The spectral spread of Hermitian matrices

Pedro Massey, Demetrio Stojanoff and Sebastián Zárate *
Centro de Matemática, FCE-UNLP, La Plata and IAM-CONICET, Argentina

To the memory of “el Diego” Maradona

Abstract

Let A be an $n \times n$ complex Hermitian matrix and let $\lambda(A) = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ denote the eigenvalues of A, counting multiplicities and arranged in non-increasing order. Motivated by problems arising in the theory of low rank matrix approximation, we study the spectral spread of A, denoted $\operatorname{Spr}^+(A)$, given by $\operatorname{Spr}^+(A) = (\lambda_1 - \lambda_n, \lambda_2 - \lambda_{n-1}, \dots, \lambda_k - \lambda_{n-k+1}) \in \mathbb{R}^k$, where $k = [n/2]$ (integer part). The spectral spread is a vector-valued measure of dispersion of the spectrum of A, that allows one to obtain several submajorization inequalities. In the present work we obtain inequalities that are related to Tao's inequality for anti-diagonal blocks of positive semidefinite matrices, Zhan's inequalities for the singular values of differences of positive semidefinite matrices, extremal properties of direct rotations between subspaces, generalized commutators and distances between matrices in the unitary orbit of a Hermitian matrix.

AMS subject classification: 42C15, 15A60.
Keywords: spectral spread, submajorization, principal angles, generalized commutators.

Contents

1 Introduction **2**

2 Spectral spread **4**
2.1 Preliminaries 4
2.2 Basic properties of the spectral spread 5
2.3 A key inequality for the spectral spread 8

3 Reformulations of the key inequality **14**
3.1 Generalized commutators and unitary conjugates 14
3.2 Equivalences of the inequalities 17
3.3 Concluding remarks 19

4 Appendix **19**

*Partially supported by CONICET (PIP 0152 CO), FONCyT (PICT-2015-1505) and UNLP (11X829)
e-mail addresses: massey@mate.unlp.edu.ar , demetrio@mate.unlp.edu.ar , seb4.zarate@gmail.com

1 Introduction

In this work we develop several results related to the spectral spread of Hermitian matrices (for its definition see Eq. (1) below). The study of this notion is motivated by some recent problems related to the absolute variation of Ritz values, which is one of the many aspects of low rank matrix approximation of Hermitian matrices (see [20, 22]). Indeed, given an $n \times n$ complex Hermitian matrix A and a k-dimensional subspace $\mathcal{S}$ of $\mathbb{C}^n$, then the Ritz values of A corresponding to $\mathcal{S}$ are the eigenvalues (counting multiplicities and arranged in non-increasing order) $\lambda(S^*AS) = (\lambda_i(S^*AS))_{i\in\mathbb{I}_k} \in \mathbb{R}^k$, where S is an $n \times k$ isometry with range $\mathcal{S}$ (here $\mathbb{I}_k = \{1, \ldots, k\}$ is an index set). If $\mathcal{T} \subset \mathbb{C}^n$ is another k-dimensional subspace then, the absolute variation of the Ritz values of A related to $\mathcal{S}$ and $\mathcal{T}$ is the vector

$$(|\lambda_i(S^*AS) - \lambda_i(T^*AT)|)_{i\in\mathbb{I}_k} \in \mathbb{R}^k$$

where T is an $n \times k$ isometry with range $\mathcal{T}$. This topic has been extensively studied (see [1, 5, 15, 13, 14, 16, 17, 24, 29]). One of the major problems in this context is to obtain upper bounds for the variation of the Ritz values in terms of some measure of the spread of the spectrum of A and some measure of the distance between the subspaces $\mathcal{S}$ and $\mathcal{T}$. As a natural vector-valued measure of distance between the subspaces it is usually considered $\Theta(\mathcal{S}, \mathcal{T}) = (\theta_i)_{i\in\mathbb{I}_k} \in \mathbb{R}^k$ the so-called vector of principal angles between $\mathcal{S}$ and $\mathcal{T}$ (for details see Section 2.1). On the other hand, as a measure of the spread of the spectrum of A, many authors have considered the diameter i.e., $\lambda_{\max}(A) - \lambda_{\min}(A) \geq 0$.

In [15], A. Knyazev et.al. considered the vector valued measure of the spread of the spectrum of A denoted by $\text{Spr}^+(A)$ given by

$$\text{Spr}^+(A) = (\lambda_i(A) - \lambda_{n-i+1}(A))_{i\in\mathbb{I}_{[n/2]}} = \big(\lambda_i(A) - \lambda_i^{\uparrow}(A)\big)_{i\in\mathbb{I}_{[n/2]}} \in \mathbb{R}_{\geq 0}^{[n/2]}, \tag{1}$$

where $\lambda^{\uparrow}(A) = (\lambda_i^{\uparrow}(A))_{i\in\mathbb{I}_n}$ is the vector of eigenvalues, but arranged in non-decreasing order, and $[n/2]$ denotes the integer part of $n/2$. Similarly, we can consider

$$\text{Spr}(A) = \lambda(A) + \lambda(-A) = \big(\lambda_i(A) - \lambda_i^{\uparrow}(A)\big)_{i\in\mathbb{I}_n} \in \mathbb{R}^n. \tag{2}$$

Notice that $\text{Spr}(A) \in \mathbb{R}^n$ is a symmetric vector, that is $\text{Spr}_i(A) = -\text{Spr}_{n-i+1}(A)$, for $i \in \mathbb{I}_n$. Moreover, using Weyl's inequality it turns out that $\text{Spr}(A)$ is a vector-valued measure of the diameter of the unitary orbit of A. With these notions, in [15] the authors conjectured that

$$(|\lambda_i(S^*AS) - \lambda_i(T^*AT)|)_{i\in\mathbb{I}_k} \prec_w (\sin(\theta_i)\ \text{Spr}_i^+(A))_{i\in\mathbb{I}_m} \tag{3}$$

where the previous inequality is with respect to submajorization and $m = \min\{k, [n/2]\}$. It is well known that submajorization relations (as that conjectured in Eq. (3)) imply inequalities with respect to arbitrary unitarily invariant norms, and tracial inequalities involving convex non-decreasing functions (see Sections 2.1 and 4 for details).

In [17] we obtained some inequalities related to the variation of Ritz values that are weak versions of Eq. (3). It turns out that Eq. (3) encodes some subtle aspects of the spectral spread $\text{Spr}^+(A)$ that are still not understood. Indeed, at that time we realized that although natural, the spectral spread seemed not to have been considered in the literature. Thus, on the one hand we consider it is interesting to develop some of its basic features. On the other hand, motivated by the seminal ideas from [1, 15, 29, 30] in this work we propose some inequalities involving the spectral spread. For example, given the k-dimensional subspace $\mathcal{S} \subset \mathbb{C}^n$ and $n \times n$ complex Hermitian matrix A as before, if we let $\mathcal{T} = e^{i\,A}\,\mathcal{S}$ then numerical experiments supported the submajorization inequality

$$\Theta(\mathcal{S}\,,\,\mathcal{T}) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(A)\,. \tag{4}$$

It turns out that Eq. (4) (that reflects some extremal properties of direct rotations, as introduced by Davis and Kahan in [6]) is equivalent to the following submajorization inequality: if $n = k + r$ and we let A be the $n \times n$ complex Hermitian matrix with blocks

$$A = \left[\begin{array}{cc} A_1 & B \\ B^* & A_2 \end{array}\right] \begin{array}{c} \mathbb{C}^k \\ \mathbb{C}^r \end{array} \quad \text{then} \quad 2\,s(B) \prec_w \mathrm{Spr}^+(A)\ , \tag{5}$$

where $s(B) = \lambda(\,(B^*B)^{1/2}) \in \mathbb{R}^r_{\geq 0}$ denotes the vector of singular values of the $k \times r$ matrix B, i.e. the eigenvalues of the modulus $|B| = (B^*B)^{1/2}$.

In this paper we prove Eq. (5) (hence, also Eq. (4)), which we consider as a key inequality for the spectral spread; we point out that this inequality is sharp. In turn, Eq. (5) connects our work with Tao's work [23], where he showed that for a positive semidefinite matrix A with blocks

$$A = \left[\begin{array}{cc} A_1 & B \\ B^* & A_2 \end{array}\right] \begin{array}{c} \mathbb{C}^k \\ \mathbb{C}^r \end{array} \quad \text{it holds that} \quad 2\,s_i(B) \leq s_i(A \oplus A) \quad \text{for} \quad i \in \mathbb{I}_k\ . \tag{6}$$

Notice that although Eq. (5) provides a spectral relation that is weaker than the entry-wise inequalities in Eq. (6), our upper bound for positive semidefinite A satisfies

$$\mathrm{Spr}^+_i(A) = \lambda_i(A) - \lambda_{n-i+1}(A) \leq \lambda_i(A) \leq \lambda_i(A \oplus A) = s_i(A \oplus A)\ , \quad \text{for} \quad i \in \mathbb{I}_{[n/2]}\ .$$

For example, in case $A = a\,I$ then $B = 0$ a fact that is reflected by $\mathrm{Spr}^+(A) = 0$, while $s_i(A \oplus A) = a$, for $i \in \mathbb{I}_n$. On the other hand, Eq. (5) is valid in the (general) Hermitian case. Again, by [2, 23] it turns out that our work is connected with Zhan's inequality for the singular values of the difference of positive semidefinite matrices in [26, 27, 28]. Motivated by these facts, we show that Eq. (5) is equivalent to the inequality

$$s(A_1 - A_2) \prec_w \mathrm{Spr}^+(A_1 \oplus A_2) \tag{7}$$

for arbitrary complex Hermitian matrices A_1 and A_2 of the same size. In this case, Eq. (7) and Zhan's inequality can be compared in a way similar to the comparison between Eq. (5) and Eq. (6). We point out that the upper bounds obtained in Eq. (5) and Eq. (7) are vectors that are invariant by translations $M \mapsto M + \lambda\,I$ for the matrices M involved, as opposed to previous upper bounds that are based on singular values, i.e. Eq, (6) and Zhan's inequality. Since the vectors that are being bounded in these theorems are also invariant under the corresponding translations, we consider that the upper bounds in terms of the spread are particularly well suited in this context.

On the other hand, it turns out that Eq. (7) can be extended to the context of generalized commutators as follows: given $n \times n$ complex matrices A_1, A_2, X such that A_1 and A_2 are Hermitian then

$$s(A_1\,X - X\,A_2) \prec_w \big(\,s_i(X)\ \mathrm{Spr}^+_i(A_1 \oplus A_2)\,\big)_{i \in \mathbb{I}_n}\,. \tag{8}$$

This last inequality connects our work with a series of papers dealing with (even more general) inequalities for singular values of generalized commutators [7, 8, 10, 11, 12, 25]. We point out that in the previous works, the authors obtain entry-wise upper bounds for the singular values of generalized commutators in terms of singular values and some measures of the spread of related matrices (among other type of inequalities). Our results are obtained in terms of weaker submajorization relations, but the upper bound in Eq. (8) involves the complete list of singular values of X and the full spectral spread of $A_1 \oplus A_2$. On the other hand, we point out that Eq. (8) holds for arbitrary Hermitian matrices A_1 and A_2.

