hexagon:2611.00179v1

# THE LU–WENZEL CONJECTURE

QIYUE TANG

Abstract. We prove the Lu–Wenzel conjecture for real and complex matrices.

## 1. Introduction

For a square complex matrix $X$, write $X^*$ for its conjugate transpose and $\|X\|_F^2 = \mathrm{Tr}(X^*X)$, where Tr is the sum of the diagonal entries. The commutator of matrices of the same size is $[X,Y] = XY - YX$. Submultiplicativity and the triangle inequality give

$$\|[X,Y]\|_F \leq 2\|X\|_F\|Y\|_F.$$

In 2005, Böttcher and Wenzel asked for the optimal constant and conjectured the sharper inequality

$$\|[X,Y]\|_F^2 \leq 2\|X\|_F^2\|Y\|_F^2 \tag{1}$$

for real square matrices [3]. They proved the two-dimensional case. Vong and Jin established the inequality in every real dimension [14], and Lu obtained an independent proof as part of his work on normal scalar curvature [9]. Böttcher and Wenzel subsequently proved the complex inequality and studied related estimates for unitarily invariant norms [2]. Lu's eigenvalue proof [10] and the account by Lu and Wenzel [12] explain several connections among these approaches.

A second source of commutator inequalities is submanifold geometry. De Smet, Dillen, Verstraelen and Vrancken proved a curvature inequality in codimension two and proposed its extension to arbitrary codimension in 1999 [5]. This became the normal scalar curvature conjecture. Through the Gauss and Ricci equations, it reduces to the matrix inequality

$$\sum_{a,b=1}^{m} \|[B_a,B_b]\|_F^2 \leq \left(\sum_{a=1}^{m}\|B_a\|_F^2\right)^2 \tag{2}$$

for every family of real symmetric matrices $B_1,\dots,B_m$ of the same size, with $m \geq 1$. Ge and Tang proved this inequality and classified its equality cases [7]. Lu independently proved the inequality and used it to obtain geometric pinching results [9]. Thus (1) controls one commutator for arbitrary matrices, whereas (2) controls the total commutator energy of a symmetric family.

Lu and Wenzel brought these questions together in their study of the normal Ricci curvature inequality [11]. Their starting point was an estimate arising in a Simons–Bochner formula, in which one symmetric matrix is distinguished from an orthogonal family. For real square matrices $X, Y_1, \dots, Y_m$ of the same size, with $m \geq 1$, they imposed the bilinear conditions

$$\mathrm{Tr}(Y_a[X,Y_b]) = 0 \qquad (1 \leq a,b \leq m).$$

Two matrices $A, B$ are Frobenius-orthogonal when $\mathrm{Tr}(A^*B) = 0$. For mutually orthogonal $X, Y_1, \dots, Y_m$, their fundamental conjecture predicts

$$\sum_{a=1}^{m}\|[X,Y_a]\|_F^2 \leq \left(\sum_{a=1}^{m}\|Y_a\|_F^2 + \max_{1\leq a\leq m}\|Y_a\|_F^2\right)\|X\|_F^2. \tag{3}$$

*Date*: November 7, 2026.

2020 *Mathematics Subject Classification.* Primary 15A42; Secondary 15A45, 35B50.

*Key words and phrases.* Lu–Wenzel conjecture, matrix commutator, spectral sums, maximum principle.

This is Conjecture 8 of [11]. Proposition 10 there shows that the fundamental conjecture implies (2) for arbitrary real matrix families satisfying $\mathrm{Tr}(B_a[B_b, B_c]) = 0$ for all indices. Symmetric matrices satisfy these constraints because the trace of a product of a symmetric matrix and a skew-symmetric matrix is zero. The constraints therefore preserve the geometric case while allowing nonsymmetric matrices.

Ge, Li, Lu and Zhou [6] reformulated the fundamental conjecture in terms of the spectrum of the commutator operator. For an integer $n \geq 2$, equip $\mathcal{H} = M_n(\mathbb{C})$ with

$$\langle A, B\rangle = \mathrm{Tr}(A^*B),$$

linear in the second variable. The adjoint of a complex linear operator on $\mathcal{H}$ is also denoted by a star. For $X \in \mathcal{H}$, set

$$D_X(Y) = [X, Y], \qquad D_X^*(Y) = [X^*, Y], \qquad T_X = D_X^* D_X.$$

Cyclicity of trace gives the adjoint formula, so $T_X$ is Hermitian and positive semidefinite. Write $\lambda_1(T_X) \geq \cdots \geq \lambda_{n^2}(T_X) \geq 0$ for its eigenvalues counted with multiplicity. The largest eigenvalue is the maximum of $\|[X, Y]\|_F^2$ over unit $Y$, so (1) gives $\lambda_1(T_X) \leq 2$ for unit $X$.

Ge, Li, Lu and Zhou proved that every positive eigenvalue occurs with even multiplicity. They showed that the fundamental conjecture is equivalent to bounding the first $2k$ eigenvalues by $2k + 2$, and extended these equivalences to complex matrices [6]. They established the spectral inequalities when $X$ is normal, when $X$ has rank one, and when $n = 2$ or 3. In general they obtained

$$\lambda_1(T_X) + \lambda_3(T_X) \leq \frac{4 + \sqrt{10}}{2}, \qquad \sum_{j=1}^{2k} \lambda_j(T_X) \leq 2k + 1 + 2\sqrt{k}$$

for $\|X\|_F = 1$ and $1 \leq k \leq \lfloor n^2/2 \rfloor$. These results and their geometric context are reviewed in [13, Section 4.3]. The spectral conjecture asks for the bound $2k + 2$ in every dimension. A preprint of Liu computes the spectrum when $X^2 = 0$ and obtains all the predicted partial sums for this class and its scalar translates [8]. Carptopus proves the inequalities on coordered alignment and isoclinic subfamilies of rank-two matrices, together with stability estimates under a positive lower bound on the smaller singular value [4].

**Theorem 1.** *Let $n \geq 2$, let $I_n$ be the $n \times n$ identity matrix, and let $X \in M_n(\mathbb{C})$. For every integer $1 \leq k \leq \lfloor n^2/2 \rfloor$,*

$$\sum_{j=1}^{2k} \lambda_j(T_X) \leq (2k + 2) \left\| X - \frac{\mathrm{Tr}\, X}{n} I_n \right\|_F^2 .$$

*For each $1 \leq k \leq n - 1$, equality is attained by a nonscalar matrix.*

The squared norm on the right is $\|X\|_F^2 - |\mathrm{Tr}\, X|^2/n$. Theorem 1 proves the complex version of [6, Conjecture 4], numbered Conjecture 7 in the preprint. For real $X$, the complex operator is the complexification of the real one, with the same eigenvalues and multiplicities. When $n = 1$, all commutators vanish.

The proof maximizes the total excess of the eigenvalues above one on the traceless unit sphere. At a maximizer, a strict gap separates the eigenvalues above one from the others. Their sum is smooth in a neighborhood, and its second derivative contains the variation of its spectral projection. Contracting this derivative against commutator directions gives a nonnegative sum of squares. A local maximum principle then permits minimization of the commutator energy within the maximum set. The critical equation, together with membership in the selected spectral subspace, reduces the maximizer to a normal or rank-one matrix.

