Introducing Discrepancy Values of Matrices and their Application to Bounding Norms of Commutators

Pourya Habib Zadeh  
Suvrit Sra  
Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

We introduce discrepancy values, quantities that are inspired by the notion of spectral spread of Hermitian matrices. In particular, the discrepancy values capture the difference between two consecutive (Ky-Fan-like) pseudo-norms that we also introduce. As a result, discrepancy values share many properties with singular values and eigenvalues, and yet are substantially different to merit their own study. We describe several key properties of discrepancy values and establish a set of useful tools (e.g., representation theorems, majorization inequalities, convex optimization formulations, etc.) for working with them. As an important application, we illustrate the role of discrepancy values in deriving tight bounds on the norms of commutators.

1 Introduction

There exist many useful objects that quantify properties of linear transformations, each of which extracts certain aspects of a linear map. To name two such notions, we can point out eigenvalues and singular values of linear operators. Eigenvectors are vectors whose direction is invariant under the linear transformation, and eigenvalues are the corresponding scaling of those vectors under the transformation. There are many possible definitions for singular values; however, they can be best described geometrically. Consider a unit sphere, a (full rank) linear map would transform the sphere into an ellipsoid. The length of the semi-axes of this ellipsoid is the non-zero singular values, and the directions of the semi-axes are the singular vectors. The eigenvalue and singular values are extremely useful tools in linear algebra and are prevalent in engineering and science [10].

The spectral spread of Hermitian matrices is another fundamental example. It was first introduced in [12] as a measure of the dispersion of the eigenvalues of Hermitian matrices. For an $n \times n$ Hermitian matrix $A$, this quantity can be represented as follows:

$$\text{spr}^+(A) = (\lambda_i(A) - \lambda_{n-i+1}(A))_{i=1,\ldots,\lfloor n/2 \rfloor},$$

where we assume that the eigenvalues $\lambda_i$ are ordered decreasingly. This measure can be viewed as a vector extension of the spread of the Hermitian operators [5]; i.e., $\text{spd}(A) := \lambda_{\max} - \lambda_{\min}$. It has been shown that the spectral spread of Hermitian matrices satisfies many useful inequalities [14, 15, 16]. However, for arbitrary square matrices, there is no similar quantity that satisfies the same inequalities.

We introduce a new quantity defined for the square matrices, which we call discrepancy values. The name arises because we define these values via differences between certain Ky-Fan-like consecutive pseudo-norms. We will show that discrepancy values are essentially equivalent to (two copies of) the spectral spread for Hermitian matrices. We will also show that they can be thought of as the vector measure of a linear operator to the class of scalar operators, a topic that has been studied in the literature—see e.g., [6, 17, 2].
After defining discrepancy values, we will prove several important theorems about them, including the connection that they have with singular values, Ky-Fan norms, and the principal angles between subspaces. We will also discover the invariances they satisfy, as well as several majorization inequalities. We will see that many important results about the spectral spread also hold for discrepancy values. We then employ the discrepancy values and the set of tools that we develop to prove a collection of other results, of which the most important consequence is a tight majorization inequality satisfied by the singular values of the commutator of two square matrices.

Commutators are pivotal objects in subjects such as Lie algebra and group theory. Algebraically, the commutator of two matrices \( A \) and \( B \) is defined using the formula \( [A, B] = AB - BA \). It can be viewed as a quantity that measures how much two operators are (non-)commuting. Finding bounds on the norms of such quantities has been investigated extensively in the literature \([11, 7, 9, 13, 19, 18, 17]\). As we will show in the paper, we can obtain sharp bounds on the norm of commutators using discrepancy values, and find when two Hermitian matrices with fixed eigenvalues are maximally non-commutative.

Moreover, we propose approaches for calculating discrepancy values; in particular, we formulate the computation of the discrepancy values as semi-definite programming problems. Finally, we discuss that our result can be easily be extended to the case of compact operators on Hilbert space. We conclude by conjecturing a stronger statement regarding the majorization inequality that we prove for commutators.

**Contributions** The main contributions of this paper are as follows:

- Introducing the notion of discrepancy values as the generalization of spectral spread of Hermitian matrices to any square matrices (Section 3).
- Proving several important results about this object such as the equivalence of two representation of the discrepancy values and many majorization bounds (Section 4).
- Using this new notion to come up with tight bounds for the commutators and consequently finding when two Hermitian matrices with fixed eigenvalues become maximally non-commutative. (Section 5).
- Providing methods to efficiently compute the discrepancy values of matrices (Section 6).

**2 Notation and Preliminaries**

Let \( \sigma(A) \) and \( \lambda(A) \) respectively denote the vector of singular values and eigenvalues of the \( m \times n \) matrix \( A \) with complex-valued entries, denoted by \( A \in M_{m,n}(\mathbb{C}) \). We assume that the singular values of any matrix and the eigenvalues of Hermitian matrices are ordered in a decreasing manner. We use \( I_n \) and \( J_n \) to denote \( n \times n \) identity and exchange matrix: a square matrix with ones on the antidiagonal and zeros elsewhere. The notation \( 1_k \) denotes a vector with \( k \) ones, followed by \( n-k \) zeros. Given \( x \in \mathbb{R}^n \), we use \( x^\downarrow \) to denote a vector which is arranged in a nonincreasing order; i.e. \( x_i^\downarrow \geq x_j^\downarrow \) for \( 1 \leq i \leq j \leq n \). The Schatten norm of the matrix \( A \in M_{m,n}(\mathbb{C}) \) is defined by \( \|A\|_p = \left( \sum_{i=1}^n \sigma_i^p \right)^{1/p} \), for \( p \geq 1 \). The Ky-Fan norm of the matrix \( A \in M_{m,n}(\mathbb{C}) \) is defined by \( \|A\|_{(k)} = \sum_{i=1}^k \sigma_i^\downarrow \), for \( k = 1, \ldots, n \). It is shown that the Ky-Fan norms has the following maximal description:

\[
\|A\|_{(k)} = \max_{\{x_j\}_{j=1}^k \text{ o.n.}, \{y_j\}_{j=1}^k \text{ o.n.}} \left| \sum_{i=1}^k (Ax_i, y_i) \right|,
\]

where o.n. means the set of vectors \( \{x_j\}_{j=1}^k \) is orthonormal.
The matrix $A \in M_{n,r}(\mathbb{C})$ is called isometry if $A^* A = I_r$ for some $r \leq n$, unitary if $n = r$, and partial isometry of order $k \leq n = r$ if $A = BC^*$ where $B \in M_{n,k}(\mathbb{C})$ and $C \in M_{n,k}(\mathbb{C})$ are both isometry. A norm $\| \cdot \|$ on $M_n$\(^1\) is unitarily invariant if $\|A\| = \|UAV\|$ for all $A \in M_n$ and all unitary matrices $U, V \in M_n$. In the sequel, we use $\| \cdot \|$ to denote such norms.

Given $x, y \in \mathbb{R}^n$, we say that $x$ is weakly majorized by $y$, denoted by $x \prec_w y$ when the following set of inequalities hold:

$$\sum_{i=1}^k y_i^\frac{1}{k} \leq \sum_{i=1}^k x_i^\frac{1}{k},$$

for $k = 1, \cdots, n$. Moreover, $y$ majorizes $x$ i.e. $x \prec y$, when the additional condition $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ also holds. Many results in matrix analysis are stated using the language of majorization; for instance $k$-itudinal vectors $\{\delta_i\}_{i=0}^n$ (assuming that pseudo-norms $\| \cdot \|$ of inequalities hold:

$$\|A + B\| \prec_w \|A\| + \|B\|,$$

$$\|AB\| \prec_w \|A\|\|B\|.$$

We refer the readers to [1, 3, 4, 8, 10, 20], for more information about the majorization inequalities, unitarily invariant norms, and variational characterization of Ky-Fan norms.

### 3 Definition of discrepancy values

Now we are ready to define the main object of this paper.

**Definition 3.1.** For the square matrix $A \in M_n$, let us define the $k\textsuperscript{th}$ discrepancy (pseudo-)norm by

$$\|A\|_D^k := \max_{\{x_i\}_{i=1}^k \text{ o.n.}, \{y_i\}_{i=1}^k \text{ o.n.}} \sum_{i=1}^k \langle A x_i, y_i \rangle,$$

(3.1)

for $k = 1, \ldots, n$, and let $\|A\|_D^0 = 0$. Now define

$$\delta_k(A) := \|A\|_D^k - \|A\|_D^{k-1}.$$

(3.2)

We call this quantity the $k\textsuperscript{th}$ discrepancy value of matrix $A$ as it is the discrepancy of two consecutive pseudo-norms.\(^2\). Let also $\delta(A) = (\delta_1(A), \ldots, \delta_n(A))$; ordered in a decreasing manner. Moreover, the set of vectors $x_i$ and $y_i$ in the definition of $\|A\|_D^k$ are called the left and right discrepancy vectors of matrix $A$ corresponding to the $k\textsuperscript{th}$ discrepancy norm.

Note that the optimization problem in formula 3.1 is always feasible and $\|A\|_D^k \leq \|A\|_D^k$ (equivalently $\delta(A) \prec_w \sigma(A)$); hence, $\|A\|_D^k$ is well-defined. We can see that $\sigma_1(A) \geq \delta_1(A) \geq \delta_2(A) \geq \cdots \geq \delta_n(A) \geq 0$, hence $\delta_i$ is always finite and unique, thus well-defined. In fact, we can define the singular values as $\sigma_i(A) = \|A\|_D^i - \|A\|_D^{i-1}$, and use the maximal representation of the Ky-Fan norm 2.1 (assuming that $\|A\|_D^{(0)} = 0$). Note that unlike the singular values, there do not exist two sets of orthonormal vectors $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ with the property $x_i \perp y_i$, for all $i = 1, \ldots, n$, such that $\delta_i(A) = \|A x_i, y_i\|$. Now let us propose more compact forms for the discrepancy norms and Ky-Fan norms.