We point out that Eq. (5) together with some of its equivalent forms allow one to develop inequalities related to Eq. (3) (see [18]). In the last section of the paper we show the equivalence of the inequalities in Eqs. (4), (5), (7) and (8).

2 Spectral spread

Although natural, the spectral spread of Hermitian matrices seems not to have been considered in the literature, after being introduced in [15]. Thus, we begin with a preliminary section (with notations and basic definitions), and then we present some basic results related to this notion. After this, we consider a submajorization inequality for the spectral spread that plays a key role in our work. We obtain some consequences of this inequality related to Zhan's [26] inequality and Davis-Kahan's notion of direct rotation between subspaces [6] (see also [21]).

2.1 Preliminaries

In this section we give the basic notation and definitions that we use throughout our work. In the Appendix (Section 4) we state several well known results of Matrix Analysis involving the notions described below.

Notation and terminology. We let $\mathcal{M}_{n,k}(\mathbb{C})$ be the space of complex $n \times k$ matrices and write $\mathcal{M}_{n,n}(\mathbb{C}) = \mathcal{M}_n(\mathbb{C})$ for the algebra of $n \times n$ complex matrices. We denote by $\mathcal{H}(n) \subset \mathcal{M}_n(\mathbb{C})$ the real subspace of Hermitian matrices and by $\mathcal{M}_n(\mathbb{C})^+$, the cone of positive semi-definite matrices. Also, $\mathcal{G}l(n) \subset \mathcal{M}_n(\mathbb{C})$ and $\mathcal{U}(n)$ denote the groups of invertible and unitary matrices respectively, and $\mathcal{G}l(n)^+ = \mathcal{G}l(n) \cap \mathcal{M}_n(\mathbb{C})^+$. A norm N in $\mathcal{M}_n(\mathbb{C})$ is unitarily invariant (briefly u.i.n.) if $N(UAV) = N(A)$, for every $A \in \mathcal{M}_n(\mathbb{C})$ and $U, V \in \mathcal{U}(n)$.

For $n \in \mathbb{N}$, let $\mathbb{I}_n = \{1, \ldots, n\}$. Given a vector $x \in \mathbb{C}^n$ we denote by D_x the diagonal matrix in $\mathcal{M}_n(\mathbb{C})$ whose main diagonal is x. Given $x = (x_i)_{i\in\mathbb{I}_n} \in \mathbb{R}^n$ we denote by $x^\downarrow = (x_i^\downarrow)_{i\in\mathbb{I}_n}$ the vector obtained by rearranging the entries of x in non-increasing order. We also use the notation $(\mathbb{R}^n)^\downarrow = \{x \in \mathbb{R}^n \ : \ x = x^\downarrow\}$ and $(\mathbb{R}^n_{\geq 0})^\downarrow = \{x \in \mathbb{R}^n_{\geq 0} \ : \ x = x^\downarrow\}$. Similarly we define $x^\uparrow$ and $(\mathbb{R}^n)^\uparrow$. For $r \in \mathbb{N}$, we let $\mathbb{1}_r = (1, \ldots, 1) \in \mathbb{R}^r$.

Given a matrix $A \in \mathcal{H}(n)$ we denote by $\lambda(A) = (\lambda_i(A))_{i\in\mathbb{I}_n} \in (\mathbb{R}^n)^\downarrow$ the eigenvalues of A counting multiplicities and arranged in non-increasing order. Similarly, we denote by $\lambda^\uparrow(A) = (\lambda_i^\uparrow(A))_{i\in\mathbb{I}_n} \in (\mathbb{R}^n)^\uparrow$. For $B \in \mathcal{M}_{k\,,\,r}(\mathbb{C})$ we let $s(B) = \lambda(|B|) \in (\mathbb{R}^r_{\geq 0})^\downarrow$ denote the singular values of B, i.e. the eigenvalues of $|B| = (B^*B)^{1/2} \in \mathcal{M}_r(\mathbb{C})^+$.

Arithmetic operations with vectors are performed entry-wise in the following sense: in case $x = (x_i)_{i\in\mathbb{I}_k} \in \mathbb{C}^k$, $y = (y_i)_{i\in\mathbb{I}_r} \in \mathbb{C}^r$ then $x+y = (x_i+y_i)_{i\in\mathbb{I}_m}$, $x\,y = (x_i\,y_i)_{i\in\mathbb{I}_m}$ and (assuming that $y_i \neq 0$, for $i \in \mathbb{I}_r$) $x/y = (x_i/y_i)_{i\in\mathbb{I}_m}$, where $m = \min\{k, r\}$. Moreover, if we assume further that $x, y \in \mathbb{R}^k$ then we write $x \leqslant y$ ($\leqslant$, different from the notation $\leq$) whenever $x_i \leq y_i$, for $i \in \mathbb{I}_k$.

Given $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, and $z = (z_i)_{i\in\mathbb{I}_k} \in I^k$ we denote $f(z) = (f(z_i))_{i\in\mathbb{I}_k} \in \mathbb{R}^k$. For example, $|z| = (|z_i|)_{i\in\mathbb{I}_k}$, $\sin(z) = (\sin(z_i))_{i\in\mathbb{I}_k}$. △

Next we recall the notion of majorization between vectors, that will play a central role throughout our work.

Definition 2.1. Let $x, y \in \mathbb{R}^k$. We say that x is *submajorized* by y, and write $x \prec_w y$, if

$$\sum_{i=1}^{j} x_i^\downarrow \leq \sum_{i=1}^{j} y_i^\downarrow \quad \text{for} \quad j \in \mathbb{I}_k .$$

If $x \prec_w y$ and $\operatorname{tr} x \overset{\text{def}}{=} \sum_{i=1}^{k} x_i = \operatorname{tr} y$, then we say that x is *majorized* by y, and write $x \prec y$.

We point out that (sub)majorization is a preorder relation in $\mathbb{R}^k$ that plays a central role in matrix analysis (see Section 4).

Remark 2.2. Let $x \in \mathbb{R}^k_{\geq 0}$ and $y \in \mathbb{R}^h_{\geq 0}$ be two vector with *non-negative* entries (of different sizes). We extend the notion of submajorization between x and y in the following sense:

$$x \prec_w y \quad \text{if} \quad \begin{cases} (x\,,\,0_{h-k}) \prec_w \quad y & \text{for} \quad k<h \\ \quad x \quad \prec_w (y\,,\,0_{k-h}) & \text{for} \quad h<k \end{cases} , \tag{9}$$

where 0_n denotes the zero vector of $\mathbb{R}^n$. △

2.2 Basic properties of the spectral spread

In this section we present several basic properties of the spectral spread, and describe the relationship between this notion and singular values (and other usual notions of matrix analysis).

Definition 2.3. Let $A \in \mathcal{H}(n)$. Consider the *full spectral spread* of A, given by

$$\operatorname{Spr}(A) \overset{\text{def}}{=} \lambda(A) + \lambda(-A) = \big(\lambda_i^\downarrow(A) - \lambda_i^\uparrow(A)\big)_{i\in\mathbb{I}_n} \in (\mathbb{R}^n)^\downarrow . \tag{10}$$

Denote by $k = [\frac{n}{2}]$ (integer part). We also consider the *spectral spread of* A, that is the non-negative part of $\operatorname{Spr}(A)$:

$$\operatorname{Spr}^+(A) = \big(\operatorname{Spr}_i(A)\big)_{i\in\mathbb{I}_k} = \big(\lambda_i(A) - \lambda_i^\uparrow(A)\big)_{i\in\mathbb{I}_k} \in (\mathbb{R}^k_{\geq 0})^\downarrow . \tag{11}$$

Remark 2.4. Let $n = 2k$ or $2k+1$, and $A \in \mathcal{H}(n)$. Let $\lambda(A) = (\lambda_i)_{i\in\mathbb{I}_n}$ and $\lambda^\uparrow(A) = (\lambda_i^\uparrow)_{i\in\mathbb{I}_n}$.

1. For every $t \in \mathbb{R}$ we have that

$$\operatorname{Spr}(A - tI) = \operatorname{Spr}(A) \quad \text{and} \quad \operatorname{Spr}^+(A - tI) = \operatorname{Spr}^+(A) . \tag{12}$$

2. For $r \in \mathbb{I}_k$, it is well known (see [3]) that

$$\sum_{i\in\mathbb{I}_r} \lambda_i(A) = \max \left\{ \sum_{i\in\mathbb{I}_r} \langle A\, x_i \,,\, x_i\rangle : \{x_i\}_{i\in\mathbb{I}_r} \text{ is an ONS } \right\} \quad \text{and}$$

$$-\sum_{i\in\mathbb{I}_r} \lambda_i^\uparrow(A) = \sum_{i\in\mathbb{I}_r} \lambda_i(-A) = \max \left\{ -\sum_{i\in\mathbb{I}_r} \langle A\, y_i \,,\, y_i\rangle : \{y_i\}_{i\in\mathbb{I}_r} \text{ is an ONS } \right\}$$

Then, for each $r \in \mathbb{I}_k$ we have that

$$\sum_{i\in\mathbb{I}_r} \mathrm{Spr}_i(A) = \max \left\{ \sum_{i\in\mathbb{I}_r} \langle A\, x_i \,,\, x_i\rangle - \langle A\, y_i \,,\, y_i\rangle : \{x_i\} \text{ and } \{y_i\} \text{ are ONS's } \right\} \tag{13}$$

3. It is straightforward to check that $\max\{|\lambda_i(A)|, |\lambda_i^\uparrow(A)|\} \le s_i(A)$, for $i \in \mathbb{I}_k$; hence, we conclude that

$$\mathrm{Spr}_i^+(A) \le |\lambda_i| + |\lambda_i^\uparrow| \le 2s_i(A) \quad \text{for every} \quad i \in \mathbb{I}_k \ . \tag{14}$$

In the positive case, we have that:

$$A \in \mathcal{M}_n(\mathbb{C})^+ \implies \mathrm{Spr}_i(A) \le \lambda_i = s_i(A) \quad \text{for every} \quad i \in \mathbb{I}_k \ . \tag{15}$$

4. Conversely, notice that

$$\mathrm{Spr}_i^+(A) = \lambda_i - \lambda_{n-i+1} = \underbrace{\lambda_i - \lambda_{k+1}}_{\ge 0} + \underbrace{\lambda_{k+1} - \lambda_{n-i+1}}_{\ge 0} \quad \text{for} \quad i \in \mathbb{I}_k \, .$$

Hence, it follows that

$$s(A - \lambda_{k+1}\, I) = \big(\, (\lambda_i - \lambda_{k+1}\,)_{i\in\mathbb{I}_k} \,,\, (\lambda_{k+1} - \lambda_{k+i}\,)_{i\in\mathbb{I}_{n-k}} \,\big)^\downarrow \prec \mathrm{Spr}^+(A) \in \mathbb{R}^k_{\ge 0} \ . \tag{16}$$

5. On the other hand, if $B \in \mathcal{H}(n)$ then

$$\lambda(A) \prec \lambda(B) \in \mathbb{R}^n \implies \mathrm{Spr}(A) \prec \mathrm{Spr}(B) \implies \mathrm{Spr}^+(A) \prec_w \mathrm{Spr}^+(B) \ , \tag{17}$$

which are direct consequences of Definition 2.3 together with Lemma 4.3.

