entropy

MDPI

Article

On a Matrix Inequality Related to the Distillability Problem

Yi Shen [1,†] **and Lin Chen** [1,2,†,*]

[1] School of Mathematics and Systems Science, Beihang University, Beijing 100191, China; yishen@buaa.edu.cn
[2] International Research Institute for Multidisciplinary Science, Beihang University, Beijing 100191, China
* Correspondence: linchen@buaa.edu.cn; Tel.: +86-188-0010-9142
† These authors contributed equally to this work.

Received: 25 May 2018; Accepted: 4 August 2018; Published: 8 August 2018

Abstract: We investigate the distillability problem in quantum information in $\mathbb{C}^d \otimes \mathbb{C}^d$. One case of the problem has been reduced to proving a matrix inequality when $d = 4$. We investigate the inequality for three families of non-normal matrices. We prove the inequality for the first two families with $d = 4$ and for the third family with $d \geq 5$. We also present a sufficient condition for the fulfillment of the inequality with $d = 4$.

Keywords: matrix inequality; singular value; eigenvalue; Kronecker product; quantum information; distillability problem

MSC: 15A18; 15A21; 15A45; 15A69

1. Introduction

Extracting pure entanglement from mixed states is a basic task in quantum information. This task is called entanglement distillation formally. We first briefly introduce the physical motivation of the distillability problem. In quantum physics, a quantum state is mathematically described by a positive semidefinite matrix. The state is pure when it has rank one, otherwise the state is mixed. Pure entangled states play an essential role in most quantum-information tasks such as quantum computation. Nevertheless, there is no pure state in nature due to the inevitable decoherence between the state and environment. Therefore, asymptotically converting initially bipartite entangled mixed states into bipartite pure entangled states under local operations and classical communications (LOCC) is a key step in quantum information processing. The distillability problem [1,2] asks whether the above-mentioned conversion succeeds for any mixed states. It has been a main open problem in quantum information [1] for a long time, since the distillability problem lies at the heart of entanglement theory [3–5] and is related to the separability problem extensively studied by the quantum information community recently [6–8]. There have been some attempts at the problem in the past years [1,2,9–16].

In the following, we will show a mathematical description of the distillability problem. Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ be the bipartite Hilbert space with $\operatorname{Dim} \mathcal{H}_A = M$ and $\operatorname{Dim} \mathcal{H}_B = N$. Recall that a quantum state is a positive semidefinite matrix. For the sake of normalization in quantum physics, it is required that every quantum state be a unit vector. However, the requirement does not play an essential role in the distillability problem and often causes inconvenience in the mathematical expressions and discussion. In this paper, unless stated otherwise, the states will not be normalized.

We shall work with the quantum state ρ on $\mathcal{H}$. Such a ρ is called a bipartite state of system A and B. We have $\rho = \sum_{i,j=1}^{M} E_{ij} \otimes \rho_{ij}$, where E_{ij} is an $M \times M$ matrix, the elements of which are all zero, except that the (i,j) entry is one. The partial transpose of ρ with respect to the system A is defined as $\rho^\Gamma := \sum_{i,j=1}^{M} E_{ji} \otimes \rho_{ij}$. We say that ρ is the positive partial transpose (PPT) if $\rho^\Gamma \geq 0$. Otherwise, ρ is the negative partial transpose (NPT), i.e., ρ^Γ has at least one negative eigenvalue. The NPT states are entangled states due to the Peres–Horodecki criterion in quantum information [17,18]. We say that a quantum state is pure when it has rank one.

Since the distillability problem requires many copies of the same states, we further need the concept of a composite system. Let $\rho_{A_iB_i}$ be an $M_i \times N_i$ state of rank r_i acting on the Hilbert space $\mathcal{H}_{A_i} \otimes \mathcal{H}_{B_i}$, $i = 1, 2$, with $\mathrm{Dim}\,\mathcal{H}_{A_i} = M_i$ and $\mathrm{Dim}\,\mathcal{H}_{B_i} = N_i$. Suppose ρ of systems A_1, A_2 and B_1, B_2 is a state on the Hilbert space $\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2}$, such that the partial trace $\mathrm{Tr}_{A_1B_1}\rho = \rho_{A_2B_2}$ and $\mathrm{Tr}_{A_2B_2}\rho = \rho_{A_1B_1}$. By switching the two middle factors, we can regard ρ as a composite bipartite state on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ where $\mathcal{H}_A = \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ and $\mathcal{H}_B = \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}$. We write $\rho = \rho_{A_1A_2:B_1B_2}$. One can verify that ρ is an $M_1M_2 \times N_1N_2$ state of rank at most r_1r_2. For example, the tensor product $\rho = \rho_{A_1B_1} \otimes \rho_{A_2B_2}$ is an $M_1M_2 \times N_1N_2$ state of rank r_1r_2. The above definition can be generalized to the tensor product of N states $\rho_{A_iB_i}, i = 1, \ldots, N$. They form a bipartite state on the Hilbert space $\mathcal{H}_{A_1,\cdots,A_N} \otimes \mathcal{H}_{B_1,\cdots,B_N}$. It is written as $\mathcal{H}^{\otimes n}$ with $\mathcal{H}_{A_i} \otimes \mathcal{H}_{B_i} = \mathcal{H}$.

Third, we shall refer to the notations $|\psi\rangle$ and $\langle\psi|$ in quantum physics respectively as a column vector and its conjugate transpose in linear algebra. In quantum information, the well-known Werner state in $\mathbb{C}^d \otimes \mathbb{C}^d$ is defined as $\frac{I+\alpha\sum_{i,j=1}^{d} E_{ij}\otimes E_{ji}}{d^2+\alpha d}$, where the real number $\alpha \in [-1, 1]$ [19]. We introduce the definition of distillable states as follows [1].

Definition 1. *A bipartite state ρ is n-distillable under local operations and classical communications if there exists a Schmidt-rank-two bipartite state $|\psi\rangle \in \mathcal{H}^{\otimes n}$ such that $\langle\psi|(\rho^{\otimes n})^{\Gamma}|\psi\rangle < 0$. Otherwise, we say that ρ is n-undistillable. We say that ρ is distillable if it is n-distillable for some $n \geq 1$.*

The definition shows that PPT states are not distillable. It has been shown [1] that all NPT bipartite states can be locally converted into NPT Werner states. Using Definition 1, one can show that the distillability of NPT Werner states is equivalent to that with $\alpha = -1/2$. Therefore, the distillability problem indeed asks whether Werner states with $\alpha = -1/2$ are distillable. In this paper, we investigate one case of the distillability problem, i.e., the two-undistillability of Werner states in $\mathbb{C}^4 \otimes \mathbb{C}^4$ with $\alpha = -1/2$. It is known that this case is equivalent to the following mathematical problem [9].

Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. Denote $A^\dagger$ and $B^\dagger$ as the conjugate transpose of A and B, respectively. The Kronecker sum of A and B denoted by $A \oplus_K B$ is defined as $A \otimes I_m + I_n \otimes B$; see more facts in [20] (Section 7.2). The work in [9] has presented the following conjecture on the Kronecker sum when A and B have the same size.

Conjecture 1. *Let $A, B, I \in \mathbb{C}^{d\times d}$, $d \geq 4$ and the matrix:*

$$X = A \otimes I + I \otimes B \tag{1}$$

where:

$$\mathrm{Tr}A = \mathrm{Tr}B = 0, \quad \mathrm{Tr}A^\dagger A + \mathrm{Tr}B^\dagger B = \frac{1}{d}. \tag{2}$$

Define set $\mathcal{X}_d$ whose elements are given by Equations (1) and (2). Let $\sigma_1, \cdots, \sigma_{d^2}$ be the singular values of $X \in \mathcal{X}_d$ in the descending order. Then:

$$\sup_{X\in\mathcal{X}_d} \ (\sigma_1^2 + \sigma_2^2) \leq \frac{1}{2}. \tag{3}$$

The condition $d \geq 4$ is essential as one can show that Conjecture 1 fails for $d = 3$ from Lemma A1. It has been shown [9] that Conjecture 1 for $d = 4$ is the mathematical description of a two-copy distillability problem. They are equivalent through a given state-operator isomorphism ([9], Equation (43)). It is also known that Conjecture 1 holds for all normal matrices with $d \geq 4$ [9]. Evidently, the matrix X in (1) is normal if and only if A and B in (1) are both normal. The remaining work on

Conjecture 1 is to prove it when X is non-normal. It turns out to be a hard problem and there is no progress so far, as far as we know.

In this paper, we investigate Conjecture 1 in terms of three families of non-normal matrices X. They are respectively constructed in Definitions 2–4. We prove Conjecture 1 for the first family $\mathcal{P}_1$ of non-normal X in Theorem 1, based on Propositions 1 and 2, and for the second family $\mathcal{P}_2$ of non-normal X in Theorem 2. For the third family $\mathcal{P}_3$ of non-normal X, we prove Conjecture 1 with $d \geq 5$ in Theorem 3. We also present a sufficient condition for Conjecture 1 with $d = 4$ in Lemma 10. It is an idea different from the above three families of matrices to deal with Conjecture 1. Our results carry out the first step to prove Conjecture 1 for non-normal matrices and, thus, the distillability problem in quantum information.

The rest of this paper is organized as follows. We introduce some notations and preliminary results in linear algebra in Section 2. We investigate Conjecture 1 for three families of non-normal matrices in Sections 3.1–3.3, respectively. We also show a sufficient condition for Conjecture 1 with $d = 4$ in Section 3.4. Finally, we provide some discussion on the distillability problem and conclude our work in Section 4.

2. Preliminaries

We first present some mathematical notations. We shall denote A^*, A^T and $A^\dagger$ as the conjugate, transpose and conjugate transpose of matrix A, respectively. We refer to $\mathbb{C}^{n\times n}$ as the set of $n \times n$ matrices with entries in the complex field and $\mathbb{H}^{n\times n}$ as the set of $n \times n$ Hermitian matrices. Let I_n be the identity matrix in $\mathbb{C}^{n\times n}$. We shall omit the subscript of the identity matrix when it is clear in the paper. Let $\sigma(A)$ be the spectrum of matrix A, $\lambda_i(A)$ be an eigenvalue of A and $A_{(i,j)}$ be the (i,j) entry of A. We shall say that X is unitarily similar to Y, i.e., $X \sim Y$ when there exists a unitary matrix W such that $X = WYW^\dagger$. In particular, they are locally similar when there exist unitary matrices $U, V \in \mathbb{C}^{d\times d}$ such that $W = U \otimes V$. We denote this as $X \sim_L Y$.

We then post some lemmas, which will be used in the following section. The following lemma is essential to simplify Conjecture 1.

Lemma 1. *The following four statements are equivalent.*

(i) Conjecture 1 holds.

(ii) Conjecture 1 holds when X is replaced by X^T, X^* or $X^\dagger$.

(iii) Conjecture 1 holds when X is replaced by any matrix unitarily similar to X.

(iv) Conjecture 1 holds when X is replaced by $I \otimes A + B \otimes I$.

Proof of Lemma 1. The equivalences between (i) and (ii) and between (i) and (iii) follow from straightforward computation. We next prove that (i) is equivalent to (iv). It follows from a known fact of the Kronecker sum, i.e., for $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$, there exists a permutation $P \in \mathbb{C}^{mn\times mn}$ such that $P(A \oplus_K B)P^{-1} = B \oplus_K A$. Hence, $A \oplus_K B$ and $B \oplus_K A$ are unitarily similar when $m = n$, and thus, (i) is equivalent to (iv) by (iii). □

Using statement (iii), we can generally assume that A and B in Conjecture 1 are both upper-triangular. In particular, we can assume that they are diagonal if and only if they are normal.

Then, let us recall the result on normal matrices in [9]. We will briefly show its proof and some related new findings in Appendix A.

Lemma 2. *Let $\mathcal{N}_d$ be a subset of normal operators X in* (1) *satisfying the constraints* (2)*. Then, Conjecture 1 holds for $X \in \mathcal{N}_d$ where $d \geq 4$.*

The following lemmas are useful to prove the result of the third non-normal family.

