

MATRIX TRIANGLE INEQUALITIES: PAST, PRESENT, AND FUTURE

JEAN-CHRISTOPHE BOURIN AND TENG ZHANG

ABSTRACT. We revisit several classical matrix inequalities related to the celebrated Thompson triangle inequality. We then present the second author's version of Thompson's inequality and related results. Finally, we survey a number of recent and elegant matrix triangle inequalities that may be viewed as cousins of Thompson's inequality. Taken together, these results illustrate the central role played by various notions of modulus in the noncommutative setting. Their interplay opens new perspectives in matrix analysis and operator theory.

CONTENTS

1. Introduction	1
2. Thompson's inequality	3
3. Symmetric norms	4
4. Majorization and unitary orbits	10
5. The operator right modulus and triangle inequalities	13
5.1. Operator version	14
5.2. Unitarily invariant norms	15
5.3. Schatten p -norms	15
6. The quadratic symmetric modulus and triangle inequalities	16
6.1. Operator version	17
6.2. Unitarily invariant norms	18
6.3. Schatten p -norms	19
7. The arithmetic symmetric modulus and triangle inequalities	21
7.1. Operator version	21
7.2. Unitarily invariant norms	22
7.3. Schatten p -norms	23
8. Some hybrid triangle inequalities and further problems	23
References	25

1. INTRODUCTION

Within the vast literature on matrix and operator inequalities, one may distinguish two special classes illustrating how matrix analysis can be viewed as a noncommutative counterpart of classical analysis. The first class transforms scalar identities into matrix inequalities, whereas the second provides matrix analogues of scalar inequalities. We briefly review a few representative examples in this introductory section.

2020 *Mathematics Subject Classification.* Primary 47A30; Secondary 15A60.

Key words and phrases. matrix inequalities, unitarily invariant norms, majorization, Thompson's triangle inequality.

The second author is supported by the China Scholarship Council, the Young Elite Scientists Sponsorship Program for PhD Students of the China Association for Science and Technology, and the Fundamental Research Funds for the Central Universities at Xi'an Jiaotong University (Grant No. xzy022024045).

Let $\mathbb{M}_n$ denote the algebra of complex $n \times n$ matrices (also regarded as operators on an n -dimensional Hilbert space), let $\mathbb{M}_n^+$ be its positive semidefinite cone, and let $\mathbb{M}_n^{\text{sa}}$ be its Hermitian part. A celebrated example of the first class is the Golden–Thompson inequality [Gol65, Tho65],

$$\text{Tr } e^{A+B} \leq \text{Tr}(e^A e^B), \quad A, B \in \mathbb{M}_n^{\text{sa}}.$$

Turning to the second class, one of the earliest examples is the trace inequality for convex functions. Let $g(t)$ be a convex function defined on an interval containing the spectra of two Hermitian matrices A and B in $\mathbb{M}_n$. Then

$$\text{Tr } g\left(\frac{A+B}{2}\right) \leq \frac{\text{Tr } g(A) + \text{Tr } g(B)}{2}. \quad (1.1)$$

With $g(t) = t \log t$, this brings the concavity of von Neumann’s entropy [Neu27], and thus we call it von Neumann’s convexity inequality, although inequalities of this type were likely known earlier. Another famous example in our second class is Thompson’s triangle inequality [Tho76]. This theorem states that if $A, B \in \mathbb{M}_n$, then there exist unitary matrices $U, V \in \mathbb{M}_n$ such that

$$|A+B| \leq U|A|U^* + V|B|V^*, \quad (1.2)$$

where $|A| = (A^*A)^{1/2}$ denotes the right modulus of A . One of Thompson’s motivations was to provide a simplified approach to the Rotfel’d trace inequality [Rot69], which states that, for any concave function $f : [0, \infty) \rightarrow [0, \infty)$,

$$\text{Tr } f(|A+B|) \leq \text{Tr } f(|A|) + \text{Tr } f(|B|). \quad (1.3)$$

Over the last 25 years, a host of elegant inequalities have been established that extend or complement these two beautiful subadditivity theorems of Rotfel’d (1.3) and Thompson (1.2).

This article, mostly of expository nature, is concerned with classical results related to Thompson’s inequality, as well as recent symmetrized versions. In Section 2, we establish Thompson’s inequality (1.2). Ky Fan k -norms and Schatten p -norms are studied in Section 3, and we show that the triangle inequality for these classical norms can be derived from Thompson’s inequality. In Section 4, we develop a rather original approach of majorization theory, thereby completing our study of symmetric norms. Section 5 examines what remains of Thompson-type estimates when the unitary conjugations on the right-hand side are removed, and explains why additional constants or averaging procedures are then unavoidable. We then turn to symmetrized versions of the modulus. Section 6 is devoted to the quadratic symmetric modulus

$$|Z|_{\text{qsym}} := \left(\frac{|Z|^2 + |Z^*|^2}{2} \right)^{1/2}.$$

A remarkable feature of this modulus is that it still satisfies a Thompson-type triangle inequality: for every $A, B \in \mathbb{M}_n$, there exist unitary matrices $U, V \in \mathbb{M}_n$ such that

$$|A+B|_{\text{qsym}} \leq U|A|_{\text{qsym}}U^* + V|B|_{\text{qsym}}V^*.$$

Section 7 studies the arithmetic symmetric modulus

$$|Z|_{\text{sym}} := \frac{|Z| + |Z^*|}{2},$$

for which the direct Thompson-type inequality fails in general, but several weaker or averaged forms remain available. Finally, Section 8 collects recent hybrid inequalities which compare these different moduli and may be viewed as further variations on Thompson’s triangle inequality. Hence the topic of matrix triangle inequalities is, 50 years after Thompson, a very active field of research.

2. THOMPSON'S INEQUALITY

We begin with Thompson's inequality itself. The proof is based on the Fan–Hoffman lemma, and the same method also gives a useful Cartesian version.

If $A \in \mathbb{M}_n$ is Hermitian, we denote by $\lambda_1^\downarrow(A) \geq \dots \geq \lambda_n^\downarrow(A)$ its eigenvalues arranged in nonincreasing order, and by $A^\downarrow$ the diagonal matrix in $\mathbb{M}_n$ having these eigenvalues on the diagonal.

We assume the reader is familiar with the min-max principle. This principle shows that, for Hermitian matrices $A, B \in \mathbb{M}_n$, the eigenvalue order is equivalent to the unitary orbit order: $\lambda_j^\downarrow(A) \leq \lambda_j^\downarrow(B)$ for all $j = 1, \dots, n$, if and only if $A \leq VBV^*$ for some unitary V ; see [Bha97, p. 74, Exercise III.5.5]. If $A = VBV^*$, we write $A \simeq B$, and say that A and B are unitarily equivalent.

For $Z \in \mathbb{M}_n$, we denote its real, or Hermitian part by

$$\operatorname{Re} Z := \frac{Z + Z^*}{2}.$$

Similarly, its imaginary part is also the Hermitian matrix

$$\operatorname{Im} Z := \frac{Z - Z^*}{2i} = \operatorname{Re}(-iZ).$$

Thus $Z = \operatorname{Re} Z + i\operatorname{Im} Z$ which is called the decomposition in real and imaginary parts, or the Toeplitz decomposition, or the Cartesian decomposition. The Fan-Hoffman lemma [FH55] states:

Lemma 2.1. *Let $Z \in \mathbb{M}_n$. Then, there exists a unitary matrix $V \in \mathbb{M}_n$ such that*

$$\operatorname{Re} Z \leq V|Z|V^*.$$

Proof. Denote by Λ_j the set of all j -dimensional subspaces of $\mathbb{C}^n$, and let S^1 be the unit sphere in the subspace S . Then,

$$\lambda_j^\downarrow(\operatorname{Re} Z) = \max_{S \in \Lambda_j} \min_{h \in S^1} \langle h, \operatorname{Re} Zh \rangle = \max_{S \in \Lambda_j} \min_{h \in S^1} \operatorname{Re} \langle h, Zh \rangle \leq \max_{S \in \Lambda_j} \min_{h \in S^1} \|Zh\|.$$

We might stop our proof there, or detail a little bit more, hence establishing the min-max principle for singular values from the min-max principle for the eigenvalues:

$$\max_{S \in \Lambda_j} \min_{h \in S^1} \|Zh\| = \max_{S \in \Lambda_j} \min_{h \in S^1} \langle h, |Z|^2 h \rangle^{1/2} = (\lambda_j^\downarrow(|Z|^2))^{1/2} = \lambda_j^\downarrow(|Z|).$$

Combining the inequality in the first part of the proof with the equality of the second part completes the proof. $\square$

Using this lemma, Thompson [Tho76] derived in a simple but clever way his famous inequality:

Theorem 2.2. *Let $A, B \in \mathbb{M}_n$. Then there exist unitary matrices $U, V \in \mathbb{M}_n$ such that*

$$|A + B| \leq U|A|U^* + V|B|V^*.$$

Proof. Let $A + B = Q|A + B|$ be a polar decomposition, and extend the polar partial isometry Q to a unitary matrix. With $W = Q^*$, we have $|A + B| = W(A + B)$. Since the left-hand side is Hermitian,

$$|A + B| = \operatorname{Re}(WA) + \operatorname{Re}(WB).$$

Applying the Fan–Hoffman lemma to WA and WB , and using $|WA| = |A|$ and $|WB| = |B|$, proves the result. $\square$

Taking $A = \operatorname{Re} Z$ and $B = i\operatorname{Im} Z$ in Theorem 2.2 gives the following Cartesian consequence.

Corollary 2.3. *Let $Z \in \mathbb{M}_n$. Then, there exist unitary matrices $U, V \in \mathbb{M}_n$ such that*

$$|Z| \leq U|\operatorname{Re} Z|U^* + V|\operatorname{Im} Z|V^*.$$

A subadditivity inequality of Aujla and Bourin [AB07], for positive matrices $A, B \in \mathbb{M}_n^+$ and nonnegative concave functions $f(t)$ on the positive half-line, states that

$$f(A + B) \leq Uf(A)U^* + Vf(B)V^*. \quad (2.1)$$

Such a function $f(t)$ is necessarily nondecreasing, so using Thompson's inequality we have a refinement of Rotfel'd inequality:

Theorem 2.4. *Let $A, B \in \mathbb{M}_n$ and let $f : [0, \infty) \rightarrow [0, \infty)$ be concave. Then, there exist unitary matrices U, V such that*

$$f(|A + B|) \leq Uf(|A|)U^* + Vf(|B|)V^*.$$

Proof. By Thompson's inequality, for some unitaries U_0, V_0 ,

$$|A + B| \leq P := U_0|A|U_0^* + V_0|B|V_0^*.$$

Since f is nonnegative and concave on $[0, \infty)$, it is nondecreasing. Hence the eigenvalue order gives $\lambda_j(f(|A + B|)) \leq \lambda_j(f(P))$ for all j . By the unitary orbit order, this means that $f(|A + B|) \leq Wf(P)W^*$ for some unitary W . Applying the Aujla–Bourin inequality to P gives

$$f(P) \leq U_1f(U_0|A|U_0^*)U_1^* + V_1f(V_0|B|V_0^*)V_1^*.$$

Absorbing the unitaries into U and V completes the proof. $\square$

Applying Theorem 2.4 to the Cartesian decomposition $Z = \operatorname{Re} Z + i\operatorname{Im} Z$ yields the next corollary.

Corollary 2.5. *Let $Z \in \mathbb{M}_n$ and let $f : [0, \infty) \rightarrow [0, \infty)$ be concave. Then, there exist unitary matrices $U, V \in \mathbb{M}_n$ such that*

$$f(|Z|) \leq Uf(|\operatorname{Re} Z|)U^* + Vf(|\operatorname{Im} Z|)V^*.$$

Note that we always have $\|Z\|_\infty \leq \| |\operatorname{Re} Z| + |\operatorname{Im} Z| \|_\infty$ and so, since $f(t)$ is necessarily nondecreasing,

$$\|f(|Z|)\|_\infty \leq \|f(|\operatorname{Re} Z| + |\operatorname{Im} Z|)\|_\infty.$$

However, for every unitarily invariant norm, or symmetric norm (to be defined in the next subsection), that is not a scalar multiple of $\|\cdot\|_\infty$, the reverse strict inequality can occur:

$$\|f(|Z|)\| > \|f(|\operatorname{Re} Z| + |\operatorname{Im} Z|)\|$$

for some $Z \in \mathbb{M}_2$ and some concave function f , see [Bou10]. Thus Corollary 2.5 is a subtle statement, and Thompson's inequality highlights a genuinely noncommutative phenomenon.

