Mathematical Inequalities & Applications
Volume 24, Number 2 (2021), 431–444 doi:10.7153/mia-2021-24-30

# MATRIX REARRANGEMENT INEQUALITIES REVISITED

VICTORIA M. CHAYES

(*Communicated by J.-C. Bourin*)

*Abstract.* Let $||X||_p = \mathrm{Tr}[(X^*X)^{p/2}]^{1/p}$ denote the $p$-Schatten norm of a matrix $X \in M_{n\times n}(\mathbb{C})$, and $\sigma(X)$ the singular values with $\uparrow\ \downarrow$ indicating its increasing or decreasing rearrangements. We wish to examine inequalities between $||A+B||_p^p + ||A-B||_p^p$, $||\sigma_\downarrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\downarrow(A)-\sigma_\downarrow(B)||_p^p$, and $||\sigma_\uparrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\uparrow(A)-\sigma_\downarrow(B)||_p^p$ for various values of $1 \leqslant p < \infty$. It was conjectured in [6] that a universal inequality $||\sigma_\downarrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\downarrow(A)-\sigma_\downarrow(B)||_p^p \leqslant ||A+B||_p^p + ||A-B||_p^p \leqslant ||\sigma_\uparrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\uparrow(A)-\sigma_\downarrow(B)||_p^p$ might hold for $1 \leqslant p \leqslant 2$ and reverse at $p \geqslant 2$, potentially providing a stronger inequality to the generalization of Hanner's Inequality to complex matrices $||A+B||_p^p + ||A-B||_p^p \geqslant (||A||_p + ||B||_p)^p + |||A||_p - ||B||_p|^p$. We extend some of the cases in which the inequalities of [6] hold, but offer counterexamples to any general rearrangement inequality holding. We simplify the original proofs of [6] with the technique of majorization. This also allows us to characterize the equality cases of all of the inequalities considered. We also address the commuting, unitary, and $\{A,B\}=0$ cases directly, and expand on the role of the anticommutator. In doing so, we extend Hanner's Inequality for self-adjoint matrices to the $\{A,B\}=0$ case for all ranges of $p$.

## 1. Introduction

It has been of great interest to extend Hanner's Inequality for $L^p$ spaces

$$||f+g||_p^p + ||f-g||_p^p \geqslant (||f||_p + ||g||_p)^p + |||f||_p - ||g||_p|^p \tag{1}$$

for $1 \leqslant p \leqslant 2$ to the non-communative analogue in $C^p$ of matrix operators under the $p$-Schatten norm

$$||A+B||_p^p + ||A-B||_p^p \geqslant (||A||_p + ||B||_p)^p + |||A||_p - ||B||_p|^p. \tag{2}$$

In [6], Carlen and Lieb proposed the following two conjectures for their potential pertinence to proving (2):

CONJECTURE 1. *For all $1 \leqslant p \leqslant 2$, and all complex-valued $n\times n$ matrices $A$ and $B$,*

$$||A+B||_p^p + ||A-B||_p^p \geqslant ||\sigma_\downarrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\downarrow(A)-\sigma_\downarrow(B)||_p^p. \tag{3}$$

*For $p > 2$, the inequality reverses.*

*Mathematics subject classification* (2010): 15A42.
*Keywords and phrases*: Matrix inequality, Hanner's inequality, singular value inequalities, $p$-Schatten norm, majorization.
This research was funded by the NDSEG Fellowship, Class of 2017. Thank you to my advisor, Professor Eric Carlen, for bringing my attention to the problem and providing me with a background to the subject, and to Jean-Christophe Bourin for discussions on related inequalities.

CONJECTURE 2. *For all $1 \leqslant p \leqslant 2$, and all complex-valued $n \times n$ matrices $A$ and $B$,*

$$||A+B||_p^p + ||A-B||_p^p \leqslant ||\sigma_\uparrow(A) + \sigma_\downarrow(B)||_p^p + ||\sigma_\uparrow(A) - \sigma_\downarrow(B)||_p^p. \tag{4}$$

*For $p > 2$, the inequality reverses.*

For these, the authors proved Conjecture 1 in the case $A \geqslant B \geqslant 0$ and $1 \leqslant p \leqslant 2$; and proved Conjecture 2 in the case $A \geqslant |B| \geqslant 0$ and $1 \leqslant p \leqslant 2$. We note a missing requirement in [6] used in the proof for Conjecture 2 in those conditions is also that $\sigma_n(A) \geqslant \sigma_1(B)$. Lemma 2.1 in [16] proves that Conjecture 1 holds for all matrices and $p = 2k$, $k \in \mathbb{N}$. To the best of our knowledge, no further work has been done on the subject.

If Conjecture 1 were true in general, with an additional application of Hanner's Inequality on the sequences of singular values, the non-commutative Hanner's Inequality for matrices would be proven in general; currently, it is only known for $A+B, A-B \geqslant 0$ for all $p$, or general $A$ and $B$ in the ranges $1 \leqslant p \leqslant \frac{4}{3}$ and $p \geqslant 4$ [3].

In this paper we extend the range of Conjecture 1 with the requirements $A \geqslant B \geqslant 0$ to $2 \leqslant p \leqslant 3$, and we prove Conjecture 2 in the $A+B, A-B \geqslant 0$, $\sigma_n(A) \geqslant \sigma_1(B)$ case for the range $1 \leqslant p \leqslant 3$. We prove both conjectures for the full range of $p$ in the commuting case. We prove Conjecture 1 in the case that $A$ and $B$ are both unitary, and in the case when $A$ and $B$ are self-adjoint and $\{A, B\} = 0$. However, we demonstrate that both conjectures are false in general. Section 2 gives a background to majorization, which is the primary technique that we use in our proofs. Section 3 presents the extensions of the conjectures' requirements and ranges, and general counterexamples.

The key observation as to why the conjectures cannot hold in general is that if the matrix $B$ is taken to be unitary, all its singular values are equal to 1, and therefore there is no distinction between the "aligned" and the "up-down" rearrangements. If both conjectures were true, this would imply equality everywhere when $B$ is unitary, which can easily be numerically confirmed as false.

