Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
ELSEVIER Linear Algebra and its Applications 377 (2004) 181–194
www.elsevier.com/locate/laa

An operator inequality and self-adjointness☆

Bojan Magajna [a], Marko Petkovšek [a], Aleksej Turnšek [b,*]

[a] Department of Mathematics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia
[b] Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva 6, 1000 Ljubljana, Slovenia

Received 1 April 2003; accepted 11 August 2003

Submitted by R. Bhatia

Abstract

Given bounded positive invertible operators A and B on a Hilbert space $\mathscr{H}$, it is shown that the inequality $\|AXA^{-1}\| + \|B^{-1}XB\| \geqslant 2\|X\|$ holds for all bounded operators X of rank 1 if and only if $B = f(A)$ for some increasing function f satisfying a certain simple inequality, which in the case when the spectrum of A is connected implies that B is a scalar multiple of A. As an application some consequences of the Corach–Porta–Recht type inequality in operator ideals are studied.

AMS classification: 47A63; 47B10; 47B15

Keywords: Operator inequality; Unitarily invariant norm; Schur product; Normal operators

1. Introduction

Corach et al. proved in [4] that if S and X are operators in $\mathscr{B}(\mathscr{H})$ (the algebra of all bounded linear operators on a complex Hilbert space $\mathscr{H}$) with S invertible and self-adjoint, then

$$\|SXS^{-1} + S^{-1}XS\| \geqslant 2\|X\|, \tag{1}$$

☆ The first author was supported by the Ministry of Science and Education of Slovenia, the third author was supported in part by the Grant J1-3125.

* Corresponding author.

E-mail addresses: bojan.magajna@fmf.uni-lj.si (B. Magajna), marko.petkovsek@fmf.uni-lj.si (M. Petkovšek), aleksej.turnsek@fs.uni-lj.si (A. Turnšek).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2003.08.007

where $\|\cdot\|$ is the usual operator norm. In [8] Kittaneh deduced the inequality (1) for all unitarily invariant norms as a consequence of the arithmetic–geometric mean inequality

$$|||AA^*X + XBB^*||| \geqslant 2|||A^*XB|||,$$

where $A, B \in \mathscr{B}(\mathscr{H})$ and $X \in \mathscr{J}$, see [2,3,7,8]. Recently Seddik (see [11]) proved that $\|SXS^{-1} + S^{-1}XS\| \geqslant 2\|X\|$ for all $X \in \mathscr{B}(\mathscr{H})$ if and only if S is a scalar multiple of some self-adjoint operator. Thus in the case of the operator norm (1) is in fact a characterization of self-adjoint operators. So it is natural to ask whether the same is true for other unitarily invariant norms. Thus, if

$$|||SXS^{-1} + S^{-1}XS||| \geqslant 2|||X||| \quad \text{for all } X \in \mathscr{J}, \tag{2}$$

is then γS necessarily self-adjoint for some $\gamma \in \mathbb{C}$? As a consequence of our main result here, we show that S is necessarily normal even under weaker assumption, namely that condition (2) holds only for operators of rank 1, but in the case of von Neumann–Schatten classes $\mathscr{C}_p$ $(1 < p < \infty)$ (2) does not imply that S is self-adjoint. However, a stronger version of the Corach–Porta–Recht inequality, namely

$$\inf_{t>0} \left|\left|\left| tSXS^{-1} + \frac{1}{t}S^{-1}XS \right|\right|\right| \geqslant 2|||X||| \quad \text{for all } X \text{of rank 1},$$

is shown to be equivalent to γS being self-adjoint for some $\gamma \in \mathbb{C}$.

In Theorem 2.1 we will prove that for two positive invertible operators $A, B \in \mathscr{B}(\mathscr{H})$ the inequality

$$\|AXA^{-1}\| + \|B^{-1}XB\| \geqslant 2\|X\| \tag{3}$$

holds for all $X \in \mathscr{B}(\mathscr{H})$ of rank 1 if and only if $B = f(A)$ for some increasing function f on the spectrum $\sigma(A)$ of A satisfying a certain condition, which in the case when $\sigma(A)$ is an interval implies that f is of the form $f(t) = ct$ for some constant c. The proof is based on an observation that condition (3) can be expressed in terms of the spectra of A and B, which enables us to assume that (in a suitable representation) the spectra of A and B coincide with the point spectra.

We end this introduction by recalling that a unitarily invariant norm defined on an ideal $\mathscr{J}$ of $\mathscr{B}(\mathscr{H})$ (contained in the ideal of compact operators) is a norm $|||\cdot|||$ satisfying the following two conditions: (i) if $U, V \in \mathscr{B}(\mathscr{H})$ are unitary then $|||UXV||| = |||X|||$ for all $X \in \mathscr{B}(\mathscr{H})$, and (ii) $|||X||| = \|X\|$ for all rank one operators. Especially important examples of these norms are the von Neumann–Schatten p-norms defined on a finite rank operator X by

$$\|X\|_p = \left(\sum_j s_j^p(X)\right)^{1/p} \qquad \text{for } 1 \leqslant p < \infty,$$

where $s_j(X)$ are the singular values of X arranged in decreasing order $s_1(X) \geqslant s_2(X) \geqslant \cdots$, counted according to the multiplicities. The closure of finite rank operators in the norm $\|\cdot\|_p$ is the von Neumann–Schatten class $\mathscr{C}_p$ (we refer to [5] for more).