6. Consider an isometry $Z \in \mathcal{M}_{n,r}(\mathbb{C})$ i.e., such that $Z^*Z = I_r$, for some $r \in \mathbb{I}_n$. Then, it is easy to see that $Z^*AZ \in \mathcal{M}_r(\mathbb{C})$ is a principal submatrix of a unitary conjugate of the matrix A. Therefore, we can apply the interlacing inequalities (see [3]) and get that $\lambda_i^\uparrow(A) \le \lambda_i^\uparrow(Z^*AZ)$ and $\lambda_i(Z^*AZ) \le \lambda_i(A)$ for $i \in \mathbb{I}_r$. As a consequence,

$$\mathrm{Spr}_i^+(Z^*AZ) \le \mathrm{Spr}_i^+(A) \quad \text{for} \quad i \in \mathbb{I}_{[r/2]} \overset{(9)}{\implies} \mathrm{Spr}^+(Z^*AZ) \prec_w \mathrm{Spr}^+(A) \ . \tag{18}$$

7. If $U \in \mathcal{U}(n)$ is a unitary matrix then, using Weyl's inequality (item 1 in Theorem 4.2),

$$\lambda(A - U^*AU) \prec \lambda(A) + \lambda(-U^*AU) = \lambda(A) - \lambda^\uparrow(A) = \mathrm{Spr}(A)\, .$$

Moreover, if N is a u.i.n. then

$$\max \big\{ \, N(A - U^*AU) \ : \ U \in \mathcal{U}(n) \, \big\} = N(D_{\mathrm{Spr}(A)})\, .$$

That is, $|\mathrm{Spr}(A)|^\downarrow = \mathrm{Spr}^+(A \oplus A) = (\mathrm{Spr}^+ A \,,\, \mathrm{Spr}^+ A)^\downarrow$ (with an extra 0 if n is odd) allows us to compute the diameter of the unitary orbit of A (with respect to any unitarily invariant norm). In this sense, $\mathrm{Spr}(A)$ can be considered as a vector valued measure of the diameter of the unitary orbit of A. △

Proposition 2.5. *Let $A \in \mathcal{H}(n)$. Then*

$$\frac{1}{2}\,|\mathrm{Spr}(A)|^{\downarrow} = \frac{1}{2}\,\mathrm{Spr}^{+}(A \oplus A) \prec_w s(A)\ . \tag{19}$$

Proof. Denote by $\lambda(A) = (\lambda_i)_{i \in \mathbb{I}_n} \in (\mathbb{R}^n)^{\downarrow}$. Then for $1 \le i \le k = [\frac{n}{2}]$,

$$\mathrm{Spr}^{+}_{2i-1}(A \oplus A) = \mathrm{Spr}^{+}_{2i}(A \oplus A) = \mathrm{Spr}^{+}_{i}(A) = \lambda_i - \lambda_{n-i+1} \le |\lambda_i| + |\lambda_{n-i+1}|\,.$$

On the one hand we have that $\mathrm{Spr}^{+}(A \oplus A)$ equals

$$\mathrm{Spr}^{+}(A \oplus A) = (\mathrm{Spr}^{+}A\,,\ \mathrm{Spr}^{+}A)^{\downarrow} \in \mathbb{R}^n \quad \text{if} \quad n = 2k$$

$$\text{or} \quad \mathrm{Spr}^{+}(A \oplus A) = (\mathrm{Spr}^{+}A\,,\ \mathrm{Spr}^{+}A\,,\ 0)^{\downarrow} \in \mathbb{R}^n \quad \text{if} \quad n = 2k+1\,.$$

Therefore, in order to check that $\mathrm{Spr}^{+}(A \oplus A) \prec_w 2\,s(A)$ it suffices to check that

$$\sum_{i \in \mathbb{I}_{2r}} \mathrm{Spr}^{+}(A \oplus A) \le 2 \sum_{i \in \mathbb{I}_{2r}} s(A) \quad \text{for} \quad r \in \mathbb{I}_k\,,$$

for the even cases $2r$. Hence, if $r \in \mathbb{I}_k$, we have that

$$\sum_{i \in \mathbb{I}_{2r}} \mathrm{Spr}^{+}(A \oplus A) \le 2 \sum_{i \in \mathbb{I}_{r}} |\lambda_i| + |\lambda_{n-i+1}| \le 2 \sum_{i \in \mathbb{I}_{2r}} s_i(A)$$

since $s(A) = (|\lambda_i|)^{\downarrow}_{i \in \mathbb{I}_n}$, and $\sum\limits_{i \in \mathbb{I}_{2r}} s_i(A) = \max\{\sum\limits_{j \in \mathbb{F}} |\lambda_j| \,:\, |\mathbb{F}| = 2r\}$. Notice that Eq. (19) follows from this fact. □

Notice that if $A \in \mathcal{H}(n)$ then, by items 1. and 4. in Remark 2.4 and Proposition 2.5,

$$\frac{1}{2}\,\mathrm{Spr}^{+}(A \oplus A) \prec s\big(\,A - \lambda_k(A)\,I\,\big) \overset{(16)}{\prec} \mathrm{Spr}^{+}(A)\ . \tag{20}$$

Our next result is a spectral spread version of Lidskii's inequality.

Proposition 2.6. *Let $A, B \in \mathcal{H}(n)$. Then*

$$\mathrm{Spr}(A) - \mathrm{Spr}(B) \prec \mathrm{Spr}(A - B) \prec \mathrm{Spr}(A) - \mathrm{Spr}^{\uparrow}(B) = \mathrm{Spr}(A) + \mathrm{Spr}(B)\ . \tag{21}$$

Proof. By Lidskii's additive inequality and item 3 of Lemma 4.3 (see the Appendix),

$$\begin{aligned} \mathrm{Spr}(A) - \mathrm{Spr}(B) \quad &= \quad \lambda(A) - \lambda(B) + \lambda(-A) - \lambda(-B) \\ &\overset{\text{Lidskii}}{\prec} \quad \lambda(A - B) + \lambda(-(A - B)) = \mathrm{Spr}(A - B)\ . \end{aligned}$$

For the other inequality, note that $\mathrm{Spr}^{\uparrow}(B) = \lambda^{\uparrow}(B) - \lambda(B) = -\mathrm{Spr}(B)$. Therefore

$$\begin{aligned} \mathrm{Spr}(A - B) \quad &= \quad \lambda(A - B) + \lambda(B - A) \\ &\overset{\text{Lidskii}}{\prec} \quad \lambda(A) - \lambda^{\uparrow}(B) + \lambda(B) - \lambda^{\uparrow}(A) = \mathrm{Spr}(A) - \mathrm{Spr}^{\uparrow}(B)\ , \end{aligned}$$

where we have used again item 3 of Lemma 4.3. □

2.3 A key inequality for the spectral spread

The following inequality plays a central role in our present work.

Theorem 2.7. *Let* $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \in \mathcal{H}(k+r)$. *Then*

$$2\, s(B) \prec_w \operatorname{Spr}^+(A)\ . \tag{22}$$

Proof. Consider $U = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \in \mathcal{U}(k+r)$. Then

$$s(UA - AU) = s(A - U^*AU) = |\lambda(A - U^*AU)|^{\downarrow}\,.$$

By Lidskii's inequality (Theorem 4.2) we have that

$$\lambda(A - U^*A\,U) \prec \lambda(A) - \lambda^{\uparrow}(U^*AU) = \lambda(A) - \lambda^{\uparrow}(A)\,.$$

Using Remark 4.5 we get that

$$s(UA - AU) = |\lambda(A - U^*A\,U)|^{\downarrow} \prec_w |\lambda(A) - \lambda^{\uparrow}(A)|^{\downarrow} = \big(\operatorname{Spr}^+(A)\,,\, \operatorname{Spr}^+(A)\,\big)^{\downarrow}$$

(or $\big(\operatorname{Spr}^+(A)\,,\,\operatorname{Spr}^+(A)\,,\,0\big)^{\downarrow}$ if $k+r$ is odd). Using Proposition 4.7 and noticing that

$$UA - AU = \begin{bmatrix} 0 & 2\,B \\ -2\,B^* & 0 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \implies s(UA - AU) = 2\,\big(\,s(B)\,,\, s(B)\,,\, 0_{|k-r|}\,\big)^{\downarrow}\ ,$$

where $s(B)$ has size $\min\{r\,,\,k\} \le [\frac{k+r}{2}]$ (we use that $\operatorname{rk} B \le \min\{r\,,\,k\}$). We conclude that

$$2\,\big(\,s(B)\,,\, s(B)\,,\, 0_{|k-r|}\,\big)^{\downarrow} \prec_w \big(\operatorname{Spr}^+(A)\,,\, \operatorname{Spr}^+(A)\,\big)^{\downarrow} \implies 2\,s(B) \prec_w \operatorname{Spr}^+(A)\ . \qquad \square$$

Remark 2.8. We point out that the inequality in Eq. (22) is sharp. Indeed, consider ($k = r$)

$$\text{if} \quad A = \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^k \end{matrix} \implies \lambda(A) = (s(B)\,,\, -s(B))^{\downarrow} \quad \text{(see Proposition 4.7)} \quad .$$

Thus, in this case we have the equality $2\,s(B) = \operatorname{Spr}^+(A)$. △

Remark 2.9. Let $A \in \mathcal{H}(k+r)$ be as in Theorem 2.7. In [23], Y. Tao proved that if $A \in \mathcal{M}_{k+r}(\mathbb{C})^+$, then

$$2s_j(B) \le \lambda_j(A) = s_j(A) \quad \text{for every} \quad j \in \mathbb{I}_{k+r}\ . \tag{23}$$

In the positive case we have that $\operatorname{Spr}^+(A) \leqslant \lambda(A)$. Nevertheless, the inequality

$$2s_j(B) \le \operatorname{Spr}^+_j(A) \quad \text{for every} \quad j \in \mathbb{I}_{[\frac{k+r}{2}]} \tag{24}$$

is not true, even for positive semidefinite matrices A. For example take

$$A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 1 & 3 \end{pmatrix} \in \mathcal{M}_{2+2}(\mathbb{C})^+$$

Then $s(B) = (1,1)$, and $\lambda(A) = (4.61\,,\, 2.61\,,\, 2.38\,,\, 0.39)$. Hence, Eq. (24) fails for $j = 2$. In particular, Theorem 2.7 does not imply Tao's inequality. Neither Tao's inequality implies Theorem 2.7, as the Eq. (20) could suggests, because $A - \lambda_{[\frac{k+r}{2}]}(A)\, I \notin \mathcal{M}_{k+r}(\mathbb{C})^+$.