## 2. Normalization and spectral pairs

Subtracting $(\operatorname{Tr} X)I_n/n$ does not change $D_X$. If the resulting matrix is nonzero, dividing it by its Frobenius norm divides $T_X$ by the square of that norm. For scalar matrices the theorem is immediate. We therefore work on the sphere

$$\mathbb{S} = \{X \in \mathcal{H} : \operatorname{Tr} X = 0,\ \|X\|_F = 1\},$$

with the real metric $\operatorname{Re}\langle \cdot, \cdot \rangle$. Use Id for the identity on $\mathcal{H}$ and $\operatorname{Tr}_{\mathcal{H}}$ for the trace of a complex linear operator on $\mathcal{H}$. From now on, norms of matrices are Frobenius norms. For a Hermitian operator $A$, $A_+$ is obtained by replacing each eigenvalue of $A$ by its maximum with zero. Set

$$G(X) = \operatorname{Tr}_{\mathcal{H}}(T_X - \mathrm{Id})_+ = \sum_{j=1}^{n^2} \max\{\lambda_j(T_X) - 1, 0\}.$$

We shall prove $G(X) \le 2$ on $\mathbb{S}$. This implies the theorem since each sum of $2k$ eigenvalues minus $2k$ is bounded above by $G(X)$.

Unitary conjugation preserves the spectrum, since

$$T_{UXU^*}(UYU^*) = UT_X(Y)U^*$$

for a unitary $n \times n$ matrix $U$, and $Y \mapsto UYU^*$ is an isometry. Let $E_{ij}$, $1 \le i, j \le n$, denote the matrix unit with a one in position $(i, j)$ and zeros elsewhere. The trace identity

$$\operatorname{Tr}_{\mathcal{H}} T_X = 2n\|X\|^2 - 2|\operatorname{Tr} X|^2 \tag{4}$$

follows from

$$\sum_{i,j} \|XE_{ij}\|^2 = \sum_{i,j} \|E_{ij}X\|^2 = n\|X\|^2,$$

$$\sum_{i,j} \langle XE_{ij}, E_{ij}X\rangle = \operatorname{Tr}\left(\left(\sum_{i,j} E_{ji}X^*E_{ij}\right)X\right) = |\operatorname{Tr} X|^2.$$

Here $\sum_{i,j} E_{ji}X^*E_{ij} = (\operatorname{Tr} X^*)I_n$.

Let $\mathcal{J}(Y) = Y^*$, an antiunitary involution of $\mathcal{H}$, so it is conjugate linear and $\langle \mathcal{J}U, \mathcal{J}V\rangle = \langle V, U\rangle$. When the base matrix $X$ is fixed or understood as the variable in a local spectral neighborhood, write $D = D_X$ and $T = T_X$. Since $([X, Y^*])^* = -[X^*, Y]$, one has

$$\mathcal{J}D\mathcal{J} = -D^*, \qquad K := \mathcal{J}D, \qquad K^2 = -T. \tag{5}$$

Consequently $K$ commutes with $T$. On the eigenspace $E_\alpha$ for a positive eigenvalue $\alpha$ of $T$, one has $\langle Ku, Kv\rangle = \alpha\langle v, u\rangle$. Thus $K/\sqrt{\alpha}$ is antiunitary there and squares to $-\mathrm{Id}$. An orthonormal basis can therefore be chosen in pairs $e, f$ with

$$Ke = \sqrt{\alpha}f, \quad Kf = -\sqrt{\alpha}e, \qquad De = \sqrt{\alpha}f^*, \quad Df = -\sqrt{\alpha}e^*. \tag{6}$$

For an antiunitary map $Q$ with $Q^2 = -\mathrm{Id}$, antiunitarity gives $\langle Qe, Q^2e\rangle = \langle Qe, e\rangle$, and $Q^2e = -e$ makes both sides zero. Thus $e \perp Qe$. The orthogonal complement of their span is invariant under $Q$, so induction produces the pairs in (6) and proves even multiplicity.

Conversely, the spectral inequalities imply $G(X) \le 2$ on $\mathbb{S}$. Indeed, the number of eigenvalues above one is even; when it is positive, apply the inequality at this cutoff.

Fix an integer $1 \le r < n^2$. Suppose there is a strict gap between the $r$-th and $(r+1)$-st eigenvalues of $T$. Let $P$ be the orthogonal projection onto the eigenspaces for the $r$ largest eigenvalues, which are positive. With an orthonormal eigenbasis write

$$Tq_i = \alpha_i q_i \quad (i \in I), \qquad Tq_j = \beta_j q_j \quad (j \in O), \qquad \alpha_i > \beta_j \ge 0. \tag{7}$$

The sets $I$ and $O$ partition the $n^2$ basis indices into selected and omitted indices. A sum over all $j$ means $j \in I \cup O$. Entire repeated eigenspaces are selected. For a complex

subspace $U \subset \mathcal{H}$, write $U^* = \{u^* : u \in U\}$. Put $W = \operatorname{ran} P$. Then $D(W) = W^*$, and (6) proves the following identity in the complex tensor product $\mathcal{H} \otimes_{\mathbb{C}} \mathcal{H}$

$$\sum_{i \in I} Dq_i \otimes q_i^* \quad \text{is skew under interchange of its factors.} \tag{8}$$

Indeed each pair contributes $\sqrt{\alpha}(f^* \otimes e^* - e^* \otimes f^*)$. The tensor is independent of the orthonormal selected basis, since the second factor transforms with the conjugate change of coordinates.

## 3. The Hessian of a separated spectral sum

Work on a neighborhood in $\mathbb{S}$ where the strict cut (7) persists. At each point choose a fixed positively oriented contour $\Gamma$ enclosing the selected eigenvalues and excluding the others, also at nearby points. Diagonalizing $T_X$ and integrating the scalar resolvents gives

$$P(X) = \frac{1}{2\pi i} \int_\Gamma (z \operatorname{Id} - T_X)^{-1} \, dz.$$

The formula gives smoothness even at repeated eigenvalues. Define

$$F(X) = \operatorname{Tr}_{\mathcal{H}}(P(X) T_X), \qquad L = \sum_{i \in I} D_{q_i}^* D_{q_i}, \quad J = LX - FX, \quad B = DPD^*.$$

Fix an orthonormal basis $e_1, \dots, e_{n^2}$ of $\mathcal{H}$. Expanding the selected basis vectors in this basis gives

$$L = \sum_{a,b=1}^{n^2} \langle e_b, Pe_a \rangle \, D_{e_a}^* D_{e_b}.$$

Thus $L$ is independent of the selected frame and smooth as a function of $P$. For a tangent vector $H$, differentiating $P^2 = P$ shows that $P'$ has zero diagonal blocks in $\operatorname{ran} P \oplus \ker P$. Since $T$ is block diagonal, $\operatorname{Tr}_{\mathcal{H}}(P'T) = 0$. The term with $P$ fixed is the derivative of $\sum_{i \in I} \|[X, q_i]\|^2$, so $dF(H) = 2 \operatorname{Re} \langle LX, H \rangle$. The identity $\langle X, LX \rangle = \sum_{i \in I} \|[q_i, X]\|^2 = F$ gives the radial component of the gradient. Each value of $D_{q_i}^*$ is a commutator and has trace zero, so $LX$ is traceless. Therefore the real sphere gradient is

$$\nabla_{\mathbb{S}} F = 2J, \qquad \langle X, J \rangle = 0. \tag{9}$$

Define the complex orthogonal projection $\Pi_X$, written $\Pi$ at a fixed base matrix $X$, by

$$\Pi H = H - \frac{\operatorname{Tr} H}{n} I_n - \langle X, H \rangle X.$$

Its image lies in the real tangent space of $\mathbb{S}$ and is closed under multiplication by $i$.