3. SYMMETRIC NORMS

The polar decomposition invites us to consider norms on $\mathbb{M}_n$ such that for all $Z \in \mathbb{M}_n$,

$$\|Z\| = \||Z|\| = \|Z^*\|.$$

These norms are characterized by an invariance property, for all unitaries $U, V \in \mathbb{M}_n$,

$$\|Z\| = \|UZV\|.$$

Hence, with the singular value decomposition $Z = U|Z|^\downarrow V^*$, we have $\|Z\| = \||Z|^\downarrow\|$, where $|Z|^\downarrow := \operatorname{diag}(\lambda_1^\downarrow(|Z|), \dots, \lambda_n^\downarrow(|Z|))$. Such a norm is called a symmetric norm on $\mathbb{M}_n$, or a unitarily invariant one. Furthermore, if $A, B \in \mathbb{M}_n$ are two contractions, then

$$\|AZB\| \leq \|Z\|.$$

Indeed, it suffices to consider the case $A, B \in \mathbb{M}_n^+$ and to note that a positive contraction is a convex combination of Hermitian unitary matrices. In particular, for two positive matrices in $\mathbb{M}_n$ such that $0 \leq X \leq Y$, we have $X = Y^{1/2} K^2 Y^{1/2} \simeq K Y K$ for some positive contraction K , so symmetric norms respect the order relation in $\mathbb{M}_n^+$.

The three simplest examples of symmetric norms on $\mathbb{M}_n$ are the operator norm $\|\cdot\|_\infty$, the Hilbert-Schmidt (or Frobenius) norm $\|\cdot\|_2$ and the trace norm $\|\cdot\|_1$. For $A \in \mathbb{M}_n$, its operator norm $\|A\|_\infty = \max\{\|Ax\| : \|x\| = 1\}$, where the symbol $\|\cdot\|$ also stands for the standard norm on $\mathbb{C}^n$.

Definition 3.1. The Hilbert-Schmidt norm, or Frobenius norm $\|\cdot\|_2$ on $\mathbb{M}_n$ is the norm associated to the canonical scalar product $\langle X, Y \rangle = \text{Tr } X^* Y$. Thus, for all $A \in \mathbb{M}_n$,

$$\|A\|_2 := \{\text{Tr}|A|^2\}^{1/2} = \left\{ \sum_{j=1}^n \lambda_j^\downarrow(|A|)^2 \right\}^{1/2}.$$

The trace norm $\|\cdot\|_1$ on $\mathbb{M}_n$ is defined by

$$\|A\|_1 := \text{Tr}|A| = \sum_{j=1}^n \lambda_j^\downarrow(|A|).$$

While the triangle inequality is obvious for the operator norm and the Hilbert-Schmidt norm, the case of the trace norm deserves a small proof. With Thompson's inequality at our disposal, the triangle inequality for $\|\cdot\|_1$ becomes obvious. Given $X, Y \in \mathbb{M}_n$, there exist two unitary matrices $U, V \in \mathbb{M}_n$ such that:

$$\|X + Y\|_1 = \text{Tr}|X + Y| \leq \text{Tr } U|X|U^* + \text{Tr } V|Y|V^* = \|X\|_1 + \|Y\|_1.$$

Remark 3.2. The Cauchy-Schwarz inequality for $\|\cdot\|_2$ entails that $\|\cdot\|_1$ is a norm. Indeed, given a pair of matrices $X, Y \in \mathbb{M}_n$ we have a unitary U such that

$$\|X + Y\|_1 = \text{Re } \text{Tr } U(X + Y) \leq |\text{Tr } UX| + |\text{Tr } UY| \leq \|X\|_1 + \|Y\|_1.$$

Here the last inequality follows from $|\text{Tr } Z| \leq \text{Tr } |Z|$ for all $Z \in \mathbb{M}_n$. This is a consequence of the Cauchy-Schwarz inequality for the decomposition $Z = (V|Z|^{1/2})|Z|^{1/2}$.

Another important class of symmetric norms is given in the following definition. These norms form a natural scale between the operator norm and the trace norm.

Definition 3.3. The Ky Fan k -norms on $\mathbb{M}_n$, $k = 1, \dots, n$ are the symmetric norms defined by

$$\|X\|_{(k)} := \max_{E \in \Pi_k} \text{Tr } E|X|E = \sum_{j=1}^k \lambda_j^\downarrow(|X|)$$

where Π_k denotes the collection of all rank k projections in $\mathbb{M}_n$.

Hence $\|\cdot\|_1 = \|\cdot\|_{(n)}$ and $\|\cdot\|_\infty = \|\cdot\|_{(1)}$, where $\|\cdot\|_\infty$ stands for the operator norm. Let us show that Definition 3.3 is consistent:

Proof. The second equality follows from the min-max principle, for any k -dimensional subspace $\mathcal{S} \subset \mathbb{C}^n$ (here we take $\mathcal{S}$ as the range of E) we have $\lambda_j^\downarrow(|X|_{\mathcal{S}}) \leq \lambda_j^\downarrow(|X|)$. It is clear that $\|X\|_{(k)}$ is a homogeneous unitarily invariant functional. To prove that this is a symmetric norm, we use Thompson's inequality to obtain the triangle inequality for $\|\cdot\|_{(k)}$ as follows: For some rank k projection E ,

$$\|X + Y\|_{(k)} = \text{Tr } E|X + Y|E,$$

and Thompson's inequality yields two unitary matrices U and V such that

$$\begin{aligned}\|X + Y\|_{(k)} &\leq \operatorname{Tr} EU|X|U^*E + \operatorname{Tr} EV|Y|V^*E \\ &= \operatorname{Tr} U^*EU|X|U^*EU + \operatorname{Tr} V^*EV|Y|V^*EV \\ &\leq \|X\|_{(k)} + \|Y\|_{(k)}.\end{aligned}$$

So the Ky Fan functionals are symmetric norms. $\square$

The following two propositions are other ways to look at Ky Fan k -norms. These propositions also show that the Ky Fan k -norms are indeed norms.

Proposition 3.4. *The Ky Fan k -norms on $\mathbb{M}_n$, $k = 1, \dots, n$ satisfy*

$$\|X\|_{(k)} = \max_{A, B \in \Gamma_k} |\operatorname{Tr} AXB|$$

where Γ_k denotes all rank $\leq k$ contractions in $\mathbb{M}_n$. Equivalently,

$$\|X\|_{(k)} = \max_{C \in \Gamma_k} |\operatorname{Tr} CX|.$$

Proof. A form of the polar decomposition of X shows that, thanks to Cauchy-Schwarz,

$$\begin{aligned}|\operatorname{Tr} AXB|^2 &= \left| \operatorname{Tr} A|X^*|^{1/2}U|X|^{1/2}B \right|^2 \\ &\leq \operatorname{Tr} |X^*|^{1/2}A^*A|X^*|^{1/2} \operatorname{Tr} |X|^{1/2}BB^*|X|^{1/2} \\ &\leq \operatorname{Tr} |X^*|^{1/2}E|X^*|^{1/2} \operatorname{Tr} |X|^{1/2}F|X|^{1/2},\end{aligned}$$

where E is the support projection of A^*A and F is the support projection of BB^* . From the min-max principle it follows that, for $j \leq k \leq n$ and $Z \in \mathbb{M}_n$, $\lambda_j^\downarrow(E|Z|E) \leq \lambda_j^\downarrow(|Z|)$ since $E|Z|E$ corresponds to a compression of $|Z|$ on a k -dimensional subspace. Hence

$$|\operatorname{Tr} AXB| \leq \sum_{j=1}^k \lambda_j^\downarrow(|X|).$$

If $B = P$ is a projection corresponding to the k largest eigenvalues of $|X|$, and $A = PU^*$, we have equality in the first maximum stated in the proposition,

$$|\operatorname{Tr} AXB| = \operatorname{Tr} AXB = \operatorname{Tr} PU^*XP = \sum_{j=1}^k \lambda_j^\downarrow(|X|).$$

So equality holds in the second maximum with $C = PU^*$. Thus Proposition 3.4 is equivalent to the initial Definition 3.3 of the k -norms. $\square$

The proof shows that it suffices to consider partial isometries in Γ_k . So we can state the next proposition, where I_k denotes the identity of $\mathbb{M}_k$ and Tr_k the trace on $\mathbb{M}_k$.

Proposition 3.5. *The Ky Fan k -norms on $\mathbb{M}_n$, $k = 1, \dots, n$ satisfy*

$$\|X\|_{(k)} := \max_{V, W \in \Theta(k, n)} |\operatorname{Tr}_k V^*XW|.$$

where $\Theta(k, n)$ stands for the set of $n \times k$ isometry matrices J , i.e., $J^*J = I_k$.

Proof. Note that with $(\operatorname{Tr} = \operatorname{Tr}_n)$ we have

$$\operatorname{Tr}_k V^*XW = \operatorname{Tr} V_0^*XW_0$$

where $W_0, V_0 \in \mathbb{M}_n$ are obtained from $W, V \in \mathbb{M}_{n, k}$ by adding $n - k$ zero columns. Hence

$$|\operatorname{Tr}_k V^*XW| \leq \|X\|_{(k)}.$$

On the other hand, observe that

$$\mathrm{Tr}_k V^* X W = \mathrm{Tr}_n X W V^* (= \mathrm{Tr} X W V^*)$$

and choose $W \in \mathbb{M}_{n,k}$ any isometry with range projection $P = W W^*$, a spectral projection of $|X|$ corresponding to the k largest eigenvalues. Then pick $V = U W$, where U still holds for the unitary factor of X . We then have

$$\mathrm{Tr}_k V^* X W = \mathrm{Tr} X W W^* U^* = \mathrm{Tr} X P U^* = \sum_{j=1}^k \lambda_j^\downarrow(|X|)$$

and the proof is complete. $\square$

Since $\|\cdot\|_2$ is a very nice norm, we may expect special results for it. Recall that $\mathbb{M}_n^{\mathrm{sa}}$ is the Hermitian part of $\mathbb{M}_n$ and $\mathbb{M}_n^+$ is the positive semidefinite cone.

Definition 3.6. A pair (A, B) in $\mathbb{M}_n^{\mathrm{sa}}$ is a *monotone pair* if $A = f(C)$ and $B = g(C)$ for two non-decreasing functions $f(t)$, $g(t)$ and some $C \in \mathbb{M}_n^{\mathrm{sa}}$. If $f(t)$ is non-decreasing and $g(t)$ is non-increasing, then (A, B) is *antimonotone*.

This notion leads to the following Hilbert–Schmidt norm comparison.

Theorem 3.7. *Let $Z \in \mathbb{M}_n$ be a normal matrix. If (A, B) is a monotone pair in $\mathbb{M}_n^+$ then,*

$$\|AZB\|_2 \leq \|ZAB\|_2.$$

If (A, B) is antimonotone, then the inequality reverses.

Since (A, B) is a monotone (antimonotone) pair in $\mathbb{M}_n^+$ if and only if so is $(A^{1/2}, B^{1/2})$, squaring both sides, the theorem is equivalent to the trace inequalities in the next corollary.

Corollary 3.8. *Let $Z \in \mathbb{M}_n$ be a normal matrix. If (A, B) is a monotone pair in $\mathbb{M}_n^{\mathrm{sa}}$ then,*

$$\mathrm{Tr} Z^* A Z B \leq \mathrm{Tr} Z^* Z A B.$$

If (A, B) is antimonotone, then the inequality reverses.