---

\(^1\) $M_n$ is short for square matrices $M_{n,n}(\mathbb{C})$.

\(^2\) Equivalently we can define $\delta_i(A)$ first and then say $\|A\|_D^k = \sum_{i=1}^k \delta_i(A)$. 

3
Proposition 3.2. For arbitrary $n \times n$ matrix $A$, we have the following formulas:

$$\|A\|_{(k)} = \max_{M \in P_k(n)} \mathrm{Re} \left( \text{tr}(AM) \right),$$

(3.3)

$$\|A\|_{(k)}^\delta = \max_{M \in P_k^0(n)} \mathrm{Re} \left( \text{tr}(AM) \right),$$

(3.4)

where $P_k(n) := \{X \in M_n : X = VU^*, U^*U = I_k, V^*V = I_k\}$ denotes the set of $n \times n$ partial isometries with rank $k$ and $P_k^0(n) := \{X \in M_n : X = VU^*, U^*U = I_k, V^*V = I_k, \text{tr}(X) = 0\}$ denotes the set of $n \times n$ traceless (i.e. commutators) partial isometries with rank $k$.

Proof. Given a positive integer $k \leq n$, the sets $\{x_i^*Ay_i \mid \{x_j\}_{j=1}^k \text{ o.n.}, \{y_j\}_{j=1}^k \text{ o.n.}, \sum_{j=1}^k \langle x_j, y_j \rangle = 0\}$, for any $i = 1, \ldots, n$, form circles in the complex plane whose centers are located at the origin. Therefore, we can replace the modulus in the definitions of $\| \cdot \|_{(k)}$ and $\| \cdot \|_{(k)}^\delta$ with the real part of the complex number. Thus

$$\|A\|_{(k)} = \max_{\|x_i\|=\|y_i\|=1} \mathrm{Re} \left( \sum_{i=1}^k \langle Ax_i, y_i \rangle \right) = \max_{U^*U = I_k, V^*V = I_k} \mathrm{Re} \left( \text{tr}(AVU^*) \right),$$

$$\|A\|_{(k)}^\delta = \max_{\|x_i\|=\|y_i\|=1} \mathrm{Re} \left( \sum_{i=1}^k \langle Ax_i, y_i \rangle \right) = \max_{U^*U = I_k, V^*V = I_k, \text{tr}(U^*V) = 0} \mathrm{Re} \left( \text{tr}(AVU^*) \right).$$

Now let $M = VU^*$ and get the desired result. \hfill \Box

Using the definition of the discrepancy values, one can trivially observe the following set of invariances for these objects.

Proposition 3.3 (The invariances of $\delta$). Given the matrix $A \in M_n$, the discrepancy values are invariant under the following transformations:

- (Unitary similarity mapping) $\delta(UAU^*) = \delta(A)$, for all unitary matrix $U$.
- (Conjugate transpose) $\delta(A^*) = \delta(A)$.
- (Eigenvalue rotation) $\delta(e^{i\theta}A) = \delta(A)$, for all $\theta \in [0, 2\pi)$.
- (Eigenvalue shift invariance) $\delta(A - \alpha I_n) = \delta(A)$, for all $\alpha \in \mathbb{C}$.

4 Main results

In this section, we study the main properties of the discrepancy values, including a dual definition for the discrepancy norms, several majorization bounds, and the discrepancy values of the direct sums of matrices.

4.1 The relation between the discrepancy norms and Ky-Fan norms

Now we aim to obtain the connection between discrepancy values and the Ky-Fan norm. Before mentioning the theorem, we need the following lemma.
Lemma 1. For any two \( n \times m \) isometries \( U, V \), where \( n \geq m \), there exists an \( n \times n \) unitary matrix \( Q \) such that \( QU = V \).

Proof. This can be seen using the fact that for any isometry \( U \), there exists the appropriate unitary matrix \( R \) such that:

\[
U_{n \times m} = R_{n \times n} \begin{bmatrix}
I_m \\
0_{(n-m) \times m}
\end{bmatrix}.
\]

\[\square\]

Theorem 4.1. For the square matrix \( A \), we have

\[
\|A\|_{(k)} = \max_{Q \in U(n)} \|AQ\|_{(k)},
\]

where \( U(n) \) denotes the set of \( n \times n \) unitary matrices.

Proof. We want to show that \( \max_{Q \in U(n)} \sum_{i=1}^{k} \delta_i(AQ) = \|A\|_{(k)} \). Trivially, \( \delta(AQ) \preceq \sigma(A) \) for any unitary matrix \( Q \); i.e. \( \max_{Q \in U(n)} \sum_{i=1}^{k} \delta_i(AQ) \leq \|A\|_{(k)} \). Using the variational formula of the discrepancy norms and the fact that \( \delta(R^*AR) = \delta(A) \) for any unitary matrix \( R \), we can see that the RHS of the formula 4.1 is equal to

\[
\max_{Q,R \in U(n)} \max_{U^*U = I_k, V^*V = I_k} \text{Re} \left( (A(QV)(RU)^*) \right).
\]

Also, by the variational formula for the Ky-Fan norms, the LHS of 4.1 is

\[
\max_{M^*M = I_k, N^*N = I_k} \text{Re} \left( (ANM^*) \right).
\]

Assume that the optimum occurs at \( \hat{M} \) and \( \hat{N} \). By lemma 1, for any isometries \( U \) and \( V \) there exist \( Q, R \in U(n) \) such that \( \hat{N} = QV \) and \( \hat{M} = RU \). Hence, \( \max_{Q \in U(n)} \sum_{i=1}^{k} \delta_i(AQ) \geq \|A\|_{(k)} \).

\[\square\]

In the previous section, we give a maximal representation for the discrepancy norms. Now we want to give a minimal representation for the discrepancy norms. This dual representation reveals another aspect of the connection between the discrepancy and Ky-Fan norms. Before mentioning the relationship, let us state two results.

Lemma 2. The set \( P_k(n) \) is the set of extreme points of the compact convex set \( \text{Conv}(P_k(n)) := \{ X \in M_n : \|X\|_1 \leq 1, \|X\|_\infty \leq k \} \). Similarly, one can observe that the set \( P_k^0(n) \) is the set of extreme points of the compact convex set \( \text{Conv}(P_k^0(n)) := \{ X \in M_n : \|X\|_1 \leq 1, \|X\|_\infty \leq k, \text{tr}(X) = 0 \} \).

This lemma, together with Formulas 3.3 and 3.4 implies that

\[
\|A\|_{(k)} = \max_{M \in \text{Conv}(P_k(n))} \text{Re} \left( (AM) \right)
\]

\[
\|A\|_{(k)}^\delta = \max_{M \in \text{Conv}(P_k^0(n))} \text{Re} \left( (AM) \right)
\]

because the cost is linear and we are substituting the set \( P_k(n) \) and \( P_k^0(n) \) with their convex hulls.

Theorem 4.2 (Sion’s minimax theorem). Let \( X \) be a compact convex subset of a linear topological space and \( Y \) a convex subset of a linear topological space. If \( f \) is a real-valued function on \( X \times Y \) such that \( f(x, \cdot) \) is upper semicontinuous and quasi-concave on \( Y \) for any \( x \in X \) and \( f(\cdot, y) \) is lower semicontinuous and quasi-convex on \( X \) for any \( y \in Y \), then we have

\[
\min_x \sup_y f(x,y) = \sup_y \min_x f(x,y)
\]
Theorem 4.3. Given $A \in M_n$, one has

$$\|A\|_{(k)} = \min_{\alpha \in \mathbb{C}} \|A - \alpha I_n\|_{(k)}. \quad (4.6)$$

Proof. Using the maximal representation 4.4 for the sum of discrepancy values, we have

$$\|A\|_{(k)} = \max_{M \in \text{Conv}(P_k^0(n))} \text{Re tr}(AM)$$

$$= \max_{M \in \text{Conv}(P_k(n))} \min_{\alpha \in \mathbb{C}} \text{Re tr}(AM) - \text{Re tr}(M)$$

$$= \min_{\alpha \in \mathbb{C}} \max_{M \in \text{Conv}(P_k(n))} \text{Re tr}(A - \alpha I_n)M$$

$$= \min_{\alpha \in \mathbb{C}} \|A - \alpha I_n\|_{(k)}, \quad (4.7)$$

where the second equality follows by the fact that Conv($P_k^0(n)$) is the intersection of the traceless matrices with Conv($P_k(n)$). Also, the third equality follows by Sion’s minimax theorem, and the last equality follows by the maximal representation 4.3.

Therefore, the discrepancy values can be thought of as the vector-valued extension of the (spectral norm) distance of a linear operator to the class of scalar operators, which is defined as follows:

$$\delta_k(A) = \min_{\alpha \in \mathbb{C}} \sigma_k(A - \alpha I_n) = \max_{\|x\| = \|y\| = 1} |\langle Ax, y \rangle|. \quad (4.8)$$

The problem of projecting a linear operator onto the subspace of scalar linear operator has been investigated in the literature (see [2, 6]). Finally, the two relations between discrepancy and Ky-Fan norms would imply the following identity.