The fact that these re-arrangement inequalities do not hold in general is notable, because the analogue to Conjecture 1 with complex functions and the spherically symmetric decreasing rearrangement is shown in [6] to hold. Therefore, we see directly that the non-commutativity of the matrices ruins a commutative identity. In disproving Conjecture 1 in general, we also rule it out as a method to attempt to extend Hanner's inequality to $C^p$.

We will use the following notation throughout this paper: $\sigma(X)$ denotes the vector of singular values of a matrix $X$, assumed to be in descending order unless $\sigma_\uparrow(X)$ is specified; $\sigma_\downarrow(X)$ may then be used for emphasis. The norm $||\cdot||_p$ may either indicate the vector $p$-norm or the $p$-Schatten norm dependent on context. We use for a vector $\mathbf{v}$ the notation $[\mathbf{v}]$ to indicate the matrix $[\mathrm{Diag}(\mathbf{v})]$.

## 2. Majorization

Let $\mathbf{a},\mathbf{b} \in \mathbb{R}^n$ with components labeled in descending order $a_1 \geqslant \ldots \geqslant a_n$ and $b_1 \geqslant \ldots \geqslant b_n$. It is said that $\mathbf{b}$ weakly majorizes $\mathbf{a}$, written $\mathbf{a} \prec_w \mathbf{b}$, when

$$\sum_{i=1}^{k} a_i \leqslant \sum_{i=1}^{k} b_i, \qquad 1 \leqslant k \leqslant n \tag{5}$$

and $\mathbf{b}$ majorizes $\mathbf{a}$, written $\mathbf{a} \prec \mathbf{b}$, when the final inequality is an equality. Weak log majorization $\mathbf{a} \prec_{w(\log)} \mathbf{b}$ is similarly defined for non-negative vectors as

$$\prod_{i=1}^{k} a_i \leqslant \prod_{i=1}^{k} b_i, \qquad 1 \leqslant k \leqslant n \tag{6}$$

with log majorization $\mathbf{a} \prec_{(\log)} \mathbf{b}$ when the final inequality is an equality. An important fact is that log majorization and weak log majorization both imply weak majorization [2] [Lemma 1.8].

Note that it is not necessary that the vectors $\mathbf{a}$ and $\mathbf{b}$ be in descending order—majorization is explicitly defined with respect the the rearrangements of the values in descending order. We define all of the above majorization for matrices, i.e. $A \prec B$ and all variations, when the singular values considered as a vector are majorized $\sigma(A) \prec \sigma(B)$. All operators stated for majorization (i.e. $f(\mathbf{a})$ or $\mathbf{ab}$) should be considered to be applied entrywise to the vectors (i.e. $(f(a_1),\ldots,f(a_n))$ or $(a_1b_1,\ldots,a_nb_n)$.

Majorization holds the following vital property:

THEOREM 3. (Hardy, Littlewood, and Pólya [9] [10]; Tomić, Weyl [17] [18]) *Suppose $\mathbf{a} \prec_w \mathbf{b}$. Then for any function $f : \mathbb{R} \to \mathbb{R}$ that is increasing and convex on the domain containing all elements of $\mathbf{a}$ and $\mathbf{b}$,*

$$\sum_{i=1}^{n} f(a_i) \leqslant \sum_{i=1}^{n} f(b_i). \tag{7}$$

*If $\mathbf{a} \prec \mathbf{b}$, the 'increasing' requirement can be dropped.*

An immediate yet highly useful lemma follows:

LEMMA 4. *Let $\mathbf{a},\mathbf{b} \in \mathbb{R}_n^+$. Suppose $\mathbf{a} \prec_w \mathbf{b}$. Then $\mathbf{a}^s \prec_w \mathbf{b}^s$ for all $s \geqslant 1$.*

Log majorization also allows us to characterize equality cases:

LEMMA 5. (Hiai [11] [Lemma 2.2]) *Let $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ be a strictly convex increasing function. Then $\mathbf{a} \prec_{(\log)} \mathbf{b}$ and $\sum_{i=1}^n \Phi(a_i) = \sum_{i=1}^n \Phi(b_i)$ imply $\mathbf{a} = \Theta\mathbf{b}$ for some permutation matrix $\Theta$.*

As exponentiating is strictly convex, an immediate corollary is $\mathbf{a} \prec_{(\log)} \mathbf{b}$ and $\mathbf{a} \prec \mathbf{b}$ imply $\mathbf{a} = \Theta\mathbf{b}$.

Majorization is an incredibly powerful technique in matrix analysis used to prove numerous inequalities about eigenvalues and singular values of matrices, powers of products of positive semidefinite matrices, Golden-Thompson-like inequalities, and more. A good overview of the techniques and important results can be found in [12], [14]. The two results that we will need for this paper will regard the eigenvalues of the sums of Hermitian matrices and the singular values of products of general matrices:

THEOREM 6. (Fan [7]) *Let* $A,B \in M_{n\times n}(\mathbb{C})$ *be self-adjoint. Then*

$$\lambda(A+B) \prec \lambda(A)+\lambda(B). \tag{8}$$

THEOREM 7. (Horn [13]; Gel'fand and Naimark [8]) *Let* $A,B \in M_{n\times n}(\mathbb{C})$. *Then*

$$\sigma(AB) \prec_{(\log)} \sigma(A)\sigma(B). \tag{9}$$

We will also need a fairly intuitive lemma that to our knowledge has not yet been addressed in existing literature, characterizing the concatenation of majorized vectors:

LEMMA 8. *Let* $\mathbf{x} \prec_w \mathbf{y}$, *and* $\mathbf{a} \prec_w \mathbf{b}$ *be non-negative vectors labeled in descending order. Then* $\mathbf{xa} \prec_w \mathbf{yb}$.