Proof of Corollary 3.8. Replacing A, B by $A + kI, B + kI$ for k large enough, we may assume $A, B \in \mathbb{M}_n^+$. It suffices to prove the monotone case, the antimonotone case follows by replacing B with $cI - B$ for $c > 0$ large enough. By layer decomposition of C , hence of A and B , it suffices to show that

$$\mathrm{Tr} Z^* E Z F \leq \mathrm{Tr} Z^* Z E F$$

for every pair of projections E and F such that $E \leq F$ or $F \leq E$.

Suppose $E \leq F$. Then, since $Z^* Z = Z Z^*$,

$$\mathrm{Tr} Z^* E Z F \leq \mathrm{Tr} Z^* E Z = \mathrm{Tr} Z Z^* E = \mathrm{Tr} Z^* Z E F.$$

Suppose $F \leq E$. Then, $\mathrm{Tr} Z^* E Z F \leq \mathrm{Tr} Z^* Z F = \mathrm{Tr} Z^* Z E F$. $\square$

Letting $Z = |h\rangle\langle h|$ be a rank one projection, Corollary 3.8 says that, for a monotone pair (A, B) in $\mathbb{M}_n$,

$$\langle h, Ah \rangle \langle h, Bh \rangle \leq \langle h, ABh \rangle.$$

Computing the scalar products in a basis diagonalizing A, B we see that we obtain Chebyshev's inequality for pairs of non-decreasing continuous functions on $(0, 1)$,

$$\int_0^1 f(t) dt \int_0^1 g(t) dt \leq \int_0^1 f(t)g(t) dt.$$

Thus Theorem 3.7 is a matrix Chebyshev inequality. Another interesting special case appears by letting Z be a unitary matrix: We may then obtain von Neumann trace inequalities [Neu37]:

Lemma 3.9. *Let $X, Y \in \mathbb{M}_n$. Then*

$$|\mathrm{Tr} XY| \leq \sum_{j=1}^n \lambda_j^\downarrow(|X|) \lambda_j^\downarrow(|Y|)$$

Proof. First if X and Y are both positive then, for some unitary Z , ZXZ^* and Y form a monotone pair, and Corollary 3.8 yields

$$\mathrm{Tr} XY = \mathrm{Tr} (ZXZ^*)ZY Z^* \leq \mathrm{Tr} ZXZ^*Y = \sum_{j=1}^n \lambda_j^\downarrow(|X|) \lambda_j^\downarrow(|Y|).$$

For general X and Y , we consider polar decompositions $X = U|X|$, $Y = V|Y|$ and we note that

$$\begin{aligned} |\mathrm{Tr} XY| &= \left| \mathrm{Tr} |X^*|^{1/2} U |X|^{1/2} V |Y| \right| \\ &= \left| \mathrm{Tr} |Y|^{1/2} |X^*|^{1/2} U |X|^{1/2} V |Y|^{1/2} \right| \\ &\leq \{\mathrm{Tr} |Y| |X^*|\}^{1/2} \{\mathrm{Tr} |X| V |Y| V^*\}^{1/2} \end{aligned}$$

by using the Cauchy-Schwarz inequality $|\mathrm{Tr} A^* B| \leq \{\mathrm{Tr} A^* A\}^{1/2} \{\mathrm{Tr} B^* B\}^{1/2}$. Hence the general case follows from the positive one. $\square$

The Hermitian version follows from the same rearrangement argument.

Lemma 3.10. *Let $A, B \in \mathbb{M}_n^{\mathrm{sa}}$. Then*

$$\mathrm{Tr} AB \leq \sum_{j=1}^n \lambda_j^\downarrow(A) \lambda_j^\downarrow(B).$$

Proof. Exactly the same proof as the previous one in the case of $X, Y \in \mathbb{M}_n^+$. $\square$

The operator norm is denoted by $\|\cdot\|_\infty$ as it is the limit of Schatten p -norms, $p \rightarrow \infty$, given in the next definition.

Definition 3.11. The Schatten p -norms on $\mathbb{M}_n$, $1 \leq p < \infty$, are the symmetric norms defined by

$$\|A\|_p := \{\mathrm{Tr} |A|^p\}^{1/p} = \left\{ \sum_{j=1}^n \lambda_j^\downarrow(|A|)^p \right\}^{1/p}.$$

We can prove that the Schatten p -norms are indeed symmetric norms in two different ways. First we may note the matrix Hölder inequality:

Proposition 3.12. *Let $X, Y \in \mathbb{M}_n$ and let $1 \leq p, q \leq \infty$ be conjugate, i.e., $p^{-1} + q^{-1} = 1$. Then*

$$|\mathrm{Tr} X^* Y| \leq \|X\|_p \|Y\|_q$$

The matrix Hölder inequality follows from the classical one by using Lemma 3.9. As in the classical case we then infer that $\|\cdot\|_p$ and $\|\cdot\|_q$ are dual norms. Thus we have the matrix Minkowski inequality:

Proposition 3.13. *Let $A, B \in \mathbb{M}_n$ and $1 \leq p \leq \infty$. Then*

$$\|A + B\|_p \leq \|A\|_p + \|B\|_p.$$

We can give a proof of this triangle inequality by using Thompson's inequality. It is our second proof that the Schatten p -norms are indeed norms.

Proof. The case $p = \infty$ is well-known, so let $p < \infty$. Suppose first that $\|A\|_p = \|B\|_p = 1$ and let $0 \leq t \leq 1$. Then, there exist two unitaries U, V such that

$$\|tA + (1-t)B\|_p^p = \text{Tr}|tA + (1-t)B|^p \leq \text{Tr}(tU|A|U^* + (1-t)V|B|V^*)^p.$$

so that, using the convexity inequality (1.1),

$$\|tA + (1-t)B\|_p^p \leq \text{Tr}(tU|A|^pU^* + (1-t)V|B|^pV^*) = 1$$

and so $\|tA + (1-t)B\|_p \leq 1$. Now for the general case, $A, B \neq 0$, apply the previous case with

$$\frac{1}{\|A\|_p}A, \quad \frac{1}{\|B\|_p}B, \quad t = \frac{\|A\|_p}{\|A\|_p + \|B\|_p}.$$

This yields $\|A + B\|_p \leq \|A\|_p + \|B\|_p$. $\square$

Remark 3.14. A third proof of Proposition 3.13 is possible by using only the convexity inequality (1.1). To do so, consider the convex function $t \mapsto |t|^p$ and the Hermitian operators

$$A_0 = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}, \quad B_0 = \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}.$$

Arguing as in the second part of the previous proof with A_0, B_0 and $t \mapsto |t|^p$ yields Proposition 3.13.

Proposition 3.13 is a special case of the next proposition (with the trace norm).

Proposition 3.15. *Let $\|\cdot\|$ be a symmetric norm on $\mathbb{M}_n$ and let $p > 1$. Then,*

$$X \mapsto \| |X|^p \|^{1/p}$$

is a symmetric norm on $\mathbb{M}_n$.

Proof. The proof is similar to the previous one. Alternatively, let Φ be the symmetric gauge function associated with $\|\cdot\|$. Thus $\|X\| = \Phi(s(X))$, where $s(X)$ is the vector of singular values of X in nonincreasing order. Thompson's inequality, or equivalently the Ky Fan inequalities, gives

$$s(A+B) \prec_w s(A) + s(B).$$

For nonnegative vectors x, y , define $\Psi(x) = \Phi(x^p)^{1/p}$. We first note that Ψ is a norm. Indeed, using the dual gauge function Φ^* and the scalar Minkowski inequality,

$$\begin{aligned} \Phi((x+y)^p)^{1/p} &= \sup_{\Phi^*(z) \leq 1, z \geq 0} \left(\sum_i z_i (x_i + y_i)^p \right)^{1/p} \\ &\leq \Phi(x^p)^{1/p} + \Phi(y^p)^{1/p}. \end{aligned}$$

Moreover, if $0 \leq x \prec_w y$, then $x^p \prec_w y^p$ for $p > 1$, and hence $\Phi(x^p) \leq \Phi(y^p)$. Therefore,

$$\| |A+B|^p \|^{1/p} = \Psi(s(A+B)) \leq \Psi(s(A) + s(B)) \leq \Psi(s(A)) + \Psi(s(B)),$$

which is the required triangle inequality. The other norm properties and the unitary invariance are immediate. $\square$

If $A \in \mathbb{M}_{n_1+n_2}$ is partitioned as

$$A = \begin{bmatrix} A_1 & X \\ Y & A_2 \end{bmatrix}$$

then

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = \frac{1}{2} \left\{ \begin{bmatrix} A_1 & X \\ Y & A_2 \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} A_1 & X \\ Y & A_2 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \right\}.$$

Hence, arguing by induction or recursively, we have the basic fact:

Proposition 3.16. *If $\Delta(A)$ is the diagonal part of a matrix $A \in \mathbb{M}_n$, then, for all symmetric norms,*

$$\|\Delta(A)\| \leq \|A\|.$$

We close this section with a historical comment and a related conjecture. Monotone pairs were studied by the first author; see [Bou06] for von Neumann trace rearrangement inequalities and further references.

Remark 3.17. If (A, B) is a monotone pair in $\mathbb{M}_n^+$, it is natural to conjecture [Bou99] that, for every symmetric norm,

$$\|AZB\| \leq \sqrt{2}\|ZAB\|.$$

The remaining material of this section is rather classical, but our proofs are original. In particular, the proof of the Minkowski inequality in Proposition 3.13, mentioned in Remark 3.14, is a new and simple application of the trace Jensen inequality (1.1).

4. MAJORIZATION AND UNITARY ORBITS

We now recall the majorization language needed to pass from operator inequalities to all symmetric norms. The emphasis is on convex hulls and unitary orbits.

For a pair of Hermitian matrices $A, B \in \mathbb{M}_n$, the weak majorization, or submajorization, relation $A \prec_w B$ means that

$$\sum_{j=1}^k \lambda_j^\downarrow(A) \leq \sum_{j=1}^k \lambda_j^\downarrow(B) \quad (4.1)$$

for all $k = 1, \dots, n$. If, furthermore, equality holds for $k = n$, that is, if A and B have the same trace, then we say that A is majorized by B , and we write

$$A \prec B. \quad (4.2)$$

The notation in (4.2) will be used throughout this section. For each k , the sum of the k largest eigenvalues occurring in (4.1),

$$A \mapsto \sum_{j=1}^k \lambda_j^\downarrow(A), \quad (4.3)$$

is a subadditive functional on $\mathbb{M}_n^{\text{sa}}$. This follows from the Ky Fan variational principle

$$\sum_{j=1}^k \lambda_j^\downarrow(A) = \max_{\substack{P=P^*=P^2 \\ \text{Tr } P=k}} \text{Tr}(PA). \quad (4.4)$$

On the positive cone $\mathbb{M}_n^+$, the subadditivity of (4.3) is exactly the subadditivity of the Ky Fan k -norms.

Majorization and weak majorization relations lie at the heart of matrix analysis; see, for example, Ando's survey [And89]. A simple example is Schur's theorem, a direct consequence of (4.4):

Proposition 4.1. *If $\Delta(A)$ is the diagonal part of a Hermitian matrix $A \in \mathbb{M}_n$, then*

$$\Delta(A) \prec A.$$

Despite its simplicity, this result has been extremely influential in the development of matrix analysis. Consideration of Hermitian or positive two-by-two matrices suggests that majorization and weak majorization are related to convexity properties. To state these convexity results, we need a few notations.

