ELSEVIER

Linear Algebra and its Applications 341 (2002) 151–169

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Inequalities involving unitarily invariant norms and operator monotone functions

Fumio Hiai [a,∗,1], Xingzhi Zhan [b,2]

[a] *Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan*
[b] *Institute of Mathematics, Peking University, Beijing 100871, China*

Received 1 February 2001; accepted 4 April 2001

Submitted by K. Okubo

Dedicated to Professor Tsuyoshi Ando on the occasion of his 70th birthday

Abstract

Let $\|\cdot\|$ be a unitarily invariant norm on matrices. For matrices A, B, X with A, B positive semidefinite and X arbitrary, we prove that the function $t \mapsto \|\,|A^t X B^{1-t}|^r\,\| \cdot \|\,|A^{1-t} X B^t|^r\,\|$ is convex on $[0, 1]$ for each $r > 0$. This convexity result interpolates the matrix Cauchy–Schwarz inequality $\|\,|A^{1/2} X B^{1/2}|^r\,\|^2 \leqslant \|\,|AX|^r\,\| \cdot \|\,|XB|^r\,\|$ due to R. Bhatia and C. Davis [Linear Algebra Appl. 223/224 (1995) 119], and also it generalizes A.W. Marshall and I. Olkin's [Pacific J. Math. 15 (1965) 241] result that the condition number $\|A^s\| \cdot \|A^{-s}\|$ is increasing in $s > 0$. We prove that if $f(t)$ is a nonnegative operator monotone function on $[0, \infty)$ and $\|\cdot\|$ is a normalized unitarily invariant norm, then $f(\|X\|) \leqslant \|f(|X|)\|$ for every matrix X. The special case when $f(t) = t^r$ $(0 < r \leqslant 1)$ is used to consider the monotonicity of $p \mapsto \|A^p + B^p\|^{1/p}$ as well as $p \mapsto \|(A^p + B^p)^{1/p}\|$. Furthermore, we obtain some norm inequalities of Hölder and Minkowski types related to the expression $\|\,|A|^p + |B|^p\,\|^{1/p}$. For example, comparisons are made between $\|C^*A + D^*B\|$ and $\|\,|A|^p + |B|^p\,\|^{1/p} \cdot \|\,|C|^q + |D|^q\,\|^{1/q}$, where $p^{-1} + q^{-1} = 1$. © 2002 Elsevier Science Inc. All rights reserved.

AMS classification: 15A60; 15A45; 47A30

Keywords: Unitarily invariant norms; Operator monotone functions; Cauchy–Schwarz inequality; Convexity; Monotonicity; Hölder inequality; Schatten norms

∗ Corresponding author.
E-mail addresses: hiai@math.is.tohoku.ac.jp (F. Hiai), zhan@math.pku.edu.cn (X. Zhan).

1 Supported by Grant-in-Aid for Exploratory Research 12874019.

2 Supported by the National Science Foundation of China under Grant 19801004 and Grant-in-Aid for JSPS Fellows 99216.

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 3 5 3 - 6

1. Introduction

Let M_n be the space of $n \times n$ complex matrices. A norm $\|\cdot\|$ on M_n is called *unitarily invariant* if $\|UAV\| = \|A\|$ for all A, U, V with U, V unitary. Examples in this class are Schatten p-norms and Ky Fan k-norms. For basic properties of unitarily invariant norms see [4,8]. The matrix Cauchy–Schwarz inequality proved by Horn and Mathias [9,10] is

$$\| \, |A^*B|^r \, \|^2 \leqslant \|(A^*A)^r\| \cdot \|(B^*B)^r\| \tag{1}$$

for all $A, B \in M_n$, any real number $r > 0$ and every unitarily invariant norm, where $|Y| \equiv (Y^*Y)^{1/2}$. Bhatia and Davis [5] (see also [4, Theorem IX.5.2] and [7, p. 174]) generalized this to the form

$$\| \, |A^*XB|^r \, \|^2 \leqslant \| \, |AA^*X|^r \, \| \cdot \| \, |XBB^*|^r \, \|$$

for all $A, B, X \in M_n$ and any $r > 0$, which is obviously equivalent to

$$\| \, |A^{1/2}XB^{1/2}|^r \, \|^2 \leqslant \| \, |AX|^r \, \| \cdot \| \, |XB|^r \, \| \tag{2}$$

for positive semidefinite A, B and arbitrary X. We remark that the following more general Hölder type inequality was proved in [11, Theorem 3]:

$$\| \, |AXB|^r \, \| \leqslant \| \, |A^pX|^r \, \|^{1/p} \cdot \| \, |XB^q|^r \, \|^{1/q} \tag{3}$$

for positive semidefinite A, B, arbitrary X, and positive real numbers r, p, q with $p^{-1} + q^{-1} = 1$.

On the other hand, for positive real numbers a_i, b_i $(i = 1, \ldots, n)$ and real numbers u, x, Callebaut [6] gave the following interesting refinement of the classical Cauchy–Schwarz inequality: The expression

$$\left(\sum_{i=1}^{n} a_i^{u+x} b_i^{u-x}\right)\left(\sum_{i=1}^{n} a_i^{u-x} b_i^{u+x}\right) \tag{4}$$

increases as $|x|$ increases. See [13] for a simple proof. To see the effect just consider the case $u = 1$, $-1 \leqslant x \leqslant 1$ in (4). In the same spirit, in Section 2 we will prove the matrix analog by considering the convexity of a norm function (Theorem 1).

Given a norm $\|\cdot\|$ on M_n, the condition number of an invertible matrix A is defined as

$$c(A) = \|A\| \cdot \|A^{-1}\|.$$

This is one of the basic concepts in numerical analysis; it serves as measures of the difficulty in solving a system of linear equations. Marshall and Olkin [12, Theorem 3.2] proved that for positive definite A and every unitarily invariant norm

$$c(A^s) = \|A^s\| \cdot \|A^{-s}\| \tag{5}$$

is increasing in $s > 0$. We generalize this result in Section 2.

In Section 3 we first prove that the inequality

$$f(\|A\|) \leqslant \|f(|A|)\| \tag{6}$$

holds when $f(t)$ is a nonnegative operator monotone function on $[0, \infty)$ and $\|\cdot\|$ is a unitarily invariant norm normalized as $\|\mathrm{diag}(1, 0, \ldots, 0)\| = 1$. Furthermore, the reverse inequality is shown when $\|\cdot\|$ is normalized as $\|I\| = 1$. We next use the special case of (6) when $f(t) = t^r$ $(0 < r \leqslant 1)$ to discuss the monotonicity of $p \mapsto \|A^p + B^p\|^{1/p}$.