*Proof.* We can write the components of $\mathbf{y}$ as $y_{n-1} = y_n + \varepsilon_1$, $y_{n-2} = y_n + \varepsilon_1 + \varepsilon_2$, $\ldots$, $y_1 = y_n + \varepsilon_1 + \ldots + \varepsilon_{n-1}$ where $\varepsilon_i \geqslant 0$. Then applying $\mathbf{a} \prec_w \mathbf{b}$

$$\varepsilon_{n-1}a_1 \leqslant \varepsilon_{n-1}b_1 \tag{10}$$

$$\varepsilon_{n-2}\left(\sum_{i=1}^{2} a_i\right) \leqslant \varepsilon_{n-2}\left(\sum_{i=1}^{2} b_i\right) \tag{11}$$

$$\vdots \tag{12}$$

$$y_n\left(\sum_{i=1}^{n} a_i\right) \leqslant y_n\left(\sum_{i=1}^{n} b_i\right) \tag{13}$$

and summing them all together,

$$\sum_{i=1}^{n} y_i a_i \leqslant \sum_{i=1}^{n} y_i b_i. \tag{14}$$

Applying the same splitting argument to $a_i$ with $\mathbf{x} \prec_w \mathbf{y}$ gives

$$\sum_{i=1}^{n} x_i a_i \leqslant \sum_{i=1}^{n} y_i a_i, \tag{15}$$

and stringing the two inequalities together

$$\sum_{i=1}^{n} x_i a_i \leqslant \sum_{i=1}^{n} y_i b_i. \tag{16}$$

Finally, nothing that when $\mathbf{a} \prec_w \mathbf{b}$, the first $k$th components maintain the weak majorization relationship $(a_1,\dots,a_k) \prec_w (b_1,\dots,b_k)$, applying the argument to the $k$th components gives the desired result. □

Note that the above technique can be expressed compactly as the weigted sum of Ky-Fan norms for matrices $[\mathbf{a}],[\mathbf{b}],[\mathbf{x}]$, and $[\mathbf{y}]$, and leveraging the matrix majorization result that $A \prec_w B$ implies $|||A||| \leqslant |||B|||$ for every unitarily invariant norm $|||\cdot|||$.

## 3. Extensions and counterexamples

First, we address rearrangement of commuting matrices:

THEOREM 9. *Let $A,B \in M_{n\times n}(\mathbb{C})$ be two self-adjoint matrices that commute. Then*

$$\begin{aligned} ||\sigma(A)+\sigma(B)||_p^p + ||\sigma(A)-\sigma(B)||_p^p &\leqslant ||A+B||_p^p + ||A-B||_p^p \\ &\leqslant ||\sigma_\uparrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\uparrow(A)-\sigma_\downarrow(B)||_p^p \end{aligned} \tag{17}$$

*for $1 \leqslant p \leqslant 2$, and the inequality reverses for $p \geqslant 2$. Furthermore, there is equality in either inequality for $p \neq 1,2$ if and only if the singular values of $A$ and $B$ are aligned in descending or ascending-descending order respectively.*

*Proof.* In the simultaneously diagonalizable basis, we can write

$$A = \begin{bmatrix} \lambda_1(A) & & \\ & \ddots & \\ & & \lambda_n(A) \end{bmatrix}, \qquad B = \begin{bmatrix} \lambda_{i_1}(B) & & \\ & \ddots & \\ & & \lambda_{i_n}(B) \end{bmatrix}. \tag{18}$$

Then we note that the singular values of $A+B$ and $A-B$ can be grouped as

$$\{|\lambda_j(A) \pm \lambda_{i_j}(B)|\} = \{||\lambda_j(A)| \pm |\lambda_{i_j}(B)||\} = \{|\sigma_i(A) \pm \sigma_{k_i}(B)|\}. \tag{19}$$

Re-labeling to preserve the pairings above, we consider the functions

$$f(x) = \sum_{i=1}^n \sigma_i(A)\chi_{[i-1,i)}(x) \tag{20}$$

$$g(x) = \sum_{i=1}^n \sigma_{k_i}(B)\chi_{[i-1,i)}(x) \tag{21}$$

and let $f^*$ and $g^*$ denote the symmetric decreasing rearrangements.

We will need an extension of the Riesz rearrangement inequality:

LEMMA 10. (Almgren, Lieb [1] Theorem 2.2) *Let $F : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous function such that $F(0,0)=0$ and*

$$F(u_2,v_2)+F(u_1,v_1) \geqslant F(u_2,v_1)+F(u_1,v_2) \tag{22}$$

*whenever* $u_1 \geqslant u_2 \geqslant 0$ *and* $v_1 \geqslant v_2 \geqslant 0$. *Let* $\mathscr{L}^n$ *denote the* $n$*-dimensional Lebesgue measure on* $\mathbb{R}^n$. *Let* $f,g : \mathbb{R}^n \to \mathbb{R}^+$ *be measurable functions such that spherically symmetric decreasing rearrangements* $f^*$, $g^*$ *are well-defined (i.e.* $\mu_{\mathscr{L}^n}(f(x) > y) < \infty$ *for all* $y$*). Then the inequality*

$$\int F(f(x),g(x))\, d\mathscr{L}^n x \leqslant \int F(f^*(x),g^*(x))\, d\mathscr{L}^n x \tag{23}$$