2. Main results

In the special case of the usual operator norm on $\mathscr{B}(\mathscr{H})$ and under a stronger assumption that X varies over all (not just rank 1) operators, a part of the following theorem is proved in [11]; our proof here is different even in this case.

Theorem 2.1. *Let $A, B \in \mathscr{B}(\mathscr{H})$ be positive invertible operators. Then the inequality*

$$\|AXA^{-1}\| + \|B^{-1}XB\| \geqslant 2\|X\| \tag{4}$$

holds for all operators $X \in \mathscr{B}(\mathscr{H})$ of rank 1 if and only if $B = f(A)$, where $f: \sigma(A) \to \sigma(B)$ is a strictly increasing positive continuous function satisfying

$$\frac{f(s)}{t} \leqslant \frac{f(t) - f(s)}{t - s} \leqslant \frac{f(t)}{s} \tag{5}$$

for all $s < t$ in the spectrum of A.

In particular, for two unitarily equivalent operators A and B (4) implies that $A = B$.

Each function f satisfying (5) is differentiable at any interior point of $\sigma(A)$ and if $\sigma(A)$ is an interval, then f is of the form $f(t) = ct$ for some constant c.

Proof. Setting $X = \xi \otimes \bar{\eta}$, where $\|\xi\| = \|\eta\| = 1$ (the rank 1 operator mapping η into ξ) and squaring both sides of (4) we get

$$\|A\xi\|^2\|A^{-1}\eta\|^2 + 2\|A\xi\|\|B\eta\|\|A^{-1}\eta\|\|B^{-1}\xi\| + \|B^{-1}\xi\|^2\|B\eta\|^2 \geqslant 4. \tag{6}$$

Since $\inf_{t>0}\left(ta + \frac{1}{t}b\right) = 2\sqrt{ab}$, (6) is equivalent to

$$\|A\xi\|^2\|A^{-1}\eta\|^2 + \inf_{t>0}\left(t^2\|A\xi\|^2\|B\eta\|^2 + \frac{1}{t^2}\|A^{-1}\eta\|^2\|B^{-1}\xi\|^2\right) + \|B^{-1}\xi\|^2\|B\eta\|^2 \geqslant 4,$$

which can be rewritten as

$$\inf_{t>0}\left(t\|A\xi\|^2 + \frac{1}{t}\|B^{-1}\xi\|^2\right)\left(\frac{1}{t}\|A^{-1}\eta\|^2 + t\|B\eta\|^2\right) \geqslant 4, \tag{7}$$

hence as

$$\inf_{t>0}\left\langle\left(tA^2 + \frac{1}{t}B^{-2}\right)\xi, \xi\right\rangle\left\langle\left(\frac{1}{t}A^{-2} + tB^2\right)\eta, \eta\right\rangle \geqslant 4.$$

Since the closure of the numerical range of a positive operator is equal to the convex hull of its spectrum, it follows that condition (4) is equivalent to

$$\inf_{t>0}\left(\min\sigma\left(tA^2 + \frac{1}{t}B^{-2}\right)\min\sigma\left(\frac{1}{t}A^{-2} + tB^2\right)\right) \geqslant 4.$$

Since this is merely a spectral condition, we may replace $\mathscr{H}$ by the Hilbert space of any representation of a C*-algebra $\mathscr{A}$ containing A and B. Thus, by choosing the representation obtained from Berberian's construction [1,12] (or the universal representation), we may assume that all the points in the approximate point spectra of all operators in $\mathscr{A}$ are eigenvalues. Note that the spectrum of a self-adjoint operator always coincides with the approximate point spectrum.

By substitution $B^{-1}XB = Y$ (4) is equivalent to the requirement that

$$\|ABY(AB)^{-1}\| + \|Y\| \geqslant 2\|BYB^{-1}\|$$

for all $Y \in \mathscr{B}(\mathscr{H})$ of rank 1. Let $AB = U|AB|$ be the polar decomposition. Since U is unitary, we get

$$\||AB|Y|AB|^{-1}\| + \|Y\| \geqslant 2\|BYB^{-1}\| \tag{8}$$

for all $Y \in \mathscr{B}(\mathscr{H})$ of rank 1. The operator $|AB|$ is positive and belongs to $\mathscr{A}$, thus its spectrum consists entirely of eigenvalues. Let ξ be an eigenvector of $|AB|$, $\|\xi\| = 1$, and $Y = \xi \otimes \bar{\xi}$. Then by (8) we have

$$1 \geqslant \|B\xi\|\|B^{-1}\xi\|.$$

Since $\langle B\xi, B^{-1}\xi\rangle = 1$, we must have

$$\langle B\xi, B^{-1}\xi\rangle = \|B\xi\|\|B^{-1}\xi\|.$$

Using the fact that equality holds in the Cauchy–Schwarz inequality if and only if the two vectors are linearly dependent, we deduce that ξ is an eigenvector of B. It follows that B and $|AB|$ are diagonal in the same orthonormal basis of $\mathscr{H}$, hence the same holds also for $BA^2B = |AB|^2$, $A^2 = B^{-1}(BA^2B)B^{-1}$ and A.

