Matrix Trace Inequalities Related to Uncertainty Principle

Kosaki, Hideki
Faculty of Mathematics, Kyushu University

出版情報：International Journal of Mathematics. 16 (6), pp.629-645, 2005-07. World Scientific
バージョン：
権利関係：

MHF Preprint Series

Kyushu University
21st Century COE Program
Development of Dynamic Mathematics with High Functionality

Matrix trace inequalities related to uncertainty principle

H. Kosaki

MHF 2005-1 [1]

(Received January 11, 2005)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

[1]This paper is also published in Kyushu University Preprint Series in Mathematics as No. 2005-1.

MATRIX TRACE INEQUALITIES RELATED TO UNCERTAINTY PRINCIPLE

HIDEKI KOSAKI

Abstract. In their recent article S. Luo and Z. Zhang conjectured the matrix trace inequality mentioned in the introduction below, which is motivated by uncertainty principle. We present a proof for the conjectured inequality.

1. Introduction

For a fixed density matrix ρ the covariance

$$\mathrm{Cov}_\rho(X,Y) = \mathrm{Tr}(\rho XY^*) - \mathrm{Tr}(\rho X)\cdot\overline{\mathrm{Tr}(\rho Y)}$$

gives rise to a positive sesquilinear form on the space of matrices. Heisenberg's uncertainty principle (in the matrix setting) states

$$\tfrac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2 \le \mathrm{Var}_\rho(A)\cdot\mathrm{Var}_\rho(B)$$

for observables (i.e., self-adjoint matrices) A, B. Here, $\mathrm{Var}_\rho(A)$ $(= \mathrm{Cov}_\rho(A,A) = \mathrm{Tr}(\rho A^2) - (\mathrm{Tr}(\rho A))^2)$ is the variance of A relative to ρ. This estimate is a consequence of the Cauchy-Schwarz inequality, and the left side here (i.e., the commutator part) arises from the imaginary part $\frac{1}{2i}\mathrm{Tr}(\rho[A,B])$ of the covariance $\mathrm{Cov}_\rho(A,B)$. The real part can be expressed in terms of a certain anti-commutator, and a stronger estimate including its effect (known as Schrödinger's uncertainty principle) is also possible (see §2, especially Remark 2). On the other hand,

$$\mathrm{Corr}_\rho(A,B) = \mathrm{Tr}(\rho AB) - \mathrm{Tr}(\rho^{1/2}A\rho^{1/2}B)$$

is called the Wigner-Yanase correlation ([7]), and the (Wigner-Yanase) skew information $\mathrm{I}(\rho,A) = \mathrm{Corr}_\rho(A,A)$ $(= -\frac{1}{2}\mathrm{Tr}([\rho^{1/2},A]^2))$ has been investigated by many authors as a certain measurement of information content of ρ relative to A. Quite a thorough account on these and related subjects can be found in [6].

In the recent article [5] S. Luo and Z. Zhang discussed various aspects of quantum measurement and the above-mentioned uncertainty principle with strong emphasis on the use of the skew information $\mathrm{I}(\rho,A)$. Among other things the inequality

$$\tfrac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2 + \tfrac{1}{16}|\mathrm{I}(\rho,A+B) - \mathrm{I}(\rho,A-B)|^2 \le \mathrm{I}(\rho,A)\cdot\mathrm{I}(\rho,B)$$

1991 *Mathematics Subject Classification.* 15A45; Secondary 47A30, 47A50, 81P15.

Key words and phrases. trace inequality, uncertainty principle, Wigner-Yanase skew information.

was claimed as a main result in their article (see [5, Theorem 2 in p.1571]), which is equivalent to a certain Cauchy-Schwarz type estimate (see §2 and §4.1 for details). Motivated by this claim, they conjectured

$$\begin{aligned}\mathrm{Var}_\rho A \cdot \mathrm{Var}_\rho B - \tfrac{1}{4}|\mathrm{Cov}_\rho(A,B) + \mathrm{Cov}_\rho(B,A)|^2 \\ \geq \mathrm{I}(\rho,A)\cdot \mathrm{I}(\rho,B) - \tfrac{1}{4}|\mathrm{Corr}_\rho(A,B) + \mathrm{Corr}_\rho(B,A)|^2\end{aligned}$$

(see [5, p.1572]), and its validity for 2×2 matrices and some other cases was checked. The conjecture means that the claimed inequality is stronger than Schrödinger's uncertainty principle in the sense that the former (together with the conjectured inequality) implies the latter.

The purpose of the article is to present a proof for their conjecture (i.e., the above second inequality). More generally we will deal with the famous one-parameter family $\{\mathrm{I}(\rho,A;\alpha)\}_{0<\alpha<1}$ introduced by Dyson: $\mathrm{I}(\rho,A;\alpha) = -\frac{1}{2}\mathrm{Tr}([\rho^\alpha,A][\rho^{1-\alpha},A])$. Our main result (Theorem 5 in §3) actually covers this family, and we will also clarify the equality condition for the inequalities in question (Proposition 6). In §4.1 we will point out that the above first inequality (i.e., the one claimed as [5, Theorem 2] by S. Luo and Z. Zhang) is unfortunately false.

The author would like to thank H. Araki for informing him of the conjecture studied in the article and also F. Hiai for discussions on presented materials.

2. Notations and preliminaries

We will denote the set of all $n\times n$ matrices (resp. all $n\times n$ self-adjoint matrices) by $M_n(\mathbf{C})$ (resp. $M_n(\mathbf{C})_{sa}$). General matrices will be denoted by $X, Y, \cdots$ while letters $A, B, \cdots$ will be used to express self-adjoint ones. We assume that $\rho \in M_n(\mathbf{C})$ is a (distinguished) density matrix, i.e., a positive matrix satisfying $\mathrm{Tr}(\rho) = 1$.

Lemma 1. *We assume* $0 < \alpha < 1$.

(i) *We have*

$$|\mathit{Tr}(\rho X)|^2 \leq \mathit{Tr}(\rho^\alpha X \rho^{1-\alpha} X^*), \;\; X \in M_n(\mathbf{C}).$$

In particular, $A \in M_n(\mathbf{C})_{sa}$ *satisfies* $(\mathit{Tr}(\rho A))^2 \leq \mathit{Tr}(\rho^\alpha A \rho^{1-\alpha} A)$.

(ii) *A normal matrix* $X \in M_n(\mathbf{C})$ *satisfies*

$$\mathit{Tr}(\rho^\alpha X \rho^{1-\alpha} X^*) \leq \mathit{Tr}(\rho |X|^2),$$

and hence $A \in M_n(\mathbf{C})_{sa}$ *satisfies* $\mathit{Tr}(\rho^\alpha A \rho^{1-\alpha} A) \leq \mathit{Tr}(\rho A^2)$.

Proof. Let us consider the map

$$(X,Y) \in M_n(\mathbf{C}) \times M_n(\mathbf{C}) \to \mathrm{Tr}(\rho^\alpha X \rho^{1-\alpha} Y^*) \in \mathbf{C}.$$

Since

$$\overline{\mathrm{Tr}(\rho^\alpha Y \rho^{1-\alpha} X^*)} = \mathrm{Tr}(X\rho^{1-\alpha} Y^* \rho^\alpha) = \mathrm{Tr}(\rho^\alpha X \rho^{1-\alpha} Y^*),$$

the above gives rise to a sesquilinear form. It is also positive, i.e., $\mathrm{Tr}(\rho^\alpha X \rho^{1-\alpha} X^*) \geq 0$, and the Cauchy-Schwarz inequality tells

$$|\mathrm{Tr}(\rho X)|^2 = |\mathrm{Tr}(\rho^\alpha X \rho^{1-\alpha} 1^*)|^2 \leq \mathrm{Tr}(\rho^\alpha X \rho^{1-\alpha} X^*) \cdot \mathrm{Tr}(\rho^\alpha 1 \rho^{1-\alpha} 1^*),$$

showing (i) because of $\mathrm{Tr}(\rho) = 1$.

