^{1-1/p} A + (1-\alpha)^{1-1/p} B.$$

Proof. We may assume $0 < \alpha < 1$ because the result for $\alpha = 0, 1$ is obtained by taking limit (or obvious by operator monotonity of $t^{1/p}$). We have

$$\begin{aligned}(A^p + B^p)^{1/p} &= \{\alpha(\alpha^{-1/p}A)^p + (1-\alpha)((1-\alpha)^{-1/p}B)^p\}^{1/p} \\ &\geqslant \alpha^{1-1/p}A + (1-\alpha)^{1-1/p}B\end{aligned}$$

by operator concavity of $t^{1/p}$. □

THEOREM 1.2. *Let $2 \leqslant p,q < \infty$ and $1 < r \leqslant \infty$ with $1/p + 1/q = 1 - 1/r$. Then for any A, B, C, D and $0 \leqslant \alpha \leqslant 1$,*

$$(|A|^p + |B|^p)^{2/p} \geqslant |\alpha^{1/r} CA + (1-\alpha)^{1/r} DB|^2$$

whenever $|C^*|^q + |D^*|^q \leqslant I$.

Proof. Since the case $r = \infty$ (hence $p = q = 2$) is (1.1), we may assume $1 < r < \infty$. It suffices to show that for every A, B, C, D

$$\begin{bmatrix} (|A|^p + |B|^p)^{2/p} & \alpha^{1/r}A^*C^* + (1-\alpha)^{1/r}B^*D^* \\ \alpha^{1/r}CA + (1-\alpha)^{1/r}DB & (|C^*|^q + |D^*|^q)^{2/q} \end{bmatrix} \geqslant 0. \tag{1.4}$$

Since $(1/2 - 1/p) + (1/2 - 1/q) = 1/r$, we get

$$\begin{aligned}&\begin{bmatrix} \alpha^{1-2/p}A^*A + (1-\alpha)^{1-2/p}B^*B & \alpha^{1/r}A^*C^* + (1-\alpha)^{1/r}B^*D^* \\ \alpha^{1/r}CA + (1-\alpha)^{1/r}DB & \alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^* \end{bmatrix} \\ &= \begin{bmatrix} \alpha^{1/2-1/p}A^* & (1-\alpha)^{1/2-1/p}B^* \\ \alpha^{1/2-1/q}C & (1-\alpha)^{1/2-1/q}D \end{bmatrix} \begin{bmatrix} \alpha^{1/2-1/p}A & \alpha^{1/2-1/q}C^* \\ (1-\alpha)^{1/2-1/p}B & (1-\alpha)^{1/2-1/q}D^* \end{bmatrix} \\ &\geqslant 0.\end{aligned}$$

Since Lemma 1.1 gives

$$(|A|^p + |B|^p)^{2/p} \geqslant \alpha^{1-2/p}A^*A + (1-\alpha)^{1-2/p}B^*B,$$

$$(|C^*|^q + |D^*|^q)^{2/q} \geqslant \alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^*,$$

we obtain (1.4). □

In particular, let $r = 2$ and replace p, q by $2p, 2q$. Then the theorem says that if $1 \leqslant p, q < \infty$, $1/p + 1/q = 1$, and $0 \leqslant \alpha \leqslant 1$, then

$$(|A|^{2p} + |B|^{2p})^{1/p} \geqslant |\sqrt{\alpha}\, CA + \sqrt{1-\alpha}\, DB|^2$$

whenever $|C^*|^{2q} + |D^*|^{2q} \leqslant I$.

It is known that for every $A, B \geqslant 0$ the increasing limit of $\{(A^p + B^p)/2\}^{1/p}$ as $p \to \infty$ exists. So $(A^p + B^p)^{1/p}$ converges to the same limit as $p \to \infty$. We write

$$A \vee B = \lim_{p\to\infty} \left(\frac{A^p + B^p}{2} \right)^{1/p} = \lim_{p\to\infty} (A^p + B^p)^{1/p}. \tag{1.5}$$

In fact, this $A \vee B$ is the supremum of A, B with respect to some spectral order among Hermitian matrices (see [9], [2, Lemma 6.5]). For instance, if P and Q are orthogonal projections, then $P \vee Q$ coincides with the usual supremum of projections.

COROLLARY 1.3. *Let $2 \leqslant p < \infty$ and $1 < p' \leqslant 2$ with $1/p + 1/p' = 1$. Then for any A, B, C, D,*

$$(\|C\|^{p'} + \|D\|^{p'})^{2/p'} (|A|^p + |B|^p)^{2/p} \geqslant |CA + DB|^2. \tag{1.6}$$

Moreover,

$$(\|C\| + \|D\|)^2 (|A| \vee |B|)^2 \geqslant |CA + DB|^2. \tag{1.7}$$

Proof. Taking the limit of (1.4) as $q \to \infty$ with p fixed (hence $r \to p'$) we get

$$\begin{bmatrix} (|A|^p + |B|^p)^{2/p} & \alpha^{1/p'} A^* C^* + (1-\alpha)^{1/p'} B^* D^* \\ \alpha^{1/p'} CA + (1-\alpha)^{1/p'} DB & |C^*| \vee |D^*| \end{bmatrix} \geqslant 0.$$

Note that $|C^*| \vee |D^*| \leqslant I$ if and only if $|C^*| \leqslant I$ and $|D^*| \leqslant I$, that is, $\|C\| \leqslant 1$ and $\|D\| \leqslant 1$. Hence, if $\|C\| \leqslant 1$ and $\|D\| \leqslant 1$, then for every $0 \leqslant \alpha \leqslant 1$

$$(|A|^p + |B|^p)^{2/p} \geqslant |\alpha^{1/p'} CA + (1-\alpha)^{1/p'} DB|^2.$$

By replacing $\alpha^{1/p'} C$ and $(1-\alpha)^{1/p'} D$ by C and D respectively, this means that if $\|C\|^{p'} + \|D\|^{p'} \leqslant 1$ then $(|A|^p + |B|^p)^{2/p} \geqslant |CA + DB|^2$. This is equivalent to (1.6). (1.7) is the limit of (1.6) as $p \to \infty$. $\square$

Here let us give a simple but worthwhile remark. Let $A = U|A|$ and $B = V|B|$ be the polar decompositions with unitaries U, V. If we set $C_1 = CU$ and $D_1 = DV$, then

$$|C^*| = |C_1^*|, \quad |D^*| = |D_1^*|, \quad |CA + DB| = |C_1\,|A| + D_1\,|B|\,|.$$

This says that when one discusses Hölder type inequalities such as (1.3) as well as similar trace inequalities, it may be assumed without loss of generality that A, B are positive semi-definite.

2. Trace inequalities and weak majorizations of Hölder type

Hereafter, according to the remark above, we shall assume that $A,B \geqslant 0$. Another way of weakening (1.3) is to consider it under trace. A tracial Hölder inequality may be proposed as follows: If $1 < p,q < \infty$, $1/p + 1/q = 1$, and $|C^*|^q + |D^*|^q = I$ (or $|C^*|^q + |D^*|^q \leqslant I$), then does the inequality

$$\mathrm{Tr}(A^p + B^p)^{1/p} \geqslant \mathrm{Tr}\,|CA + DB| \tag{2.1}$$

hold true? To consider this question, it is convenient to use the following variational expressions associated with $A,B \geqslant 0$:

$$V_p(A,B) = \max\{\mathrm{Tr}\,|CA + DB| : |C^*|^q + |D^*|^q \leqslant I\}, \tag{2.2}$$

$$\tilde{V}_p(A,B) = \max\{\mathrm{Tr}\,|CA + DB| : |C^*|^q + |D^*|^q = I\}, \tag{2.3}$$

where $1 < p \leqslant \infty$ and $1 \leqslant q < \infty$ with $1/p + 1/q = 1$. Then the problem consists of the comparison of $\mathrm{Tr}(A^p + B^p)^{1/p}$ with $V_p(A,B)$ (or $\tilde{V}_p(A,B)$). Here $\mathrm{Tr}(A^p + B^p)^{1/p}$ when $p = \infty$ means $\mathrm{Tr}(A \vee B)$ according to (1.5). Also, in case of $p = 1$ we may define

$$\begin{aligned} V_1(A,B) &= \lim_{p \to 1} V_p(A,B) \\ &= \max\{\mathrm{Tr}\,|CA + DB| : \|C\|,\ \|D\| \leqslant 1\} \end{aligned} \tag{2.4}$$

due to (1.5) and equivalence of $|C^*| \vee |D^*| \leqslant I$ and $\|C\|, \|D\| \leqslant 1$. But the problem is trivial in this case because $\mathrm{Tr}(A + B) = V_1(A,B)$ is obvious.

From the argument at the beginning of Sec. 1 it is clear that

$$\mathrm{Tr}(A^2 + B^2)^{1/2} = V_2(A,B) = \tilde{V}_2(A,B). \tag{2.5}$$

But it seems that this variational equality fails to hold for any $1 < p \leqslant \infty$ except $p = 2$. Indeed, we shall explicitly show in Sec. 3 that for each $2 < p < \infty$ there is no relation between $\mathrm{Tr}(A^p + B^p)^{1/p}$ and $V_p(A,B)$; more precisely, both $\mathrm{Tr}(A^p + B^p)^{1/p} > V_p(A,B)$ and $\mathrm{Tr}(A^p + B^p)^{1/p} < \tilde{V}_p(A,B)$ can occur. So one needs to modify (2.1) to get some trace inequality of Hölder type (at least when $2 < p < \infty$). In this section we shall obtain a weaker version of (2.1) as well as related inequalities.

First, let us give equivalent expressions of $V_p(A,B)$ and $\tilde{V}_p(A,B)$, while they are not essential in our later discussions.

PROPOSITION 2.1. *Let $1 < p \leqslant \infty$ and $1/p + 1/q = 1$. Then for $A,B \geqslant 0$,*

$$\begin{aligned} V_p(A,B) &= \max\{\mathrm{Re}\,\mathrm{Tr}(CA + DB) : |C^*|^q + |D^*|^q \leqslant I\} \\ &= \max\{\mathrm{Tr}(|CA| + |DB|) : C,D \geqslant 0,\ C^q + D^q \leqslant I\}. \end{aligned}$$

The analogous expressions hold for $\tilde{V}_p(A,B)$ too. Moreover, $V_p(A,B) = \tilde{V}_p(A,B)$ when $1 < p \leqslant 2$.

Proof. If $|C^*|^q + |D^*|^q \leqslant I$ and $CA + DB = U|CA + DB|$ is the polar decomposition, then we have

$$\begin{aligned} &\mathrm{Tr}\,|CA + DB| = \mathrm{Tr}(U^*CA + U^*DB), \\ &|(U^*C)^*|^q + |(U^*D)^*|^q = U^*(|C^*|^q + |D^*|^q)U \leqslant I. \end{aligned}$$

Hence the first equality is obtained.

