Linear Algebra and its Applications 592 (2020) 134–164

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

On some conjectures by Lu and Wenzel

Jianquan Ge [a,*,1], Fagui Li [a], Zhiqin Lu [b,2], Yi Zhou [a]

[a] *School of Mathematical Sciences, Laboratory of Mathematics and Complex Systems, Beijing Normal University, Beijing 100875, PR China*
[b] *Department of Mathematics, University of California, Irvine, Irvine, CA 92697, USA*

A R T I C L E   I N F O

Article history:
Received 19 August 2019
Accepted 21 January 2020
Available online 23 January 2020
Submitted by A. Böttcher

MSC:
15A45
15B57
53C42

Keywords:
DDVV-type inequalities
Böttcher-Wenzel inequality
Commutator

A B S T R A C T

In order to give a unified generalization of the BW inequality and the DDVV inequality, Lu and Wenzel proposed several conjectures in 2016. In this paper we discuss further these conjectures and put forward a couple of new conjectures which will be shown equivalent to one of the conjectures. In particular, we prove this conjecture and hence all conjectures in some special cases. Another conjecture is shown with a slightly bigger upper bound. In addition, the developed methods give some new simple proofs of the complex BW inequality and its condition for equality.

1. Introduction

In 2005, Böttcher and Wenzel [4] raised the so-called BW conjecture that if X, Y are real square matrices, then

* Corresponding author.
E-mail addresses: jqge@bnu.edu.cn (J. Ge), faguili@mail.bnu.edu.cn (F. Li), zlu@math.uci.edu (Z. Lu), zhou_yi@mail.bnu.edu.cn (Y. Zhou).
[1] The first author is partially supported by Beijing Natural Science Foundation (Z190003), NSFC (No. 11522103, 11331002) and by the Fundamental Research Funds for the Central Universities of China.
[2] The third author is partially supported by the National Science Foundation of USA (DMS-19-08513).

https://doi.org/10.1016/j.laa.2020.01.027
0024-3795/

$$\|XY - YX\|^2 \le 2\|X\|^2\|Y\|^2,$$

where $\|X\| = \sqrt{\operatorname{Tr} XX^*}$ is the Frobenius norm (with X^* denoting the conjugate transpose of X). After some steps and several fundamentally different proofs (cf. [2,5,21,26]), the inequality is now known to be true in the complex case. Also, analogues with, e.g., Schatten norms were investigated (cf. [6,9,27,28]). We recommend [7] and [24] for a comprehensive overview on the developments.

In comparison with the BW inequality that estimates the Frobenius norm of the commutator between two arbitrary matrices, the DDVV inequality involves commutators of arbitrary many, but only real symmetric matrices. The original formulation in the language of submanifold theory was posed in 1999 by De Smet, Dillen, Verstraelen and Vrancken [11], but later transformed into the matrix-algebraic inequality

$$\sum_{\alpha,\beta=1}^{m} \| [B_\alpha, B_\beta] \|^2 \le c \left(\sum_{\alpha=1}^{m} \|B_\alpha\|^2 \right)^2,$$

(see [12]). The classical result is $c = 1$. It was shown by the authors' groups independently and differently (cf. [14,21,10,15,20]). Also other classes of matrices were analyzed (cf. [13, 16,17]), resulting into various DDVV-type inequalities with other values of c depending on the matrix structure.

With the BW inequality and the DDVV inequality on both hands, Lu and Wenzel ([23,24]) summarized the commutator estimates and considered a unified generalization of them. They started with the following three conjectures and an open question in the space $M(n, \mathbb{K})$ of $n \times n$ matrices over the field $\mathbb{K}$.

Conjecture 1. *Let $B_1, \cdots, B_m \in M(n, \mathbb{R})$ subject to*

$$\operatorname{Tr}\Big(B_\alpha[B_\gamma, B_\beta]\Big) = 0$$

for any $1 \le \alpha, \beta, \gamma \le m$, then

$$\sum_{\alpha,\beta=1}^{m} \| [B_\alpha, B_\beta] \|^2 \le \left(\sum_{\alpha=1}^{m} \|B_\alpha\|^2 \right)^2. \tag{1.1}$$

Conjecture 2. (*LW Conjecture*). *Let $B, B_2, \cdots, B_m \in M(n, \mathbb{R})$ with*

(i) $\operatorname{Tr}(B_\alpha B_\beta^*) = 0$ *(i.e., $B_\alpha \perp B_\beta$) for any $\alpha \ne \beta$;*
(ii) $\operatorname{Tr}\Big(B_\alpha[B, B_\beta]\Big) = 0$ *for any $2 \le \alpha, \beta \le m$.*

Then

$$\sum_{\alpha=2}^{m} \| [B, B_\alpha] \|^2 \le \left(\max_{2\le\alpha\le m} \|B_\alpha\|^2 + \sum_{\alpha=2}^{m} \|B_\alpha\|^2 \right) \|B\|^2. \tag{1.2}$$

Note that the number m cannot be arbitrarily large.

Conjecture 3. *For $X \in M(n,\mathbb{R})$ with $\|X\| = 1$, let T_X be the linear map on $M(n,\mathbb{R})$ defined by $T_X(Y) = [X^*, [X, Y]]$ and $\lambda_1(T_X) \geq \lambda_2(T_X) \geq \lambda_3(T_X) \cdots$ be its eigenvalues. Then*

$$\lambda_1(T_X) + \lambda_3(T_X) \leq 3.$$

Question 1. What is the upper bound of $\sum\limits_{i=1}^{k} \lambda_{2i-1}(T_X)$?

If $k = 1$, the bound is 2 by the BW inequality, i.e., $\lambda_1(T_X) \leq 2$, since we have

$$\lambda_1(T_X) = \max_{\|Y\|=1} \langle T_X Y, Y\rangle = \max_{\|Y\|=1} \|[X, Y]\|^2 \leq 2.$$

If $k = 2$, the bound is supposed to be 3 by Conjecture 3. How are all these conjectures and the known inequalities connected? When restricted to real symmetric matrices, Conjecture 1 reduces to the DDVV inequality. It turns out that not only the BW inequality and the DDVV inequality but also both Conjectures 1 and 3 are implied by Conjecture 2 (cf. [23]). Moreover, we will show that Conjecture 2 is equivalent to assigning $k+1$ as the upper bound of $\sum_{i=1}^{k} \lambda_{2i-1}(T_X)$ for $k \geq 1$, which is nothing but the new Conjecture 4 due to the fact that $\lambda_{2i-1}(T_X) = \lambda_{2i}(T_X)$ for any i (see Proposition 2.2 (c)). Hence, Conjecture 2 (as well as the equivalent Conjectures 4–6) takes exactly the role of a unified generalization of the BW inequality and the DDVV inequality for real matrices. We call it the Fundamental Conjecture of Lu and Wenzel, or simply the (real) LW Conjecture.

First consider $\mathbb{K} = \mathbb{R}$.

Conjecture 4. *For $X \in M(n,\mathbb{K})$ with $\|X\| = 1$, we have*

$$\sum_{i=1}^{2k} \lambda_i(T_X) \leq 2k + 2, \quad k = 1, \cdots, \left[\frac{n^2}{2}\right]. \tag{1.3}$$

In fact, the sum $\sum_{i=1}^{2k} \lambda_i(T_X)$ in Conjecture 4 cannot exceed $2n$. We explain this by introducing the following Conjecture 5 which looks stronger but in fact is equivalent to Conjecture 4.

Before continuing, we adopt some notation from [29] and [19]. Let $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, where $x_1^{\downarrow} \geq x_2^{\downarrow} \geq \cdots \geq x_n^{\downarrow}$. For $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n)$ in $\mathbb{R}^n$, we say that x is *weakly majorized* by y, written as $x \prec y$, if

$$\sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow}, \quad k = 1, 2, \cdots, n.$$

Multisets are an extension of the set definition by allowing elements to appear more than one time. A finite multiset is often represented as $\{a_1^{m(a_1)}, a_2^{m(a_2)}, \ldots, a_n^{m(a_n)}\}$, where $m(a_k) \in \mathbb{N}$ is the multiplicity, that is, the number of occurrences. For example, the multiset $\{a, a, b\}$ is written as $\{a^2, b\}$. We use ordered vectors and abbreviating multisets synonymously.

Conjecture 5. *For $X \in M(n, \mathbb{K})$ with $\|X\| = 1$, the set $\lambda(T_X)$ of decreasingly ordered eigenvalues of T_X is weakly majorized by the multiset $\{2^2, 1^{2n-4}, 0^{(n-1)^2+1}\}$.*

As promised, this becomes

$$\sum_{i=1}^{2k} \lambda_i(T_X) \leq 2n, \quad \text{for } k \geq n,$$

seemingly strengthening the assertion of Conjecture 4. Another equivalent conjecture that also appears to be stronger is the following Conjecture 6. It omits the second assumption of Conjecture 2, at the price of a factor 2 in the bound.

Conjecture 6. *Let $B, B_2, \cdots, B_m \in M(n, \mathbb{K})$ with $\mathrm{Tr}(B_\alpha B_\beta^*) = 0$ for any $2 \leq \alpha \neq \beta \leq m$. Then*

$$\sum_{\alpha=2}^{m} \| [B, B_\alpha] \|^2 \leq \left(2 \max_{2 \leq \alpha \leq m} \|B_\alpha\|^2 + \sum_{\alpha=2}^{m} \|B_\alpha\|^2 \right) \|B\|^2.$$

We summarize the relations of these conjectures in the following theorem. Remember $\mathbb{K} = \mathbb{R}$ so far.

Theorem 1.1. *The following relations hold in these conjectures.*

(i) *Conjectures 2, 4, 5, and 6 are equivalent to each other.*
(ii) *If one of the conjectures above is true, then Conjectures 1 and 3 hold.*

Since the BW inequality (resp. the DDVV inequality) holds also for complex (resp. complex symmetric) matrices (cf. [5], [17]), we can also consider the same conjectures as above in the complex version $\mathbb{K} = \mathbb{C}$. In fact we will prove the relations of Theorem 1.1 between these conjectures in the complex version. Hence we call Conjecture 2 for complex matrices the complex LW Conjecture. Obviously, the complex LW Conjecture implies the real LW Conjecture.

In this paper, we prove the complex LW Conjecture in some special cases which we conclude in the following.

Theorem 1.2. *The complex LW Conjectures 4 and 5 (and because of Theorem 1.1 all conjectures of this paper) are true in one of the following cases:*

(i) $X \in M(n, \mathbb{C})$ *is a normal matrix;*
(ii) $\operatorname{rank} X = 1$*;*
(iii) $n = 2$ *or* $n = 3$*.*

For the Conjectures 4 and 5 in general we are able to get some weakened results as follows.

Theorem 1.3. *For* $X \in M(n, \mathbb{C})$ *with* $\|X\| = 1$*, we have*

$$\lambda_1(T_X) + \lambda_3(T_X) \leq \frac{4 + \sqrt{10}}{2} \approx 3.58.$$

Theorem 1.4. *For* $X \in M(n, \mathbb{C})$ *with* $\|X\| = 1$*, we have*

$$\sum_{t=1}^{2k} \lambda_i(T_X) \leq 2k + 1 + 2\sqrt{k}, \quad k = 1, \cdots, \left[\frac{n^2}{2}\right].$$

In Section 2 we prepare several useful lemmas and properties of T_X. It turns out that the methods we developed in the study of the conjectures above lead us to some new simple proofs of the complex BW inequality and its condition for equality, which we will discuss in Section 3. It is just the first eigenvalue estimate $\lambda_1(T_X) \leq 2$, the basic case $k = 1$ of the complex LW Conjecture 4. In Section 4 we prove the equivalences in Theorem 1.1 in the complex version. In Section 5 we prove the special cases given in Theorem 1.2 and the partial results of Theorems 1.3 and 1.4.