Corollary 4.4. Using Lemma 4.1 and Theorem 4.3, we have

$$\max_{U \in U(n)} \min_{\alpha \in \mathbb{C}} \|A - \alpha U\|_{(k)} = \|A\|_{(k)}$$

4.2 Discrepancy values and vectors for the Hermitian matrices

Now we can illustrate the connection between discrepancy values and the spectral spread for the Hermitian matrices. Note that when $A$ is Hermitian, we can solve the optimization problem 4.6 in closed form and get $\delta_{k-1}(A) = |\lambda_k(A) - \lambda_{n-k+1}(A)|/2$ for $k = 1, \ldots, [n/2]$. Moreover, if $n$ is odd, then $\delta_n(A)$ would be zero. On the other hand, in Def. 3.1, for a Hermitian matrix with eigenvectors $v_1, v_2, \ldots, v_n$ corresponding to the non-increasing eigenvalues, we can check that the vectors $x_{2k-1} = (-v_k + v_{n-k+1})/\sqrt{2}$ and $y_{2k-1} = (-v_k - v_{n-k+1})/\sqrt{2}$ for the odd terms, and $x_{2k} = y_{2k-1}$ and $y_{2k} = x_{2k-1}$ for the even terms are a maximizer of the problem 3.1. More compactly, for a Hermitian matrix $A$ we have

$$\delta(A) = \frac{|\lambda_k(A) - \lambda_{n-k+1}(A)|}{2}. \quad (4.9)$$

Remark 4.5. The discrepancy value of Hermitian matrices is invariant w.r.t. a particular transformation. We call this invariance looseness or slackness. It states that if we fix the outer eigenvalues, we can shift the inner eigenvalues and the discrepancy values remain the same so long as we do not pass the fixed eigenvalues. We are going to exploit this property later in this paper.

We have the interlacing property for the Hermitian matrices, which implies the stability of the discrepancy values under perturbation.
**Proposition 4.6** (Interlacing theorem). Let \( A \) be a Hermitian matrix of order \( n \), and \( B \) be a principal submatrix of \( A \). Using the Cauchy interlacing theorem one can easily show that \( \delta_1(A) = \delta_2(A) \geq \delta_1(B) = \delta_2(B) \geq \delta_3(A) = \delta_4(A) \geq \delta_3(B) = \delta_4(B) \geq \ldots \).

Let us consider the \( n \times n \) Hermitian matrix \( A \). One can observe that when \( n \) is even, the spectral spread of \( A \) has the following relationship with the discrepancy values:

\[
\delta(A) = \text{spr}^+(A \oplus A)/2.
\]

And when \( n \) is odd, \( \delta(A) \) has an extra zero attached to the end of the vector \( \text{spr}^+(A \oplus A)/2 \). Look at Appendix A to see one decomposition for Hermitian matrices using their discrepancy values and vectors.

### 4.3 Important classes of matrices based on their discrepancy values

We know that the unitary matrices as the set of square matrices with unit singular values, and partial isometries as the set of square matrices with zero and one singular values play an important role in the characterization of singular values. We believe that in order to understand the discrepancy values better, we need to understand the equivalent classes for the discrepancy values. In this section, we explore the structure of the classes of matrices with unit discrepancy values and talk about matrices with or zero and one discrepancy values. Before that, we need some tools:

**Proposition 4.7.** For the square matrix \( A \) we have \( \delta(A) = \sigma(A) \) iff it has a singular value decomposition \( A = U \Sigma V^* \), with \( U, V \in U_n \) and \( \text{Diag}(U^* V) = 0 \); i.e. \( \langle u_i, v_i \rangle = 0 \) for \( i = 1, \ldots, n \).

**Proof.** Note that the “if” part follows trivially by the definition of discrepancy values. We first prove the “only if” for the case of full-rank matrices without repeated singular values. In this case, we know that the singular vectors are unique up to multiplication by \( e^{i\theta} \). We have

\[
\max_{\| x_{11} \| = \| y_{11} \| = 1} |\langle Ax_{11}, y_{11} \rangle| = \max_{\| u_{11} \| = \| v_{11} \| = 1} |\langle Au_{11}, v_{11} \rangle|.
\]

Therefore, we must have the optimizers \( u_{11}^* = e^{i\theta_1} x_{11}^* \) and \( v_{11}^* = e^{i\theta_1} y_{11}^* \), i.e., \( \langle u_{11}^*, v_{11}^* \rangle = 0 \). Next, we have

\[
\max_{\| x_{12} \| = \| y_{12} \| = 1} |\langle Ax_{12}, y_{12} \rangle| + |\langle Ax_{22}, y_{22} \rangle| = \max_{\| u_{12} \| = \| v_{12} \| = 1} |\langle Au_{12}, v_{12} \rangle| + |\langle Au_{22}, v_{22} \rangle|,
\]

which implies that \( \langle u_{12}^*, v_{12}^* \rangle + \langle u_{22}^*, v_{22}^* \rangle = 0 \). On the other hand, we know that \( u_{12}^* = e^{i\theta_2} u_{11}^* \) and \( v_{12}^* = e^{i\theta_2} v_{11}^* \); thus \( \langle u_{22}^*, v_{22}^* \rangle = 0 \). The result would follow inductively. For matrices with repeated singular values or zero singular values, not all possible singular value decompositions of the matrix \( A \) have the property that \( \text{Diag}(U^* V) = 0 \), but a decomposition with such property belongs to the set of possible singular value decompositions of that matrix. \( \square \)

**Remark 4.8.** In other words, matrices of the form \( A = \sum_{i=1}^n \delta_i X_i + \alpha I_n \), where \( \delta_i \geq 0 \), and \( X_i \) are nilpotent rank-1 isometries and mutually orthogonal, i.e. \( \langle X_i, X_j \rangle = 0 \), have discrepancy values equal to \( \delta_i \).

**Corollary 4.9.** If \( A = UV^* + \alpha I_n \), where \( U \) and \( V \) are unitary matrices with \( \text{tr}(U^* V) = 0 \), then \( \delta_i(A) = 1 \).

**Proof.** It follows by the fact that there exists a unitary matrix \( Q \) such that for \( \tilde{U} = UQ \) and \( \tilde{V} = VQ \) we have \( \text{Diag}(\tilde{U}^* \tilde{V}) = 0 \) and \( A = \tilde{U} \tilde{V}^* + \alpha I_n \). \( \square \)
**Corollary 4.10.** The singular values and the discrepancy values of matrices with the form $B = QAQ^*$, where $A$ is an antidiagonal matrix and $Q$ unitary, are equal to the absolute value of the anti-diagonal entries of $A$.

**Proposition 4.11.** For a $2n \times 2n$ Hamiltonian matrix $H$, we have

$$\delta(H) = \sigma(H).$$

**Proof.** We know that any Hamiltonian matrix can be represented as the multiplication of $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$ – not to confuse with the exchange matrix – and a symmetric matrix, thus we have $H = JQ\Lambda Q^*$. Let $\Lambda = \Sigma D$ where $D$ is just a diagonal matrix with \pm 1 entries. In the singular value decomposition of $H = U\Sigma V^*$ we have $\mu = JQ$ and $V = QD$. To show that $\text{Diag}(DQ^*JQ) = 0$ one only needs to prove $\text{Diag}(Q^*JQ) = 0$, which is immediate after recognizing that $\delta(Q^*JQ) = \sigma(Q^*JQ)$.

Note that the same argument can be used to prove that the singular and discrepancy values of the multiplication of $J$ and any normal matrix are equal.

Let us adopt the notation $\Psi(n)$ to refer to the set of matrices with unit discrepancy values. The following theorem states that this class is equivalent to the scalar shifts of the traceless unitary matrices.

**Theorem 4.12.** A square matrix belongs to $\Psi(n)$ iff it has the form $M - \alpha I_n$, where $M$ is a unitary matrix with zero trace (real part of) and $\alpha$ an arbitrary complex number.

**Proof.** We already proved in Corollary 4.9 that $M - \alpha I_n \in \Psi(n)$ when $M$ is a traceless unitary matrix. By the definition of discrepancy values, there exists $\hat{\alpha}$ such that $\sigma_1(A - \hat{\alpha}I_n) = \delta_1(A) = 1$, and $\sigma_1(A - \hat{\alpha}I_n) + \sigma_2(A - \hat{\alpha}I_n) \geq \delta_1(A) + \delta_2(A) = 2$. Therefore, $\sigma_2(A - \hat{\alpha}I_n) \geq 1$, but we assumed that $\sigma_2(A - \hat{\alpha}I_n) \leq \sigma_1(A - \hat{\alpha}I_n)$; hence $\sigma_2(A - \hat{\alpha}I_n) = \delta_2(A) = 1$. Using the same argument we can show that $\sigma_k(A - \hat{\alpha}I_n) = \delta_k(A) = 1$, for $1 \leq k \leq n$. Finally, by proposition 4.7 we know that matrix $A - \hat{\alpha}I_n$ has a decomposition $UV^*$ where $\text{Diag}(U^*V) = 0$.

In other words, a matrix is in $\Psi(n)$ iff for some $\alpha$ its eigenvalues satisfy $|\lambda_i - \alpha| = 1$ and $\sum_{i=0}^{n} \text{Re} \lambda_i = 0$, which always has roots for $n \geq 2$. Thus, $\Psi(n)$ is not empty.