On the other hand, when $A \in \mathcal{M}_{k+r}(\mathbb{C})^+$ then both Tao's and our result are applicable. In this case, if N is a unitarily invariant norm and $\mu = (\text{Spr}^+(A),\, 0) \in \mathbb{R}^{k+r}$, then

$$2\, N \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \leq N(D_\mu) \leq N(A)\ , \tag{25}$$

where D_μ is the diagonal matrix with main diagonal μ. Indeed, since $\mu \leqslant \lambda(A) = s(A)$, then

$$2\,(s(B),\, 0) \overset{(2.7)}{\prec_w} (\text{Spr}^+(A),\, 0) \prec_w s(A)$$

and Eq. (25) follows from these relations. We also point out that in (the generic) case $A \in \mathcal{G}l(k+r)^+$ we get a strict inequality $N(D_\mu) < N(A)$, for an arbitrary strictly convex u.i.n. N. On the other hand, notice that Theorem 2.7 also applies in case A is an arbitrary (not necessarily positive semidefinite) Hermitian matrix. △

Corollary 2.10. *Let* $A = \left[\begin{array}{cc} A_1 & B \\ B^* & A_2 \end{array}\right] \begin{array}{c} \mathbb{C}^k \\ \mathbb{C}^r \end{array} \in \mathcal{H}(k+r)$. *Then*

$$\text{Spr}^+ \left[\begin{array}{cc} 0 & B \\ B^* & 0 \end{array}\right] \prec_w \text{Spr}^+(A) \quad \text{and} \quad \text{Spr}^+ \left[\begin{array}{cc} A_1 & 0 \\ 0 & A_2 \end{array}\right] \prec_w \text{Spr}^+(A)\ . \tag{26}$$

Proof. By Proposition 4.7 we get that $\text{Spr}^+ \left[\begin{array}{cc} 0 & B \\ B^* & 0 \end{array}\right] = 2\, s(B)$, (modulo some zeros at the end to equate sizes). Applying Eq. (22), we get the first inequality in Eq. (26). The second follows from Eq.'s (17) and (60). □

In what follows we develop two different applications of the key inequality in Theorem 2.7, that are related to a companion to Zhan's inequality from [26] and with results related to direct rotations between subspaces from [6, 21] in the finite dimensional context. We begin with the following

Remark 2.11. In [26] Zhan showed that given $A_1,\, A_2 \in \mathcal{M}_n(\mathbb{C})^+$ then we have that

$$s_i(A_1 - A_2) \leq s_i(A_1 \oplus A_2) \quad \text{for every} \quad i \in \mathbb{I}_n\ . \tag{27}$$

These fundamental inequalities are known to be equivalent to other central results in matrix analysis (see [2, 23, 27, 28]). On the other hand, these entry-wise inequalities between singular values allow one to get operator inequalities the form $A_1 - A_2 \leq V^*(A_1 \oplus A_2)V$, for suitable contractions $V \in \mathcal{M}_{2n,n}(\mathbb{C})$. Moreover, Eq. (27) imply the (weaker) submajorization relation $s(A_1 - A_2) \prec_w s(A_1 \oplus A_2)$, as in Eq. (9).

In case $C,\, D \in \mathcal{M}_d(\mathbb{C})$ are arbitrary, then the previous inequality does not hold. Nevertheless, a small modification of arguments in [26] imply the inequalities

$$s_i(C - D)\ \leq\ 2\ s_i(C \oplus D) \quad \text{for every} \quad i \in \mathbb{I}_n\ . \tag{28}$$

The previous inequalities are sharp, even for Hermitian matrices C, $D \in \mathcal{H}(n)$ (take $C \in \mathcal{M}_n(\mathbb{C})^+$ and let $D = -C$). As before, these entry-wise inequalities between singular values allow us to get some related operator inequalities and submajorization relations.

On the other hand, if we are interested in norm inequalities with respect to unitarily invariant norms then we can improve the upper bounds derived from Eqs. (27) and (28) for arbitrary Hermitian matrices as follows. △

Theorem 2.12. *Let A_1, $A_2 \in \mathcal{H}(n)$ be Hermitian matrices. Then*

$$s(A_1 - A_2) \prec_w \operatorname{Spr}^+(A_1 \oplus A_2) \ . \tag{29}$$

Proof. We assume that A_1 , $A_2 \in \mathcal{H}(n)$. Let

$$Z = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \in \mathcal{U}(2n) \quad \text{and} \quad T = \frac{1}{2} \begin{bmatrix} A_1 + A_2 & A_1 - A_2 \\ A_1 - A_2 & A_1 + A_2 \end{bmatrix} \in \mathcal{H}(2n) \ .$$

Then $ZTZ^* = A_1 \oplus A_2$; hence, by Theorem 2.7 we see that $s(A_1 - A_2) \prec_w \operatorname{Spr}^+(T) = \operatorname{Spr}^+(A_1 \oplus A_2)$. □

Remark 2.13. Let A_1, $A_2 \in \mathcal{M}_n(\mathbb{C})^+$. As we have already mentioned, in this case we have that $\operatorname{Spr}^+_i(A_1 \oplus A_2) \le s_i(A_1 \oplus A_2)$, for $i \in \mathbb{I}_n$. Using this fact and Theorem 2.12 we get that

$$N(A_1 - A_2) \le N(D_{\operatorname{Spr}^+(A_1 \oplus A_2)}) \le N(D_{(s_i(A_1 \oplus A_2))_{i \in \mathbb{I}_n}}) \ ,$$

for every u.i.n. N. On the other hand, we point out that the entry-wise inequalities

$$s_i(A_1 - A_2) \le \operatorname{Spr}^+_i (A_1 \oplus A_2) \quad \text{for} \quad i \in \mathbb{I}_n \ ,$$

are false, even in the positive semidefinite case. Indeed, take $A_1 = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ and $A_2 = 3\,I$. Then $s(A_1 - A_2) = (2\,,\,2)$ but

$$\lambda\,(A_1 \oplus A_2) = (5\,,\,3\,,\,3\,,\,1) \implies \operatorname{Spr}^+ (A_1 \oplus A_2) = (4\,,\,0)\,.$$

We point out that the upper bounds obtained in Theorems 2.7 and 2.12 are vectors that are invariant by translations $M \mapsto M + \lambda\,I$ for the matrices M involved, as opposed to previous upper bounds that are based on singular values, i.e. Eqs. (23) and (27). Since the vectors that are being bounded in these theorems are also invariant under the corresponding translations, we consider that the upper bounds in terms of the spread are particularly well suited in this context.

Finally, we point out that the inequality in Eq. (29) is sharp; indeed, take an arbitrary $A_1 \in \mathcal{M}_n(\mathbb{C})^+$ and let $A_2 = -A_1 \in \mathcal{H}(n)$. Then $s(A_1 - A_2) = 2\,\lambda(A_1) = \operatorname{Spr}^+(A_1 \oplus A_2)$. △

Remark 2.14. Let $\mathcal{S}, \mathcal{T} \subset \mathbb{C}^n$ be k-dimensional subspaces and let $S, T \in \mathcal{M}_{n,k}(\mathbb{C})$ and $S_\perp \in \mathcal{M}_{n\,,\,n-k}(\mathbb{C})$ be isometries with ranges $R(S) = \mathcal{S}$, $R(T) = \mathcal{T}$and $R(S_\perp) = \mathcal{S}^\perp$. The principal angles $\Theta(\mathcal{S}\,,\,\mathcal{T}) = (\theta_j)_{j \in \mathbb{I}_k} \in ([0, \pi/2]^k)^\downarrow$ between $\mathcal{S}$ and $\mathcal{T}$ are defined by

$$\cos(\theta_j) = s_{k-j+1}(S^*T) \quad \text{or also by} \quad \sin(\theta_j) = s_j(T^*S_\perp) \quad \text{for} \quad j \in \mathbb{I}_k \ , \tag{30}$$

(see, for example, [21] for details). The principal angles $\Theta(\mathcal{S},\mathcal{T})$ completely describe the relative position of these subspaces. On the other hand, they provide a natural notion of distance in the Grassmann manifold of all k-dimensional subspaces in $\mathbb{C}^n$.

In [6] Davis and Kahan introduce the fundamental notion of direct rotation from $\mathcal{S}$ onto $\mathcal{T}$. Briefly, a unitary $U \in \mathcal{U}(n)$ is a direct rotation from $\mathcal{S}$ to $\mathcal{T}$ if $U\mathcal{S} = \mathcal{T}$ and there exist $C_0 \in M_k(\mathbb{C})^+$, $C_1 \in M_{n-k}(\mathbb{C})^+$ and $S_0 \in \mathcal{M}_{n-k,k}(\mathbb{C})$ such that

$$W^*U\,W = \begin{bmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{bmatrix} \in \mathcal{U}(n)$$

for some unitary matrix $W \in \mathcal{U}(n)$, whose first k columns form an orthonormal basis of $\mathcal{S}$. In this case we can write $U = e^{i\,Z}$, where $Z \in \mathcal{H}(n)$ is such that

$$\lambda(Z) = (\Theta(\mathcal{S}\,,\,\mathcal{T})^*, -\,\Theta(\mathcal{S}\,,\,\mathcal{T})^*, 0)^{\downarrow} \in (\mathbb{R}^n)^{\downarrow}\ ,$$

where $\Theta(\mathcal{S}\,,\,\mathcal{T})^* = (\theta_i)_{i\in\mathbb{I}_r} \in ((0,\pi/2]^r)^{\downarrow}$ denotes the vector with the positive principal angles, and hence $r \le [n/2]$. Direct rotations enjoy some extremal properties that play a central role in the study of metric properties in the Grassmann manifold of all k-dimensional subspaces in $\mathbb{C}^n$. Indeed, Davis and Kahan showed in [6] that if $V \in \mathcal{U}(n)$ is such that $V\,\mathcal{S} = \mathcal{T}$ and U is a direct rotation from $\mathcal{S}$ onto $\mathcal{T}$ then

$$2\,\sin(\theta_j/2) = s_j((1-U)|_{\mathcal{S}}) \le s_j((1-V)|_{\mathcal{S}}) \quad \text{for} \quad j \in \mathbb{I}_k\,. \tag{31}$$