Although the inequalities we study in this paper are matrix inequalities, it is not hard to generalize them as inequalities of bounded operators on separable Hilbert spaces. In quantum physics, these inequalities are related to the *Uncertainty Principle*, or more precisely, the Robertson-Schrödinger relations. The classical Uncertainty Principle, in our notations, can be formulated by

$$\|[A, B]\|_{OP}^2 \leq 4\, \|A\|_{OP}^2 \cdot \|B\|_{OP}^2,$$

where $\|\cdot\|_{OP}$ is the operator norm and A, B are self-adjoint. In this context, the BW-type inequality can be interpreted as a mathematical generalization of the Uncertainty Principle, because we don't need to assume A, B are self-adjoint. There are some literatures in physics providing various of generalizations of the Uncertainty Principle; see [25] for example. In our paper, we study the optimal version of all these inequalities in the Frobenius norm.

2. Preliminaries

In this section, we will introduce some necessary notations and lemmas which are interesting in themselves. To avoid needless duplication, we discuss the complex version directly so as to include the real version.

Let T be a linear mapping on a complex N-dimensional vector space V with Hermitian inner product $\langle \cdot, \cdot \rangle$. In this paper, we always denote by

$$\lambda(T) := \{\lambda_1(T) \geq \cdots \geq \lambda_N(T)\}, \quad \sigma(T) := \{\sigma_1(T) \geq \cdots \geq \sigma_N(T) \geq 0\}$$

the ordered sets of real eigenvalues (if available) and singular values of T respectively, where singular values are square roots of eigenvalues of T^*T.

By elementary linear algebra, we have the following statements.

Lemma 2.1. *Suppose T be self-dual and positive semi-definite.*

(a) *The (geometric) multiplicity of each positive eigenvalue of T is even if and only if there exists a mapping S such that*

$$T = S^*S = -S^2.$$

In addition, $Tx = 0$ if and only if $Sx = 0$.

Assume T furthermore has even multiplicities (i.e., $\lambda_{2i-1}(T) = \lambda_{2i}(T)$ for any i with $\lambda_{2i-1}(T) > 0$), and S be the unitary skew-symmetric mapping as in (a).

(b) *Let $y \in V$ with $|y| = 1$. Then*

$$\left\langle T\frac{Sy}{|Sy|}, \frac{Sy}{|Sy|} \right\rangle \geq \langle Ty, y \rangle.$$

(c) *Let $W \subseteq V$ be a complex m-dimensional isotropic subspace of S, i.e., $S(W) \subset W^\perp$ ($\langle Sw_1, w_2 \rangle = 0$ for any $w_1, w_2 \in W$). Then we have*

$$\operatorname{Tr} T\big|_W \leq \operatorname{Tr} T\big|_{S(W)}, \quad \operatorname{Tr} T\big|_W \leq \sum_{i=1}^{m} \lambda_{2i-1}(T).$$

Proof. First we prove (a), the sufficiency is clear. Now suppose that there are g distinct positive eigenvalues $\lambda(T) = \{t_1 = s_1^2 > \cdots > t_g = s_g^2 > 0\}$ with multiplicities $2n_1, \cdots, 2n_g$, and denote by $r = 2\sum_{j=1}^{g} n_j$ the rank of T. Then we can diagonalize T by a unitary matrix U as

$$T = U \operatorname{diag}\Big(t_1 I_{2n_1}, \cdots, t_g I_{2n_g}, O_{N-r}\Big) U^*,$$

where O_{N-r} denotes the zero matrix of order $N - r$. Then the required unitary skew-symmetric matrix can be defined as

$$S := U \operatorname{diag}\left(\begin{pmatrix} O & -s_1 I_{n_1} \\ s_1 I_{n_1} & O \end{pmatrix}, \cdots, \begin{pmatrix} O & -s_g I_{n_g} \\ s_g I_{n_g} & O \end{pmatrix}, O_{N-r}\right) U^*.$$

Next we prove (b). Since $T = S^*S = -S^2$, the inequality in (b) is equivalent to

$$\langle T^2 y, y\rangle \geq \langle Ty, y\rangle^2.$$

Let $\{e_i\}_{i=1}^N$ be an orthonormal basis of V such that e_i is a unit eigenvector corresponding to $\lambda_i(T)$. Setting $y = \sum_{i=1}^N y_i e_i$, then $\sum_{i=1}^N y_i^2 = 1$ and we have

$$\begin{aligned}\langle T^2 y, y\rangle &= \sum_{i=1}^N y_i^2 \lambda_i^2(T) = \left(\sum_{i=1}^N y_i^2 \lambda_i^2(T)\right)\left(\sum_{i=1}^N y_i^2\right) \\ &\geq \left(\sum_{i=1}^N y_i^2 \lambda_i(T)\right)^2 = \langle Ty, y\rangle^2.\end{aligned}$$

To prove (c), we will find a suitable basis to compare the traces by using (b). Let $\{E_i\}_{i=1}^N$ be an orthonormal basis of V such that $\{E_i\}_{i=1}^m$ is a basis of W, and under this basis we identify $V \cong \mathbb{C}^N$. Denote

$$\operatorname{rank}(SE_1, \cdots, SE_m) = \dim S(W) =: k \leq m.$$

Assume $k \geq 1$, otherwise we have $S\big|_W = 0$ and thus $\operatorname{Tr} T\big|_W = 0$ by (a). By singular value decomposition, there exist $P \in U(N)$ and $Q \in U(m)$ such that

$$P^*(SE_1, \cdots, SE_m)Q = \Lambda =: \begin{pmatrix} \widetilde{\Lambda}_{k\times k} & O \\ O & O \end{pmatrix}_{N\times m},$$

where $\widetilde{\Lambda} =: \operatorname{diag}(\Lambda_1, \cdots, \Lambda_k)$, $\Lambda_i > 0$ for $1 \leq i \leq k$. Setting

$$P\Lambda =: (F_1, \cdots, F_m),$$

we have $\langle F_i, F_j\rangle = \Lambda_i \Lambda_j \delta_{ij}$ for $1 \leq i, j \leq k$ and $F_i = 0$ for $i > k$. Thus $\widetilde{F}_i := \Lambda_i^{-1} F_i$ gives an orthonormal basis of $S(W)$. Let

$$(\widetilde{E}_1, \cdots, \widetilde{E}_m) := (E_1, \cdots, E_m)Q,$$

then $\{\widetilde{E}_i\}_{i=1}^m$ is an orthonormal basis of W and satisfies

$$(F_1, \cdots, F_m) = P\Lambda = (SE_1, \cdots, SE_m)Q = (S\widetilde{E}_1, \cdots, S\widetilde{E}_m).$$

Therefore, (b) implies

$$\operatorname{Tr} T\big|_W = \sum_{i=1}^m \langle T\widetilde{E}_i, \widetilde{E}_i\rangle \leq \sum_{i=1}^k \langle T\widetilde{F}_i, \widetilde{F}_i\rangle = \operatorname{Tr} T\big|_{S(W)}.$$

Since $S(W) \subset W^\perp$, $\{\widetilde{E}_i\}_{i=1}^m \bigcup \{\widetilde{F}_i\}_{i=1}^k$ is an orthonormal basis of $W \oplus S(W)$. Hence,

$$\operatorname{Tr} T\big|_W + \operatorname{Tr} T\big|_{S(W)} = \operatorname{Tr} T\big|_{W\oplus S(W)} \leq \sum_{i=1}^{m+k} \lambda_i(T) \leq \sum_{i=1}^{2m} \lambda_i(T),$$

$$\operatorname{Tr} T\big|_W \leq \frac{1}{2}\sum_{i=1}^{2m} \lambda_i(T) = \sum_{i=1}^{m} \lambda_{2i-1}(T).$$

The proof is complete. □

Now we consider the linear operator T_X as in Conjecture 3. More specifically, for any $n \times n$ complex matrix X with $\|X\| = 1$, we define

$$T_X : M(n, \mathbb{C}) \longrightarrow M(n, \mathbb{C}),\ Y \longmapsto [X^*, [X, Y]]. \tag{2.1}$$

It turns out that T_X is exactly an operator of the same type as T in the preceding lemma with $V = M(n, \mathbb{C})$, $\dim V = n^2 =: N$ (cf. [21]). For the sake of completeness, we repeat some properties of T_X, which was the key to Lu's method for proving the BW and DDVV inequalities.

Proposition 2.2. *T_X has the following properties:*

(a) T_X is an self-dual and positive semi-definite linear map.
(b) The set of eigenvalues $\lambda(T_X) := \{\lambda_1(T_X) \geq \cdots \geq \lambda_N(T_X)\}$ is invariant under unitary congruences of X.
(c) The multiplicity of each positive eigenvalue of T_X is even, i.e., $\lambda_{2i-1}(T_X) = \lambda_{2i}(T_X)$ for any i with $\lambda_{2i-1}(T_X) > 0$.

Proof. For (a), the real version proof given in [21, Lemma 3] is easily turned complex by using adjoints instead of transposes.

Part (b) follows immediately from

$$T_{U^*XU}(U^*YU) = U^*(T_X Y)U, \quad \text{for } U \in U(n). \tag{2.2}$$

Finally we prove (c). Let $\lambda > 0$ be a positive eigenvalue of T_X and E_λ be its eigenspace. We will show that the complex dimension of E_λ is even.

Define a quasi-linear map by

$$\widetilde{S}_X : M(n, \mathbb{C}) \longrightarrow M(n, \mathbb{C}),\ Y \longmapsto [X, Y]^*.$$

Then it follows easily that $\widetilde{S}_X(zY) = \bar{z}\widetilde{S}_X(Y)$ for $z \in \mathbb{C}$, $\widetilde{S}_X$ is anti-self-dual and $T_X = -\widetilde{S}_X^2$ because

$$\mathrm{Tr}\Big(Y_2(\widetilde{S}_X Y_1)^*\Big) = \mathrm{Tr}\Big(Y_2[X,Y_1]\Big) = \mathrm{Tr}\Big(X[Y_1,Y_2]\Big) = -\,\mathrm{Tr}\Big(Y_1(\widetilde{S}_X Y_2)^*\Big),$$
$$-\widetilde{S}_X^2 Y = -[X,[X,Y]^*]^* = [X^*,[X,Y]] = T_X Y.$$

Now for any eigenvector $Y \in E_\lambda$, i.e., $T_X Y = \lambda Y$, we claim that $\widetilde{S}_X Y$ is also an eigenvector in E_λ which is $\mathbb{C}$-independent (even $\mathbb{C}$-orthogonal) to Y. In fact, since $T_X = -\widetilde{S}_X^2$ we have

$$T_X \widetilde{S}_X Y = \widetilde{S}_X T_X Y = \lambda \widetilde{S}_X Y, \quad \|\widetilde{S}_X Y\|^2 = \langle T_X Y, Y\rangle = \lambda \|Y\|^2 > 0,$$
$$\mathrm{Tr}\Big(Y(\widetilde{S}_X Y)^*\Big) = \mathrm{Tr}\Big(Y[X,Y]\Big) = 0, \quad \text{and thus} \quad \langle Y, \widetilde{S}_X Y\rangle = \langle \mathbf{i}\, Y, \widetilde{S}_X Y\rangle = 0,$$

where $\mathbf{i}$, as usual, is the imaginary unit.