Let $C,D \geqslant 0$ and $C^q + D^q \leqslant I$. Taking the polar decompositions $CA = V|CA|$ and $DB = W|DB|$ we get

$$\begin{aligned}\mathrm{Tr}(|CA| + |DB|) &= \mathrm{Tr}(V^*CA + W^*DB) \\ &= \mathrm{Tr}(ACV + BDW) \\ &= \mathrm{Tr}(CVA + DWB),\end{aligned}$$

$$|(CV)^*|^q + |(DW)^*|^q = |C^*|^q + |D^*|^q \leqslant I.$$

Moreover, for any C,D with the polar decompositions $C^* = V|C^*|$ nad $D^* = W|D^*|$ we get

$$\begin{aligned}\mathrm{Re}\,\mathrm{Tr}(CA + DB) &= \mathrm{Re}\,\mathrm{Tr}(A\,|C^*|\,V^* + B\,|D^*|\,W^*) \\ &\leqslant \mathrm{Tr}(|A\,|C^*|\,| + |B\,|D^*|\,|) \\ &= \mathrm{Tr}(|\,|C^*|\,A| + |\,|D^*|\,B|).\end{aligned}$$

Hence the second equality is obtained.

In the same way as above we have also

$$\begin{aligned}\tilde{V}_p(A,B) &= \max\{\mathrm{Re}\,\mathrm{Tr}(CA + DB) : |C^*|^q + |D^*|^q = I\} \\ &= \max\{\mathrm{Tr}(|CA| + |DB|) : C,D \geqslant 0,\ C^q + D^q = I\}.\end{aligned}$$

Assume $1 < p \leqslant 2$ and so $2 \leqslant q < \infty$. When $C,D \geqslant 0$ and $C^q + D^q \leqslant I$, if we set $C_1 = (I - D^q)^{1/q}$, then $C_1^q + D^q = I$ and $C^q \leqslant C_1^q$ so that $C^2 \leqslant C_1^2$ by operator monotonity of $t^{2/q}$. Hence $|CA|^2 = AC^2A \leqslant AC_1^2A = |C_1A|^2$ so that $\mathrm{Tr}\,|CA| \leqslant \mathrm{Tr}\,|C_1A|$. Thanks to the equivalent expressions of $V_p(A,B)$ and $\tilde{V}_p(A,B)$ this implies that $V_p(A,B) = \tilde{V}_p(A,B)$. □

The following is our trace inequality of Hölder type, where C,D as well as A,B are restricted to positive semi-definite matrices.

THEOREM 2.2. *Let* $1 < p,q < \infty$ *and* $1/p + 1/q = 1$. *If* $C,D \geqslant 0$ *and* $C^q + D^q \leqslant I$, *then*

$$\mathrm{Tr}(A^p + B^p)^{1/p} \geqslant \mathrm{Tr}(CA + DB)$$

for every $A,B \geqslant 0$.

Proof. When $C,D \geqslant 0$ and $C^q + D^q \leqslant I$, if $C_1 = (I - D^q)^{1/q}$, then $C_1^q + D^q = I$ and $C^q \leqslant C_1^q$, implying $C \leqslant C_1$ by operator monotonity of $t^{1/q}$ and so $\mathrm{Tr}\,CA \leqslant \mathrm{Tr}\,C_1A$. Hence we may assume that $C^q + D^q = I$ so that C,D are commuting. Since Tr is invariant under unitary similarity, it suffices to assume that

$$\begin{aligned}C &= \mathrm{diag}(\alpha_1^{1/q}, \ldots, \alpha_n^{1/q}), \\ D &= \mathrm{diag}((1-\alpha_1)^{1/q}, \ldots, (1-\alpha_n)^{1/q}),\end{aligned}$$

where $0 \leqslant \alpha_i \leqslant 1$ $(i = 1, \dots, n)$. With the canonical basis $e_1, \dots, e_n$ we have

$$\begin{aligned}\mathrm{Tr}(CA + DB) &= \sum_{i=1}^{n} \{ \langle Ae_i, Ce_i \rangle + \langle Be_i, De_i \rangle \} \\ &= \sum_{i=1}^{n} \{ \alpha_i^{1/q} \langle Ae_i, e_i \rangle + (1-\alpha_i)^{1/q} \langle Be_i, e_i \rangle \} \\ &\leqslant \sum_{i=1}^{n} \langle (A^p + B^p)^{1/p} e_i, e_i \rangle \\ &= \mathrm{Tr}(A^p + B^p)^{1/p}.\end{aligned}$$

Lemma 1.1 was used for the above inequality. □

COROLLARY 2.3. *If $A = [a_{ij}] \geqslant 0$ and $B = [b_{ij}] \geqslant 0$, then for every $1 < p < \infty$*

$$\mathrm{Tr}(A^p + B^p)^{1/p} \geqslant \sum_{i=1}^{n} (a_{ii}^p + b_{ii}^p)^{1/p}.$$

Proof. The proof of Theorem 2.2 implies that

$$\mathrm{Tr}(A^p + B^p)^{1/p} \geqslant \sum_{i=1}^{n} \{ \alpha_i^{1/q} a_{ii} + (1-\alpha_i)^{1/q} b_{ii} \} \tag{2.6}$$

for all $0 \leqslant \alpha_i \leqslant 1$ $(i = 1, \dots, n)$. Take the maximum on the right-hand side of (2.6) for $0 \leqslant \alpha_i \leqslant 1$ to obtain the inequality required. □

Indeed, the above corollary can be stated in a bit stronger form. To do so, we need to introduce the notion of weak majorization. Let $\vec{a} = (a_1, \dots, a_n)$ and $\vec{b} = (b_1, \dots, b_n)$ be real vectors. We say that $\vec{a}$ weakly majorizes $\vec{b}$, written as $\vec{a} \succ_w \vec{b}$, if the inequalities

$$\sum_{i=1}^{k} a_i^{\downarrow} \geqslant \sum_{i=1}^{k} b_i^{\downarrow} \tag{2.7}$$

hold for all $k = 1, \dots, n$, where $a_1^{\downarrow} \geqslant \cdots \geqslant a_n^{\downarrow}$ and $b_1^{\downarrow} \geqslant \cdots \geqslant b_n^{\downarrow}$ are the decreasing rearrangements of the components of $\vec{a}$ and $\vec{b}$. Also, we write $\vec{a} \succ^w \vec{b}$ if $-\vec{a} \succ_w -\vec{b}$, equivalently

$$\sum_{i=1}^{k} a_{n-i+1}^{\downarrow} \leqslant \sum_{i=1}^{k} b_{n-i+1}^{\downarrow} \quad \text{for } k = 1, \dots, n.$$

Furthermore, it is said that $\vec{a}$ majorizes $\vec{b}$, written as $\vec{a} \succ \vec{b}$, if $\vec{a} \succ_w \vec{b}$ and equality holds in (2.7) for $k = n$. An important fact is that $\vec{a} \succ_w \vec{b}$ (resp. $\vec{a} \succ^w \vec{b}$) implies $\sum_{i=1}^{n} f(a_i) \geqslant \sum_{i=1}^{n} f(b_i)$ (resp. $\sum_{i=1}^{n} f(a_i) \leqslant \sum_{i=1}^{n} f(b_i)$) for any increasing convex (resp. concave) function f on an interval containing all a_i, b_i (see [13] and [2]).

For a Hermitian matrix H let

$$\vec{\lambda}(H) = (\lambda_1(H), \dots, \lambda_n(H))$$

denote the eigenvalues (with multiplicities) of H in decreasing order, i.e. $\lambda_1(H) \geqslant \cdots \geqslant \lambda_n(H)$. Many (weak) majorization results are known so far for eigenvalues and singular values of matrices (see [13] and [2, 3]), which are useful in deriving various norm and trace inequalities for matrices.

Now a stronger form of Corollary 2.3 is the following:

COROLLARY 2.4. *If $A, B \geqslant 0$, then for every $1 < p < \infty$*

$$\vec{\lambda}((A^p + B^p)^{1/p}) \succ_w ((a_{11}^p + b_{11}^p)^{1/p}, \ldots, (a_{nn}^p + b_{nn}^p)^{1/p}).$$

Proof. For any $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ we have

$$\begin{aligned} \sum_{j=1}^k (a_{i_j i_j}^p + b_{i_j i_j}^p)^{1/p} &\leqslant \sum_{j=1}^k \langle (A^p + B^p)^{1/p} e_{i_j}, e_{i_j} \rangle \\ &\leqslant \sum_{i=1}^k \lambda_i((A^p + B^p)^{1/p}). \end{aligned}$$

In the above, the first inequality follows from the proofs of Theorem 2.2 and Corollary 2.3, and the latter is due to the well-known theorem of Ky Fan [13, p. 511]. Hence the stated weak majorization is obtained. □

This corollary shows that the inequality

$$\operatorname{Tr} f((A^p + B^p)^{1/p}) \geqslant \sum_{i=1}^n f((a_{ii}^p + b_{ii}^p)^{1/p})$$

holds for every $1 < p < \infty$ whenever f is an increasing convex function on $[0, \infty)$. In particular,

$$\operatorname{Tr}(A^p + B^p)^r \geqslant \sum_{i=1}^n (a_{ii}^p + b_{ii}^p)^r$$

for every $1 < p < \infty$ and $1/p \leqslant r < \infty$ (the case $r = 1/p$ is Corollary 2.3).

Although we are mostly concerned with $\operatorname{Tr}(A^p + B^p)^{1/p}$ for $p > 1$, the above theorem and corollaries can be reversed when $1/2 \leqslant p < 1$.

THEOREM 2.5. *Let $1/2 \leqslant p < 1$ and $q \leqslant -1$ with $1/p + 1/q = 1$. Let $A = [a_{ij}] \geqslant 0$ and $B = [b_{ij}] \geqslant 0$. If $C, D \geqslant 0$ are invertible and $C^q + D^q \leqslant I$, then*

$$\operatorname{Tr}(A^p + B^p)^{1/p} \leqslant \operatorname{Tr}(CA + DB).$$

Furthermore,

$$\vec{\lambda}((A^p + B^p)^{1/p}) \succ^w ((a_{11}^p + b_{11}^p)^{1/p}, \ldots, (a_{nn}^p + b_{nn}^p)^{1/p}). \tag{2.8}$$

Proof. First, note the elementary fact that for any $a, b \geqslant 0$ and $0 < p < 1$

$$(a^p + b^p)^{1/p} = \inf_{0<\alpha<1} \{\alpha^{1/q} a + (1-\alpha)^{1/q} b\}.$$

Since $t^{1/p}$ is operator convex thanks to $1 < 1/p \leqslant 2$, the reversed version of Lemma 1.1 holds as follows:

$$\begin{aligned}(A^p + B^p)^{1/p} &= \{\alpha(\alpha^{-1/p}A)^p + (1-\alpha)((1-\alpha)^{-1/p}B)^p\}^{1/p} \\ &\leqslant \alpha^{1/q}A + (1-\alpha)^{1/q}B \qquad (2.9)\end{aligned}$$

for all $0 < \alpha < 1$.