Lemma 3 ([20], Fact 4.10.16.). *(Gershgorin circle theorem) Let $A \in \mathbb{C}^{n\times n}$. Then,*

$$\sigma(A) \subset G(A) = \bigcup_{i=1}^{n} \{s \in \mathbb{C} : \left|s - A_{(i,i)}\right| \leq \sum_{\substack{j=1\\j\neq i}}^{n} \left|A_{(i,j)}\right| \}, \tag{4}$$

and a corollary is:

$$\sigma(A) \subset \bigcup_{i=1}^{n} \{s \in \mathbb{C} : |s| \leq \sum_{j=1}^{n} \left|A_{(i,j)}\right| \}. \tag{5}$$

Remark 1. *Let $R_i = \sum_{\substack{j=1\\j\neq i}}^{n} \left|A_{(i,j)}\right|$ and $D(A_{(i,i)}, R_i)$ be the closed disc centered at $A_{(i,i)}$ with radius R_i. Every eigenvalue of A lies within at least one of the Gershgorin discs $D(A_{(i,i)}, R_i)$.*

Lemma 4 ([20], Fact 4.10.21.). *(Brauer theorem) Let $A \in \mathbb{C}^{n\times n}$. Then,*

$$\sigma(A) \subset \bigcup_{\substack{i,j=1\\i\neq j}}^{n} \{s \in \mathbb{C} : \left|s - A_{(i,i)}\right| \left|s - A_{(j,j)}\right| \leq \sum_{\substack{k=1\\k\neq i}}^{n} \left|A_{(i,k)}\right| \sum_{\substack{k=1\\k\neq j}}^{n} \left|A_{(j,k)}\right| \}. \tag{6}$$

Remark 2. *The eigenvalues of A lie in the union of $n(n-1)/2$ ovals of Cassini, which is contained in the union of Gershgorin discs (4). Hence, the Brauer theorem is stronger than the Gershgorin circle theorem.*

Lemma 5 ([21], Corollary 4.3.15.). *Let $A, B \in \mathbb{H}^{n\times n}$. Let the eigenvalues of A, B be in the increasing order, that is $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then, for all $i = 1, \cdots n$, we have:*

$$\lambda_i(A) + \lambda_1(B) \leq \lambda_i(A+B) \leq \lambda_i(A) + \lambda_n(B). \tag{7}$$

3. Results

In this section, we will present our main results on non-normal matrices. We prove Conjecture 1 with $d = 4$ for the first two families we construct and prove Conjecture 1 with $d \geq 5$ for the third family we construct. We also show a sufficient condition for Conjecture 1 with $d = 4$.

3.1. Conjecture 1 with Non-Normal Matrices X: Family 1

In this subsection, we prove Conjecture 1 with a family of non-normal matrices X in Definition 2. We will formulate our main result in Theorem 1, followed by two preliminary facts, i.e., Proposition 1 and 2.

Definition 2. *Let $\mathcal{P}_1$ be the subset of matrices X with $d = 4$, such that A is normal or $A \sim \begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix}$, and B is normal or $B \sim \begin{bmatrix} 0 & b_1 & 0 & 0 \\ b_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_3 \\ 0 & 0 & b_4 & 0 \end{bmatrix}$. A and B should satisfy the constraints in (2).*

One can show $\mathcal{N}_4 \subset \mathcal{P}_1$, and $\begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix} \sim \begin{bmatrix} \tilde{a_{11}} & \tilde{a_{12}} & 0 & 0 \\ 0 & \tilde{a_{22}} & 0 & 0 \\ 0 & 0 & \tilde{a_{33}} & \tilde{a_{34}} \\ 0 & 0 & 0 & \tilde{a_{44}} \end{bmatrix}$ with $\tilde{a_{11}} + \tilde{a_{22}} = \tilde{a_{33}} + \tilde{a_{44}} = 0$. We present the main result of this section as follows.

Theorem 1. *The inequality* (3) *holds when* $X \in \mathcal{P}_1$.

Proof of Theorem 1. If $A \sim \begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix}$ and $B \sim \begin{bmatrix} 0 & b_1 & 0 & 0 \\ b_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_3 \\ 0 & 0 & b_4 & 0 \end{bmatrix}$, then the assertion follows from Proposition 1. If one of A and B is normal, then we may assume that it is diagonal by Lemma 1 (iii). Therefore, the assertion follows from Proposition 2 and the switch of A, B (if any), as well, because of Lemma 1 (iv). If A and B are both normal, then the assertion follows from Lemma 2. This completes the proof. □

Proposition 1. *The inequality* (3) *holds when* $X \in \mathcal{P}_1$, $A \sim \begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix}$ *and* $B \sim \begin{bmatrix} 0 & b_1 & 0 & 0 \\ b_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_3 \\ 0 & 0 & b_4 & 0 \end{bmatrix}$.

Proof of Proposition 1. For Lemma 1 (iii), we assume $A = \begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & b_1 & 0 & 0 \\ b_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_3 \\ 0 & 0 & b_4 & 0 \end{bmatrix}$. By computing, one can show that $X^\dagger X = Y_1 \oplus Y_2$. We can partition Y_1 as $\left[\begin{array}{c|c} Z_1 & Z_2 \\ \hline Z_2^\dagger & Z_4 \end{array}\right]$, where $Z_1 = \mathrm{diag}(|a_2|^2 + |b_2|^2, |a_2|^2 + |b_1|^2, |a_2|^2 + |b_4|^2, |a_2|^2 + |b_3|^2)$, $Z_2 = \begin{bmatrix} 0 & a_1 b_2^* + b_1 a_2^* & 0 & 0 \\ a_1 b_1^* + b_2 a_2^* & 0 & 0 & 0 \\ 0 & 0 & 0 & a_1 b_4^* + b_3 a_2^* \\ 0 & 0 & a_1 b_3^* + b_4 a_2^* & 0 \end{bmatrix}$, and $Z_4 = \mathrm{diag}(|a_1|^2 + |b_2|^2, |a_1|^2 + |b_1|^2, |a_1|^2 + |b_4|^2, |a_1|^2 + |b_3|^2)$. One can formulate the characteristic polynomial of Y_1 as follows.

$$\det(\lambda I - Y_1) = f_1(\lambda) \cdot f_2(\lambda) \cdot f_3(\lambda) \cdot f_4(\lambda), \tag{8}$$

where:

$$\begin{aligned} f_1(\lambda) &= \left(\lambda - (|a_1|^2 + |b_2|^2)\right)\left(\lambda - (|a_2|^2 + |b_1|^2)\right) - |a_1 b_1^* + a_2^* b_2|^2, \\ f_2(\lambda) &= \left(\lambda - (|a_1|^2 + |b_1|^2)\right)\left(\lambda - (|a_2|^2 + |b_2|^2)\right) - |a_1 b_2^* + a_2^* b_1|^2, \\ f_3(\lambda) &= \left(\lambda - (|a_1|^2 + |b_4|^2)\right)\left(\lambda - (|a_2|^2 + |b_3|^2)\right) - |a_1 b_3^* + a_2^* b_4|^2, \\ f_4(\lambda) &= \left(\lambda - (|a_1|^2 + |b_3|^2)\right)\left(\lambda - (|a_2|^2 + |b_4|^2)\right) - |a_1 b_4^* + a_2^* b_3|^2. \end{aligned} \tag{9}$$

We first claim the larger root of $f_i(\lambda) = 0, \forall i = 1,2,3,4$, is not greater than $\frac{1}{4}$. Take $f_1(\lambda)$ as an example. From the Vieta theorem, we have the sum of two roots of $f_1(\lambda) = 0$ is $|a_1|^2 + |a_2|^2 + |b_1|^2 + |b_2|^2$. Since $X^\dagger X = Y_1 \oplus Y_2$ is a semipositive definite matrix, all the eigenvalues of Y_1 and Y_2 are non-negative. This implies that all the roots of $f_i(\lambda) = 0, i = 1,2,3,4$ are non-negative. So we have the larger root of $f_1(\lambda) = 0$ is not greater than the sum of two roots, i.e., $|a_1|^2 + |a_2|^2 + |b_1|^2 + |b_2|^2$.