On the other hand, if $X \in \mathcal{H}(n)$ is such that $\lambda_j(X) \in [-\pi,\pi]$, for $j \in \mathbb{I}_n$, and $V = e^{-iX}$ then, the interlacing inequalities show that $s_j((1-V)|_{\mathcal{S}}) \le s_j(1-V)$, for $j \in \mathbb{I}_k$. Since $\lambda_j(X)/2 \in [-\pi/2,\pi/2]$ we see that $|\sin(\lambda_j(X)/2)| = \sin(|\lambda_j(X)|/2)$; then, straightforward computations show that $s_j(1-V) = 2\,\sin(s_j(X)/2)$, for $j \in \mathbb{I}_n$. Hence, from Davis and Kahan results it can be deduced that

$$\theta_j = s_j(Z) \le s_j(X) \quad \text{for} \quad j \in \mathbb{I}_k\,. \tag{32}$$

The following result is related to the inequalities in Eq. (32) in a similar sense as previous cases: better bounds (by Proposition 2.5) with respect to a weaker order (weak majorization instead of entrywise inequalities). △

Theorem 2.15. *Let $\mathcal{S},\mathcal{T} \subseteq \mathbb{C}^n$ be subspaces and let $X \in \mathcal{H}(n)$ be such that $e^{iX}\mathcal{S} = \mathcal{T}$. Then*

$$\Theta(\mathcal{S}\,,\,\mathcal{T}) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X)\ . \tag{33}$$

Proof. Let $S \in \mathcal{M}_{n,k}(\mathbb{C})$ be an isometry such that $R(S) = \mathcal{S}$. We consider the smooth curve $T(\cdot) : [0,1] \to M_{n,k}(\mathbb{C})$ given by $T(t) = e^{i\,t\,X}S$, for $t \in [0,1]$; we also set $\mathcal{T}(t) = R(T(t)) \subseteq \mathbb{C}^n$, for $t \in [0,1]$. Notice that $T(0) = S$, $\mathcal{T}(0) = \mathcal{S}$ and $\mathcal{T}(1) = \mathcal{T}$. Since each $T(t)$ is an isometry, the function $\Theta(\cdot) : [0,1] \to [0,\pi/2]^k$ given by $\Theta(t) = \Theta(\mathcal{S}\,,\,\mathcal{T}(t)) = \arccos(s^{\uparrow}(S^*T(t)\,)\,)$ for $t \in [0,1]$, is continuous and $\Theta(0) = 0$. Then, as a consequence of the triangle inequality for principal angles [21, Theorem 1] we see that

$$\Theta(\mathcal{S}\,,\,\mathcal{T}) \ \prec_w \sum_{j=0}^{m-1} \Theta(\mathcal{T}(\tfrac{j}{m})\,,\,\mathcal{T}(\tfrac{j+1}{m})) \tag{34}$$

Notice that $T(t+h) = e^{i\,t\,X}\,T(h)$ with $e^{i\,t\,X} \in \mathcal{U}(n)$, for $t,\ h,\ t+h \in [0,1]$; thus, we see that $\Theta(\mathcal{T}(\frac{j}{m})\,,\,\mathcal{T}(\frac{j+1}{m})) = \Theta(\mathcal{T}(0)\,,\,\mathcal{T}(\frac{1}{m}))$, for each $j \in \mathbb{I}_{n-1}$. Next, we show that

$$\Theta(\mathcal{S}\,,\,\mathcal{T}(\tfrac{1}{m})) \quad \prec_w \tfrac{1}{2m}\,\operatorname{Spr}^+(X) + O(m) \quad \text{ with } \quad \lim_{m\to\infty} m\,O(m) = 0\,. \tag{35}$$

Indeed, since $\frac{d}{dt}\sin(t)|_{t=0} = 1$ and $\sin(0) = 0$ we see that

$$\Theta(\mathcal{S},\,\mathcal{T}(\tfrac{1}{m})) \quad = \sin(\Theta(\mathcal{S},\,\mathcal{T}(\tfrac{1}{m}))) + O_1(m) \quad \text{ with } \quad \lim_{m\to\infty} m\,O_1(m) = 0\,. \tag{36}$$

Let $S_\perp \in \mathcal{M}_{n,n-k}(\mathbb{C})$ be an isometry such that $R(S_\perp) = \mathcal{S}^\perp$. For every $m \in \mathbb{N}$, since $T(\frac{1}{m})$ is an isometry, then Eq. (30) assures that

$$\sin(\Theta(\mathcal{S}\,,\,\mathcal{T}(\tfrac{1}{m}))) \quad = s(T(\tfrac{1}{m})^* S_\perp) = s(S\,e^{\frac{i}{m}\,X}\,S_\perp) \in ([0,1]^k)^\downarrow\,.$$

Since $S\,e^{i\,t\,X}\,S_\perp|_{t=0} = 0$, and $\frac{d}{dt}(S\,e^{i\,t\,X}\,S_\perp)|_{t=0} = i\,S\,X\,S_\perp$ then, we have that

$$S\,e^{\frac{i}{m}\,X}\,S_\perp = \frac{i}{m}\,S\,X\,S_\perp + O_2(m) \quad \text{ with } \quad \lim_{m\to\infty} m\,O_2(m) = 0\,. \tag{37}$$

Eqs. (22), (36) and (37) together with Weyl's inequality (Theorem 4.1) imply that

$$\begin{aligned}\Theta(\mathcal{S},\,\mathcal{T}(\tfrac{1}{m})) \quad &\prec_w \tfrac{1}{2m} s\big(\,2\,S\,X\,S_\perp\,\big) + s\big(\,O_2(m)\,\big) + O_1(m)\\ &\overset{(22)}{\prec_w} \tfrac{1}{2m}\,\operatorname{Spr}^+(X) + O(m)\,,\end{aligned}$$

which shows that Eq. (35) holds. Then, by Eq. (34) we get that

$$\Theta(\mathcal{S}\,,\,\mathcal{T}) \prec_w \frac{1}{2}\,\operatorname{Spr}^+(X) + m\,O(m)\,.$$

The result now follows by taking the limit when $m \to \infty$. □

As a consequence of Theorem 2.15 we strengthen [21, Theorem 6].

Corollary 2.16. *Let $\mathcal{S}, \mathcal{T} \subseteq \mathbb{C}^n$ be k-dimensional subspaces and let $X \in \mathcal{H}(n)$ be such that $e^{iX}\mathcal{S} = \mathcal{T}$. If $U = e^{i\,Z} \in \mathcal{U}(n)$ is a direct rotation from $\mathcal{S}$ onto $\mathcal{T}$, for $Z \in \mathcal{H}(n)$, then*

$$s(Z) \prec_w \frac{1}{2}|\operatorname{Spr}(X)| \prec_w s(X)\,.$$

Proof. If $U \in \mathcal{U}(n)$ is a direct rotation from $\mathcal{S}$ onto $\mathcal{T}$ then, we have seen that $U = e^{i\,Z}$ for $Z \in \mathcal{H}(n)$ such that

$$\lambda(Z) = (\Theta(\mathcal{S}\,,\,\mathcal{T})^*, -\,\Theta(\mathcal{S}\,,\,\mathcal{T})^*, 0)\ , \qquad \text{where} \qquad \Theta(\mathcal{S}\,,\,\mathcal{T})^* = (\theta_i)_{i\in\mathbb{I}_r} \in \mathbb{R}^r$$

denotes the vector with the positive principal angles, for some $r \le [n/2]$. Hence, $s(Z) = (\Theta(\mathcal{S}\,,\,\mathcal{T})^*, \Theta(\mathcal{S}\,,\,\mathcal{T})^*, 0)^\downarrow \in (\mathbb{R}^n)^\downarrow$. On the other hand, by Theorem 2.15 we get that

$$s(Z) = (\Theta(\mathcal{S}\,,\,\mathcal{T})^*, \Theta(\mathcal{S}\,,\,\mathcal{T})^*, 0) \prec_w \frac{1}{2}(\operatorname{Spr}^+(X)\,,\,\operatorname{Spr}^+(X))^\downarrow = \frac{1}{2}|\operatorname{Spr}(X)|^\downarrow\,.$$

By Proposition 2.5 we get that $\frac{1}{2}|\operatorname{Spr}(X)| \prec_w s(X)$ and the result now follows from these two submajorization relations. □

Remark 2.17. Let $A = \begin{bmatrix} a_1 & b \\ \bar{b} & a_2 \end{bmatrix} \in \mathcal{H}(2)$. Then

$$\mathrm{Spr}^+(A) = \lambda_1(A) - \lambda_2(A) = \big[\,(\mathrm{tr}\ A)^2 - 4\ \det\ A\big]^{1/2} = \big(\,(a_1 - a_2)^2 + 4|b|^2\,\big)^{1/2} \in \mathbb{R}_{\geq 0}\ . \quad (38)$$

More generally, if we consider $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^k \end{matrix} \in \mathcal{H}(2k)$, it is natural to wonder whether an inequality of the form

$$\lambda\Big(\big[\,(A_1 - A_2)^2 + 4\,B^*B\big]^{1/2}\Big) \prec_w \mathrm{Spr}^+(A) \quad (39)$$

holds true. Notice that this inequality would be an improvement of both Theorems 2.7 and 2.12. Nevertheless, it turns out that Eq. (39) does not hold in general. For example,

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{matrix} \mathbb{C}^2 \\ \\ \mathbb{C}^2 \end{matrix} \qquad \text{has} \quad \mathrm{Spr}^+(A) = (6.2714\,,\ 1.6339)\ ,$$

and $\lambda\Big(\big[\,(A_1 - A_2)^2 + 4\,B^*B\big]^{1/2}\Big) = (4.7599\,,\ 3.3680)$. But

$$\mathrm{tr}\ \mathrm{Spr}^+(A) = 7.9053 < 8.1279 = \mathrm{tr}\Big(\big[\,(A_1 - A_2)^2 + 4\,B^*B\big]^{1/2}\Big)\ .$$

In what follows we state a weak version of Eq. (39) that holds true in the general case. △

Proposition 2.18. *Let* $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^k \end{matrix} \in \mathcal{H}(2k)$. *Then*

$$\lambda\big(\,(A_1 - A_2)^2 + 4\,\mathrm{Re}(B)^2\,\big) \prec_w \mathrm{Spr}^2(A)\ . \quad (40)$$

Proof. We conjugate by the unitary matrix $U = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \mathcal{U}(2k)$,

$$U^*AU = \begin{bmatrix} A_2 & -B^* \\ -B & A_1 \end{bmatrix} \quad \text{and get} \quad D \overset{\text{def}}{=} A - U^*AU = \begin{bmatrix} A_1 - A_2 & 2\,\mathrm{Re}(B) \\ 2\,\mathrm{Re}(B) & A_2 - A_1 \end{bmatrix}\ .$$

By Weyl's inequality (item 1 in Theorem 4.2),

$$\lambda(D) \prec \lambda(A) + \lambda(-A) = (\mathrm{Spr}(A)\,,\ -\mathrm{Spr}^{\uparrow}(A)\,) \in (\mathbb{R}^{2k}_{\geq 0})^{\downarrow}\ .$$

Note that

$$D^2 = \begin{bmatrix} (A_1 - A_2)^2 + 4\,\mathrm{Re}(B)^2 & * \\ * & (A_1 - A_2)^2 + 4\,\mathrm{Re}(B)^2 \end{bmatrix}\ .$$

If $E \overset{\text{def}}{=} (A_1 - A_2)^2 + 4\mathrm{Re}(B)^2$, then using Remark 4.5 and Theorem 4.2 (see the Appendix),

$$(\lambda(E)\,,\ \lambda(E)\,)^{\downarrow} \overset{(60)}{\prec} \lambda(D^2) \prec_w (\mathrm{Spr}^+(A)^2, \mathrm{Spr}^+(A)^2)\,, \quad (41)$$

since $t \mapsto t^2$ is a convex map. Clearly, Eq. (40) follows from Eq. (41). □

We point out that Proposition 2.18 does not imply neither Theorem 2.7 nor Theorem 2.12 (since the relation $\prec_w$ is not preserved by taking square roots).