For $Y \in \operatorname{ran} \Pi$, differentiation gives $T_Y' = D^* D_Y + D_Y^* D$. For $i \in I, j \in O$, put

$$A_{ji}(Y) = \langle [Dq_j, q_i^*], Y \rangle, \qquad C_{ji}(Y) = \langle Y, [Dq_i, q_j^*] \rangle.$$

Then $\langle q_j, T_Y' q_i \rangle = A_{ji}(Y) + C_{ji}(Y)$. The adjoint formula and $D_Y q = -D_q Y$ give these two expressions. The first term is complex linear in $Y$, the second conjugate linear. The sphere Hessian $\operatorname{Hess}_{\mathbb{S}} F(Y, Y)$ is the second derivative of $F$ along the sphere geodesic starting with velocity $Y$. It is

$$\operatorname{Hess}_{\mathbb{S}} F(Y, Y) = 2\langle Y, (L - F \operatorname{Id}) Y \rangle + 2 \sum_{i \in I, j \in O} \frac{|A_{ji}(Y) + C_{ji}(Y)|^2}{\alpha_i - \beta_j}. \tag{10}$$

Indeed, differentiating $T_X P = P T_X$ gives the off-diagonal entry

$$\langle q_j, P_Y' q_i \rangle = \langle q_j, T_Y' q_i \rangle / (\alpha_i - \beta_j).$$

The two off-diagonal blocks give

$$\operatorname{Tr}_{\mathcal{H}}(P_Y' T_Y') = 2 \sum_{i \in I, j \in O} \frac{|\langle q_j, T_Y' q_i \rangle|^2}{\alpha_i - \beta_j}.$$

Along the sphere geodesic with initial velocity $Y$, the acceleration is $-\|Y\|^2 X$, and $T'' = 2D_Y^* D_Y - 2\|Y\|^2 T$. Thus $\mathrm{Tr}_{\mathcal{H}}(PT'') = 2\langle Y, LY\rangle - 2F\|Y\|^2$. Differentiating $F' = \mathrm{Tr}_{\mathcal{H}}(PT')$ proves (10).

The identities $A_{ji}(iY) = iA_{ji}(Y)$ and $C_{ji}(iY) = -iC_{ji}(Y)$ give

$$\begin{aligned}|A_{ji}(Y) + C_{ji}(Y)|^2 + |A_{ji}(iY) + C_{ji}(iY)|^2 \\ = 2|A_{ji}(Y)|^2 + 2|C_{ji}(Y)|^2.\end{aligned}$$

Fix a selected eigenvalue $\alpha > 0$ and an omitted eigenvalue $\beta \geq 0$, and sum over $i, j$ in their respective eigenspaces $E_\alpha, E_\beta$. The vectors $Dq_i/\sqrt{\alpha}$ form an orthonormal basis of $E_\alpha^*$ by (6). For $\beta > 0$, the vectors $Dq_j/\sqrt{\beta}$ form one of $E_\beta^*$. Changing orthonormal basis in these two sums gives

$$\sum_{\substack{i:\, q_i \in E_\alpha \\ j:\, q_j \in E_\beta}} |A_{ji}(Y)|^2 = \beta \sum_{\substack{i:\, q_i \in E_\alpha \\ j:\, q_j \in E_\beta}} |\langle [q_i^*, q_j^*], Y\rangle|^2,$$

$$\sum_{\substack{i:\, q_i \in E_\alpha \\ j:\, q_j \in E_\beta}} |C_{ji}(Y)|^2 = \alpha \sum_{\substack{i:\, q_i \in E_\alpha \\ j:\, q_j \in E_\beta}} |\langle [q_i^*, q_j^*], Y\rangle|^2.$$

For $\beta = 0$, $Dq_j = 0$, so the first identity still holds, and the second follows by changing the selected basis alone. The denominator $\alpha - \beta$ is constant in each such sum. For $U \in \mathcal{H}$, let $|U\rangle\langle U|$ denote the operator $H \mapsto U\langle U, H\rangle$. Define the Hermitian operator

$$\mathcal{L} = L - F\mathrm{Id} + \sum_{i \in I, j \in O} \frac{\alpha_i + \beta_j}{\alpha_i - \beta_j} |[q_i^*, q_j^*]\rangle\langle [q_i^*, q_j^*]|. \tag{11}$$

Then

$$\frac{\mathrm{Hess}_{\mathbb{S}} F(Y, Y) + \mathrm{Hess}_{\mathbb{S}} F(iY, iY)}{4} = \langle Y, \mathcal{L}Y\rangle. \tag{12}$$

## 4. Contraction with the commutator covariance

Let $A_i, B_i$ be two matrix families indexed by a common finite set, with all sums over $a, i$ in that set. Suppose that $\sum_i B_i \otimes A_i$ is skew. Put

$$I_0 = \sum_{a,i} \mathrm{Tr}([B_a, A_i^*][A_a, B_i^*]).$$

Then

$$2I_0 = \left\| \sum_i [B_i, A_i] \right\|_F^2. \tag{13}$$

Set $R = \sum_i B_i A_i$. Skewness implies $\sum_i A_i B_i = -R$ and $\sum_i [B_i, A_i] = 2R$. Expand the two commutators in $I_0$ into four terms. The terms with a minus sign are

$$-\sum_{a,i} \mathrm{Tr}(B_a A_i^* B_i^* A_a) = -\mathrm{Tr}((-R)R^*) = \|R\|^2,$$

$$-\sum_{a,i} \mathrm{Tr}(A_i^* B_a A_a B_i^*) = -\mathrm{Tr}((-R^*)R) = \|R\|^2.$$

For the two remaining terms, skewness in the $a$ index and cyclicity give

$$\sum_{a,i} \mathrm{Tr}(A_i^* B_a B_i^* A_a) = -\sum_{a,i} \mathrm{Tr}(A_i^* A_a B_i^* B_a) = -\sum_{a,i} \mathrm{Tr}(B_a A_i^* A_a B_i^*).$$

They therefore cancel. Thus $I_0 = 2\|R\|^2 = \frac{1}{2}\|\sum_i [B_i, A_i]\|^2$.

Apply (13) with $A_i = q_i^*$, $B_i = Dq_i$. Identity (8) supplies skewness, and

$$\sum_{i \in I} [Dq_i, q_i^*] = LX. \tag{14}$$

For any three basis indices $a, i, j$, define

$$u_{aij} = \mathrm{Tr}((Dq_a)[q_i, q_j]).$$

We restrict $a, i$ to $I$ in the following sums, but the definition also gives $u_{jai}$ when $j$ is omitted. Using $u_{aij} = \mathrm{Tr}([Dq_a, q_i]q_j)$, Parseval over the orthonormal basis $q_j^*$ gives

$$\sum_j u_{aij}\overline{u_{iaj}} = \mathrm{Tr}([Dq_a, q_i][q_a^*, (Dq_i)^*]).$$

Therefore (13), (14), and (9) imply the exact identity

$$2\,\mathrm{Re}\sum_{a,i\in I,j\ \text{all}} u_{aij}\overline{u_{iaj}} = \|LX\|^2 = F^2 + \|J\|^2. \tag{15}$$

Since $D$ is a derivation and every commutator has trace zero,

$$0 = \mathrm{Tr}\big(D(q_a[q_i, q_j])\big) = \mathrm{Tr}((Dq_a)[q_i, q_j]) + \mathrm{Tr}(q_a[Dq_i, q_j]) + \mathrm{Tr}(q_a[q_i, Dq_j]).$$

The last two terms equal $-u_{iaj}$ and $u_{jai}$, respectively, by cyclicity. Thus

$$u_{aij} - u_{iaj} + u_{jai} = 0. \tag{16}$$

On $I^3$, permutation of the indices makes the sums of $|u_{aij}|^2$, $|u_{iaj}|^2$, and $|u_{jai}|^2$ equal. Summing $|u_{jai}|^2 = |u_{iaj} - u_{aij}|^2$ gives

$$\sum_{a,i,j\in I} |u_{aij}|^2 = 2\,\mathrm{Re}\sum_{a,i,j\in I} u_{aij}\overline{u_{iaj}}. \tag{17}$$

Introduce

$$K_{II} = \sum_{i,l\in I} |[q_i^*, q_l^*]\rangle\langle[q_i^*, q_l^*]|, \quad R_{IO} = \sum_{i\in I, j\in O} \frac{\beta_j}{\alpha_i - \beta_j} |[q_i^*, q_j^*]\rangle\langle[q_i^*, q_j^*]|.$$