Finally in Section 4 we obtain some norm inequalities of Hölder and Minkowski types. Similar kinds of norm inequalities related to the expression $\|(|A|^p + |B|^p)^{1/p}\|$ were discussed in [2]. However, in this paper we mostly treat the different expression $\|\, |A|^p + |B|^p \,\|^{1/p}$. It seems that the latter expression is somewhat easier to handle than the former. The forms of the inequalities obtained are

$$\alpha\|C^*A + D^*B\| \leqslant \|\, |A|^p + |B|^p \,\|^{1/p} \cdot \|\, |C|^q + |D|^q \,\|^{1/q}$$
$$\quad (\text{for } p^{-1} + q^{-1} = 1),$$
$$\beta\|\, |A_1 + A_2|^p + |B_1 + B_2|^p \,\|^{1/p}$$
$$\quad \leqslant \|\, |A_1|^p + |B_1|^p \,\|^{1/p} + \|\, |A_2|^p + |B_2|^p \,\|^{1/p},$$

where α, β are constants depending on p (also the norm $\|\cdot\|$). Unlike the scalar case it turns out that the constants α, β strictly smaller than 1 are indispensable except the case $p = 2$.

2. Convexity of certain functions involving unitarily invariant norms

In this section we treat some functions of a real variable involving unitarily invariant norms. We prove the convexity of those functions refining the known norm inequalities of Cauchy–Schwarz type.

Theorem 1. *Let $A, B, X \in M_n$ with A, B positive semidefinite and X arbitrary. For every positive real number r and every unitarily invariant norm, the function*

$$\phi(t) = \|\, |A^t X B^{1-t}|^r \,\| \cdot \|\, |A^{1-t} X B^t|^r \,\|$$

is convex on the interval $[0, 1]$ *and attains its minimum at* $t = 1/2$. *Consequently, it is decreasing on* $[0, 1/2]$ *and increasing on* $[1/2, 1]$.

Proof. Since $\phi(t)$ is continuous and symmetric with respect to $t = 1/2$, all the conclusions will follow after we show that

$$\phi(t) \leqslant \{\phi(t + s) + \phi(t - s)\}/2 \tag{7}$$

for $t \pm s \in [0, 1]$. By (2) we have

$$\begin{aligned}\| \, |A^t X B^{1-t}|^r \, \| &= \| \, |A^s (A^{t-s} X B^{1-t-s}) B^s|^r \, \| \\ &\leqslant \left\{ \| \, |A^{t+s} X B^{1-(t+s)}|^r \, \| \cdot \| \, |A^{t-s} X B^{1-(t-s)}|^r \, \| \right\}^{1/2}\end{aligned}$$

and

$$\begin{aligned}\| \, |A^{1-t} X B^t|^r \, \| &= \| \, |A^s (A^{1-t-s} X B^{t-s}) B^s|^r \, \| \\ &\leqslant \left\{ \| \, |A^{1-(t-s)} X B^{t-s}|^r \, \| \cdot \| \, |A^{1-(t+s)} X B^{t+s}|^r \, \| \right\}^{1/2}.\end{aligned}$$

Upon multiplication of the above two inequalities we obtain

$$\| \, |A^t X B^{1-t}|^r \, \| \cdot \| \, |A^{1-t} X B^t|^r \, \| \leqslant \{\phi(t+s)\phi(t-s)\}^{1/2}. \tag{8}$$

Applying the arithmetic–geometric mean inequality to the right-hand side of (8) yields (7). This completes the proof. □

An immediate consequence of Theorem 1 interpolates the Cauchy–Schwarz inequality (2) as follows.

Corollary 2. *Let* $A, B, X \in M_n$ *be as in Theorem* 1. *For every* $r > 0$ *and every unitarily invariant norm*,

$$\begin{aligned}\| \, |A^{1/2} X B^{1/2}|^r \, \|^2 &\leqslant \| \, |A^t X B^{1-t}|^r \, \| \cdot \| \, |A^{1-t} X B^t|^r \, \| \\ &\leqslant \| \, |AX|^r \, \| \cdot \| \, |XB|^r \, \|\end{aligned}$$

holds for $0 \leqslant t \leqslant 1$.

Corollary 3. *Let* $A, B, X \in M_n$ *with* A, B *positive definite and* X *arbitrary. For every* $r > 0$ *and every unitarily invariant norm, the function*

$$g(s) = \| \, |A^s X B^s|^r \, \| \cdot \| \, |A^{-s} X B^{-s}|^r \, \| \tag{9}$$

is convex on $(-\infty, \infty)$, *attains its minimum at* $s = 0$, *and hence it is decreasing on* $(-\infty, 0)$ *and increasing on* $(0, \infty)$.

Proof. In Theorem 1, replacing A, B, X and t by $A^2, B^{-2}, A^{-1}XB$ and $(1+s)/2$ respectively, we see that $g(s)$ is convex on $(-1, 1)$, decreasing on $(-1, 0)$, increasing on $(0, 1)$ and attains its minimum at $s = 0$ when $-1 \leqslant s \leqslant 1$. Next replacing A, B by their appropriate powers it is easily seen that the above convexity and monotonicity of $g(s)$ on those intervals are equivalent to the same properties on $(-\infty, \infty)$, $(-\infty, 0)$ and $(0, \infty)$ respectively. □

Note that Marshall and Olkin's [12] monotonicity result on the condition number in (5) corresponds to the case $r = 1$, $X = B = I$ (the identity matrix) of $g(s)$ in (9).

To see that Callebaut's result on (4) is indeed a special case of Corollary 3, in (9) put $r = 1$, $s = x$, $A = \mathrm{diag}(a_1, \ldots, a_n)$, $B = \mathrm{diag}(b_1^{-1}, \ldots, b_n^{-1})$, $X = \mathrm{diag}((a_1 b_1)^u, \ldots, (a_n b_n)^u)$, and let $\| \cdot \|$ be the trace norm.

The following is another example of convex functions involving unitarily invariant norms.

Theorem 4. *Let $A_i \in M_n$ $(i = 1, \ldots, k)$ be positive semidefinite. For every positive real number r and every unitarily invariant norm, the function* $t \mapsto \|(\sum_{i=1}^k A_i^t)^r\|$ *is convex on* $(0, \infty)$.

Proof. It suffices to show

$$\left\| \left(\sum_{i=1}^k A_i^{(s+t)/2} \right)^r \right\| \leqslant \frac{\left\| \left(\sum_{i=1}^k A_i^s \right)^r \right\| + \left\| \left(\sum_{i=1}^k A_i^t \right)^r \right\|}{2} \tag{10}$$

for all $s, t > 0$. In inequality (1) setting

$$A = \begin{bmatrix} A_1^{s/2} & 0 & \cdots & 0 \\ A_2^{s/2} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ A_k^{s/2} & 0 & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} A_1^{t/2} & 0 & \cdots & 0 \\ A_2^{t/2} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ A_k^{t/2} & 0 & \cdots & 0 \end{bmatrix},$$

we obtain

$$\left\| \left(\sum_{i=1}^k A_i^{(s+t)/2} \right)^r \right\| \leqslant \left\{ \left\| \left(\sum_{i=1}^k A_i^s \right)^r \right\| \cdot \left\| \left(\sum_{i=1}^k A_i^t \right)^r \right\| \right\}^{1/2}. \tag{11}$$

Applying the arithmetic–geometric mean inequality to the right-hand side of (11) gives (10). □

3. A norm inequality for operator monotone functions with applications

For Hermitian matrices H, K we write $H \leqslant K$ or $K \geqslant H$ to mean that $K - H$ is positive semidefinite. A real-valued continuous function $f(t)$ on $[0, \infty)$ is said to be *operator monotone* if $0 \leqslant A \leqslant B$ implies $f(A) \leqslant f(B)$ for any $A, B \in M_n$ of all orders n. Here $f(A)$ is defined by the usual functional calculus on A. Familiar examples of operator monotone functions are t^p $(0 < p \leqslant 1)$ and $\log(t + 1)$.