For $k \geq 1$, suppose that $\mathrm{Span}_{\mathbb{C}}\{Y_i, \widetilde{S}_X Y_i\}_{i=1}^k \subset E_\lambda$ and $Y_{k+1} \in E_\lambda$ is orthogonal to $\mathrm{Span}_{\mathbb{C}}\{Y_i, \widetilde{S}_X Y_i\}_{i=1}^k$. Then it suffices to prove

$$\widetilde{S}_X Y_{k+1} \perp \mathrm{Span}_{\mathbb{C}}\{Y_i, \widetilde{S}_X Y_i\}_{i=1}^k.$$

This is easily verified as follows:

$$\mathrm{Tr}\Big(Y_i(\widetilde{S}_X Y_{k+1})^*\Big) = -\,\mathrm{Tr}\Big(Y_{k+1}(\widetilde{S}_X Y_i)^*\Big) = 0,$$
$$\mathrm{Tr}\Big(\widetilde{S}_X Y_i(\widetilde{S}_X Y_{k+1})^*\Big) = \mathrm{Tr}\Big(Y_{k+1}(T_X Y_i)^*\Big) = \lambda\,\mathrm{Tr}\Big(Y_{k+1}(Y_i)^*\Big) = 0.$$

The proof is complete. □

Specifying T and S in Lemma 2.1, we can define a unitary skew-symmetric linear operator S_X on $V = M(n,\mathbb{C})$ such that $T_X = S_X^* S_X = -S_X^2$ as follows.[3] Taking an orthonormal basis $\{v_i\}_{i=1}^N$ of V such that v_i is an eigenvector of the eigenvalue $\lambda_i(T_X)$, we define S_X on this basis by $S_X(v_i) := \widetilde{S}_X v_i = [X, v_i]^*$ and then extend it linearly to the whole space as

$$S_X : M(n,\mathbb{C}) \longrightarrow M(n,\mathbb{C}),\ Y = \sum_{i=1}^N y_i v_i \longmapsto \sum_{i=1}^N y_i [X, v_i]^*. \tag{2.3}$$

As in Lemma 2.1 (a), suppose that there are g distinct positive eigenvalues $\lambda(T_X) = \{t_1 = s_1^2 > \cdots > t_g = s_g^2 > 0\}$ with multiplicities $2n_1, \cdots, 2n_g$, and denote by $r = 2\sum_{j=1}^g n_j$ the rank of T_X. By the proof of Proposition 2.2 (c), we see that $\widetilde{S}_X$ preserves the eigenspaces of T_X and we can choose the orthonormal basis such that the first r vectors are ordered as:

$$v_1^j, \cdots, v_{n_j}^j, \widetilde{S}_X v_1^j / \|\widetilde{S}_X v_1^j\|, \cdots, \widetilde{S}_X v_{n_j}^j / \|\widetilde{S}_X v_{n_j}^j\|, \quad j = 1, \cdots, g,$$

[3] Notice that $\widetilde{S}_X$ is not $\mathbb{C}$-linear and hence we need to define S_X instead.

where v_i^j's are unit eigenvectors of T_X corresponding to the eigenvalue t_j and thus $\|\widetilde{S}_X v_1^j\| = \cdots = \|\widetilde{S}_X v_{n_j}^j\| = s_j$. Under the special basis above, the linear operator S_X can be represented by the real skew-symmetric matrix

$$S_X = \mathrm{diag}\left(\begin{pmatrix} O & -s_1 I_{n_1} \\ s_1 I_{n_1} & O \end{pmatrix}, \cdots, \begin{pmatrix} O & -s_g I_{n_g} \\ s_g I_{n_g} & O \end{pmatrix}, O_{N-r}\right),$$

while T_X is represented by

$$T_X = \mathrm{diag}\left(t_1 I_{2n_1}, \cdots, t_g I_{2n_g}, O_{N-r}\right).$$

One can also reorder the basis in the way $v_{2i} = \widetilde{S}_X v_{2i-1}/\|\widetilde{S}_X v_{2i-1}\|$ such that

$$S_X = \mathrm{diag}\left(I_{n_1} \otimes \begin{pmatrix} 0 & -s_1 \\ s_1 & 0 \end{pmatrix}, \cdots, I_{n_g} \otimes \begin{pmatrix} 0 & -s_g \\ s_g & 0 \end{pmatrix}, O_{N-r}\right). \tag{2.4}$$

Hence, Lemma 2.1 (c) is suitable for the pair (T_X, S_X) and will be applied in the proof of the equivalence of Conjecture 2 and Conjecture 4.

We will also need the following notations and useful lemmas. Let Vec be the canonical isomorphism from $M(n,\mathbb{C})$ to $\mathbb{C}^N$, i.e.,

$$\mathrm{Vec} : M(n,\mathbb{C}) \longrightarrow \mathbb{C}^N,\ X = (x_{ij}) \longmapsto (x_{11}, \cdots, x_{n1}, x_{12}, \cdots, x_{n2}, \cdots, x_{1n}, \cdots, x_{nn})^t,$$

where X^t is the transpose of X. Using Kronecker product of matrices, we have

Lemma 2.3. *[18]* $\mathrm{Vec}(AYB) = (B^t \otimes A)\,\mathrm{Vec}(Y)$.

Moreover, Vec is an isometry since $\langle X, Y\rangle = \langle \mathrm{Vec}(X), \mathrm{Vec}(Y)\rangle$, and thus we can calculate the eigenvalues of T_X by

$$\lambda(T_X) = \lambda\left(\mathrm{Vec} \circ T_X \circ (\mathrm{Vec})^{-1}\right).$$

Proposition 2.4. *$\lambda(T_X) = \lambda(K_X^* K_X) = \lambda(K_1 + K_2)$, where $K_X = I \otimes X - X^t \otimes I$ and $K_1 = I \otimes X^* X + \overline{X} X^t \otimes I,\ K_2 = -X^t \otimes X^* - \overline{X} \otimes X$.*

Proof. Define $\Phi_X(Y) := [X, Y]$, by Lemma 2.3, we have

$$\mathrm{Vec}\left(\Phi_X(Y)\right) = K_X\,\mathrm{Vec}(Y),$$

where K_X is regarded as a linear operator on $\mathbb{C}^N$, or equivalently as a $N \times N$ matrix. Then it is easily seen that $K_{X^*} = K_X^*$ and as compositions of the operators we have

$$\mathrm{Vec} \circ \Phi_X \circ (\mathrm{Vec})^{-1} = K_X, \quad T_X = \Phi_{X^*} \circ \Phi_X.$$

In particular, we have

$$\mathrm{Vec} \circ T_X \circ (\mathrm{Vec})^{-1} = K_{X^*} K_X = K_X^* K_X,$$

hence

$$\lambda(T_X) = \lambda(K_X^* K_X).$$

By direct calculation, we have $K_X^* K_X = K_1 + K_2$, where K_1 and K_2 are Hermitian matrices. □

Corollary 2.5. *For $X \in M(n, \mathbb{C})$ with $\|X\| = 1$, we have $\operatorname{Tr} T_X = 2n - 2|\operatorname{Tr} X|^2$. In particular, for $n = 2$, $\lambda_1(T_X) = \lambda_2(T_X) = 2 - |\operatorname{Tr} X|^2$ and $\lambda_3(T_X) = \lambda_4(T_X) = 0$.*

Proof. It follows immediately from Proposition 2.4 that

$$\operatorname{Tr} T_X = \operatorname{Tr} K_1 + \operatorname{Tr} K_2 = 2n\|X\|^2 - 2|\operatorname{Tr} X|^2 = 2n - 2|\operatorname{Tr} X|^2.$$

For $n = 2$, the conclusion follows from Proposition 2.2 (c) and the fact that $T_X X = 0$ and thus T_X must have at least one zero eigenvalue. □

To end this section, we cite two useful properties about eigenvalues of Kronecker product and sum of two matrices.

Lemma 2.6. *[29] Let A and B be $m \times m$ and $n \times n$ complex matrices with eigenvalues $\alpha_1, \cdots, \alpha_m$ and $\beta_1, \cdots, \beta_n$, respectively.*

(a) *([29, Theorem 4.8]) The eigenvalues of $A \otimes B$ are*

$$\alpha_i \beta_j, 1 \le i \le m, 1 \le j \le n,$$

and the eigenvalues of $A \otimes I_n + I_m \otimes B$ are

$$\alpha_i + \beta_j, 1 \le i \le m, 1 \le j \le n.$$

(b) *([29, Theorem 8.18]) Suppose A, B are $n \times n$ Hermitian matrices and $\alpha_1 \ge \cdots \ge \alpha_n$, $\beta_1 \ge \cdots \ge \beta_n$. Let $C = A + B$ with eigenvalues $\gamma_1 \ge \cdots \ge \gamma_n$. Then for any sequence $1 \le i_1 < \cdots < i_k \le n$,*

$$\sum_{t=1}^{k} \alpha_{i_t} + \sum_{t=1}^{k} \beta_{n-k+t} \le \sum_{t=1}^{k} \gamma_{i_t} \le \sum_{t=1}^{k} \alpha_{i_t} + \sum_{t=1}^{k} \beta_t.$$

3. Some new proofs of the complex BW inequality

In this section, we will give some new simple proofs of the complex BW inequality by eigenvalue estimates of T_X for $X \in M(n,\mathbb{C})$ with $\|X\| = 1$. Each estimate implies $\lambda_1(T_X) \leq 2$ and thus the complex BW inequality, since for $\|Y\| = 1$,

$$\|[X,Y]\|^2 \leq \max_{\|Y\|=1} \|[X,Y]\|^2 = \max_{\|Y\|=1} \langle T_X Y, Y\rangle = \lambda_1(T_X) \leq 2 = 2\|X\|^2\|Y\|^2.$$

As a matter of fact, the core of our approach lies in the fact that the multiplicity of positive eigenvalues of T_X is even (Proposition 2.2 (c)).

Theorem 3.1. *Let $X = A + B \in M(n,\mathbb{C})$ be the canonical decomposition and $\|X\| = 1$, where A is Hermitian, B is skew-Hermitian. Then*

$$\lambda_1(T_X) \leq 2\Big(\max_{i,j}\{-a_ia_j\} + \max_{i,j}\{-b_ib_j\}\Big) + \Big(\sigma_1^2(X) + \sigma_2^2(X)\Big) \leq 2,$$

where $\sigma_1(X) \geq \cdots \geq \sigma_n(X)$ are singular values of X and $\lambda(A) = \{a_1, \cdots, a_n\}$, $\lambda(B) = \{b_1\mathbf{i}, \cdots, b_n\mathbf{i}\}$ are eigenvalues of A, B respectively.