For fixed $i$ and every $H \in \mathcal{H}$, cyclicity and Parseval give

$$\sum_{j\in I\cup O} |\langle[q_i^*, q_j^*], H\rangle|^2 = \sum_{j\in I\cup O} |\langle q_j^*, [q_i, H]\rangle|^2 = \|[q_i, H]\|^2.$$

Summing over $i \in I$ proves the operator identity $\sum_{i\in I, j\in I\cup O} |[q_i^*, q_j^*]\rangle\langle[q_i^*, q_j^*]| = L$, by equality of Hermitian quadratic forms. Using $(\alpha_i + \beta_j)/(\alpha_i - \beta_j) = 1 + 2\beta_j/(\alpha_i - \beta_j)$, we can rewrite (11) as

$$\mathcal{L} = 2L - F\mathrm{Id} - K_{II} + 2R_{IO}. \tag{18}$$

Use the direct covariance expression $B = \sum_{a\in I} |Dq_a\rangle\langle Dq_a|$. Taking its trace against the terms in (18) gives

$$\begin{aligned} \mathrm{Tr}_{\mathcal{H}}(BL) &= \sum_{a,i\in I,j\ \text{all}} |u_{aij}|^2, \\ \mathrm{Tr}_{\mathcal{H}}(BK_{II}) &= \sum_{a,i,j\in I} |u_{aij}|^2, \\ \mathrm{Tr}_{\mathcal{H}}(BR_{IO}) &= \sum_{a,i\in I,j\in O} \frac{\beta_j}{\alpha_i - \beta_j} |u_{aij}|^2. \end{aligned} \tag{19}$$

For the first identity, expand $\sum_{a,i} \|[q_i, Dq_a]\|^2$ in the complete orthonormal starred basis $q_j^*$. For the other two use $\mathrm{Tr}_{\mathcal{H}}(|U\rangle\langle U|\,|V\rangle\langle V|) = |\langle U, V\rangle|^2$ and $\langle[q_i^*, q_j^*], Dq_a\rangle = u_{aij}$. Also $\mathrm{Tr}_{\mathcal{H}} B = F$. Thus

$$\mathrm{Tr}_{\mathcal{H}}(B\mathcal{L}) = \sum_{a,i,j\in I} |u_{aij}|^2 + 2\sum_{a,i\in I,j\in O} \frac{\alpha_i}{\alpha_i - \beta_j} |u_{aij}|^2 - F^2. \tag{20}$$

Subtracting $\|J\|^2$ in (20), substituting (15), and canceling the terms with $j \in I$ by (17) gives

$$\begin{aligned}
&\operatorname{Tr}_{\mathcal{H}}(B\mathcal{L}) - \|J\|^2 \\
&\quad = \sum_{a,i\in I, j\in O} \left( \frac{2\alpha_i}{\alpha_i - \beta_j} |u_{aij}|^2 - 2\operatorname{Re}(u_{aij}\overline{u_{iaj}}) \right) \\
&\quad = \sum_{a,i\in I, j\in O} \left( \frac{\alpha_i}{\alpha_i - \beta_j} |u_{aij}|^2 + \frac{\alpha_a}{\alpha_a - \beta_j} |u_{iaj}|^2 - 2\operatorname{Re}(u_{aij}\overline{u_{iaj}}) \right).
\end{aligned}$$

In the last line we interchanged $a, i$ in one half of the first sum. Define

$$\mathscr{S} = \sum_{a,i\in I, j\in O} \left( |u_{aij} - u_{iaj}|^2 + \frac{\beta_j}{\alpha_i - \beta_j} |u_{aij}|^2 + \frac{\beta_j}{\alpha_a - \beta_j} |u_{iaj}|^2 \right) \geq 0. \tag{21}$$

Expanding each factor $\alpha/(\alpha - \beta)$ as $1 + \beta/(\alpha - \beta)$ in the preceding expression gives

$$\operatorname{Tr}_{\mathcal{H}}(B\mathcal{L}) = \mathscr{S} + \|J\|^2. \tag{22}$$

The strict cut and $\beta_j \geq 0$ suffice for nonnegativity.

For $i \in I, j \in O$,

$$\langle [q_i^*, q_j^*], X \rangle = -\langle q_j^*, Dq_i \rangle = 0,$$

since $Dq_i \in W^*$. Consequently $R_{IO}X = 0$, $K_{II}X = LX$, and $\mathcal{L}X = J$. The projection onto the scalar matrices makes no contribution because $B(I_n) = 0$ and $\mathcal{L}(I_n) = -FI_n$. Expanding the projection off $X$, with $\|X\| = 1$, gives

$$\begin{aligned}
\operatorname{Tr}_{\mathcal{H}}(B\Pi\mathcal{L}\Pi) &= \operatorname{Tr}_{\mathcal{H}}(B\mathcal{L}) - \langle X, \mathcal{L}BX \rangle - \langle X, B\mathcal{L}X \rangle \\
&\quad + \langle X, BX \rangle \langle X, \mathcal{L}X \rangle \\
&= \operatorname{Tr}_{\mathcal{H}}(B\mathcal{L}) - 2\operatorname{Re}\langle J, BX \rangle,
\end{aligned}$$

since $\mathcal{L}X = J$ and $\langle X, J \rangle = 0$. Hence

$$\operatorname{Tr}_{\mathcal{H}}(B\Pi\mathcal{L}\Pi) = \mathscr{S} + \|J\|^2 - 2\operatorname{Re}\langle J, BX \rangle. \tag{23}$$

Using the fixed basis $e_1, \ldots, e_{n^2}$, define the real tangent fields $Z_a(X) = \Pi_X D_X P(X) e_a$, $1 \leq a \leq n^2$. Their covariance satisfies $\sum_{a=1}^{n^2} |Z_a\rangle\langle Z_a| = \Pi B \Pi$. Therefore

$$\sum_{a=1}^{n^2} \langle Z_a, \mathcal{L}Z_a \rangle = \operatorname{Tr}_{\mathcal{H}}(B\Pi\mathcal{L}\Pi), \qquad 4dF(\Pi BX) = 8\operatorname{Re}\langle J, BX \rangle.$$

Together with (12) and (23), these give

$$\sum_{a=1}^{n^2} \big( \operatorname{Hess}_{\mathbb{S}} F(Z_a, Z_a) + \operatorname{Hess}_{\mathbb{S}} F(iZ_a, iZ_a) \big) + 4\, dF(\Pi BX) = 4(\mathscr{S} + \|J\|^2) \geq 0. \tag{24}$$