To show (ii), we consider the bounded continuous function $f(z) = \mathrm{Tr}(\rho^z X \rho^{1-z} X^*)$ on the strip $0 \leq \Re z \leq 1$, which is analytic on the interior. Boundary values admit the following upper bounds:

$$\begin{aligned}
|f(it)| &= |\mathrm{Tr}(\rho^{it} X \rho^{1/2} \rho^{-it} \rho^{1/2} X^*)| \\
&\leq ||\rho^{it} X \rho^{1/2} \rho^{-it}||_{HS} ||X \rho^{1/2}||_{HS} = ||X\rho^{1/2}||_{HS}^2, \\
|f(1+it)| &= |\mathrm{Tr}(\rho^{it} \rho X \rho^{-it} X^*)| = |\mathrm{Tr}(\rho^{it} \rho^{1/2} X \rho^{-it} X^* \rho^{1/2})| \\
&\leq ||\rho^{it} \rho^{1/2} X \rho^{-it}||_{HS} ||\rho^{1/2} X||_{HS} = ||\rho^{1/2} X||_{HS}^2.
\end{aligned}$$

Here, $||\cdot||_{HS}$ means the Hilbert-Schmidt norm and $t \in \mathbf{R}$. The normality of X implies $||X\rho^{1/2}||_{HS}^2 = ||\rho^{1/2} X||_{HS}^2 = \mathrm{Tr}(\rho |X|^2)$ so that the desired estimate follows from the maximum modulus principle. □

The estimate in Lemma 1,(ii) means $||\rho^{\alpha/2} X \rho^{(1-\alpha)/2}||_{HS} \leq ||\rho^{1/2} X||_{HS}$. Extensive investigation on more general estimates in this nature (for general unitarily invariant norms) was carried out in [2].

For a fixed $\alpha \in (0,1)$ and $X, Y \in M_n(\mathbf{C})$ we set

$$\mathrm{Corr}_\rho(X, Y; \alpha) = \mathrm{Tr}(\rho X Y^*) - \mathrm{Tr}(\rho^\alpha X \rho^{1-\alpha} Y^*), \text{ correlation.}$$

Hence, for $A, B \in M_n(\mathbf{C})_{sa}$ we have

$$\begin{aligned}
\mathrm{Corr}_\rho(A, B; \alpha) &= \mathrm{Tr}(\rho A B) - \mathrm{Tr}(\rho^\alpha A \rho^{1-\alpha} B), \\
\mathrm{Corr}_\rho(A, A; \alpha) &= -\tfrac{1}{2} \mathrm{Tr}([\rho^\alpha, A][\rho^{1-\alpha}, A]).
\end{aligned}$$

In the literature (see [1, 6, 7] for instance) the quantities

$$\mathrm{I}(\rho, A; 1/2) = \mathrm{Corr}_\rho(A, A; 1/2)$$

and $\mathrm{Corr}_\rho(A, B; 1/2)$ are often called the Wigner-Yanase skew information and correlation, and the former is regarded as a certain measurement of information content of the density ρ with respect to A. Its generalization to

$$\mathrm{I}(\rho, A; \alpha) = \mathrm{Corr}_\rho(A, A; \alpha) \ (\geq 0)$$

was suggested by Dyson, and the convexity of $\mathrm{I}(\rho, A; \alpha)$ as a function of ρ was established in the celebrated work [4] by Lieb. We will denote $\mathrm{Corr}_\rho(A, B; 1/2)$ and $\mathrm{I}(\rho, A; 1/2)$ simply by $\mathrm{Corr}_\rho(A, B)$ and $\mathrm{I}(\rho, A)$ respectively.

We have just observed $\mathrm{Corr}_\rho(X, X; \alpha) \geq 0$ for X normal (Lemma 1,(ii)), but let us emphasize that the sesquilinear form

$$(X, Y) \in M_n(\mathbf{C}) \times M_n(\mathbf{C}) \to \mathrm{Corr}_\rho(X, Y; \alpha) \in \mathbf{C} \tag{1}$$

is not positive, i.e., $\mathrm{Corr}_\rho(X,X;\alpha) \not\geq 0$. Indeed, for

$$\rho = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad X = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$$

we easily compute

$$\mathrm{Corr}_\rho(X,X;\alpha) = \lambda_1^\alpha \left(\lambda_1^{1-\alpha} - \lambda_2^{1-\alpha}\right)|a|^2 + \lambda_2^\alpha \left(\lambda_2^{1-\alpha} - \lambda_1^{1-\alpha}\right)|b|^2$$

(which is negative for either $(a,b) = (1,0)$ or $(0,1)$ as long as $\lambda_1 \neq \lambda_2$). On the other hand, although the quantity $\mathrm{Corr}_\rho(A,B;\alpha)$ may not be real, the correspondence

$$(A,B) \in M_n(\mathbf{C})_{sa} \times M_n(\mathbf{C})_{sa} \to \Re\,\mathrm{Corr}_\rho(A,B;\alpha) \in \mathbf{R}$$

certainly gives rise to a positive bilinear form on the real vector space $M_n(\mathbf{C})_{sa}$ so that the Cauchy-Schwarz inequality yields

$$(\Re\,\mathrm{Corr}_\rho(A,B;\alpha))^2 \leq \mathrm{I}(\rho,A;\alpha)\cdot \mathrm{I}(\rho,B;\alpha). \tag{2}$$

S. Luo and Z. Zhang derived the inequality mentioned in §1 (i.e., [5, Theorem 2]) from the stronger estimate $|\mathrm{Corr}_\rho(A,B)|^2 \leq \mathrm{I}(\rho,A)\cdot\mathrm{I}(\rho,B)$ (for $\alpha = 1/2$). However, (1) is not a positive sesquilinear form on the vector space $M_n(\mathbf{C})$ and the author sees no a priori reason for the validity of such a strong Cauchy-Schwarz type estimate (on the absolute value). The above (non-) positivity seems to be the point overlooked in [5], and [5, Theorem 2] (or equivalently, the above estimate on the absolute value) is indeed false as will be clarified in §4.1.

Remark 2. *The correspondence*

$$(X,Y) \in M_n(\mathbf{C}) \times M_n(\mathbf{C}) \to \langle X,Y\rangle = Tr(\rho XY^*) - Tr(\rho X)\cdot \overline{Tr(\rho Y)} \in \mathbf{C}$$

is a positive sesquilinear form. Indeed, the positivity is a consequence of the Cauchy-Schwarz inequality (*applied to* $(X,Y) \to Tr(\rho XY^*)$):

$$|\,Tr(\rho X)|^2 = |\,Tr(\rho X1^*)|^2 \leq Tr(\rho XX^*)\cdot Tr(\rho 11^*) = Tr(\rho XX^*).$$

For $A,B \in M_n(\mathbf{C})_{sa}$ *we set*

$$Cov_\rho(A,B) = \langle A,B\rangle \;\left(= Tr(\rho AB) - Tr(\rho A)\cdot Tr(\rho B)\right), \text{ covariance},$$
$$Var_\rho A = \langle A,A\rangle \;\left(= Tr(\rho A^2) - (Tr(\rho A))^2\right), \text{ variance}.$$

Lemma 1,(*i*) *means*

$$Var_\rho(A) \geq I(\rho,A;\alpha)$$

(*see* [5, Theorem 1]), *and the Cauchy-Schwarz inequality* (*applied to* $\langle\cdot,\cdot\rangle$) *tells*

$$|\,Cov_\rho(A,B)|^2 \leq Var_\rho A\cdot Var_\rho B. \tag{3}$$

The product $Tr(\rho A) \cdot Tr(\rho B)$ being real, we have

$$\begin{aligned} \Re\, Cov_\rho(A,B) &= \Re\, Tr(\rho AB) - Tr(\rho A)\cdot Tr(\rho B) \\ &= \tfrac{1}{2}\left(Tr(\rho AB) + Tr(\rho BA) - 2\,Tr(\rho A)\cdot Tr(\rho B)\right), \\ \Im\, Cov_\rho(A,B) &= \Im\, Tr(\rho AB) = \tfrac{1}{2i}\left(Tr(\rho AB) - Tr(\rho BA)\right). \end{aligned}$$

By setting $A_0 = A - Tr(\rho A)1$, $B_0 = B - Tr(\rho B)1$, we easily observe

$$Tr(\rho AB) + Tr(\rho BA) - 2\,Tr(\rho A)\cdot Tr(\rho B) = Tr(\rho\{A_0, B_0\})$$

with the anti-commutator $\{A_0, B_0\} = A_0B_0 + B_0A_0$ so that (3) *means*

$$\tfrac{1}{4}|\,Tr(\rho\{A_0,B_0\})|^2 + \tfrac{1}{4}|\,Tr(\rho[A,B])|^2 \le Var_\rho A \cdot Var_\rho B.$$

This inequality is known as Schrödinger's uncertainty principle whereas the weaker (*commutator*) *estimate*

$$\tfrac{1}{4}|\,Tr(\rho[A,B])|^2 \le Var_\rho A \cdot Var_\rho B$$

is referred to as Heisenberg's uncertainty principle in the literature.