Proof. Let $\sigma_1(X) \geq \cdots \geq \sigma_n(X)$ be singular values of X, then

$$\lambda\left(X^*X\right) = \lambda(\overline{X}X^t) = \{\sigma_1^2(X), \cdots, \sigma_n^2(X)\}.$$

We decompose T_X with help of Proposition 2.4. Hence for $K_1 = I \otimes X^*X + \overline{X}X^t \otimes I$, we have by Lemma 2.6 ($a$)

$$\lambda(K_1) = \{\sigma_i^2(X) + \sigma_j^2(X) : 1 \leq i,j \leq n\}. \tag{3.1}$$

In particular, $\lambda_2(K_1) = \sigma_1^2(X)+\sigma_2^2(X)$. And for K_2, we calculate by inserting $X = A+B$ that

$$K_2 = -X^t \otimes X^* - \overline{X} \otimes X = 2\left(B^t \otimes B - A^t \otimes A\right).$$

Then by Lemma 2.6 (a),

$$\lambda(-A^t \otimes A) = \{-a_ia_j : 1 \leq i,j \leq n\},$$
$$\lambda(B^t \otimes B) = \{-b_ib_j : 1 \leq i,j \leq n\},$$

where $\lambda(A) = \{a_i : a_1 \geq \cdots \geq a_n\}$; $\lambda(B) = \{b_j\mathbf{i} : b_1 \geq \cdots \geq b_n\}$. Therefore

$$\begin{aligned}\lambda_1(-A^t \otimes A) &= \max_{i,j}\{-a_i a_j\} = \max\{\max_{i\neq j}\{-a_i a_j\}, \max_i\{-a_i^2\}\} \\ &\leq \max\{\max_{i\neq j}\{-a_i a_j\}, 0\} \leq \max_{i\neq j}\{|a_i a_j|\} \\ &\leq \frac{1}{2}\max_{i\neq j}\{a_i^2 + a_j^2\} \leq \frac{1}{2}\|A\|^2.\end{aligned} \tag{3.2}$$

Similarly

$$\lambda_1(B^t \otimes B) = \max_{i,j}\{-b_i b_j\} \leq \frac{1}{2}\max_{i\neq j}\{b_i^2 + b_j^2\} \leq \frac{1}{2}\|B\|^2. \tag{3.3}$$

Since $B^t \otimes B$ and $-A^t \otimes A$ are Hermitian, by Lemma 2.6 (b), we have

$$\begin{aligned}\lambda_1(K_2) &\leq 2\left(\lambda_1(B^t \otimes B) + \lambda_1(-A^t \otimes A)\right) = 2\Big(\max_{i,j}\{-a_i a_j\} + \max_{i,j}\{-b_i b_j\}\Big) \\ &\leq \|A\|^2 + \|B\|^2 = \|X\|^2 = 1.\end{aligned} \tag{3.4}$$

Moreover, for $K_X^* K_X = K_1 + K_2$ in Proposition 2.4, again by Lemma 2.6 (b) we have

$$\lambda_2(K_X^* K_X) \leq \lambda_2(K_1) + \lambda_1(K_2) \leq \sigma_1^2(X) + \sigma_2^2(X) + \|X\|^2 \leq 2\|X\|^2.$$

Finally by Proposition 2.2 (c), we have the desired estimate

$$\lambda_1(T_X) = \lambda_2(T_X) = \lambda_2(K_X^* K_X) \leq 2\|X\|^2 = 2.$$

The proof is complete. □

For $X \in M(n, \mathbb{C})$ with $\|X\| = 1$, we have the following characterization of when $\lambda_1(T_X)$ attains the upper bound 2.

Theorem 3.2. *$\lambda_1(T_X) = 2$ if and only if $X = U\,\mathrm{diag}(X_0, O_{n-2})U^*$ for some $U \in U(n)$, where $X_0 \in M(2, \mathbb{C})$ and $\mathrm{Tr}(X_0) = 0$.*

Proof. We first prove the necessity. All the inequalities in the proof of Theorem 3.1 achieve equality when $\lambda_1(T_X) = 2$. Thus by the equality conditions of (3.2) and (3.3), we have $a_1 = -a_n =: a \geq 0$, $b_1 = -b_n =: b \geq 0$, and $a_i = b_i = 0$ for $1 < i < n$. Therefore,

$$\begin{aligned}\lambda(-A^t \otimes A) &= \{a^2, a^2, 0, \cdots, 0, -a^2, -a^2\}, \\ \lambda(B^t \otimes B) &= \{b^2, b^2, 0, \cdots, 0, -b^2, -b^2\},\end{aligned}$$

and there exist $U, V \in U(n)$ such that

$$\begin{aligned}U^* A U &= \mathrm{diag}(a, -a, 0, \cdots, 0), \\ V^* B V &= \mathrm{diag}(b\mathbf{i}, -b\mathbf{i}, 0, \cdots, 0).\end{aligned}$$

Hence

$$\mathrm{Tr}(X) = \mathrm{Tr}(A) + \mathrm{Tr}(B) = 0.$$

Because (3.4) achieves equality, the eigenspaces of $\lambda_1(B^t \otimes B)$ and $\lambda_1(-A^t \otimes A)$ have a nontrivial intersection. Let $U = (u_1, u_2, \cdots, u_n)$, $V = (v_1, v_2, \cdots, v_n)$, we have

$$Au_1 = au_1,\ Au_2 = -au_2,\ Au_j = 0,\ 3 \le j \le n;$$
$$Bv_1 = b\mathbf{i}v_1,\ Bv_2 = -b\mathbf{i}v_2,\ Bv_j = 0,\ 3 \le j \le n.$$

Since A is Hermitian and B is skew-Hermitian, we have

$$A^t\overline{u_1} = a\overline{u_1},\ A^t\overline{u_2} = -a\overline{u_2},\ A^t\overline{u_j} = 0,\ 3 \le j \le n;$$
$$B^t\overline{v_1} = b\mathbf{i}\overline{v_1},\ B^t\overline{v_2} = -b\mathbf{i}\overline{v_2},\ B^t\overline{v_j} = 0,\ 3 \le j \le n.$$

By the properties of the Kronecker product, the eigenspace of $\lambda_1(-A^t \otimes A)$ is $\mathrm{Span}_{\mathbb{C}}\{\overline{u_1} \otimes u_2, \overline{u_2} \otimes u_1\}$, and the eigenspace of $\lambda_1(B^t \otimes B)$ is $\mathrm{Span}_{\mathbb{C}}\{\overline{v_1} \otimes v_2, \overline{v_2} \otimes v_1\}$. So, there exist $k_1, k_2, l_1, l_2 \in \mathbb{C}$ and $|k_1|^2 + |k_2|^2 = |l_1|^2 + |l_2|^2 \ne 0$ such that

$$k_1\overline{u_1} \otimes u_2 + k_2\overline{u_2} \otimes u_1 = l_1\overline{v_1} \otimes v_2 + l_2\overline{v_2} \otimes v_1. \tag{3.5}$$

Recall that $U, V \in U(n)$, so we have

$$\overline{k_2}u_2 = \overline{l_1(u_1^* v_2)}v_1 + \overline{l_2(u_1^* v_1)}v_2,$$

by left-multiplication with $I \otimes u_1^*$ and conjugating (3.5). Similarly, $\overline{k_1}u_1 = \overline{l_1(u_2^* v_2)}v_1 + \overline{l_2(u_2^* v_1)}v_2$, $\overline{l_1}v_1 = \overline{k_1(v_2^* u_2)}u_1 + \overline{k_2(v_2^* u_1)}u_2$, $\overline{l_2}v_2 = \overline{k_1(v_1^* u_2)}u_1 + \overline{k_2(v_1^* u_1)}u_2$.

If $k_1 k_2 \ne 0$, it is easy to see that $\mathrm{Span}_{\mathbb{C}}\{u_1, u_2\} = \mathrm{Span}_{\mathbb{C}}\{v_1, v_2\}$. If one of k_1, k_2 is zero, we can assume without loss of generality that $k_1 \ne 0$ and $k_2 = 0$. Then one of l_1, l_2 is also zero, otherwise v_1 and v_2 would be linearly dependent. Assume without loss of generality that $l_1 \ne 0, l_2 = 0$, thus

$$k_1\overline{u_1} \otimes u_2 = l_1\overline{v_1} \otimes v_2.$$

Since $U, V \in U(n)$, we have $|k_1/l_1| = 1$ and

$$1 = (v_1^t \otimes v_2^*)(\overline{v_1} \otimes v_2) = (k_1/l_1)(v_1^t \otimes v_2^*)(\overline{u_1} \otimes u_2) = (k_1/l_1)(v_1^t\overline{u_1} \otimes v_2^* u_2).$$

The equality condition of Cauchy-Schwartz inequality implies that u_1, v_1 are linearly dependent and u_2, v_2 are linearly dependent.

In any case, we have $\mathrm{Span}_{\mathbb{C}}\{u_1, u_2\} = \mathrm{Span}_{\mathbb{C}}\{v_1, v_2\}$. Therefore

$$U^*XU = U^*AU + U^*BU = \mathrm{diag}(a, -a, 0, \cdots, 0) + \mathrm{diag}(B_0, O_{n-2}),$$

where $B_0 \in M(2, \mathbb{C})$. Setting $X_0 := \mathrm{diag}(a, -a) + B_0$, we have the necessity.

To prove the sufficiency, since $X_0 \in M(2,\mathbb{C})$ and $\operatorname{Tr} X_0 = 0$, it follows from Proposition 2.2 (b) and Corollary 2.5 that

$$\lambda_1(T_X) = \lambda_1(T_{\operatorname{diag}(X_0,O_{n-2})}) = \lambda_1(T_{X_0}) = 2 - |\operatorname{Tr}(X_0)|^2 = 2.$$

This completes the proof. □

Now we give a new proof of the equality condition for the complex BW inequality. With the notion introduced in [5], a pair (X,Y) of $M(n,\mathbb{C})$ is said to be maximal if $X \neq O$, $Y \neq O$ and $\|XY - YX\|^2 = 2\|X\|^2\|Y\|^2$ is satisfied.

Corollary 3.3. *Let $X, Y \in M(n,\mathbb{C})$ be nonzero matrices. Then (X,Y) is maximal if and only if there exists a unitary matrix $U \in U(n)$ such that*

$$X = U \operatorname{diag}(X_0, 0)U^* \quad \text{and} \quad Y = U \operatorname{diag}(Y_0, 0)U^*$$

with a maximal pair (X_0, Y_0) in $M(2,\mathbb{C})$, i.e., $X_0 \perp_{\mathbb{C}} Y_0$ and $\operatorname{Tr} X_0 = \operatorname{Tr} Y_0 = 0$.

Proof. Without loss of generality, we assume $\|X\| = \|Y\| = 1$. If (X,Y) is maximal, by definition, we have

$$\langle T_X Y, Y\rangle = \langle T_Y X, X\rangle = \|[X,Y]\|^2 = 2.$$

Thus $\lambda_1(T_X) = \lambda_1(T_Y) = 2$ and hence by Theorem 3.2, there exist unitary matrices $U_1, U_2 \in U(n)$ such that

$$X = U_1 \operatorname{diag}(X_0, 0)U_1^* \quad \text{and} \quad Y = U_2 \operatorname{diag}(\widetilde{Y_0}, 0)U_2^*$$

with $\operatorname{Tr} X = \operatorname{Tr} Y = 0$. Since Y is an eigenvector of the maximal eigenvalue $\lambda_1(T_X) = 2$ and X is an eigenvector of the zero eigenvalue of T_X, we know immediately $X \perp_{\mathbb{C}} Y$. Moreover, by (2.2) and Proposition 2.2 (b) we know $U_1^* Y U_1$ is an eigenvector of the maximal eigenvalue $\lambda_1(T_{U_1^* X U_1}) = \lambda_1(T_{X_0}) = 2$, which implies $U_1^* Y U_1 = \operatorname{diag}(Y_0, 0)$ for some $Y_0 \in M(2,\mathbb{C})$. This completes the proof of the necessity.

The sufficiency can be verified by direct computation (cf. [5,8]). □

Let $\|X\|_{(2),2}$ be the Ky Fan-mix (2, 2)-norm defined by (see [2,27])

$$\|X\|_{(2),2} = \sqrt{\sigma_1^2(X) + \sigma_2^2(X)}.$$

For $X \in M(n,\mathbb{R})$, Lu [22] has already proved

$$\lambda_1(T_X) \leq 2\|X\|_{(2),2}^2.$$

Though this proof can be technically extended to the complex version, we want to give an alternative proof using the approach of Kronecker products.