Set $C(X) = [X^*, X]$ on $\mathbb{S}$, writing $C$ at the current base matrix. Suppose that $X$ is a local maximum of $F$ on $\mathbb{S}$, with $F > 0$. Then $J = 0$ and the real Hessian is nonpositive. Equation (24) forces $\mathscr{S} = 0$. For every omitted $\beta_j > 0$, (21) gives $u_{aij} = u_{iaj} = 0$, and (16) gives $u_{jai} = 0$. Let $O_+$ be the orthogonal sum of the omitted eigenspaces for positive eigenvalues. For subspaces $U, V$, write $[U, V]$ for the linear span of the commutators $[u, v]$, $u \in U, v \in V$. The vectors $(Dq_j)^*$ for omitted $\beta_j > 0$ span $O_+$, by (5)–(6), and $\langle (Dq_j)^*, [q_a, q_i] \rangle = u_{jai} = 0$. Hence

$$[W, W] \perp O_+. \tag{25}$$

Taking matrix adjoints gives $[W^*, W^*] \perp O_+^*$. By (14) and stationarity, $FX = \sum_{i\in I} [Dq_i, q_i^*]$ belongs to $[W^*, W^*]$, so $X \perp O_+^*$. The output singular-space decomposition is

$$\mathcal{H} = W^* \oplus O_+^* \oplus \ker D^*.$$

The map $D^*$ takes $W^*$ into $W$ and kills $\ker D^*$. It follows that the Hermitian matrix satisfies

$$C = D^*X \in W. \tag{26}$$

## 5. Extremal matrices

*Proof of Theorem 1.* On a real coordinate neighborhood of $\mathbb{S}$ with a strict spectral gap, list the $2n^2$ fields $Z_a, iZ_a$, $1 \le a \le n^2$, from (24) as $Z_s$, $1 \le s \le 2n^2$. For smooth tangent fields $U, V$, the sphere's Levi–Civita derivative $\nabla_U V$ is the real orthogonal projection of the ambient derivative of $V$ in direction $U$ onto the tangent space of $\mathbb{S}$. Define

$$Y = 4\Pi BX - \sum_{s=1}^{2n^2} \nabla_{Z_s} Z_s.$$

For a real $C^2$ function $h$ on this neighborhood, set

$$\mathscr{A}h = \sum_{s=1}^{2n^2} Z_s(Z_s h) + Yh.$$

Vector fields act on functions by directional differentiation. Since $\mathrm{Hess}_{\mathbb{S}}\, h(Z_s, Z_s) = Z_s(Z_s h) - (\nabla_{Z_s} Z_s)h$, (24) gives

$$\mathscr{A}F = 4(\mathscr{S} + \|J\|^2) \ge 0.$$

Lemma 5 below applies to $\mathscr{A}$ in the chosen coordinates.

The function $G$ is continuous, by continuity of the ordered eigenvalues, and attains its maximum $M$ on $\mathbb{S}$. Suppose $M > 2$. Among the matrices at which $G = M$, choose $X_0$ minimizing $E(X) = \|[X^*, X]\|^2$. This is possible because the maximum set is compact. Let $r$ count the eigenvalues strictly above one at $X_0$. Then $0 < r < n^2$, since $G(X_0) > 0$ and $D_{X_0}(I_n) = 0$. Choose a connected real coordinate neighborhood $\Omega$ of $X_0$ on which the gap at $r$ remains strict and $\lambda_r(T_X) > 1$. Take $P$ and $F$ for this cutoff. On $\Omega$,

$$F - r \le G \le M,$$

and equality holds throughout this chain at $X_0$. Thus the set where $F = M + r$ in the neighborhood lies in the maximum set of $G$. At each of its points (26) gives $PC = C$, with $C = [X^*, X]$.

Define the smooth real tangent field $V_0 = -\Pi DPC$ on the whole neighborhood. It is a smooth real linear combination of the fields $Z_a, iZ_a$. Indeed, if $C = \sum_a c_a e_a$, its coefficients are $-\operatorname{Re} c_a$ and $-\operatorname{Im} c_a$, respectively. By Lemma 5, the short integral curve of $V_0$ through $X_0$ stays in the maximum set of $F$. On that set $PC = C$ and $\langle X, DC\rangle = \langle D^*X, C\rangle = E$, so

$$V_0 = -DC + EX.$$

For a tangent vector $H$, the derivative of $C$ is $[H^*, X] + [X^*, H]$. Cyclicity gives

$$\operatorname{Tr}(C[H^*, X]) = \operatorname{Tr}(H^*[X, C]), \qquad \operatorname{Tr}(C[X^*, H]) = \operatorname{Tr}([C, X^*]H).$$

Since $C = C^*$, these two scalars are conjugates, and hence

$$dE(H) = 2\operatorname{Tr}\big(C([H^*, X] + [X^*, H])\big) = 4\operatorname{Re}\langle DC, H\rangle.$$

Moreover,

$$\|DC - EX\|^2 = \|DC\|^2 - E^2, \qquad dE(-DC+EX) = -4(\|DC\|^2 - E^2) = -4\|DC - EX\|^2.$$

If $V_0(X_0) \neq 0$, this derivative is strictly negative at time zero along the propagated curve. For small positive time the curve remains in $\{G = M\}$ and decreases $E$, contradicting the choice of $X_0$. Thus $V_0(X_0) = 0$. In particular, at $X_0$,

$$DC = EX_0, \qquad C \in \operatorname{ran} P. \tag{27}$$

All matrix-dependent quantities in the two cases below are evaluated at $X_0$.

If $C = 0$, then $X_0$ is normal. In a unitary basis in which it is diagonal, write $x_1, \dots, x_n$ for its diagonal entries and $x$ for their column vector. The eigenvalues of $T_{X_0}$ are $|x_i - x_j|^2$ over ordered pairs $(i, j)$. Form the simple graph with an edge $ij$ when $|x_i - x_j|^2 > 1$, and let $m$ be its edge count. For $m = 0$ we have $G(X_0) = 0$. For $m > 0$, orient the edges arbitrarily and let $A$ be the vertex-edge incidence matrix, with entries $+1, -1$ at the endpoints of each column and zero elsewhere. Then

$$G(X_0)/2 = \|A^* x\|^2 - m.$$

Write $\deg i$ for the number of edges incident with vertex $i$. For the row indexed by an edge $e = ij$, the diagonal entry of $A^*A$ is two. The absolute off-diagonal entry is one exactly for edges sharing an endpoint with $e$, and, summing over all graph edges $f$,

$$\sum_f |(A^*A)_{ef}| = 2 + (\deg i - 1) + (\deg j - 1) = \deg i + \deg j \le m + 1.$$

The last bound follows because the two incident edge sets overlap only in $ij$. For an eigenvector of $A^*A$, evaluating its eigenvalue equation at a coordinate of largest absolute value shows that the absolute eigenvalue is at most the largest absolute row sum. Multiplication by $A$ and $A^*$ transfers the positive eigenspaces of $A^*A$ and $AA^*$ to each other, so their largest eigenvalues coincide. Hence $\|A^*x\|^2 = \langle x, AA^*x\rangle \le (m+1)\|x\|^2 = m + 1$, and $G(X_0) \le 2$.

Suppose instead that $C \neq 0$. By (27), $TC = ED^*X_0 = EC$. The eigenvalue $E$ exceeds one because $\operatorname{ran} P$ at $X_0$ consists of the eigenspaces above one. Diagonalize $C$ unitarily, with real diagonal entries $c_1, \dots, c_n$. Each nonzero entry $(X_0)_{ij}$ satisfies $c_j - c_i = E$, by the commutator equation in (27). Write $s = \operatorname{rank} X_0$, and choose a nonzero minor of size $s$ with row set $R$ and column set $S$. A nonzero term of its determinant matches the rows to the columns along nonzero entries. Summing the corresponding relations gives

$$sE = \sum_{j \in S} c_j - \sum_{i \in R} c_i.$$

The coefficient of each $c_i$ in this difference is $0, 1$, or $-1$, and at most $2s$ coefficients are nonzero. Since $\sum_i c_i^2 = E$, Cauchy–Schwarz gives $s^2E^2 \le 2sE$. Thus $sE \le 2$, and $E > 1$ forces $s = 1$.

A traceless rank-one unit matrix is unitarily similar to $E_{12}$. In fact, writing it as $uv^*$ with $\|u\| = \|v\| = 1$, the trace condition gives $v^*u = 0$. Completing $u, v$ to an orthonormal basis gives the stated unitary similarity. In the resulting unitary basis $X_0 = E_{12}$, and multiplication gives

$$\begin{aligned} &D(E_{21}) = E_{11} - E_{22}, \quad D((E_{11} - E_{22})/\sqrt{2}) = -\sqrt{2}E_{12}, \\ &D(E_{2j}) = E_{1j}, \qquad\qquad D(E_{j1}) = -E_{j2} \qquad (3 \le j \le n). \end{aligned}$$

The displayed inputs are orthonormal and their outputs are orthogonal. All remaining matrix units except $E_{11}, E_{22}$ are killed by $D$, and $D(E_{11} + E_{22}) = 0$. The resulting spectrum, in decreasing order, is

$$\big(2, 2, \underbrace{1, \dots, 1}_{2n-4}, \underbrace{0, \dots, 0}_{(n-1)^2+1}\big). \tag{28}$$

In particular $G(X_0) = 2$ in the rank-one case.