3. Main result

In [5] the inequality

$$\begin{aligned} \text{Var}_\rho A \cdot \text{Var}_\rho B - \tfrac{1}{4}|\text{Cov}_\rho(A,B) + \text{Cov}_\rho(B,A)|^2 & \\ \ge \text{I}(\rho,A)\cdot \text{I}(\rho,B) - \tfrac{1}{4}|\text{Corr}_\rho(A,B) + \text{Corr}_\rho(B,A)|^2 & \end{aligned} \tag{4}$$

was conjectured (and its validity was checked for 2×2 matrices and some other cases). We prove the conjectured inequality (4) in this section. We will actually obtain a one-parameter version involving quantities such as $\text{I}(\cdot;\alpha)$ and $\text{Corr}_\rho(\cdot,\cdot;\alpha)$, $\alpha \in (0,1)$.

We may and do assume that the density matrix ρ is diagonalized so that in the rest we will assume

$$\rho = \text{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

with $\lambda_i > 0$. The general case $\lambda_i \ge 0$ can be obtained by the limiting argument thanks to the obvious continuity of relevant quantities. We begin by expressing these quantities in terms of matrix components.

Lemma 3. *For self-adjoint matrices $A = [a_{ij}]$, $B = [b_{ij}] \in M_n(\mathbf{C})_{sa}$ we have*

$$\Re\, Tr(\rho AB) = \sum_{1\le i<j\le n} \alpha_{ij}\, \Re\left(a_{ij}\overline{b_{ij}}\right) + \sum_{i=1}^n \lambda_i a_{ii} b_{ii} \quad \text{with} \quad \alpha_{ij} = \lambda_i + \lambda_j,$$

and

$$\Re\, Tr(\rho^\alpha A \rho^{1-\alpha} B) = \sum_{1\le i<j\le n} \beta_{ij}\, \Re\left(a_{ij}\overline{b_{ij}}\right) + \sum_{i=1}^n \lambda_i a_{ii} b_{ii} \quad \text{with} \quad \beta_{ij} = \lambda_i^\alpha \lambda_j^{1-\alpha} + \lambda_i^{1-\alpha}\lambda_j^\alpha.$$

In particular, with $A = B$ we have

$$\begin{aligned} Tr(\rho A^2) &= \sum_{i<j} \alpha_{ij}|a_{ij}|^2 + \sum_i \lambda_i a_{ii}^2, \\ Tr(\rho B^2) &= \sum_{i<j} \alpha_{ij}|b_{ij}|^2 + \sum_i \lambda_i b_{ii}^2, \\ Tr(\rho^\alpha A \rho^{1-\alpha} A) &= \sum_{i<j} \beta_{ij}|a_{ij}|^2 + \sum_i \lambda_i a_{ii}^2, \\ Tr(\rho^\alpha B \rho^{1-\alpha} B) &= \sum_{i<j} \beta_{ij}|b_{ij}|^2 + \sum_i \lambda_i b_{ii}^2. \end{aligned}$$

Proof. It is plain to see

$$\mathrm{Tr}(\rho^\alpha A \rho^{1-\alpha} B) = \sum_{i,j=1}^n \lambda_i^\alpha \lambda_j^{1-\alpha} a_{ij}\overline{b_{ij}} = \sum_{i \neq j} \lambda_i^\alpha \lambda_j^{1-\alpha} a_{ij}\overline{b_{ij}} + \sum_i \lambda_i a_{ii} b_{ii}$$

because diagonal components a_{ii}, b_{ii} are real. Then, the trivial fact $\Re\left(a_{ji}\overline{b_{ji}}\right) = \Re\left(\overline{a_{ij}\overline{b_{ij}}}\right) = \Re\left(a_{ij}\overline{b_{ij}}\right)$ yields

$$\begin{aligned} \Re\,\mathrm{Tr}(\rho^\alpha A \rho^{1-\alpha} B) &= \sum_{i \neq j} \lambda_i^\alpha \lambda_j^{1-\alpha} \Re\left(a_{ij}\overline{b_{ij}}\right) + \sum_i \lambda_i a_{ii} b_{ii} \\ &= \sum_{i<j} \left(\lambda_i^\alpha \lambda_j^{1-\alpha} + \lambda_j^\alpha \lambda_i^{1-\alpha}\right) \Re\left(a_{ij}\overline{b_{ij}}\right) + \sum_i \lambda_i a_{ii} b_{ii}, \end{aligned}$$

which is exactly the second equation in the lemma. The first is just a special case of the second (with $\alpha = 1$). □

It is easy to see

$$\left\{ \begin{array}{l} \mathrm{Corr}_\rho(A + a1, B + b1; \alpha) = \mathrm{Corr}_\rho(A, B; \alpha), \\ \mathrm{Var}_\rho(A + a1, B + b1) = \mathrm{Var}_\rho(A, B) \end{array} \right. \tag{5}$$

for $a, b \in \mathbf{R}$. Thanks to this invariance, to prove (4) or its one-parameter version, (by considering $A - \mathrm{Tr}(\rho A)1$ and $B - \mathrm{Tr}(\rho B)1$) we may and do assume $\mathrm{Tr}(\rho A) = \mathrm{Tr}(\rho B) = 0$. This assumption means

$$\mathrm{Var}_\rho A = \mathrm{Tr}(\rho A^2),\ \mathrm{Var}_\rho B = \mathrm{Tr}(\rho B^2),\ \mathrm{Cov}_\rho(A, B) = \mathrm{Tr}(\rho AB)$$

and hence

$$\begin{aligned} &\mathrm{Var}_\rho A \cdot \mathrm{Var}_\rho B - \tfrac{1}{4}|\mathrm{Cov}_\rho(A, B) + \mathrm{Cov}_\rho(B, A)|^2 \\ &= \mathrm{Tr}(\rho A^2) \cdot \mathrm{Tr}(\rho B^2) - \left(\Re\,\mathrm{Tr}(\rho AB)\right)^2. \end{aligned}$$

On the other hand, we compute

$$\begin{aligned}
&\mathrm{I}(\rho,A;\alpha)\cdot\mathrm{I}(\rho,B;\alpha)-\tfrac{1}{4}|\mathrm{Corr}_\rho(A,B;\alpha)+\mathrm{Corr}_\rho(B,A;\alpha)|^2\\
&\quad=\Big(\mathrm{Tr}(\rho A^2)-\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}A\Big)\Big(\mathrm{Tr}(\rho B^2)-\mathrm{Tr}(\rho^\alpha B\rho^{1-\alpha}B\Big)\\
&\qquad\quad-\Big(\Re\,\mathrm{Tr}(\rho AB)-\Re\,\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}B)\Big)^2\\
&\quad=\mathrm{Tr}(\rho A^2)\cdot\mathrm{Tr}(\rho B^2)+\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}A)\cdot\mathrm{Tr}(\rho^\alpha B\rho^{1-\alpha}B)\\
&\qquad\quad-\mathrm{Tr}(\rho A^2)\cdot\mathrm{Tr}(\rho^\alpha B\rho^{1-\alpha}B)-\mathrm{Tr}(\rho B^2)\cdot\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}A)\\
&\qquad\quad-(\Re\,\mathrm{Tr}(\rho AB))^2-\Big(\Re\,\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}B)\Big)^2\\
&\qquad\quad+2\Big(\Re\,\mathrm{Tr}(\rho AB)\Big)\Big(\Re\,\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}B)\Big).
\end{aligned}$$

We set

$$\begin{aligned}
&I=\mathrm{Tr}(\rho A^2)\cdot\mathrm{Tr}(\rho^\alpha B\rho^{1-\alpha}B)+\mathrm{Tr}(\rho B^2)\cdot\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}A),\\
&J=2\Big(\Re\,\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}B)\Big)\Big(\Re\,\mathrm{Tr}(\rho AB)\Big),\\
&K=\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}A)\cdot\mathrm{Tr}(\rho^\alpha B\rho^{1-\alpha}B)-\Big(\Re\,\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}B)\Big)^2,
\end{aligned}$$

and observe that (4) (or rather its one-parameter version) means $I-J-K\geq 0$.