*holds. If condition (22) is reversed, the inequality (23) is reversed.*

This lemma results from the following technique taking use of the Riesz rearrangement inequality: we can consider $F$ as the limit of twice-differentiable functions with $W(x-y) = \varepsilon^{-1}\exp(|x-y|/\varepsilon)$ and take the limit as $\varepsilon \to 0$. One can express

$$\begin{aligned}\int\int F(f(x),g(y))W(x-y)\,\mathrm{d}\mathscr{L}^n x\,\mathrm{d}\mathscr{L}^n y \\ = \int\int F_{12}(s,t)\left[\int\int \chi_{\{f>s\}}(x)\chi_{\{g>t\}}(y)W(x-y)\,\mathrm{d}\mathscr{L}^n x\,\mathrm{d}\mathscr{L}^n y\right]\mathrm{d}\mathscr{L}^1 s\,\mathrm{d}\mathscr{L}^1 t\end{aligned} \tag{24}$$

and apply the Riesz rearrangement inequality to the interior integral. When we take the limit, $W$ converges to the delta distribution, and we have

$$\begin{aligned}\int\int F_{12}(s,t)\left[\int\int \chi_{\{f>s\}}(x)\chi_{\{g>t\}}(y)\delta(x-y)\,\mathrm{d}\mathscr{L}^n x\,\mathrm{d}\mathscr{L}^n y\right]\mathrm{d}\mathscr{L}^1 s\,\mathrm{d}\mathscr{L}^1 t \\ \geqslant \int\int F_{12}(s,t)\left[\int\int \chi_{\{f^*>s\}}(x)\chi_{\{g^*>t\}}(y)\delta(x-y)\,\mathrm{d}\mathscr{L}^n x\,\mathrm{d}\mathscr{L}^n y\right]\mathrm{d}\mathscr{L}^1 s\,\mathrm{d}\mathscr{L}^1 t\end{aligned} \tag{25}$$

The following technique is inspired by [6] [Lemma 1.1], which in fact proves a more general theorem on symmetric decreasing arrangements of general complex functions. For the left half of our inequality, we choose

$$F(x,y) = |x+y|^p + |x-y|^p. \tag{26}$$

We see that $\partial^2 F(x,y)/\partial x \partial y \leqslant 0$ when $1 < p \leqslant 2$, with the inequality switching at $p = 2$, satisfying the condition of Equation (22). Then

$$||f+g||_p^p + ||f-g||_p^p \geqslant ||f^*+g^*||_p^p + ||f^*-g^*||_p^p \tag{27}$$

for $1 < p \leqslant 2$ (and taking the limit for $p = 1$), with the inequality switching for $p \geqslant 2$. As $||f \pm g||_p^p = ||A \pm B||_p^p$ and $||f^* \pm g^*||_p^p = ||\sigma(A) \pm \sigma(B)||_p^p$, the left half of Equation (17) is proven.

For the right half, without loss of generality, we can assume that $B$ is invertible; otherwise, we could consider a limit of perturbations. As the inequality for matrices $A,B$ holds if and only if it holds for $cA, cB$ for some scaling constant $c$, we can further assume that the largest singular value of $B$ is equal to 1. We define piecewise functions such as

$$F(x,y)=\begin{cases}\left|x+\frac{1}{y}\right|^p+\left|x-\frac{1}{y}\right|^p & x\geqslant 0,\ y\geqslant 1\\ e^{1-\frac{1}{y}}\left(\left|x+\frac{1}{y}\right|^p+\left|x-\frac{1}{y}\right|^p\right) & x\geqslant 0,\ 0\leqslant y<1\end{cases}\tag{28}$$

for $1\leqslant p\leqslant 2$ and

$$F(x,y)=\begin{cases}\left|x+\frac{1}{y}\right|^p+\left|x-\frac{1}{y}\right|^p & x\geqslant 0,\ y\geqslant 1\\ e^{p(1-y)}\left(|x+y|^p+|x-y|^p\right) & x\geqslant 0,\ 0\leqslant y<1\end{cases}\tag{29}$$

for $p\geqslant 2$. The values of the function $F$ that we care about will be in the $y\geqslant 1$ range; these are merely examples of functions existing that satisfy the necessary conditions to apply Lemma 10.

It can be readily confirmed that $F(x,y)$ is continuous, and by exponential domination in the limit $F(0,0)=0$. We therefore calculate the partial derivative on each piece and see that $\partial^2F(x,y)/\partial x\partial y\geqslant 0$ when $1<p\leqslant 2$, with the inequality reversing at $p=2$. Then letting

$$f(x)=\sum_{i=1}^{n}\sigma_i(A)\chi_{[i-1,i)}(x)\tag{30}$$

$$g(x)=\sum_{i=1}^{n}(\sigma_{k_i}(B))^{-1}\chi_{[i-1,i)}(x)\tag{31}$$

and comparing $\int F(f(x),g(x))\,\mathrm{d}x$ and $\int F(f^*(x),g^*(x))\,\mathrm{d}x$ (and taking the limit for $p=1$), the full inequality is proven.

To characterize the equality cases, we look closer at Line (25). For piecewise $f$ and $g$ as defined in Equations (20) and (21) or (30) and (31), there is strict inequality in application of the Riesz rearrangement inequality for each $W(x-y)=\varepsilon^{-1}\exp(|x-y|/\varepsilon)$ if the functions are not aligned [5]. For the choice of $f$ and $g$ as defined,

$$\int\int\chi_{\{f>s\}}(x)\chi_{\{g>t\}}(y)\delta(x-y)\,\mathrm{d}\mathscr{L}^nx\,\mathrm{d}\mathscr{L}^ny=\ell\tag{32}$$

where $\ell$ is the number of pairs $\{\sigma_i(A),\sigma_{k_i}(B)\}$ (or $\{\sigma_i(A),\sigma_{k_i}(B)^{-1}\}$) that satisfy $\sigma_i(A)>s$ and $\sigma_{k_i}(B)>t$ (or the $\sigma_i(A)>s,\sigma_{k_i}(B)^{-1}>t$). By the Riesz rearrangement inequality as applied in Line (25), we know that

$$\int\int\chi_{\{f^*>s\}}(x)\chi_{\{g^*>t\}}(y)\delta(x-y)\,\mathrm{d}\mathscr{L}^nx\,\mathrm{d}\mathscr{L}^ny=\ell'\geqslant\ell\tag{33}$$

for some other integer $\ell'$, which now represents the number of pairs $\{\sigma_i(A),\sigma_i(B)\}$ (or similarly $\{\sigma_i(A),\sigma_i(B)^{-1}\}$) that satisfy $\sigma_i(A)>s$ and $\sigma_i(B)>t$ (or likewise the inverse). We note that for both of our choices of $F$, we have $F_{12}(s,t)>0$ for $s,t>0$. For the inequality to be strict, there simply must exist intervals of $[s_0,s_1]$ and $[t_0,t_1]$ where $\ell'>\ell$, which upon inspection for any pair of unaligned matrices will be true, as aligning the pairs of singular values will increase the number of pairs that satisfy those desired conditions in at least some range. This means that the interior integral will be strictly greater when $f$ and $g$ are in decreasing rearrangements, and we conclude that the inequality (25) must be strict. $\square$

Next, we address the case when anticommutator $\{A,B\}=0$.