A norm on M_n is said to be *normalized* if $\|\mathrm{diag}(1, 0, \ldots, 0)\| = 1$. All the Ky Fan k-norms $(k = 1, \ldots, n)$ and Schatten p-norms $(1 \leqslant p \leqslant \infty)$ are normalized. Given a norm $\| \cdot \|$ on M_n, the dual norm of $\| \cdot \|$ with respect to the Frobenius inner product is

$$\|A\|^{\mathrm{D}} \equiv \max \left\{ |\mathrm{tr}\, AX^*| : X \in M_n, \ \|X\| = 1 \right\}.$$

If $\| \cdot \|$ is a unitarily invariant norm on M_n and $A \geqslant 0$, then by the duality theorem we have (see [8, Theorem 3.5.5] for an equivalent result)

$$\|A\| = \max\left\{\operatorname{tr} AB : B \geqslant 0,\ \|B\|^{\mathrm{D}} = 1,\ B \in M_n\right\}. \tag{12}$$

The following result is a norm inequality for operator monotone functions. The special case when $f(t) = t^r$ $(0 < r \leqslant 1)$ will be used later.

Theorem 5. *Let $f(t)$ be a nonnegative operator monotone function on* $[0, \infty)$ *and $\|\cdot\|$ be a normalized unitarily invariant norm on M_n. Then for every $A \in M_n$,*

$$f(\|A\|) \leqslant \|f(|A|)\|. \tag{13}$$

Proof. Since $\|A\| = \|\,|A|\,\|$, it suffices to prove (13) for the case when A is positive semidefinite. We now make this assumption. By (12) there exists a $B \geqslant 0$ with $\|B\|^{\mathrm{D}} = 1$ such that

$$\|A\| = \operatorname{tr} AB. \tag{14}$$

Denote by $\|\cdot\|_\infty$ and $\|\cdot\|_1$ the operator (spectral) norm and the trace norm respectively. Every normalized unitarily invariant norm satisfies

$$\|X\|_\infty \leqslant \|X\| \leqslant \|X\|_1 \quad (X \in M_n)$$

(see [4, p. 93]). Since $\|\cdot\|$ is normalized, $\|\cdot\|^{\mathrm{D}}$ is also a normalized unitarily invariant norm. Hence

$$1 = \|B\|^{\mathrm{D}} \leqslant \|B\|_1 = \operatorname{tr} B. \tag{15}$$

From $\|A\|_\infty \leqslant \|A\|$ and (14) we have

$$\begin{aligned} \frac{s\|A\|}{s + \|A\|} \leqslant \frac{s\|A\|}{s + \|A\|_\infty} &= \operatorname{tr} \frac{sAB}{s + \|A\|_\infty} \\ &= \operatorname{tr} B^{1/2} \frac{sA}{s + \|A\|_\infty} B^{1/2} \\ &\leqslant \operatorname{tr} B^{1/2}\{sA(sI + A)^{-1}\}B^{1/2} \\ &= \operatorname{tr} sA(sI + A)^{-1}B \end{aligned} \tag{16}$$

for any real number $s > 0$. In the above latter inequality we have used the fact that $sA/(s + \|A\|_\infty) \leqslant sA(sI + A)^{-1}$.

It is well known (e.g., [4]) that for each nonnegative operator monotone function $f(t)$ on $[0, \infty)$ there are unique constants $\alpha, \beta \geqslant 0$ and a positive measure $\mu(\cdot)$ on $(0, \infty)$ such that

$$f(t) = \alpha + \beta t + \int_0^\infty \frac{st}{s + t}\,\mathrm{d}\mu(s) \qquad (0 \leqslant t < \infty). \tag{17}$$

Using this integral representation, (15), (14), (16) and (12) we compute

$$\begin{aligned} f(\|A\|) &= \alpha + \beta\|A\| + \int_0^\infty \frac{s\|A\|}{s + \|A\|}\,\mathrm{d}\mu(s) \\ &\leqslant \alpha \operatorname{tr} B + \beta \operatorname{tr} AB + \int_0^\infty \operatorname{tr} sA(sI + A)^{-1}B\,\mathrm{d}\mu(s) \end{aligned}$$

$$
\begin{aligned}
&= \operatorname{tr} \left\{ \alpha I + \beta A + \int_0^\infty sA(sI + A)^{-1} \mathrm{d}\mu(s) \right\} B \\
&= \operatorname{tr} f(A)B \\
&\leqslant \|f(A)\|,
\end{aligned}
$$

completing the proof. □

We say that a norm $\|\cdot\|$ is *strictly increasing* if $0 \leqslant A \leqslant B$ and $\|A\| = \|B\|$ imply $A = B$. For instance, the Schatten p-norm $\|\cdot\|_p$ is strictly increasing for all $1 \leqslant p < \infty$. We now consider the equality case of (13).

Theorem 6. *Let $f(t)$ be a nonnegative operator monotone function on* $[0, \infty)$ *and assume that $f(t)$ is non-linear. Let $\|\cdot\|$ be a strictly increasing normalized unitarily invariant norm and $A \in M_n$ with $n \geqslant 2$. Then $f(\|A\|) = \|f(|A|)\|$ if and only if* $f(0) = 0$ *and* rank $A \leqslant 1$.

Proof. First assume that $f(0) = 0$ and $|A| = \lambda P$ with a projection P of rank 1. Then $\|A\| = \lambda\|P\| = \lambda$ by the normalization assumption and $\|f(|A|)\| = \|f(\lambda)P\| = f(\lambda) = f(\|A\|)$. Conversely, assume $f(\|A\|) = \|f(|A|)\|$. If $A = 0$, then since $\|\cdot\|$ is normalized and strictly increasing we must have $f(0) = 0$. Next suppose $A \neq 0$. Let μ be the measure in the integral representation (17) of $f(t)$. Since $f(t)$ is non-linear, $\mu \neq 0$. From the proof of Theorem 5 we know that $f(\|A\|) = \|f(|A|)\|$ implies $\|A\|_\infty = \|A\|$ or equivalently

$$\|\operatorname{diag}(s_1, 0, \ldots, 0)\| = \|\operatorname{diag}(s_1, s_2, \ldots, s_n)\|,$$

where $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_n$ are the singular values of A. Now the strict increasingness of $\|\cdot\|$ forces $s_2 = \cdots = s_n = 0$, that is, rank $A = 1$. So write $|A| = \lambda P$ with a projection P of rank 1. Since $f(\|A\|) = \|f(|A|)\|$ means

$$\|f(\lambda)P\| = \|f(\lambda)P + f(0)(I - P)\|,$$

we have $f(0) = 0$ due to $I - P \neq 0$ and the strict increasingness of $\|\cdot\|$ again. □

Theorem 5 can be complemented by the following reverse inequality for unitarily invariant norms with different normalization.