For an eigenvalue s of A let $\tilde{s}$ be an eigenvalue of B that corresponds to the same common eigenvector ξ (that is, $A\xi = s\xi$ and $B\xi = \tilde{s}\xi$). If s, t are eigenvalues of A and ξ, η the corresponding unit eigenvectors, then with $X = \xi \otimes \bar{\eta}$, (4) shows that

$$\frac{s}{t} + \frac{\tilde{t}}{\tilde{s}} \geqslant 2. \tag{9}$$

If $t = s$ then (9) implies that $\tilde{t} \geqslant \tilde{s}$; interchanging the roles of s and t, it follows that $\tilde{t} = \tilde{s}$. This shows that $f(s) := \tilde{s}$ defines a function on $\sigma(A)$ and we can rewrite (9) as

$$\frac{s}{t} + \frac{f(t)}{f(s)} \geqslant 2. \tag{10}$$

If $s < t$, (10) implies that $f(s) < f(t)$, which proves that f is an increasing function. Moreover, (10) implies that

$$\frac{f(s)}{t} \leqslant \frac{f(t) - f(s)}{t - s} \quad \text{if } s < t \quad \text{and} \quad \frac{f(t) - f(s)}{t - s} \leqslant \frac{f(s)}{t} \quad \text{if } s > t.$$

Thus, if $s < t$, we see (interchanging the roles of s and t in the second inequality) that the estimates (5) hold, from which the continuity of f is evident. Further, if s is an interior point of $\sigma(A)$, then (5) implies that f is differentiable with $f'(s) = f(s)/s$, hence $f(s)/s$ is a constant on each interval contained in $\sigma(A)$. If A and B are unitarily equivalent, then, since A and B have a common complete set of reducing eigenspaces, they must coincide on each such eigenspace, hence $A = B$.

Suppose now conversely that f is an increasing (positive, continuous) function on the spectrum of A satisfying (5) and $B = f(A)$. Choose an orthonormal basis (e_s) of $\mathscr{H}$ consisting of eigenvectors of A. By an approximation, it suffices to show that (4) is satisfied for each rank 1 operator X supported in a subspace $\mathscr{K}$ spanned by a finite subset of basic vectors e_s; in other words, the problem is reduced to the case when X is a finite matrix $X = [x_{ij}]$ with the norm $\|X\| = 1 = \|X\|_2$ (since X is of rank 1) and A and B are diagonal matrices with the entries s_i and $t_i = f(s_i)$ along the diagonal. Since condition (5) implies that $t_i/t_j + s_j/s_i = f(s_i)/f(s_j) + s_j/s_i \geqslant 2$ and for rank 1 operators the Hilbert–Schmidt norm coincides with the operator norm, it follows that

$$\begin{aligned}\|AXA^{-1}\| + \|B^{-1}XB\| &\geqslant \|AXA^{-1} + B^{-1}XB\|_2 \\ &= \left\|\left[(s_i/s_j + t_j/t_i)x_{ij}\right]\right\|_2 \geqslant 2\|X\|. \qquad \square\end{aligned}$$

We remark that if $\sigma(A)$ is discrete it suffices to check condition (5) for each pair of neighboring points, by a convexity argument.

The following generalizes [11, Lemma 4.2]

Corollary 2.2. *All invertible operators $S \in \mathscr{B}(\mathscr{H})$ satisfying the condition*

$$|||SXS^{-1} + S^{-1}XS||| \geqslant 2|||X|||$$

for all $X \in \mathscr{B}(\mathscr{H})$ of rank 1 *are normal.*

Proof. Let $A = (S^*S)^{1/2}$, $B = (SS^*)^{1/2}$, and let $S = UA$ and $S^* = U^*B$ be the polar decompositions. Then by the unitary invariance of the norm

$$2|||X||| \leqslant |||UAXA^{-1}U^*||| + |||U^*B^{-1}XBU||| = |||AXA^{-1}||| + |||B^{-1}XB|||.$$

Since A and B are unitarily equivalent, we have by Theorem 2.1 that $A = B$, hence S is normal. $\square$

As already mentioned, Seddik [11] proved that the condition $\|SXS^{-1} + S^{-1}XS\| \geqslant 2\|X\|$ for all $X \in \mathscr{B}(\mathscr{H})$ implies that γS is self-adjoint for some nonzero $\gamma \in \mathbb{C}$. Now we will study this question for other unitarily invariant norms. To get counterexamples, we consider 2×2 matrices.