Let $C, D \geqslant 0$ be invertible and $C^q + D^q \leqslant I$. If $C_1 = (I - D^q)^{1/q}$, then $C_1^q + D^q = I$ and $C \geqslant C_1$ so that $\mathrm{Tr}\, CA \geqslant \mathrm{Tr}\, C_1 A$. Hence, as in the proof of Theorem 2.2, we may assume that

$$\begin{aligned}C &= \mathrm{diag}(\alpha_1^{1/q}, \dots, \alpha_n^{1/q}), \\ D &= \mathrm{diag}((1-\alpha_1)^{1/q}, \dots, (1-\alpha_n)^{1/q}),\end{aligned}$$

where $0 < \alpha_i < 1$ ($i = 1, \dots, n$). Then

$$\begin{aligned}\mathrm{Tr}(CA + DB) &= \sum_{i=1}^{n} \{\alpha_i^{1/q}\langle Ae_i, e_i\rangle + (1-\alpha_i)^{1/q}\langle Be_i e_i\rangle\} \\ &\geqslant \mathrm{Tr}(A^p + B^p)^{1/p}\end{aligned}$$

by (2.9). Next, for any $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ we have

$$\begin{aligned}\sum_{j=1}^{k}(a_{i_j i_j}^p + b_{i_j i_j}^p)^{1/p} &= \inf_{0<\alpha_j<1} \sum_{j=1}^{k} \{\alpha_j^{1/q}\langle Ae_{i_j}, e_{i_j}\rangle + (1-\alpha_j)^{1/q}\langle Be_{i_j}, e_{i_j}\rangle\} \\ &\geqslant \sum_{j=1}^{k} \langle (A^p + B^p)^{1/p} e_{i_j}, e_{i_j}\rangle \\ &\geqslant \sum_{i=1}^{k} \lambda_{n-i+1}((A^p + B^p)^{1/p})\end{aligned}$$

by the theorem of Ky Fan. This means the weak majorization (2.8). □

When $1/2 \leqslant p < 1$ and $0 < r \leqslant 1/p$, apply (2.8) to the increasing concave function t^{pr} on $[0, \infty)$ to get

$$\mathrm{Tr}((A^p + B^p)^r) \leqslant \sum_{i=1}^{n} (a_{ii}^p + b_{ii}^p)^r.$$

We do not know whether Theorem 2.5 is true or not when $0 < p < 1/2$.

In the rest of this section let us consider some other variational expressions and their relations to $\mathrm{Tr}(A^p + B^p)^{1/p}$ in particular when $1 < p \leqslant 2$.

LEMMA 2.6. *Let $1 < p \leqslant 2$ and $2 \leqslant q < \infty$ with $1/p + 1/q = 1$, and $A, B \geqslant 0$. Then there exist C, D such that*

$$CA + DB = (A^p + B^p)^{1/p} \qquad (2.10)$$

and

$$\alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^* \leqslant I \ \textit{for all}\ \ 0 \leqslant \alpha \leqslant 1. \tag{2.11}$$

Proof. Let

$$C = (A^p + B^p)^{-1/q}A^{p-1}, \quad D = (A^p + B^p)^{-1/q}B^{p-1}.$$

Here, when $A^p + B^p$ is not invertible, $(A^p + B^p)^{-1/q}$ is defined as $(A^p + B^p)^{-1/q}P$ where P is the projection onto the range of $A^p + B^p$. Then (2.10) is clear. For any $0 \leqslant \alpha \leqslant 1$ we have

$$\begin{aligned}
&\alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^* \\
&\quad = (A^p + B^p)^{-1/q}\{\alpha^{1-2/q}A^{2(p-1)} + (1-\alpha)^{1-2/q}B^{2(p-1)}\}(A^p + B^p)^{-1/q} \\
&\quad \leqslant (A^p + B^p)^{-1/q}\{(A^{2(p-1)})^{q/2} + (B^{2(p-1)})^{q/2}\}^{2/q}(A^p + B^p)^{-1/q} \\
&\quad \leqslant I.
\end{aligned}$$

Above we used Lemma 1.1. □

THEOREM 2.7. *Let* $1 < p \leqslant 2$ *and* $2 \leqslant q < \infty$ *with* $1/p + 1/q = 1$. *Then for every* $A, B \geqslant 0$,

$$\begin{aligned}
&\max\{\mathrm{Tr}(CA + DB) : C, D \geqslant 0,\ C^q + D^q \leqslant I\} \\
&\quad \leqslant \begin{cases} \mathrm{Tr}(A^p + B^p)^{1/p} \\ V_p(A,B) \end{cases} \\
&\quad \leqslant \max\{\mathrm{Tr}\,|CA + DB| : \alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^* \leqslant I\,(0 \leqslant \alpha \leqslant 1)\} \\
&\quad \leqslant \min\left\{\sum_{i=1}^{n}(\|Ae_i\|^p + \|Be_i\|^p)^{1/p} : \{e_i\}\ \textit{is an orthonormal basis}\right\}.
\end{aligned}$$

Proof. We denote the two maximum expressions and the minimum expression in the theorem by M_1, M_2 and M_3, respectively. Theorem 2.2 means $M_1 \leqslant \mathrm{Tr}(A^p + B^p)^{1/p}$, and $M_1 \leqslant V_p(A,B)$ is trivial. Lemma 2.6 implies $\mathrm{Tr}(A^p + B^p)^{1/p} \leqslant M_2$. Since by Lemma 1.1

$$(|C^*|^q + |D^*|^q)^{2/q} \geqslant \alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^* \ \text{ for } 0 \leqslant \alpha \leqslant 1,$$

we get $V_p(A,B) \leqslant M_2$. As in the proof of Proposition 2.1 it is immediately seen that M_2 is equal to the maximum of $|\,\mathrm{Tr}(CA + DB)|$ over C, D satisfying (2.11). Moreover, condition (2.11) means that

$$\alpha^{1-2/q}\|C^*e\|^2 + (1-\alpha)^{1-2/q}\|D^*e\| \leqslant 1$$

for all $0 \leqslant \alpha \leqslant 1$ and $\|e\| = 1$, which is equivalent to

$$\|C^*e\|^q + \|D^*e\|^q \leqslant 1 \ \text{ for } \|e\| = 1. \tag{2.12}$$

When C, D satisfy (2.11) and $\{e_i\}$ is an orthonormal basis, we have

$$
\begin{aligned}
|\operatorname{Tr}(CA + DB)| &= \left| \sum_{i=1}^{n} \{ \langle Ae_i, C^* e_i \rangle + \langle Be_i, D^* e_i \rangle \} \right| \\
&\leqslant \sum_{i=1}^{n} (\|Ae_i\| \, \|C^* e_i\| + \|Be_i\| \, \|D^* e_i\|) \\
&\leqslant \sum_{i=1}^{n} (\|Ae_i\|^p + \|Be_i\|^p)^{1/p} (\|C^* e_i\|^q + \|D^* e_i\|^q)^{1/q} \\
&\leqslant \sum_{i=1}^{n} (\|Ae_i\|^p + \|Be_i\|^p)^{1/p}
\end{aligned}
$$

by (2.12). This completes the proof. □

The above theorem implies in particular that

$$
\operatorname{Tr}(A^p + B^p)^{1/p} \leqslant \sum_{i=1}^{n} (\|Ae_i\|^p + \|Be_i\|^p)^{1/p} \tag{2.13}
$$

for any orthonormal basis $\{e_i\}$. Indeed, this can be strengthened to a weak majorization as Corollary 2.4 and Theorem 2.5.

PROPOSITION 2.8. *Let $A, B \geqslant 0$ and $\{e_i\}$ be an orthonormal basis. Then for every $1 < p \leqslant 2$,*

$$
\begin{aligned}
&\vec{\lambda}((A^p + B^p)^{1/p}) \\
&\quad \succ^w ((\|Ae_1\|^p + \|Be_1\|^p)^{1/p}, \dots, (\|Ae_n\|^p + \|Be_n\|^p)^{1/p}).
\end{aligned} \tag{2.14}
$$

Proof. Set a linear map Φ on the $n \times n$ matrices into itself by

$$
\Phi(X) = \operatorname{diag}(\langle Xe_1, e_1 \rangle, \dots, \langle Xe_n, e_n \rangle).
$$

Then Φ is a doubly stochastic map on the $n \times n$ matrices; namely Φ is positive, $\Phi(I) = I$, and preserves Tr. In fact, Φ can be written as $\Phi(X) = \sum_{j=1}^{k} \alpha_j U_j X U_j^*$ where $\alpha_j > 0$, $\sum_{j=1}^{k} \alpha_j = 1$, and U_j are unitaries. Since $t^{2/p}$ is operator convex, we have for $X \geqslant 0$

$$
\begin{aligned}
\Phi(X^p)^{2/p} &= \left(\sum_{j=1}^{k} \alpha_j U_j X^p U_j^* \right)^{2/p} \\
&\leqslant \sum_{j=1}^{k} \alpha_j (U_j X^p U_j^*)^{2/p} = \Phi(X^2),
\end{aligned}
$$

and hence thanks to operator monotonity of $t^{p/2}$

$$
\Phi(X^p) \leqslant \Phi(X^2)^{p/2}. \tag{2.15}
$$

For $X \geqslant 0$, since $\vec{\lambda}(X) \succ \vec{\lambda}(\Phi(X))$ (see [2, Theorem 7.1]) and $t^{1/p}$ is concave, we get

$$\vec{\lambda}(X^{1/p}) \succ^w \vec{\lambda}(\Phi(X))^{1/p}.$$

Therefore for $A, B \geqslant 0$,

$$\begin{aligned}\vec{\lambda}((A^p+B^p)^{1/p}) &\succ^w \vec{\lambda}(\Phi(A^p+B^p))^{1/p} \\ &\leqslant \vec{\lambda}(\Phi(A^2)^{p/2}+\Phi(B^2)^{p/2})^{1/p}\end{aligned}$$

by (2.15). The latter vector is nothing but the right-hand side of (2.14) (up to a component permutation), and we have the result because $\vec{a} \succ^w \vec{b} \leqslant \vec{c}$ implies $\vec{a} \succ^w \vec{c}$. □

The above weak majorization implies in particular that

$$\operatorname{Tr}((A^p+B^p)^r) \leqslant \sum_{i=1}^{n} (\|Ae_i\|^p + \|Be_i\|^p)^r$$

whenever $1 < p \leqslant 2$ and $0 < r \leqslant 1/p$.

When $p > 2$, inequality (2.13) is false even when A, B are commuting. In fact, let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$e_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad e_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

Then

$$\operatorname{Tr}(A^p+B^p)^{1/p} = \operatorname{Tr} I = 2$$

while

$$\sum_{i=1}^{2} (\|Ae_i\|^p + \|Be_i\|^p)^{1/p} = 2^{1/2+1/p}.$$

3. Counter-examples from two projections

The aim of this section is to give counter-examples which show that Hölder inequalities as stated in (1.3) and (2.1) are false. To do so, we take as A, B two non-commuting orthogonal projections of rank one and examine the validity of those inequalities in detail.

Now let

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad Q_t = \begin{bmatrix} t^2 & t\sqrt{1-t^2} \\ t\sqrt{1-t^2} & 1-t^2 \end{bmatrix}$$

for $0 < t < 1$. Then P and Q_t are rank one projections and

$$(P^p + Q_t^p)^{1/p} = (P + Q_t)^{1/p}.$$

The diagonalization of $P+Q_t$ is

$$P+Q_t = U_t \operatorname{diag}(1+t, 1-t)\, U_t$$

where

$$U_t = \begin{bmatrix} \sqrt{\frac{1+t}{2}} & \sqrt{\frac{1-t}{2}} \\ \sqrt{\frac{1-t}{2}} & -\sqrt{\frac{1+t}{2}} \end{bmatrix},$$

a Hermitian unitary. Set

$$C_t = (P+Q_t)^{-1/2}P\,, \quad D_t = (P+Q_t)^{-1/2}Q_t\,,$$

so that $C_tP + D_tQ_t = (P+Q_t)^{1/2}$.