THEOREM 11. *Let $A,B\in M_{n\times n}(\mathbb{C})$ be self-adjoint such that $AB+BA=0$. Then*

$$||A+B||_p^p+||A-B||_p^p \geqslant ||\sigma(A)+\sigma(B)||_p^p+||\sigma(A)-\sigma(B)||_p^p \tag{34}$$

*for $1\leqslant p\leqslant 2$, with the inequality reversing for $p\geqslant 2$.*

*Proof.* We note that as $\lambda(X^2)=\sigma(X^2)$ for sef-adjoint $X$,

$$||A+B||_p^p=\sum_{i=1}^n \lambda_i((A+B)^2)^{p/2}=\sum_{i=1}^n \lambda_i(A^2+B^2)^{p/2}=\sum_{i=1}^n \lambda_i((A-B)^2)^{p/2}=||A-B||_p^p. \tag{35}$$

When $1\leqslant p\leqslant 2$, we make use of the majorization identity of Theorem 6 of $\lambda(A+B)\prec\lambda(A)+\lambda(B)$ and the fact that $f(x)=x^{p/2}$ is concave to conclude that

$$\begin{aligned}
||A+B||_p^p+||A-B||_p^p &= 2\sum_{i=1}^n \lambda_i(A^2+B^2)^{p/2} && (36)\\
&\geqslant 2\sum_{i=1}^n (\lambda_i(A^2)+\lambda_i(B^2))^{p/2} && (37)\\
&= 2\sum_{i=1}^n (\sigma_i(A)^2+\sigma_i(B)^2)^{p/2} && (38)\\
&= 2\sum_{i=1}^n \left(\frac{(\sigma_i(A)+\sigma_i(B))^2}{2}+\frac{(\sigma_i(A)-\sigma_i(B))^2}{2}\right)^{p/2} && (39)\\
&\geqslant \sum_{i=1}^n ((\sigma_i(A)+\sigma_i(B))^2)^{p/2}+((\sigma_i(A)-\sigma_i(B))^2)^{p/2} && (40)\\
&= ||\sigma(A)+\sigma(B)||_p^p+||\sigma(A)-\sigma(B)||_p^p. && (41)
\end{aligned}$$

An identical argument for $p\geqslant 2$ with reversed inequalities can be made now leveraging convexity of $x^{p/2}$. Note that this proof extends to general $A,B$ when $AB^*+BA^*=0$. $\square$

The unitary case gives some insight to the role of the anticommutator.

THEOREM 12. *Let $U,V\in M_{n\times n}(\mathbb{C})$ be unitary. Then*

$$||U+V||_p^p+||U-V||_p^p \geqslant 2^p n \tag{42}$$

*for $1\leqslant p\leqslant 2$, with the inequality switching for $p\geqslant 2$. There is equality for $p\neq 2$ if and only if $U=V$. The extremization of the inequality is directly dependent on $\sigma(UV+VU)$, with greatest difference when $\{U,V\}=0$.*

*Proof.* Note that Equation (42) can be directly derived from [15] the Clarkson type inequalities

$$2(||A||_p^p+||B||_p^p)\leqslant ||A+B||_p^p+||A-B||_p^p\leqslant 2^{p-1}(||A||_p^p+||B||_p^p) \tag{43}$$

for $p\geqslant 2$ and reversing for $1\leqslant p\leqslant 2$; and in fact can be seen from direct matrix inequalities of Theorems 2.1 and 2.5 of [4]. However, we can use majorization to examine this inequality on the level of the eigenvalues to see the direct role of the anticommutator.

We can assume without loss of generality that $U$ and $V$ are also self-adjoint; otherwise, consider the unitary matrices

$$\widehat{U}=\begin{bmatrix}0 & U\\ U^* & 0\end{bmatrix},\qquad \widehat{V}=\begin{bmatrix}0 & V\\ V^* & 0\end{bmatrix} \tag{44}$$

then the inequality holds for $U,V$ if and only if it holds for $\widehat{U},\widehat{V}$ by dividing by the appropriate factor of 2.

Once more we will make use of $\lambda((U\pm V)^2)=\sigma((U\pm V)^2)$. We note that $UV+VU$ is a Hermitian matrix, and as $||UV+VU||\leqslant ||UV||+||VU||\leqslant 2||U||||V||=2$, the eigenvalues of $UV+VU$ must be within the interval $[-2,2]$, and can be written as $2\cos(\theta_j)$. Then

$$\begin{aligned}
||U+V||_p^p+||U-V||_p^p &= ||(U+V)^2||_{p/2}^{p/2}+||(U-V)^2||_{p/2}^{p/2} && (45)\\
&= ||\lambda(2I+UV+VU)||_{p/2}^{p/2}+||\lambda(2I-UV-VU)||_{p/2}^{p/2} && (46)\\
&= ||2+\lambda(UV+VU)||_{p/2}^{p/2}+||2-\lambda(UV+VU)||_{p/2}^{p/2} && (47)\\
&= \sum_{j=1}^n 2^{p/2}|1+\cos(\theta_j)|^{p/2}+2^{p/2}|1-\cos(\theta_j)|^{p/2}. && (48)
\end{aligned}$$