Both cases contradict $M > 2$. The normalization in Section 2 gives the inequality, and (28) gives equality at $E_{12}$ for the stated cutoffs. □

## 6. Inequalities for matrix families

Let $X$ be any unit matrix. The number $r$ of eigenvalues of $T_X$ strictly above one is even by (6). If $r = 0$, the positive-part trace is zero. Otherwise Theorem 1, applied with

$2k = r$, gives

$$\mathrm{Tr}_{\mathcal{H}}(T_X - \mathrm{Id})_+ = \sum_{j=1}^{r} \lambda_j(T_X) - r \le (r+2)\left\|X - \frac{\mathrm{Tr}\, X}{n} I_n\right\|^2 - r \le 2.$$

The first eigenvalue is at most two by the pairing and the theorem at $k = 1$. At every cutoff $\ell \ge 2$, the partial sum is at most $\ell + 2$ by the positive-part bound and at most $2n$ by (4). The bounds 2 at cutoff one and $\min\{\ell + 2, 2n\}$ thereafter are the partial sums of (28). Append zero eigenvalues at indices beyond $n^2$; the same bounds then apply.

**Corollary 2.** *Let $X, Y_1, \dots, Y_m \in M_n(\mathbb{C})$, where $m \ge 1$ and $Y_1, \dots, Y_m$ are mutually Frobenius-orthogonal. Then*

$$\sum_{a=1}^{m} \|[X, Y_a]\|^2 \le \left(\sum_{a=1}^{m} \|Y_a\|^2 + 2 \max_{1\le a\le m} \|Y_a\|^2\right) \|X\|^2. \tag{29}$$

*If additionally $\mathrm{Tr}(Y_a[X, Y_b]) = 0$ for all $1 \le a, b \le m$, then*

$$\sum_{a=1}^{m} \|[X, Y_a]\|^2 \le \left(\sum_{a=1}^{m} \|Y_a\|^2 + \max_{1\le a\le m} \|Y_a\|^2\right) \|X\|^2. \tag{30}$$

*Each inequality implies the spectral conjecture when imposed for all $X$ and all families satisfying its hypotheses.*

*Proof.* If $X = 0$ or all $Y_a = 0$, the inequalities hold directly. Otherwise take $\|X\| = 1$ by homogeneity, remove the zero members, and relabel the remaining family as $Y_1, \dots, Y_m$. Set $w_a = \|Y_a\|^2$ and $e_a = Y_a/\sqrt{w_a}$, with $w_1 \ge \cdots \ge w_m > 0$ after reordering. The operator $A = \sum_a w_a |e_a\rangle\langle e_a|$ is positive semidefinite and satisfies $A \le w_1 \mathrm{Id}$, in the order of Hermitian operators. For positive semidefinite operators $R, S$, $\mathrm{Tr}_{\mathcal{H}}(RS) = \mathrm{Tr}_{\mathcal{H}}(S^{1/2} R S^{1/2}) \ge 0$. Apply this to $A$, $w_1\mathrm{Id} - A$, and the spectral positive part below. Since $T_X \le \mathrm{Id} + (T_X - \mathrm{Id})_+$, we obtain

$$\sum_a \|[X, Y_a]\|^2 = \mathrm{Tr}_{\mathcal{H}}(T_X A) \le \mathrm{Tr}_{\mathcal{H}}\, A + w_1\, \mathrm{Tr}_{\mathcal{H}}(T_X - \mathrm{Id})_+ \le \sum_a w_a + 2w_1.$$

Under the trace conditions, the span of the $Y_a$ is orthogonal to its image under $K = \mathcal{J} D_X$, since $\langle Ku, v\rangle = -\mathrm{Tr}(u[X, v])$. Let $W$ be a subspace of this span of dimension $d \ge 1$. Write $P_W$ for its orthogonal projection. Choose an orthonormal basis $u_1, \dots, u_d$ diagonalizing the compression $P_W T_X|_W$, the operator that applies $T_X$ and projects back to $W$. Let $\mu_i \ge 0$ be the corresponding eigenvalues. For each $\mu_i > 0$, set $v_i = Ku_i/\sqrt{\mu_i}$. The vectors $v_i$ are orthonormal because $\langle Ku_i, Ku_j\rangle = \langle u_j, T_X u_i\rangle$. They lie in $W^\perp$ and satisfy

$$\langle v_i, T_X v_i\rangle = \frac{\langle u_i, T_X^2 u_i\rangle}{\mu_i} \ge \mu_i,$$

using $KT_X = T_X K$ and Cauchy–Schwarz. For any orthonormal set of size $b$, the sum of its diagonal entries for a positive semidefinite operator is at most the sum of the first $b$ eigenvalues. Indeed, expanding the set in an eigenbasis gives weights in $[0, 1]$ that sum to $b$, and their weighted eigenvalue sum is maximized by filling the largest entries first. Apply this to the $u_i$ and the available $v_i$, at most $2d$ vectors. The partial sum bound then yields $2\sum_i \mu_i \le 2d + 2$.

Use this estimate on the spans of $e_1, \dots, e_j$ for each $1 \le j \le m$. Setting $w_{m+1} = 0$, summation by parts gives

$$\sum_{a=1}^{m} w_a \langle e_a, T_X e_a\rangle = \sum_{j=1}^{m} (w_j - w_{j+1}) \sum_{a=1}^{j} \langle e_a, T_X e_a\rangle \le \sum_{j=1}^{m} (w_j - w_{j+1})(j+1) = \sum_{a=1}^{m} w_a + w_1.$$

Rescaling proves (30).

For the converse to (29), use $2k$ orthonormal eigenvectors for the largest eigenvalues. For (30), use one vector from each of the first $k$ positive spectral pairs in (6), or from every positive pair if there are fewer. Their span is orthogonal to its image under $K$, and the sum of the energies is half of the first $2k$ eigenvalue sum. With $h$ chosen vectors, the bound is $h+1 \le k+1$, and when $h=0$ all energies vanish. These chosen vectors belong to positive eigenspaces and are orthogonal to $X$, since $T_X X = 0$. □

**Corollary 3.** *Let $m \ge 1$. If $B_1, \dots, B_m \in M_n(\mathbb{C})$ satisfy $\operatorname{Tr}(B_a[B_b, B_c]) = 0$ for all $1 \le a, b, c \le m$, then*

$$\sum_{a,b=1}^{m} \|[B_a, B_b]\|^2 \le \left( \sum_{a=1}^{m} \|B_a\|^2 \right)^2 .$$

*Proof.* The all-zero family satisfies the inequality. Otherwise normalize $\sum_a \|B_a\|^2 = 1$ and maximize the left side on the compact set of normalized families satisfying the trace conditions. For a unitary $m \times m$ matrix $U = (U_{ab})$, replace the family by $B'_a = \sum_b U_{ab} B_b$. Orthogonality of the columns of $U$ gives $\sum_a \|B'_a\|^2 = \sum_a \|B_a\|^2$. The commutators transform by

$$[B'_a, B'_b] = \sum_{c,d} U_{ac} U_{bd} [B_c, B_d],$$

so the same orthogonality on both indices preserves the total energy. Every triple trace transforms as a linear combination of the original triple traces, so it remains zero. If $\Gamma_{ab} = \langle B_a, B_b \rangle$ is the Gram matrix, then the new one is $\overline{U} \Gamma U^{\mathsf{T}}$, where $^{\mathsf{T}}$ is transpose. The spectral theorem chooses a unitary $U$ making this diagonal. Thus a maximizing family may be chosen orthogonal. Reorder it so that $w_1 \ge w_2 \ge \cdots$, where $w_a = \|B_a\|^2$. If only one matrix is nonzero the energy is zero.