Notations. If $R \in \mathbb{M}_n$ is a diagonal matrix, $R = \text{diag}(r_1, \dots, r_n)$ and $\sigma \in \mathfrak{S}_n$ is a permutation of $\{1, \dots, n\}$, we define the eigenvalue vectors:

$$r = (r_1, \dots, r_n), \quad r_\sigma := (r_{\sigma(1)}, \dots, r_{\sigma(n)})$$

In particular, If $A \in \mathbb{M}_n$ is Hermitian, we define the eigenvalue vectors:

$$a^\downarrow = (a_1^\downarrow, \dots, a_n^\downarrow) := (\lambda_1^\downarrow(A), \dots, \lambda_n^\downarrow(A))$$

and

$$a_\sigma^\downarrow = (a_{\sigma(1)}^\downarrow, \dots, a_{\sigma(n)}^\downarrow) := (\lambda_{\sigma(1)}^\downarrow(A), \dots, \lambda_{\sigma(n)}^\downarrow(A)).$$

Further we denote by $\mathcal{E}_n$ the set of vectors of $\mathbb{R}^n$ with 0-1 entries, thus $\varepsilon \in \mathcal{E}_n$ if and only if $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ where $\varepsilon_i \in \{0, 1\}$. For such a vector ε and an eigenvalue vector c_σ associated to a diagonal matrix $C \in \mathbb{M}_n$ and a permutation σ , we define the product

$$\varepsilon c_\sigma := (\varepsilon_1 c_{\sigma(1)}, \dots, \varepsilon_n c_{\sigma(n)}).$$

We may then state a fundamental proposition: majorization is equivalent to the property that an eigenvalue vector belongs to the convex hull of the permutations of another eigenvalue vector.

Proposition 4.2. *Let A, B be two Hermitian matrices in $\mathbb{M}_n$. Then,*

$$A \prec B \iff a^\downarrow \in \text{conv}\{b_\sigma^\downarrow : \sigma \in \mathfrak{S}_n\}.$$

Proof. Assume first that $A \prec B$. If $a^\downarrow \notin \text{conv}\{b_\sigma^\downarrow : \sigma \in \mathfrak{S}_n\}$, then we would have a vector $x = (x_1, \dots, x_n)$ such that, for any permutation σ ,

$$\sum_{i=1}^n x_i a_i^\downarrow > \sum_{i=1}^n x_i b_{\sigma(i)}^\downarrow.$$

and so, with a suitable choice of σ ,

$$\sum_{i=1}^n x_i a_i^\downarrow > \sum_{i=1}^n x_i^\downarrow b_i^\downarrow$$

where $x^\downarrow$ denotes the vector obtained from x by nonincreasing rearrangement of its components. By the commutative case of von Neumann trace inequality (Lemma 3.10), this would entail:

$$\sum_{i=1}^n x_i^\downarrow a_i^\downarrow > \sum_{i=1}^n x_i^\downarrow b_i^\downarrow. \tag{4.5}$$

Since $\sum_{i=1}^n a_i^\downarrow = \sum_{i=1}^n b_i^\downarrow$, we would also have

$$\sum_{i=1}^n y_i^\downarrow a_i^\downarrow > \sum_{i=1}^n y_i^\downarrow b_i^\downarrow.$$

where $y_i = x_i + c$, the constant $c > 0$ being large enough to ensure $y_n^\downarrow > 0$. Hence we may assume that $x_n^\downarrow > 0$. Observe that (4.5) implies

$$\begin{aligned} x_n^\downarrow \left\{ \sum_{i=1}^n (a_i^\downarrow - b_i^\downarrow) \right\} + (x_{n-1}^\downarrow - x_n^\downarrow) \left\{ \sum_{i=1}^{n-1} (a_i^\downarrow - b_i^\downarrow) \right\} + \dots \\ + \dots + (x_1^\downarrow - x_2^\downarrow) \{a_1^\downarrow - b_1^\downarrow\} > 0. \end{aligned}$$

As $x_n^\downarrow > 0$, this contradicts $A \prec B$. Therefore $a^\downarrow \in \text{conv}\{b_\sigma^\downarrow : \sigma \in \mathfrak{S}_n\}$.

Conversely, suppose that

$$a^\downarrow = \sum_{r=1}^N t_r b_{\sigma_r}^\downarrow, \quad t_r \geq 0, \quad \sum_{r=1}^N t_r = 1.$$

For each k , the sum of the k largest coordinates is a convex function. Hence

$$\sum_{j=1}^k a_j^\downarrow \leq \sum_{r=1}^N t_r \sum_{j=1}^k b_j^\downarrow = \sum_{j=1}^k b_j^\downarrow.$$

For $k = n$, equality holds. Thus $A \prec B$. $\square$

The convex-hull characterization immediately yields a unitary-orbit formulation of weak majorization.

Corollary 4.3. *Let A, B be two Hermitian matrices in $\mathbb{M}_n$. If $A \prec_w B$ then there exist unitary matrices $\{U_i\}_{i=1}^{n+1}$ such that*

$$A \leq \sum_{i=1}^{n+1} t_i U_i B U_i^*$$

for a convex combination with weights $0 \leq t_i \leq 1$, $\sum_{i=1}^{n+1} t_i = 1$.

Proof. Consider the spectral decomposition $B = \sum_{k=1}^n \lambda_k^\downarrow(B) E_k$, and define

$$B' := \left(\sum_{k=1}^{n-1} \lambda_k^\downarrow(B) E_k \right) + \mu E_n$$

where $\mu := \lambda_n^\downarrow(B) - \text{Tr } B + \text{Tr } A$ is chosen so that

$$A \prec B'.$$

The weak-majorization inequality for $k = n$ gives $\mu \leq \lambda_n^\downarrow(B)$, and hence $B' \leq B$. By Proposition 4.2 and Carathéodory's theorem in $\mathbb{R}^n$, A can be written as a convex combination of at most $n + 1$ elements in the unitary orbit of B' :

$$A = \sum_{i=1}^{n+1} t_i U_i B' U_i^*,$$

where some coefficients may be zero. This completes the proof. $\square$

Taking diagonal parts in the preceding orbit inequality gives a useful refinement.

Corollary 4.4. *Let A, B be two Hermitian matrices in $\mathbb{M}_n$. If $A \prec_w B$ then there exists a diagonal matrix $\Delta \in \mathbb{M}_n^+$ such that*

$$A^\downarrow + \Delta \prec B^\downarrow$$

Proof. By the inequality in the previous corollary,

$$A^\downarrow \leq V \left\{ \sum_{i=1}^{n+1} t_i U_i B U_i^* \right\} V^*$$

for some unitary V . Letting

$$C = V \left\{ \sum_{i=1}^{n+1} t_i U_i B U_i^* \right\} V^* - A^\downarrow$$

we have $C \geq 0$. The matrix $A^\downarrow + C$ is a convex combination of unitary conjugates of B , and so

$$A^\downarrow + C \prec B^\downarrow.$$

By Schur's theorem,

$$A^\downarrow + \Delta(C) = \Delta(A^\downarrow + C) \prec A^\downarrow + C.$$

Therefore, with $\Delta := \Delta(C)$ and by transitivity, we have $A^\downarrow + \Delta \prec B^\downarrow$. $\square$

Combining this diagonal description with the monotonicity of symmetric norms gives Ky Fan's dominance principle.

Corollary 4.5. *Let $A, B \in \mathbb{M}_n$. Then,*

$$|A| \prec_w |B| \iff \|A\| \leq \|B\| \text{ for all symmetric norms.}$$

Proof. The implication $\Rightarrow$ is a direct consequence of Corollary 4.3, the other implication is trivial. $\square$

For positive matrices, the same criterion has a purely eigenvalue-vector formulation.

Corollary 4.6. *Let A, B be two positive matrices in $\mathbb{M}_n$. Then,*

$$A \prec_w B \iff a^\downarrow \in \text{conv}\{\varepsilon b_\sigma^\downarrow : \varepsilon \in \mathcal{E}_n, \sigma \in \mathfrak{S}_n\}.$$

Proof. By Corollary 4.4 there exists a diagonal matrix $\Delta \in \mathbb{M}_n^+$ such that

$$A^\downarrow + \Delta \prec B^\downarrow.$$

By Proposition 4.2,

$$a^\downarrow + \delta \in \text{conv}\{b_\sigma^\downarrow : \sigma \in \mathfrak{S}_n\}.$$

Note that $a^\downarrow \in \text{conv}\{\varepsilon(a^\downarrow + \delta) : \varepsilon \in \mathcal{E}_n\}$ and so for each $\nu \in \mathcal{E}_n$, we have

$$\nu(a^\downarrow + \delta) \in \text{conv}\{\nu b_\sigma^\downarrow : \sigma \in \mathfrak{S}_n\}.$$

Therefore $a^\downarrow \in \text{conv}\{\varepsilon b_\sigma^\downarrow : \varepsilon \in \mathcal{E}_n, \sigma \in \mathfrak{S}_n\}$. Conversely, each vector $\varepsilon b_\sigma^\downarrow$ is weakly majorized by $b^\downarrow$, and weak majorization is preserved under convex combinations. Hence the convex-hull condition implies $A \prec_w B$. $\square$

For classical proofs of Ky Fan's principle, see [Bha97, HP14, HJ91]; for the unitary-orbit viewpoint used here, see [BL12, BL26a].

Remark 4.7. Corollary 4.5 is Ky Fan's principle, a fundamental fact. The statement of Corollary 4.4 is somewhat original, and its proof is new. It is based on a systematic approach to majorization through inequalities related to unitary orbits. Stronger results than Corollary 4.3, together with several other unitary-orbit results, will be introduced in Section 8.

5. THE OPERATOR RIGHT MODULUS AND TRIANGLE INEQUALITIES

In this section, we survey some triangle-type estimates for the operator right modulus. In contrast with Thompson's inequality, where unitary conjugations are allowed on the right-hand side, we compare $|\sum_k A_k|$ directly with $\sum_k |A_k|$. Such direct comparisons are much more delicate: in general they fail without additional assumptions, suitable constants, or averaging over unitary orbits. We begin with elementary examples that explain these obstructions, and then review several positive results under natural hypotheses.

5.1. Operator version. As we know, the operator inequality

$$|A + B| \leq |A| + |B|, \quad A, B \in \mathbb{M}_n,$$

does not hold in general; for example, consider $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, and $B = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$. This elementary example explains why the unitary conjugations in Thompson's triangle inequality are indispensable.

Naturally, one may ask whether there exists a finite constant $C(n)$ such that

$$|A + B| \leq C(n)(|A| + |B|), \quad A, B \in \mathbb{M}_n.$$

Unfortunately, no such finite constant exists when $n \geq 2$; equivalently, $C(n) = \infty$. Indeed, it suffices to consider the matrices

$$A_t = \begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix}, \quad B_t = \begin{pmatrix} 1 & -t \\ 0 & 0 \end{pmatrix}, \quad t > 0.$$

A direct computation gives

$$|A_t + B_t| = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad |A_t| + |B_t| = \frac{2}{\sqrt{1+t^2}} \begin{pmatrix} 1 & 0 \\ 0 & t^2 \end{pmatrix}.$$

Thus any such constant must satisfy $C(n) \geq \sqrt{1+t^2}$ for all $t > 0$.

If the operators are invertible, however, the situation changes substantially. Before presenting the corresponding results, we recall that the condition number of an invertible matrix A is defined by

$$\kappa(A) := \|A\|_\infty \|A^{-1}\|_\infty.$$

Lee [Lee12] proved the following condition-number-dependent triangle inequality, in which the constant depends only on a common upper bound ω for the condition numbers of the matrices. Zhang [Zha26a] later showed, by an explicit example, that this dependence on ω is sharp.

Theorem 5.1 [Lee12, Zha26a]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be invertible matrices whose condition numbers are bounded by $\omega \geq 1$. Then*

$$|A_1 + \dots + A_m| \leq \frac{\omega + 1}{2\sqrt{\omega}} (|A_1| + \dots + |A_m|), \quad (5.1)$$

where the constant is sharp as a function of ω .

Moreover, Zhang [Zha26a] obtained the following more refined version of (5.1) in terms of the condition numbers.