Theorem 3.4. *For $X \in M(n, \mathbb{C})$ with $\|X\| = 1$, we have*

$$\lambda_1(T_X) \le 2\|X\|_{(2),2}^2 \le 2.$$

Proof. For $Y \in M(n, \mathbb{C})$, by Proposition 2.4 we have

$$\langle WW^* \widetilde{v}, \widetilde{v} \rangle = \langle T_X Y, Y \rangle, \tag{3.6}$$

where

$$W = \begin{pmatrix} I \otimes X^* & O \\ -\overline{X} \otimes I & O \end{pmatrix}_{2N \times 2N}, \ \widetilde{v} = \begin{pmatrix} \mathrm{Vec}\, Y \\ \mathrm{Vec}\, Y \end{pmatrix}.$$

Notice that

$$WW^* = \begin{pmatrix} I \otimes X^*X & -X^t \otimes X^* \\ -\overline{X} \otimes X & \overline{X}X^t \otimes I \end{pmatrix}_{2N \times 2N},$$
$$W^*W = \begin{pmatrix} I \otimes XX^* + X^t\overline{X} \otimes I & O \\ O & O \end{pmatrix}_{2N \times 2N}.$$

Let L_1 and E_1 be the first eigenspace of T_X and WW^*, respectively. And let

$$\widetilde{L_1} = \left\{ \begin{pmatrix} \mathrm{Vec}\, Y \\ \mathrm{Vec}\, Y \end{pmatrix} : Y \in L_1 \right\}, \ E_1^{\perp} = \left\{ \xi \in \mathbb{C}^{2N} : \xi \perp E_1 \right\}.$$

We can assume without loss of generality that $\dim E_1 = 1$, since $\dim E_1 \ge 2$ implies that $\lambda_1(WW^*) = \lambda_2(WW^*)$, then by Lemma 2.6 (a) and (3.6) we have

$$\lambda_1(T_X) \le 2\lambda_1(WW^*) = 2\lambda_2(WW^*) = 2\lambda_2(W^*W) = 2(\sigma_1^2(X) + \sigma_2^2(X)) = 2\|X\|_{(2),2}^2.$$

Thus by Proposition 2.2 (c) we have $\dim \widetilde{L_1} \ge 2$ and

$$\begin{aligned} \dim(\widetilde{L_1} \cap E_1^{\perp}) &= \dim \widetilde{L_1} + \dim E_1^{\perp} - \dim(\widetilde{L_1} \cup E_1^{\perp}) \\ &\ge 2 + 2N - 1 - 2N = 1. \end{aligned}$$

Therefore $\widetilde{L_1} \cap E_1^{\perp}$ is non-empty and by Lemma 2.6 (a) and (3.6), for any $Y \in L_1$ with $\|Y\| = 1$ and $\xi = \left(\begin{smallmatrix} \mathrm{Vec}\, Y \\ \mathrm{Vec}\, Y \end{smallmatrix}\right) \in \widetilde{L_1}$ ($\|\xi\| = \sqrt{2}$), we have

$$\begin{aligned} \lambda_1(T_X) = \langle T_X Y, Y \rangle = \langle WW^* \xi, \xi \rangle &= 2 \max_{\substack{\xi \in \widetilde{L_1} \cap E_1^{\perp} \\ \|\xi\| = 1}} \langle WW^* \xi, \xi \rangle \\ &\le 2 \max_{\substack{\xi \in E_1^{\perp} \\ \|\xi\| = 1}} \langle WW^* \xi, \xi \rangle = 2\lambda_2(WW^*) = 2\|X\|_{(2),2}^2. \end{aligned}$$

This completes the proof. □

Denote the upper bound in Theorem 3.1 by

$$C_X := 2(\max_{i,j}\{-a_ia_j\} + \max_{i,j}\{-b_ib_j\}) + \|X\|^2_{(2),2}.$$

It is worth remarking that $C_X \leq 2\|X\|^2_{(2),2}$ if $\operatorname{rank}(X) \leq 2$. In general, C_X is not necessarily less than $2\|X\|^2_{(2),2}$. However, we are able to obtain $C_X \leq 3\|X\|^2_{(2),2}$, since $\{|a_j - \mathbf{i}b_{n-j+1}|^2\}_{j=1}^n$ is majorized by $\{\sigma_j^2(X)\}_{j=1}^n$ due to Ando-Bhatia [1]. Therefore these two upper bounds are strictly different. Combining Theorems 3.1 and 3.4, we have the following estimate.

Corollary 3.5. *For $X \in M(n, \mathbb{C})$ with $\|X\| = 1$, we have*

$$\lambda_1(T_X) \leq \min\{C_X, 2\|X\|^2_{(2),2}\} \leq 2.$$

Furthermore, our approach can be used to estimate all eigenvalues of T_X by these of K_1 (Proposition 2.4). Recall that the set of eigenvalues of K_1 is given in (3.1).

Theorem 3.6. *For $X \in M(n, \mathbb{C})$ with $\|X\| = 1$, we have $\lambda_i(T_X) \leq 2\lambda_i(K_1)$ for all i.*

Proof. Recall that $K_1 = I \otimes X^*X + \overline{X}X^t \otimes I,\ K_2 = -\left(X^t \otimes X^* + \overline{X} \otimes X\right)$, and

$$\mathrm{Vec} \circ T_X \circ (\mathrm{Vec})^{-1} = K_1 + K_2.$$

Let $\widehat{K}_X := I \otimes X + X^t \otimes I$. Then we observe

$$2K_1 - \mathrm{Vec} \circ T_X \circ (\mathrm{Vec})^{-1} = K_1 - K_2 = \widehat{K}_X^* \widehat{K}_X \geq 0,$$

which implies $\lambda_i(T_X) \leq 2\lambda_i(K_1)$ for all i. □

In particular, Theorem 3.6 implies Theorem 3.4 since

$$\lambda_1(T_X) = \lambda_2(T_X) \leq 2\lambda_2(K_1) = 2(\sigma_1^2 + \sigma_2^2) = 2\|X\|^2_{(2),2}.$$

4. Equivalence of the conjectures with the LW Conjecture

In this section, we prove the equivalence between Conjectures 4–6 and Conjecture 2, i.e., Theorem 1.1 in the complex version. The proof will be split into the following four propositions.

Proposition 4.1. *Conjecture 2 is equivalent to Conjecture 4.*

Proof. Assume Conjecture 2 is true at first. Setting $B = X$ and B_α be a unit eigenvector of $\lambda_{2\alpha-3}(T_X)$ for $\alpha = 2, \cdots, m$, by the representation of S_X in (2.3, 2.4) we know

$S_X B_\alpha = [B, B_\alpha]^*$ is exactly an eigenvector of $\lambda_{2\alpha-2}(T_X)$. Therefore the conditions (i, ii) of Conjecture 2 are satisfied and thus we have the inequality (1.2). Then the inequality (1.3) of Conjecture 4 for $k = m - 1$ follows by Proposition 2.2 (c) and the following calculation:

$$\begin{aligned}\sum_{i=1}^{2k} \lambda_i(T_X) &= 2\sum_{\alpha=2}^{m} \lambda_{2\alpha-3}(T_X) = 2\sum_{\alpha=2}^{m} \langle T_B B_\alpha, B_\alpha\rangle = 2\sum_{\alpha=2}^{m} \|[B, B_\alpha]\|^2 \\ &\leq 2\left(\max_{2\leq\alpha\leq m} \|B_\alpha\|^2 + \sum_{\alpha=2}^{m} \|B_\alpha\|^2\right) \|B\|^2 = 2m = 2(k+1).\end{aligned}$$

Now we assume Conjecture 4 is true. Without loss of generality, we assume $1 = \|B\| \geq \|B_2\| \geq \cdots \geq \|B_m\| > 0$. Using summation by parts, we can write

$$\begin{aligned}\sum_{\alpha=2}^{m} \|\,[B, B_\alpha]\,\|^2 &= \sum_{\alpha=2}^{m} \langle T_B B_\alpha, B_\alpha\rangle = \sum_{\alpha=2}^{m} \left\langle T_B \frac{B_\alpha}{\|B_\alpha\|}, \frac{B_\alpha}{\|B_\alpha\|}\right\rangle \|B_\alpha\|^2 \\ &= \sum_{\beta=2}^{m} \left(\|B_\beta\|^2 - \|B_{\beta+1}\|^2\right) \sum_{\alpha=2}^{\beta} \left\langle T_B \frac{B_\alpha}{\|B_\alpha\|}, \frac{B_\alpha}{\|B_\alpha\|}\right\rangle,\end{aligned} \tag{4.1}$$

where $B_{m+1} = 0$. Setting $X = B$, the conditions (i, ii) of Conjecture 2 show that the subspace $W := \mathrm{Span}_{\mathbb{C}}\{B_\alpha\}_{\alpha=2}^{m}$ is isotropic about S_X, i.e., $S_X(W) \perp_{\mathbb{C}} W$. Then by the formula above, Lemma 2.1 (c) and the inequality (1.3) of Conjecture 4, we have

$$\begin{aligned}\sum_{\alpha=2}^{m} \|\,[B, B_\alpha]\,\|^2 &\leq \sum_{\beta=2}^{m} \left(\|B_\beta\|^2 - \|B_{\beta+1}\|^2\right) \sum_{\alpha=2}^{\beta} \lambda_{2\alpha-3}(T_X) \\ &\leq \sum_{\beta=2}^{m} \left(\|B_\beta\|^2 - \|B_{\beta+1}\|^2\right) \beta \\ &= \|B_2\|^2 + \sum_{\alpha=2}^{m} \|B_\alpha\|^2,\end{aligned}$$

which is the inequality (1.2) of Conjecture 2. The proof is complete. □

Proposition 4.2. *Conjecture 4 is equivalent to Conjecture 5.*

Proof. Obviously Conjecture 5 implies Conjecture 4 by definition. Suppose Conjecture 4 to be true. To prove Conjecture 5, we only need to prove the following four parts:

(i) $\lambda_1(T_X) \leq 2$;
(ii) $\sum_{i=1}^{2k} \lambda_i(T_X) \leq 2k + 2$;
(iii) $\sum_{i=1}^{2k-1} \lambda_i(T_X) \leq 2k + 1$;
(iv) $\sum_{i=1}^{N} \lambda_i(T_X) = 2n - 2|\operatorname{Tr} X|^2 \leq 2n$,

where (i) and (iv) are ensured by the complex BW inequality (e.g., Theorem 3.1) and Corollary 2.5, and (ii) is assumed by Conjecture 4, respectively. We are left to show the inequality (iii). We prove it by contradiction in the following.

Assume that there is a positive number $m \geq 2$ such that

$$\sum_{i=1}^{2m-1} \lambda_i(T_X) > 2m + 1.$$

Then

$$2m + 1 < \sum_{i=1}^{2m-1} \lambda_i(T_X) = \lambda_{2m-1}(T_X) + \sum_{i=1}^{2m-2} \lambda_i(T_X) \leq \lambda_{2m-1}(T_X) + 2m.$$

Thus

$$\lambda_{2m}(T_X) = \lambda_{2m-1}(T_X) > 1,$$

and

$$\sum_{i=1}^{2m} \lambda_i(T_X) = \lambda_{2m}(T_X) + \sum_{i=1}^{2m-1} \lambda_i(T_X) > 1 + 2m + 1 = 2m + 2.$$

This leads to the contradiction to (ii) and completes the proof. □

Proposition 4.3. *Conjecture 4 is equivalent to Conjecture 6.*

Proof. The proof is somehow similar to that of Proposition 4.1. However, we are unable to use Lemma 2.1 (c), since we have no condition (ii) of Conjecture 2. Therefore we estimate the full sum of λ_i's here instead of only half of it.