Before studying the class of matrices with zero or one discrepancy values, let us define the notion of Discrepancy-rank for a square matrix as follows:

$$r^\delta(A) := \text{number of non-zero entries of } \delta(A).$$

(4.10)

One can verify that $r^\delta(A) = r(A - \hat{\alpha}I_n)$ where $\hat{\alpha} = \text{argmin}_{\alpha \in \mathbb{C}} \|A - \alpha I_n\|_{(n)}$. We can also see that $r^\delta(A) = 0$ iff $A = \alpha I_n$ for a complex number $\alpha$. And $r^\delta(A) = 1$ iff $A = \beta uv^* + \alpha I_n$ for some complex-values $\alpha$ and $\beta$, and orthogonal unitary vectors $u$ and $v$.

**Definition 4.13.** Let $\Psi_k(n)$ denote the set of $n \times n$ matrices with discrepancy-rank $k$ with unit non-zero discrepancy values (i.e. $\delta_i = 1$ for $i = 1, \ldots, k$).

Now we want to explore the relation between $\Psi_k(n)$ and the set $\mathcal{P}_k^0(n)$. Note that $\mathcal{P}_k^0(n) \neq \Psi_k(n)$; for instance, consider $M = \text{Diag}(1,3,4)^3$ which belongs to $\Psi_2(3)$ but not to $\mathcal{P}_2^0(3)$. However, we can state the following result:

**Lemma 3.** $\mathcal{P}_k^0(n) \subset \Psi_k(n)$.

**Proof.** Consider matrix $M \in \mathcal{P}_k^0(n)$. We can observe that $\|M\|_k^\delta = k$ by using the left and right singular vectors of $M$ as the vectors $x$ and $y$ in the definition 3.1. On the other hand, we know that $\delta(M) \prec_{\text{Diag}} 1_k$. Therefore the only possibility is that $\delta(M) = 1_k$.\[\square\]
4.4 Basic majorization inequalities

A plethora of majorization inequalities for singular values exist in the literature, which tell us about many properties of these quantities. In this part, we investigate some majorization results for discrepancy values. We also see that the discrepancy values satisfy many of the known majorization bounds known for the spectral spread of Hermitian matrices; thus, suggesting that these objects are the true generalization of the spectral spread. First, let us state two trivial inequalities.

**Proposition 4.14.** For $n \times n$ matrices $A$ and $B$, one has

1. $|\delta(A) - \delta(B)| \prec_w \delta(A \pm B) \prec_w \delta(A) + \delta(B)$.
2. $\delta(A) \prec_w \sigma(A)$.

Now let us look at some of the non-trivial majorization inequalities. The first inequality is about the pinching of matrices. A similar majorization inequality holds for singular values.

**Proposition 4.15.** Let $A$ be any square matrix and $C(A)$ denote any pinching of this matrix, then we have

$$
\delta(C(A)) \prec_w \delta(A)
$$

*Proof.* We know that $C(A) = \frac{1}{2}(A + UAU^*)$ for some unitary matrix $U$. Then we have

$$
\sum_{i=1}^{k} \delta_i(C(A)) = \frac{1}{2} \sum_{i=1}^{k} \delta_i(A + UAU^*) 
\leq \sum_{i=1}^{k} \delta_i(A) \quad (4.11)
$$

□

**Proposition 4.16.** Let $A$ be a square matrix and $L(A)$ denote the anti-pinching of this matrix; namely $L(A) = A - C(A)$, then we have

$$
\delta(L(A)) \prec_w \delta(A)
$$

*Proof.* We know that $C(A) = \frac{1}{2}(A + UAU^*)$ for some unitary matrix $U$. Then we have

$$
\sum_{i=1}^{k} \delta_i(L(A)) = \frac{1}{2} \sum_{i=1}^{k} \delta_i(A - UAU^*) 
\leq \sum_{i=1}^{k} \delta_i(A) \quad (4.12)
$$

□

**Corollary 4.17.** For any $n \times n$ matrices $A_1, A_2, A_3, A_4$, we have

$$
\sigma\left(\begin{bmatrix} 0 & A_2 \\ A_1 & 0 \end{bmatrix}\right) = \delta\left(\begin{bmatrix} 0 & A_2 \\ A_1 & 0 \end{bmatrix}\right) \prec_w \delta\left(\begin{bmatrix} A_3 & A_2 \\ A_1 & A_4 \end{bmatrix}\right)
$$
Proof. Assume that matrices $A_1$ and $A_2$ have the singular value decompositions $A_1 = U_1 \Sigma_1 V_1^*$ and $A_2 = U_2 \Sigma_2 V_2^*$, then we have
\[
\begin{bmatrix}
0 & A_2 \\
A_1 & 0
\end{bmatrix} = 
\begin{bmatrix}
0 & U_2 \\
U_1 & 0
\end{bmatrix} 
\begin{bmatrix}
\Sigma_1 & 0 \\
0 & \Sigma_2
\end{bmatrix} 
\begin{bmatrix}
V_1 & 0 \\
0 & V_2
\end{bmatrix}^*.
\]
We can see that $Q$ and $R$ are also unitary and $\text{Diag}(Q^* R) = 0$; therefore by Proposition 4.7 we have the equality. The weakly majorization follows by Proposition 4.16.

This inequality generalizes a similar inequality for the spectral spread. It implies the following inequality:

**Corollary 4.18.** For the isometry $S \in M_{n,k}(\mathbb{C})$ and arbitrary matrix $X$ we have
\[
\sigma(S^* X S^\perp) \prec_w \delta(X)\big|_{\min(k,n-k)}.
\]

**Lemma 4.** For the $n \times n$ Hermitian matrix $X$ we have
\[
\delta(e^{it} X) \prec_w \sigma(X).
\]

Proof. We first show that
\[
\delta(e^{it} X) \prec_w \int_0^1 \delta(iX e^{it} X) dt.
\]
Thus using the fact that $\delta(i e^{it} X) \prec_w \sigma(X)$ we have
\[
\delta(e^{it} X) \prec_w \int_0^1 \sigma(X) dt = \sigma(X).
\]
To prove Inequality 4.13, let us define function $f(t) = e^{it}X$. We have
\[
e^{it}X = I_n - \sum_{j=1}^{m-1} f\left(\frac{j}{m}\right) - f\left(\frac{j+1}{m}\right).
\]
Hence
\[
\delta(e^{it} X) \prec_w \sum_{j=1}^{m-1} \delta\left(f\left(\frac{j}{m}\right) - f\left(\frac{j+1}{m}\right)\right).
\]
If we let $m$ goes to infinity the LHS would converge to
\[
\int_0^1 \delta\left(\frac{d}{dt} f(t)\right) dt.
\]
Finally, we have $df(t)/dt = iX e^{it} X$.

**Corollary 4.19.** For any square matrix $A$ we have
\[
\delta(e^{it} |A|) \prec_w \sigma(A),
\]
where $|A| = (A^* A)^{1/2}$.

We also have another similar relationship between the discrepancy of the absolute values of a matrix and the discrepancy of the matrix. For any matrix $A$,
\[
\delta(|A|) = \frac{||\sigma(A) - \sigma^\perp(A)||^\perp}{2} \prec_w \delta(A).
\]
In order to prove this inequality, we need the following statement:
**Proposition 4.20.** For the PSD matrix $P$ and unitary matrix $Q$, we have
\[ \delta(P) \prec_w \delta(QP). \]

**Proof.** It suffices to show that for any diagonal matrix with non-negative entries $\Lambda$ and unitary matrix $Q$ we have
\[ \delta(\Lambda) \prec_w \delta(Q\Lambda). \]
Note that $\sigma(\Lambda) - |\alpha|\sigma(Q) \prec_w \sigma(\Lambda - \alpha Q)$; hence for any unitary matrix $Q$, we have $\min_{\alpha \in \mathbb{R}^+} \|\Lambda - \alpha I\| \leq \min_{\alpha \in \mathbb{C}} \|\Lambda - \alpha Q\|$, which implies the desired inequality. \qed

Lastly, we use the above results to prove an upper bound for the principal angles between two subspaces.

**Theorem 4.21.** Given the Hermitian matrix $X$, the principal angles between $k$ dimensional subspaces $S \subset \mathbb{C}^n$ and $T = e^{iX}S \subset \mathbb{C}^n$ satisfies
\[ \Theta(S, T) \prec_w \delta(X)_{\min(k, n-k)}, \]
where $\delta(X)_{|k}$ denotes a $k$ dimensional vector with the first $k$ entries of $\delta(X)$.

**Proof.** We follow similar steps as in [16] to prove the theorem. Define $T(t) = e^{itX}S$. We have $T(0) = S$ and $T(1) = T$. By the triangle inequality for the principal angles we have
\[
\Theta(S, T) \prec_w \sum_{j=0}^{m-1} \Theta(T\left(\frac{j}{m}\right), T\left(\frac{j+1}{m}\right)) \\
= \sum_{j=0}^{m-1} \arcsin \left( \sigma(S^*e^{i\frac{jX}{m}}S_\perp) \right) \\
= m \arcsin \left( \sigma(S^*e^{i\frac{X}{m}}S_\perp) \right) \\
\xrightarrow{\text{L'Hôpital's rule}} \sigma(S^*XS_\perp) \\
\prec_w \delta(X)_{\min(k, n-k)},
\]
where the last majorization follows by Corollary 4.18. The limit can be justified by L'Hôpital's rule and the continuity of singular values. \qed

### 4.5 The discrepancy values of the direct sum of matrices

In the last part of this section, we briefly discuss the discrepancy values of the direct sum of matrices. These results would particularly become useful when comparing the results for discrepancy values with ones for the spread of Hermitian matrices. Inspired by the minimal representation of the discrepancy pseudo-norm, we will introduce for $A, B \in M_n(\mathbb{C})$,
\[
\sum_{i=1}^{k} \delta_i(A, B) = \min_{\alpha \in \mathbb{C}} \frac{\|A - \alpha I_n\|_{(k)} + \|B - \alpha I_n\|_{(k)}}{2}.
\]
Note that $\delta(A, Q^*AQ) = \delta(A)$, for any unitary matrix $Q$.