Note that I consists of products of "first terms" (i.e., off-diagonal entries), products of "second terms" (i.e., diagonal entries), and "cross terms". More explicitly, thanks to Lemma 3 we have

$$I=I_{11}+I_{12}+I_{22}$$

with

$$\begin{aligned}
I_{11}&=\left(\sum_{i<j}\alpha_{ij}|a_{ij}|^2\right)\left(\sum_{i<j}\beta_{ij}|b_{ij}|^2\right)+\left(\sum_{i<j}\alpha_{ij}|b_{ij}|^2\right)\left(\sum_{i<j}\beta_{ij}|a_{ij}|^2\right),\\
I_{12}&=\left(\sum_{i<j}\alpha_{ij}|a_{ij}|^2\right)\left(\sum_{i}\lambda_i b_{ii}^2\right)+\left(\sum_{i<j}\beta_{ij}|b_{ij}|^2\right)\left(\sum_{i}\lambda_i a_{ii}^2\right)\\
&\qquad+\left(\sum_{i<j}\alpha_{ij}|b_{ij}|^2\right)\left(\sum_{i}\lambda_i a_{ii}^2\right)+\left(\sum_{i<j}\beta_{ij}|a_{ij}|^2\right)\left(\sum_{i}\lambda_i b_{ii}^2\right)\\
&=\left(\sum_{i<j}(\alpha_{ij}+\beta_{ij})\,|a_{ij}|^2\right)\left(\sum_{i}\lambda_i b_{ii}^2\right)+\left(\sum_{i<j}(\alpha_{ij}+\beta_{ij})\,|b_{ij}|^2\right)\left(\sum_{i}\lambda_i a_{ii}^2\right),\\
I_{22}&=2\left(\sum_{i}\lambda_i a_{ii}^2\right)\left(\sum_{i}\lambda_i b_{ii}^2\right).
\end{aligned}$$

It is plain to see

$$I_{11}=\sum_{i<j,\,k<\ell}(\alpha_{ij}\beta_{k\ell}+\alpha_{k\ell}\beta_{ij})\,|a_{ij}|^2|b_{k\ell}|^2.$$

The coefficients here are unchanged under the transpositions $(i,j) \leftrightarrow (k,\ell)$, which enables us to rewrite I_{11} in the following "more symmetric" fashion:

$$I_{11} = \frac{1}{2} \sum_{i<j,\, k<\ell} (\alpha_{ij}\beta_{k\ell} + \alpha_{k\ell}\beta_{ij}) \left(|a_{ij}|^2 |b_{k\ell}|^2 + |a_{k\ell}|^2 |b_{ij}|^2 \right). \tag{6}$$

Similarly, we express J, K as

$$J = J_{11} + J_{12} + J_{22}, \quad K = K_{11} + K_{12} + K_{22},$$

and (to prove $I - J - K \geq 0$) we will look at

$$I_{11} - J_{11} - K_{11},\ I_{12} - J_{12} - K_{12},\ I_{22} - J_{22} - K_{22}$$

separately. Let us recall

$$J = 2 \left(\sum_{i<j} \beta_{ij} \Re \left(a_{ij} \overline{b_{ij}} \right) + \sum_i \lambda_i a_{ii} b_{ii} \right) \left(\sum_{i<j} \alpha_{ij} \Re \left(a_{ij} \overline{b_{ij}} \right) + \sum_i \lambda_i a_{ii} b_{ii} \right)$$

(see Lemma 3). From this expression we conclude

$$\begin{aligned}
J_{11} &= 2 \left(\sum_{i<j} \beta_{ij} \Re \left(a_{ij} \overline{b_{ij}} \right) \right) \left(\sum_{i<j} \alpha_{ij} \Re \left(a_{ij} \overline{b_{ij}} \right) \right), \\
J_{12} &= 2 \left(\sum_{i<j} \beta_{ij} \Re \left(a_{ij} \overline{b_{ij}} \right) \right) \left(\sum_i \lambda_i a_{ii} b_{ii} \right) + 2 \left(\sum_{i<j} \alpha_{ij} \Re \left(a_{ij} \overline{b_{ij}} \right) \right) \left(\sum_i \lambda_i a_{ii} b_{ii} \right) \\
&= 2 \left(\sum_{i<j} (\alpha_{ij} + \beta_{ij}) \Re \left(a_{ij} \overline{b_{ij}} \right) \right) \left(\sum_i \lambda_i a_{ii} b_{ii} \right), \\
J_{22} &= 2 \left(\sum_i \lambda_i a_{ii} b_{ii} \right)^2.
\end{aligned}$$

We have

$$\begin{aligned}
J_{11} &= 2 \sum_{i<j,\, k<\ell} \alpha_{ij}\beta_{k\ell} \Re \left(a_{ij} \overline{b_{ij}} \right) \Re \left(a_{k\ell} \overline{b_{k\ell}} \right) \\
&= \sum_{i<j,\, k<\ell} (\alpha_{ij}\beta_{k\ell} + \alpha_{k\ell}\beta_{ij}) \Re \left(a_{ij} \overline{b_{ij}} \right) \Re \left(a_{k\ell} \overline{b_{k\ell}} \right)
\end{aligned} \tag{7}$$

because the products $\Re \left(a_{ij} \overline{b_{ij}} \right) \Re \left(a_{k\ell} \overline{b_{k\ell}} \right)$ are unchanged under the transpositions $(i,j) \leftrightarrow (k,\ell)$.

It remains to compute K_{11}, K_{12}, K_{22}. Lemma 3 implies

$$\begin{aligned} K &= \left(\sum_{i<j}\beta_{ij}|a_{ij}|^2+\sum_i \lambda_i a_{ii}^2\right)\left(\sum_{i<j}\beta_{ij}|b_{ij}|^2+\sum_i \lambda_i b_{ii}^2\right) \\ &\qquad -\left(\sum_{i<j}\beta_{ij}\,\Re\left(a_{ij}\overline{b_{ij}}\right)+\sum_i \lambda_i a_{ii}b_{ii}\right)^2, \end{aligned}$$

and we observe

$$\begin{aligned} K_{11} &= \left(\sum_{i<j}\beta_{ij}|a_{ij}|^2\right)\left(\sum_{i<j}\beta_{ij}|b_{ij}|^2\right)-\left(\sum_{i<j}\beta_{ij}\,\Re\left(a_{ij}\overline{b_{ij}}\right)\right)^2, \\ K_{12} &= \left(\sum_{i<j}\beta_{ij}|a_{ij}|^2\right)\left(\sum_i \lambda_i b_{ii}^2\right)+\left(\sum_{i<j}\beta_{ij}|b_{ij}|^2\right)\left(\sum_i \lambda_i a_{ii}^2\right) \\ &\qquad -2\left(\sum_{i<j}\beta_{ij}\,\Re\left(a_{ij}\overline{b_{ij}}\right)\right)\left(\sum_i \lambda_i a_{ii}b_{ii}\right), \\ K_{22} &= \left(\sum_i \lambda_i a_{ii}^2\right)\left(\sum_i \lambda_i b_{ii}^2\right)-\left(\sum_i \lambda_i a_{ii}b_{ii}\right)^2. \end{aligned}$$

We note

$$\begin{aligned} K_{11} &= \sum_{i<j,\,k<\ell}\beta_{ij}\beta_{k\ell}|a_{ij}|^2|b_{k\ell}|^2-\sum_{i<j,\,k<\ell}\beta_{ij}\beta_{k\ell}\,\Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right) \\ &= \frac{1}{2}\sum_{i<j,\,k<\ell}\beta_{ij}\beta_{k\ell}\left(|a_{ij}|^2|b_{k\ell}|^2+|a_{k\ell}|^2|b_{ij}|^2\right) \\ &\qquad -\sum_{i<j,\,k<\ell}\beta_{ij}\beta_{k\ell}\,\Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right) \\ &= \frac{1}{2}\sum_{i<j,\,k<\ell}\beta_{ij}\beta_{k\ell}\left(|a_{ij}|^2|b_{k\ell}|^2+|a_{k\ell}|^2|b_{ij}|^2-2\,\Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right)\right), \end{aligned} \tag{8}$$

where the second equality follows from the invariance of the products $\beta_{ij}\beta_{k\ell}$ under the transpositions $(i,j)\leftrightarrow(k,\ell)$.