The function $f(\theta)=(1+\cos(\theta))^s+(1-\cos(\theta))^s$ can be examined on the interval $(0,\frac{\pi}{2})$. It has derivative $s\sin(\theta)[(1-\cos(\theta))^{s-1}-(1+\cos(\theta))^{s-1}]$, which can only be 0 at $\theta=0,\theta=\frac{\pi}{2}$. It is immediately confirmed that the function monotone for all $s\geqslant 0$ is minimized at $\theta=0$ and maximized at $\theta=\frac{\pi}{2}$ for $0\leqslant s\leqslant 1$, with the maximum and minimum reversing for $s\geqslant 1$. Therefore,

$$\sum_{j=1}^n 2^{p/2}|1+\cos(\theta_j)|^{p/2}+2^{p/2}|1-\cos(\theta_j)|^{p/2}\geqslant \sum_{j=1}^n 2^{p/2}|2|^{p/2}=n2^p \tag{49}$$

and the rearrangement inequality holds as desired for $1\leqslant p\leqslant 2$, with the inequality reversing for $p\geqslant 2$. As the desired extrema are reached only at $\theta=0$, then if there is equality for $p>2$, we must have $\theta_j=0$ for all $j$, and hence $UV=VU=I$. As $U$ is self-adjoint and unitary, we know that $U^{-1}=U$, and hence we conclude $V=U$. The alternative extrema are reached when $\theta_j=\frac{\pi}{2}$ for all $j$, and hence $UV+VU=0$. $\square$

We finally expand upon the ranges of Conjectures 1 and 2 as originally seen in [6], and comment on how this can lead to counterexamples.

THEOREM 13. *Let $A,B \in M_{n\times n}(\mathbb{C})$ be self-adjoint with $A \geqslant B \geqslant 0$. Then*

$$||A+B||_p^p + ||A-B||_p^p \geqslant ||\sigma(A)+\sigma(B)||_p^p + ||\sigma(A)-\sigma(B)||_p^p \tag{50}$$

*for $1 \leqslant p \leqslant 2$, with the inequality reversing for $2 \leqslant p \leqslant 3$. There is equality for $p \neq 1,2$ if and only if there is equality in the entire range $1 \leqslant p \leqslant 3$.*

*Proof.* For a positive matrix $C$ and $1 < p < 2$, for positive normalization constant $k_p$ we have

$$C^p = k_p \int_0^\infty \left( \frac{C}{t^2} - \frac{1}{t} + \frac{1}{t+C} \right) t^p \, dt. \tag{51}$$

We can therefore express the difference between sides in Equation (50) for $1 < p < 2$ by the integral representation after cancellation as

$$k_p \operatorname{Tr} \left[ \int_0^\infty \left( \frac{1}{A+B+t} + \frac{1}{A-B+t} - \frac{1}{\sigma(A)+\sigma(B)+t} - \frac{1}{\sigma(A)-\sigma(B)+t} \right) t^p \, dt \right]. \tag{52}$$

In [6] it is proven that when $A \geqslant B \geqslant 0$, this integrand is always positive semidefinite. Therefore, the integral is zero is and only if it is zero everywhere, if and only if Equation (52) is zero. This would happen independent of $p$, and hence if there is equality for some $1 < p < 2$, there must be equality for all $1 \leqslant p \leqslant 2$.

To extend the range to $2 \leqslant p \leqslant 3$, we see that

$$C^p = k_p C \int_0^\infty \left( \frac{C}{t^2} - \frac{1}{t} + \frac{1}{t+C} \right) t^p \, dt = k_p \int_0^\infty \left( \frac{C^2}{t^2} - \frac{C}{t} + \frac{C}{t+C} \right) t^p \, dt \tag{53}$$

$$= k_p \int_0^\infty \left( \frac{C^2}{t^3} - \frac{C}{t^2} + \frac{1}{t} - \frac{1}{t+C} \right) t^{p+1} \, dt. \tag{54}$$

The first three terms of the integral cancel completely between each side of Equation (50), and now as the sign of the final term is reversed, the argument for $1 < p < 2$ is reversed. □

The obvious question is whether or not it is possible to relax the requirement that $A \geqslant B \geqslant 0$, perhaps even to $A+B, A-B \geqslant 0$. The answer is: it is not.

COUNTEREXAMPLE 14. *The matrices*

$$A = \begin{bmatrix} 6 & 0 \\ 0 & 5 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{55}$$

*have the property $A+B, A-B \geqslant 0$, and*

$$||A+B||_p^p + ||A-B||_p^p \leqslant ||\sigma(A)+\sigma(B)||_p^p + ||\sigma(A)-\sigma(B)||_p^p \tag{56}$$

*for $1 \leqslant p \leqslant 2$ and $p \geqslant 3$, with the inequality reversing between $2 \leqslant p \leqslant 3$.*

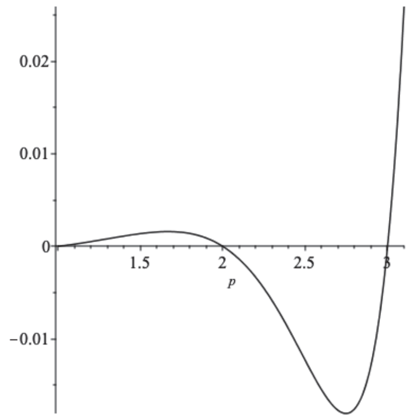


Figure 1: $||\sigma(A)+\sigma(B)||_p^p + ||\sigma(A)-\sigma(B)||_p^p - ||A+B||_p^p - ||A-B||_p^p$ *for* $1 \leqslant p \leqslant 3.1$*, demonstrating the opposite expected behavior on the intervals* $1 \leqslant p \leqslant 2$ *and* $2 \leqslant p \leqslant 3$.

A plot of Counterexample 14 can be seen in Figure 1. This counterexample hinges on the fact that we chose $B$ to be unitary, so the "up-down" rearrangement and the "aligned" rearrangements were the same. In this case, as $A$ and $B$ satisfied the requirements of Theorem 15 (the extension of Conjecture 2) but not of Theorem 13, $||\sigma(A) \pm \sigma(B)||_p^p$ was treated as the "up-down" and not the "aligned" case.