**Proposition 4.22.** We have
\[ (\delta(A, B), \delta(A, B)) \prec_w \delta(A \oplus B). \]
Proof. First note that for any \(x, y \in \mathbb{R}^n\), we have
\[
\left( \frac{x^+ + y^+}{2}, \frac{x^+ + y^+}{2} \right) \prec_w (x, y)^+. 
\]
Then for any scalar \(\alpha\) we have
\[
\left( \frac{\sigma(A - \alpha I_n)^+ + \sigma(B - \alpha I_n)^+}{2}, \frac{\sigma(A - \alpha I_n)^+ + \sigma(B - \alpha I_n)^+}{2} \right) \prec_w \sigma(A \oplus B - \alpha I_n)^+.
\]
The result would follow by minimizing the RHS of the above inequality w.r.t. \(\alpha\).

Let us define an averaging operator on vectors. Let’s assume \(x \in \mathbb{R}^{kn}\) where \(k\) and \(n\) are some positive integers. Then we define \(\mu_k(x) \in \mathbb{R}^n\) by
\[
\mu_k(x) = \begin{bmatrix}
x_1 + \ldots + x_k \\
x_{k+1} + \ldots + x_{2k} \\
\vdots \\
x_{n-k+1} + \ldots + x_n
\end{bmatrix}.
\]
(4.16)
Using this operator we can state the above result as \(\delta(A, B) \prec_w \mu_2 \left( \delta(A \oplus B) \right)\), or state the connection between the spectral spread of the Hermitian matrix \(A \in M_n\) with even \(n\) and discrepancy values as \(\mu_2(\delta(A)) = \text{sp}^+(A)/2\). It is also trivial to see the following inequality:

**Proposition 4.23.** For the PSD matrix \(A\), we have
\[
\mu_2(\delta(A)) \leq \mu_2(\sigma(A)).
\]

**Lemma 5.** Let \(A\) be \(n \times n\) matrix, then
\[
\mu_k \left( \delta(\overbrace{A \oplus \cdots \oplus A}^{k}) \right) = \delta(A)
\]

Proof. Since we have
\[
\left\| \overbrace{A \oplus \cdots \oplus A}^{k} \right\|_{(km)} = k \left\| A \right\|_{(m)} , \quad m = 1, \ldots, n,
\]
we can see that
\[
\min_{\alpha \in \mathbb{C}} \left\| \overbrace{A \oplus \cdots \oplus A - \alpha I_n}^{k} \right\|_{(km)} = \min_{\alpha \in \mathbb{C}} \left\| (A - \alpha I_n) \oplus \cdots \oplus (A - \alpha I_n) \right\|_{(km)}
\]
\[
= k \min_{\alpha \in \mathbb{C}} \left\| A - \alpha I_n \right\|_{(m)},
\]
for \(m = 1, \ldots, n\). Hence we have the desired result.

We can similarly prove this lemma if matrices \(A\) are arbitrarily replaced by \(A^*\).

**5 Application: bounding the norms of the commutators**

In this section, the main goal is to employ the tools we developed in the previous sections to obtain upper bounds on the Ky-Fan norms of the commutators. We have the following invariance for the generalized commutator \(AX - XB\):
\[
(A - \alpha I_n)X - X(B - \alpha I_n) = AX - XB, \tag{5.1}
\]
for any complex number $\alpha$. An special case of this invariance is $[A, B] = [A, B - \alpha I_n]$. We now want to use invariance 5.1 together with the properties of the discrepancy values to amplify the already known majorization inequalities such as $\sigma([A, B]) \prec_w 2\sigma(A)\sigma(B)$ and achieve sharp bounds for similar objects. For example, we have the following result:

**Lemma 6.** For any square matrices $A, B$ and partial isometry $X$ we have

$$\sigma \left( AX - XB \right) \prec_w 2\delta(A, B)\sigma(X)$$

*Proof.* Using invariance 5.1, we know that $\sum_{i=1}^{k} \sigma_i(AX - XB) \leq \sum_{i=1}^{k} (\sigma_i(A - \alpha I) + \sigma_i(B - \alpha I))\sigma_i(X)$ for any $\alpha \in \mathbb{C}$. Now we minimize the RHS for any $k = 1, \ldots, n$. Considering the fact that $\sigma_i(X)$ is either 1 or 0, we get the desired result. $\square$

Now let us propose a decomposition that enables us to extend the above lemma.

**Lemma 7.** Any square matrix $A$ has the following decomposition

$$A = \sum_{i=1}^{n} \alpha_i X_i,$$

where $X_i$ is a rank $i$ isometry matrix; i.e. with $i$ non-zero singular values which are equal to one. And $\sum_{i=1}^{n} \alpha_i = \sigma_1, \sum_{i=2}^{n} \alpha_i = \sigma_2, \ldots, \alpha_n = \sigma_n$, where the singular values are ordered in a decreasing manner.

*Proof.* This easily follows by the singular value decomposition $A = \sum_{i=1}^{n} \sigma_i x_i y_i^*$, and using summation by part formula. We can rewrite it as follows:

$$A = (\sigma_1 - \sigma_2)x_1 y_1^* + (\sigma_2 - \sigma_3)(x_1 y_1^* + x_2 y_2^*) + \cdots + \sigma_n(x_1 y_1^* + \cdots + x_n y_n^*).$$

Now let $\alpha_i = \sigma_i - \sigma_{i+1}$ and $X_i = x_1 y_1^* + \cdots + x_i y_i^*$. $\square$

Note that if the matrix $A$ in the previous theorem is positive definite, then each $X_i$ is an orthogonal projection matrix. Now we are able to demonstrate the first main inequality of this section.

**Theorem 5.1.** For the $n \times n$ square matrices $A, B$ and $X$ we have

$$\sigma \left( AX - XB \right) \prec_w 2\delta(A, B)\sigma(X).$$

*Proof.* Using the previous decomposition for $X = \sum_{i=1}^{n} \alpha_i X_i$, we have

$$\sigma(AX - XB) = \sigma \left( \sum_{i=1}^{n} \alpha_i (AX_i - X_i B) \right) \prec_w \sum_{i=1}^{n} \alpha_i \sigma(AX_i - X_i B) \prec_w 2\delta(A, B)\sum_{i=1}^{n} \alpha_i \sigma(X_i) = 2\delta(A, B)\sigma(X),$$

where the first majorization follows by the fact that $\alpha_i \geq 0$, and the second majorization from Lemma 6. $\square$
Using the fact that $\delta(A, A) = \delta(A)$, we have the following:

**Corollary 5.2.** For the $n \times n$ square matrices $A$ and $B$ we have

$$\sigma([B, A]) \prec_w 2\delta(B)\sigma(A). \quad (5.4)$$

**Remark 5.3.** Note that if either $A$ or $B$ belongs to $\Psi_k(n)$ (i.e. it only has discrepancy values zero or one) then we can prove $\sigma([B, A]) \prec_w 2\delta(B)\delta(A)$ by minimizing the RHS of the above inequalities.

Let us continue with inequalities for the generalized commutators.

**Corollary 5.4.** For the $n \times n$ square matrices $A$ and $B$ we have

$$\sigma(AX - XB) \prec_w 2\mu_2(\delta(A \oplus B))\sigma(X), \quad (5.5)$$

and

$$\sigma(A - B) \prec_w 2\delta(A, B) \prec_w 2\mu_2(\delta(A \oplus B)). \quad (5.6)$$

Note that if $A$ and $B$ are Hermitian then $2\mu_2(\delta(A \oplus B)) = \text{Spr}^+(A \oplus B)$. This inequality has been previously shown for the spectral spread.

For $\delta(A, 0) = \min_{\alpha \in \mathbb{C}} \frac{\|A - \alpha I_n\|_{(k)}}{2^{|k|}}$, we have

$$\sigma(A) = 2\delta(A, 0). \quad (5.7)$$

Since by triangle inequality we have for any $\alpha$

$$\|A\|_{(k)} \leq \|A - \alpha I_n\|_{(k)} + \|\alpha I_n\|_{(k)},$$

and if we let $\alpha = 0$, the inequality becomes equality.

We also have $\delta(A, A^*) = \delta(A)$, thus $\sigma(A - A^*) \prec_w 2\delta(A)$, which means that $|\lambda(S(A))| \prec_w \delta(A)$ and $\sigma(A^2 - AA^*) \prec_w 2\sigma(A)\delta(A)$. Using a similar argument we can deduce:

**Proposition 5.5.** For any square matrix $X \in M_n(C)$ we have

$$\sum_{i=1}^k \delta_i(X, X^*) = \min_{\alpha \in \mathbb{R}} \|X + \alpha I_n\|_{(k)},$$

and

$$\sum_{i=1}^k \delta_i(-X, X^*) = \min_{\alpha \in \mathbb{R}} \|X + i\alpha I_n\|_{(k)}.$$

They enable us obtaining majorization bounds for $\sigma(AB^* - BA)$ and $\sigma(AB^* + BA)$. Now we want to derive similar inequalities for the commutators. First, we need the following lemma:

**Lemma 8.** The Hermitian matrix $A$ has the decomposition

$$A = \omega Y + \sum_{i=1}^n \beta_i X_i = \omega Y + (\beta_1 - \omega)X_1 + (\beta_2 - \omega)X_2 + \cdots + \beta_n X_n,$$

where $\sum_{i=1}^n \beta_i = \delta_k(A)$, and $\delta(X_i)$ is a zero-one vector with $k$ ones, $\omega$ is a complex scalar, and $\delta(Y) = 0$.