We are now ready to look at $I_{ij}-J_{ij}-K_{ij}$ $((i,j)=(1,1),(1,2),(22))$. The easiest one to deal with is $I_{22}-J_{22}-K_{22}$. Indeed, it is plain to see

$$I_{22}-J_{22}-K_{22}=\left(\sum_i \lambda_i a_{ii}^2\right)\left(\sum_i \lambda_i b_{ii}^2\right)-\left(\sum_i \lambda_i a_{ii}b_{ii}\right)^2. \tag{9}$$

From (6) and (7) we get

$$I_{11}-J_{11}=\frac{1}{2}\sum_{i<j,\,k<\ell}(\alpha_{ij}\beta_{k\ell}+\alpha_{k\ell}\beta_{ij})\left(|a_{ij}|^2|b_{k\ell}|^2+|a_{k\ell}|^2|b_{ij}|^2-2\,\Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right)\right).$$

Therefore, (8) implies

$$I_{11} - J_{11} - K_{11} = \frac{1}{2} \sum_{i<j,\, k<\ell} \left(\beta_{ij}\left(\alpha_{k\ell} - \tfrac{1}{2}\beta_{k\ell}\right) + \beta_{k\ell}\left(\alpha_{ij} - \tfrac{1}{2}\beta_{ij}\right)\right) \times \left(|a_{ij}|^2|b_{k\ell}|^2 + |a_{k\ell}|^2|b_{ij}|^2 - 2\,\Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right)\right). \tag{10}$$

Finally, all the cross terms sum up to

$$\begin{aligned} & I_{12} - J_{12} - K_{12} \\ &= \left(\sum_{i<j}(\alpha_{ij}+\beta_{ij})\,|a_{ij}|^2\right)\left(\sum_i \lambda_i b_{ii}^2\right) + \left(\sum_{i<j}(\alpha_{ij}+\beta_{ij})\,|b_{ij}|^2\right)\left(\sum_i \lambda_i a_{ii}^2\right) \\ &\quad -2\left(\sum_{i<j}(\alpha_{ij}+\beta_{ij})\,\Re\left(a_{ij}\overline{b_{ij}}\right)\right)\left(\sum_i \lambda_i a_{ii}b_{ii}\right) \\ &\quad -\left(\sum_{i<j}\beta_{ij}|a_{ij}|^2\right)\left(\sum_i \lambda_i b_{ii}^2\right) - \left(\sum_{i<j}\beta_{ij}|b_{ij}|^2\right)\left(\sum_i \lambda_i a_{ii}^2\right) \\ &\quad +2\left(\sum_{i<j}\beta_{ij}\Re\left(a_{ij}\overline{b_{ij}}\right)\right)\left(\sum_i \lambda_i a_{ii}b_{ii}\right) \\ &= \left(\sum_{i<j}\alpha_{ij}|a_{ij}|^2\right)\left(\sum_i \lambda_i b_{ii}^2\right) + \left(\sum_{i<j}\alpha_{ij}|b_{ij}|^2\right)\left(\sum_i \lambda_i a_{ii}^2\right) \\ &\quad -2\left(\sum_{i<j}\alpha_{ij}\Re\left(a_{ij}\overline{b_{ij}}\right)\right)\left(\sum_i \lambda_i a_{ii}b_{ii}\right). \end{aligned}$$

The first term in the above last expression can be rewritten as

$$\left(\sum_{i<j}\alpha_{ij}|a_{ij}|^2\right)\left(\sum_i \lambda_i b_{ii}^2\right) = \sum_k \lambda_k \left(\sum_{i<j}\alpha_{ij}b_{kk}^2|a_{ij}|^2\right).$$

Note that the first factor in each of the last three terms contains the identical coefficients α_{ij}, which enables us to rearrange the above sum into

$$\begin{aligned} I_{12} - J_{12} - K_{12} &= \sum_k \lambda_k \left(\sum_{i<j}\alpha_{ij}\left(b_{kk}^2|a_{ij}|^2 + a_{kk}^2|b_{ij}|^2 - 2a_{kk}b_{kk}\Re\left(a_{ij}\overline{b_{ij}}\right)\right)\right) \\ &= \sum_k \lambda_k \left(\sum_{i<j}\alpha_{ij}|b_{kk}a_{ij} - a_{kk}b_{ij}|^2\right). \end{aligned} \tag{11}$$

Summing up the computations so far (see (9), (10) and (11)) and recalling the definitions of α_{ij}, β_{ij} in Lemma 3, we conclude

Lemma 4. *We have*

$$
\begin{aligned}
&I - J - K \\
&\quad = \frac{1}{2}\sum_{i<j,\,k<\ell}\Big\{\left(\lambda_i^{\alpha}\lambda_j^{1-\alpha}+\lambda_i^{1-\alpha}\lambda_j^{\alpha}\right)\left(\lambda_k+\lambda_\ell-\tfrac{1}{2}\left(\lambda_k^{\alpha}\lambda_\ell^{1-\alpha}+\lambda_k^{1-\alpha}\lambda_\ell^{\alpha}\right)\right) \\
&\qquad\qquad + \left(\lambda_k^{\alpha}\lambda_\ell^{1-\alpha}+\lambda_k^{1-\alpha}\lambda_\ell^{\alpha}\right)\left(\lambda_i+\lambda_j-\tfrac{1}{2}\left(\lambda_i^{\alpha}\lambda_j^{1-\alpha}+\lambda_i^{1-\alpha}\lambda_j^{\alpha}\right)\right)\Big\} \\
&\qquad\qquad \times\left(|a_{ij}|^2|b_{k\ell}|^2+|a_{k\ell}|^2|b_{ij}|^2-2\,\Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right)\right) \\
&\quad + \sum_k \lambda_k\left(\sum_{i<j}(\lambda_i+\lambda_j)\,|b_{kk}a_{ij}-a_{kk}b_{ij}|^2\right) \\
&\quad + \left(\sum_i \lambda_i a_{ii}^2\right)\left(\sum_i \lambda_i b_{ii}^2\right)-\left(\sum_i \lambda_i a_{ii}b_{ii}\right)^2 .
\end{aligned}
$$

The last part $\left(\sum_i \lambda_i a_{ii}^2\right)\left(\sum_i \lambda_i b_{ii}^2\right)-\left(\sum_i \lambda_i a_{ii}b_{ii}\right)^2$ is non-negative by the Cauchy-Schwarz inequality while the arithmetic-geometric mean inequality guarantees

$$
\begin{aligned}
\frac{|a_{ij}|^2|b_{k\ell}|^2+|a_{k\ell}|^2|b_{ij}|^2}{2} &\geq |a_{ij}|\,|b_{k\ell}|\times|a_{k\ell}|\,|b_{ij}| = |a_{ij}\overline{b_{ij}}|\times|a_{k\ell}\overline{b_{k\ell}}| \\
&\geq \left|\Re\left(a_{ij}\overline{b_{ij}}\right)\right|\times\left|\Re\left(a_{k\ell}\overline{b_{k\ell}}\right)\right| \\
&\geq \Re\left(a_{ij}\overline{b_{ij}}\right)\Re\left(a_{k\ell}\overline{b_{k\ell}}\right).
\end{aligned}
\tag{12}
$$

It remains to check behavior of the coefficient

$$
\begin{aligned}
\tilde f(\alpha) &= \left(\lambda_i^{\alpha}\lambda_j^{1-\alpha}+\lambda_i^{1-\alpha}\lambda_j^{\alpha}\right)\left(\lambda_k+\lambda_\ell-\tfrac{1}{2}\left(\lambda_k^{\alpha}\lambda_\ell^{1-\alpha}+\lambda_k^{1-\alpha}\lambda_\ell^{\alpha}\right)\right) \\
&\quad + \left(\lambda_k^{\alpha}\lambda_\ell^{1-\alpha}+\lambda_k^{1-\alpha}\lambda_\ell^{\alpha}\right)\left(\lambda_i+\lambda_j-\tfrac{1}{2}\left(\lambda_i^{\alpha}\lambda_j^{1-\alpha}+\lambda_i^{1-\alpha}\lambda_j^{\alpha}\right)\right)
\end{aligned}
$$

(for each fixed i, j, k, ℓ) appearing in the first sum. We obviously have the symmetry $\tilde f(\alpha)=\tilde f(1-\alpha)$ ($\alpha\in(0,1)$), and for the special value $\alpha=1/2$ we have

$$
\tilde f(1/2) = 2\sqrt{\lambda_i\lambda_j}\left(\lambda_k+\lambda_\ell-\sqrt{\lambda_k\lambda_\ell}\right)+2\sqrt{\lambda_k\lambda_\ell}\left(\lambda_i+\lambda_j-\sqrt{\lambda_i\lambda_j}\right)>0,
$$

showing $I-J-K\geq 0$ for $\alpha=1/2$, i.e., the conjectured inequality (4) is valid. With $x=\lambda_j/\lambda_i$ and $y=\lambda_\ell/\lambda_k$ we have

$$
\begin{aligned}
\frac{\tilde f(\alpha)}{\lambda_i\lambda_k} &= \left(x^{1-\alpha}+x^{\alpha}\right)\left(1+y-\tfrac{1}{2}\left(y^{1-\alpha}+y^{\alpha}\right)\right) \\
&\quad + \left(y^{1-\alpha}+y^{\alpha}\right)\left(1+x-\tfrac{1}{2}\left(x^{1-\alpha}+x^{\alpha}\right)\right) \\
&= \sqrt{xy}\,\Big(\left(x^{1/2-\alpha}+x^{\alpha-1/2}\right)\left(y^{1/2}+y^{-1/2}-\tfrac{1}{2}\left(y^{1/2-\alpha}+y^{\alpha-1/2}\right)\right) \\
&\quad + \left(y^{1/2-\alpha}+y^{\alpha-1/2}\right)\left(x^{1/2}+x^{-1/2}-\tfrac{1}{2}\left(x^{1/2-\alpha}+x^{\alpha-1/2}\right)\right)\Big).
\end{aligned}
$$