Theorem 7. *Let $f(t)$ be a nonnegative operator monotone function on* $[0, \infty)$ *and $\|\cdot\|$ be a unitarily invariant norm on M_n with $\|I\| = 1$. Then for every $A \in M_n$,*

$$\|f(|A|)\| \leqslant f(\|A\|).$$

Proof. We may assume that A is positive semidefinite. Since

$$f(A) = \alpha I + \beta A + \int_0^\infty sA(sI + A)^{-1}\,\mathrm{d}\mu(s)$$

as in the proof of Theorem 5, we have

$$\|f(A)\| \leqslant \alpha + \beta\|A\| + \int_0^\infty \|sA(sI + A)^{-1}\|\,\mathrm{d}\mu(s)$$

due to $\|I\| = 1$. Hence it suffices to show

$$\|A(sI + A)^{-1}\| \leqslant \frac{\|A\|}{s + \|A\|} \qquad (s > 0). \tag{18}$$

For each $s > 0$, since

$$\frac{x}{s + x} \leqslant t^2 + (1 - t)^2 \frac{x}{s}$$

for all $x > 0$ and $0 < t < 1$, we get

$$A(sI + A)^{-1} \leqslant t^2 I + (1 - t)^2 s^{-1} A$$

so that

$$\|A(sI + A)^{-1}\| \leqslant \|t^2 I + (1 - t)^2 s^{-1} A\| \leqslant t^2 + (1 - t)^2 s^{-1}\|A\|. \tag{19}$$

Minimize the right-hand side of (19) over $t \in (0, 1)$ to obtain (18). This completes the proof. □

Denote $E \equiv \mathrm{diag}(1, 0, \ldots, 0)$. Combining the inequalities in Theorems 5 and 7, we can write

$$\|E\| \cdot f\left(\frac{\|A\|}{\|E\|}\right) \leqslant \|f(|A|)\| \leqslant \|I\| \cdot f\left(\frac{\|A\|}{\|I\|}\right)$$

for every nonnegative operator monotone function $f(t)$ on $[0, \infty)$ and for every unitarily invariant norm $\|\cdot\|$. As an immediate consequence of this we have:

Corollary 8. *Let $g(t)$ be a strictly increasing function on $[0, \infty)$ such that $g(0) = 0$, $g(\infty) = \infty$ and the inverse function g^{-1} on $[0, \infty)$ is operator monotone. Let $\|\cdot\|$ be an arbitrary unitarily invariant norm. Then for every $A \in M_n$,*

$$\|I\| \cdot g\left(\frac{\|A\|}{\|I\|}\right) \leqslant \|g(|A|)\| \leqslant \|E\| \cdot g\left(\frac{\|A\|}{\|E\|}\right).$$

Given a unitarily invariant norm $\|\cdot\|$ on M_n, for $p > 0$ define

$$\|X\|^{(p)} \equiv \|\,|X|^p\,\|^{1/p} \quad (X \in M_n). \tag{20}$$

Then it is known [4, p. 95] (or [7, Lemma 2.13]) that when $p \geqslant 1$, $\|\cdot\|^{(p)}$ is also a unitarily invariant norm.

Corollary 9. *Let* $\|\cdot\|$ *be a normalized unitarily invariant norm on* M_n. *Then for any* $A \in M_n$, *the function* $p \mapsto \|A\|^{(p)}$ *is decreasing on* $(0, \infty)$ *and*

$$\lim_{p\to\infty} \|A\|^{(p)} = \|A\|_\infty.$$

The above limit formula remains valid without the normalization condition on $\|\cdot\|$.

Proof. The monotonicity part is the special case of Theorem 5 when $f(t) = t^r$, $0 < r \leqslant 1$, but we may give a short direct proof. It suffices to consider the case when A is positive semidefinite, and now we make this assumption. We first show

$$\|A^r\| \geqslant \|A\|^r \quad (0 < r \leqslant 1), \tag{21}$$

$$\|A^r\| \leqslant \|A\|^r \quad (1 \leqslant r < \infty). \tag{22}$$

Since $\|A\|_\infty \leqslant \|A\|$, for $r \geqslant 1$ we get

$$\begin{aligned}\|A^r\| &= \|AA^{r-1}\| \leqslant \|A\|\,\|A^{r-1}\|_\infty \\ &= \|A\|\,\|A\|_\infty^{r-1} \leqslant \|A\|\,\|A\|^{r-1} = \|A\|^r,\end{aligned}$$

proving (22). Inequality (21) follows from (22): For $0 < r \leqslant 1$, $\|A\| = \|(A^r)^{1/r}\| \leqslant \|A^r\|^{1/r}$.

If $0 < p < q$, then

$$\|A^p\| = \|(A^q)^{p/q}\| \geqslant \|A^q\|^{p/q}$$

so that $\|A^p\|^{1/p} \geqslant \|A^q\|^{1/q}$. Moreover,

$$\|A\|_\infty = \|A^p\|_\infty^{1/p} \leqslant \|A^p\|^{1/p} \leqslant \|A^p\|_1^{1/p} = \|A\|_p \longrightarrow \|A\|_\infty$$

as $p \to \infty$, where $\|\cdot\|_p$ is the Schatten p-norm. When $\|\cdot\|$ is not normalized, we just apply the normalized case to $\|\cdot\|/\|\mathrm{diag}(1, 0, \ldots, 0)\|$. □

When $f(t)$ is a nonnegative operator monotone function on $[0, \infty)$, the inequality

$$\|f(X+Y)\| \leqslant \|f(X) + f(Y)\|$$

was proved in [3] for all positive semidefinite $X, Y \in M_n$ and for every unitarily invariant norm. Also, for a function $g(t)$ as in Corollary 8 the reverse inequality

$$\|g(X+Y)\| \geqslant \|g(X) + g(Y)\|$$

was proved there. Special cases of these are

$$\|(X+Y)^r\| \leqslant \|X^r + Y^r\| \quad (0 < r \leqslant 1), \tag{23}$$

$$\|(X+Y)^r\| \geqslant \|X^r + Y^r\| \quad (1 \leqslant r < \infty), \tag{24}$$

which will be repeatedly used in the sequel of the paper.

Next we consider the monotonicity of the functions $p \mapsto \|(A^p + B^p)^{1/p}\|$ and $p \mapsto \|A^p + B^p\|^{1/p}$. We denote by $A \vee B$ the supremum of two positive semidefinite matrices A, B in the sense of Kato (see [1, Lemma 6.15]).

Theorem 10. *Let $A, B \in M_n$ be positive semidefinite. For every unitarily invariant norm, the function $p \mapsto \|(A^p + B^p)^{1/p}\|$ is decreasing on $(0, 1]$. For every normalized unitarily invariant norm, the function $p \mapsto \|A^p + B^p\|^{1/p}$ is decreasing on $(0, \infty)$ and*

$$\lim_{p\to\infty} \|A^p + B^p\|^{1/p} = \|A \vee B\|_\infty.$$

The above limit formula remains valid without the normalization condition.