Note that the last two conditions imply that $\sum_{i=1}^n \delta(X_i)\beta_i = \delta(A)$. We now want to employ the slackness property of the discrepancy values of Hermitian matrices to prove the result.
Proof. Note that for the Hermitian matrix $A$ we have $\beta_{2k-1} = 0$ since $\delta_{2k-1} = \delta_{2k}$ for $k = 1, \ldots, \lfloor \frac{n}{2} \rfloor$. First we can verify that we have 

$$\Lambda = (\delta_2 - \delta_3) \operatorname{Diag} \left( \frac{\alpha_3 - \alpha_2}{\delta_2 - \delta_3}, \ldots, \frac{\alpha_3 - \alpha_2}{\delta_2 - \delta_3}, -1 \right)$$

$$\Lambda = (\delta_4 - \delta_5) \operatorname{Diag} \left( \frac{\alpha_5 - \alpha_4}{\delta_4 - \delta_5}, \ldots, \frac{\alpha_5 - \alpha_4}{\delta_4 - \delta_5}, -1, -1 \right)$$

$$\Lambda = \cdots + \delta_{n-1} \operatorname{Diag} \left( 1, \ldots, 1, -1, \ldots, -1 \right) + \alpha_1 (I + O),$$

where we assume that $\Lambda = \operatorname{Diag}(\lambda(A))$, $\alpha_1 = \frac{\lambda_1 + \lambda_2}{2}$, $\alpha_3 = \frac{\lambda_2 + \lambda_{n-1}}{2}$, and $\alpha_4 = \frac{\lambda_1 + \lambda_n}{2}$.

Matrix $O$ is zero when $n$ is even, otherwise only its Central entry is nonzero, and we have $\delta(Y) = 0$. To show that $\delta(X_2) = 1_2, \delta(X_4) = 1_4, \cdots$, we need to prove that

$$|\alpha_i - \alpha_{i-1}| \leq |\delta_i - \delta_{i-1}|$$

for all odd $i$ greater than or equal to 3. For simplicity of notation, we only prove this fact for one case and the other cases would follow similarly. We want to show that $|\alpha_3 - \alpha_2| \leq |\delta_3 - \delta_2|$ which is equivalent to

$$|(\lambda_1 - \lambda_2) - (\lambda_{n-1} - \lambda_n)| \leq |(\lambda_1 - \lambda_2) + (\lambda_{n-1} - \lambda_n)|.$$

This would follow immediately after recognizing that $\lambda_1 - \lambda_2 \geq 0$ and $\lambda_{n-1} - \lambda_n \geq 0$. Finally, the result of the theorem would follow because

$$A = Q\Lambda Q^* = (\delta_2 - \delta_3)QX_2Q^* + (\delta_4 - \delta_5)QX_4Q^* + \cdots + \delta_{n-1}QX_{n-1}Q^* + \alpha_1QYQ^*,$$

and the fact that discrepancy values are invariant under unitarily similarity transforms.

Now we have enough tools to prove the second main majorization inequality in this section.

**Theorem 5.6.** For the $n \times n$ square matrix $B$ and Hermitian matrix $A$ we have

$$\sigma([A, B]) \lesssim_w 2\delta(A) \delta(B).$$

**Proof.** Using the previous decomposition for $A$, we have

$$\sigma([B, A]) = \sigma \left( \left[ B, \omega Y + \sum_{i=1}^n \beta_i X_i \right] \right)$$

$$\lesssim_w |\omega| \sigma([B, Y]) + \sum_{i=1}^n \beta_i \sigma([B, X_i])$$

$$\lesssim_w 2\delta(B) \sum_{i=1}^n \beta_i \delta(X_i)$$

$$= 2\delta(B) \delta(A),$$

where the first majorization follows by the fact that $\beta_i \geq 0$, and the second majorization from Coroll. 5.2, and the fact that $\delta(Y) = 0$. 

**Corollary 5.7.** For the $n \times n$ square matrix $A$ and a normal matrix $B$ whose eigenvalues lie on a straight line in the complex plane, we have

$$\sigma([B, A]) \lesssim_w 2\delta(B) \delta(A).$$

(5.10)
Corollary 5.8. Let $A$ be an arbitrary $n$ by $n$ matrix, and $P$ be an orthogonal projection matrix with rank $r$. Then for $k \in [n]$ we have
\[
\| [A, P] \|_{(k)} \leq \delta_k(A),
\]
where $q = \min(2r, 2n - 2r)$.

Corollary 5.9. Given matrix $X \in M_n(C)$ and the orthogonal projection matrix $P$; i.e., $P = P^* = P^2$, we have
\[
\text{Corollary 5.8.} \quad \sigma(\text{PX}(I_n - P)) \prec_s \delta(X).
\]

Proof. It follows by the fact that $\sigma(\text{PX}(I_n - P)) \prec_s \sigma([P, X])$, which follows by the fact that the set of the singular values of $[P, X]$ includes all the nonzero singular values of $\text{PX}(I_n - P)$.

The inequality (5.11) has been proven for the spectral of spread. Here we showed that it holds for any square matrices. The inequality 5.8 implies majorization bounds for other operations:
\[
\sigma(A^2B^2 - (AB)^2) = \sigma(A[A, B]B) \prec_w 2\delta(A)\delta(B)\sigma(A)\sigma(B).
\]

5.1 Maximally non-commutative Hermitian matrices

In this section, we utilize the derived majorization bounds in the previous part, together with Fan’s dominance theorem to solve some optimization problems. These results indicate that the inequalities are sharp. Before mentioning the results, let us recall the celebrated Fan’s dominance theorem.

Theorem 5.10 (Fan’s dominance theorem). Given $A, B \in M_{m,n}(C)$, $\|A\| \leq \|B\|$ for any unitarily invariant norm on $M_{m,n}(C)$ if and only if $\sigma(A) \prec_w \sigma(B)$.

The diameter of the unitary orbit of a linear operator $A \in M_n(C)$ w.r.t. the $k$th Ky-Fan norm is defined by
\[
d_k(A) := \max_{U, V \in U(n)} \|VAV^* - UAU^*\|_{(k)}.
\]
By inequality 5.2, we know that $d_k(A) \leq 2\|A\|_{(k)}$. Now for the case of Hermitian matrices we can show that this bound is actually tight.

Proposition 5.11 (The generalization of Thm 1.2 in [6]). Let $A$ be a Hermitian matrix, then $d_k(A) = 2\|A\|_{(k)}^\delta$.

Proof. Let $A$ have the eigenvalue decomposition $A = Q\Lambda Q^*$, such that the eigenvalues are in the nonincreasing order and set $Q^*UQ$ to be the exchange matrix, then we have
\[
\|A - UAU^*\|_{(k)} = \|\Lambda - (Q^*UQ)\Lambda(Q^*UQ)^*\|_{(k)} = \sum_{i=1}^k |\lambda_i - \lambda_{n-i+1}|^\delta.
\]

A special case of the inequality 5.8 states that for the Hermitian matrices $A$ and $B$, we have
\[
\sigma([A, B]) \prec_w \frac{|\lambda^\delta(A) - \lambda^\delta(B)|}{2}.
\]

Before stating the sharpness of the above inequality (which implies that inequality 5.8 is sharp), let us define a family of rotation matrices.
Definition 5.12. Given $\theta \in [0, 2\pi)$, define the rotation matrix $R_n(\theta)$ for even $n$ and odd $n$ as

\[
\begin{bmatrix}
\cos \theta & 0 & \cdots & 0 & -\sin \theta \\
0 & \cdots & & \cdots & 0 \\
\vdots & & \ddots & & \vdots \\
0 & \cdots & \cdots & \cos \theta & \sin \theta \\
\sin \theta & 0 & \cdots & 0 & \cos \theta
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
\cos \theta & 0 & \cdots & 0 & -\sin \theta \\
0 & \cdots & & \cdots & 0 \\
\vdots & & \ddots & & \vdots \\
\sin \theta & 0 & \cdots & 0 & \cos \theta \\
0 & \cdots & \cdots & \cos \theta & -\sin \theta
\end{bmatrix}
\]

respectively. This family can be viewed as the direct sum of equivalent two by two rotation matrices.