Therefore, (with $s = \frac{1}{2}\log x$ and $t = \frac{1}{2}\log y$) we need to consider the function

$$\begin{aligned} f(\alpha) &= \cosh\left((2\alpha-1)\,s\right)\Big(2\cosh\left(t\right) - \cosh\left((2\alpha-1)\,t\right)\Big) \\ &\qquad + \cosh\left((2\alpha-1)\,t\right)\Big(2\cosh\left(s\right) - \cosh\left((2\alpha-1)\,s\right)\Big). \end{aligned}$$

It is elementary to see

$$\begin{aligned} f'(\alpha) &= 4s\,\sinh((2\alpha-1)s)\Big(\cosh(t) - \cosh((2\alpha-1)t)\Big) \\ &\qquad +4t\,\sinh((2\alpha-1)t)\Big(\cosh(s) - \cosh((2\alpha-1)s)\Big), \end{aligned}$$

and we note

$$\begin{cases} \cosh(t) \geq \cosh((2\alpha-1)t) \ \text{ and } \ \cosh(s) \geq \cosh((2\alpha-1)s) & \text{for } \alpha \in (0,1), \\ s\,\sinh((2\alpha-1)s) \geq 0 \ \text{ and } \ t\,\sinh((2\alpha-1)t) \geq 0 & \text{for } \alpha \in [1/2,1), \end{cases}$$

showing that $f(\alpha)$ (and hence $\tilde{f}(\alpha)$ as well) is increasing on the interval $[1/2,1)$.

From Lemma 4 and the discussions so far we obtain the following main result in the article:

Theorem 5. *We assume that A, B are self-adjoint matrices and ρ is a density matrix. Then, the quantity*

$$\begin{aligned} F(\alpha) = &\Big(Var_\rho A \cdot Var_\rho B - \tfrac{1}{4}|Cov_\rho(A,B) + Cov_\rho(B,A)|^2\Big) \\ &- \Big(I(\rho,A;\alpha)\cdot I(\rho,B;\alpha) - \tfrac{1}{4}|Corr_\rho(A,B;\alpha) + Corr_\rho(B,A;\alpha)|^2\Big) \end{aligned}$$

defined for $\alpha \in (0,1)$ satisfies $F(1/2) \geq 0$, that is, the conjectured inequality (4) *holds true. Moreover, the function $F(\alpha)$ $(= F(1-\alpha))$ is monotone increasing on the right half interval* $[1/2,1)$.

The equality condition for the inequality in the theorem can be also determined (under the additional assumption that ρ is invertible) from the explicit expression given by Lemma 4.

Proposition 6. *We assume that the density matrix ρ is invertible and $\alpha \in (0,1)$. Then, the inequality in the preceding theorem reduces to the equality*

$$\begin{aligned} &Var_\rho A \cdot Var_\rho B - \tfrac{1}{4}|Cov_\rho(A,B) + Cov_\rho(B,A)|^2 \\ &\qquad = I(\rho,A;\alpha)\cdot I(\rho,B;\alpha) - \tfrac{1}{4}|Corr_\rho(A,B;\alpha) + Corr_\rho(B,A;\alpha)|^2 \end{aligned}$$

for $A, B \in M_n(\mathbf{C})_{sa}$ if and only if the two matrices $A - Tr(\rho A)1$ and $B - Tr(\rho B)1$ are proportional.

Proof. Thanks to the invariance (5) we may and do assume $\mathrm{Tr}(\rho A) = \mathrm{Tr}(\rho B) = 0$ in what follows. When A, B are proportional, the expression in Lemma 4 readily shows the equality in question. We conversely assume this equality in the rest (and will show that A, B are proportional), which forces that all the three parts of the expression in Lemma 4 must be zero.

Case 1. $(a_{11}, a_{22}, \cdots, a_{nn}) = 0$ (as a vector) and $(b_{11}, b_{22}, \cdots, b_{nn}) \neq 0$
The second part $\sum_k \lambda_k \left(\sum_{i<j} (\lambda_i + \lambda_j) |b_{kk}a_{ij} - a_{kk}b_{ij}|^2\right)$ of the expression in Lemma 4 must be zero. Thus, by choosing $k \in \{1, 2, \cdots, n\}$ with $b_{kk} \neq 0$ we must have $\sum_{i<j} (\lambda_i + \lambda_j) b_{kk}^2 |a_{ij}|^2 = 0$. Thus, we get $a_{ij} = 0$ for each $i < j$, showing $A = 0$.
Case 2. $(b_{11}, b_{22}, \cdots, b_{nn}) = 0$ and $(a_{11}, a_{22}, \cdots, a_{nn}) \neq 0$
By exchanging the roles of a's and b's in the previous case, we get $B = 0$.
Case 3. $(a_{11}, a_{22}, \cdots, a_{nn}) \neq 0$ and $(b_{11}, b_{22}, \cdots, b_{nn}) \neq 0$
The last part $(\sum_i \lambda_i a_{ii}^2)(\sum_i \lambda_i b_{ii}^2) - (\sum_i \lambda_i a_{ii} b_{ii})^2$ of the expression in Lemma 4 is zero so that we must have $a_{kk} = \gamma b_{kk}$ $(k = 1, 2, \cdots, n)$ for some $\gamma \in \mathbf{R} \setminus \{0\}$. By choosing $b_{kk} \neq 0$, from the second part we get

$$0 = \sum_{i<j} (\lambda_i + \lambda_j) |b_{kk}a_{ij} - a_{kk}b_{ij}|^2 = \sum_{i<j} (\lambda_i + \lambda_j) b_{kk}^2 |a_{ij} - \gamma b_{ij}|^2,$$

showing $a_{ij} = \gamma b_{ij}$ for each $i < j$ and hence $A = \gamma B$.
Case 4. $(a_{11}, a_{22}, \cdots, a_{nn}) = (b_{11}, b_{22}, \cdots, b_{nn}) = 0$
We can assume $A \neq 0$ and $B \neq 0$ since we have nothing to prove otherwise. Only the first part of the expression in Lemma 4 provides us meaning information in the present case. Namely, all the three inequalities in (12) have to be equalities (for each $i < j$ and $k < \ell$). Let us assume $a_{ij} \neq 0$ for some $i < j$. The equality condition for the first inequality in (12) (i.e., the arithmetic-geometric mean inequality) says

$$|a_{ij}b_{k\ell}| = |a_{k\ell}b_{ij}| \quad (\text{for each } k < \ell). \tag{13}$$

By choosing $k < \ell$ with $b_{k\ell} \neq 0$, we observe $b_{ij} \neq 0$ (and $a_{k\ell} \neq 0$). By the symmetric reasoning, $b_{ij} \neq 0$ also implies $a_{ij} \neq 0$. For convenience we set

$$\Lambda = \{(i,j);\ 1 \leq i < j \leq n \text{ and } a_{ij} \neq 0\} \Big(= \{(i,j);\ 1 \leq i < j \leq n \text{ and } b_{ij} \neq 0\}\Big).$$

The above equality condition (13) also yields that ratios $|a_{ij}|/|b_{ij}|$ are constant, i.e., $|a_{ij}| = \gamma |b_{ij}|$ (for each $(i,j) \in \Lambda$) for some $\gamma > 0$. It is elementary to see

$$|z_1 \overline{z_2}| = |\Re(z_1 \overline{z_2})| \ \text{ if and only if } \ \arg z_1 = \arg z_2 \text{ or } \arg z_1 = \arg z_2 + \pi.$$

Thus, the requirement that the second inequality in (12) must be the equality further forces $a_{ij} = \pm \gamma b_{ij}$ (for each $(i,j) \in \Lambda$). The parity here depends upon $(i,j) \in \Lambda$ at this stage, but γ and $-\gamma$ cannot actually mix because the third inequality in (12) is also the equality. Indeed, if both of $a_{ij} = \gamma b_{ij}$ and $a_{k\ell} = -\gamma b_{k\ell}$ occured for some $(i,j), (k,\ell) \in \Lambda$, then we would get two different values

$$\Re\left(a_{ij}\overline{b_{ij}}\right) \Re\left(a_{k\ell}\overline{b_{k\ell}}\right) = \left(\gamma |b_{ij}|^2\right)\left(-\gamma |b_{k\ell}|^2\right) = -\gamma^2 |b_{ij}|^2 |b_{k\ell}|^2,$$
$$\left|\Re\left(a_{ij}\overline{b_{ij}}\right)\right| \times \left|\Re\left(a_{k\ell}\overline{b_{k\ell}}\right)\right| = \gamma^2 |b_{ij}|^2 |b_{k\ell}|^2,$$

a contradiction. Thus, we conclude either $A = \gamma B$ or $A = -\gamma B$, and we are done. □

Each of the quantities appearing in the theorem and the proposition is known to make a perfect sense in the general von Neumann algebra setting, and the notion of skew information was indeed useful in [3] (where the homogeneity of the state space of a type III_1 factor was established). Our arguments presented in this section depend heavily upon manipulations of matrix components so that our method cannot be employed in the general setting. A suitable way to avoid this difficulty has to be found, and the author hopes to be able to come back to this problem.