Assume Conjecture 6 is true. Setting $B = X$ and B_α be a unit eigenvector of $\lambda_{\alpha-1}(T_X)$ for $\alpha = 2, \cdots, m$, we know or can ensure that B_α's are $\mathbb{C}$-orthogonal, and therefore we have the inequality of Conjecture 6. Then the inequality (1.3) of Conjecture 4 for $m = 2k + 1$ follows by

$$\sum_{i=1}^{m-1} \lambda_i(T_X) = \sum_{\alpha=2}^{m} \lambda_{\alpha-1}(T_X) = \sum_{\alpha=2}^{m} \langle T_B B_\alpha, B_\alpha \rangle = \sum_{\alpha=2}^{m} \|[B, B_\alpha]\|^2$$
$$\leq \left(2 \max_{2 \leq \alpha \leq m} \|B_\alpha\|^2 + \sum_{\alpha=2}^{m} \|B_\alpha\|^2 \right) \|B\|^2 = m + 1.$$

Now we assume Conjecture 4 is true (and hence Conjecture 5 is true by Proposition 4.2). We have

$$\sum_{i=1}^{m} \lambda_i(T_X) \leq m+2 \quad \text{for any } m.$$

Without loss of generality, we assume $1 = \|B\| \geq \|B_2\| \geq \cdots \geq \|B_m\| > 0$ and set $B_{m+1} = 0$. Then by (4.1), we have

$$\begin{aligned}\sum_{\alpha=2}^{m} \| [B, B_\alpha] \|^2 &\leq \sum_{\beta=2}^{m} \left(\|B_\beta\|^2 - \|B_{\beta+1}\|^2\right) \sum_{\alpha=1}^{\beta-1} \lambda_\alpha(T_X) \\ &\leq \sum_{\beta=2}^{m} \left(\|B_\beta\|^2 - \|B_{\beta+1}\|^2\right)(\beta+1) \\ &= 2\|B_2\|^2 + \sum_{\alpha=2}^{m} \|B_\alpha\|^2,\end{aligned}$$

which is the inequality of Conjecture 6. $\square$

Proposition 4.4. *[23] The LW Conjecture 2 implies Conjectures 1 and 3.*

Proof. Conjecture 3 is trivially implied by Conjecture 4 due to Proposition 4.1 and thus by Conjecture 2.

As for Conjecture 1, the proof is similar to the real version in [23] with only the inner products replaced by traces for our complex version now. $\square$

5. Partial results on the complex LW Conjecture

In this section, we prove the complex LW Conjecture separately for a couple of special cases (Theorem 1.2). For the general case, we give some non-sharp upper bounds for the inequalities of Conjectures 3 and 4 (Theorems 1.3 and 1.4).

Firstly we prove the complex version of Conjecture 3 for the first special case of Theorem 1.2. We remind that Conjecture 3 is the first instance of the complex LW Conjecture 4 after the solution of the BW inequality (i.e., $\lambda_1(T_X) \leq 2$).

Normality often plays a special role, and also the BW inequality could be treated much simpler with this property. In the context of geometry as with the DDVV inequality, it naturally emerges.

Theorem 5.1. *Conjecture 3 is true when X is a normal matrix.*

Proof. Since X is a normal matrix, there exists a unitary matrix U such that

$$U^* X U = \operatorname{diag}(x_1, \cdots, x_n), \quad \textit{for some } x_1, \cdots, x_n \in \mathbb{C} \textit{ with } \sum_i |x_i|^2 = 1.$$

Direct calculations show that for any $1 \leq i, j \leq n$,

$$T_{U^*XU}(E_{ij}) = |x_i - x_j|^2 E_{ij},$$

where $E_{ij} \in M(n, \mathbb{C})$ is the standard basis matrix with the (i, j)-element being 1 and the others being 0. Then by the identity (2.2):

$$T_{U^*XU}(U^*YU) = U^* T_X(Y) U,$$

we have

$$T_X(UE_{ij}U^*) = UT_{U^*XU}(E_{ij})U^* = |x_i - x_j|^2 UE_{ij}U^*.$$

It follows that

$$\lambda(T_X) = \left\{|x_i - x_j|^2 : 1 \le i, j \le n\right\} = \{\lambda_1 \ge \cdots \ge \lambda_{n^2}\}.$$

Suppose $\lambda_1 = \lambda_2 = |x_a - x_b|^2$, $\lambda_3 = \lambda_4 = |x_c - x_d|^2$, where $1 \le a, b, c, d \le n$. There are two cases need to be discussed:

- If a, b, c, d are four different integers, then

$$\lambda_1 + \lambda_3 = |x_a - x_b|^2 + |x_c - x_d|^2 \le 2(|x_a|^2 + |x_b|^2 + |x_c|^2 + |x_d|^2) \le 2.$$

- If one of a, b is equal to one of c, d, we can assume $a = c$, $b \ne d$. Then

$$\begin{aligned} \lambda_1 + \lambda_3 &= |x_a - x_b|^2 + |x_a - x_d|^2 \\ &= |x_a|^2 - x_a\overline{x_b} - \overline{x_a}x_b + |x_b|^2 + |x_a|^2 - x_a\overline{x_d} - \overline{x_a}x_d + |x_d|^2 \\ &\le 2|x_a|^2 + 2|x_a|\left(|x_b| + |x_d|\right) + |x_b|^2 + |x_d|^2 \\ &\le 3|x_a|^2 + \left(|x_b| + |x_d|\right)^2 + |x_b|^2 + |x_d|^2 \\ &\le 3(|x_a|^2 + |x_b|^2 + |x_d|^2) \le 3. \end{aligned}$$

  The equality holds if and only if $|x_a| = \frac{\sqrt{6}}{3}$, $x_b = x_d = -\frac{1}{2}x_a$, other $x_e = 0$.

The proof is complete. □

For more general cases, we need Lu's lemma in the complex version. By dividing each complex number into real and imaginary parts, the complex version follows immediately from the real version.

Lemma 5.2. *[21, Lemma 1] Suppose $\eta_1, \cdots, \eta_n$ are complex numbers and*

$$\eta_1 + \cdots + \eta_n = 0, \quad |\eta_1|^2 + \cdots + |\eta_n|^2 = 1.$$

Let $r_{ij} \geq 0$ be nonnegative numbers for $i < j$. Then we have

$$\sum_{i<j} |\eta_i - \eta_j|^2 r_{ij} \leq \sum_{i<j} r_{ij} + \max_{i<j}(r_{ij}). \tag{5.1}$$

Corollary 5.3. *The complex LW Conjecture 4 is true when X is a normal matrix.*

Proof. Let X be a normal matrix with $\|X\| = 1$, and let

$$\hat{X} = X - \frac{\operatorname{Tr} X}{n} I.$$

Then $\operatorname{Tr} \hat{X} = 0$, $[\hat{X}, Y] = [X, Y]$ and

$$\|\hat{X}\|^2 = \|X\|^2 - \frac{|\operatorname{Tr} X|^2}{n} \leq \|X\|^2 = 1.$$

It follows from the proof of Theorem 5.1 that

$$\lambda(T_X) = \lambda(T_{\hat{X}}) = \left\{ |\eta_i - \eta_j|^2 : 1 \leq i, j \leq n \right\},$$

where $\lambda(\hat{X}) = \{\eta_1, \eta_2, \cdots, \eta_n\}$. Let $\widetilde{X} = \hat{X}/\|\hat{X}\|$ and $r_{ij} \in \{0, 1\}$. Then $\operatorname{Tr} \widetilde{X} = 0$, $\|\widetilde{X}\|^2 = 1$ and $T_X = T_{\hat{X}} = \|\hat{X}\|^2 T_{\widetilde{X}}$. Thus Lemma 5.2 is applicable and tells us

$$\sum_{\alpha=1}^{k} \lambda_{2\alpha-1}(T_X) = \sum_{\alpha=1}^{k} \lambda_{2\alpha-1}(T_{\hat{X}}) = \|\hat{X}\|^2 \sum_{\alpha=1}^{k} \lambda_{2\alpha-1}(T_{\widetilde{X}}) \leq k + 1,$$

where $\lambda_{2\alpha-1}(T_{\widetilde{X}})$ equals some $\left| \eta_i/\|\hat{X}\| - \eta_j/\|\hat{X}\| \right|^2$ and $r_{ij} = 1$ for k pairs of $(i < j)$. This completes the proof. □

Corollary 5.4. *Let $B_1, \cdots, B_m \in M(n, \mathbb{C})$ be Hermitian matrices. Assume that*

$$\operatorname{Tr} \Big(B_\alpha [B_\gamma, B_\beta] \Big) = 0 \tag{5.2}$$

for any $1 \leq \alpha, \beta, \gamma \leq m$, we have

$$\sum_{\alpha,\beta=1}^{m} \| [B_\alpha, B_\beta] \|^2 \leq \left(\sum_{\alpha=1}^{m} \|B_\alpha\|^2 \right)^2.$$

Proof. As Hermitian matrices are normal matrices, by Corollary 5.3 above, the complex LW Conjecture 4 holds for this case. This in turn by Theorem 1.1 implies the complex version of Conjecture 1. □

Remark 5.5. When $B_1, \cdots, B_m$ are real symmetric matrices, (5.2) is valid for all α, β, γ. Thus the corollary generalizes the DDVV inequality. We remind that for general Hermitian matrices without the trace condition, the optimal constant $c = \frac{4}{3}$ is bigger than 1 here (cf. Section 1, [16], [17]).

Next we prove Conjecture 3 for the second special case $\text{rank}(X) = 1$.

Theorem 5.6. *The complex LW Conjecture 4 is true when* $\text{rank}(X) = 1$.

Proof. Recall Proposition 2.4 that we have $K_X^* K_X = K_1 + K_2$, where $K_1 = I \otimes X^*X + \overline{X}X^t \otimes I$, $K_2 = -\left(X^t \otimes X^* + \overline{X} \otimes X\right)$. Denote $K_3 = -X^t \otimes X^*$, then $K_2 = K_3^* + K_3$. In [3], we can find that for any $M \in M(n, \mathbb{C})$,

$$\lambda_i\left(\frac{M^* + M}{2}\right) \le \sigma_i(M), \quad 1 \le i \le n.$$

Thus

$$\lambda_i(K_2) = 2\lambda_i\left(\frac{K_3^* + K_3}{2}\right) \le 2\sigma_i(K_3), \quad 1 \le i \le n.$$

Let $\sigma_1(X) \ge \cdots \ge \sigma_n(X)$ be the singular values of X, then by Lemma 2.6 (a),

$$\sigma(K_3) = \{\sigma_i(X)\sigma_j(X) : 1 \le i, j \le n\}.$$

In particular, now $\text{rank}(X) = 1$ implies $\sigma_1(X) = 1$ and $\sigma_i(X) = 0$ for $2 \le i \le n$. Thus we have $\sigma(K_3) = \{1^1, 0^{N-1}\}$ and by (3.1)

$$\lambda(K_1) = \{\sigma_i(X)^2 + \sigma_j(X)^2 : 1 \le i, j \le n\} = \{2^1, 1^{2(n-1)}, 0^{(n-1)^2}\}.$$

Finally by Propositions 2.2 (c), 2.4 and Lemma 2.6 (b), we have

$$\begin{aligned}\sum_{i=1}^k \lambda_{2i-1}(T_X) = \sum_{i=1}^k \lambda_{2i}(T_X) &\le \sum_{i=1}^k \lambda_i(K_1) + \sum_{i=1}^k \lambda_{2i}(K_2) \\ &\le \sum_{i=1}^k \lambda_i(K_1) + \sum_{i=1}^k 2\sigma_{2i}(K_3) \\ &= \sum_{i=1}^k \lambda_i(K_1) \le k + 1,\end{aligned}$$

which completes the proof. □

Furthermore, we can get the characteristic polynomial of T_X if $\text{rank}(X) = 1$.