Write $Q$ for the maximum and $S_1 = \sum_{b \ne 1} \|[B_1, B_b]\|^2$ for the contribution with first index one. For $t > 0$, multiply $B_1$ by $t$ and divide the whole family by $\sqrt{1 + (t^2 - 1) w_1}$. The total squared norm becomes

$$\frac{t^2 w_1 + \sum_{a=2}^{m} w_a}{1 + (t^2 - 1) w_1} = 1.$$

Each triple trace is multiplied by the product of the scalars applied to its three matrices, so the trace constraints are preserved. The energy along this admissible path is

$$\frac{Q + 2(t^2 - 1) S_1}{(1 + (t^2 - 1) w_1)^2}.$$

Its derivative at $t = 1$ equals $4S_1 - 4Qw_1$. Maximality makes it zero, giving $S_1 = Q w_1$. Applying (30) with $X = B_1$ gives $S_1 \le w_1(1 - w_1 + w_2) \le w_1$. As $w_1 > 0$, this gives $Q \le 1$. □

## 7. A rank bound for the bracket on bivectors

Let $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ be the subspace of traceless matrices with the Frobenius metric. The exterior square $\Lambda^2 \mathfrak{g}$ is spanned by the bilinear alternating products $U \wedge V$, with $U \wedge V = -V \wedge U$. Choose a real orthonormal basis $F_1, \dots, F_{n^2-1}$ of the traceless skew-Hermitian matrices, which satisfy $F_i^* = -F_i$. It is a complex orthonormal basis of $\mathfrak{g}$, since every $X \in \mathfrak{g}$ has the decomposition

$$X = \frac{X - X^*}{2} + i \, \frac{X + X^*}{2i}$$

into complex multiples of traceless skew-Hermitian matrices. The exterior metric is defined, for $U_1, U_2, V_1, V_2 \in \mathfrak{g}$, by

$$\langle U_1 \wedge U_2, V_1 \wedge V_2 \rangle = \langle U_1, V_1 \rangle \langle U_2, V_2 \rangle - \langle U_1, V_2 \rangle \langle U_2, V_1 \rangle.$$

Thus $F_i \wedge F_j$, $1 \le i < j \le n^2-1$, form an orthonormal basis, and the metric is preserved by a unitary change of basis. Write $b : \Lambda^2\mathfrak{g} \to \mathfrak{g}$ for the linear map $b(U \wedge V) = [U, V]$, and $b^* : \mathfrak{g} \to \Lambda^2\mathfrak{g}$ for its adjoint with respect to these metrics. For $\eta = \sum_{1\le i<j\le n^2-1} H_{ij} F_i \wedge F_j$, extend its coefficients to a skew-symmetric matrix $H$ by $H_{ji} = -H_{ij}$ and $H_{ii} = 0$. The rank of $H$ is independent of the basis, since a change of basis acts by invertible congruence.

**Corollary 4.** *For $n \ge 2$ and $1 \le k \le \lfloor n^2/2 \rfloor$, one has*

$$\|b\eta\|^2 \le (k+1)\|\eta\|^2 \qquad \text{if } \operatorname{rank} H \le 2k.$$

*These inequalities are equivalent to Theorem 1.*

*Proof.* For $1 \le i, j, \ell \le n^2-1$, put

$$c_{ij\ell} = \langle [F_i, F_j], F_\ell \rangle = -\operatorname{Tr}([F_i, F_j] F_\ell) = -\operatorname{Tr}(F_i [F_j, F_\ell]).$$

These numbers are real because the Frobenius inner product of skew-Hermitian matrices is real. The first bracket makes them antisymmetric in $i, j$, and the last bracket makes them antisymmetric in $j, \ell$; thus they are fully alternating. For $X = \sum_{\ell=1}^{n^2-1} x_\ell F_\ell$, the full skew coefficient matrix $C$ of $b^*X$, with $C_{ji} = -C_{ij}$ and $C_{ii} = 0$, has entries, for $i < j$,

$$C_{ij} = \langle F_i \wedge F_j, b^*X \rangle = \langle [F_i, F_j], X \rangle = \sum_{\ell=1}^{n^2-1} c_{ij\ell} x_\ell.$$

The matrix of $D_X$ restricted to $\mathfrak{g}$ has $i, j$ entry $\sum_{\ell=1}^{n^2-1} c_{\ell ji} x_\ell = -C_{ij}$. Consequently the squared singular values of $C$, with one extra zero for the scalar direction, give the spectrum of $T_X$.

The exterior metric gives $\|H\|_F^2 = 2\|\eta\|^2$ and $\langle b^*X, \eta \rangle = \operatorname{Tr}(C^*H)/2$. If $P$ is the orthogonal projection onto the image of $H$, then $PH = H$, so

$$|\langle b^*X, \eta \rangle|^2 \le \tfrac{1}{4}\|PC\|_F^2 \|H\|_F^2 \le \tfrac{1}{2}\|\eta\|^2 \sum_{j=1}^{2k} \lambda_j(T_X).$$

The last step bounds the trace of $CC^*$ on a subspace of dimension at most $2k$ by the corresponding eigenvalue sum, as in the proof of Corollary 2.

For fixed $X$ the estimate is attained on some unit $\eta$ of rank at most $2k$. To see this when $C \ne 0$, the conjugate-linear map $v \mapsto \overline{Cv}$ squares to $-C^*C$ since $C^{\mathsf{T}} = -C$, where $^{\mathsf{T}}$ is the ordinary transpose. It commutes with $C^*C$, because its square is $-C^*C$, and its squared norm at $v$ is $\langle v, C^*Cv \rangle$. Thus, on an eigenspace of $C^*C$ for a positive eigenvalue $\alpha$, division by $\sqrt{\alpha}$ gives an antiunitary map of square $-\mathrm{Id}$. The construction leading to (6) supplies orthonormal pairs there. If $U$ is the unitary matrix with columns given by these pairs and an orthonormal kernel basis, $U^{\mathsf{T}}CU$ consists of two-by-two skew blocks and zeros. This follows because the $i, j$ entry of $U^{\mathsf{T}}CU$ is the conjugate of the $i$-th coordinate of $\overline{CU_j}$ in the columns of $U$, with $U_j$ the $j$-th column. Retain the $k$ blocks of largest singular values, or all if fewer, and undo the congruence. The resulting skew matrix has squared Frobenius norm $\sum_{j=1}^{2k} \lambda_j(T_X)$, and its Frobenius inner product with $C$ equals that same sum. Multiplying this matrix by $\sqrt{2/\sum_{j=1}^{2k} \lambda_j(T_X)}$ makes it the coefficient matrix of a unit bivector $\eta$ for which

$$|\langle b^*X, \eta \rangle|^2 = \frac{1}{4} \frac{2}{\sum_{j=1}^{2k} \lambda_j(T_X)} \left( \sum_{j=1}^{2k} \lambda_j(T_X) \right)^2 = \frac{1}{2} \sum_{j=1}^{2k} \lambda_j(T_X).$$

If $C = 0$, the estimate is identically zero.

Taking the supremum over unit $X \in \mathfrak{g}$ recovers $\|b\eta\|^2$ on the left by duality. Conversely the inequality for $b\eta$ bounds the maximum over rank-constrained unit $\eta$ by $(k+1)\|X\|^2$. The attainment just proved gives the spectral inequality for traceless $X$, and subtracting the scalar part gives Theorem 1. □

## Appendix A. Local maximum propagation

The following local form of maximum propagation is part of the maximum principle of Bony [1].