The next proposition says that matrix Hölder inequality (1.3) is false for any $1 < p < \infty$ except 2.

PROPOSITION 3.1. *For every* $0 < q < \infty$*,*

$$|C_t^*|^q + |D_t^*|^q = I\,. \tag{3.1}$$

But for every $0 < p < \infty$ *with* $p \neq 2$ *and* $0 < t < 1$*,*

$$(P+Q_t)^{1/p} \not\geqslant |C_tP + D_tQ_t|\,.$$

Proof. A direct computation gives

$$U_tC_tC_t^*U_t = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad U_tD_tD_t^*U_t = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}.$$

Hence $|C_t^*|$ and $|D_t^*|$ are projections such that $|C_t^*| + |D_t^*| = I$ and (3.1) holds for any $q > 0$. On the other hand, the eigenvalues of $(P+Q_t)^{1/p} - |C_tP + D_tQ_t|$ are

$$(1 \pm t)^{1/p} - (1 \pm t)^{1/2}\,,$$

one of which is strictly negative for any $p > 0$ with $p \neq 2$. □

Next, to examine trace inequality (2.1), for $1 \leqslant p \leqslant \infty$ and $0 < t < 1$ put

$$f_t(p) = \operatorname{Tr}(P^p + Q_t^p)^{1/p} = \operatorname{Tr}(P+Q_t)^{1/p},$$

and compare it with $V_p(P, Q_t)$ (or $\tilde{V}_p(P, Q_t)$) defined in (2.2)–(2.4). The following are clear by (2.4) and (2.5):

$$f_t(1) = f_t(\infty) = V_1(P, Q_t) = 2\,, \tag{3.2}$$

$$f_t(2) = V_2(P, Q_t)\,. \tag{3.3}$$

LEMMA 3.2. *$V_p(P,Q_t)$ is decreasing in $1 \leqslant p \leqslant 2$ and $V_p(P,Q_t) = \tilde{V}_p(P,Q_t) = V_2(P,Q_t)$ for all $2 \leqslant p \leqslant \infty$.*

Proof. When $1 < p < p' \leqslant \infty$ and $1/p' + 1/q' = 1$ as well as $1/p + 1/q = 1$, since $|C^*|^{q'} + |D^*|^{q'} \leqslant I$ implies $|C^*|^q + |D^*|^q \leqslant I$, we get $V_p(P,Q_t) \geqslant V_{p'}(P,Q_t)$. Hence $V_p(P,Q_t)$ is decreasing in $1 \leqslant p \leqslant \infty$. Moreover we have by (3.1) and (3.3)

$$V_p(P,Q_t) \geqslant \tilde{V}_p(P,Q_t) \geqslant \operatorname{Tr}|C_tP + D_tQ_t| = V_2(P,Q_t)$$

for all $1 \leqslant p \leqslant \infty$. $\square$

LEMMA 3.3. *There exist $2 < p_0(t) < p_1(t) < \infty$ such that*

(*i*) *$f_t(p)$ is strictly decreasing in $p \in [1, p_0(t)]$ and strictly increasing in $p \in [p_0(t), \infty]$,*

(*ii*) *$f_t(2) = f_t(p_1(t))$,*

(*iii*) *$p_0(t) \to \infty$ as $t \to 1$ and $p_1(t) \to 2$ as $t \to 0$.*

Proof. For simplicity write

$$g_t(x) = f_t(1/x) = (1+t)^x + (1-t)^x$$

for $x \geqslant 0$ and $0 < t < 1$. Since

$$g_t'(x) = (1+t)^x \log(1+t) + (1-t)^x \log(1-t),$$

we get

$$g_t'(0) = \log(1-t^2) < 0.$$

To show $g_t'(1/2) > 0$ let

$$\phi(t) = (1+t)^{1/2}\log(1+t) + (1-t)^{1/2}\log(1-t) \quad \text{for } 0 \leqslant t < 1.$$

By differentiating ϕ up to the fourth degree one can easily show that $\phi(t) > \phi(0) = 0$ for $0 < t < 1$. Hence $g_t'(1/2) > 0$. Since

$$g_t''(x) = (1+t)(\log(1+t))^2 + (1-t)^x(\log(1-t))^2 > 0,$$

there exists a unique $x_0 \in (0, 1/2)$ (depending on t) such that $g_t'(x_0) = 0$. Since $g_t(0) = g_t(1) = 2$ by (3.2), there exists a unique $x_1 \in (0, x_0)$ (depending on t) such that $g_t(x_1) = g_t(1/2)$. Then (*i*) and (*ii*) hold with $p_0(t) = 1/x_0$ and $p_1(t) = 1/x_1$. Now it remains to show that $x_0 \to 0$ as $t \to 1$ and $x_1 \to 1/2$ as $t \to 0$. The first is seen because $\lim_{t\to 1} g_t'(x) > 0$ for any $x > 0$. To see the second, take the Taylor expansions in t up to the second order of both sides of the equation

$$(1+t)^{x_1} + (1-t)^{x_1} = (1+t)^{1/2} + (1-t)^{1/2}.$$

Though x_1 depends on t, since the Taylor coefficients of the left-hand side are uniformly bounded for $0 < x_1 < 1/2$, we get

$$x_1(x_1-1)t^2 + O(t^3) = -\frac{1}{4}t^2 + O(t^3)$$

and hence $(x_1 - 1/2)^2 = O(t)$. This implies that $x_1 \to 1/2$ as $t \to 0$. $\square$

By Lemmas 3.2 and 3.3 together with (3.3) we have

$$\begin{cases} f_t(p) < \tilde{V}_p(P, Q_t) & \text{if } 2 < p < p_1(t), \\ f_t(p) > V_p(P, Q_t) & \text{if } p_1(t) < p \leqslant \infty. \end{cases}$$

Thanks to Lemma 3.3 *(iii)* this shows the following:

PROPOSITION 3.4. *For each* $2 < p < \infty$*, both* $\mathrm{Tr}(P + Q_t)^{1/p} < \tilde{V}_p(P, Q_t)$ *and* $\mathrm{Tr}(P + Q_t)^{1/p} > V_p(P, Q_t)$ *can occur when* t *varies in* $(0, 1)$.

In this way trace inequality (2.1) does not hold when $2 < p < \infty$.

Though we could not settle (2.1) for $1 < p < 2$, numerical computations strongly suggest that

$$\mathrm{Tr}(P + Q_t)^{1/p} > V_p(P, Q_t) \ \text{ for } \ 1 < p < 2.$$

The next proposition is concerned with another variational expression in Theorem 2.7.

PROPOSITION 3.5. *For each* $1 < p < 2$*, strict inequality*

$$\begin{aligned} &\mathrm{Tr}(P + Q_t)^{1/p} \\ &\quad < \max\{\mathrm{Tr}\,|CP + DQ_t| : \alpha^{1-2/q}CC^* + (1-\alpha)^{1-2/q}DD^* \leqslant I\,(0 \leqslant \alpha \leqslant 1)\} \end{aligned}$$

can occur when t *varies in* $(0, 1)$.

Proof. Let $1 < p < 2$ and $1/p + 1/q = 1$. By Lemma 2.6 and its proof, if we set for $0 < t < 1$

$$C_t^{(p)} = (P + Q_t)^{-1/q}P, \quad D_t^{(p)} = (P + Q_t)^{-1/q}Q_t,$$

then $C_t^{(p)}P + D_t^{(p)}Q_t = (P + Q_t)^{1/p}$ and $C_t^{(p)}, D_t^{(p)}$ satisfy (2.11), so that

$$\mathrm{Tr}(P + Q_t)^{1/p} = \mathrm{Tr}(C_t^{(p)}P + D_t^{(p)}Q_t).$$

Now define

$$F_t(x) = \mathrm{Tr}(C_x^{(p)}P + D_x^{(p)}Q_t) \ \text{ for } \ 0 < x < 1.$$

Since $\mathrm{Tr}(P + Q_t)^{1/p} = F_t(t)$, the result follows if $F_t'(t) \neq 0$ is shown. A simple but a bit tedious computation yields

$$2F_t'(t) = \left(\frac{1}{p} - \frac{1}{q}\right)\{(1+t)^{-1/q} - (1-t)^{-1/q}\}$$

which is strictly negative due to $p < 2 < q$. $\square$

One may expect to get more information from two projections of higher rank. But this is not so. In fact, when P, Q are two projections such that $P \vee Q = I$ and $P \wedge Q = 0$, it is readily seen from the structure theorem for two projections (see [15, pp. 306–308]) that the eigenvalues of $P + Q$ except 1 are $1 \pm \alpha_1, \ldots, 1 \pm \alpha_k$ with

$0 < \alpha_i < 1$. So we can have no better results than the above propositions from two projections.

We end this section with a brief discussion about joint convexity of the function $(A,B) \mapsto \mathrm{Tr}(A^p + B^p)^{1/p}$. Note that variational expressions such as (2.2) are jointly convex in (A,B) because so is $\mathrm{Tr}\,|CA + DB|$ (the trace norm of $CA + DB$). Our motivation of comparing $\mathrm{Tr}(A^p + B^p)^{1/p}$ with such variational expressions in Sec. 2 came from the search for joint convexity of $\mathrm{Tr}(A^p + B^p)^{1/p}$, namely

$$\mathrm{Tr}((A_1 + A_2)^p + (B_1 + B_2)^p)^{1/p} \leqslant \mathrm{Tr}(A_1^p + B_1^p)^{1/p} + \mathrm{Tr}(A_2^p + B_2^p)^{1/p} \tag{3.4}$$

for $A_i, B_i \geqslant 0$ ($i = 1,2$). However, the proposition below says that this is negative for any $2 < p \leqslant \infty$, while of course it is true for $p = 2$ by (2.5) and trivial for $p = 1$. So it may be conjectured for $1 < p < 2$ only. But the negative results shown above in this section as well as numerical computations suggest us that the way of proof via variational expression is hopeless.