Suppose that S is an invertible normal 2×2 matrix, say

$$S = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \tag{11}$$

Then for $X = [x_{ij}]$,

$$SXS^{-1} + S^{-1}XS = \begin{bmatrix} 2x_{11} & \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1}\right)x_{12} \\ \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1}\right)x_{21} & 2x_{22} \end{bmatrix}. \tag{12}$$

Given $\alpha \in \mathbb{C}$, let

$$\mathscr{S}_\alpha : M_2 \to M_2, \quad \mathscr{S}_\alpha(X) = A \circ X$$

be the Schur (that is, entry-wise) multiplication by the matrix

$$A = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$$

and denote by $|||\mathscr{S}_\alpha|||$ the operator norm of $\mathscr{S}_\alpha$ on $(M_2, |||\cdot|||)$, thus

$$|||\mathscr{S}_\alpha||| = \sup\left\{|||\mathscr{S}_\alpha(X)||| : |||X||| = 1\right\}.$$

In particular $\|\mathscr{S}_\alpha\|_p$ denotes the operator norm of $\mathscr{S}_\alpha$ on $(M_2, \|\cdot\|_p)$ and $\|\mathscr{S}_\alpha\|$ the operator norm of $\mathscr{S}_\alpha$ on $(M_2, \|\cdot\|)$. By (12), in this notation, the condition

$$|||SXS^{-1} + S^{-1}XS||| \geqslant 2|||X||| \quad \forall X \in M_2$$

can be written as

$$|||\mathscr{S}_{\alpha^{-1}}(X)||| \geqslant |||X||| \quad \forall X \in M_2, \tag{13}$$

where

$$\alpha = \frac{2}{\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1}}. \tag{14}$$

Since $\mathscr{S}_{\alpha^{-1}}$ is invertible with the inverse $\mathscr{S}_\alpha$, this is equivalent to

$$|||\mathscr{S}_\alpha(X)||| \leqslant |||X||| \quad \forall X \in M_2.$$

If γS is self-adjoint for some $\gamma \in \mathbb{C}$, then λ_1 and λ_2 lie on a straight line through the origin in the complex plane, hence $\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1}$ (and then also α) is a real number. So, we can formulate our problem as follows:

Given a diagonal invertible 2×2 matrix S as in (11), and if α is defined by (14), does the condition

$$|||\mathscr{S}_\alpha(X)||| \leqslant |||X||| \quad \forall X \in M_2, \text{ or equivalently } |||\mathscr{S}_\alpha||| = 1,$$

imply that α is real?

In the case of the operator norm we have

$$\|\mathscr{S}_\alpha\| = \tfrac{1}{2}(|1+\alpha| + |1-\alpha|)$$

by [6], which shows that $\|\mathscr{S}_\alpha\| = 1$ does imply that $\alpha \in [-1, 1]$. By an easy duality argument the same can be deduced for the trace class norm. For the Hilbert–Schmidt norm, however, we compute that

$$\|\mathscr{S}_\alpha(X)\|_2^2 = |x_{11}|^2 + |\alpha|^2|x_{12}|^2 + |\alpha|^2|x_{21}|^2 + |x_{22}|^2 \leqslant \max\{1, |\alpha|^2\}\|X\|_2^2,$$

hence

$$\|\mathscr{S}_\alpha\|_2 = \max\{1, |\alpha|\}.$$

This shows that each $\alpha \in \mathbb{C}$ with $|\alpha| \leqslant 1$ satisfies $\|\mathscr{S}_\alpha\|_2 = 1$ and so in the case of the Hilbert-Schmidt norm (2) does not imply that S is self-adjoint. Analogously, but with much more computational effort, we will show that for each $p > 1$ there exists a disc $\Delta_p \subset \mathbb{C}$ centered at the origin, such that $\|\mathscr{S}_\alpha\|_p = 1$ for all $\alpha \in \Delta_p$. More precisely:

Theorem 2.3. *If $p \geqslant 2$, then $\|\mathscr{S}_\alpha\|_p = 1$ for all $\alpha \in \mathbb{C}$ satisfying $|\alpha| \leqslant 1/\sqrt{n-1}$, where n is the smallest even integer greater than or equal to p. If $1 < p < 2$, then $\|\mathscr{S}_\alpha\|_p = 1$ for all $\alpha \in \mathbb{C}$ satisfying $|\alpha| \leqslant 1/\sqrt{n-1}$, where n is the smallest even integer greater than or equal to $p/(p-1)$.*

Since the proof is rather long and technical it will be given in the last section.

Thus the condition $\|SXS^{-1} + S^{-1}XS\|_p \geqslant 2\|X\|_p$ for all $X \in \mathscr{C}_p$ implies that S is normal by Corollary 2.2 but not necessarily self-adjoint. So we need a stronger condition than the Corach–Porta–Recht inequality to ensure self-adjointness. We begin with an elementary observation.