3 Reformulations of the key inequality

In what follows we obtain a series of results that are consequences of the key inequality (Theorem 2.7). Then we show that the key inequality is actually equivalent to several of these derived inequalities (see Theorem 3.7 below).

3.1 Generalized commutators and unitary conjugates

We begin this section with the following inequality for singular values of generalized commutators in terms of the spectral spread.

Theorem 3.1. *Let* $A_1\,,\,A_2\in\mathcal{H}(n)$ *and* $X\in\mathcal{M}_n(\mathbb{C})$*. Then*

$$s(A_1\,X-X A_2)\prec_w s(X)\ \mathrm{Spr}^+(A_1\oplus A_2)\,. \tag{42}$$

Proof. Consider the matrices $C=A_1X-XA_2\in\mathcal{M}_n(\mathbb{C})$,

$$B=A_1\oplus A_2=\begin{bmatrix} A_1 & 0\\ 0 & A_2\end{bmatrix}\in\mathcal{H}(2n)\quad\text{ and }\quad Z=\begin{bmatrix} 0 & X\\ -X^* & 0\end{bmatrix}\in i\,\mathcal{H}(2n)\ .$$

Note that

$$B\ Z-Z\ B=\begin{bmatrix} 0 & A_1X-XA_2\\ X^*A_1-A_2X^* & 0\end{bmatrix}=\begin{bmatrix} 0 & C\\ C^* & 0\end{bmatrix}\in\mathcal{H}(2n)$$

Therefore, by Proposition 4.7 we get that

$$\lambda\Big(B\ Z-Z\ B\Big)=\Big(s(C)\,,\,-s(C)\Big)^{\downarrow}\ .$$

Fix $1\le k\le n$. Then, the previous identity shows that

$$\sum_{j=1}^k s_j(A_1X-XA_2)=\sum_{j=1}^k\lambda_j^{\downarrow}\Big(B\ Z-Z\ B\Big)\ .$$

Then, there exists a projection $P\in\mathcal{H}(2n)$ with $\mathrm{rk}\,P=k$ such that

$$\sum_{j=1}^k s_j(A_1X-XA_2)=\mathrm{tr}\,\Big(\,\big[\,B\ Z-Z\ B\,\big]\ P\Big)\ . \tag{43}$$

Using that $\mathrm{tr}\,\Big(\,\big[\,B\ Z-Z\ B\,\big]\ P\Big)\ =\mathrm{tr}\,\Big(\,\big[\,P\ B-B\ P\,\big]\ Z\Big)$, and that

$$|\,\mathrm{tr}(VW)|\le\mathrm{tr}\,\big(\,s(VW)\big)\ \le\mathrm{tr}\,\big(\,s(V)\ s(W)\big)$$

for every $V\,,\,W\in\mathcal{M}_{2n}(\mathbb{C})$, from Eq. (43) we get

$$\sum_{j=1}^k s_j(A_1X-XA_2)\le\mathrm{tr}\,\big(\,s(P\ B-B\ P\,)\ s(Z)\big)\ \ .$$

Now, notice that if we consider the block matrix representation (obtained by changing the orthogonal decomposition of $\mathbb{C}^{2n}$)

$$P=\begin{bmatrix} I & 0\\ 0 & 0\end{bmatrix}\begin{matrix} k\\ 2n-k\end{matrix}\qquad\text{and}\qquad B=A_1\oplus A_2=\begin{bmatrix} S_{11} & S_{12}\\ S_{12}^* & S_{22}\end{bmatrix}$$

then, using these new representations we have that

$$P\ B - B\ P = \begin{bmatrix} 0 & S_{12} \\ -S_{12}^* & 0 \end{bmatrix} .$$

Hence $s\big(PB - BP\big) = \big(s(S_{12}^*)\,,\, s(S_{12})\big)^\downarrow$. Thus, by Theorem 2.7,

$$s(S_{12}^*) \prec_w \frac{1}{2}\operatorname{Spr}^+ B \implies s\big(PB-BP\big) \prec_w \frac{1}{2}\big(\operatorname{Spr}^+ B\,,\,\operatorname{Spr}^+ B\big)\ . \tag{44}$$

Using these facts, that $s(Z) = \big(s(X)\,,\, s(X)\big)^\downarrow$ and item 5 in Lemma 4.3, we can now see that

$$\begin{aligned} s\big(PB-BP\big)\ s(Z) &\prec_w \tfrac{1}{2}\ \big(\operatorname{Spr}^+ B\,,\,\operatorname{Spr}^+ B\big)^\downarrow\ s(Z) \\ &= \tfrac{1}{2}\ \big(\operatorname{Spr}^+ B\ s(X)\,,\,\operatorname{Spr}^+ B\ \ s(X)\big)^\downarrow \stackrel{\text{def}}{=} \rho\ . \end{aligned} \tag{45}$$

Note that $\operatorname{rk} P = k \implies \operatorname{rk}(PB-BP) \le 2k \implies s_{2k+1}\big(PB-BP\big) = 0$. Then

$$\begin{aligned} \sum_{j=1}^{k} s_j(A_1X - XA_2) &\le \operatorname{tr}\big(s\big(PB-BP\big)\ s(Z)\big) \\ &= \sum_{j=1}^{2k} s_j\big(PB-BP\big)\,s_j(Z) \\ &\stackrel{(45)}{\le} \sum_{j=1}^{2k}\rho_j = \sum_{j=1}^{k} s_j(X)\ \operatorname{Spr}_j^+(A_1\oplus A_2)\ , \end{aligned}$$

for every $k \in \mathbb{I}_n$. This shows Eq. (42). □

Remark 3.2. By inspection of the proof of Theorem 3.1, we see that this result follows from the application of Theorem 2.7 in Eq. (44). On the other hand, Theorem 3.1 extends Theorem 2.12 (by taking $X = I$ in Eq. (42)); in particular, the inequality in Eq. (42) is sharp (see the end of Remark 2.13). In the next section we will see that all these results are actually equivalent. △

Corollary 3.3. *Let $A\,,\,X \in \mathcal{H}(n)$. Then*

$$s(AX - XA) \prec_w \operatorname{Spr}^+(X)\ \operatorname{Spr}^+(A\oplus A)\ . \tag{46}$$

If $X \in \mathcal{M}_n(\mathbb{C})^+$ then also

$$s(AX - XA) \prec_w \frac{1}{2}\,\|X\|\operatorname{Spr}^+(A\oplus A)\ . \tag{47}$$

Proof. If we let $Y = X - \lambda I \in \mathcal{H}(n)$ for some $\lambda \in \mathbb{R}$, then $AY - YA = AX - XA$. Hence, by Theorem 3.1 we see that

$$s(AX - XA) \prec_w s(X - \lambda I)\ \operatorname{Spr}^+(A\oplus A)\,.$$

Take $\lambda = \lambda_{k+1}(X)$, for $k = [\frac{n}{2}]$. By Eq. (16) we get that $s(X-\lambda) \prec_w \operatorname{Spr}^+(X)$, and Eq. (46) follows from these facts together with item 5 in Lemma 4.3. If $X \in \mathcal{M}_n(\mathbb{C})^+$ we take $\lambda = \frac{\|X\|}{2}$, so we have that $s(Y) \leqslant \|Y\|\mathbb{1} = \frac{\|X\|}{2}\,\mathbb{1}$ and hence $s(Y) \prec_w \frac{\|X\|}{2}\,\mathbb{1}$. This fact together with item 5 in Lemma 4.3 imply Eq. (47) above. □

Corollary 3.4. *Let A , $B \in \mathcal{M}_n(\mathbb{C})^+$ and $X \in \mathcal{M}_n(\mathbb{C})$. Then*

$$s(AX - XB) \prec_w s(X)\ s(A \oplus B)\,. \tag{48}$$

In particular we get that, for every unitary invariant norm N,

$$N(AX - XB) \leq \|X\|\, N\,(A \oplus B)\ . \tag{49}$$

Proof. Use Theorem 3.1 and recall that if $C \in \mathcal{M}_{2n}(\mathbb{C})^+$, then $\operatorname{Spr}(C) \leqslant \lambda(C) = s(C)$. □

Let A, $X \in \mathcal{H}(n)$ and let $U := e^{i\,X} \in \mathcal{U}(n)$. As mentioned in item 7 in Remark 2.4 (or as a consequence of Eq. (29) in Theorem 2.12) we get that

$$s(A - UAU^*) \prec_w \operatorname{Spr}^+(A \oplus UAU^*) = \operatorname{Spr}^+(A \oplus A) = |\operatorname{Spr}(A)|\,.$$

Nevertheless, in case X is close to a (real) multiple of the identity then $U = e^{i\,X}$ is close to a multiple of the identity as well; hence, in this case we would expect A and U^*AU to be close, too. Similarly, in case A is close to a (real) multiple of the identity, then we would expect A and U^*AU to be close. The following result provides quantitative estimates that deal with these situations.