PROPOSITION 3.6. *Let*

$$A_1 = B_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2\varepsilon & 0 \\ 0 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 \\ 0 & 2\varepsilon \end{bmatrix}$$

for $\varepsilon > 0$. *If* $2 < p \leqslant \infty$, *then (3.4) fails to hold for* ε *small enough, that is,*

$$\mathrm{Tr}((A_1 + A_2)^p + (B_1 + B_2)^p)^{1/p} > \mathrm{Tr}(A_1^p + B_1^p)^{1/p} + \mathrm{Tr}(A_2^p + B_2^p)^{1/p}. \tag{3.5}$$

Proof. First let $p < \infty$. The right-hand side of (3.5) can be immediately computed and it is equal to $2^{1+1/p} + 4\varepsilon$. The eigenvalues of $A_1 + A_2$ (also $B_1 + B_2$) are

$$\alpha_1 = 1 + \varepsilon + \sqrt{1 + \varepsilon^2}, \quad \alpha_2 = 1 + \varepsilon - \sqrt{1 + \varepsilon^2},$$

and $A_1 + A_2$ is diagonalized as $A_1 + A_2 = U\,\mathrm{diag}(\alpha_1, \alpha_2)\,U$ with a Hermitian unitary

$$U = \frac{1}{\sqrt{2}(1 + \varepsilon^2)^{1/4}} \begin{bmatrix} \{(1 + \varepsilon^2)^{1/2} + \varepsilon\}^{1/2} & \{(1 + \varepsilon^2)^{1/2} - \varepsilon\}^{1/2} \\ \{(1 + \varepsilon^2)^{1/2} - \varepsilon\}^{1/2} & -\{(1 + \varepsilon^2)^{1/2} + \varepsilon\}^{1/2} \end{bmatrix}.$$

Hence we get

$$(A_1 + A_2)^p = \begin{bmatrix} \frac{\alpha_1^p + \alpha_2^p}{2} + \frac{\varepsilon(\alpha_1^p - \alpha_2^p)}{2\sqrt{1+\varepsilon^2}} & \frac{\alpha_1^p - \alpha_2^p}{2\sqrt{1+\varepsilon^2}} \\ \frac{\alpha_1^p - \alpha_2^p}{2\sqrt{1+\varepsilon^2}} & \frac{\alpha_1^p + \alpha_2^p}{2} - \frac{\varepsilon(\alpha_1^p - \alpha_2^p)}{2\sqrt{1+\varepsilon^2}} \end{bmatrix}.$$

Also $(B_1 + B_2)^p$ is analogously computed, where the diagonal entries are just exchanged. Therefore

$$(A_1 + A_2)^p + (B_1 + B_2)^p = \begin{bmatrix} \alpha_1^p + \alpha_2^p & \frac{\alpha_1^p - \alpha_2^p}{\sqrt{1+\varepsilon^2}} \\ \frac{\alpha_1^p - \alpha_2^p}{\sqrt{1+\varepsilon^2}} & \alpha_1^p + \alpha_2^p \end{bmatrix}$$

and its eigenvalues are

$$\beta_1 = \alpha_1^p + \alpha_2^p + \frac{\alpha_1^p - \alpha_2^p}{\sqrt{1 + \varepsilon^2}}, \quad \beta_2 = \alpha_1^p + \alpha_2^p - \frac{\alpha_1^p - \alpha_2^p}{\sqrt{1 + \varepsilon^2}}.$$

Since

$$\beta_1^{1/p} \geqslant \left(1+\frac{1}{\sqrt{1+\varepsilon^2}}\right)^{1/p}\alpha_1 \geqslant 2^{1+1/p}+O(\varepsilon)\,,$$

$$\beta_2^{1/p} \geqslant \left(1-\frac{1}{\sqrt{1+\varepsilon^2}}\right)^{1/p}\alpha_1 \geqslant 2^{1-1/p}\varepsilon^{2/p}+O(\varepsilon)\,,$$

the left-hand side of (3.5) is bounded below by $2^{1+1/p}+2^{1-1/p}\varepsilon^{2/p}+O(\varepsilon)$. This proves (3.5) for small ε; hence (3.4) does not hold for $2<p<\infty$.

Next when $p=\infty$, since $A_2\vee B_2=2\varepsilon I$, (3.5) means that

$$\mathrm{Tr}((A_1+A_2)\vee(B_1+B_2))>2+4\varepsilon\,. \tag{3.6}$$

Set $A=(A_1+A_2)\vee(B_1+B_2)$ and $A-A_1=\begin{bmatrix} a & c\\ \bar{c} & b\end{bmatrix}$. Since $A-A_1\geqslant A_2$ and $A-A_1=A-B_1\geqslant B_2$, we get

$$a\geqslant 2\varepsilon\,,\quad (a-2\varepsilon)b\geqslant |c|^2\,,$$

$$b\geqslant 2\varepsilon\,,\quad a(b-2\varepsilon)\geqslant |c|^2\,.$$

If (3.6) does not hold, then $a+b=\mathrm{Tr}(A-A_1)\leqslant 4\varepsilon$, so that $a=b=2\varepsilon$ and $c=0$. Hence $A=A_1+2\varepsilon I$. Moreover we have $A^2\geqslant(B_1+B_2)^2$ by definition of $\vee$ in [9], which implies $\begin{bmatrix} 4\varepsilon+4\varepsilon^2 & 2\varepsilon\\ 2\varepsilon & 0\end{bmatrix}\geqslant 0$, a contradiction. So (3.6) holds for any $\varepsilon>0$. $\square$

For the operator norm $\|(A^p+B^p)^{1/p}\|$ we have the following positive result.

PROPOSITION 3.7. *If* $1\leqslant p\leqslant 2$, *then* $\|(A^p+B^p)^{1/p}\|$ *is jointly convex in* $A,B\geqslant 0$.

Proof. Let $A_i,B_i\geqslant 0$ ($i=1,2$) and show that

$$\|((A_1+A_2)^p+(B_1+B_2)^p)^{1/p}\|\leqslant\|(A_1^p+B_1^p)^{1/p}\|+\|(A_2^p+B_2^p)^{1/p}\|\,. \tag{3.7}$$

Put $a_i=\|(A_i^p+B_i^p)^{1/p}\|$ and $\alpha=a_1/(a_1+a_2)$. Then a_i is the largest eigenvalue of $(A_i^p+B_i^p)^{1/p}$. Since t^p is operator convex and $A_i^p+B_i^p\leqslant a_i^pI$, we get as (2.9)

$$\begin{aligned}(A_1+A_2)^p+(B_1+B_2)^p &\leqslant \alpha^{1-p}(A_1^p+B_1^p)+(1-\alpha)^{1-p}(A_2^p+B_2^p)\\ &\leqslant\{\alpha^{1-p}a_1^p+(1-\alpha)^{1-p}a_2^p\}I\\ &=(a_1+a_2)^pI\,,\end{aligned}$$

implying (3.7). $\square$

4. Multiplicative weak majorizations of Hölder type

The aim of this section is to search how we can generalize joint concavity stated in (0.3) to matrices. First the next proposition shows that the map $(A,B)\mapsto(B^{1/2q}A^{1/p}B^{1/2q})^{1/r}$ has a rather poor property of operator concavity.

PROPOSITION 4.1. *Let $p,q,r>0$. The map $A\geqslant 0\mapsto (BA^{1/p}B)^{1/r}$ is operator concave for every $B\geqslant 0$ if and only if $p,r\geqslant 1$. For every $q,r>0$ and for some $A\geqslant 0$ the map $B\mapsto (B^{1/2q}AB^{1/2q})^{1/r}$ is not operator concave.*

Proof. When $p,r\geqslant 1$ operator concavity of $A\mapsto (BA^{1/p}B)^{1/r}$ is clear because $t^{1/r}$ is operator concave and operator monotone. Conversely, suppose it is operator concave for every $B\geqslant 0$. Since

$$\left(\begin{bmatrix} I & I \\ I & I \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^{1/p}\begin{bmatrix} I & I \\ I & I \end{bmatrix}\right)^{1/r} = 2^{1/r-1}\begin{bmatrix} X & X \\ X & X \end{bmatrix}$$

where $X=(A^{1/p}+B^{1/p})^{1/r}$, it follows that $(A,B)\mapsto (A^{1/p}+B^{1/p})^{1/r}$ is jointly operator concave. Hence if P,Q are projections, then we get

$$(P+Q)^{1/r} = 2^{1/pr}\left\{\left(\frac{P+0}{2}\right)^{1/p}+\left(\frac{0+Q}{2}\right)^{1/p}\right\}^{1/r} \geqslant 2^{1/pr-1}(P+Q).$$

Since any $x\in(0,1)$ can be an eigenvalue of $P+Q$, we must have $x^{1/r-1}\geqslant 2^{1/pr-1}$ for all $0<x<1$. This imples $r\geqslant 1$. Moreover, since the numerical function $(x^{1/p}+1)^{1/r}$ is concave in $x\geqslant 0$, we get

$$\begin{aligned}&\frac{d^2}{dx^2}(x^{1/p}+1)^{1/r}\\ &\quad = \frac{1}{pr}x^{1/p-2}(x^{1/p}+1)^{1/r-2}\left\{\left(\frac{1}{pr}-1\right)x^{1/p}+\left(\frac{1}{p}-1\right)\right\}\leqslant 0\end{aligned}$$

for every $x>0$. This implies $p\geqslant 1$, and hence the first assertion is shown.

Next, let

$$A=\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix},\quad B_1=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\quad B_2=\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then

$$\lambda_2((B_1+B_2)^{1/2q}A(B_1+B_2)^{1/2q})=0$$

while

$$\lambda_2((B_1^{1/2q}AB_1^{1/2q})^{1/r}+(B_2^{1/2q}AB_2^{1/2q})^{1/r})>0.$$

These prove the second assertion. □

Now we are concerned with a Hölder type inequality for positive semi-definite matrices under taking eigenvalue products $\prod_{i=1}^{k}\lambda_{n-i+1}(\cdot)$. It is some kind of multiplicative weak majorization and also considered as a matrix version of joint concavity (0.3).

In the sequel of this paper let A,B,A_j,B_j be positive semi-definite $n\times n$ matrices. We may write $\lambda_i(AB)$ instead of $\lambda_i(B^{1/2}AB^{1/2})$ because AB is similar to $B^{1/2}AB^{1/2}$.

LEMMA 4.2. *For every* $1 \leqslant p,q < \infty$,

$$\lambda_n((A_1^p + A_2^p)^{1/p}(B_1^q + B_2^q)^{1/q})^{pq/(p+q)} \geqslant \lambda_n(A_1B_1)^{pq/(p+q)} + \lambda_n(A_2B_2)^{pq/(p+q)}. \tag{4.1}$$

Proof. Put $a_j = \lambda_n(A_jB_j)^{pq/(p+q)}$ and $\alpha = a_1/(a_1+a_2)$, $a_1 + a_2 > 0$ being assumed. Note that if $A,B \geqslant 0$ and $a \geqslant 0$, then $\lambda_n(AB) \geqslant a$ (i.e. $A^{1/2}BA^{1/2} \geqslant aI$) is equivalent to $\begin{bmatrix} A & a^{1/2}I \\ a^{1/2}I & B \end{bmatrix} \geqslant 0$. Since

$$\alpha^{2-(p+q)/pq}A_1^{1/2}B_1A_1^{1/2} \geqslant \alpha^{2-(p+q)/pq}a_1^{(p+q)/pq}I = \alpha^2(a_1+a_2)^{(p+q)/pq}I,$$

we get

$$\begin{bmatrix} \alpha^{1-1/p}A_1 & \alpha(a_1+a_2)^{(p+q)/2pq}I \\ \alpha(a_1+a_2)^{(p+q)/2pq}I & \alpha^{1-1/q}B_1 \end{bmatrix} \geqslant 0,$$

and similarly

$$\begin{bmatrix} (1-\alpha)^{1-1/p}A_2 & (1-\alpha)(a_1+a_2)^{(p+q)/2pq}I \\ (1-\alpha)(a_1+a_2)^{(p+q)/2pq}I & (1-\alpha)^{1-1/q}B_2 \end{bmatrix} \geqslant 0,$$

so that

$$\begin{bmatrix} \alpha^{1-1/p}A_1 + (1-\alpha)^{1-1/p}A_2 & (a_1+a_2)^{(p+q)/2pq}I \\ (a_1+a_2)^{(p+q)/2pq}I & \alpha^{1-1/q}B_1 + (1-\alpha)^{1-1/q}B_2 \end{bmatrix} \geqslant 0.$$

Therefore by Lemma 1.1

$$\begin{bmatrix} (A_1^p + A_2^p)^{1/p} & (a_1+a_2)^{(p+q)/2pq}I \\ (a_1+a_2)^{(p+q)/2pq}I & (B_1^q + B_2^q)^{1/q} \end{bmatrix} \geqslant 0,$$

which is equivalent to

$$\lambda_n((A_1^p + A_2^p)^{1/p}(B_1^q + B_2^q)^{1/q}) \geqslant (a_1+a_2)^{(p+q)/pq},$$

as desired. □

In the following proof we use the technique of antisymmetric tensor powers. For each $n \times n$ matrix X and $k = 1,\dots,n$, the k-fold antisymmetric tensor power (or the kth compound) of X is denoted by $\wedge^k X$. This is the restriction of the k-fold tensor product $\otimes^k X$ on $\otimes^k \mathbb{C}^n$ to the antisymmetric tensor product subspace $\wedge^k \mathbb{C}^n$ of dimension $\binom{n}{k}$. What we need are the following basic properties (see [13, 19.F] and [5, I.5]):

(*i*) $\wedge^k(XY) = (\wedge^k X)(\wedge^k Y)$ and $\wedge^k(X^*) = (\wedge^k X)^*$.