Lemma 2.4. *If $\inf_{t>0} |tz + \frac{1}{tz}| \geqslant 2$, then z is real.*

Proof. Let $z = |z|\mathrm{e}^{\mathrm{i}\varphi}$. Then the assumption can be written as

$$\inf_{t>0}\left(t^2|z|^2 + \frac{1}{t^2|z|^2}\right) + 2\cos 2\varphi \geqslant 4,$$

from which it follows that $\cos 2\varphi = 1$ (since the "inf" is 2), hence z is real. □

Theorem 2.5. *Let $|||\cdot|||$ be any unitarily invariant norm and suppose that $S \in \mathscr{B}(\mathscr{H})$ is invertible. Then*

$$\inf_{t>0}\left|\left|\left| tSXS^{-1} + \frac{1}{t}S^{-1}XS \right|\right|\right| \geqslant 2|||X||| \tag{15}$$

for all $X \in \mathscr{B}(\mathscr{H})$ of rank 1 if and only if γS is self-adjoint for some nonzero complex number γ.

Proof. ($\Leftarrow$) The arithmetic–geometric mean inequality, see [2,3,7], states that for every unitarily invariant norm we have

$$|||AA^*Y + YBB^*||| \geqslant 2|||A^*YB|||. \tag{16}$$

Then (15) follows from (16) by letting $A = t\gamma S$, $B = \gamma S$, and $Y = A^{-1}XB^{-1}$.

($\Rightarrow$) By Corollary 2.2 S is normal. If S can be diagonalized, then for every two points $\lambda, \mu \in \sigma(S)$ we can choose a pair of corresponding unit eigenvectors ξ and η and put in (15) the rank one operator $X = \xi \otimes \bar{\eta}$, which shows that $\inf_{t>0} |t\frac{\lambda}{\mu} + \frac{1}{t}\frac{\mu}{\lambda}| \geqslant 2$. Thus $\lambda/\mu \in \mathbb{R}$ by Lemma 2.4 and the spectrum of S lies on a straight line through the origin in the complex plane. In general (if S can not be diagonalized), by the spectral theorem S can be approximated by invertible normal operators with finite spectra (hence diagonal) and the above argument for the diagonal case can be adapted to this approximation. We omit the routine details (they are the same as in [11, Lemma 4.3]). □

3. Proof of Theorem 2.3

In this section q will be a positive integer.

Note that the singular values of a 2×2 matrix X are given by

$$s_{1,2} = \left(\frac{1}{2}\left(\operatorname{tr}(X^*X) \pm \sqrt{\operatorname{tr}^2(X^*X) - 4|\det X|^2}\right)\right)^{1/2}$$

and by definition

$$\|X\|_{2q}^{2q} = s_1^{2q} + s_2^{2q}.$$

Recall that for a 2×2 matrix

$$X = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$$

we have

$$\mathscr{S}_\alpha(X) = \begin{bmatrix} x & \alpha y \\ \alpha z & t \end{bmatrix}.$$

Then

$$|\det X| \leqslant |xt| + |yz| \quad \text{and} \quad |\det(\mathscr{S}_\alpha(X))| \geqslant ||xt| - |\alpha|^2|yz||. \tag{17}$$

Further, note that the function

$$f(u) = (C+u)^q + (C-u)^q \text{ is increasing for } C, u > 0. \tag{18}$$

Thus using the second estimate in (17) we have

$$\begin{aligned} u_1 &= \operatorname{tr}^2(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) - 4|\det \mathscr{S}_\alpha(X)|^2 \\ &\leqslant \operatorname{tr}^2(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) - 4(|xt| - |\alpha|^2|yz|)^2 = u_2. \end{aligned}$$

Hence using (18) with $C = \operatorname{tr}(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X))$ we have

$$\begin{aligned} &2^q\|\mathscr{S}_\alpha(X)\|_{2q}^{2q} \\ &\quad\leqslant \left(\operatorname{tr}(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) + \sqrt{\operatorname{tr}^2(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) - 4(|xt| - |\alpha|^2|yz|)^2}\right)^q \\ &\quad+\left(\operatorname{tr}(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) - \sqrt{\operatorname{tr}^2(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) - 4(|xt| - |\alpha|^2|yz|)^2}\right)^q. \end{aligned} \tag{19}$$

Similarly, using the first estimate in (17) and (18) we get

$$2^q\|X\|_{2q}^{2q} \geqslant \left(\operatorname{tr}(X^*X) + \sqrt{\operatorname{tr}^2(X^*X) - 4(|xt| + |yz|)^2}\right)^q + \left(\operatorname{tr}(X^*X) - \sqrt{\operatorname{tr}^2(X^*X) - 4(|xt| + |yz|)^2}\right)^q. \quad (20)$$

Put

$$|x|^2 + |t|^2 = a, \quad |y|^2 + |z|^2 = c, \quad |x|^2 - |t|^2 = b,$$
$$|y|^2 - |z|^2 = d \quad \text{and} \quad |\alpha|^2 = \lambda.$$

Furthermore, with no loss of generality we may assume that $|x| \geqslant |t|$ and $|y| \geqslant |z|$ (otherwise put $b = |t|^2 - |x|^2$ and similarly for d). Hence a, b, c, d and λ are nonnegative real numbers. Note that with this notation

$$\operatorname{tr}(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) = a + \lambda c$$

and

$$\operatorname{tr}^2(\mathscr{S}_\alpha(X)^*\mathscr{S}_\alpha(X)) - 4(|xt| - |\alpha|^2|yz|)^2 = b^2 + \lambda^2 d^2 + 2\lambda(ac + \sqrt{a^2 - b^2}\sqrt{c^2 - d^2}).$$