Theorem 5.13. Let $A, B$ be two Hermitian matrices with the eigenvalue decompositions $A = Q\Lambda Q^*$ and $B = VDV^*$. Assume that $\Lambda = \text{diag} (\lambda_1, \ldots, \lambda_n)$ and $\Lambda = \text{diag} (d_1, \ldots, d_n)$ where $\lambda_1 \geq \cdots \geq \lambda_n$ and $d_1 \geq \cdots \geq d_n$. Then we have

\[
QR_n(\frac{\pi}{4}) V^* \in \text{argmax}_{U \in \mathbb{U}(n)} \||AUBU^* - UBU^* A||. \tag{5.14}
\]

Proof. Note that the matrix $\tilde{U} := QR_n(\frac{\pi}{4}) V^*$ is indeed a unitary matrix and it is independent of the eigenvalues. At the maximum point, we have

\[
\sigma(A \tilde{U} B \tilde{U}^* - \tilde{U} B \tilde{U}^* A) = \sigma(\Lambda R_n DR_n^* - R_n(\frac{\pi}{4}) DR_n(\frac{\pi}{4})^* \Lambda).
\]

And the matrix $M = \Lambda R_n(\frac{\pi}{4}) DR_n(\frac{\pi}{4})^* - R_n(\frac{\pi}{4}) DR_n(\frac{\pi}{4})^* \Lambda$ is an anti-diagonal skew Hermitian matrix, hence $(MM^*)^{1/2}$ is diagonal. We can observe that

\[
\sigma_i(M), \sigma_{n-i+1}(M) = \frac{|\lambda_i - \lambda_{n-i+1}| |d_i - d_{n-i+1}|}{2}, \quad \text{for } i = 1, \ldots, n.
\]

By inequality 5.13, we know that these specific singular values majorize the singular values of $[A, UBU^*]$ for the arbitrary unitary matrix $U$. Hence, by Fan dominance theorem $\tilde{U}$ is a maximizer of the above optimization problem.

Indeed we can solve similar optimization problems using the same technique. First, let us generalize the inequalities in the last part.

Proposition 5.14. Let $A_1, \ldots, A_k$ be square matrices of the same size, then

\[
\sigma\left([A_1, [A_2, \ldots, [A_{k-1}, A_k] \ldots]]\right) \prec_w 2^{k-1} \sigma(A_1) \delta(A_2) \cdots \delta(A_k).
\]

This follows by repeated application of $\sigma([X, Y]) \prec_w 2 \delta(X) \sigma(Y)$. If $A_1$ is Hermitian, in the very last inequality we can use $\sigma([X, Y]) \prec_w 2 \delta(X) \delta(Y)$ and have

\[
\sigma\left([A_1, [A_2, \ldots, [A_{k-1}, A_k] \ldots]]\right) \prec_w 2^{k-1} \delta(A_1) \delta(A_2) \cdots \delta(A_k).
\]

Corollary 5.15. Let $A_1, \ldots, A_k$ be Hermitian matrices

\[
\sigma\left([A_1, [A_2, \ldots, [A_{k-1}, A_k] \ldots]]\right) \prec_w \frac{|\lambda^\dagger(A_1) - \lambda^\dagger(A_1)| \cdots |\lambda^\dagger(A_k) - \lambda^\dagger(A_k)|}{2}.
\]
Proposition 5.16. For Hermitian matrices $A, B \in M_n$ with the eigenvalue decompositions $A = Q\Lambda Q^*$ and $B = VDV^*$, we have

$$QR_n\left(\frac{\pi}{4}\right)V^* \in \arg\max\|\left[A, [A, UBU^*]\right]\| = \arg\max\|\left[A, [A, [A, UBU^*]]\right]\| U \in U(n)$$

$$= \arg\max\|\left[A, [UBU^*, [UBU^*, [UBU^*, A]]]\right]\| U \in U(n)$$

5.2 Majorization bound for other operations with scalar invariance

We now aim to show that the discrepancy values could be useful when there is a multiplicative scalar invariance.

Lemma 9. Given $A, B \in M_n$ we have

$$\sigma(e^A Be^{-A}) \prec_w \sigma(B) \exp(2\delta(A)),$$

where $\exp(\cdot)$ denotes the elementwise exponential.

Proof. We know that [3]

$$e^A Be^{-A} = B + [A, B] + [A, [A, B]]/2! + \cdots.$$

Therefore

$$\sigma(e^A Be^{-A}) \prec_w \sigma(B) + \sigma([A, B]) + \sigma([A, [A, B]]/2!) + \cdots.$$

Thus

$$\sigma(e^A Be^{-A}) \prec_w \sigma(B) \sum_{j=0}^{\infty} \frac{(2\delta(A))^j}{j!}.$$

Using the Taylor expansion of $\exp(\cdot)$ we get the result.

The previous lemma implies the following:

Corollary 5.17. Let $A \succ 0$ be a strictly positive matrix and $B \in M_n$, then

$$\sigma(ABA^{-1}) \prec_w \sigma(B) \frac{\lambda^\dagger(A)}{\lambda(\lambda^\dagger(A))}.$$
Similarly, we can solve maximization problems using these inequalities.

**Corollary 5.18.** Let $A$ be a positive definite matrix and $B \in M_n$ the eigendecomposition $A = V \Lambda V^*$ and singular value decompositions $B = QD P^*$. Assume that $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_n)$ and $D = \text{diag}(d_1, \cdots, d_n)$ where $\lambda_1 \geq \cdots \geq \lambda_n$ and $d_1 \geq \cdots \geq d_n$. Then we have

$$
(V R_n(\frac{\pi}{4}) Q, V R_n(\frac{\pi}{4}) P^*) \in \arg\max_{U_1, U_2 \in U(n)} \|AU_1 B U_2^* A^{-1}\|.
$$

(5.15)

### 6 Calculation of the discrepancy values

As we showed earlier, for the Hermitian matrix $A$, we know the discrepancy values in closed form:

$$
\delta_1(A), \delta_2(A) = \frac{\lambda_1 - \lambda_n}{2}, \delta_3(A), \delta_4(A) = \frac{\lambda_2 - \lambda_{n-1}}{2},
$$

etc. In fact, we can still find the discrepancy of some classes of matrices easily.

**Proposition 6.1.** Given that $\delta(e^{i\theta}A) = \delta(A)$ and the formula for the discrepancy values of Hermitian matrices, one can easily show that for any normal matrix whose eigenvalues lie on a straight line in the complex plane and are symmetric about the origin we have $\delta(A) = \sigma(A)$.

**Corollary 6.2.** The singular and discrepancy values of a real anti-symmetric matrix are equal. Furthermore, if $A$ and $B$ are real symmetric matrices, we have

$$
\delta(|A, B|) = \sigma(|A, B|).
$$

In the next parts, we talk about computing such values for other classes of matrices.

#### 6.1 Calculation for normal matrices

We know that $\delta(A^*) = \delta(A)$ and $\delta(U^* AU) = \delta(A)$ for any unitary matrix $U$; hence, if $A$ is a normal matrix with the eigenvalue decomposition $A = U^* \Lambda U$ then $\delta(A) = \delta(A)$. We then have

$$
\sum_{i=1}^k \delta_i(A) = \min_{a \in \mathbb{C}} \max_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \sum_{j=1}^k |\lambda_{i_j} - a|.
$$

(6.1)

This problem can be reformulated as the following second-order cone programming problem:

$$
\begin{align*}
\min_{a \in \mathbb{C}, q \in \mathbb{R}, u, x \in \mathbb{R}^n} & \quad 1^T_n u + kq, \\
\text{subject to} & \quad u \geq x - q1_n \\
& \quad x \geq |\lambda - a1_n| \\
& \quad u \geq 0
\end{align*}
$$

(6.2)

where $|\cdot|$ denotes the element-wise modulus of a complex vector. Moreover, for some special cases such as Hermitian and skew-Hermitian matrices, the optimization problem has a closed form solution.

We can see that for any normal matrix $A$ we have $\delta_1(A) = \delta_2(A) = R$, where $R$ corresponds to the radius of the smallest circle containing all the eigenvalues; see Fig. 2. Also, we can see that the optimum $a$ which minimizes $\|A - a1_n\|_1$ is just the geometric median of the eigenvalues in the two-dimensional complex plane.
6.2 SDP representation for the discrepancy norms of square matrices

By Theorem 4.3, we observed that $\|A\|_{(k)}^\delta = \min_{\alpha \in \mathbb{C}} \|A - \alpha I_n\|_{(k)}$, which is a convex (and continuous) function in $\alpha$. Hence, we can find $\delta(A)$ by employing a series of non-smooth convex optimization methods. In this section, we propose a more efficient way to compute the discrepancy values of arbitrary square matrices.

Note that we can propose a SDP formulation for $\sum_{j=1}^n \delta_j(A)$. We know that

$$\sum_{i=1}^n \delta_i(A) = \max_{\substack{U^*U=I_n \\|V^*V=I_n \\| \text{tr}(U^*V)=0 \\text{tr}(K)=0}} \text{Re tr}(AV^*).$$

(6.3)

If we let $M = \begin{bmatrix} U^* & U \end{bmatrix} \begin{bmatrix} V^* & V \end{bmatrix}$ then

$$\sum_{i=1}^n \delta_i(A) = \max_{M=\begin{bmatrix} I_n & K \\ K^* & I_n \end{bmatrix}\succeq 0} \frac{1}{2} \text{Re tr} \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} M.$$

(6.4)

In order to derive SDP formulations for other $\sum_{j=1}^k \delta_j(A)$, we use the fact that for any Hermitian matrix $X$ with eigenvalues $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$ we have

$$t \geq \lambda_1(A) + \cdots + \lambda_k(A) \iff \exists (Z, s) : Z \succeq 0, Z + sI_n \succeq X, t \geq \text{Re } \left( \text{tr}(Z) + ks \right),$$

where $t$ is a non-negative real number, and $s \in \mathbb{C}$ and $Z$ can be complex valued in general. For an arbitrary complex square matrix $A$, by using the Hermitian dilation transform $X = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$, we have
the following SDP formulation for the Ky-Fan norm:

\[
\|A\|_k = \min_{Z \geq 0, t \geq 0, \sigma \in \mathbb{C}} t, \quad (6.5)
\]

\[
\begin{bmatrix}
0 & A \\
A^* & 0
\end{bmatrix} \succeq Z + sI_{2n}
\]

\[
t \geq \Re (\text{tr}(Z) + ks)
\]

where \(k = 1, \ldots, n\) and \(Z\) is a \(2n \times 2n\) complex matrix. Therefore, based on the definition of \(\sum_{i=1}^{k} \delta_i(A)\), we have

\[
\sum_{i=1}^{k} \delta_i(A) = \min_{Z \geq 0, t \geq 0, \sigma \in \mathbb{C}} t, \quad (6.6)
\]

\[
\begin{bmatrix}
0 & A - \sigma I_n \\
A^* - \sigma I_n & 0
\end{bmatrix} \succeq Z + sI_{2n}
\]

\[
t \geq \Re (\text{tr}(Z) + ks)
\]

where \(k = 1, \ldots, n\) and \(Z\) is a \(2n \times 2n\) complex matrix. Note that if matrix \(A\) is real, we can let \(Z\) and other optimization variables be real.