Recall that $\sum_{i=1}^{4}(|a_i|^2+|b_i|^2)=\frac{1}{4}$. Therefore, we conclude that the larger root of $f_1(\lambda)=0$ is not greater than $\frac{1}{4}$. One can draw the same conclusion for $f_2(\lambda), f_3(\lambda), f_4(\lambda)$. Then, our claim holds. This implies the largest eigenvalue of Y_1 is no greater than $\frac{1}{4}$. One can show that Y_2 can be evolved from Y_1 by replacing a_1 with a_3 and replacing a_2 with a_4 in Y_1. Hence, we obtain the expression of $\det(\lambda I - Y_2)$ by replacing a_1 with a_3 and replacing a_2 with a_4 in (9). In the same way, we conclude that the largest eigenvalue of Y_2 is no greater than $\frac{1}{4}$. Since $X^\dagger X = Y_1 \oplus Y_2$, the sum of the largest two eigenvalues of $X^\dagger X$ is at most $\frac{1}{2}$. This completes the proof. □

We proceed with the proof of the upcoming Proposition 2. For this purpose, we need two preliminary results. The first result is known as one of the basic inequalities.

Lemma 6. *If $a,b,x,y \in \mathbb{R}$ then $ab(x+y)^2 \le (a+b)(ax^2+by^2)$.*

Lemma 7. *Suppose a_1,a_2,b_1,b_2 are nonnegative real numbers and $a_1^2+a_2^2+b_1^2+b_2^2=1/4$. Then:*

$$\sqrt{(a_1^2-a_2^2)^2+4b_1^2(a_1+a_2)^2}+\sqrt{(a_1^2-a_2^2)^2+4b_2^2(a_1+a_2)^2} \le 1/2. \tag{10}$$

Proof of Lemma 7. Using the basic inequality $x+y \le \sqrt{2(x^2+y^2)}$ for any real x,y, we obtain that the lhs of (10) is upper bounded by:

$$\begin{aligned}
& \sqrt{2\Big(2(a_1^2-a_2^2)^2+4(b_1^2+b_2^2)(a_1+a_2)^2\Big)} \\
= & \sqrt{2(a_1+a_2)^2\Big(1-2(a_1+a_2)^2\Big)} \\
\le & \ 1/2.
\end{aligned} \tag{11}$$

The equality follows from the equation $a_1^2+a_2^2+b_1^2+b_2^2=1/4$. This completes the proof. □

Proposition 2. *The inequality (3) holds when $X \in \mathcal{P}_1$, for which $A \sim \begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix}$ and B is normal, i.e., $B \sim \mathrm{diag}(b_1,b_2,b_3,b_4)$.*

Proof of Proposition 2. For Lemma 1 (iii), we assume $A = \begin{bmatrix} 0 & a_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & a_4 & 0 \end{bmatrix}$ and $B = \mathrm{diag}(b_1,b_2,b_3,b_4)$. Since A and B satisfy (2), we have:

$$\sum_{i=1}^{4} b_i = 0, \tag{12}$$

$$\sum_{j=1}^{4}(|a_j|^2+|b_j|^2) = 1/4. \tag{13}$$

By computing, one can show that $X^\dagger X = H_1 \oplus H_2$. We can partition H_1 and H_2 as follows.

$$H_1 = \left[\begin{array}{c|c} D_{11} & D_{12} \\ \hline D_{12}^\dagger & D_{14} \end{array}\right], H_2 = \left[\begin{array}{c|c} D_{21} & D_{22} \\ \hline D_{22}^\dagger & D_{24} \end{array}\right], \tag{14}$$

where:

$$
\begin{aligned}
D_{11} &= \mathrm{diag}(|a_2|^2+|b_1|^2, |a_2|^2+|b_2|^2, |a_2|^2+|b_3|^2, |a_2|^2+|b_4|^2),\\
D_{12} &= \mathrm{diag}(a_1 b_1^* + a_2^* b_1, a_1 b_2^* + a_2^* b_2, a_1 b_3^* + a_2^* b_3, a_1 b_4^* + a_2^* b_4),\\
D_{14} &= \mathrm{diag}(|a_1|^2+|b_1|^2, |a_1|^2+|b_2|^2, |a_1|^2+|b_3|^2, |a_1|^2+|b_4|^2),\\
D_{21} &= \mathrm{diag}(|a_4|^2+|b_1|^2, |a_4|^2+|b_2|^2, |a_4|^2+|b_3|^2, |a_4|^2+|b_4|^2),\\
D_{22} &= \mathrm{diag}(a_3 b_1^* + a_4^* b_1, a_3 b_2^* + a_4^* b_2, a_3 b_3^* + a_4^* b_3, a_3 b_4^* + a_4^* b_4),\\
D_{24} &= \mathrm{diag}(|a_3|^2+|b_1|^2, |a_3|^2+|b_2|^2, |a_3|^2+|b_3|^2, |a_3|^2+|b_4|^2).
\end{aligned}
\tag{15}
$$

There exists a permutation $P = \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&1 \end{bmatrix} \oplus \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&1 \end{bmatrix}$ such that $P^\dagger(X^\dagger X)P = \oplus_{j=1}^4 (Y_j \oplus Z_j)$, where Y_j and Z_j are order-two submatrices such that:

$$
Y_j = \begin{bmatrix} |a_2|^2 + |b_j|^2 & b_j^* a_1 + a_2^* b_j \\ b_j a_1^* + a_2 b_j^* & |a_1|^2 + |b_j|^2 \end{bmatrix}, \tag{16}
$$

and:

$$
Z_j = \begin{bmatrix} |a_4|^2 + |b_j|^2 & b_j^* a_3 + a_4^* b_j \\ b_j a_3^* + a_4 b_j^* & |a_3|^2 + |b_j|^2 \end{bmatrix}. \tag{17}
$$

Let λ and μ be two arbitrary eigenvalues of $X^\dagger X$. Then, proving the inequality (3) is equivalent to proving $\lambda + \mu \le 1/2$. There are five cases for λ and μ.