Theorem 5.2 [Zha26a]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be invertible matrices whose condition numbers are $\kappa_1, \dots, \kappa_m$. Then*

$$|A_1 + \dots + A_m| \leq \frac{\kappa_1 + 1}{2\sqrt{\kappa_1}} |A_1| + \dots + \frac{\kappa_m + 1}{2\sqrt{\kappa_m}} |A_m|,$$

where the coefficient function $c(\kappa) := \frac{\kappa+1}{2\sqrt{\kappa}}$ is individually sharp for each term: for every $\kappa \geq 1$, the coefficient $c(\kappa)$ in front of a term of condition number κ cannot be replaced by any smaller constant, even if all the other coefficients are kept unchanged.

If one allows a single unitary conjugation on the right-hand side, Bourin and Lee [BL26e] obtained the following direct comparison between $|A + B|$ and $|A| + |B|$ for normal matrices.

Proposition 5.3 [BL26e]. *Let $A, B \in \mathbb{M}_n$ be normal. Then, for some unitary $V \in \mathbb{M}_n$,*

$$|A + B| \leq |A| + |B| + \frac{1}{4}V(|A| + |B|)V^*.$$

If $n \geq 3$, the constant $1/4$ is the smallest possible one.

It is natural to ask what remains true if the normality assumption in Proposition 5.3 is removed and arbitrary complex matrices are allowed. More precisely, we pose the following problem.

Problem 5.4. *Let $n \geq 2$. Determine all pairs $(C_1(n), C_2(n)) \in [0, \infty)^2$ with the following property: for every $A, B \in \mathbb{M}_n$, there exists a unitary matrix $V \in \mathbb{M}_n$ such that*

$$|A + B| \leq C_1(n)(|A| + |B|) + C_2(n)V(|A| + |B|)V^*.$$

Similarly, allowing two unitary conjugations on the right-hand side leads to the following natural variant.

Problem 5.5. *Let $n \geq 2$. Determine all pairs $(C_1(n), C_2(n)) \in [0, \infty)^2$ with the following property: for every $A, B \in \mathbb{M}_n$, there exist unitary matrices $U, V \in \mathbb{M}_n$ such that*

$$|A + B| \leq C_1(n)U(|A| + |B|)U^* + C_2(n)V(|A| + |B|)V^*.$$

It follows easily from Thompson's inequality that every pair in $[1, \infty)^2$ has the property stated in Problem 5.5.

More generally, allowing an average over n unitary conjugates, where n is the size of the matrices, Bourin and Lee [BL26a] established the following comparison.

Theorem 5.6 [BL26a]. *If $A_1, \dots, A_m \in \mathbb{M}_n$ are normal matrices, then, for some unitary matrices $V_1, \dots, V_n \in \mathbb{M}_n$,*

$$\left| \sum_{k=1}^m A_k \right| \leq \frac{1}{n} \sum_{i=1}^n V_i \left\{ \sum_{k=1}^m |A_k| \right\} V_i^*.$$

5.2. Unitarily invariant norms. An interesting result of Lee [Lee12] gives the following bound for sums of matrices under any unitarily invariant norm, Zhang [Zha26b] subsequently proved that this constant is sharp when $n \geq m$.

Theorem 5.7 [Lee12, Zha26b]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left\| \sum_{k=1}^m A_k \right\| \leq \sqrt{m} \left\| \sum_{k=1}^m |A_k| \right\|, \quad (5.2)$$

where the constant is sharp when $n \geq m$.

Zhang [Zha26b] also established the following sharp estimate, which strengthens (5.2).

Theorem 5.8 [Zha26b]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left\| \sum_{k=1}^m A_k \right\| \leq \sqrt{\min\{m, n\}} \left\| \sum_{k=1}^m |A_k| \right\|.$$

Moreover, the constant $\sqrt{\min\{m, n\}}$ is sharp.

5.3. Schatten p -norms. Lee [Lee10] posed the problem of determining the optimal constant c_p such that

$$\|A + B\|_p \leq c_p \| |A| + |B| \|_p, \quad \text{for all } A, B \in \mathbb{M}_n,$$

which is now commonly referred to as *Lee's problem*. Lee's problem is notoriously difficult. The cases $p = 1$ and $p = \infty$ are relatively easy to handle, whereas even the case $p = 2$ is already quite challenging; the remaining cases are still open. In particular, Lee [Lee10] conjectured that $c_2 = \sqrt{\frac{1+\sqrt{2}}{2}}$. This conjecture was first confirmed by Lin and Zhang [LZ22]. Subsequently, Zhang [Zha25a] provided an alternative proof, and [Zha25b] further refined the estimate by deriving a sharper bound in which the constant is expressed more precisely in terms of the singular values of A and B .

Motivated by Lee's problem, Tang and Zhang [TZ26] proposed the following generalization for several matrices $A_k \in \mathbb{M}_n$, $1 \leq k \leq m$: determine the optimal constant $c_p(m)$ depending only on p and m such that

$$\left\| \sum_{k=1}^m A_k \right\|_p \leq c_p(m) \left\| \sum_{k=1}^m |A_k| \right\|_p. \quad (5.3)$$

For the above problem, Tang and Zhang [TZ26] stated several sharp values and general bounds. More precisely, they asserted that

- (1) $c_1(m) = 1$;
- (2) $c_2(m) = \sqrt{\frac{1+\sqrt{m}}{2}}$, and they exhibited m matrices of size $m \times m$ for which equality is attained in (5.3);
- (3) $c_\infty(m) = \sqrt{m}$, and they exhibited m matrices of size $m \times m$ for which equality is attained;
- (4) $c_p(m) \leq m^{\frac{1}{2} - \frac{1}{2p}}$ for all $1 \leq p \leq \infty$.

Since the constant $c_p(m)$ in (5.3) is required to work uniformly over the matrix size, the $m \times m$ extremizers in [TZ26] are sufficient to prove the optimality of any such dimension-independent constant. They do not, however, determine the best constant in each fixed dimension $n < m$. This motivates introducing the dimension-sensitive constant $c_p(m, n)$ and studying the sharper problem stated below.

Problem 5.9 [Zha26b]. *Let $1 \leq p \leq \infty$ and let $A_1, \dots, A_m \in \mathbb{M}_n$. Determine the optimal constant $c_p^{\text{abs}}(m, n)$ such that*

$$\left\| \sum_{k=1}^m A_k \right\|_p \leq c_p^{\text{abs}}(m, n) \left\| \sum_{k=1}^m |A_k| \right\|_p.$$

For Problem 5.9, Zhang [Zha26b] proved the following dimension-dependent conclusions:

- (1) $c_1^{\text{abs}}(m, n) = 1$;
- (2) $c_2^{\text{abs}}(m, n) = \sqrt{\frac{1+\sqrt{\min\{m, n\}}}{2}}$;
- (3) $c_\infty^{\text{abs}}(m, n) = \sqrt{\min\{m, n\}}$;
- (4) $c_p^{\text{abs}}(m, n) \leq (\min\{m, n\})^{\frac{1}{2} - \frac{1}{2p}}$ for all $1 \leq p \leq \infty$.

In particular, these bounds are dimension-sensitive and, when $m > n$, they strictly sharpen the corresponding estimates in Tang and Zhang [TZ26].

6. THE QUADRATIC SYMMETRIC MODULUS AND TRIANGLE INEQUALITIES

For a matrix $Z \in \mathbb{M}_n$, its left and right moduli $|Z^*|$ and $|Z|$ occur in a number of important operator inequalities. There is no compelling reason to favor either the left or the right modulus. We will see that the consideration of symmetrized versions, or means of these two moduli, leads to several interesting results. In this section, we discuss the quadratic symmetric modulus.

The quadratic symmetric modulus of $Z \in \mathbb{M}_n$ is defined as

$$|Z|_{\text{qsym}} := \sqrt{\frac{|Z|^2 + |Z^*|^2}{2}}.$$

Let $\text{Re } Z = (Z + Z^*)/2$ and $\text{Im } Z = (Z - Z^*)/2i$ denote the real and imaginary parts of $Z \in \mathbb{M}_n$ and note that

$$|Z|_{\text{qsym}} = \sqrt{|\text{Re } Z|^2 + |\text{Im } Z|^2}.$$

Hence the quadratic symmetric modulus appears to be a very natural extension of the modulus in $\mathbb{C}$.

6.1. Operator version. The following result is a striking symmetrized version of Thompson's inequality.

Theorem 6.1 [Zha26a]. *Let $A, B \in \mathbb{M}_n$. Then there exist two unitaries $U, V \in \mathbb{M}_n$ such that*

$$|A + B|_{\text{qsym}} \leq U|A|_{\text{qsym}}U^* + V|B|_{\text{qsym}}V^*.$$

We adopt a proof of Bourin and Lee in [BL26e].

Proof of Theorem 6.1. Note that (1.2) is equivalent to the existence of a pair of contractions $K, L \in \mathbb{M}_n$ such that

$$|A + B| \leq K|A|K^* + L|B|L^*. \quad (6.1)$$

Clearly (1.2) $\Rightarrow$ (6.1). Conversely, $K|A|K^* = XX^*$ with $X = K|A|^{1/2}$. So

$$K|A|K^* = U|A|^{1/2}K^*K|A|^{1/2}U^*$$

for some unitary $U \in \mathbb{M}_n$. Since $K^*K \leq I$, we have $K|A|K^* \leq U|A|U^*$. Similarly, $L|B|L^* \leq V|B|V^*$ for some unitary V . Hence (6.1) $\Rightarrow$ (1.2).

Now, given four matrices $A_1, A_2, B_1, B_2 \in \mathbb{M}_n$ and letting

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

Thompson's theorem for A and B reads as

$$\begin{bmatrix} \sqrt{|A_1 + B_1|^2 + |A_2 + B_2|^2} & 0 \\ 0 & 0 \end{bmatrix} \leq U_0 \begin{bmatrix} \sqrt{|A_1|^2 + |A_2|^2} & 0 \\ 0 & 0 \end{bmatrix} U_0^* + V_0 \begin{bmatrix} \sqrt{|B_1|^2 + |B_2|^2} & 0 \\ 0 & 0 \end{bmatrix} V_0^*$$

for some unitary matrices $U_0, V_0 \in \mathbb{M}_{2n}$. Thus

$$\sqrt{|A_1 + B_1|^2 + |A_2 + B_2|^2} \leq K\sqrt{|A_1|^2 + |A_2|^2}K^* + L\sqrt{|B_1|^2 + |B_2|^2}L^*$$

for some contractions $K, L \in \mathbb{M}_n$. By the first part of the proof, we have

$$\sqrt{|A_1 + B_1|^2 + |A_2 + B_2|^2} \leq U\sqrt{|A_1|^2 + |A_2|^2}U^* + V\sqrt{|B_1|^2 + |B_2|^2}V^*$$

for some unitaries $U, V \in \mathbb{M}_n$. Then, following Zhang, we pick $A_2 = A_1^*$ and $B_2 = B_1^*$ to get

$$\sqrt{|A_1 + B_1|^2 + |(A_1 + B_1)^*|^2} \leq U\sqrt{|A_1|^2 + |A_1^*|^2}U^* + V\sqrt{|B_1|^2 + |B_1^*|^2}V^*.$$

Dividing by $\sqrt{2}$,

$$|A_1 + B_1|_{\text{qsym}} \leq U|A_1|_{\text{qsym}}U^* + V|B_1|_{\text{qsym}}V^*.$$

This proves Theorem 6.1. □

An interesting special case [BL26c] is:

Corollary 6.2. *Let $Z \in \mathbb{M}_n$. Then, for some unitaries $U, V \in \mathbb{M}_n$,*

$$|Z|_{\text{qsym}} \leq U|\operatorname{Re} Z|U^* + V|\operatorname{Im} Z|V^*.$$

Applying Weyl's inequalities, we infer:

Corollary 6.3. *Let $Z \in \mathbb{M}_n$. Then, for all $j, k = 0, 1, 2, \dots$ such that $1 + j + k \leq n$,*

$$\lambda_{1+j+k}^\downarrow(|Z|_{\text{qsym}}) \leq \lambda_{1+j}^\downarrow(|\operatorname{Re} Z|) + \lambda_{1+k}^\downarrow(|\operatorname{Im} Z|).$$

This is a sharp result, in the sense that we cannot take $U = V$, even up to a scalar multiple. Indeed:

Proposition 6.4. *Fix $c > 0$. There exists $C \in \mathbb{M}_3$ such that*

$$\det \frac{|C| + |C^*|}{2} > c \det(|\operatorname{Re} C| + |\operatorname{Im} C|).$$

Proof. Take

$$C_t = \begin{pmatrix} 0 & 2t & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad t > 0.$$

Then

$$\frac{|C_t| + |C_t^*|}{2} = \text{diag}(t, t+1, 1),$$

and

$$|\text{Re } C_t| + |\text{Im } C_t| = \frac{2}{\sqrt{t^2 + 1}} \text{diag}(t^2, t^2 + 1, 1).$$

Therefore

$$\frac{\det \frac{|C_t| + |C_t^*|}{2}}{\det(|\text{Re } C_t| + |\text{Im } C_t|)} = \frac{(t+1)\sqrt{t^2 + 1}}{8t} \rightarrow \infty,$$

which proves the assertion. $\square$

Thompson's inequality for quadratic symmetric modulus immediately induces a norm inequality for every symmetric norm.