Proof. Let $0 < p < q \leqslant 1$. Set $r = q/p \ (> 1)$, $X = A^p$, $Y = B^p$ in (24) to get

$$\|(A^p + B^p)^{q/p}\| \geqslant \|A^q + B^q\|. \tag{25}$$

Using a majorization principle [8, Lemma 3.3.8] together with Ky Fan's dominance principle [4, 8], we can apply a convex and increasing function $t^{1/q}$ on $[0, \infty)$ to (25) and get

$$\|(A^p + B^p)^{1/p}\| \geqslant \|(A^q + B^q)^{1/q}\|$$

which shows the first assertion.

To show the second assertion we must prove

$$\|A^p + B^p\|^{1/p} \geqslant \|A^q + B^q\|^{1/q}, \tag{26}$$

for $0 < p < q$. It is easily seen that (26) is equivalent to

$$\|A + B\|^r \geqslant \|A^r + B^r\|$$

for all $r \geqslant 1$ and all positive semidefinite $A, B \in M_n$, which follows from (22) and (24):

$$\|A + B\|^r \geqslant \|(A + B)^r\| \geqslant \|A^r + B^r\|.$$

For $p \geqslant 1$,

$$\begin{aligned}
\|(A^p + B^p)^{1/p}\|_\infty &= \|A^p + B^p\|_\infty^{1/p} \\
&\leqslant \|A^p + B^p\|^{1/p} \\
&\leqslant \|A^p + B^p\|_1^{1/p} \\
&= \|(A^p + B^p)^{1/p}\|_p \\
&\leqslant \|(A^p + B^p)^{1/p} - (A \vee B)\|_p + \|A \vee B\|_p \\
&\leqslant \|(A^p + B^p)^{1/p} - (A \vee B)\|_1 + \|A \vee B\|_p.
\end{aligned}$$

Since

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

(see [1, Lemma 6.15]) and

$$\lim_{p\to\infty}\|A \vee B\|_p = \|A \vee B\|_\infty,$$

we obtain

$$\lim_{p\to\infty}\|A^p + B^p\|^{1/p} = \|A \vee B\|_\infty.$$

This completes the proof. □

We remark that there are some unitarily invariant norms for which $p \mapsto \|(A^p + B^p)^{1/p}\|$ is not decreasing on $(1, \infty)$. Consider the trace norm (here it is just trace since the matrices involved are positive semidefinite). In fact, for the 2×2 matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_t = \begin{bmatrix} t^2 & t\sqrt{1-t^2} \\ t\sqrt{1-t^2} & 1-t^2 \end{bmatrix} \quad (0 < t < 1),$$

it was proved in [2, Lemma 3.3] that for any $p_0 > 2$ there exists a $t \in (0, 1)$ such that $p \mapsto \operatorname{tr}\left\{(A^p + B_t^p)^{1/p}\right\}$ is strictly increasing on $[p_0, \infty)$. Also consider the example $\psi(p) = \operatorname{tr}\left\{(A^p + B^p)^{1/p}\right\}$ with

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -6 \\ -6 & 50 \end{bmatrix}.$$

Then $\psi(1.5) - \psi(8) \approx -1.5719$. Thus $\psi(1.5) \leqslant \psi(8)$. Hence $\psi(p)$ is not decreasing on $[1.5,\ 8]$.

4. Norm inequalities of Hölder and Minkowski types

Let $1 \leqslant p, q \leqslant \infty$ with $p^{-1} + q^{-1} = 1$. It is known [4, p. 95] (also [7, p. 174]) that the Hölder inequality

$$\|X^*Y\| \leqslant \|\,|X|^p\,\|^{1/p} \cdot \|\,|Y|^q\,\|^{1/q} \ \left(= \|X\|^{(p)} \cdot \|Y\|^{(q)}\right) \tag{27}$$

holds for all $X, Y \in M_n$ and every unitarily invariant norm, where $\|\,|X|^p\,\|^{1/p}$ for $p = \infty$ is understood as the operator norm $\|X\|_\infty$ (see Corollary 9). Actually, this is a special case of the Hölder inequality (3) mentioned in Section 1. Here note that $\|\,|X^*|^r\,\| = \|\,|X|^r\,\|$ for $r > 0$. The results in this section may be regarded as applications of inequalities (23), (24) and (27).

In what follows $\|\,|A|^p + |B|^p\,\|^{1/p}$ for $p = \infty$ will be understood as $\|\,|A| \vee |B|\,\|_\infty$ due to Theorem 10. We will use the following simple fact several times: Let A and B be positive semidefinite matrices having the eigenvalues $\alpha_1 \geqslant \cdots \geqslant \alpha_n$ $(\geqslant 0)$ and $\beta_1 \geqslant \cdots \geqslant \beta_n$ $(\geqslant 0)$, respectively. If $\alpha_i \leqslant \beta_i$ $(i = 1, \ldots, n)$ (in particular, if $A \leqslant B$), then there exists a unitary U such that $A^r \leqslant UB^rU^*$ for all $r > 0$.

Theorem 11. *Let* $1 \leqslant p, q \leqslant \infty$ *with* $p^{-1} + q^{-1} = 1$. *For all* $A, B, C, D \in M_n$ *and every unitarily invariant norm*,

$$2^{-|1/p-1/2|}\|C^*A + D^*B\| \leqslant \| \, |A|^p + |B|^p \, \|^{1/p} \cdot \| \, |C|^q + |D|^q \, \|^{1/q} . \quad (28)$$

Moreover, the constant $2^{-|1/p-1/2|}$ *is best possible.*

Proof. Since

$$\left|\frac{1}{p} - \frac{1}{2}\right| = \left|\frac{1}{q} - \frac{1}{2}\right|$$

and the inequality is symmetric with respect to p and q, we may assume $1 \leqslant p \leqslant 2 \leqslant q \leqslant \infty$. Note that

$$\begin{bmatrix} C & 0 \\ D & 0 \end{bmatrix}^* \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} = \begin{bmatrix} C^*A + D^*B & 0 \\ 0 & 0 \end{bmatrix}.$$

From (27) it follows that

$$\begin{aligned} \|C^*A + D^*B\| &= \left\| \begin{bmatrix} C & 0 \\ D & 0 \end{bmatrix}^* \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} \right\| \\ &\leqslant \left\| \left| \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} \right|^p \right\|^{1/p} \cdot \left\| \left| \begin{bmatrix} C & 0 \\ D & 0 \end{bmatrix} \right|^q \right\|^{1/q} \\ &= \|(|A|^2 + |B|^2)^{p/2}\|^{1/p} \cdot \|(|C|^2 + |D|^2)^{q/2}\|^{1/q}. \end{aligned}$$

Since $1 \leqslant p \leqslant 2$, (23) implies

$$\|(|A|^2 + |B|^2)^{p/2}\| \leqslant \| \, |A|^p + |B|^p \|.$$

Since the operator concavity of $t^{2/q}$ gives

$$\frac{|C|^2 + |D|^2}{2} \leqslant \left(\frac{|C|^q + |D|^q}{2}\right)^{2/q},$$

by the remark proceding the theorem we get

$$\left(\frac{|C|^2 + |D|^2}{2}\right)^{q/2} \leqslant U \left(\frac{|C|^q + |D|^q}{2}\right) U^*$$

for some unitary U. Therefore, we have

$$\begin{aligned} \|(|C|^2 + |D|^2)^{q/2}\|^{1/q} &\leqslant 2^{1/2-1/q} \| \, |C|^q + |D|^q \, \|^{1/q} \\ &= 2^{1/p-1/2} \| \, |C|^q + |D|^q \, \|^{1/q}. \end{aligned}$$

Thus the desired inequality (28) follows.