(*ii*) For $A \geqslant 0$ and $r > 0$, $\wedge^k A \geqslant 0$ and $\wedge^k(A^r) = (\wedge^k A)^r$.

(*iii*) For $A \geqslant 0$, $\lambda_1(\wedge^k A) = \prod_{i=1}^k \lambda_i(A)$ and $\lambda_N(\wedge^k A) = \prod_{i=1}^k \lambda_{n-i+1}(A)$ where $N = \binom{n}{k}$.

For every $A,B \geqslant 0$ these properties give

$$\lambda_N((\wedge^k A)(\wedge^k B)) = \prod_{i=1}^k \lambda_{n-i+1}(AB) \ \text{ for } N = \binom{n}{k}. \tag{4.2}$$

It is known [1, Corollary 6.2] (also [12]) that $A \geqslant 0 \mapsto \otimes^k A^{1/k}$ is operator concave for any $k \in \mathbb{N}$. Since $\wedge^k A^{1/k} = \otimes^k A^{1/k}|_{\wedge^k \mathbb{C}^n}$, we state another important property:

(iv) $A \geqslant 0 \mapsto \wedge^k A^{1/k}$ is operator concave.

THEOREM 4.3. *For every* $1 \leqslant p, q < \infty$ *and* $k = 1, \ldots, n$,

$$\left\{ \prod_{i=1}^{k} \lambda_{n-i+1}((A_1^p + A_2^p)^{1/pk}(B_1^q + B_2^q)^{1/qk}) \right\}^{pq/(p+q)} \geqslant \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(A_1^{1/k} B_1^{1/k}) \right\}^{pq/(p+q)} + \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(A_2^{1/k} B_2^{1/k}) \right\}^{pq/(p+q)}. \tag{4.3}$$

Proof. For $k = 1, \ldots, n$ and $N = \binom{n}{k}$, by (4.1) applied to $\wedge^k A_j^{1/k}, \wedge^k B_j^{1/k}$ instead of A_j, B_j we have

$$\lambda_N((\wedge^k A_1^{p/k} + \wedge^k A_2^{p/k})^{1/p}(\wedge^k B_1^{q/k} + \wedge^k B_2^{q/k})^{1/q})^{pq/(p+q)} \geqslant \lambda_N((\wedge^k A_1^{1/k})(\wedge^k B_1^{1/k}))^{pq/(p+q)} + \lambda_N((\wedge^k A_2^{1/k})(\wedge^k B_2^{1/k}))^{pq/(p+q)}. \tag{4.4}$$

By property (iv) stated above we get

$$\wedge^k A_1^{p/k} + \wedge^k A_2^{p/k} \leqslant \wedge^k (A_1^p + A_2^p)^{1/k},$$

$$\wedge^k B_1^{q/k} + \wedge^k B_2^{q/k} \leqslant \wedge^k (B_1^q + B_2^q)^{1/k}.$$

Thanks to operator monotonity of $t^{1/p}$ and $t^{1/q}$, it follows that $\lambda_i(A^{1/p} B^{1/q})$ is jointly increasing in $A, B \geqslant 0$. So the above inequalities imply that

$$\begin{aligned} &\lambda_N((\wedge^k A_1^{p/k} + \wedge^k A_2^{p/k})^{1/p}(\wedge^k B_1^{q/k} + \wedge^k B_2^{q/k})^{1/q}) \\ &\leqslant \lambda_N(\wedge^k (A_1^p + A_2^p)^{1/pk} \wedge^k (B_1^q + B_2^q)^{1/qk}) \\ &= \prod_{i=1}^{k} \lambda_{n-i+1}((A_1^p + A_2^p)^{1/pk}(B_1^q + B_2^q)^{1/qk}) \end{aligned}$$

by (4.2). Also by (4.2) the right-hand side of (4.4) is equal to that of (4.3). Hence we obtain (4.3). $\square$

By taking the limit of (4.3) as $p \to \infty$ we see that (4.3) holds true for $p = \infty$ and $1 \leqslant q < \infty$ too, that is,

$$\left\{ \prod_{i=1}^{k} \lambda_{n-i+1}((A_1 \vee A_2)^{1/k}(B_1^q + B_2^q)^{1/qk}) \right\}^{q} \geqslant \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(A_1^{1/k} B_1^{1/k}) \right\}^{q} + \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(A_2^{1/k} B_2^{1/k}) \right\}^{q}.$$

So, weak majorization (4.3) becomes a proper form of Hölder type when $1 \leqslant p,q \leqslant \infty$ and $1/p+1/q=1$.

Let $p=q=1/r$ and replace A_j^p, B_j^q by A_j, B_j in (4.3). Then for every $0<r\leqslant 1$ and $k=1,\dots,n$ we have the superadditivity

$$\left\{\prod_{i=1}^k \lambda_{n-i+1}((A_1+A_2)^{r/k}(B_1+B_2)^{r/k})\right\}^{1/2r} \geqslant \left\{\prod_{i=1}^k \lambda_{n-i+1}(A_1^{r/k}B_1^{r/k})\right\}^{1/2r} + \left\{\prod_{i=1}^k \lambda_{n-i+1}(A_2^{r/k}B_2^{r/k})\right\}^{1/2r}, \tag{4.5}$$

which means that when $0<r\leqslant 1$ the map $(A,B)\mapsto \{\prod_{i=1}^k \lambda_{n-i+1}(A^{r/k}B^{r/k})\}^{1/2r}$ is jointly concave (and jointly increasing). In particular, when $A_1=B_1=A$ and $A_2=B_2=B$, this reads as

$$\left\{\prod_{i=1}^k \lambda_{n-i+1}(A+B)\right\}^{1/k} \geqslant \left\{\prod_{i=1}^k \lambda_{n-i+1}(A)\right\}^{1/k} + \left\{\prod_{i=1}^k \lambda_{n-i+1}(B)\right\}^{1/k}, \tag{4.6}$$

which is the Oppenheim inequality [14] (or [13, p. 475]). Moreover, (4.5) can be stated in a formally more general form that if $f,g:[0,\infty)\to[0,\infty)$ are operator monotone (hence operator concave, [7]), then for every $0<r\leqslant 1$ and $k=1,\dots,n$ the function

$$(A,B)\mapsto \left\{\prod_{i=1}^k \lambda_{n-i+1}(f(A)^{r/k}g(B)^{r/k})\right\}^{1/2r}$$

is jointly concave. This is obvious because for $0\leqslant\alpha\leqslant 1$

$$\lambda_i(f(\alpha A_1+(1-\alpha)A_2)^{r/k}g(\alpha B_1+(1-\alpha)B_2)^{r/k}) \geqslant \lambda_i((\alpha f(A_1)+(1-\alpha)f(A_2))^{r/k}(\alpha g(B_1)+(1-\alpha)g(B_2))^{r/k}).$$

Since $\log t$ ($t>0$) is concave and increasing, (4.5) shows that for every $0<r\leqslant 1$ and $k=1,\dots,n$ the function

$$(A,B)\mapsto \log\left\{\prod_{i=1}^k \lambda_{n-i+1}(A^{r/k}B^{r/k})\right\}$$

is jointly concave. In particular, when $A=B$, this implies Ky Fan's result [6] (or [13, p. 476]) that $A\mapsto\log\{\prod_{i=1}^k \lambda_{n-i+1}(A)\}$ is concave.

When $A,B\geqslant 0$ are invertible, the (continuous parameter version of) Lie product formula says that

$$\lim_{r\to 0}(B^{r/2}A^rB^{r/2})^{1/r} = \exp(\log A+\log B). \tag{4.7}$$

Also, note [8, Sec. 4] that for general $A,B\geqslant 0$ the expression $\exp(\log A+\log B)$ is meaningful and given as

$$\exp(\log A+\log B) = \lim_{\varepsilon\downarrow 0}\exp(\log(A+\varepsilon I)+\log(B+\varepsilon I)). \tag{4.8}$$

The following is one more consequence of (4.5).

COROLLARY 4.4. *For every* $k = 1, \ldots, n$ *the function*

$$(A, B) \mapsto \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(\exp(\log A + \log B)) \right\}^{1/2k}$$

is jointly concave.

Proof. Thanks to (4.8) we may assume that A_j, B_j are invertible. Then by (4.7) we get

$$\lim_{r \to 0} \lambda_i (A_j^{r/k} B_j^{r/k})^{1/r} = \lambda_i (\exp(\log A_j + \log B_j))^{1/k}.$$

Hence the limit of (4.5) as $r \to 0$ is

$$\begin{aligned} &\left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(\exp(\log(A_1 + A_2) + \log(B_1 + B_2))) \right\}^{1/2k} \\ &\quad \geqslant \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(\exp(\log A_1 + \log B_1)) \right\}^{1/2k} \\ &\qquad + \left\{ \prod_{i=1}^{k} \lambda_{n-i+1}(\exp(\log A_2 + \log B_2)) \right\}^{1/2k}, \end{aligned}$$

which means the stated joint concavity. □

The next proposition illustrates that the assumption $0 < r \leqslant 1$ is essential for (4.5).