Our proof of Theorem 15 is very similar to our proof of Theorem 13 which drew heavy inspiration from the proofs in [6]. However, it diverges from [6] in a very important manner: in [6], the rearrangement inequalities in the integral representation required both $A,B \geqslant 0$. Therefore for Conjecture 2, they first proved $||A+B||_p^p + ||A-B||_p^p \leqslant ||A+|B|||_p^p + ||A-|B|||_p^p$ for $1 \leqslant p \leqslant 2$, then working with positive matrices $A$ and $|B|$ addressed the rearrangement. As monotonicity of $X^p$ was required, this does not extend as easily to $2 \leqslant p \leqslant 3$ as the proof of Theorem 13 did. We instead use majorization in the integral representation, removing the need to consider $|B|$ at all, which then allows us to extend the range without trouble:

THEOREM 15. *Let* $A,B \in M_{n\times n}(\mathbb{C})$ *be self-adjoint with* $A+B, A-B \geqslant 0$ *and* $\sigma_n(A) \geqslant \sigma_1(B)$. *Then*

$$||A+B||_p^p + ||A-B||_p^p \leqslant ||\sigma_\uparrow(A)+\sigma_\downarrow(B)||_p^p + ||\sigma_\uparrow(A)-\sigma_\downarrow(B)||_p^p \tag{57}$$

*for* $1 \leqslant p \leqslant 2$ *with the inequality reversing for* $2 \leqslant p \leqslant 3$. *There is equality for* $p \neq 1,2$ *if and only if* $A$ *and* $B$ *commute and they have simultaneous diagonalizations with diagonals* $\sigma_\uparrow(A)$ *and* $\sigma_\downarrow(B)$*, and hence there is equality in the entire range* $1 \leqslant p \leqslant 3$.

*Proof.* Once more, we use the integral representation. We can express the difference between sides in Equation (57) for $1 < p < 2$ by the integral representation after cancellation as

$$k_p \operatorname{Tr}\left[\int_0^\infty \left(\frac{1}{A+B+t} + \frac{1}{A-B+t} - \frac{1}{\sigma_\uparrow(A)+\sigma_\downarrow(B)+t} - \frac{1}{\sigma_\uparrow(A)-\sigma_\downarrow(B)+t}\right) t^p dt\right]. \tag{58}$$

We will show that the integrand is always negative. We make the substitution $H = A+t$, $K = H^{-1/2}BH^{-1/2}$, then

$$(A \pm B+t)^{-1} = H^{-1/2}(I \pm K)^{-1}H^{-1/2} = H^{-1/2}\left(\sum_{m=0}^\infty (-1)^n (\pm K)^m\right)H^{-1/2} \tag{59}$$

and

$$\frac{1}{A+B+t}+\frac{1}{A-B+t}=2H^{-1/2}\left(\sum_{m=0}^{\infty}K^{2m}\right)H^{-1/2}. \tag{60}$$

For each $m$, we notice that $K^{2m}$ is a positive matrix, and hence $H^{-1/2}K^{2m}H^{-1/2}$ is positive, and the eigenvalues and singular values are the same. Therefore,

$$\begin{aligned}
\mathrm{Tr}[H^{-1/2}K^{2m}H^{-1/2}] &= \sum_{i=1}^{n}\sigma_i(H^{-1/2}K^{2m}H^{-1/2}) && (61)\\
&= \sum_{i=1}^{n}\sigma_i(H^{-1/2}(H^{-1/2}BH^{-1/2})^{2m}H^{-1/2}) && (62)\\
&= \sum_{i=1}^{n}\sigma_i(H^{-1/2}(H^{-1/2}BH^{-1/2})^{m})^2 && (63)\\
&\leqslant \sum_{i=1}^{n}\sigma_i(H^{-1/2})^2\sigma_i((H^{-1/2}BH^{-1/2})^{m})^2 && (64)\\
&\leqslant \sum_{i=1}^{n}\sigma_i(H^{-1})\sigma_i(H^{-1/2})^{2m}\sigma_i(B)^{2m}\sigma_i(H^{-1/2})^{2m} && (65)\\
&= \sum_{i=1}^{n}\sigma_{n+1-i}(H)^{-2m-1}\sigma_i(B)^{2m}. && (66)
\end{aligned}$$

This string makes repeated use of the majorization inequalities from Theorem 7 and Lemmas 4 and 8. Furthermore, there is equality for $p \neq 1,2$ if and only if the integrand is always 0, and there is equality throughout. As we made use of log majorization $\sigma(AB) \prec_{(\log)} \sigma(A)\sigma(B)$, by Lemma 5 this must imply that $\sigma(AB)=\sigma(A)\sigma(B)$, which happens if and only if $A$ and $B$ commute with singular values aligned: by variational argument, $\sigma_1(AB)=\sigma_1(A)\sigma_1(B)$ if and only if the eigenvectors of $\sigma_1(A)$ and $\sigma_1(B)$ are the same, and then reduction of dimension repeats the argument for the other singular values. Reversing the expansion trick from Line (66) gives $\frac{1}{\sigma_\uparrow(A)\pm\sigma_\downarrow(B)+t}$ as desired, completing our proof for $1<p<2$. The same integral representation for $2 \leqslant p \leqslant 3$ as in the proof of Theorem 13 now extends the range. □

An obvious counterexample to Conjecture 2 for all ranges are any pair of unitary matrices, as shown by Thoerem 12. However, there are matrices that hold in the range $1 \leqslant p \leqslant 2$, but not in the range $2 \leqslant p \leqslant 3$, as demonstrated by Counterexample 16 and Figure 2. In fact, these matrices $C$ and $D$ also provide a counterexample to Conjecture 1, as seen in Figure 3.