The best possibility of the constant is seen from the following example:

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

with the operator norm $\|\cdot\|_\infty$. $\square$

In particular, the case $p=q=2$ of (28) is

$$\|C^*A+D^*B\| \leqslant \|\,|A|^2+|B|^2\|^{1/2}\cdot\|\,|C|^2+|D|^2\|^{1/2}.$$

We now consider Schatten norms.

Theorem 12. *Let* $1\leqslant r\leqslant\infty$ *and* $1\leqslant p,q\leqslant\infty$ *with* $p^{-1}+q^{-1}=1$. *For all* A, B, C, $D\in M_n$,

$$2^{1/r-1}\|C^*A+D^*B\|_r \leqslant \|\,|A|^p+|B|^p\|_r^{1/p}\cdot\|\,|C|^q+|D|^q\|_r^{1/q}\,. \tag{29}$$

Proof. By (27) for every unitarily invariant norm we have

$$\begin{aligned}\|C^*A+D^*B\| &\leqslant \|C^*A\|+\|D^*B\|\\ &\leqslant \|\,|A|^p\|^{1/p}\cdot\|\,|C|^q\|^{1/q}+\|\,|B|^p\|^{1/p}\cdot\|\,|D|^q\|^{1/q}\\ &\leqslant \left(\|\,|A|^p\|+\|\,|B|^p\|\right)^{1/p}\cdot\left(\|\,|C|^q\|+\|\,|D|^q\|\right)^{1/q}.\end{aligned}$$

When $\|\cdot\|=\|\cdot\|_r$,

$$\begin{aligned}\|\,|A|^p\|_r+\|\,|B|^p\|_r &= \left(\operatorname{tr}|A|^{pr}\right)^{1/r}+\left(\operatorname{tr}|B|^{pr}\right)^{1/r}\\ &\leqslant 2^{1-1/r}\left(\operatorname{tr}(|A|^{pr}+|B|^{pr})\right)^{1/r}\\ &= 2^{1-1/r}\|(|A|^{pr}+|B|^{pr})^{1/r}\|_r\\ &\leqslant 2^{1-1/r}\|\,|A|^p+|B|^p\,\|_r\end{aligned}$$

and similarly for $\|\,|C|^q\|_r+\|\,|D|^q\|_r$. In the last inequality above we have used (23). Thus we get the required inequality (29). $\square$

Theorem 12 is meaningful only for $1\leqslant r\leqslant 2$ because in the case $2<r\leqslant\infty$ it is weaker than Theorem 11. In particular, for $r=1$,

$$\|C^*A+D^*B\|_1 \leqslant \|\,|A|^p+|B|^p\|_1^{1/p}\cdot\|\,|C|^q+|D|^q\|_1^{1/q}\,.$$

Next we consider norm inequalities of Minkowski type.

Theorem 13. *Let* $1\leqslant p<\infty$. *For* $A_i, B_i\in M_n$ $(i=1,2)$ *and every unitarily invariant norm*,

$$\begin{aligned}&2^{-|1/p-1/2|}\|\,|A_1+A_2|^p+|B_1+B_2|^p\,\|^{1/p}\\ &\quad\leqslant \|\,|A_1|^p+|B_1|^p\,\|^{1/p}+\|\,|A_2|^p+|B_2|^p\,\|^{1/p}.\end{aligned}$$

Proof. Since

$$\|(\,|A|^2+|B|^2)^{p/2}\|^{1/p} = \left\|\,\left|\begin{bmatrix}A & 0\\ B & 0\end{bmatrix}\right|^p\,\right\|^{1/p}$$

is a norm in (A, B), we have

$$\begin{aligned}&\|(|A_1+A_2|^2+|B_1+B_2|^2)^{p/2}\|^{1/p}\\&\quad\leqslant \|(|A_1|^2+|B_1|^2)^{p/2}\|^{1/p}+\|(|A_2|^2+|B_2|^2)^{p/2}\|^{1/p}.\end{aligned} \tag{30}$$

When $1 \leqslant p \leqslant 2$, (23) implies

$$\|(|A_i|^2+|B_i|^2)^{p/2}\| \leqslant \|\,|A_i|^p+|B_i|^p\| \quad (i=1,2). \tag{31}$$

By the operator concavity of $t^{p/2}$ we get

$$\frac{|A_1+A_2|^p+|B_1+B_2|^p}{2} \leqslant \left(\frac{|A_1+A_2|^2+|B_1+B_2|^2}{2}\right)^{p/2} \tag{32}$$

so that

$$\begin{aligned}&2^{p/2-1}\|\,|A_1+A_2|^p+|B_1+B_2|^p\|\\&\quad\leqslant \|(|A_1+A_2|^2+|B_1+B_2|^2)^{p/2}\|.\end{aligned} \tag{33}$$

Combining (33), (30) and (31) we have

$$\begin{aligned}&2^{1/2-1/p}\|\,|A_1+A_2|^p+|B_1+B_2|^p\|^{1/p}\\&\quad\leqslant \|\,|A_1|^p+|B_1|^p\|^{1/p}+\|\,|A_2|^p+|B_2|^p\|^{1/p}.\end{aligned}$$

When $p \geqslant 2$, (24) implies

$$\|\,|A_1+A_2|^p+|B_1+B_2|^p\| \leqslant \|(|A_1+A_2|^2+|B_1+B_2|^2)^{p/2}\|. \tag{34}$$

Since, as in the proof of Theorem 11,

$$\left(\frac{|A_i|^2+|B_i|^2}{2}\right)^{p/2} \leqslant U_i\left(\frac{|A_i|^p+|B_i|^p}{2}\right)U_i^*$$

for some unitary U_i, we have

$$2^{1-p/2}\|(|A_i|^2+|B_i|^2)^{p/2}\| \leqslant \|\,|A_i|^p+|B_i|^p\| \quad (i=1,2). \tag{35}$$

Combining (34), (30) and (35) yields

$$\begin{aligned}&2^{1/p-1/2}\|\,|A_1+A_2|^p+|B_1+B_2|^p\,\|^{1/p}\\&\quad\leqslant \|\,|A_1|^p+|B_1|^p\,\|^{1/p}+\|\,|A_2|^p+|B_2|^p\,\|^{1/p}.\end{aligned}$$

This completes the proof. □

The inequality in Theorem 13 holds for $p=\infty$ as well; however the sharper inequality

$$\|\,|A_1+A_2| \vee |B_1+B_2|\,\|_\infty \leqslant \|\,|A_1| \vee |B_1|\,\|_\infty + \|\,|A_2| \vee |B_2|\,\|_\infty \tag{36}$$

is valid. This is seen from Theorem 14 below, but a direct proof is also easy since $\|\,|A| \vee |B|\,\|_\infty = \max\{\|A\|_\infty, \|B\|_\infty\}$.