PROPOSITION 4.5. *If* $1 < r < \infty$, *then* $(A, B) \mapsto \lambda_n (A^r B^r)^{1/2r}$ *is not jointly concave.*

Proof. Let $r > 0$ and suppose that $(A, B) \mapsto \lambda_n (A^r B^r)^{1/2r}$ is jointly concave. Then for any invertible A_1, A_2 we get

$$\lambda_n((A_1 + A_2)^r (A_1^{-1} + A_2^{-1})^r)^{1/2r} \geqslant \lambda_n (A_1^r A_1^{-r})^{1/2r} + \lambda_n (A_2^r A_2^{-r})^{1/2r} = 2,$$

which implies that

$$(A_1^{-1} + A_2^{-1})^{r/2} (A_1 + A_2)^r (A_1^{-1} + A_2^{-1})^{r/2} \geqslant 2^{2r} I$$

and hence

$$\left(\frac{A_1 + A_2}{2} \right)^r \geqslant \left(\frac{A_1^{-1} + A_2^{-1}}{2} \right)^{-r}.$$

Now let $A = A_1^{1/2}$ and $B = A_1^{-1/2} A_2 A_1^{-1/2}$. Then the above inequality is written as

$$\{A(I + B)A/2\}^r \geqslant \{2AB(I + B)^{-1} A\}^r,$$

which holds for all $A, B \geqslant 0$ by continuity. For instance, take

$$A = \frac{1}{(1 + a^2)^{1/4}} \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix} \ (0 < a < 1), \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

and put $X = A(I+B)A/2$ and $Q = 2AB(I+B)^{-1}A$. Then

$$X = \frac{1}{\sqrt{1+a^2}}\begin{bmatrix} 1+\frac{a^2}{2} & \frac{3a}{2} \\ \frac{3a}{2} & \frac{1}{2}+a^2 \end{bmatrix}, \quad Q = \frac{1}{\sqrt{1+a^2}}\begin{bmatrix} 1 & a \\ a & a^2 \end{bmatrix},$$

so that Q is a rank one projection, $X \geqslant Q$, $\det(X-Q) = 0$, and $X^r \geqslant Q$. Transforming by unitary similarity we can get

$$\tilde{X} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \ (\alpha > \beta > 0), \quad \tilde{Q} = \begin{bmatrix} t^2 & t\sqrt{1-t^2} \\ t\sqrt{1-t^2} & 1-t^2 \end{bmatrix} \ (0 < t < 1)$$

such that $\tilde{X} \geqslant \tilde{Q}$, $\det(\tilde{X}-\tilde{Q}) = 0$, and $\tilde{X}^r \geqslant \tilde{Q}$. Therefore

$$\alpha\beta = \alpha(1-t^2) + \beta t^2,$$
$$\alpha^r\beta^r \geqslant \alpha^r(1-t^2) + \beta^r t^2,$$

so that

$$\{\alpha(1-t^2)+\beta t^2\}^r \geqslant \alpha^r(1-t^2) + \beta^r t^2.$$

This does not hold for $r > 1$ and we must have $r \leqslant 1$. □

We end this section with a small remark. For any $n \times n$ matrix X let $\vec{s}(X) = (s_1(X), \dots, s_n(X))$ denote the singular values (with multiplicities) of X (i.e. the eigenvalues of $|X|$) in decreasing order. When $r = 1/2$, (4.5) may be viewed as a sort of Cauchy-Schwarz type inequality. As was discussed in [10], there is a more natural weak majorization of Cauchy-Schwarz type as follows: For every X_j, Y_j and $k = 1, \dots, n$,

$$\prod_{i=1}^{k} s_i(X_1^*X_1 + X_2^*X_2)^{1/2} s_i(Y_1^*Y_1 + Y_2^*Y_2)^{1/2} \geqslant \prod_{i=1}^{k} s_i(X_1^*Y_1 + X_2^*Y_2).$$

This can be proved by a nice application of 2×2 trick and the Horn theorem on majorization.

5. Weak majorizations involving α-power mean

In this section let us discuss inequalities similar to (4.3) and (4.5) involving α-power mean. When A, B are invertible, the parallel sum $A : B$ is defined by

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

and for $0 \leqslant \alpha \leqslant 1$ the α-power mean $A \#_\alpha B$ is defined by

$$A \#_\alpha B = A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}.$$

For general $A, B \geqslant 0$, $A : B$ is given as the decreasing limit of $(A+\varepsilon I) : (B+\varepsilon I)$ as $\varepsilon \downarrow 0$, and similarly for $A \#_\alpha B$. We already treated parallel sum in the proof of Proposition 4.5. The harmonic mean is $2(A : B)$ and the geometric mean is $A \# B = A \#_{1/2} B$. Also, $A \#_0 B = A$ and $A \#_1 B = B$. As properties of general operator means, the map $(A, B) \mapsto A \#_\alpha B$ is jointly operator monotone and jointly operator concave. (See [11] for general theory of operator means.)

The next proposition gives a variant of (4.3) when usual product is replaced by geometric mean.

PROPOSITION 5.1. *For every* $1 \leqslant p, q < \infty$ *and* $k = 1, \ldots, n$,

$$\left\{\prod_{i=1}^{k} \lambda_{n-i+1}((A_1^p + A_2^p)^{1/p} \# (B_1^q + B_2^q)^{1/q})\right\}^{2pq/(p+q)k}$$
$$\geqslant \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(A_1 \# B_1)\right\}^{2pq/(p+q)k} + \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(A_2 \# B_2)\right\}^{2pq/(p+q)k}.$$

Proof. Put $a_j = \{\prod_{i=1}^{k} \lambda_{n-i+1}(A_j \# B_j)\}^{2pq/(p+q)k}$ and $\alpha = a_1/(a_1 + a_2)$. Then we have

$$\left\{\prod_{i=1}^{k} \lambda_{n-i+1}((A_1^p + A_2^p)^{1/p} \# (B_1^q + B_2^q)^{1/q})\right\}^{1/k}$$
$$\geqslant \left\{\prod_{i=1}^{k} \lambda_{n-i+1}((\alpha^{1-1/p} A_1 + (1-\alpha)^{1-1/p} A_2) \right.$$
$$\left. \# (\alpha^{1-1/q} B_1 + (1-\alpha)^{1-1/q} B_2)\right\}^{1/k}$$
$$\geqslant \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(\alpha^{1-(p+q)/2pq} A_1 \# B_1 \right.$$
$$\left. + (1-\alpha)^{1-(p+q)/2pq} A_2 \# B_2)\right\}^{1/k}$$
$$\geqslant \alpha^{1-(p+q)/2pq} a_1^{(p+q)/2pq} + (1-\alpha)^{1-(p+q)/2pq} a_2^{(p+q)/2pq}$$
$$= (a_1 + a_2)^{(p+q)/2pq}.$$

Above we used Lemma 1.1, the properties of # mentioned above, the Oppenheim inequality (4.6), and positive homogeneity of $\{\prod_{i=1}^{k} \lambda_{n-i+1}(\cdot)\}^{1/k}$. Hence the desired inequality is shown. □

The following variant of (4.5) holds for general α-power mean. The proof is similar to the above and we omit the details.

PROPOSITION 5.2. *Let* $0 \leqslant \alpha \leqslant 1$. *Then for every* $0 < r \leqslant 1$ *and* $k = 1, \ldots, n$ *the function*

$$(A, B) \mapsto \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(A^r \#_\alpha B^r)\right\}^{1/rk}$$

is jointly concave.

Next, note that (4.3) can be reformulated in terms of parallel sum. When A_j, B_j

are invertible, replace A_j, B_j by A_j^{-1}, B_j^{-1} in (4.3) to get

$$\left\{\prod_{i=1}^{k} \lambda_i((A_1^{-p}+A_2^{-p})^{-1/kp}(B_1^{-q}+B_2^{-q})^{-1/kq})\right\}^{-pq/(p+q)}$$
$$\geqslant \left\{\prod_{i=1}^{k} \lambda_i(A_1^{1/k}B_1^{1/k})\right\}^{-pq/(p+q)} + \left\{\prod_{i=1}^{k} \lambda_i(A_2^{1/k}B_2^{1/k})\right\}^{-pq/(p+q)}.$$

This is equivalently written as

$$\left\{\prod_{i=1}^{k} \lambda_i((A_1^{p} : A_2^{p})^{1/kp}(B_1^{q} : B_2^{q})^{1/kq})\right\}^{pq/(p+q)}$$
$$\leqslant \left\{\prod_{i=1}^{k} \lambda_i(A_1^{1/k}B_1^{1/k})\right\}^{pq/(p+q)} : \left\{\prod_{i=1}^{k} \lambda_i(A_2^{1/k}B_2^{1/k})\right\}^{pq/(p+q)}$$

for every $1 \leqslant p, q < \infty$.

The following is a similar inequality for α-power mean, while it does not contain parameters p, q.

PROPOSITION 5.3. *For every* $0 < \alpha < 1$ *and* $k = 1, \ldots, n$,

$$\prod_{i=1}^{k} \lambda_i((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)) \leqslant \prod_{i=1}^{k} \{\lambda_i(A_1B_1)^{1-\alpha}\lambda_i(A_2B_2)^{\alpha}\} \tag{5.1}$$

and

$$\prod_{i=1}^{k} \lambda_{n-i+1}((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)) \geqslant \prod_{i=1}^{k} \{\lambda_{n-i+1}(A_1B_1)^{1-\alpha}\lambda_{n-i+1}(A_2B_2)^{\alpha}\} \tag{5.2}$$

with equality for $k = n$.

Proof. Once we can show that

$$\lambda_1((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)) \leqslant \lambda_1(A_1B_1)^{1-\alpha}\lambda_1(A_2B_2)^{\alpha}, \tag{5.3}$$

the proof of (5.1) is a standard application of the technique of antisymmetric tensors as in the proof of [4, Theorem 2.1] (also Theorem 4.3 above). For $k = n$ both sides of (5.1) are equal to $(\det A_1 \det B_1)^{1-\alpha}(\det A_2 \det B_2)^{\alpha}$. So it is enough for (5.1) to show (5.3). By positive homogeneity we have for $\gamma_1, \gamma_2 > 0$

$$\lambda_1(((\gamma_1 A_1) \#_\alpha (\gamma_2 A_2))(B_1 \#_\alpha B_2)) = \gamma_1^{1-\alpha}\gamma_2^{\alpha}\lambda_1((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)),$$
$$\lambda_1((\gamma_1 A_1)B_1)^{1-\alpha}\lambda_1((\gamma_2 A_2)B_2)^{\alpha} = \gamma_1^{1-\alpha}\gamma_2^{\alpha}\lambda_1(A_1B_1)^{1-\alpha}\lambda_1(A_2B_2)^{\alpha}.$$

So we may assume that $\lambda_1(A_1B_1) = \lambda_1(A_2B_2) = 1$. Furthermore, we may assume that A_1, A_2 are invertible. Then $B_1 \leqslant A_1^{-1}$ and $B_2 \leqslant A_2^{-1}$, so that by monotonity of α-power mean

$$B_1 \#_\alpha B_2 \leqslant A_1^{-1} \#_\alpha A_2^{-1} = (A_1 \#_\alpha A_2)^{-1}.$$

This implies that $\lambda_1((A_1 \#_\alpha A_2)(B_1 \#_\alpha B_2)) \leqslant 1$. Hence (5.3) is shown. Moreover, (5.2) immediately follows from (5.1) applied to A_j^{-1}, B_j^{-1}. $\square$

By using the notation $\underset{(\log)}{\prec}$ of log majorization introduced in [4], (5.1) is rewritten as

$$\vec{s}((A_1 \#_\alpha A_2)^{1/2}(B_1 \#_\alpha B_2)^{1/2}) \underset{(\log)}{\prec} \vec{s}(A_1^{1/2}B_1^{1/2})^{1-\alpha}\vec{s}(A_2^{1/2}B_2^{1/2})^\alpha, \tag{5.4}$$

where $\vec{s}(\cdot)$ is the vector of the singular values in decreasing order and $\vec{a}^{1-\alpha}\vec{b}^\alpha = (a_1^{1-\alpha}b_1^\alpha, \dots, a_n^{1-\alpha}b_n^\alpha)$ for $\vec{a}, \vec{b} \geqslant 0$.

A norm $|||\cdot|||$ on the $n \times n$ matrices is said to be unitarily invariant if $|||UXV||| = |||X|||$ for every X and unitaries U, V. As is well known (see e.g. [5]), there is a bijective correspondence between the set of unitarily invariant norms on $n \times n$ matrices and the set of symmetric gauge functions on the n-vectors.