Then (19) becomes

$$\begin{aligned} 2^q\|\mathscr{S}_\alpha(X)\|_{2q}^{2q} &\leqslant \left(a + \lambda c + \sqrt{b^2 + \lambda^2 d^2 + 2\lambda(ac + \sqrt{a^2 - b^2}\sqrt{c^2 - d^2})}\right)^q \\ &\quad + \left(a + \lambda c - \sqrt{b^2 + \lambda^2 d^2 + 2\lambda(ac + \sqrt{a^2 - b^2}\sqrt{c^2 - d^2})}\right)^q \\ &\leqslant \left(a + \lambda c + \sqrt{b^2 + \lambda^2 d^2 + 4\lambda ac}\right)^q \\ &\quad + \left(a + \lambda c - \sqrt{b^2 + \lambda^2 d^2 + 4\lambda ac}\right)^q, \end{aligned} \quad (21)$$

where for the last line we used the estimate

$$\sqrt{a^2 - b^2}\sqrt{c^2 - d^2} \leqslant ac$$

and (18). Similarly, (20) becomes

$$\begin{aligned} 2^q\|X\|_{2q}^{2q} &\geqslant \left(a + c + \sqrt{b^2 + d^2 + 2ac - 2\sqrt{a^2 - b^2}\sqrt{c^2 - d^2})}\right)^q \\ &\quad + \left(a + c - \sqrt{b^2 + d^2 + 2ac - 2\sqrt{a^2 - b^2}\sqrt{c^2 - d^2})}\right)^q \\ &\geqslant (a + c + b + d)^q + (a + c - (b + d))^q, \end{aligned} \quad (22)$$

where in the last estimate we used the inequality

$$\sqrt{a^2-b^2}\sqrt{c^2-d^2} \leqslant ac-bd$$

and again (18).

Let

$$L_{2q}(a,b,c,d) = \left(a+\lambda c+\sqrt{b^2+\lambda^2 d^2+4\lambda ac}\right)^q + \left(a+\lambda c-\sqrt{b^2+\lambda^2 d^2+4\lambda ac}\right)^q$$

and

$$R_{2q}(a,b,c,d) = (a+c+b+d)^q + (a+c-(b+d))^q.$$

What we already know is that by (21)

$$2^q \|\mathscr{S}_\alpha(X)\|_{2q}^{2q} \leqslant L_{2q}(a,b,c,d),$$

and by (22) that

$$R_{2q}(a,b,c,d) \leqslant 2^q \|X\|_{2q}^{2q}.$$

Now we are ready to prove the following lemma.

Lemma 3.1. *If q is a positive integer, $\lambda = 1/(2q-1)$ and $a,b,c,d \geqslant 0$, then*

$$L_{2q}(a,b,c,d) \leqslant R_{2q}(a,b,c,d).$$

Proof. Clearly $L_{2q}(a,b,c,d)$ and $R_{2q}(a,b,c,d)$ are polynomials in a,b,c,d, homogeneous of degree q. We claim that all the coefficients of $R_{2q}(a,b,c,d) - L_{2q}(a,b,c,d)$ are nonnegative. We prove our claim by first expressing these coefficients in closed form, then comparing them.

1. *The coefficients of $L_{2q}(a,b,c,d)$*. We have

$$\begin{aligned} L_{2q}(a,b,c,d) &= 2\sum_{i,j} \binom{q}{2i,\ j,\ q-2i-j} a^{q-2i-j}(\lambda c)^j (b^2+\lambda^2 d^2+4\lambda ac)^i \\ &= 2\sum_{i,j,s,t} \binom{q}{2i,\ j,\ q-2i-j}\binom{i}{s,\ t,\ i-s-t} \\ &\quad \times 4^s \lambda^{j+s+2t} a^{q-2i-j+s} b^{2i-2s-2t} c^{j+s} d^{2t} \end{aligned}$$

by using trinomial theorem twice. Each summation index ranges over all integers, with the usual convention that a multinomial coefficient vanishes whenever one of its lower indices exceeds the upper index, or becomes negative. Now replace i by $u = i-s-t$ and j by $v = j+s$ to find that

$$L_{2q}(a,b,c,d) = 2\sum_{t,u,v} \lambda^{2t+v} a^{q-2t-2u-v} b^{2u} c^v d^{2t} S(v) \tag{23}$$

where $S(v) = \sum_s F(v, s)$ and

$$F(v,s) = \binom{q}{2s+2t+2u,\, v-s,\, q-s-2t-2u-v} \binom{s+t+u}{s,t,u} 4^s.$$

The sum $S(v)$ can be evaluated in closed form as follows. Let

$$p_0(v) = (q-2t-2u-v)(2q-2t-2u-2v-1),$$
$$p_1(v) = (v+1)(2t+2u+2v+1),$$
$$G(v,s) = -\frac{g(v)h(s)}{q+1} H(v,s),$$

where

$$g(v) = q-2t-2u-2v-1,$$
$$h(s) = s(2s+2t+2u-1),$$
$$H(v,s) = \binom{q+1}{2s+2t+2u,\, v-s+1,\, q-s-2t-2u-v} \binom{s+t+u}{s,t,u} 4^s.$$