Theorem 3.5. Let A, $X \in \mathcal{H}(n)$ and $U = e^{i\,X} \in \mathcal{U}(n)$. Then

$$s(A - U^*AU) \prec_w s(X)\ \operatorname{Spr}^+(A \oplus A) \quad \text{and} \quad s(A - U^*AU) \prec_w s(A)\ \operatorname{Spr}^+(X \oplus X)\,.$$

Proof. We prove the first inequality; the proof of the second inequality is similar (it changes only after Eq. (52)) and the details are left to the reader. Let $U(\cdot) : [0,1] \to \mathcal{H}(n)$ be the smooth function given by $A(t) = e^{-i\,t\,X}\, A\, e^{i\,t\,X}$, for $t \in [0,1]$. Notice that $A(0) = A$ and $A(1) = U^*A\,U$; using Weyl's inequality for singular values (item 1 in Theorem 4.2)

$$s(A - U^*AU) \prec_w \sum_{j=0}^{m-1} s\big(\,A(\tfrac{j}{m}) - A(\tfrac{j+1}{m})\,\big) \quad \text{for every} \quad m \in \mathbb{N}\,. \tag{50}$$

Notice that $A(t+h) = e^{-i\,t\,X}\, A(h)\, e^{i\,t\,X}$ with $e^{i\,t\,X} \in \mathcal{U}(n)$, for $t,\ h,\ t+h \in [0,1]$. Thus

$$s(A(\tfrac{j}{m}) - A(\tfrac{j+1}{m})) \ = s(A - A(\tfrac{1}{m})) \quad \text{for} \quad j \in \mathbb{I}_{m-1}\ . \tag{51}$$

Since $A - A(0) = 0$ and $\frac{d}{dt}A(t)|_{t=0} = i\,(AX - XA)$ we get that

$$s(A - A(\tfrac{1}{m})) \ = \tfrac{1}{m}\, s(AX - XA) + O(m) \quad \text{with} \quad \lim_{m\to\infty} m\, O(m) = 0\,. \tag{52}$$

Hence, by Theorem 3.1 we have that

$$s(A - A(\tfrac{1}{m})) \ \prec_w \tfrac{1}{m}\, s(X)\ \ \operatorname{Spr}^+(A \oplus A) + O(m)\,. \tag{53}$$

Therefore, by Eq.'s (50) and (51) we have that, for sufficiently large m,

$$s(A - U^*AU) \prec_w s(X)\ \ \operatorname{Spr}^+(A \oplus A) + m\, O(m)\,.$$

The statement now follows by taking the limit $m \to \infty$ in the expression above. □

Remark 3.6. By inspection of the proof of Theorem 3.5, we see that this result follows from the application of Theorem 3.1 in Eq. (53).

3.2 Equivalences of the inequalities

In this section we show the equivalence of several of the main results obtained in Sections 2.3 and 3.1.

Theorem 3.7. *The following inequalities are equivalent:*

1. *Given* $A_1,\, A_2 \in \mathcal{H}(n)$ *then*
$$s(A_1 - A_2) \prec_w \operatorname{Spr}^+(A_1 \oplus A_2)\ . \tag{54}$$

2. *Given* $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^n \\ \mathbb{C}^n \end{matrix} \in \mathcal{H}(2n)$ *then*
$$2\, s(B) \prec_w \operatorname{Spr}^+(A)\ . \tag{55}$$

3. *Given* $A_1\,,\, A_2 \in \mathcal{H}(n)$ *and* $X \in \mathcal{M}_n(\mathbb{C})$ *then*
$$s(A_1\, X - X A_2) \prec_w s(X)\ \operatorname{Spr}^+(A_1 \oplus A_2)\,. \tag{56}$$

4. *Given* $A,\, X \in \mathcal{H}(n)$ *and* $U = e^{i\,X} \in \mathcal{U}(n)$ *then*
$$s(A - U^*AU) \prec_w s(X)\ \operatorname{Spr}^+(A \oplus A)\,. \tag{57}$$

5. *Given subspaces* $\mathcal{S}, \mathcal{T} \subseteq \mathbb{C}^n$ *and* $X \in \mathcal{H}(n)$ *such that* $e^{iX}\mathcal{S} = \mathcal{T}$, *then*
$$\Theta(\mathcal{S}\,,\,\mathcal{T}) \prec_w \frac{1}{2}\operatorname{Spr}^+(X)\ . \tag{58}$$

Proof. $1 \implies 2$: Let $B = U|B|$ be the polar decomposition of B, with $U \in \mathcal{U}(n)$. If
$$W = \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \in \mathcal{U}(2n) \quad \text{then} \quad W^*AW = \begin{bmatrix} U^*A_1U & U^*B \\ B^*U & A_2 \end{bmatrix} = \begin{bmatrix} U^*A_1U & |B| \\ |B| & A_2 \end{bmatrix}\,.$$
Since $\operatorname{Spr}(A) = \operatorname{Spr}(W^*AW)$ and $s(B) = s(|B|)$, in order to show Eq. (55) we can assume that $B \in \mathcal{H}(n)$. In this case, taking $R = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \in \mathcal{U}(2n) \cap \mathcal{H}(2n)$, we have that
$$RAR = \begin{bmatrix} A_2 & B \\ B & A_1 \end{bmatrix} \implies \frac{A + RAR}{2} = \begin{bmatrix} \frac{A_1+A_2}{2} & B \\ B & \frac{A_1+A_2}{2} \end{bmatrix}\,.$$
By Weyl inequality (item 1 in Theorem 4.2),
$$\lambda\left(\frac{A+RAR}{2}\right) \prec \frac{\lambda(A) + \lambda(RAR)}{2} = \lambda(A) \overset{(17)}{\implies} \operatorname{Spr}^+(\frac{A+RAR}{2}) \prec_w \operatorname{Spr}^+(A)\,.$$
Therefore, in order to show Eq. (55) we can assume that $B \in \mathcal{H}(n)$ and $A_1 = A_2$.

Take now $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \in \mathcal{U}(2n)$. Since now $A = \begin{bmatrix} A_1 & B \\ B & A_1 \end{bmatrix}$, then
$$Z^*AZ = \begin{bmatrix} A_1 - B & 0 \\ 0 & A_1 + B \end{bmatrix} \quad \text{and} \quad \operatorname{Spr}^+(Z^*AZ) = \operatorname{Spr}^+(A)\ . \tag{59}$$

Hence $2s(B) = s\big([A_1 - B] - [A_1 + B]\big) \overset{(54)}{\prec_w} \mathrm{Spr}^+\big([A_1 - B] \oplus [A_1 + B]\big) \overset{(59)}{=} \mathrm{Spr}^+(A)$.

We have already shown (see Remarks 3.2 and 3.6) that $2 \implies 3 \implies 4$.

$4 \implies 3$: We first consider the case $A_1 = A_2 = A$ and $X \in \mathcal{H}(n)$. In this case we consider $D(t) = A - e^{-itX} A\, e^{itX}$, for $t \in [0,1]$. Then, $D(\cdot)$ is a smooth function such that $D(0) = 0$ and $D'(0) = -i\,(AX - XA)$. Hence, by Weyl's inequality, we have that

$$s(t\,(AX - XA)) = s(A - e^{-itX} A\, e^{itX}) + O(t) \quad \text{with} \quad \lim_{t \to 0^+} \frac{O(t)}{t} = 0\,.$$

Using Eq. (57) we now see that

$$s(AX - XA) \prec_w s(X)\, \mathrm{Spr}^+(A \oplus A) + \frac{O(t)}{t}\,.$$

Then 3. follows by taking the limit $t \to 0^+$, when $A_1 = A_2$ and $X \in \mathcal{H}(n)$. For the general case, consider

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \in \mathcal{H}(2n) \quad \text{and} \quad \hat{X} = \begin{bmatrix} 0 & X \\ X^* & 0 \end{bmatrix} \in \mathcal{H}(2n)\,.$$

Notice that

$$A\,\hat{X} - \hat{X}\,A = \begin{bmatrix} 0 & A_1X - XA_2 \\ -(A_1X - XA_2)^* & 0 \end{bmatrix}$$

Hence, by Proposition 4.7 and the previous facts,

$$s(A\,\hat{X} - \hat{X}\,A) = (s(A_1X - XA_2)\,,\, s(A_1X - XA_2))^{\downarrow} \prec_w s(\hat{X})\, \mathrm{Spr}^+(A \oplus A)\,.$$

Notice that Eq. (56) follows from the previous submajorization relation, since

$$s(\hat{X}) = (s(X)\,,\, s(X))^{\downarrow} \quad \text{and} \quad \mathrm{Spr}^+(A \oplus A) = (\mathrm{Spr}^+(A_1 \oplus A_2)\,,\, \mathrm{Spr}^+(A_1 \oplus A_2))^{\downarrow}\,.$$

Since $3 \implies 1$ (by taking $X = I$) we see that $1 - 4$ are equivalent. We have shown (in the proof of Theorem 2.15) that $2 \implies 5$. Thus, we are left to show that $5 \implies 2$. Indeed, fix $A \in \mathcal{H}(2n)$ as in 2. Let $\mathcal{S} = \mathbb{C}^n \oplus 0 \subset \mathbb{C}^n \oplus \mathbb{C}^n$ and let $\mathcal{T}(t) = e^{itA}\mathcal{S}$, for $t \in [0,1]$. Let

$$S = \begin{bmatrix} I_n \\ 0 \end{bmatrix}\,,\ S_\perp = \begin{bmatrix} 0 \\ I_n \end{bmatrix} \in M_{2n,n}(\mathbb{C}) \quad \text{and} \quad T(t) = e^{itA} S \quad \text{for} \quad t \in [0,1]\,.$$

Hence, by Eq. (30), in this case we have that

$$\sin(\Theta(\mathcal{S}\,,\, \mathcal{T}(t))) = s(T(t)^*\, S_\perp) = s(S\, e^{itA}\, S_\perp)\,.$$

Since $T^*(0)\, S_\perp = 0$ and $\frac{d}{dt}(T^*(t)\, S_\perp)|_{t=0} = i\, S\, A\, S_\perp = i\, B \in \mathcal{M}_n(\mathbb{C})$ we get that

$$t\, s(B) = \sin(\Theta(\mathcal{S}\,,\, \mathcal{T}(t))) + O_1(t) \quad \text{with} \quad \lim_{t \to 0^+} \frac{O_1(t)}{t} = 0\,.$$

On the other hand, since $\Theta(\mathcal{S}\,,\, \mathcal{T}(t))$ is a continuous function of $t \in [0,1]$ such that $\Theta(\mathcal{S}\,,\, \mathcal{T}(0)) = 0$, then we have that

$$\sin(\Theta(\mathcal{S}\,,\, \mathcal{T}(t))) = \Theta(\mathcal{S}\,,\, \mathcal{T}(t)) + O_2(t) \quad \text{with} \quad \lim_{t \to 0^+} \frac{O_2(t)}{t} = 0\,.$$

The previous facts together with (58) in 5. show that

$$t\, s(B) \prec_w \frac{t}{2}\, \mathrm{Spr}^+(A) + O_1(t) + O_2(t) \implies 2\, s(B) \prec_w \mathrm{Spr}^+(A) + \frac{2}{t}\,(O_1(t) + O_2(t))\,.$$

Then 2. follows from the previous inequality, by taking the limit $t \to 0^+$. □

Remark 3.8. We point out that item 2 of Theorem 3.7 implies Theorem 2.7 in its general form (where the block B can be rectangular). This follows from Eq. (18) in Remark 2.4 (the details are left to the reader). △

3.3 Concluding remarks

We have developed several aspects of the spectral spread of Hermitian matrices introduced in [15], which is a natural vector valued measure of the dispersion of the spectra. We have also connected our work with well established research topics in matrix analysis. We have obtained several inequalities involving the spectral spread; in particular, we have obtained sharp inequalities for generalized commutators of the form $A_1X - XA_2$, for Hermitian matrices $A_1, A_2 \in \mathcal{H}(n)$ and arbitrary $X \in \mathcal{M}_n(\mathbb{C})$. Generalized commutators appear, in a natural way, as derivative vectors of matrix-valued smooth curves. We expect that our results will have applications in matrix perturbation bounds for Hermitian matrices, obtained from a (differential) geometrical perspective. Indeed, in [18] we have already applied the results herein and obtained some inequalities related to the bounds in Eq. (3) (see [15]) using a geometrical approach. We point out that matrix sensitivity problems related with the variation of the eigenvalues and eigenspaces of (perturbations of) Hermitian matrices have applications in Machine Learning (tracking changes in data, see [4]) while the bounds in the variation of angles between subspaces are of interest in Quantum Computing theory (see [19]).