COROLLARY 5.4. *For every $0 < \alpha < 1$ and every unitarily invariant norm $|||\cdot|||$,*

$$|||(A_1 \#_\alpha A_2)^{1/2}(B_1 \#_\alpha B_2)^{1/2}||| \leqslant |||A_1^{1/2}B_1^{1/2}|||^{1-\alpha}|||A_2^{1/2}B_2^{1/2}|||^\alpha. \tag{5.5}$$

Proof. Log majorization (5.4) implies the weak majorization

$$\vec{s}((A_1 \#_\alpha A_2)^{1/2}(B_1 \#_\alpha B_2)^{1/2}) \prec_w \vec{s}(A_1^{1/2}B_1^{1/2})^{1-\alpha}\vec{s}(A_2^{1/2}B_2^{1/2})^\alpha.$$

Let Φ be the symmetric gauge function corresponding to $|||\cdot|||$ so that $|||X||| = \Phi(\vec{s}(X))$. Note that $\vec{a} \prec_w \vec{b}$ implies $\Phi(\vec{a}) \leqslant \Phi(\vec{b})$ for $\vec{a}, \vec{b} \geqslant 0$. Furthermore, the following Hölder inequality for Φ is well known (see [5, IV.1.6]): For every $0 < \alpha < 1$ and $\vec{a}, \vec{b} \geqslant 0$,

$$\Phi(\vec{a}^{1-\alpha}\vec{b}^\alpha) \leqslant \Phi(\vec{a})^{1-\alpha}\Phi(\vec{b})^\alpha.$$

Therefore

$$\begin{aligned}
&\Phi(\vec{s}((A_1 \#_\alpha A_2)^{1/2}(B_1 \#_\alpha B_2)^{1/2})) \\
&\quad \leqslant \Phi(\vec{s}(A_1^{1/2}B_1^{1/2})^{1-\alpha}\vec{s}(A_2^{1/2}B_2^{1/2})^\alpha) \\
&\quad \leqslant \Phi(\vec{s}(A_1^{1/2}B_1^{1/2}))^{1-\alpha}\Phi(\vec{s}(A_2^{1/2}B_2^{1/2}))^\alpha,
\end{aligned}$$

which is the required inequality. $\square$

When $A_1 = B_1 = A$ and $A_2 = B_2 = B$, (5.5) becomes a known inequality

$$|||A \#_\alpha B||| \leqslant |||A|||^{1-\alpha}|||B|||^\alpha = |||A||| \#_\alpha |||B|||.$$

In [4] and [8], we obtained several logarithmic trace inequalities. The following corollary gives another one.

COROLLARY 5.5. *If A_2, B_2 are invertible, then*

$$\begin{aligned}
&\mathrm{Tr}(A_1^{1/2}B_1A_1^{1/2}\log(A_1^{1/2}B_2^{-1}A_1^{1/2})) + \mathrm{Tr}(B_1^{1/2}A_1B_1^{1/2}\log(B_1^{1/2}A_2^{-1}B_1^{1/2})) \\
&\quad \geqslant (\mathrm{Tr}\,A_1B_1)\log\frac{\mathrm{Tr}\,A_1B_1}{\mathrm{Tr}\,A_2B_2}.
\end{aligned} \tag{5.6}$$

Proof. First note that the traces in the left-hand side are well defined for all $A_1, B_1 \geqslant 0$ whenever A_2, B_2 are invertible. By continuity we may assume that A_1, B_1 are also invertible. By (5.5) applied to Frobenius norm (with A_2, B_2 exchanged) we get

$$\mathrm{Tr}((A_1 \#_\alpha B_2)(B_1 \#_\alpha A_2)) \leqslant (\mathrm{Tr}\, A_1 B_1)^{1-\alpha} (\mathrm{Tr}\, A_2 B_2)^\alpha .$$

When $\alpha = 0$, both sides of the above are equal to $\mathrm{Tr}\, A_1 B_1$. Therefore

$$\left.\frac{d}{d\alpha}\right|_{\alpha=0} \mathrm{Tr}((A_1 \#_\alpha B_2)(B_1 \#_\alpha A_2)) \leqslant \left.\frac{d}{d\alpha}\right|_{\alpha=0} (\mathrm{Tr}\, A_1 B_1)^{1-\alpha} (\mathrm{Tr}\, A_2 B_2)^\alpha .$$

Computing these derivatives yields the result. $\square$

In particular when $B_1 = I$, (5.6) becomes

$$\mathrm{Tr}\, A_1 (\log A_1^{1/2} B_2^{-1} A_1^{1/2} - \log A_2) \geqslant \mathrm{Tr}\, A_1 \log \frac{\mathrm{Tr}\, A_1}{\mathrm{Tr}\, A_2 B_2} .$$

This becomes the Peierls-Bogoliubov inequality when $B_2 = I$ too.

6. Weak majorizations involving Hadamard product

Let $X \circ Y$ denote the Hadamard product (or the Schur product), i.e. the entrywise product of matrices X, Y. When $1 \leqslant p, q < \infty$ and $1/p + 1/q = 1$, the following Hölder type inequality holds: For every $A_j, B_j \geqslant 0$,

$$(A_1^p + B_1^p)^{1/p} \circ (B_1^q + B_2^q)^{1/q} \geqslant A_1 \circ B_1 + A_2 \circ B_2 .$$

This is seen because $(A, B) \mapsto A^{1/p} \otimes B^{1/q}$ is jointly operator oncave in $A, B \geqslant 0$ ([1], [12]) and $X \circ Y$ is a principal submatrix of $X \otimes Y$.

The following variant of (4.3) for Hadamard product can be shown as in the proof of Proposition 5.1 thanks to joint monotonity and concavity of $(A, B) \mapsto A^{1/2} \circ B^{1/2}$.

PROPOSITION 6.1. *For every $1 \leqslant p, q < \infty$ and $k = 1, \dots, n$,*

$$\begin{aligned} &\left\{ \prod_{i=1}^k \lambda_{n-i+1}((A_1^p + A_2^p)^{1/2p} \circ (B_1^q + B_2^q)^{1/2q}) \right\}^{2pq/(p+q)k} \\ &\quad \geqslant \left\{ \prod_{i=1}^k \lambda_{n-i+1}(A_1^{1/2} \circ B_1^{1/2}) \right\}^{2pq/(p+q)k} \\ &\qquad + \left\{ \prod_{i=1}^k \lambda_{n-i+1}(A_2^{1/2} \circ B_2^{1/2}) \right\}^{2pq/(p+q)k} . \end{aligned} \tag{6.1}$$

Let $p = q = 1/r$ and replace A_j^p, B_j^q by A_j, B_j in (6.1) to show the superadditivity

$$\left\{\prod_{i=1}^{k} \lambda_{n-i+1}((A_1 + A_2)^{r/2} \circ (B_1 + B_2)^{r/2})\right\}^{1/rk} \geqslant \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(A_1^{r/2} \circ B_1^{r/2})\right\}^{1/rk} + \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(A_2^{r/2} \circ B_2^{r/2})\right\}^{1/rk}, \tag{6.2}$$

namely $(A,B) \mapsto \{\prod_{i=1}^{k} \lambda_{n-i+1}(A^{r/2} \circ B^{r/2})\}^{1/rk}$ is jointly concave for $0 < r \leqslant 1$ and $k = 1, \dots, n$. In particular, so is $(A,B) \mapsto \{\det(A^{r/2} \circ B^{r/2})\}^{1/rn}$ for $0 < r \leqslant 1$.

It is shown in [16] that

$$\lim_{r \downarrow 0} (A^r \circ B^r)^{1/r} = \exp(I \circ (\log A + \log B))$$

when $A, B \geqslant 0$ are invertible. This right-hand side is meaningful for all $A, B \geqslant 0$ (take the limit from $A + \varepsilon I, B + \varepsilon I$). So, letting $r \downarrow 0$ in (6.2) yields the following:

COROLLARY 6.2. *For every* $k = 1, \dots, n$ *the function*

$$(A,B) \mapsto \left\{\prod_{i=1}^{k} \lambda_{n-i+1}(\exp(I \circ (\log A + \log B)))\right\}^{1/2k}$$

is jointly concave.

When $k = n$ this says that $(A,B) \mapsto \exp\{\frac{1}{2n}\operatorname{Tr}(\log A + \log B)\}$ is jointly concave, and it may be compared with Lieb's result [12, Corollary 6.1] that so is $(A,B) \mapsto \operatorname{Tr}\exp\{\frac{1}{2}(\log A + \log B)\}$. When $A = \operatorname{diag}(a_1, \dots, a_n)$ and $B = \operatorname{diag}(b_1, \dots, b_n)$,

$$\exp\left\{\frac{1}{2n}\operatorname{Tr}(\log A + \log B)\right\} = \left(\prod_{i=1}^{n} a_i b_i\right)^{1/2n}$$

while

$$\operatorname{Tr}\exp\left\{\frac{1}{2}(\log A + \log B)\right\} = \sum_{i=1}^{n} (a_i b_i)^{1/2}.$$

This clearly explains why constant $1/n$ is necessary in the former expression.

REFERENCES

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

[2] T. ANDO, *Majorizations, doubly stochastic matrices, and comparison of eigenvalues*, Linear Algebra Appl. **118** (1989), 163–248.

[3] T. ANDO, *Majorizations and inequalities in matrix theory,* Linear Algebra Appl. **199** (1994), 17–67.

[4] T. ANDO AND F. HIAI, *Log majorization and complementary Golden-Thompson type inequalities,* Linear Algebra Appl. **197/198** (1994), 113–131.

[5] R. BHATIA, *Matrix Analysis*, Springer-Verlag, New-York, 1997.

[6] K. FAN, *On a theorem of Weyl concerning eigenvalues of linear transoformations. II,* Proc. Nat. Acad. Sci. U.S.A. **36** (1950), 31–35.
[7] F. HANSEN AND G.K. PEDERSEN, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann. **258** (1982), 229–241.
[8] F. HIAI AND D. PETZ, *The Golden-Thompson trace inequality is complemented*, Linear Algebra Appl. **181** (1993), 153–185.
[9] T. KATO, *Spectral order and a matrix limit theorem,* Linear and Multilinear Algebra **8** (1979), 15–19.
[10] H. KOSAKI, *Some remarks on the positive operator* $(a^2+b^2)^{1/2}$, unpublished notes.
[11] F. KUBO AND T. ANDO, *Means of positive linear operators*, Math. Ann. **246** (1980), 205–224.
[12] E.H. LIEB, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Adv. Math. **11** (1973), 267–288.
[13] A.W. MARSHALL AND I. OLKIN, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.
[14] A. OPPENHEIM, *Inequalities connected with definite Hermitian forms, II*, Amer. Math. Monthly **61** (1954), 463–466.
[15] M. TAKESAKI, *Theory of Operator Algebras I*, Springer-Verlag, New York-Heidelberg-Berlin, 1979.
[16] G. VISICK, *Majorizations of Hadamard products of matrix powers*, Linear Algebra Appl., to appear.

(Received April 25, 1997)

Tsuyoshi Ando
Faculty of Economics
Hokusei Gakuen University
Atsubetsu-ku, Sapporo 004-0042, Japan
e-mail: ando@hokusei.ac.jp

Fumio Hiai
Department of Mathematical Sciences
Ibaraki University
Mito 310-0056, Japan
e-mail: hiai@mito.ipc.ibaraki.ac.jp