Case 1. λ and μ are the eigenvalues of the same Y_j or Z_j. Equations (13) and (16) imply that $\lambda + \mu = \mathrm{Tr} Y_j \le 1/2$. Equations (13) and (17) imply that $\lambda + \mu = \mathrm{Tr} Z_j \le 1/2$.

Case 2. λ and μ are the eigenvalues of different Y_j's. Without loss of generality, we can assume that λ is the maximum eigenvalue of Y_1 and μ is the maximum eigenvalue of Y_2. By computation, one can obtain:

$$
\lambda = \frac{1}{2}\left(|a_1|^2 + |a_2|^2 + 2|b_1|^2 + \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4\left|b_1 a_2^* + a_1 b_1^*\right|^2} \right), \tag{18}
$$

$$
\mu = \frac{1}{2}\left(|a_1|^2 + |a_2|^2 + 2|b_2|^2 + \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4\left|b_2 a_2^* + a_1 b_2^*\right|^2} \right). \tag{19}
$$

Therefore, $\lambda + \mu$ is upper bounded by the sum of the rhs of (18) and (19), in which any a_i and b_j are replaced by $|a_i|$ and $|b_j|$, respectively, and a_3, a_4, b_3, b_4 equal zero. Using Lemma 7 and (13), we have $\lambda + \mu \le 1/2$.

Case 3. λ and μ are the eigenvalues of different Z_j's. We can prove Conjecture 1 by following the proof in Case 2, except that we switch a_1 and a_3 and switch a_2 and a_4 at the same time.

Case 4. λ is the eigenvalue of some Y_j; μ is the eigenvalue of some Z_k; and $j \neq k$. Without loss of generality, we may assume that $j = 1$ and $k = 2$. By computation, one can show that:

$$\begin{aligned}
\lambda + \mu &= \frac{1}{2}\Big(|a_1|^2 + |a_2|^2 + |a_3|^2 + |a_4|^2 + 2(|b_1|^2 + |b_2|^2) \\
&+ \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4\left|b_1 a_2^* + a_1 b_1^*\right|^2} + \sqrt{(|a_3|^2 - |a_4|^2)^2 + 4\left|b_2 a_4^* + a_3 b_2^*\right|^2}\Big) \\
&\leq \frac{1}{2}\Big(|a_1|^2 + |a_2|^2 + |a_3|^2 + |a_4|^2 + 2(|b_1|^2 + |b_2|^2) \\
&+ \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4|b_1|^2(|a_1| + |a_2|)^2} + \sqrt{(|a_3|^2 - |a_4|^2)^2 + 4|b_2|^2(|a_3| + |a_4|)^2}\Big) \\
&\leq \frac{1}{2}\Big(|a_1|^2 + |a_2|^2 + |a_3|^2 + |a_4|^2 + 2(|b_1|^2 + |b_2|^2) \\
&+ \sqrt{\Big((|a_1| + |a_2|)^2 + (|a_3| + |a_4|)^2\Big)\Big((|a_1| - |a_2|)^2 + (|a_3| - |a_4|)^2 + 4(|b_1|^2 + |b_2|^2)\Big)}\Big) \\
&:= \frac{1}{2}\Big(x + \sqrt{y(2x - y)}\Big) \\
&\leq x \\
&\leq 1/2,
\end{aligned} \tag{20}$$

where $x = |a_1|^2 + |a_2|^2 + |a_3|^2 + |a_4|^2 + 2(|b_1|^2 + |b_2|^2)$. The second inequality in (20) follows from Lemma 6 in which we have set $x = \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4|b_1|^2(|a_1| + |a_2|)^2}$, $y = \sqrt{(|a_3|^2 - |a_4|^2)^2 + 4|b_2|^2(|a_3| + |a_4|)^2}$, $a = (a_3 + a_4)^2$ and $b = (a_1 + a_2)^2$. The last inequality in (20) follows from (13).

Case 5. λ is the eigenvalue of some Y_j, and μ is the eigenvalue of Z_j. Without loss of generality, we may assume that $j = 1$. Equations (12) and (13) imply that $|b_1|^2 = |b_2 + b_3 + b_4|^2 \leq 3(|b_2|^2 + |b_3|^2 + |b_4|^2) = \frac{3}{4} - 3|b_1|^2 - 3\sum_{j=1}^{4} |a_j|^2$. Hence:

$$|b_1|^2 \leq \frac{3}{16} - \frac{3}{4}\sum_{j=1}^{4} |a_j|^2. \tag{21}$$

On the other hand, by computation, one can show that:

$$\begin{aligned}
\lambda + \mu &= \frac{1}{2}\Big(|a_1|^2 + |a_2|^2 + |a_3|^2 + |a_4|^2 + 4|b_1|^2 \\
&+ \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4\left|b_1 a_2^* + a_1 b_1^*\right|^2} + \sqrt{(|a_3|^2 - |a_4|^2)^2 + 4\left|b_1 a_4^* + a_3 b_1^*\right|^2}\Big) \\
&\leq \frac{1}{2}\Big(|a_1|^2 + |a_2|^2 + |a_3|^2 + |a_4|^2 + 4|b_1|^2 \\
&+ \sqrt{(|a_1|^2 - |a_2|^2)^2 + 4|b_1|^2(|a_1| + |a_2|)^2} + \sqrt{(|a_3|^2 - |a_4|^2)^2 + 4|b_1|^2(|a_3| + |a_4|)^2}\Big) \\
&:= \Lambda(|a_1|, |a_2|, |a_3|, |a_4|, |b_1|^2).
\end{aligned} \tag{22}$$

It monotonically increases with $|b_1|^2$. Using (13), we may assume that $|a_1| = x\cos d\cos g$, $|a_2| = x\cos d\sin g$, $|a_3| = x\sin d\cos h$ and $|a_4| = x\sin d\sin h$, where the real numbers $x \in [0,1/2]$, and $d,g,h \in [0,\pi/2]$. Equations (21) and (22) imply that:

$$\begin{aligned}\lambda+\mu &\le \Lambda(|a_1|,|a_2|,|a_3|,|a_4|,\frac{3}{16}-\frac{3}{4}\sum_{j=1}^{4}|a_j|^2)\\ &= \frac{1}{8}\left(3-8x^2+2x\cos d\sqrt{f_1(d,x,g)}+2x\sin d\sqrt{f_2(d,x,h)}\right)\end{aligned} \tag{23}$$

where:

$$f_1(d,x,g) = 3-10x^2+2x^2\cos 2d+(3-12x^2)\sin 2g+(-2x^2-2x^2\cos 2d)\sin^2 2g, \tag{24}$$

and:

$$f_2(d,x,h) = 3-10x^2-2x^2\cos 2d+(3-12x^2)\sin 2h+(-2x^2+2x^2\cos 2d)\sin^2 2h. \tag{25}$$

One can verify that $f_1(d,x,g) = f_2(\pi/2-d,x,g)$. The last equation of (23) is unchanged under the switch of d and $\pi/2-d$ and the switch of g and h at the same time. Therefore, the maximum of (23) is achieved when $x \in [0,1/2]$, $d \in [0,\pi/4]$ and $g,h \in [0,\pi/2]$.