Although these quantities are, in general, functions of s, t, u, v, for brevity we show only dependence on v and s. Denote also

$$f(v,s) = q-s-2t-2u-v.$$

Then the following identities hold for all integer values of s, t, u, v:

$$G(v,s+1) = -2f(v,s)g(v)F(v,s), \tag{24}$$

$$(v-s+1)G(v,s) = -g(v)h(s)F(v,s), \tag{25}$$

$$(v-s+1)F(v+1,s) = f(v,s)F(v,s), \tag{26}$$

$$f(v,s)p_1(v) - (v-s+1)p_0(v) = g(v)h(s) - 2(v-s+1)f(v,s)g(v). \tag{27}$$

Let us prove (24). Note that both sides of (24) vanish whenever $s > v$, or $s > q - 2t - 2u - v - 1$, or any of s, t, u is negative. Otherwise, by repeatedly using the identity $n! = n(n-1)!$ which is valid for $n > 0$, we have

$$\begin{aligned} &H(v,s+1) \\ &\quad= \frac{(q+1)!(s+t+u+1)!4^{s+1}}{(2s+2t+2u+2)!(v-s)!(q-s-2t-2u-v-1)!(s+1)!t!u!} \\ &\quad= \frac{2(q+1)(q-s-2t-2u-v)q!(s+t+u)!4^s}{(2s+2t+2u+1)(s+1)(2s+2t+2u)!(v-s)!(q-s-2t-2u-v)!s!t!u!} \\ &\quad= \frac{2(q+1)f(v,s)}{h(s+1)} F(v,s), \end{aligned}$$

and so

$$G(v, s+1) = -\frac{g(v)h(s+1)}{q+1} H(v, s+1) = -2f(v, s)g(v)F(v, s)$$

as claimed. The proofs of (25) and (26) are analogous, and (27) is just a polynomial identity.

Multiplying (27) by $F(v, s)$ and using (24)–(26) we obtain

$$\begin{aligned}&(v - s + 1)[p_1(v)F(v + 1, s) - p_0(v)F(v, s)]\\ &\quad = (v - s + 1)[G(v, s + 1) - G(v, s)].\end{aligned}$$

So if $s \neq v + 1$ we have

$$p_1(v)F(v + 1, s) - p_0(v)F(v, s) = G(v, s + 1) - G(v, s). \tag{28}$$

But this holds also when $s = v + 1$ because $F(v, v + 1) = G(v, v + 2) = 0$ and $p_1(v)F(v + 1, v + 1) = -G(v, v + 1)$, as can be readily checked.

By summing the "telescoping" recurrence (28) over all integer s we obtain

$$p_1(v)S(v + 1) - p_0(v)S(v) = 0.$$

Together with the obvious initial condition $S(0) = \binom{q}{2t+2u}\binom{t+u}{u}$, this recurrence implies that

$$S(v) = \binom{q}{2t+2u}\binom{t+u}{u}\binom{q-2t-2u}{v}\frac{\binom{q-t-u-1/2}{v}}{\binom{t+u+v-1/2}{v}}. \tag{29}$$

With (23) this gives the coefficients of $L_{2q}(a, b, c, d)$ in closed form. We note that $p_0(v)$, $p_1(v)$ and $G(v, s)$ were computed by means of Zeilberger's algorithm [10,13].

2. *The coefficients of* $R_{2q}(a, b, c, d)$. By the quadrinomial theorem we have

$$\begin{aligned}R_{2q}(a, b, c, d) &= 2\sum_{\substack{i,j,k\\ i+j=2n}} \binom{q}{i, j, k, q-i-j-k} a^{q-i-j-k} b^j c^k d^i\\ &= 2\sum_{i,k,n} \binom{q}{i, 2n-i, k, q-k-2n} a^{q-k-2n} b^{2n-i} c^k d^i. \end{aligned} \tag{30}$$

3. *Comparing the coefficients.* From (23) it follows that the coefficient of $a^{q-t-u-v} b^u c^v d^t$ in $L_{2q}(a, b, c, d)$ is zero unless t and u are even. As the coefficients of $R_{2q}(a, b, c, d)$ are nonnegative, it suffices to show that the coefficient of $a^{q-2t-2u-v} b^{2u} c^v d^{2t}$ in $L_{2q}(a, b, c, d)$ does not exceed the corresponding coefficient in $R_{2q}(a, b, c, d)$. Thus, by (23), (29) and (30), we need to show that

$$2\frac{\binom{q}{2t+2u}\binom{t+u}{u}\binom{q-2t-2u}{v}\binom{q-t-u-1/2}{v}}{(2q-1)^{2t+v}\binom{t+u+v-1/2}{v}}$$
$$\leqslant 2\binom{q}{2t,2u,v,q-2t-2u-v},$$

or equivalently, that

$$\binom{t+u}{u}\binom{q-t-u-1/2}{v}\leqslant\binom{2t+2u}{2t}\binom{t+u+v-1/2}{v}(2q-1)^{2t+v}.$$