Corollary 6.5. *For every symmetric norm on $\mathbb{M}_n$, the functional $X \mapsto \| |X|_{\text{qsym}} \|$ is a weakly unitarily invariant norm.*

Combining Theorem 6.1 with the Aujla–Bourin inequality (2.1), we obtain the following consequence.

Theorem 6.6. *Let $A, B \in \mathbb{M}_n$ and let $f : [0, \infty) \rightarrow [0, \infty)$ be concave. Then, there exist unitary matrices U, V such that*

$$f(|A + B|_{\text{qsym}}) \leq U f(|A|_{\text{qsym}}) U^* + V f(|B|_{\text{qsym}}) V^*.$$

The following condition-number estimates for the quadratic symmetric modulus were proved in [Zha26a]. We record them here as quadratic symmetric analogues of the estimates in Section 5.

Proposition 6.7 [Zha26a]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be invertible matrices whose condition numbers are bounded by $\omega \geq 1$. Then*

$$|A_1 + \dots + A_m|_{\text{qsym}} \leq \frac{\omega + 1}{2\sqrt{\omega}} (|A_1|_{\text{qsym}} + \dots + |A_m|_{\text{qsym}}),$$

where the constant is sharp as a function of ω .

Proposition 6.8 [Zha26a]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be invertible matrices whose condition numbers are $\kappa_1, \dots, \kappa_m$. Then*

$$|A_1 + \dots + A_m|_{\text{qsym}} \leq \frac{\kappa_1 + 1}{2\sqrt{\kappa_1}} |A_1|_{\text{qsym}} + \dots + \frac{\kappa_m + 1}{2\sqrt{\kappa_m}} |A_m|_{\text{qsym}},$$

where the coefficient function $c(\kappa) := \frac{\kappa+1}{2\sqrt{\kappa}}$ is individually sharp for each term: for every $\kappa \geq 1$, the coefficient $c(\kappa)$ in front of a term of condition number κ cannot be replaced by any smaller constant, even if all the other coefficients are kept unchanged.

6.2. Unitarily invariant norms. We next pass from operator inequalities to inequalities for unitarily invariant norms. In this setting, one can compare the ordinary sum directly with the sum of the quadratic symmetric moduli. The following result gives a sharp universal constant.

Theorem 6.9 [Zha26b]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left\| \sum_{k=1}^m A_k \right\| \leq \sqrt{2} \left\| \sum_{k=1}^m |A_k|_{\text{qsym}} \right\|.$$

Moreover, the constant $\sqrt{2}$ is sharp.

The preceding theorem compares the ordinary sum with the sum of quadratic symmetric moduli. Applying the quadratic symmetric modulus to the total sum leads to the following dimension-dependent estimate.

Theorem 6.10 [Zha26b]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left\| \sum_{k=1}^m A_k \right\|_{\text{qsym}} \leq \sqrt{\min\{m, 2n\}} \left\| \sum_{k=1}^m |A_k|_{\text{qsym}} \right\|.$$

Furthermore, we strongly suspect that the constant in Theorem 6.10 is not optimal; in particular, we have not been able to find any examples that attain equality. This leaves an open problem.

6.3. Schatten p -norms. For Schatten p -norms, the comparison between the ordinary sum and the sum of quadratic symmetric moduli admits sharper constants than those available for general unitarily invariant norms. In fact, the optimal behavior exhibits a phase transition at $p = 2$: no loss is needed for $1 \leq p \leq 2$, whereas a dimension-free factor appears for $p \geq 2$.

Theorem 6.11 [Zha26b]. *Let $1 \leq p \leq \infty$ and let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\begin{aligned} \left\| \sum_{k=1}^m A_k \right\|_p &\leq \left\| \sum_{k=1}^m |A_k|_{\text{qsym}} \right\|_p, & 1 \leq p \leq 2; \\ \left\| \sum_{k=1}^m A_k \right\|_p &\leq 2^{\frac{1}{2} - \frac{1}{p}} \left\| \sum_{k=1}^m |A_k|_{\text{qsym}} \right\|_p, & 2 \leq p \leq \infty. \end{aligned}$$

Moreover, the constants 1 and $2^{\frac{1}{2} - \frac{1}{p}}$ are sharp.

Although Theorem 6.11 holds for the whole Schatten scale, the Hilbert–Schmidt case $p = 2$ has a special flavor. In this case the inequality admits a particularly elegant proof, based only on two simple observations. We record these observations as lemmas before giving the proof.

Lemma 6.12. *For all $A, B \in \mathbb{M}_n$,*

$$|\text{Tr}(A^* B)|^2 \leq \text{Tr}(|A| |B|) \text{Tr}(|A^*| |B^*|). \quad (6.2)$$

Consequently,

$$|\text{Tr}(A^* B)| \leq \frac{\text{Tr}(|A| |B|) + \text{Tr}(|A^*| |B^*|)}{2}. \quad (6.3)$$

Proof. Let $A = U |A|$ and $B = V |B|$ be polar decompositions. Then

$$\begin{aligned} |\text{Tr}(A^* B)| &= |\text{Tr}(|A| U^* V |B|)| \\ &= |\text{Tr}((|A|^{\frac{1}{2}} U^* V |B|^{\frac{1}{2}})(|B|^{\frac{1}{2}} |A|^{\frac{1}{2}}))| \\ &\leq \left\| |A|^{\frac{1}{2}} U^* V |B|^{\frac{1}{2}} \right\|_2 \left\| |B|^{\frac{1}{2}} |A|^{\frac{1}{2}} \right\|_2 \quad (\text{By Cauchy-Schwarz inequality}) \\ &= (\text{Tr}(|A^*| |B^*|))^{\frac{1}{2}} (\text{Tr}(|A| |B|))^{\frac{1}{2}} \end{aligned}$$

which is equivalent to (6.2).

The bound (6.3) follows from (6.2) by the scalar AM–GM inequality. $\square$

Lieb’s theorem [Lie73] asserts that the map $(X, Y) \mapsto \text{Tr}(X^{\frac{1}{2}} Y^{\frac{1}{2}})$ is jointly concave on $\mathbb{M}_n^+ \times \mathbb{M}_n^+$; for a shorter proof, see also [Bha07, Chapter 4]. That is,

Lemma 6.13. *Let $X_1, X_2, Y_1, Y_2 \in \mathbb{M}_n^+$. Then*

$$\text{Tr} \left(\left(\frac{X_1 + X_2}{2} \right)^{\frac{1}{2}} \left(\frac{Y_1 + Y_2}{2} \right)^{\frac{1}{2}} \right) \geq \frac{\text{Tr}(X_1^{\frac{1}{2}} Y_1^{\frac{1}{2}}) + \text{Tr}(X_2^{\frac{1}{2}} Y_2^{\frac{1}{2}})}{2}.$$

Next, we present a sharp inequality for the Frobenius norm. Although it is subsumed by Theorem 6.11, the argument in this special case is particularly short and elegant.

Proof of the case $p = 2$ in Theorem 6.11. Clearly,

$$\| |A_k|_{\text{qsym}} \|_2^2 = \|A_k\|_2^2. \quad (6.4)$$

For any $i \neq j$, by using Lemma 6.12 and Lemma 6.13 with $X_1 = |A_i|^2$, $X_2 = |A_i^*|^2$, $Y_1 = |A_j|^2$ and $Y_2 = |A_j^*|^2$, we have

$$\begin{aligned} |\operatorname{Tr}(A_i^* A_j)| &\leq \frac{\operatorname{Tr}(|A_i| |A_j|) + \operatorname{Tr}(|A_i^*| |A_j^*|)}{2} \\ &\leq \operatorname{Tr} \left(\left(\frac{|A_i|^2 + |A_i^*|^2}{2} \right)^{\frac{1}{2}} \left(\frac{|A_j|^2 + |A_j^*|^2}{2} \right)^{\frac{1}{2}} \right) \\ &= \operatorname{Tr} \left(|A_i|_{\text{qsym}} |A_j|_{\text{qsym}} \right). \end{aligned}$$

Hence,

$$\begin{aligned} \left\| \sum_{k=1}^m A_k \right\|_2^2 &= \sum_{k=1}^m \|A_k\|_2^2 + 2 \sum_{i < j} \Re \operatorname{Tr}(A_i^* A_j) \\ &\leq \sum_{k=1}^m \|A_k\|_2^2 + 2 \sum_{i < j} |\operatorname{Tr}(A_i^* A_j)| \\ &\leq \sum_{k=1}^m \| |A_k|_{\text{qsym}} \|_2^2 + 2 \sum_{i < j} \operatorname{Tr} \left(|A_i|_{\text{qsym}} |A_j|_{\text{qsym}} \right) \\ &= \left\| \sum_{k=1}^m |A_k|_{\text{qsym}} \right\|_2^2, \end{aligned}$$

where we used (6.4). Taking square roots yields the desired result. $\square$

The preceding theorem compares the ordinary sum $\sum_k A_k$ with the sum of the individual quadratic symmetric moduli. A more intrinsic triangle-type question is obtained by applying the quadratic symmetric modulus to the total sum itself. This leads to the following dimension-sensitive optimization problem.

Problem 6.14 [Zha26b]. Let $A_1, \dots, A_m \in \mathbb{M}_n$. Determine the optimal constant $c_p^{\text{qsym}}(m, n)$ such that

$$\left\| \sum_{k=1}^m A_k \right\|_{\text{qsym}, p} \leq c_p^{\text{qsym}}(m, n) \left\| \sum_{k=1}^m |A_k|_{\text{qsym}} \right\|_p.$$

For Problem 6.14, Zhang [Zha26b] obtained the following results:

- (1) $c_1^{\text{qsym}}(m, n) = 1$;
- (2) $c_2^{\text{qsym}}(m, n) \leq \sqrt{\frac{1 + \sqrt{\min\{m, 2n\}}}{2}}$;
- (3) $c_\infty^{\text{qsym}}(m, n) \leq \sqrt{\min\{m, 2n\}}$;
- (4) $c_p^{\text{qsym}}(m, n) \leq (\min\{m, 2n\})^{\frac{1}{2} - \frac{1}{2p}}$ for all $1 \leq p \leq \infty$.

7. THE ARITHMETIC SYMMETRIC MODULUS AND TRIANGLE INEQUALITIES

We now turn from the quadratic symmetric modulus to its arithmetic analogue. For a matrix $Z \in \mathbb{M}_n$, the polar decomposition

$$Z = |Z^*|U = U|Z|$$

shows that the right modulus $|Z|$ and the left modulus $|Z^*|$ are two natural positive representatives of Z . It is therefore natural to average them and consider the arithmetic symmetric modulus

$$|Z|_{\text{sym}} := \frac{|Z| + |Z^*|}{2}.$$

This modulus is more closely tied to the two-sided structure of Z than the ordinary right modulus, but it is also more delicate than the quadratic symmetric modulus studied in Section 6. The results in this section illustrate both phenomena: some Thompson-type inequalities survive after suitable modifications, while the most direct analogue fails in general. For related developments, see also [BL26c].