Proposition 5.7. *Let* $K_X = I \otimes X \ - X^t \otimes I$*. Then the sets of singular values*

$$\sigma(K_X) = \sigma\Big(I \otimes \Lambda \ - (\Lambda \otimes I)\left(Q^t \otimes \ Q^*\right)\Big),$$

where $X = Q_1 \Lambda Q_2$ *is the singular value decomposition of* X *and* $Q = Q_2 Q_1$*.*

Proof. Direct calculations show

$$\begin{aligned} K_X &= I \otimes X \ - X^t \otimes I = I \otimes (Q_1 \Lambda Q_2) - \left(Q_2^t \Lambda Q_1^t\right) \otimes I \\ &= \left(Q_2^t \otimes \ Q_1\right) \left[I \otimes \Lambda \ - (\Lambda \otimes I)\left(Q^t \otimes \ Q^*\right)\right] \left(\overline{Q_2} \otimes \ Q_2\right). \end{aligned}$$

This completes the proof thanks to the invariance of singular values under congruences. □

Theorem 5.8. *Let* X *be a complex square matrix of order* n (≥ 2) *with* $\|X\| = 1$ *and* $\operatorname{rank}(X) = 1$*, then the characteristic polynomial of* T_X *is*

$$\det(\lambda I_N - T_X) = \left(\lambda - 2 + |\operatorname{Tr} X|^2\right)^2 (\lambda - 1)^{2n-4} \lambda^{(n-1)^2+1}.$$

Proof. Proposition 5.7 implies $\sigma(K_X) = \sigma(\widetilde{K_X})$, where $\widetilde{K_X} = I \otimes \Lambda \ - (\Lambda \otimes I)\left(Q \otimes \ \overline{Q}\right)$. By Proposition 2.4, we have

$$\lambda(T_X) = \lambda(K_X K_X^*) = \lambda(\widetilde{K_X} \widetilde{K_X}^*),$$

where direct calculations show

$$\widetilde{K_X} \widetilde{K_X}^* = I \otimes \Lambda^2 \ + \Lambda^2 \otimes I \ - (Q^* \Lambda) \otimes \left(\Lambda Q^t\right) - (\Lambda Q) \otimes \left(\overline{Q} \Lambda\right).$$

Since $\|X\| = 1$ and $\operatorname{rank}(X) = 1$, one has $\Lambda = \operatorname{diag}(1, 0, \cdots, 0)$. By direct calculations, we have $I \otimes \Lambda^2 \ + \Lambda^2 \otimes I \ = \operatorname{diag}(I + \Lambda, \Lambda, \cdots, \Lambda)$ and thus

$$\lambda \cdot (I \otimes I) - \widetilde{K_X} \widetilde{K_X}^* = \begin{pmatrix} A_{n \times n} & B_{n \times (N-n)} \\ C_{(N-n) \times n} & D_{(N-n) \times (N-n)} \end{pmatrix}_{N \times N},$$

where

$$A := (\lambda - 1) I - \Lambda + q_{11} \overline{Q} \Lambda + \overline{q_{11}} \Lambda Q^t, \quad B := \left(q_{12} \overline{Q} \Lambda, q_{13} \overline{Q} \Lambda, \cdots, q_{1n} \overline{Q} \Lambda\right),$$
$$C := B^*, \quad D := \operatorname{diag}(\lambda I - \Lambda, \lambda I - \Lambda, \cdots, \lambda I - \Lambda).$$

Without loss of generality, suppose that the determinant of the matrix D is not zero, then

$$\begin{aligned}\det\begin{pmatrix} A & B \\ C & D\end{pmatrix} &= \det\left(A - BD^{-1}C\right)\cdot\det D \\ &= \det\left(A - \left(1-|q_{11}|^2\right)\overline{Q}\Lambda\widehat{D}\Lambda Q^t\right)\cdot\det D \\ &= \det\left(A - \frac{1-|q_{11}|^2}{\lambda-1}\overline{Q}\Lambda Q^t\right)\cdot\det D,\end{aligned}$$

where $\widehat{D} = (\lambda I - \Lambda)^{-1} = \mathrm{diag}(\frac{1}{\lambda-1}, \frac{1}{\lambda}\cdots, \frac{1}{\lambda})$. Thus

$$A - BD^{-1}C = A - \frac{1-|q_{11}|^2}{\lambda-1}\overline{Q}\Lambda Q^t = \begin{pmatrix} \widetilde{A}_{1\times 1} & \widetilde{B}_{1\times(n-1)} \\ \widetilde{C}_{(n-1)\times 1} & \widetilde{D}_{(n-1)\times(n-1)}\end{pmatrix}_{n\times n},$$

where $\widetilde{A} := \lambda-2+2\,|q_{11}|^2 - \frac{1-|q_{11}|^2}{\lambda-1}\,|q_{11}|^2$, $\widetilde{B} := (\overline{q_{11}}q_{21}\frac{\lambda-2+|q_{11}|^2}{\lambda-1}, \cdots, \overline{q_{11}}q_{n1}\frac{\lambda-2+|q_{11}|^2}{\lambda-1})$, $\widetilde{C} := \widetilde{B}^*$, $\widetilde{D} := (\lambda-1)\,I - \frac{1-|q_{11}|^2}{\lambda-1}u^*u$, $u := (q_{21}, q_{31}, \cdots, q_{n1})$.
Similarly,

$$\begin{aligned}\det\begin{pmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D}\end{pmatrix} &= \det\left(\widetilde{A} - \widetilde{C}\widetilde{A}^{-1}\widetilde{B}\right)\cdot\det\widetilde{A} \\ &= \det\left((\lambda-1)\,I - \frac{1}{\lambda-1+|q_{11}|^2}u^*u\right)\cdot\det\widetilde{A} \\ &= \left[\lambda\frac{\lambda-2+|q_{11}|^2}{\lambda-1+|q_{11}|^2}(\lambda-1)^{n-2}\right]\cdot\left[\frac{1}{\lambda-1}\left(\lambda-1+|q_{11}|^2\right)\left(\lambda-2+|q_{11}|^2\right)\right] \\ &= \left(\lambda-2+|q_{11}|^2\right)^2(\lambda-1)^{n-3}\,\lambda.\end{aligned}$$

So we have

$$\begin{aligned}\det\begin{pmatrix} A & B \\ C & D\end{pmatrix} &= \det\left(\widetilde{A} - \widetilde{C}\widetilde{A}^{-1}\widetilde{B}\right)\cdot\det\widetilde{A}\cdot\det D \\ &= \left(\lambda-2+|q_{11}|^2\right)^2(\lambda-1)^{n-3}\,\lambda\,(\lambda-1)^{n-1}\,\lambda^{(n-1)^2} \\ &= \left(\lambda-2+|q_{11}|^2\right)^2(\lambda-1)^{2n-4}\,\lambda^{(n-1)^2+1}.\end{aligned}$$

Finally we observe that $q_{11} = \mathrm{Tr}\,X$. The proof is complete. □

Immediately, we obtain the following assertion.

Corollary 5.9. *Let X be a complex square matrix of order n (≥ 2) with $\|X\| = 1$ and $rank(X) = 1$. Then $\lambda_1(T_X) = 2$ if and only if* $\mathrm{Tr}(X) = 0$.

Remark 5.10. Actually, the conditions $\|X\| = 1$, $\text{rank}(X) = 1$ and $\text{Tr}(X) = 0$ in Corollary 5.9 imply that X is unitarily similar to $\text{diag}(X_0, O)$, where

$$X_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Here, we can give a simple calculation. Suppose $X = Q_1 \Lambda Q_2$ is the singular value decomposition of X and $Q = Q_2 Q_1$, then $Q_1^* X Q_1 = \Lambda Q$. Due to $\|X\| = 1$, $\text{rank}(X) = 1$ and $\text{Tr}(X) = 0$, we can assume

$$\Lambda Q = \begin{pmatrix} 0 & q \\ 0 & 0 \end{pmatrix},$$

where $q = (q_{12}, q_{13}, \cdots, q_{1n})$ and $\|q\| = 1$. Extend q to be a unit orthogonal basis $\{q, p_1, p_2, \cdots, p_{n-2}\}$ of $\mathbb{C}^{n-1}$ and let

$$U = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ q^* & O & p_1^* & p_2^* & \cdots & p_{n-2}^* \end{pmatrix},$$

then $U^* U = I$ and $U^* Q_1^* X Q_1 U = \text{diag}(X_0, O)$.

The last special case of Theorem 1.2 is a simple consequence of Corollary 2.5.

Theorem 5.11. *The complex LW Conjecture 4 is true when $n = 2, 3$.*

Proof. The case $n = 2$ is immediate since Corollary 2.5 implies that the set of eigenvalues $\lambda(T_X)$ is weakly majorized by $\{2^2, 0^2\}$.

The case $n = 3$ is similar, since Corollary 2.5 shows that

$$\sum_{i=1}^{2k} \lambda_i(T_X) \le \text{Tr}\, T_X = 6 - 2|\text{Tr}\, X|^2 \le 6 \le 2k + 2 \quad \text{for any } k \ge 2,$$

and for $k = 1$ it follows from the BW inequality (e.g., Theorem 3.1) that

$$\sum_{i=1}^{2k} \lambda_i(T_X) = 2\lambda_1(T_X) \le 4 = 2k + 2.$$

The proof is complete. □

Now we come to prove the partial results with increased bounds.

Proof of Theorem 1.3. Due to the previous result, we assume $n \ge 4$. Otherwise it is trivial. By Lemma 2.6 (b) and the proof of Theorem 3.1, for the fixed sequence $i_1 = 2, i_2 = 3, i_3 = 4$, we have

$$\sum_{i=2}^{4} \lambda_i(T_X) \le \sum_{i=2}^{4} \lambda_i(K_1) + \sum_{i=1}^{3} \lambda_i(K_2)$$
$$\le 3\Big(\sigma_1^2(X) + \sigma_2^2(X)\Big) + 2\sum_{i=1}^{3} \Big(\lambda_i(-A^t \otimes A) + \lambda_i(B^t \otimes B)\Big),$$

as $K_2 = -X^t \otimes X^* - \overline{X} \otimes X = 2\,(B^t \otimes B - A^t \otimes A)$ for the decomposition $X = A + B$ with A Hermitian and B skew-Hermitian. Similarly we have

$$\lambda_1(T_X) = \lambda_2(T_X) \le \sigma_1^2(X) + \sigma_2^2(X) + 2\Big(\lambda_1(-A^t \otimes A) + \lambda_1(B^t \otimes B)\Big).$$

This implies

$$\sum_{i=1}^{4} \lambda_i(T_X) \le 4\Big(\sigma_1^2(X) + \sigma_2^2(X)\Big) + \phi(X),$$

where $\phi(X) := \varphi(A) + \widetilde{\varphi}(B)$ and

$$\varphi(A) := 4\lambda_1(-A^t \otimes A) + 2\sum_{i=2}^{3} \lambda_i(-A^t \otimes A),$$
$$\widetilde{\varphi}(B) := 4\lambda_1(B^t \otimes B) + 2\sum_{i=2}^{3} \lambda_i(B^t \otimes B).$$

Let $\lambda(A) = \{a_1, \cdots, a_n\}$, $a_1 \ge \cdots \ge a_n$; $\lambda(B) = \{b_1\mathbf{i}, \cdots, b_n\mathbf{i}\}$, $b_1 \ge \cdots \ge b_n$. Then by Lemma 2.6 (a),

$$\lambda(-A^t \otimes A) = \{-a_i a_j : 1 \le i, j \le n\},$$
$$\lambda(B^t \otimes B) = \{-b_i b_j : 1 \le i, j \le n\}.$$

We claim that

$$\phi(X) = \varphi(A) + \widetilde{\varphi}(B) \le \sqrt{10}\Big(\|A\|^2 + \|B\|^2\Big).$$

We only show $\varphi(A) \le \sqrt{10}\|A\|^2$. The argumentation for $\widetilde{\varphi}(B)$ is similar. Obviously $\varphi(A)$ can only be positive if $\lambda_1 > 0$, i.e. $a_1 > 0 > a_n$, in which case we have

$$\lambda_1(-A^t \otimes A) = \lambda_2(-A^t \otimes A) = a_1|a_n| = \max_{i,j}\{-a_i a_j\}$$

and

$$\lambda_3(-A^t \otimes A) = \lambda_4(-A^t \otimes A) = a_2|a_n| \ (or\ a_1|a_{n-1}|) \ge 0.$$

Then

$$\begin{aligned}\varphi(A) &= 6a_1|a_n| + 2a_2|a_n| = 2|a_n|(3a_1 + a_2)\\ &\le 2\sqrt{10}|a_n|\sqrt{a_1^2 + a_2^2} \le \sqrt{10}(a_n^2 + a_1^2 + a_2^2)\\ &\le \sqrt{10}\|A\|^2,\end{aligned}$$

(*or* $\varphi(A) = 6a_1|a_n| + 2a_1|a_{n-1}| = 2|a_1|(3a_n + a_{n-1}) \le \sqrt{10}\|A\|^2$). In conclusion,

$$\sum_{i=1}^{4} \lambda_i(T_X) \le 4\Big(\sigma_1^2(X) + \sigma_2^2(X)\Big) + \sqrt{10}\Big(\|A\|^2 + \|B\|^2\Big) \le (4 + \sqrt{10})|X\|^2.$$

As all eigenvalues arise twice, this completes the proof. □

Remark 5.12. From the proof, one can see that the (non-sharp) upper bounds for the complex version and the real version of Conjecture 3 are both equal to $2 + \sqrt{10}/2$.