4. Discussions

4.1.

In [5, Theorem 2 in p.1671] the inequality

$$\frac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2 + \frac{1}{16}|\mathrm{I}(\rho,A+B) - \mathrm{I}(\rho,A-B)|^2 \leq \mathrm{I}(\rho,A)\cdot\mathrm{I}(\rho,B) \tag{14}$$

was claimed. The real and imaginary parts of $\mathrm{Corr}_\rho(A,B)$ are given by

$$\begin{cases} \frac{1}{2}(\mathrm{Corr}_\rho(A,B) + \mathrm{Corr}_\rho(B,A)) = \frac{1}{4}(\mathrm{I}(\rho,A+B) - \mathrm{I}(\rho,A-B)), \\ \frac{1}{2i}(\mathrm{Corr}_\rho(A,B) - \mathrm{Corr}_\rho(B,A)) = \frac{1}{2i}\mathrm{Tr}(\rho[A,B]), \end{cases}$$

and (14) is actually equivalent to the following Cauchy-Schwarz tye estimate:

$$|\mathrm{Corr}_\rho(A,B)|^2 \leq \mathrm{Corr}_\rho(A,A)\cdot\mathrm{Corr}_\rho(B,B)\ (= \mathrm{I}(\rho\, A)\cdot\mathrm{I}(\rho,B)).$$

However, the presented argument in [5] has a gap (see the paragraph right before Remark 2), and (14) (i.e., the above Cauchy-Schwarz type estimate) is false even for 2×2 matrices.

To see this, let us set

$$\rho = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a \\ \overline{a} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b \\ \overline{b} & b_{22} \end{bmatrix}$$

with $a_{ii}, b_{ii} \in \mathbf{R}$ and $\lambda_i > 0$, $\lambda_1+\lambda_2 = 1$. Lemma 3 (or direct calculations) shows

$$\begin{aligned} &\mathrm{Tr}(\rho AB) = \lambda_1 a_{11}b_{11} + \lambda_1 a\overline{b} + \lambda_2\overline{a}b + \lambda_2 a_{22}b_{22}, \\ &\mathrm{Tr}(\rho^\alpha A\rho^{1-\alpha}B) = \lambda_1 a_{11}b_{11} + \lambda_1^\alpha\lambda_2^{1-\alpha}a\overline{b} + \lambda_1^{1-\alpha}\lambda_2^\alpha\,\overline{a}b + \lambda_2 a_{22}b_{22}. \end{aligned}$$

Since

$$\begin{aligned} \mathrm{Corr}_\rho(A,B) &= \mathrm{Tr}(\rho AB) - \mathrm{Tr}(\rho^{1/2}A\rho^{1/2}B) \\ &= \lambda_1 a\overline{b} + \lambda_2\overline{a}b - \sqrt{\lambda_1\lambda_2}\,a\overline{b} - \sqrt{\lambda_1\lambda_2}\,\overline{a}b \\ &= \left(\lambda_1+\lambda_2 - 2\sqrt{\lambda_1\lambda_2}\right)\Re\left(a\overline{b}\right) + i\,(\lambda_1-\lambda_2)\,\Im\left(a\overline{b}\right), \end{aligned} \tag{15}$$

we observe

$$
\begin{aligned}
|\mathrm{Corr}_\rho(A,B)|^2 &= \left(\lambda_1+\lambda_2-2\sqrt{\lambda_1\lambda_2}\right)^2 \left(\Re\left(a\bar{b}\right)\right)^2 + (\lambda_1-\lambda_2)^2\left(\Im\left(a\bar{b}\right)\right)^2 \\
&= \left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^4 \left(\Re\left(a\bar{b}\right)\right)^2 + \left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^2\left(\sqrt{\lambda_1}+\sqrt{\lambda_2}\right)^2\left(\Im\left(a\bar{b}\right)\right)^2, \\
\mathrm{Corr}_\rho(A,A)&\cdot\mathrm{Corr}_\rho(B,B) \ \ (=\mathrm{I}(\rho,A)\cdot\mathrm{I}(\rho,B)) \\
&= \left(\lambda_1+\lambda_2-2\sqrt{\lambda_1\lambda_2}\right)^2|a|^2|b|^2 = \left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^4|a|^2|b|^2.
\end{aligned}
$$

Consequently, if (14) were valid, then it would mean

$$\left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^2\left(\sqrt{\lambda_1}+\sqrt{\lambda_2}\right)^2\left(\Im\left(a\bar{b}\right)\right)^2 \le \left(\sqrt{\lambda_1}-\sqrt{\lambda_2}\right)^4\left(\Im\left(a\bar{b}\right)\right)^2.$$

But, when $\Im\left(a\bar{b}\right)\neq 0$ and $\lambda_1\neq\lambda_2$, it would be impossible for λ_1 close to $1/2$.

4.2.

One of the reasons why the inequality in Theorem 5 (with $\alpha=1/2$) was conjectured in [5] was that this together with (14) would imply Schrödinger's uncertainty principle (explained in Remark 2). However, (14) is not valid, and the best we can hope in this direction is the estimate

$$\left(\Re\,\mathrm{Corr}_\rho(A,B)\right)^2 \le \mathrm{I}(\rho,A)\cdot\mathrm{I}(\rho,B)$$

(see (2)). Note that the difference between the left side here and that of (14) is

$$\left(\Im\,\mathrm{Corr}_\rho(A,B)\right)^2 = \left(\Im\,\mathrm{Tr}(\rho AB)\right)^2 \ \left(=\tfrac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2\right).$$

Therefore, a natural alternative question is whether the estimate

$$F(1/2) \ge \left(\Im\,\mathrm{Corr}_\rho(A,B)\right)^2$$

(which is stronger than that in Theorem 5 by the above difference) is valid.

Actually this estimate is not valid. In fact, for

$$\rho=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix},\quad A=\begin{bmatrix}\lambda_2 & a\\ \bar{a} & -\lambda_1\end{bmatrix},\quad B=\begin{bmatrix}\lambda_2 & b\\ \bar{b} & -\lambda_1\end{bmatrix}$$

(satisfying $\mathrm{Tr}(\rho A)=\mathrm{Tr}(\rho B)=0$) Lemma 4 yields

$$
\begin{aligned}
F(\alpha) &= \lambda_1\lambda_2|a-b|^2 \\
&\quad +2\left(\lambda_1^{\alpha}\lambda_2^{1-\alpha}+\lambda_1^{1-\alpha}\lambda_2^{\alpha}\right)\left(1-\tfrac{1}{2}\left(\lambda_1^{\alpha}\lambda_2^{1-\alpha}+\lambda_1^{1-\alpha}\lambda_2^{\alpha}\right)\right)\left(\Im\left(a\bar{b}\right)\right)^2.
\end{aligned}
\tag{16}
$$

In particular, we have

$$\begin{aligned} F(1/2) &= \lambda_1\lambda_2|a-b|^2 + 4\left(\sqrt{\lambda_1\lambda_2}-\lambda_1\lambda_2\right)\left(\Im\left(a\overline{b}\right)\right)^2 \\ &= \lambda_1\lambda_2|a-b|^2 + \left(1-\left(1-2\sqrt{\lambda_1\lambda_2}\right)^2\right)\left(\Im\left(a\overline{b}\right)\right)^2 \end{aligned}$$

(which is non-negative due to $\sqrt{\lambda_1\lambda_2} \leq (\lambda_1+\lambda_2)/2 = 1/2$). In [5, p.1574] this value of $F(1/2)$ was computed to confirm the conjecture (i.e., Theorem 5 with $\alpha = 1/2$) for 2×2 matrices. On the other hand, (15) shows

$$\left(\Im\,\mathrm{Corr}_\rho(A,B)\right)^2 = (\lambda_1-\lambda_2)^2\left(\Im\left(a\overline{b}\right)\right)^2,$$

and it is plain to see

$$F(1/2) - \left(\Im\,\mathrm{Corr}_\rho(A,B)\right)^2 = \lambda_1\lambda_2|a-b|^2 + \left(4\sqrt{\lambda_1\lambda_2}-1\right)\left(\Im\left(a\overline{b}\right)\right)^2$$

(thanks to $\lambda_1+\lambda_2 = 1$). This quantity cannot be non-negative for λ_1 sufficiently small (as long as $\Im\left(a\overline{b}\right) \neq 0$).