Acknowledgments. We would like to thank Professor X. Zhan and the anonymous reviewers for several useful suggestions that improved the exposition of the results in this work.

4 Appendix

Here we collect several well known results about majorization, used throughout our work. For detailed proofs of these results and general references in majorization theory see [3, 9]. We begin with the Weyl's inequalities for singular values:

Theorem 4.1. Let $C, D \in \mathcal{M}_n(\mathbb{C})$. Then,

1. Weyl's additive inequality: $s(C + D) \prec_w s(C) + s(D)$;

2. Weyl's multiplicative inequality: $s(CD) \prec_w s(C)\, s(D)$. □

Theorem 4.2. Let $C, D \in \mathcal{H}(n)$. Then,

1. Weyl's additive inequalities (for eigenvalues):
   (a) $\lambda(C + D) \prec \lambda(C) + \lambda(D)$;
   (b) if $C \leq D$ then $\lambda(C) \leqslant \lambda(D)$;

2. Lidskii's additive inequality: $\lambda(C) - \lambda(D) \prec \lambda(C - D) \prec \lambda(C) - \lambda^{\uparrow}(D)$;

3. $|\lambda(C) - \lambda(D)| \prec_w s(C - D)$;

4. Let $\mathcal{P} = \{P_j\}_{j=1}^r$ be a system of projections (i.e. they are mutually orthogonal projections on $\mathbb{C}^d$ such that $\sum_{i=1}^r P_i = I$). If

$$\mathcal{C}_{\mathcal{P}}(D) \stackrel{\text{def}}{=} \sum_{i=1}^{r} P_i\, D\, P_i \implies \lambda(\mathcal{C}_{\mathcal{P}}(D)) \prec \lambda(D) \ . \tag{60}$$

□

In the next result we describe several elementary but useful properties of (sub)majorization between real vectors.

Lemma 4.3. Let $x\,,\,y\,,\,z\,,\,w \in \mathbb{R}^k$. Then,

1. $x^\downarrow + y^\uparrow \prec x + y \prec x^\downarrow + y^\downarrow$;
2. If $x \prec_w y$ and $y,\ z \in (\mathbb{R}^k)^\downarrow$ then $x + z \prec_w y + z$;
3. Moreover, if $z\,,\,w \in (\mathbb{R}^k)^\downarrow$, $x \prec z$ and $y \prec w$ then $x + y \prec z + w$.

If we assume further that $x\,,\,y\,,\,z \in \mathbb{R}_{\geq 0}^k$ then,

4. $x^\downarrow\, y^\uparrow \prec_w x\, y \prec_w x^\downarrow\, y^\downarrow$;
5. If $x \prec_w y$ and $y,\ z \in (\mathbb{R}_{\geq 0}^k)^\downarrow$ then $x\, z \prec_w y\, z$. □

Remark 4.4. Let $x, y \in \mathbb{R}^k$. If $x \leqslant y$ then,

$$x^\downarrow \leqslant y^\downarrow \quad \text{and} \quad x \prec_w y \, .$$

Recall that given $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}$ is an interval, and $z = (z_i)_{i\in\mathbb{I}_k} \in I^k$ we denote $f(z) = (f(z_i))_{i\in\mathbb{I}_k} \in \mathbb{R}^k$.

Remark 4.5. Let $I \subset \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a convex function. Then,

1. if $x,\ y \in I^k$ satisfy $x \prec y$ then $f(x) \prec_w f(y)$.
2. If $x \prec_w y$ but f is further non-decreasing in I, then $f(x) \prec_w f(y)$. △

Recall that a norm N in $\mathcal{M}_n(\mathbb{C})$ is unitarily invariant (briefly u.i.n.) if $N(UAV) = N(A)$, for every $A \in \mathcal{M}_n(\mathbb{C})$ and $U,\ V \in \mathcal{U}(n)$. Well known examples of u.i.n. are the spectral norm $\|\cdot\|_{sp}$ and the p-norms $\|\cdot\|_p$, for $p \geq 1$.

Remark 4.6. It is well known that (sub)majorization relations between singular values of matrices are intimately related with inequalities with respect to u.i.n's. Indeed, given $A,\ B \in \mathcal{M}_n(\mathbb{C})$ the following statements are equivalent:

1. For every u.i.n. N in $\mathcal{M}_n(\mathbb{C})$ we have that $N(A) \leq N(B)$.
2. $s(A) \prec_w s(B)$. △

Proposition 4.7. Let $1 \leq k < n$, $E \in \mathcal{M}_{k\,,\,n-k}(\mathbb{C})$ and $\hat{E} = \begin{pmatrix} 0 & E \\ E^* & 0 \end{pmatrix} \in \mathcal{H}(n)$. Then

$$\lambda(\hat{E}) = \big(\, s(E)\,,\, -s(E^*)\,\big)^\downarrow = \big(\, s(E)\,,\, -s^\uparrow(E)\,\big) \in (\mathbb{R}_{\geq 0}^n)^\downarrow \ ,$$

with some zeros at the middle, in the rectangular case. □

References

[1] M.E. Argentati, A.V. Knyazev, C.C. Paige, I. Panayotov, Bounds on changes in Ritz values for a perturbed invariant subspace of a Hermitian matrix. SIAM J. Matrix Anal. Appl. 30 (2008), no. 2, 548-559.

[2] W. Audeh, F. Kittaneh, Singular value inequalities for compact operators. Linear Algebra Appl. 437 (2012), no. 10, 2516-2522.

[3] R. Bhatia, Matrix analysis, 169, Springer-Verlag, New York, 1997.

[4] C. M. Bishop, Pattern recognition and machine learning. Information Science and Statistics. Springer, New York, 2006.

[5] N. Bosner, Z. Drmač, Subspace gap residuals for Rayleigh-Ritz approximations. SIAM J. Matrix Anal. Appl. 31 (2009), no. 1, 54-67.

[6] C. Davis, W.M. Kahan, The rotation of eigenvectors by a perturbation. III. SIAM J. Numer. Anal. 7 1970 1-46.

[7] O. Hirzallah, Commutator inequalities for Hilbert space operators. Linear Algebra Appl. 431 (2009), no. 9, 1571-1578.

[8] O. Hirzallah, F. Kittaneh, Singular values, norms, and commutators. Linear Algebra Appl. 432 (2010), no. 5, 1322-1336.

[9] R.A. Horn, C.R: and Johnson, Matrix analysis. Second Edition, Cambridge University Press, Cambridge, 2012.

[10] F. Kittaneh, Inequalities for commutators of positive operators. J. Funct. Anal. 250 (2007), no. 1, 132-143.

[11] F. Kittaneh, Norm inequalities for commutators of Hermitian operators. Integral Equations Operator Theory 62 (2008), no. 1, 129-135.

[12] F. Kittaneh, Singular value inequalities for commutators of Hilbert space operators. Linear Algebra Appl. 430 (2009), no. 8-9, 2362-2367.

[13] A.V. Knyazev, M.E. Argentati, Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra. SIAM J. Matrix Anal. Appl. 29 (2006/07), no. 1, 15-32.

[14] A.V. Knyazev, M.E. Argentati, On proximity of Rayleigh quotients for different vectors and Ritz values generated by different trial subspaces. Linear Algebra Appl. 415 (2006), no. 1, 82-95.

[15] A.V. Knyazev, M.E. Argentati, Rayleigh-Ritz majorization error bounds with applications to FEM. SIAM J. Matrix Anal. Appl. 31 (2009), no. 3, 1521-1537.

[16] C.-K. Li, R.-C. Li, A note on eigenvalues of perturbed Hermitian matrices. Linear Algebra Appl. 395 (2005), 183-190.

[17] P. Massey, D. Stojanoff, S. Zarate, Majorization bounds for Ritz values of Hermitian matrices. SIAM J. Matrix Anal. Appl. 41 (2020), no. 2, 554-572.

[18] P. Massey, D. Stojanoff and S.Zarate, Absolute variation of Ritz values, principal angles and spectral spread (available at arXiv:2012.09018 [math.FA]).

[19] M. A. Nielsen, I.L. Chuang, Quantum computation and quantum information. Cambridge University Press, Cambridge, 2000.

[20] B.N. Parlett, The symmetric eigenvalue problem. Corrected reprint of the 1980 original, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[21] L. Qiu, Y. Zhang, C.K. Li, Unitarily invariant metrics on the Grassmann space. SIAM J. Matrix Anal. Appl. 27 (2005), no. 2, 507-531.

[22] Y. Saad, Numerical Methods for Large Eigenvalue Problems, Revised Edition, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[23] Y. Tao, More results on singular value inequalities of matrices, Linear Algebra Appl. 416 (2006) 724-729.

[24] Z. Teng, L. Lu, R.-C. Li, Cluster-robust accuracy bounds for Ritz subspaces. Linear Algebra Appl. 480 (2015), 11-26.

[25] Y.Q. Wang, H.K. Du, Norms of commutators of Hermitian operators. J. Math. Anal. Appl. 342 (2008), no. 1, 747-751.

[26] X. Zhan, Singular values of differences of positive semidefinite matrices. SIAM J. Matrix Anal. Appl. 22 (2000), no. 3, 819-823.

[27] X. Zhan, Matrix inequalities. Lecture Notes in Mathematics, 1790. Springer-Verlag, Berlin, 2002.

[28] X. Zhan, On some matrix inequalities, Linear Algebra and its Applications 376 (2004) 299-303.

[29] P. Zhu, M.E. Argentati, A.V. Knyazev, Bounds for the Rayleigh quotient and the spectrum of Hermitian operators. SIAM J. Matrix Anal. Appl. 34 (2013), no. 1, 244-256.

[30] P. Zhu, A.K. Knyazev, Rayleigh-Ritz majorization error bounds of mixed type. SIAM J. Matrix Anal. Appl. 38 (2017), no. 1, 30-49.