### 7 Discussion and extensions

Most of the results proven in the previous sections can be extended to any linear operator of the form \(A + B\), where \(A\) is a compact operator between Hilbert spaces, and \(\gamma \in \mathbb{C}\). For any compact operator \(A : H \to K\) between Hilbert spaces \(H\) and \(K\), we know that \(A\) always has countably many non-negative singular values, among which \(\sigma = 0\) is the only possible point of accumulation. For any finite \(k\), the operator norm \(\|A\|_k\) is finite, hence \(\delta_k(A + \gamma I)\) is well-defined. Furthermore, if the rank of the operator is not finite, we have \(\delta_k(A + \gamma I) \to 0\) as \(k \to \infty\); regardless of the fact that \(A\) belongs to the trace-class or not. This follows by the fact that \(\|A + \gamma I\|_\infty \leq \|A\|_k\) for any finite \(k \in \mathbb{N}\) and \(\sigma_k \to 0\) as \(k \to \infty\).

Let us end the paper with a conjecture for the discrepancy values.

**Conjecture 7.1.** For the \(n \times n\) square matrices \(A\) and \(B\) we have

\[
\sigma([A, B]) \preceq_w 2 \delta(A)\delta(B). \quad (7.1)
\]

Using Lemma 4.1, this conjecture can be equivalently stated that for any unitary matrix \(U\), we have

\[
\delta(U[A, B]) \preceq_w 2 \delta(A)\delta(B).
\]

For the \(2 \times 2\) matrix \(A = [A_{ij}]\), we can show that

\[
\delta_i(A) = \sqrt{\frac{1}{2}(A_{11} - A_{22})^2 + A_{12}^2 + A_{21}^2 + |A_{12} - A_{21}| \sqrt{(A_{12} + A_{21})^2 + (A_{11} - A_{22})^2}}
\]

for \(i = 1, 2\). By inspection, one can verify that \(\delta(A) = \sigma\left(\begin{bmatrix} A - \frac{A_{11} + A_{22}}{2} I_2 \end{bmatrix}\right)\). Thus we have

\[
\sigma([A, B]) = \sigma\left(\begin{bmatrix} A - \frac{A_{11} + A_{22}}{2} I_2, B - \frac{B_{11} + B_{22}}{2} I_2 \end{bmatrix}\right) \preceq_w 2 \delta(A)\delta(B).
\]

Therefore the conjecture is true for any two by two matrices. Furthermore, the conjecture has been proven for some general cases in this paper (see Remark 5.3 and Corollary 5.7). However, the conjecture in its full generality is an open problem.
References

1. Tsuyoshi Ando. Majorizations and inequalities in matrix theory. *Linear algebra and its applications*, 199:17–67, 1994.

2. C. Apostol, L.A. Fialkow, D.A. Herrero, and D. Voiculescu. *Approximation of Hilbert space operators II*. Pitman, 1984.

3. Richard Bellman. *Introduction to matrix analysis*. SIAM, 1997.

4. Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013.

5. Rajendra Bhatia and Fuad Kittaneh. Commutators, pinchings, and spectral variation. *Operators and matrices*, 2, 2008.

6. Rajendra Bhatia and Peter Šemrl. Orthogonality of matrices and some distance problems. *Linear algebra and its applications*, 287(1-3):77–85, 1999.

7. Albrecht Böttcher and David Wenzel. How big can the commutator of two matrices be and how big is it typically? *Linear algebra and its applications*, 403:216–228, 2005.

8. Fumio Hiai and Dénes Petz. *Introduction to matrix analysis and applications*. Springer Science & Business Media, 2014.

9. Omar Hirzallah. Commutator inequalities for hilbert space operators. *Linear algebra and its applications*, 431(9):1571–1578, 2009.

10. Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

11. Fuad Kittaneh. Inequalities for commutators of positive operators. *Journal of Functional Analysis*, 250(1):132–143, 2007.

12. Andrew V Knyazev and Merico E Argentati. Rayleigh–ritz majorization error bounds with applications to fem. *SIAM journal on matrix analysis and applications*, 31(3):1521–1537, 2010.

13. Zhiqin Lu and David Wenzel. Commutator estimates comprising the frobenius norm–looking back and forth. In *Large truncated Toeplitz matrices, Toeplitz operators, and related topics*, pages 533–559. Springer, 2017.

14. Pedro Massey, Demetrio Stojanoff, and Sebastián Zárate. Absolute variation of ritz values, principal angles and spectral spread. *arXiv preprint arXiv:2012.09018*, 2020.

15. Pedro Massey, Demetrio Stojanoff, and Sebastian Zarate. Norm inequalities for the spectral spread of hermitian operators. *arXiv preprint arXiv:2106.09092*, 2021.

16. Pedro Massey, Demetrio Stojanoff, and Sebastián Zárate. The spectral spread of hermitian matrices. *Linear algebra and its applications*, 616:19–44, 2021.

17. Joseph Stampfli. The norm of a derivation. *Pacific journal of mathematics*, 33(3):737–747, 1970.

18. David Wenzel. A strange phenomenon for the singular values of commutators with rank one matrices. *The electronic journal of linear algebra*, 30:649–669, 2015.

19. David Wenzel and Koenraad MR Audenaert. Impressions of convexity: an illustration for commutator bounds. *Linear algebra and its applications*, 433(11-12):1726–1759, 2010.

20. Xingzhi Zhan. *Matrix theory*, volume 147. American Mathematical Soc., 2013.
A The $X$ decomposition for Hermitian matrices

In this part, we propose a decomposition for the Hermitian matrices based on their discrepancy values and vectors. First, let us introduce a family of matrices.

**Definition A.1** ($X$ matrices). We call a square matrix which is allowed to have nonzero entries only on the main diagonal, or on the anti-diagonal, an $X$ matrix (or star matrix).

**Remark A.2.** The set of $X$ matrices are unitarily similar to matrices of the form $B_1 \oplus B_2 \oplus \cdots \oplus B_n$, where $B_i$ are two by two matrices, plus one scalar if it has odd dimensionality.

Let the notations $CX(a_1, \ldots, a_n)(b_1, \ldots, b_n)$ and $CX(a_1, \ldots, a_n)(c)(b_1, \ldots, b_n)$ denote the following centrosymmetric $X$ matrices, respectively:

\[
\begin{bmatrix}
    a_1 & 0 & \cdots & 0 & b_1 \\
    0 & \ddots & \ddots & \ddots & \ddots \\
    \vdots & \ddots & \ddots & \ddots & \ddots \\
    a_n & \ddots & \ddots & \ddots & 0 \\
    b_1 & 0 & \cdots & 0 & a_1
\end{bmatrix},
\begin{bmatrix}
    a_1 & 0 & \cdots & 0 & b_1 \\
    0 & \ddots & \ddots & \ddots & \ddots \\
    \vdots & \ddots & \ddots & \ddots & \ddots \\
    a_n & \ddots & \ddots & \ddots & 0 \\
    b_n & 0 & \cdots & 0 & a_1
\end{bmatrix}.
\]

**Proposition A.3** ($X$ decomposition). Any $n \times n$ Hermitian matrix $A$ can be decomposed as $A = UXV^*$, where $X$ is a centrosymmetric $X$ matrix; moreover, $U$ and $V$ are unitary matrices such that $U^*V = J_n$, the exchange matrix.

**Proof.** Let $R_n(\cdot)$ be the $n \times n$ orthogonal matrix defined in 5.12, then using the eigenvalue decomposition of $A$ we have

\[
A = Q\Lambda Q^* = QR_n(\frac{\pi}{4})(R_n(\frac{\pi}{4})^* \Lambda J_n R_n(\frac{\pi}{4})^*)(QJ_n R_n(\frac{\pi}{4})^*)^*.
\]

Now let $U = QR_n(\frac{\pi}{4})$, $V = QJ_n R_n(\frac{\pi}{4})^*$, and $X = R_n(\frac{\pi}{4})^* \Lambda J_n R_n(\frac{\pi}{4})^*$. One can easily verify all the conditions in the proposition are satisfied. \hfill \Box

In the previous proposition, we can show that the columns of $U$ and $V$ are vectors $x_i$ and $y_i$ in the definition of $\delta$ in 3.1. Also if $n$ is an even number we have

\[
X = CX(\delta_1(A), \delta_3(A), \ldots)(a_1^*(A), a_3^*(A), \ldots),
\]

and otherwise we have

\[
X = CX(\delta_1(A), \delta_3(A), \ldots)(a_n^*(A), a_3^*(A), \ldots),
\]

where $\delta_i(A) = (\lambda_i - \lambda_{n-i+1})/2$ and $a_i^*(A) = (\lambda_i + \lambda_{n-i+1})/2$. 

23