The situation cannot be improved even if $F(\alpha)$ ($\geq F(1/2)$), $\alpha \in (0,1)$, is used. In the limit case ($\alpha = 0$ or 1) we have $\mathrm{I}(\rho,A,\alpha) = \mathrm{I}(\rho,B,\alpha) = \Re\,\mathrm{Corr}_\rho(A,B;\alpha) = 0$ so that

$$F(0) = F(1) = \mathrm{Var}_\rho A\cdot\mathrm{Var}_\rho B - \tfrac{1}{4}|\mathrm{Cov}_\rho(A,B)+\mathrm{Cov}_\rho(B,A)|^2.$$

We note that $F(1) \geq \left(\Im\,\mathrm{Corr}_\rho(A,B;1)\right)^2 = 0$ is obviously true while

$$F(0) \geq \left(\Im\,\mathrm{Corr}_\rho(A,B;0)\right)^2 = \tfrac{1}{4}|\mathrm{Tr}(\rho[A,B])|^2$$

is nothing but Schrödinger's uncertainty principle (explained in Remark 2).

However,

$$F(\alpha) - \left(\Im\,\mathrm{Corr}_\rho(A,B;\alpha)\right)^2 \ngeq 0$$

can be easily checked for each $\alpha \in (0,1)$. In fact, similar computations as (15) show

$$\Im\,\mathrm{Corr}_\rho(A,B;\alpha) = \left(\lambda_1 - \lambda_1^\alpha\lambda_2^{1-\alpha} + \lambda_1^{1-\alpha}\lambda_2^\alpha - \lambda_2\right)\Im\left(a\overline{b}\right)$$

so that with (16) we get

$$\begin{aligned} &F(\alpha) - \left(\Im\,\mathrm{Corr}_\rho(A,B;\alpha)\right)^2 \\ &\quad = \lambda_1\lambda_2|a-b|^2 + \Big(2\left(\lambda_1^\alpha\lambda_2^{1-\alpha}+\lambda_1^{1-\alpha}\lambda_2^\alpha\right)\left(1-\tfrac{1}{2}\left(\lambda_1^\alpha\lambda_2^{1-\alpha}+\lambda_1^{1-\alpha}\lambda_2^\alpha\right)\right) \\ &\qquad\qquad - \left(\lambda_1-\lambda_1^\alpha\lambda_2^{1-\alpha}+\lambda_1^{1-\alpha}\lambda_2^\alpha-\lambda_2\right)^2\Big)\left(\Im\left(a\overline{b}\right)\right)^2. \end{aligned}$$

As $\lambda_1 \searrow 0$ (and hence $\lambda_2 = 1-\lambda_1 \nearrow 1$), this quantity tends to $-\left(\Im\left(a\overline{b}\right)\right)^2$, showing that $F(\alpha) - \left(\Im\,\mathrm{Corr}_\rho(A,B;\alpha)\right)^2$ cannot be always non-negative.

References

1. H. Araki and M. M. Yanase, *Measurement of quantum mechanical operators*, Phys. Rev., **200** (1960), 622-626.
2. J.-C. Bourin, *Some inequalities for norms on matrices and operators*, Linear Algebra Appl., **292** (1999), 139-154.
3. A. Connes and E. Størmer, *Homogeneity of the state space of factors of type* III_1, J. Funct. Anal., **28** (1978), 187-196.
4. E. Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Adv. Math. **11** (1973), 267-288.
5. S. Luo and Z. Zhang, *An information charactrization of Schrödinger's uncertainty principle*, J. Stat. Phys., **114** (2004), 1557-1576.
6. A. Wehrl, *General properties of entropy*, Rev. Modern Phys., **50** (1978), 221-260.
7. E. P. Wigner and M. M. Yanase, *Information contents of distributions*, Proc. Nat. Acad. Sci. USA, **49** (1963), 910-918.

Faculty of Mathematics, Kyushu University, Higashi-ku, Fukuoka 812-8581, Japan

List of MHF Preprint Series, Kyushu University

21st Century COE Program
Development of Dynamic Mathematics with High Functionality

MHF2004-1 Koji YONEMOTO & Takashi YANAGAWA
Estimating the Lyapunov exponent from chaotic time series with dynamic noise

MHF2004-2 Rui YAMAGUCHI, Eiko TSUCHIYA & Tomoyuki HIGUCHI
State space modeling approach to decompose daily sales of a restaurant into time-dependent multi-factors

MHF2004-3 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Cubic pencils and Painlevé Hamiltonians

MHF2004-4 Atsushi KAWAGUCHI, Koji YONEMOTO & Takashi YANAGAWA
Estimating the correlation dimension from a chaotic system with dynamic noise

MHF2004-5 Atsushi KAWAGUCHI, Kentarou KITAMURA, Koji YONEMOTO, Takashi YANAGAWA & Kiyofumi YUMOTO
Detection of auroral breakups using the correlation dimension

MHF2004-6 Ryo IKOTA, Masayasu MIMURA & Tatsuyuki NAKAKI
A methodology for numerical simulations to a singular limit

MHF2004-7 Ryo IKOTA & Eiji YANAGIDA
Stability of stationary interfaces of binary-tree type

MHF2004-8 Yuko ARAKI, Sadanori KONISHI & Seiya IMOTO
Functional discriminant analysis for gene expression data via radial basis expansion

MHF2004-9 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Hypergeometric solutions to the q - Painlevé equations

MHF2004-10 Raimundas VIDŪNAS
Expressions for values of the gamma function

MHF2004-11 Raimundas VIDŪNAS
Transformations of Gauss hypergeometric functions

MHF2004-12 Koji NAKAGAWA & Masakazu SUZUKI
Mathematical knowledge browser

MHF2004-13 Ken-ichi MARUNO, Wen-Xiu MA & Masayuki OIKAWA
Generalized Casorati determinant and Positon-Negaton-Type solutions of the Toda lattice equation

MHF2004-14 Nalini JOSHI, Kenji KAJIWARA & Marta MAZZOCCO
Generating function associated with the determinant formula for the solutions of the Painlevé II equation

MHF2004-15 Kouji HASHIMOTO, Ryohei ABE, Mitsuhiro T. NAKAO & Yoshitaka WATANABE
Numerical verification methods of solutions for nonlinear singularly perturbed problem

MHF2004-16 Ken-ichi MARUNO & Gino BIONDINI
Resonance and web structure in discrete soliton systems: the two-dimensional Toda lattice and its fully discrete and ultra-discrete versions

MHF2004-17 Ryuei NISHII & Shinto EGUCHI
Supervised image classification in Markov random field models with Jeffreys divergence

MHF2004-18 Kouji HASHIMOTO, Kenta KOBAYASHI & Mitsuhiro T. NAKAO
Numerical verification methods of solutions for the free boundary problem

MHF2004-19 Hiroki MASUDA
Ergodicity and exponential β-mixing bounds for a strong solution of Lévy-driven stochastic differential equations

MHF2004-20 Setsuo TANIGUCHI
The Brownian sheet and the reflectionless potentials

MHF2004-21 Ryuei NISHII & Shinto EGUCHI
Supervised image classification based on AdaBoost with contextual weak classifiers

MHF2004-22 Hideki KOSAKI
On intersections of domains of unbounded positive operators

MHF2004-23 Masahisa TABATA & Shoichi FUJIMA
Robustness of a characteristic finite element scheme of second order in time increment

MHF2004-24 Ken-ichi MARUNO, Adrian ANKIEWICZ & Nail AKHMEDIEV
Dissipative solitons of the discrete complex cubic-quintic Ginzburg-Landau equation

MHF2004-25 Raimundas VIDŪNAS
Degenerate Gauss hypergeometric functions

MHF2004-26 Ryo IKOTA
The boundedness of propagation speeds of disturbances for reaction-diffusion systems

MHF2004-27 Ryusuke KON
Convex dominates concave: an exclusion principle in discrete-time Kolmogorov systems

MHF2004-28 Ryusuke KON
Multiple attractors in host-parasitoid interactions: coexistence and extinction

MHF2004-29 Kentaro IHARA, Masanobu KANEKO & Don ZAGIER
Derivation and double shuffle relations for multiple zeta values

MHF2004-30 Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Generalized partitioned quantum cellular automata and quantization of classical CA

MHF2005-1 Hideki KOSAKI
Matrix trace inequalities related to uncertainty principle