**Lemma 5.** *Let $d, N \geq 1$ be integers, let $\Omega \subset \mathbb{R}^d$ be open, and let $Z_1, \ldots, Z_N, Y$ be smooth real vector fields on $\Omega$. Define the differential operator*

$$Ah = \sum_{s=1}^{N} Z_s(Z_s h) + Yh$$

*for real $h \in C^2(\Omega)$. Suppose $m \in \mathbb{R}$, $u \in C^2(\Omega; \mathbb{R})$, $u \leq m$, and $Au \geq 0$ on $\Omega$. If $W = \sum_{s=1}^{N} f_s Z_s$ for smooth real functions $f_s$, then every sufficiently short integral curve of $W$ through a point of $\{u = m\}$ stays in $\{u = m\}$.*

*Proof.* Let $K = \{u = m\}$. Suppose an open ball of center $c$ and radius $R$ in $\Omega \setminus K$ touches $K$ at $p \in \Omega$. Replacing the center and radius by $p + t(c - p)$ and $tR$ with $0 < t < 1$ puts the new closed ball inside the original open ball except at $p$, by the triangle inequality and its equality case. The normal direction at $p$ is unchanged. Thus we may assume the closed ball lies in $\Omega$ and meets $K$ only at $p$.

In coordinates write $Z_s = \sum_{i=1}^{d} z_s^i \partial_i$ and $Y = \sum_{i=1}^{d} y^i \partial_i$, with $\partial_i = \partial/\partial x_i$. Set

$$a^{ij} = \sum_{s=1}^{N} z_s^i z_s^j, \qquad b^i = y^i + \sum_{s=1}^{N} \sum_{j=1}^{d} z_s^j \partial_j z_s^i \qquad (1 \leq i, j \leq d).$$

Expanding $Z_s^2$ gives

$$A = \sum_{i,j=1}^{d} a^{ij} \partial_i \partial_j + \sum_{i=1}^{d} b^i \partial_i.$$

The matrix $(a^{ij})$ is positive semidefinite. We show $Z_s(p) \cdot (p - c) = 0$ for every $s$. If their squares sum to a positive number, take $v(x) = e^{-\kappa |x-c|^2} - e^{-\kappa R^2}$ for $\kappa > 0$. Differentiation gives

$$Av(p) = e^{-\kappa R^2} \left( 4\kappa^2 \sum_s (Z_s(p) \cdot (p - c))^2 - 2\kappa \left( \sum_i a^{ii}(p) + \sum_i b^i(p)(p_i - c_i) \right) \right).$$

Choose $\kappa$ large enough that $Av(p) > 0$, and take a small ball $U$ centered at $p$ with $\overline{U} \subset \Omega$ and $Av > 0$ on $\overline{U}$. The compact part of $\partial U$ in the closed exterior ball is disjoint from $K$. On it $u$ is uniformly below $m$ if this part is nonempty. For sufficiently small $\varepsilon > 0$, $u + \varepsilon v < m$ there. Elsewhere on $\partial U$ we have $v < 0$, giving the same strict inequality. Since $u(p) + \varepsilon v(p) = m$, the function $u + \varepsilon v$ attains an interior maximum in $U$. At that point $A(u + \varepsilon v) \leq 0$ by the vanishing gradient and the nonpositive Hessian, whereas $Au + \varepsilon Av > 0$ on $U$. The contradiction proves the asserted tangency. In particular $W(p) \cdot (p - c) = 0$ for every such exterior ball.

Fix $x_0 \in K$, choose $\rho > 0$ with $\overline{B(x_0, 2\rho)} \subset \Omega$, and set $K_0 = K \cap \overline{B(x_0, 2\rho)}$. Let $\gamma$ be the integral curve of $W$ from $x_0$ on a short interval $[0, T]$ with image in $B(x_0, \rho/2)$. For each time choose a nearest point $p \in K_0$ to $\gamma(t)$, using compactness. Then $|\gamma(t) - p| \leq |\gamma(t) - x_0| < \rho/2$, so

$$|p - x_0| \leq |p - \gamma(t)| + |\gamma(t) - x_0| < \rho.$$

If $\gamma(t) \neq p$, the ball centered at $\gamma(t)$ of radius $|\gamma(t) - p|$ lies in $B(x_0, \rho)$. Every point of $K$ in that ball would belong to $K_0$ and contradict minimality of the distance. Hence the ball avoids all of $K$. Thus $W(p) \cdot (\gamma(t) - p) = 0$, also valid when $\gamma(t) = p$.

The squared distance $f(t) = \operatorname{dist}(\gamma(t), K_0)^2$ is Lipschitz on this interval, since distance to a set is 1-Lipschitz. Let $L$ be a Lipschitz constant for $W$ on $\overline{B(x_0, 2\rho)}$. At any point of

differentiability of $f$, holding the nearest point $p$ fixed for positive increments gives

$$f'(t) \le 2(\gamma(t)-p)\cdot W(\gamma(t)) = 2(\gamma(t)-p)\cdot (W(\gamma(t))-W(p)) \le 2Lf(t).$$

Since $f(0)=0$, the absolutely continuous function $e^{-2Lt}f(t)$ is nonincreasing and $f=0$. The same argument with $-W$ gives propagation in negative time. □

## References


[1] J.-M. Bony, *Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés*, Ann. Inst. Fourier (Grenoble) **19** (1969), no. 1, 277–304.

[2] A. Böttcher and D. Wenzel, *The Frobenius norm and the commutator*, Linear Algebra Appl. **429** (2008), 1864–1885.

[3] A. Böttcher and D. Wenzel, *How big can the commutator of two matrices be and how big is it typically?*, Linear Algebra Appl. **403** (2005), 216–228.

[4] Carptopus, *Coordered alignment and isoclinic stability in the rank-two complex Lu–Wenzel problem*, preprint, October 18, 2026, author's repository.

[5] P. J. De Smet, F. Dillen, L. Verstraelen, and L. Vrancken, *A pointwise inequality in submanifold theory*, Arch. Math. (Brno) **35** (1999), 115–128.

[6] J. Ge, F. Li, Z. Lu, and Y. Zhou, *On some conjectures by Lu and Wenzel*, Linear Algebra Appl. **592** (2020), 134–164. Preprint arXiv:1908.06624.

[7] J. Ge and Z. Tang, *A proof of the DDVV conjecture and its equality case*, Pacific J. Math. **237** (2008), no. 1, 87–95.

[8] Z. Liu, *The Lu–Wenzel spectral conjecture for square-zero matrices*, preprint, version 1, Zenodo record 21693383.

[9] Z. Lu, *Normal scalar curvature conjecture and its applications*, J. Funct. Anal. **261** (2011), no. 5, 1284–1308.

[10] Z. Lu, *Remarks on the Böttcher–Wenzel inequality*, Linear Algebra Appl. **436** (2012), no. 7, 2531–2535.

[11] Z. Lu and D. Wenzel, *The normal Ricci curvature inequality*, Contemp. Math. **674**, Amer. Math. Soc., Providence, RI, 2016, 99–110.

[12] Z. Lu and D. Wenzel, *Commutator estimates comprising the Frobenius norm—looking back and forth*, Oper. Theory Adv. Appl. **259**, Birkhäuser/Springer, Cham, 2017, 533–559.

[13] J. Ge, F. Li, Z. Tang, and Y. Zhou, *A survey on the DDVV-type inequalities*, Adv. Math. (China) **53** (2024), no. 3, 449–467.

[14] S.-W. Vong and X.-Q. Jin, *Proof of Böttcher and Wenzel's conjecture*, Oper. Matrices **2** (2008), no. 3, 435–442.



Zhili College, Tsinghua University, Beijing 100084, China

*Email address:* tangqy24@mails.tsinghua.edu.cn