To prove the assertion, one needs to obtain the maximum of (23). For this purpose, we need to obtain the maximum of the function f_1 in terms of g and the maximum of the function f_2 in terms of h. The two functions f_1 and f_2 are quadratic functions in terms of $\sin 2g$ and $\sin 2h$, respectively. The axises of symmetry of f_1 and f_2 are respectively $\sin 2g = \frac{3-12x^2}{4x^2+4x^2\cos 2d}$ and $\sin 2h = \frac{3-12x^2}{4x^2-4x^2\cos 2d}$. If $\sin 2g = 1$ or $\sin 2h = 1$, then we respectively obtain $x = \frac{1}{\sqrt{4+\frac{8}{3}\cos^2 d}}$ or $x = \frac{1}{\sqrt{4+\frac{8}{3}\sin^2 d}}$. We discuss three subcases in terms of the above facts, $\sin 2g \le 1$, $\sin 2h \le 1$ and $\cos d \ge \sin d$.

Subcase 5.1. $x \in [0,\frac{1}{\sqrt{4+\frac{8}{3}\cos^2 d}}]$. One can show that $\max_g f_1(d,x,g) = f_1(d,x,\pi/4)$ and $\max_h f_2(d,x,h) = f_1(d,x,\pi/4)$. Then, one can show that (23) is upper bounded by $1/2$.

Subcase 5.2. $x \in [\frac{1}{\sqrt{4+\frac{8}{3}\cos^2 d}},\frac{1}{\sqrt{4+\frac{8}{3}\sin^2 d}}]$. One can show that $\max_g f_1(d,x,g)$ is achieved when $\sin 2g = \frac{3-12x^2}{4x^2+4x^2\cos 2d}$ and $\max_h f_2(d,x,h) = f_1(d,x,\pi/4)$. Then, one can show that (23) is upper bounded by $3/8$.

Subcase 5.3. $x \in [\frac{1}{\sqrt{4+\frac{8}{3}\sin^2 d}},\frac{1}{2}]$. One can show that $\max_g f_1(d,x,g)$ is achieved when $\sin 2g = \frac{3-12x^2}{4x^2+4x^2\cos 2d}$ and $\max_h f_2(d,x,h)$ is achieved when $\sin 2h = \frac{3-12x^2}{4x^2-4x^2\cos 2d}$. Then, one can show that (23) is upper bounded by $3/8$.

We have shown that (23) is upper bounded by $1/2$, i.e., $\lambda+\mu \le 1/2$.

This completes the proof. □

3.2. Conjecture 1 with Non-Normal Matrices X: Family 2

Definition 3. *Let $\mathcal{P}_2$ be the subset of matrices X with $d = 4$, such that A is normal, i.e.,* $A \sim \begin{bmatrix} a_1 & 0 & 0 & 0\\ 0 & a_2 & 0 & 0\\ 0 & 0 & a_3 & 0\\ 0 & 0 & 0 & a_4\end{bmatrix}$, *and* $B \sim \begin{bmatrix} 0 & b_1 & 0 & 0\\ 0 & 0 & b_2 & 0\\ 0 & 0 & 0 & b_1e^{i\theta_1}\\ b_2e^{i\theta_2} & 0 & 0 & 0\end{bmatrix}$ *or A and B switch from each other. The following constraints should be satisfied.*

$$\sum_{i=1}^{4} a_i = 0, \qquad \sum_{i=1}^{4}|a_i|^2 = \frac{1}{4}-2(|b_1|^2+|b_2|^2). \tag{26}$$

We present the main result of this subsection in Theorem 2. Lemma 8 will be used in the proof of Theorem 2.

Lemma 8. *Suppose $x_1, \cdots, x_n$ are n complex variables and c is a complex constant. If the x_i's satisfy the constraint:*

$$\sum_{i=1}^{n} x_i = c, \quad \forall c \in \mathbb{C}, \tag{27}$$

the following equality holds by setting $x_i = \frac{c}{n}$.

$$\min(\sum_{i=1}^{n} |x_i|^2) = \frac{1}{n} |c|^2. \tag{28}$$

Theorem 2. *The inequality (3) holds when $X \in \mathcal{P}_2$.*

Proof of Theorem 2. For Lemma 1 (iii) and (iv), we assume $A = \begin{bmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & b_1 & 0 & 0 \\ 0 & 0 & b_2 & 0 \\ 0 & 0 & 0 & b_1 e^{i\theta_1} \\ b_2 e^{i\theta_2} & 0 & 0 & 0 \end{bmatrix}$. Applying a local unitary on X, we may assume that $b_1, b_2 \geq 0$ and $\theta_1 = 0$. Then, one can show:

$$X^\dagger X = \oplus_{i=1}^{4} M_i, \tag{29}$$

where:

$$M_i = \begin{bmatrix} |a_i|^2 + b_2^2 & b_1 a_i^* & 0 & a_i b_2 e^{-i\theta_2} \\ a_i b_1 & |a_i|^2 + b_1^2 & b_2 a_i^* & 0 \\ 0 & a_i b_2 & |a_i|^2 + b_2^2 & b_1 a_i^* \\ b_2 e^{i\theta_2} a_i^* & 0 & a_i b_1 & |a_i|^2 + b_1^2 \end{bmatrix}. \tag{30}$$

Set:

$$\tilde{M}_i = \begin{bmatrix} |a_i|^2 + b_2^2 & b_1 a_i^* & 0 & 0 \\ a_i b_1 & |a_i|^2 + b_1^2 & 0 & 0 \\ 0 & 0 & |a_i|^2 + b_2^2 & b_1 a_i^* \\ 0 & 0 & a_i b_1 & |a_i|^2 + b_1^2 \end{bmatrix}. \tag{31}$$