7.1. Operator version. As shown in the preceding sections, Thompson's inequality holds for the ordinary right modulus and for the quadratic symmetric modulus. It is therefore natural to ask whether the same phenomenon persists for the arithmetic symmetric modulus. The answer, however, is negative.

Theorem 7.1 [Zha26a]. *Let $A, B \in \mathbb{M}_n$. In general, there do not exist unitaries U, V such that*

$$|A + B|_{\text{sym}} \leq U|A|_{\text{sym}}U^* + V|B|_{\text{sym}}V^*.$$

Indeed, there exist $A, B \in \mathbb{M}_2$ such that

$$\| |A + B|_{\text{sym}} \|_{\infty} > \| |A|_{\text{sym}} \|_{\infty} + \| |B|_{\text{sym}} \|_{\infty}.$$

Although the direct Thompson-type inequality fails, it is still natural to ask whether it can be restored after allowing a uniform multiplicative constant. This leads to the following dimension-dependent problem.

Problem 7.2. *Determine the smallest constant $C(n) > 0$ such that, for every $A, B \in \mathbb{M}_n$, there exist unitary matrices $U, V \in \mathbb{M}_n$ satisfying*

$$|A + B|_{\text{sym}} \leq C(n) (U|A|_{\text{sym}}U^* + V|B|_{\text{sym}}V^*). \quad (7.1)$$

Zhang [Zha26b] conjectured that for all $n \geq 3$, $C(n) = \sqrt{2}$ in (7.1).

The condition-number estimates for the right modulus also yield corresponding bounds for the arithmetic symmetric modulus. In particular, Theorems 5.1 and 5.2 immediately imply the following two consequences.

Proposition 7.3. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be invertible matrices whose condition numbers are bounded by $\omega \geq 1$. Then*

$$|A_1 + \dots + A_m|_{\text{sym}} \leq \frac{\omega + 1}{2\sqrt{\omega}} (|A_1|_{\text{sym}} + \dots + |A_m|_{\text{sym}}),$$

where the constant is sharp as a function of ω .

Proposition 7.4. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be invertible matrices whose condition numbers are $\kappa_1, \dots, \kappa_m$. Then*

$$|A_1 + \dots + A_m|_{\text{sym}} \leq \frac{\kappa_1 + 1}{2\sqrt{\kappa_1}} |A_1|_{\text{sym}} + \dots + \frac{\kappa_m + 1}{2\sqrt{\kappa_m}} |A_m|_{\text{sym}},$$

where the coefficient function $c(\kappa) := \frac{\kappa+1}{2\sqrt{\kappa}}$ is individually sharp for each term: for every $\kappa \geq 1$, the coefficient $c(\kappa)$ in front of a term of condition number κ cannot be replaced by any smaller constant, even if all the other coefficients are kept unchanged.

Even without invertibility assumptions, one can still recover an operator comparison by averaging over unitary orbits. In this direction, Bourin and Lee [BL26a] obtained the following averaged unitary-orbit inequality.

Theorem 7.5 [BL26a]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then, for some unitaries $V_1, \dots, V_n \in \mathbb{M}_n$,*

$$\left| \sum_{k=1}^m A_k \right|_{\text{sym}} \leq \frac{\sqrt{2}}{n} \sum_{i=1}^n V_i \left\{ \sum_{k=1}^m |A_k|_{\text{sym}} \right\} V_i^*.$$

7.2. Unitarily invariant norms. Zhang [Zha26b] established a direct sharp comparison between the sum $\sum_k A_k$ and the sum of the symmetric moduli.

Theorem 7.6 [Zha26b]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left\| \sum_{k=1}^m A_k \right\| \leq 2 \left\| \sum_{k=1}^m |A_k|_{\text{sym}} \right\|.$$

Moreover, the constant 2 is sharp.

Bourin and Lee [BL26e] proved the following uniform triangle estimate for the arithmetic symmetric modulus. The constant is known to be sharp in all dimensions $n \geq 3$, while the sharpness in dimension $n = 2$ remains open; see [Zha26b].

Theorem 7.7 [BL26e]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then, for all symmetric norms,*

$$\left\| \left| \sum_{k=1}^m A_k \right|_{\text{sym}} \right\| \leq \sqrt{2} \left\| \sum_{k=1}^m |A_k|_{\text{sym}} \right\|,$$

where the constant is sharp for all $n \geq 3$.

Proposition 7.8. *There exist rank-one matrices $A, B \in \mathbb{M}_3$ with ± 1 -entries such that*

$$\left\| |A + B|_{\text{sym}} \right\|_{\infty} = \sqrt{2} \left\| |A|_{\text{sym}} + |B|_{\text{sym}} \right\|_{\infty}.$$

For instance, take

$$A = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Both matrices have rank one. A direct computation gives

$$|A|_{\text{sym}} + |B|_{\text{sym}} = 2I_3, \quad \left\| |A + B|_{\text{sym}} \right\|_{\infty} = 2\sqrt{2}.$$

Thus the constant $\sqrt{2}$ in Theorem 7.7 is sharp in dimension 3. Embedding this example as a direct summand shows sharpness in all dimensions $n \geq 3$.

Since the symmetric modulus is invariant by a congruence with a permutation matrix and by multiplication by -1 , there are several examples of ± 1 -matrices satisfying Proposition 7.8.

Taking traces in Theorem 7.7 gives only the displayed constant $\sqrt{2}$. The usual trace-norm triangle inequality is recovered separately from

$$\text{Tr } |A|_{\text{sym}} = \text{Tr } |A|.$$

The same unitary-orbit argument also yields the following eigenvalue refinement.

Proposition 7.9 [BL26e]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. If $j \geq 0$ and $1 + 3j \leq n$, then*

$$\lambda_{1+3j}^{\downarrow} \left(\left| \sum_{k=1}^m A_k \right|_{\text{sym}} \right) \leq \sqrt{2} \lambda_{1+j}^{\downarrow} \left(\sum_{k=1}^m |A_k|_{\text{sym}} \right).$$

To describe a sharper form of the preceding estimate, we introduce a structural condition on the polar factor.

Definition 7.10. We say that a matrix S is *polar Hermitian* if it has a polar decomposition $S = V|S|$ in which V is a scalar multiple of a Hermitian unitary.

The class includes all involutions, as the next example records.

Example 7.11. If $S \in \mathbb{M}_n$ is an involution (a symmetry, $S = S^{-1}$), then S is a polar Hermitian matrix. In fact, the unitary part of S is Hermitian. Here is a simple proof from [BL26b]. Since $S = U|S| = |S^*|U$ and $S = S^{-1}$ we infer $U|S| = U^*|S^*|^{-1}$ and so $U = U^*$ as the polar decomposition of an invertible matrix is unique. The main result of [BL26b] says much more on the structure of involutions.

This observation allows the constant $\sqrt{2}$ in Theorem 7.7 to be improved to 1 for polar-Hermitian sums.

Theorem 7.12 [BL26e]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$ be such that their sum is polar Hermitian. Then, for all symmetric norms,*

$$\left\| \left| \sum_{k=1}^m A_k \right|_{\text{sym}} \right\| \leq \left\| \sum_{k=1}^m |A_k|_{\text{sym}} \right\|.$$

7.3. Schatten p -norms. We finally specialize the preceding norm inequalities to the Schatten scale. For the arithmetic symmetric modulus, Zhang obtained a sharp comparison between the ordinary sum and the sum of the individual symmetric moduli. The optimal constant depends only on p , and is independent of both the number of summands and the matrix size.

Theorem 7.13 [Zha26b]. *Let $1 \leq p \leq \infty$ and let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left\| \sum_{k=1}^m A_k \right\|_p \leq 2^{1-\frac{1}{p}} \left\| \sum_{k=1}^m |A_k|_{\text{sym}} \right\|_p.$$

Moreover, the constant $2^{1-\frac{1}{p}}$ is sharp.

The preceding theorem is a “sum versus sum” estimate. A more intrinsic triangle-type problem is to apply the arithmetic symmetric modulus to the total sum itself and then compare it with the sum of the individual symmetric moduli. This leads to the following dimension-sensitive optimization problem.

Problem 7.14 [Zha26b]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Determine the optimal constant $c_p^{\text{sym}}(m, n)$ such that*

$$\left\| \left| \sum_{k=1}^m A_k \right|_{\text{sym}} \right\|_p \leq c_p^{\text{sym}}(m, n) \left\| \sum_{k=1}^m |A_k|_{\text{sym}} \right\|_p.$$

In Problem 7.14, Zhang [Zha26b] proved that

- (1) $c_1^{\text{sym}}(m, n) = 1$;
- (2) $c_\infty^{\text{sym}}(m, n) = \sqrt{2}$ for $n \geq 3$;
- (3) $c_p^{\text{sym}}(m, n) \leq \min\{2^{1-\frac{1}{p}}, \sqrt{2}\}$;
- (4) $c_p^{\text{sym}}(m, n) \geq \left(\frac{2^{\frac{p}{2}} + 2^{1-\frac{p}{2}}}{3} \right)^{\frac{1}{p}}$ for $n \geq 3$; in particular, the right-hand side increases from 1 at $p = 2$ to $\sqrt{2}$ as $p \rightarrow \infty$.

8. SOME HYBRID TRIANGLE INEQUALITIES AND FURTHER PROBLEMS

The preceding sections treated the right modulus, the quadratic symmetric modulus, and the arithmetic symmetric modulus separately. We now turn to inequalities in which these different moduli interact with one another. Such hybrid estimates often refine the norm inequalities discussed above, and they also lead naturally to questions involving weak majorization, weak log-majorization, and unitary-orbit averages.

We begin with the basic comparison between the arithmetic and quadratic symmetric moduli. By the operator concavity of the square-root function,

$$|Z|_{\text{sym}} \leq |Z|_{\text{qsym}}. \quad (8.1)$$

Indeed, this is just the inequality

$$\frac{|Z| + |Z^*|}{2} \leq \left(\frac{|Z|^2 + |Z^*|^2}{2} \right)^{1/2}.$$

In fact, a stronger squared form holds. By the operator convexity of t^2 ,

$$\left(\frac{|Z| + |Z^*|}{2} \right)^2 \leq \frac{|Z|^2 + |Z^*|^2}{2},$$

or equivalently,

$$|Z|_{\text{sym}}^2 \leq |Z|_{\text{qsym}}^2. \quad (8.2)$$

Since the square-root function is operator monotone, (8.2) implies (8.1). The next proposition [BL26d] records a geometric interpretation of the stronger inequality (8.2).

Proposition 8.1. *Let $Z \in \mathbb{M}_n$ and let $\mathcal{U}$ be the (closed) unit ball of $\mathbb{C}^n$. Then,*

$$|Z|_{\text{sym}}(\mathcal{U}) \subset |Z|_{\text{qsym}}(\mathcal{U}).$$

Proof. Thanks to (8.2) it suffices to show that, for $A, B \in \mathbb{M}_n^+$, we have the equivalence

$$A^2 \leq B^2 \iff A(\mathcal{U}) \subset B(\mathcal{U}).$$

Since each side ensures that the range of A is included in the range of B , restricting our operators to the range of B if necessary, we may assume that B is invertible. Then

$$A^2 \leq B^2 \iff B^{-1}A^2B^{-1} \leq I \iff \|B^{-1}A\|_\infty \leq 1 \iff B^{-1}A(\mathcal{U}) \subset \mathcal{U},$$

equivalently, $A(\mathcal{U}) \subset B(\mathcal{U})$. $\square$

Given $A, B \in \mathbb{M}_n^+$, we write $A \prec_{\text{wlog}} B$ to mean that the following n inequalities hold:

$$\prod_{j=1}^p \lambda_j^\downarrow(A) \leq \prod_{j=1}^p \lambda_j^\downarrow(B), \quad p = 1, \dots, n.$$

It is well known that $A \prec_{\text{wlog}} B \Rightarrow A \prec_w B$.