Remark 5.13. The reason of why we did not get the optimal upper bound 3 of Conjecture 3 mainly comes from that we divided the Hermitian matrix $K_X^* K_X$ into three parts and estimated them separately. With the following example in mind, we don't know how to sharpen our method to prove Conjecture 3. Set

$$X = \begin{pmatrix} -0.1236 & 0.0334 & 0.0647 \\ -0.4343 & 0.1029 & -0.8833 \\ 0 & 0 & 0 \end{pmatrix}.$$

By numerical calculation we see

$$\sum_{i=1}^{4} \lambda_i(T_X) \approx 5.9814 < 6 < 4\Big(\sigma_1^2(X) + \sigma_2^2(X)\Big) + \phi(X) \approx 7.0554 < 4 + \sqrt{10}.$$

To estimate higher order eigenvalues, we need the following lemma.

Lemma 5.14. *Suppose* $\eta_1, \eta_2, \cdots, \eta_{n_1}$ *and* $\omega_1, \omega_2, \cdots, \omega_{n_2}$ *are nonnegative real numbers and* $r_{ij} \in \{0, 1\}$ *such that*

$$\sum_{i=1}^{n_1} \eta_i^2 + \sum_{i=1}^{n_2} \omega_i^2 = 1, \quad \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} r_{ij} = m.$$

Then we have

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \eta_i \omega_j r_{ij} \le \frac{\sqrt{m}}{2}. \tag{5.3}$$

Proof. Suppose $\eta_1 \geq \cdots \geq \eta_{n_1} \geq 0$ and $\omega_1 \geq \cdots \geq \omega_{n_2} \geq 0$, without loss of generality we can select the following m elements with non-vanishing r_{ij}'s:

- $\eta_1\omega_1 \geq \eta_1\omega_2 \geq \cdots \geq \eta_1\omega_{p_1}$
- $\eta_2\omega_1 \geq \eta_2\omega_2 \geq \cdots \geq \eta_2\omega_{p_2}$
- $\cdots$
- $\eta_t\omega_1 \geq \eta_t\omega_2 \geq \cdots \geq \eta_t\omega_{p_t}$

where $p_1 + p_2 + \cdots + p_t = m$. Thus we complete the proof by

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \eta_i\omega_j r_{ij} = \sum_{i=1}^{t}\eta_i \sum_{j=1}^{p_i}\omega_j \leq \sqrt{\sum_{i=1}^{t}\eta_i^2}\sqrt{\sum_{i=1}^{t}(\sum_{j=1}^{p_i}\omega_j)^2} \leq \sqrt{\sum_{i=1}^{t}\eta_i^2}\sqrt{\sum_{i=1}^{t}p_i\sum_{j=1}^{p_i}\omega_j^2}$$
$$\leq \sqrt{\sum_{i=1}^{n_1}\eta_i^2}\sqrt{m\sum_{j=1}^{n_2}\omega_j^2} \leq \frac{\sqrt{m}}{2}\Big(\sum_{i=1}^{n_1}\eta_i^2 + \sum_{i=1}^{n_2}\omega_i^2\Big) = \frac{\sqrt{m}}{2}. \quad \square$$

Proof of Theorem 1.4. The proof is similar to that of Theorem 1.3. Briefly, by Lemma 2.6 (b) and Lemma 5.14, we have

$$\begin{aligned}
\sum_{i=1}^{2k}\lambda_i(T_X) &\leq \sum_{i=1}^{2k}\lambda_i(K_1) + \sum_{i=1}^{2k}\lambda_i(K_2) \\
&\leq \sum_{i=1}^{2k}\lambda_i(K_1) + 2\sum_{i=1}^{2k}\Big(\lambda_i(-A^t\otimes A) + \lambda_i(B^t\otimes B)\Big) \\
&\leq 2k+1+2\Big(\sqrt{k}\|A\|^2 + \sqrt{k}\|B\|^2\Big) \\
&= 2k+1+2\sqrt{k},
\end{aligned}$$

where $\sum_{i=1}^{2k}\lambda_i(K_1) \leq 2k+1$ follows from

$$\lambda_1(K_1) = 2\sigma_1^2(X) \leq 2,\ \lambda_i(K_1) \leq \lambda_2(K_1) = \sigma_1^2(X) + \sigma_2^2(X) \leq 1 \text{ for } i \geq 2;$$

and

$$\sum_{i=1}^{2k}\lambda_i(-A^t\otimes A) \leq 2\sum_{r=1}^{k}\lambda_{2r-1}(-A^t\otimes A) \leq \sqrt{k}\|A\|^2$$

(similar for $\sum_{i=1}^{2k}\lambda_i(B^t\otimes B) \leq \sqrt{k}\|B\|^2$) follows by setting in Lemma 5.14

$$\begin{cases} \eta_i := a_i/\|A\|, & 1 \leq i \leq n_1, \\ \omega_j := -a_{n_1+j}/\|A\|, & 1 \leq j \leq n - n_1, \end{cases}$$

for $a_1 \geq \cdots \geq a_{n_1} \geq 0 \geq a_{n_1+1} \geq \cdots \geq a_n$ and noticing that now the nonnegative eigenvalues $\lambda_{2r-1}(-A^t \otimes A) = \lambda_{2r}(-A^t \otimes A) = -a_i a_{n_1+j}$ appear in pairs. $\square$

Declaration of competing interest

There is no competing interest.

Acknowledgements

The authors sincerely thank David Wenzel and the referee for their many valuable and constructive suggestions which greatly improved the presentation of the paper.

References

[1] T. Ando, R. Bhatia, Eigenvalue inequalities associated with the Cartesian decomposition, Linear Multilinear Algebra 22 (2) (1987) 15.
[2] K.M.R. Audenaert, Variance bounds, with an application to norm bounds for commutators, Linear Algebra Appl. 432 (5) (2009) 1126–1143.
[3] R. Bhatia, Matrix Analysis, Springer, New York, 1997.
[4] A. Böttcher, D. Wenzel, How big can the commutator of two matrices be and how big is it typically?, Linear Algebra Appl. 403 (2005) 216–228.
[5] A. Böttcher, D. Wenzel, The Frobenius norm and the commutator, Linear Algebra Appl. 429 (8–9) (2008) 1864–1885.
[6] C.M. Cheng, K.S. Fong, W.F. Lei, On some norm inequalities involving the commutator and $XY - YX^T$, Linear Algebra Appl. 438 (6) (2013) 2793–2807.
[7] C.M. Cheng, X.Q. Jin, S.W. Vong, A survey on the Böttcher-Wenzel conjecture and related problems, Oper. Matrices 9 (3) (2015) 659–673.
[8] C.M. Cheng, S.W. Vong, D. Wenzel, Commutators with maximal Frobenius norm, Linear Algebra Appl. 432 (2010) 292–306.
[9] C.M. Cheng, Y. Liang, Some sharp bounds for the commutator of real matrices, Linear Algebra Appl. 521 (2017) 263–282.
[10] T. Choi, Z. Lu, On the DDVV conjecture and the comass in calibrated geometry (I), Math. Z. 260 (2008) 409–429.
[11] P.J. De Smet, F. Dillen, L. Verstraelen, L. Vrancken, A pointwise inequality in submanifold theory, Arch. Math. (Brno) 35 (1999) 115–128.
[12] F. Dillen, J. Fastenakels, J. Veken, Remarks on an inequality involving the normal scalar curvature, in: F. Dillen, I. Van deWoestyne (Eds.), Proceedings of the International Congress on Pure and Applied Differential Geometry PADGE, Brussels, Shaker, Aachen, 2007, pp. 83–92.
[13] J.Q. Ge, DDVV-type inequality for skew-symmetric matrices and Simons-type inequality for Riemannian submersions, Adv. Math. 251 (2014) 62–86.
[14] J.Q. Ge, Z.Z. Tang, A proof of the DDVV conjecture and its equality case, Pacific J. Math. 237 (1) (2008) 87–95.
[15] J.Q. Ge, Z.Z. Tang, A survey on the DDVV conjecture, in: Harmonic Maps and Differential Geometry, in: Contemp. Math., vol. 542, Amer. Math. Soc., Providence, RI, 2011, pp. 247–254.
[16] J.Q. Ge, S. Xu, H.Y. You, Y. Zhou, DDVV-type inequality for Hermitian matrices, Linear Algebra Appl. 529 (2017) 133–147.
[17] J.Q. Ge, F.G. Li, Y. Zhou, Some generalizations of the DDVV-type inequalities, arXiv:1807.07307.
[18] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[19] R.A. Brualdi, Introductory Combinatorics, 5 edition, Prentice-Hall, Upper Saddle, 2009.
[20] Z. Lu, Recent developments of the DDVV conjecture, Bull. Transilv. Univ. Braşov Ser. B (N. S.) 14 (49) (2007) 133–143.
[21] Z. Lu, Normal scalar curvature conjecture and its applications, J. Funct. Anal. 261 (2011) 1284–1308.
[22] Z. Lu, Remarks on the Böttcher-Wenzel inequality, Linear Algebra Appl. 436 (7) (2012) 2531–2535.

[23] Z. Lu, D. Wenzel, The normal Ricci curvature inequality, in: Recent Advances in the Geometry of Submanifolds—Dedicated to the Memory of Franki Dillen (1963–2013), in: Contemporary Mathematics, vol. 674, 2016, pp. 99–110.
[24] Z. Lu, D. Wenzel, Commutator estimates comparising the Frobenius norm - looking back and forth, Oper. Theory Adv. Appl. 259 (2017) 533–559.
[25] D.A. Trifonov, Generalizations of Heisenberg uncertainty relation, Eur. Phys. J. B 29 (2002) 349–353.
[26] S.W. Vong, X.Q. Jin, Proof of Böttcher and Wenzel's conjecture, Oper. Matrices 2 (3) (2008) 435–442.
[27] D. Wenzel, Dominating the commutator, in: Topics in Operator Theory, in: Oper. Theory: Adv. and Appl., vol. 202, 2010, pp. 579–600.
[28] D. Wenzel, K.M.R. Audenaert, Impressions of convexity: an illustration for commutator bounds, Linear Algebra Appl. 433 (11) (2010) 1726–1759.
[29] F. Zhang, Matrix Theory: Basic Results and Techniques, Springer, Berlin, 2011.