Then, $\tilde{\lambda}_i^1, \tilde{\lambda}_i^2$ are the double eigenvalues of $\tilde{M}_i$ as follows.

$$\begin{aligned} \tilde{\lambda}_i^1 &= \frac{2|a_i|^2 + b_1^2 + b_2^2 + \sqrt{(b_1^2 - b_2^2)^2 + 4|a_i|^2 b_1^2}}{2}, \\ \tilde{\lambda}_i^2 &= \frac{2|a_i|^2 + b_1^2 + b_2^2 - \sqrt{(b_1^2 - b_2^2)^2 + 4|a_i|^2 b_1^2}}{2}. \end{aligned} \tag{32}$$

One can show the characteristic polynomial of M_i as follows.

$$\det(\lambda I - M_i) = ((\lambda - \tilde{\lambda}_i^1)(\lambda - \tilde{\lambda}_i^2))^2 - p(\lambda^2 - (\tilde{\lambda}_i^1 + \tilde{\lambda}_i^2)\lambda + r), \tag{33}$$

where $\tilde{\lambda}_i^1, \tilde{\lambda}_i^2$ are given in Equation (32) and:

$$\begin{aligned} p &= 2\,|a_i|^2\, b_2^2, \\ r &= \frac{1}{p}\big(2\,|a_i|^2\, b_1^2 b_2^4 + 2\,|a_i|^6\, b_2^2 + 2\,|a_i|^4\, b_1^2 b_2^2 + |a_i|^4\, b_2^4 + a_i^4 (b_1^2 b_2^2 e^{i\theta_2})^* + (a_i^4)^* b_1^2 b_2^2 e^{i\theta_2}\big). \end{aligned} \tag{34}$$

We have the four roots of the equation $\det(\lambda I - M_i) = 0$ as follows.

$$\begin{aligned} \lambda_i^1 &= \frac{1}{2}(\tilde{\lambda}_i^1 + \tilde{\lambda}_i^2 + \sqrt{2p + (\tilde{\lambda}_i^1 - \tilde{\lambda}_i^2)^2 + 2\sqrt{p^2 + 4pr - 4\tilde{\lambda}_i^1 \tilde{\lambda}_i^2}}), \\ \lambda_i^2 &= \frac{1}{2}(\tilde{\lambda}_i^1 + \tilde{\lambda}_i^2 + \sqrt{2p + (\tilde{\lambda}_i^1 - \tilde{\lambda}_i^2)^2 - 2\sqrt{p^2 + 4pr - 4\tilde{\lambda}_i^1 \tilde{\lambda}_i^2}}), \\ \lambda_i^3 &= \frac{1}{2}(\tilde{\lambda}_i^1 + \tilde{\lambda}_i^2 - \sqrt{2p + (\tilde{\lambda}_i^1 - \tilde{\lambda}_i^2)^2 - 2\sqrt{p^2 + 4pr - 4\tilde{\lambda}_i^1 \tilde{\lambda}_i^2}}), \\ \lambda_i^4 &= \frac{1}{2}(\tilde{\lambda}_i^1 + \tilde{\lambda}_i^2 - \sqrt{2p + (\tilde{\lambda}_i^1 - \tilde{\lambda}_i^2)^2 + 2\sqrt{p^2 + 4pr - 4\tilde{\lambda}_i^1 \tilde{\lambda}_i^2}}). \end{aligned} \tag{35}$$

Substituting Equation (34) into Equation (35), we have the sum of the largest two eigenvalues, which belong to exact one M_i as follows.

$$\begin{aligned} &\lambda_i^1 + \lambda_i^2 \\ &= 2\,|a_i|^2 + b_1^2 + b_2^2 \\ &+ \frac{1}{2}\sqrt{(b_1^2 - b_2^2)^2 + 4\left|a_i b_1 + a_i^* b_2 e^{i\frac{\theta_2}{2}}\right|^2} \\ &+ \frac{1}{2}\sqrt{(b_1^2 - b_2^2)^2 - 4\left|a_i b_1 + a_i^* b_2 e^{i\frac{\theta_2}{2}}\right|^2}. \end{aligned} \tag{36}$$

We also have the sum of the largest two eigenvalues from two different submatrices M_i, M_j as follows.

$$\begin{aligned} &\lambda_i^1 + \lambda_j^2 \\ &= |a_i|^2 + |a_j|^2 + b_1^2 + b_2^2 \\ &+ \frac{1}{2}\sqrt{(b_1^2 - b_2^2)^2 + 4\left|a_i b_1 + a_i^* b_2 e^{i\frac{\theta_2}{2}}\right|^2} \\ &+ \frac{1}{2}\sqrt{(b_1^2 - b_2^2)^2 + 4\left|a_j b_1 + a_j^* b_2 e^{i\frac{\theta_2}{2}}\right|^2}. \end{aligned} \tag{37}$$

Next, we need to show that both Equations (36) and (37) are upper bounded by $\frac{1}{2}$.

We firstly consider Equation (36). Suppose $b_1^2 + b_2^2 = k$. It follows from Equation (26) that $\sum_{i=1}^{4} a_i = 0$ and $\sum_{i=1}^{4} |a_i|^2 = \frac{1}{4} - 2k$. It is safe to suppose $|a_1|$ is the maximum. Then, we have $\sum_{i=2}^{4} a_i = -a_1$. View a_2, a_3, a_4 as the three variables x_1, x_2, x_3 in Lemma 8 and $-a_1$ as the constant c in Lemma 8. Then, Lemma 8 implies $\min \sum_{i=2}^{4} |a_i|^2 = \frac{1}{3}|a_1|^2$ by setting $a_2 = a_3 = a_4 = -\frac{a_1}{3}$. From $\sum_{i=1}^{4} |a_i|^2 = \frac{1}{4} - 2k$, we obtain $\frac{4}{3}|a_1|^2 = \frac{1}{4} - 2k$. Therefore, we have $\max|a_i|^2 = \frac{3}{4}(\frac{1}{4} - 2k)$. Set $b_1^2 = x, b_2^2 = k - x, a_i = \sqrt{\frac{3}{4}(\frac{1}{4} - 2k)}e^{i\alpha_i}$. Then, Equation (36) can be transformed into the following function.