Clearly, Theorem 6.9 is equivalent to

$$\left| \sum_{k=1}^m A_k \right| \prec_w \sqrt{2} \sum_{k=1}^m |A_k|_{\text{qsym}}, \quad A_k \in \mathbb{M}_n.$$

The comparison in Theorem 6.9 can in fact be strengthened from weak majorization to weak log-majorization. The following result of Bourin and Lee gives such a refinement.

Theorem 8.2 [BL26d]. *Let $A_1, \dots, A_m \in \mathbb{M}_n$. Then*

$$\left| \sum_{k=1}^m A_k \right| \prec_{\text{wlog}} \sqrt{2} \sum_{k=1}^m |A_k|_{\text{qsym}}.$$

This suggests that several norm inequalities discussed earlier may admit stronger spectral forms. We therefore record the following open problem.

Problem 8.3. *Determine which of the unitarily invariant norm inequalities in Sections 5–7 can be strengthened to weak log-majorization.*

The same circle of ideas also leads to direct comparisons between the quadratic and arithmetic symmetric moduli. In particular, Bourin and Lee proved the following averaged unitary-orbit estimate.

Theorem 8.4 [BL26a]. *Let $Z \in \mathbb{M}_n$. Then, for some unitaries $V_1, \dots, V_n \in \mathbb{M}_n$,*

$$|Z|_{\text{qsym}} \leq \frac{\sqrt{2}}{n} \sum_{i=1}^n V_i |Z|_{\text{sym}} V_i^*.$$

The constant $\sqrt{2}$ cannot be diminished.

The preceding results concern the two endpoint cases $p = 1$ and $p = 2$, corresponding respectively to the arithmetic and quadratic symmetric moduli. It is natural to interpolate between them by considering the following power-mean symmetric modulus. For $p > 0$, define

$$|Z|_{p,\text{sym}} := \left(\frac{|Z|^p + |Z^*|^p}{2} \right)^{1/p}.$$

For this family, Zhang obtained a sharp range for the Cartesian Thompson-type inequality. The result shows that the quadratic case is the largest power for which such an estimate remains valid.

Theorem 8.5 [Zha26a]. *Fix $n \geq 2$ and $p > 0$. The following assertions are equivalent:*

- (1) $0 < p \leq 2$;
- (2) *for every $Z \in \mathbb{M}_n$, there exist unitaries $U, V \in \mathbb{M}_n$ such that*

$$|Z|_{p,\text{sym}} \leq U |\operatorname{Re} Z| U^* + V |\operatorname{Im} Z| V^*.$$

This theorem suggests that the moduli $|\cdot|_{p,\text{sym}}$ may exhibit substantially different triangle behavior in the ranges $0 < p \leq 2$ and $p > 2$. We therefore leave the following problem for further study.

Problem 8.6. *Study triangle inequalities and optimal constants for the power-mean symmetric moduli $|\cdot|_{p,\text{sym}}$.*

In particular, the quadratic case suggests a broader norm-theoretic question. Although Thompson-type operator inequalities may fail outside the range $0 < p \leq 2$, it is still natural to ask whether the associated symmetric-norm functionals satisfy a triangle inequality.

Conjecture 8.7. *Let $p \geq 1$, and let*

$$|X|_{p,\text{sym}} := \left(\frac{|X|^p + |X^*|^p}{2} \right)^{1/p}.$$

Then, for every symmetric norm $\|\cdot\|$ on $\mathbb{M}_n$, the functional

$$X \mapsto \| |X|_{p,\text{sym}} \|$$

is a weakly unitarily invariant norm on $\mathbb{M}_n$.

REFERENCES

- [AB07] J. S. Aujla and J.-C. Bourin, *Eigenvalue inequalities for convex and log-convex functions*, Linear Algebra Appl. **424** (2007), no. 1, 25–35. doi:[10.1016/j.laa.2006.02.027](https://doi.org/10.1016/j.laa.2006.02.027)
- [And89] T. Ando, *Majorization, doubly stochastic matrices, and comparison of eigenvalues*, Linear Algebra Appl. **118** (1989), 163–248. doi:[10.1016/0024-3795\(89\)90580-6](https://doi.org/10.1016/0024-3795(89)90580-6)
- [Bha97] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, New York, 1997. doi:[10.1007/978-1-4612-0653-8](https://doi.org/10.1007/978-1-4612-0653-8)
- [Bha07] R. Bhatia, *Positive Definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton and Oxford, 2007.
- [Bou99] J.-C. Bourin, *Some inequalities for norms on matrices and operators*, Linear Algebra Appl. **292** (1999), no. 1–3, 139–154. doi:[10.1016/S0024-3795\(99\)00024-5](https://doi.org/10.1016/S0024-3795(99)00024-5).
- [Bou06] J.-C. Bourin, *Matrix versions of some classical inequalities*, Linear Algebra Appl. **416** (2006), no. 2–3, 890–907. doi:[10.1016/j.laa.2006.01.002](https://doi.org/10.1016/j.laa.2006.01.002)
- [Bou10] J.-C. Bourin, *A matrix subadditivity inequality for symmetric norms*, Proc. Amer. Math. Soc. **138** (2010), no. 2, 495–504. doi:[10.1090/S0002-9939-09-10103-X](https://doi.org/10.1090/S0002-9939-09-10103-X)

- [BL12] J.-C. Bourin and E.-Y. Lee, *Unitary orbits of Hermitian operators with convex or concave functions*, Bull. Lond. Math. Soc. **44** (2012), no. 6, 1085–1102. doi:[10.1112/blms/bds080](https://doi.org/10.1112/blms/bds080)
- [BL26a] J.-C. Bourin and E.-Y. Lee, *Averages over matrix unitary orbits and spectral order*, preprint, arXiv:2606.15624, 2026. doi:[10.48550/arXiv.2606.15624](https://doi.org/10.48550/arXiv.2606.15624)
- [BL26b] J.-C. Bourin and E.-Y. Lee, *Involutions and angles between subspaces*, Linear Algebra Appl. **730** (2026), 124–137. doi:[10.1016/j.laa.2025.10.017](https://doi.org/10.1016/j.laa.2025.10.017)
- [BL26c] J.-C. Bourin and E.-Y. Lee, *Matrix parallelogram laws and symmetric moduli*, Internat. J. Math. **37** (2026), no. 3, 2650018. doi:[10.1142/S0129167X26500187](https://doi.org/10.1142/S0129167X26500187)
- [BL26d] J.-C. Bourin and E.-Y. Lee, *Some hybrid matrix triangle inequalities*, Linear Algebra Appl. **749** (2026), 197–207; arXiv:2606.29188. doi:[10.1016/j.laa.2026.06.033](https://doi.org/10.1016/j.laa.2026.06.033)
- [BL26e] J.-C. Bourin and E.-Y. Lee, *Triangle inequalities for the operator symmetric modulus*, Proc. Amer. Math. Soc., in press; arXiv:2602.19607, 2026. doi:[10.48550/arXiv.2602.19607](https://doi.org/10.48550/arXiv.2602.19607)
- [FH55] K. Fan and A. J. Hoffman, *Some metric inequalities in the space of matrices*, Proc. Amer. Math. Soc. **6** (1955), 111–116. doi:[10.1090/S0002-9939-1955-0067841-7](https://doi.org/10.1090/S0002-9939-1955-0067841-7)
- [Gol65] S. Golden, *Lower bounds for the Helmholtz function*, Phys. Rev. (2) **137** (1965), B1127–B1128. doi:[10.1103/PhysRev.137.B1127](https://doi.org/10.1103/PhysRev.137.B1127)
- [HJ91] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991. doi:[10.1017/CBO9780511840371](https://doi.org/10.1017/CBO9780511840371)
- [HP14] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*, Universitext, Springer, Cham; Hindustan Book Agency, New Delhi, 2014. doi:[10.1007/978-3-319-04150-6](https://doi.org/10.1007/978-3-319-04150-6)
- [Lee10] E.-Y. Lee, *Rotfel'd type inequalities for norms*, Linear Algebra Appl. **433** (2010), no. 3, 580–584. doi:[10.1016/j.laa.2010.03.029](https://doi.org/10.1016/j.laa.2010.03.029)
- [Lee12] E.-Y. Lee, *How to compare the absolute values of operator sums and the sums of absolute values?*, Oper. Matrices **6** (2012), no. 3, 613–619. doi:[10.7153/oam-06-42](https://doi.org/10.7153/oam-06-42)
- [Lie73] E. H. Lieb, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, Adv. Math. **11** (1973), 267–288. doi:[10.1016/0001-8708\(73\)90011-X](https://doi.org/10.1016/0001-8708(73)90011-X)
- [LZ22] J. Lin and Y. Zhang, *A proof of Lee's conjecture on the sum of absolute values of matrices*, J. Math. Anal. Appl. **516** (2022), no. 2, Paper No. 126542, 6 pp. doi:[10.1016/j.jmaa.2022.126542](https://doi.org/10.1016/j.jmaa.2022.126542)
- [Neu27] J. von Neumann, *Thermodynamik quantenmechanischer Gesamtheiten*, Nachr. Ges. Wiss. Göttingen Math.-Phys. Kl. **1927** (1927), 273–291.
- [Neu37] J. von Neumann, *Some matrix-inequalities and metrization of matrix-space*, Tomsk Univ. Rev. **1** (1937), 286–300.
- [Rot69] S. Yu. Rotfel'd, *The singular numbers of the sum of completely continuous operators*, in *Spectral Theory, Topics in Mathematical Physics*, vol. 3, Consultants Bureau, New York, 1969, pp. 73–78. doi:[10.1007/978-1-4684-7589-0_5](https://doi.org/10.1007/978-1-4684-7589-0_5)
- [TZ26] Q. Tang and S. Zhang, *Generalizing Lee's conjecture on the sum of absolute values of matrices*, Linear Algebra Appl. **731** (2026), 196–204. doi:[10.1016/j.laa.2025.11.015](https://doi.org/10.1016/j.laa.2025.11.015)
- [Tho65] C. J. Thompson, *Inequality with applications in statistical mechanics*, J. Math. Phys. **6** (1965), no. 11, 1812–1813. doi:[10.1063/1.1704727](https://doi.org/10.1063/1.1704727)
- [Tho76] R. C. Thompson, *Convex and concave functions of singular values of matrix sums*, Pacific J. Math. **66** (1976), no. 1, 285–290. doi:[10.2140/pjm.1976.66.285](https://doi.org/10.2140/pjm.1976.66.285)
- [Zha25a] T. Zhang, *A new proof of Lee's conjecture on the Frobenius norm via the matrix Cauchy–Schwarz inequality*, Linear Algebra Appl. **725** (2025), 355–358. doi:[10.1016/j.laa.2025.07.017](https://doi.org/10.1016/j.laa.2025.07.017)
- [Zha25b] T. Zhang, *Sharp perturbation bounds on the Frobenius norm of subunitary and positive polar factor*, preprint, arXiv:2507.14940, 2025. doi:[10.48550/arXiv.2507.14940](https://doi.org/10.48550/arXiv.2507.14940)
- [Zha26a] T. Zhang, *An operator triangle inequality for the quadratic symmetric modulus*, preprint, arXiv:2602.01463, 2026. doi:[10.48550/arXiv.2602.01463](https://doi.org/10.48550/arXiv.2602.01463)
- [Zha26b] T. Zhang, *Operator symmetric moduli and sharp triangle inequalities*, preprint, arXiv:2603.01046, 2026. doi:[10.48550/arXiv.2603.01046](https://doi.org/10.48550/arXiv.2603.01046)

UNIVERSITÉ MARIE ET LOUIS PASTEUR, CNRS, LMB (UMR 6623), F-25000 BESANÇON, FRANCE
 Email address: jcbourin@univ-fcomte.fr

SCHOOL OF MATHEMATICS AND STATISTICS, XI'AN JIAOTONG UNIVERSITY, XI'AN 710049, P. R. CHINA
 Email address: teng.zhang@stu.xjtu.edu.cn