The example

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

with the operator norm shows that when $1\leqslant p\leqslant 2$ is fixed and $\|\cdot\|$ is arbitrary, the constant $2^{1/2-1/p}$ in Theorem 13 is best possible.

When A_i, B_i $(i=1,2)$ are positive semidefinite matrices, there is a possibility to obtain a sharper inequality. When $\|\cdot\|=\|\cdot\|_\infty$ and $1\leqslant p\leqslant 2$, it is proved in [2, Proposition 3.7] that

$$\|(A_1+A_2)^p+(B_1+B_2)^p\|_\infty^{1/p}\leqslant\|A_1^p+B_1^p\|_\infty^{1/p}+\|A_2^p+B_2^p\|_\infty^{1/p}.$$

We also have

$$2^{1/p-1}\|(A_1+A_2)^p+(B_1+B_2)^p\|^{1/p}\leqslant\|A_1^p+B_1^p\|^{1/p}+\|A_2^p+B_2^p\|^{1/p}$$

for every unitarily invariant norm and $1\leqslant p\leqslant 2$. (The constant $2^{1/p-1}$ is better than $2^{-|1/p-1/2|}$ for $1\leqslant p<4/3$.) Indeed, since the operator convexity of t^p gives

$$2^{1-p}(A_1+A_2)^p\leqslant A_1^p+A_2^p,\quad 2^{1-p}(B_1+B_2)^p\leqslant B_1^p+B_2^p,$$

we get

$$\begin{aligned}&2^{1/p-1}\|(A_1+A_2)^p+(B_1+B_2)^p\|^{1/p}\\&\quad\leqslant\|A_1^p+A_2^p+B_1^p+B_2^p\|^{1/p}\\&\quad\leqslant(\|A_1^p+B_1^p\|+\|A_2^p+B_2^p\|)^{1/p}\\&\quad\leqslant\|A_1^p+B_1^p\|^{1/p}+\|A_2^p+B_2^p\|^{1/p}.\end{aligned}$$

For Schatten norms we have:

Theorem 14. *For* $1\leqslant p,r\leqslant\infty$ *and* $A_i,B_i\in M_n$ $(i=1,2)$,

$$\begin{aligned}&2^{(1/p)(1/r-1)}\|\,|A_1+A_2|^p+|B_1+B_2|^p\,\|_r^{1/p}\\&\quad\leqslant\|\,|A_1|^p+|B_1|^p\,\|_r^{1/p}+\|\,|A_2|^p+|B_2|^p\,\|_r^{1/p}.\end{aligned}$$

Proof. Both limit cases $p=\infty$ and $r=\infty$ follow by taking the limits of the cases $p<\infty$ and $r<\infty$, so we may assume $p,r<\infty$. First, the case $r=1$ is obvious since

$$\|\,|A|^p+|B|^p\,\|_1^{1/p}=\left\|\begin{bmatrix}A&0\\0&B\end{bmatrix}\right\|_p$$

is a norm in (A,B). Next, for $1<r<\infty$, since

$$\left(\frac{|A_1+A_2|^p+|B_1+B_2|^p}{2}\right)^r\leqslant U\left(\frac{|A_1+A_2|^{pr}+|B_1+B_2|^{pr}}{2}\right)U^*$$

for some unitary U by the operator concavity of $t^{1/r}$, we can apply the above trace norm case to get

$$
\begin{aligned}
&\| \, |A_1 + A_2|^p + |B_1 + B_2|^p \|_r^{1/p} \\
&\quad = \| (|A_1 + A_2|^p + |B_1 + B_2|^p)^r \|_1^{1/pr} \\
&\quad \leqslant (2^{r-1})^{1/pr} \| \, |A_1 + A_2|^{pr} + |B_1 + B_2|^{pr} \|_1^{1/pr} \\
&\quad \leqslant 2^{(1/p)(1-1/r)} \left(\| \, |A_1|^{pr} + |B_1|^{pr} \|_1^{1/pr} + \| \, |A_2|^{pt} + |B_2|^{pr} \|_1^{1/pr} \right) \\
&\quad = 2^{(1/p)(1-1/r)} \left(\| (|A_1|^{pr} + |B_1|^{pr})^{1/r} \|_r^{1/p} + \| (|A_2|^{pr} + |B_2|^{pr})^{1/r} \|_r^{1/p} \right).
\end{aligned}
$$

Since by (23)

$$
\| (|A_i|^{pr} + |B_i|^{pr})^{1/r} \|_r \leqslant \| \, |A_i|^p + |B_i|^p \|_r,
$$

we have

$$
\begin{aligned}
&\| \, |A_1 + A_2|^p + |B_1 + B_2|^p \|_r^{1/p} \\
&\quad \leqslant 2^{(1/p)(1-1/r)} \left(\| \, |A_1|^p + |B_1|^p \|_r^{1/p} + \| \, |A_2|^p + |B_2|^p \|_r^{1/p} \right),
\end{aligned}
$$

as desired. □

In the rest of this section we consider norm inequalities of Minkowski type concerning $\| (|A|^p + |B|^p)^{1/p} \|$. At first, since

$$
\| (|A|^2 + |B|^2)^{1/2} \| = \left\| \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} \right\|,
$$

it is obvious that

$$
\begin{aligned}
&\| (|A_1 + A_2|^2 + |B_1 + B_2|^2)^{1/2} \| \\
&\quad \leqslant \| (|A_1|^2 + |B_1|^2)^{1/2} \| + \| (|A_2|^2 + |B_2|^2)^{1/2} \|. \qquad (37)
\end{aligned}
$$

For Schatten norms we give the following two results.

Theorem 15. *For* $1 \leqslant p \leqslant r \leqslant \infty$ *and* $A_i, B_i \in M_n$ $(i = 1, 2)$,

$$
\begin{aligned}
&2^{-|1/p-1/2|} \| (|A_1 + A_2|^p + |B_1 + B_2|^p)^{1/p} \|_r \\
&\quad \leqslant \| (|A_1|^p + |B_1|^p)^{1/p} \|_r + \| (|A_2|^p + |B_2|^p)^{1/p} \|_r .
\end{aligned}
$$

Proof. When $p = \infty$ (hence $r = \infty$), the sharper inequality (36) is valid. So we may assume $p < \infty$ and prove the inequality for a more general unitarily invariant norm of the form $\| \cdot \| = ||| \cdot |||^{(p)}$ (see (20)) with another unitarily invariant norm $||| \cdot |||$. Here note that $\| \cdot \|_r = \| \cdot \|_{r/p}^{(p)}$. When $1 \leqslant p \leqslant 2$, it follows from (23) that

$$
\begin{aligned}
\| (|A_i|^2 + |B_i|^2)^{1/2} \| &= ||| (|A_i|^2 + |B_i|^2)^{p/2} |||^{1/p} \\
&\leqslant ||| \, |A_i|^p + |B_i|^p \, |||^{1/p} \\
&= \| (|A_i|^p + |B_i|^p)^{1/p} \|.
\end{aligned}
$$