This inequality will follow immediately from

$$\binom{t+u}{u}\leqslant\binom{2t+2u}{2t} \tag{31}$$

and

$$\binom{q-t-u-1/2}{v}\leqslant\binom{t+u+v-1/2}{v}(2q-1)^{v}. \tag{32}$$

To prove (31) rewrite it as

$$\prod_{i=1}^{u}(u+i)\prod_{j=1}^{t}(t+j)\leqslant\prod_{i=1}^{u}(t+u+i)\prod_{j=1}^{t}(t+2u+j)$$

where it becomes obvious. To prove (32) rewrite it as

$$\prod_{i=1}^{v}\left(q-t-u-\frac{2i-1}{2}\right)\leqslant\left(q-\frac{1}{2}\right)^{v}\prod_{i=1}^{v}(2t+2u+2v-2i+1)$$

which is true because each factor on the left is at most $q-1/2$, while each factor in the right-hand side product is at least 1. □

Proof of Theorem 2.3. Let n be an even integer. Then by Lemma 3.1 (and the calculation preceding the lemma) $\|\mathscr{S}_\alpha\|_n=1$ if $|\alpha|=1/\sqrt{n-1}$. Since the set of all $\alpha\in\mathbb{C}$ satisfying $\|\mathscr{S}_\alpha\|_n\leqslant 1$ is convex, it must contain the disc with the center 0 and radius $1/\sqrt{n-1}$. But it is known, see [9], that $\|\mathscr{S}_\alpha\|_p\leqslant\|\mathscr{S}_\alpha\|_n$ if $2\leqslant p\leqslant n$. This proves the first part of the theorem. The second part then follows by duality. □

The following example shows that the above estimates are sharp.

Example 3.2. Let

$$A=\begin{bmatrix}1 & ia\\ ia & 1\end{bmatrix}\quad\text{and}\quad X=\begin{bmatrix}1 & \varepsilon\\ -\varepsilon & 1\end{bmatrix},$$

where $0 < a, \varepsilon < 1$. Then the singular values of $\mathscr{S}_{ia}(X)$ are $1 \pm a\varepsilon$ and the singular values of X are both equal to $\sqrt{1+\varepsilon^2}$. Thus the condition $\|\mathscr{S}_{ia}(X)\|_p^p \leqslant \|X\|_p^p$ means

$$(1 + a\varepsilon)^p + (1 - a\varepsilon)^p \leqslant 2(1 + \varepsilon^2)^{p/2},$$

hence

$$\binom{p}{2} a^2 + \binom{p}{4} a^4 \varepsilon^2 + \cdots \leqslant \frac{p}{2} + \binom{p/2}{2} \varepsilon^2 + \cdots$$

Letting $\varepsilon \to 0$ it follows that $\binom{p}{2} a^2 \leqslant p/2$, thus (if $p \geqslant 2$) $a \leqslant 1/\sqrt{p-1}$.

References

[1] S.K. Berberian, Approximate proper vectors, Proc. Amer. Math. Soc. 13 (1962) 111–114.
[2] R. Bhatia, Matrix Analysis, Springer Verlag, New York, 1997.
[3] R. Bhatia, C. Davis, More matrix forms of the arithmetic–geometric mean inequality, SIAM J. Matrix Anal. Appl. 14 (1993) 132–136.
[4] G. Corach, H. Porta, L. Recht, An operator inequality, Linear Algebra Appl. 142 (1990) 153–158.
[5] I.C. Gohberg, M.G. Krein, Introduction to the Theory of Linear Nonselfadjoint Operators, American Mathematical Society, Providence, RI, 1969.
[6] M. Hladnik, Schur norms of bicirculant matrices, Linear Algebra Appl. 286 (1999) 261–272.
[7] F. Kittaneh, A note on the arithmetic–geometric mean inequality for matrices, Linear Algebra Appl. 171 (1992) 1–8.
[8] F. Kittaneh, On some operator inequalities, Linear Algebra Appl. 208/209 (1994) 19–28.
[9] K. Okubo, Norm of Schur multiplication for Schatten norm, Proc. Amer. Math. Soc. 111 (1991) 925–930.
[10] M. Petkovšek, H.S. Wilf, D. Zeilberger, $A = B$, A K Peters, Wellesley, MA, 1996.
[11] A. Seddik, Some results related to the Corach–Porta–Recht inequality, Proc. Amer. Math. Soc. 129 (2001) 3009–3015.
[12] D. Xia, Spectral Theory of Hyponormal Operators, Birkhäuser Verlag, Basel, 1983.
[13] D. Zeilberger, The method of creative telescoping, J. Symbolic Comput. 11 (1991) 195–204.