COUNTEREXAMPLE 16. *The matrices*

$$C=\begin{bmatrix}6 & 0\\ 0 & -1\end{bmatrix}, \qquad D=\begin{bmatrix}-1.97035 & 1.72243\\ 1.72243 & 1.79035\end{bmatrix} \tag{67}$$

*are a counterexample for both Conjecture 1 and Conjecture 2, with contrary behavior for Conjecture 1 within the interval $1 \leqslant p \leqslant 2$; and contrary behavior for Conjecture 2 within the interval $2 \leqslant p \leqslant 3$.*

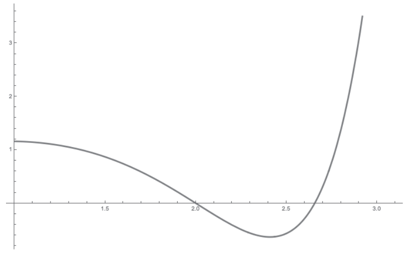


Figure 2: $||\sigma_{\uparrow}(C)+\sigma_{\downarrow}(D)||_p^p + ||\sigma_{\uparrow}(C)-\sigma_{\downarrow}(D)||_p^p - ||C+D||_p^p - ||C-D||_p^p$ *for* $1 \leqslant p \leqslant 3$*, demonstrating the expected behavior on the interval* $1 \leqslant p \leqslant 2$*, and contrary behavior within* $2 \leqslant p \leqslant 3$*.*

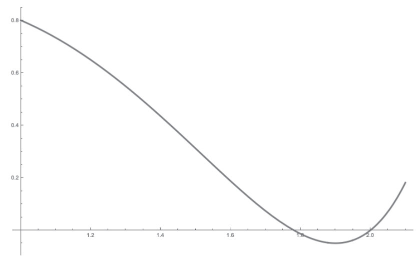


Figure 3: $||\sigma(C)+\sigma(D)||_p^p + ||\sigma(C)-\sigma(D)||_p^p - ||C+D||_p^p - ||C-D||_p^p$ *with contrary behavior within* $1 \leqslant p \leqslant 2$*.*

## REFERENCES


[1] F. J. ALMGREN JR, E. H. LIEB, *Symmetric decreasing rearrangement is sometimes continuous* Journal of the American Mathematical Society, pp. 683–773 (1989).

[2] T. ANDO, *Majorization, doubly stochastic matrices, and comparison of eigenvalues*, Linear Algebra and its Applications **118**, 163–248 (1989), `https://doi.org/10.1016/0024-3795(89)90580-6`, `http://www.sciencedirect.com/science/article/pii/0024379589905806`.

[3] K. BALL, E. A. CARLEN, E. H. LIEB, *Sharp uniform convexity and smoothness inequalities for trace norms*, Inventiones mathematicae **115** (1), 463–482 (1994), `https://doi.org/10.1007/BF01231769`.

[4] J. C. BOURIN, E. Y. LEE, *Clarkson-McCarthy inequalities with unitary and isometry orbits*, Linear Algebra and its Applications **601**, 170–179 (2020), `https://doi.org/10.1016/j.laa.2020.04.019`, `http://www.sciencedirect.com/science/article/pii/S0024379520302135`.

[5] A. BURCHARD, *Cases of equality in the riesz rearrangement inequality*, Annals of Mathematics **143** (3), 499–527 (1996). `http://www.jstor.org/stable/2118534`.

[6] E. CARLEN, E. H. LIEB, *Some matrix rearrangement inequalities*, Annali di Matematica Pura ed Applicata **185** (5), S315–S324 (2006), `https://doi.org/10.1007/s10231-004-0147-z`.

[7] K. FAN, *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proceedings of the National Academy of Sciences of the United States of America **37** (11), 760–766 (1951), `10.1073/pnas.37.11.760`, `https://www.ncbi.nlm.nih.gov/pubmed/16578416`.

[8] I. M. GEL'FAND, M. A. NAIMARK, *The relation between the unitary representations of the complex unimodular group and its unitary subgroup*, Izv. Akad. Nauk SSSR Ser. Mat. **14** (3), 239–260 (1950).

[9] G. H. HARDY, J. E. LITTLEWOOD, G. PÓLYA, *Some simple inequalities satisfied by convex functions*, Messenger Math. **58**, 145–152 (1929), `https://ci.nii.ac.jp/naid/10009422169/en/`.

[10] G. H. HARDY, G. PÓLYA, *Inequalities*, Cambridge: Cambridge University Press (1934), Bibliography: p. 300–314.

[11] F. HIAI, *Equality cases in matrix norm inequalities of golden-thompson type*, Linear and Multilinear Algebra **36** (4), 239–249 (1994), `doi:10.1080/03081089408818297`, `https://doi.org/10.1080/03081089408818297`.

[12] F. HIAI, D. PETZ, *Introduction To Matrix Analysis And Applications*, 1 edn., chap. 6, pp. 227–271, Springer International Publishing, Cham (2014).

[13] A. HORN, *On the singular values of a product of completely continuous operators*, Proceedings of the National Academy of Sciences of the United States of America **36** (7), 374 (1950).

[14] A. W. MARSHALL, I. OLKIN, B. C. ARNOLD, *Inequalities: Theory of Majorization and Its Applications*, 2 edn., Springer, New York (2011).

[15] C. MCCARTHY, *c_p cp*, Isr. J. Math. **5**, 249–271 (1967).

[16] N. TOMCZAK-JAEGERMANN, *The moduli of smoothness and convexity and the Rademacher averages of the trace classes* $S_p(1 \leqslant p \leqslant \infty)^*$, Studia Mathematica **50** (2), 163–182 (1974), http://eudml.org/doc/217886.

[17] M. TOMIĆ, *Théoreme de gauss relatif au centre de gravité et son application*, Bull. Soc. Math. Phys. Serbie **1**, 31–40 (1949).

[18] H. WEYL, *Inequalities between two types of eigenvalues of a linear transformation*, Proceedings of the National Academy of Sciences of the United States of America **35** (7), 408–411 (1949).

(Received September 21, 2020)

*Victoria M. Chayes*
*Department of Mathematics*
*Rutgers University*
*Piscataway, NJ 08854*
*e-mail:* vc362@math.rutgers.edu