By (32) we get

$$2^{1/2-1/p}\|(|A_1+A_2|^p+|B_1+B_2|^p)^{1/p}\| \\ \leqslant \|(|A_1+A_2|^2+|B_1+B_2|^2)^{1/2}\|.$$

Combining the above two inequalities and (37) yields

$$2^{1/2-1/p}\|(|A_1+A_2|^p+|B_1+B_2|^p)^{1/p}\| \\ \leqslant \|(|A_1|^p+|B_1|^p)^{1/p}\|+\|(|A_2|^p+|B_2|^p)^{1/p}\|.$$

The proof for the case $2 \leqslant p < \infty$ is similar. Applying (24) to $|||(|A_1+A_2|^2+|B_1+B_2|^2)^{p/2}|||$ gives

$$\|(|A_1+A_2|^2+|B_1+B_2|^2)^{1/2}\| \geqslant \|(|A_1+A_2|^p+|B_1+B_2|^p)^{1/p}\|.$$

Also, applying (35) to the norm $|||\cdot|||$ gives

$$2^{1/p-1/2}\|(|A_i|^2+|B_i|^2)^{1/2}\| \leqslant \|(|A_i|^p+|B_i|^p)^{1/p}\|.$$

Finally use (37) again to obtain the required inequality. □

Theorem 16. *For* $1 \leqslant p, r \leqslant \infty$ *and* $A_i, B_i \in M_n$ $(i=1,2)$,

$$2^{-|1/p-1/r|}\|(|A_1+A_2|^p+|B_1+B_2|^p)^{1/p}\|_r \\ \leqslant \|(|A_1|^p+|B_1|^p)^{1/p}\|_r+\|(|A_2|^p+|B_2|^p)^{1/p}\|_r\,.$$

Proof. We may assume $p, r < \infty$ as in the proof of Theorem 14. When $1 \leqslant r \leqslant p < \infty$, by (23) and Theorem 14 (the trace norm case) we have

$$\begin{aligned}&\|(|A_1+A_2|^p+|B_1+B_2|^p)^{1/p}\|_r \\ &= \|(|A_1+A_2|^p+|B_1+B_2|^p)^{r/p}\|_1^{1/r} \\ &\leqslant \|\,|A_1+A_2|^r+|B_1+B_2|^r\|_1^{1/r} \\ &\leqslant \|\,|A_1|^r+|B_1|^r\|_1^{1/r}+\|\,|A_2|^r+|B_2|^r\|_1^{1/r}\,.\end{aligned}$$

By the operator concavity of $t^{r/p}$ we get

$$|A_i|^r+|B_i|^r \leqslant 2^{1-r/p}(|A_i|^p+|B_i|^p)^{r/p}.$$

Therefore,

$$\begin{aligned}&\|(|A_1+A_2|^p+|B_1+B_2|^p)^{1/p}\|_r \\ &\leqslant 2^{1/r-1/p}\left\{\|(|A_1|^p+|B_1|^p)^{r/p}\|_1^{1/r}+\|(|A_2|^p+|B_2|^p)^{r/p}\|_1^{1/r}\right\} \\ &= 2^{1/r-1/p}\left\{\|(|A_1|^p+|B_1|^p)^{1/p}\|_r+\|(|A_2|^p+|B_2|^p)^{1/p}\|_r\right\}.\end{aligned}$$

When $1 \leqslant p \leqslant r < \infty$, we can apply Theorem 14 with r/p in place of r to get

$$2^{1/r-1/p} \| \, |A_1 + A_2|^p + |B_1 + B_2|^p \, \|_{r/p}^{1/p} \leqslant \| \, |A_1|^p + |B_1|^p \, \|_{r/p}^{1/p} + \| \, |A_2|^p + |B_2|^p \, \|_{r/p}^{1/p}.$$

Since

$$\| \, |A|^p + |B|^p \, \|_{r/p}^{1/p} = \| (|A|^p + |B|^p)^{1/p} \|_r \, ,$$

the desired inequality follows. □

When $1 \leqslant p \leqslant 2$, the example following the proof of Theorem 13 shows the best possibility of the constant in the inequality in Theorem 15 (attained in the case $r = \infty$). On the other hand, when $2 < p \leqslant \infty$, it is known [2, Proposition 3.6] that there exist positive semidefinite A_i, B_i such that

$$\mathrm{tr} \left\{ \left((A_1 + A_2)^p + (B_1 + B_2)^p \right)^{1/p} \right\} > \mathrm{tr} \left\{ (A_1^p + B_1^p)^{1/p} \right\} + \mathrm{tr} \left\{ (A_2^p + B_2^p)^{1/p} \right\}.$$

Thus, for any $1 \leqslant p \leqslant \infty$ except $p = 2$, the inequality

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

cannot generally hold.

Acknowledgment

This work was done while the second-named author was at the Graduate School of Information Sciences, Tohoku University as a postdoctoral fellow of the Japan Society for the Promotion of Science. He thanks JSPS for the support.

References

[1] T. Ando, Majorization, doubly stochastic matrices, and comparison of eigenvalues, Linear Algebra Appl. 118 (1989) 163–248.
[2] T. Ando, F. Hiai, Hölder type inequalities for matrices, Math. Inequal. Appl. 1 (1998) 1–30.
[3] T. Ando, X. Zhan, Norm inequalities related to operator monotone functions, Math. Ann. 315 (1999) 771–780.
[4] R. Bhatia, Matrix Analysis, Springer, New York, 1997.
[5] R. Bhatia, C. Davis, A Cauchy–Schwarz inequality for operators with applications, Linear Algebra Appl. 223/224 (1995) 119–129.
[6] D.K. Callebaut, Generalization of the Cauchy–Schwarz inequality, J. Math. Anal. Appl. 12 (1965) 491–494.

[7] F. Hiai, Log-majorizations and norm inequalities for exponential operators, in: Linear Operators, Banach Center Publications, vol. 38, Polish Academy of Sciences, Warszawa, 1997, pp. 119–181.
[8] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[9] R.A. Horn, R. Mathias, Cauchy–Schwarz inequalities associated with positive semidefinite matrices, Linear Algebra Appl. 142 (1990) 63–82.
[10] R.A. Horn, R. Mathias, An analog of the Cauchy–Schwarz inequality for Hadamard products and unitarily invariant norms, SIAM J. Matrix Anal. Appl. 11 (1990) 481–498.
[11] R.A. Horn, X. Zhan, Inequalities for C–S seminorms and Lieb functions, Linear Algebra Appl. 291 (1999) 103–113.
[12] A.W. Marshall, I. Olkin, Norms and inequalities for condition numbers, Pacific J. Math. 15 (1965) 241–247.
[13] H.W. McLaughlin, F.T. Metcalf, Remark on a recent generalization of the Cauchy–Schwarz inequality, J. Math. Anal. Appl. 18 (1967) 522–523.