$$\begin{aligned} f(x) = & \frac{3}{8} - 2k \\ & + \frac{1}{2}\sqrt{(2x-k)^2 + 3(\frac{1}{4} - 2k)(2\sqrt{x(k-x)}\cos\beta + k)} \\ & + \frac{1}{2}\sqrt{(2x-k)^2 + 3(\frac{1}{4} - 2k)(-2\sqrt{x(k-x)}\cos\beta + k)}, \quad \beta = 2\alpha_i - \frac{\theta_2}{2}. \end{aligned} \tag{38}$$

Then, we have the derivative of f as follows.

$$\frac{\partial f}{\partial x} = \frac{f_1(x)}{f_2(x)}, \tag{39}$$

where:

$$\begin{aligned} f_1(x) &= 2(k-2x)|a_i|^2\cos\beta(t_1 - t_2) - 2(k-2x)\sqrt{(k-x)x}(t_1 + t_2), \\ f_2(x) &= \sqrt{(k-x)x}t_1t_2, \\ t_1 &= \sqrt{(k-2x)^2 + 4k|a_i|^2 - 8|a_i|^2\cos\beta\sqrt{(k-x)x}} \\ &\quad - \sqrt{(k-2x)^2 + 4k|a_i|^2 + 8|a_i|^2\cos\beta\sqrt{(k-x)x}}, \\ t_2 &= \sqrt{(k-2x)^2 + 4k|a_i|^2 - 8|a_i|^2\cos\beta\sqrt{(k-x)x}} \\ &\quad + \sqrt{(k-2x)^2 + 4k|a_i|^2 + 8|a_i|^2\cos\beta\sqrt{(k-x)x}}. \end{aligned} \tag{40}$$

One can show that $x = 0, \frac{k}{2}, k$ are three extreme points of $f(x)$. We have $f(0) = f(k) = \frac{3}{8} - 2k + \sqrt{\frac{3k}{4} - 5k^2}$. The maximum of $f(0)$ is $\frac{9}{20}$ when $x = \frac{1}{40}$. We have $f(\frac{k}{2}) = \frac{3}{8} - 2k + \sqrt{\frac{3}{4}(\frac{1}{4} - 2k)(k + k\cos\beta)} + \sqrt{\frac{3}{4}(\frac{1}{4} - 2k)(k - k\cos\beta)}$. The maximum of $f(\frac{k}{2})$ is $\frac{4+\sqrt{10}}{16}(< \frac{9}{20})$ when $\cos\beta = 0, k = \frac{5-\sqrt{10}}{80}$. Therefore, Equation (36) is upper bounded by $\frac{1}{2}$.

We next consider Equation (37) and suppose $b_1^2 + b_2^2 = k, b_1^2 = x, b_2^2 = k - x, a_i = |a_i|^2 e^{i\alpha_i}, a_j = |a_j|^2 e^{i\alpha_j}$ in the same way. One can show that the maximum of Equation (37) with $|a_i|^2 + |a_j|^2 = c$ can be reached when $|a_i| = |a_j|$ and $\cos(2\alpha_i - \frac{\theta_2}{2}) = \cos(2\alpha_j - \frac{\theta_2}{2}) = 0$. Therefore, in order to maximize (37), we can set $|a_i|^2 = |a_j|^2 = \frac{1}{8} - k$. Then, Equation (37) can be transformed into the following function.

$$g(x) = \frac{1}{4} - k + \sqrt{(2x-k)^2 + (\frac{1}{2} - 4k)(k + 2\sqrt{x(k-x)})}. \tag{41}$$

One can show that the maximum of $g(x)$ can only appear when $x = 0, k, \frac{4k-\sqrt{-1+16k-48k^2}}{8}, \frac{4k+\sqrt{-1+16k-48k^2}}{8}$. Then, we have $g(0) = g(k) = \frac{1}{4} - k + \sqrt{\frac{k}{2} - 3k^2}$ and $g(\frac{4k-\sqrt{-1+16k-48k^2}}{8}) = g(\frac{4k+\sqrt{-1+16k-48k^2}}{8}) = \frac{1}{4}\sqrt{16k^2 - 8k + 1}$. One can show that both $\frac{1}{4} - k + \sqrt{\frac{k}{2} - 3k^2}$ and $\frac{1}{4}\sqrt{16k^2 - 8k + 1}$ are upper bounded by $\frac{1}{2}$ when $k \in [0, \frac{1}{8}]$.

Therefore, we conclude that the sum of the largest two eigenvalues of $X^\dagger X$ in Equation (29) is upper bounded by $\frac{1}{2}$. This completes the proof. □

3.3. Conjecture 1 with Non-Normal Matrices X: Family 3

In this subsection, we investigate Conjecture 1 with X for $d \geq 5$ defined as follows.

Definition 4. *Let $\mathcal{P}_3$ be the subset of matrices X with $d \geq 5$, such that $A \sim D_A P_A, B \sim D_B P_B$, where $D_A = \mathrm{diag}(a_1, a_2, \cdots, a_d)$, $D_B = \mathrm{diag}(b_1, b_2, \cdots, b_d)$, and P_A, P_B are permutation matrices with zero-diagonals. The parameters a_i's, b_i's satisfy the following constraint:*

$$\sum_{i=1}^{d}(|a_i|^2 + |b_i|^2) = \frac{1}{d}. \tag{42}$$

It is not an easy task to characterize the characteristic polynomial of $X^\dagger X$, especially when the dimension is large. In Theorem 3, we use the Gershgorin circle theorem and Brauer theorem to study Conjecture 1. They are two important theorems in the field of the localization of eigenvalues. The following fact will be used in the proof of Theorem 3.

It follows from (1) that $X^\dagger X = H_1 + H_2$ where:

$$\begin{aligned} H_1 &:= A^\dagger A \otimes I + I \otimes B^\dagger B, \\ H_2 &:= A^\dagger \otimes B + A \otimes B^\dagger. \end{aligned} \tag{43}$$

Furthermore, the first equation of (2) implies that $\mathrm{Tr} H_2 = 0$, and the second equation of (2) implies that $\mathrm{Tr} X^\dagger X = \mathrm{Tr} H_1 = \sum_{j=1}^{d^2} \sigma_j^2 = 1$. Therefore, $X^\dagger X$ can be regarded as a normalized quantum state in terms of quantum physics.

Theorem 3. *The inequality (3) holds for $d \geq 5$ when $X \in \mathcal{P}_3$.*

Proof of Theorem 3. Recall that H_1 in Equation (43) is diagonal and H_2 in Equation (43) is a Hermitian matrix with zero-diagonal. There exist two permutations σ and τ with $\sigma(k) \neq k, \tau(k) \neq k, \forall k \in \{1, \cdots, d\}$, which are respectively equivalent to P_A and P_B. We find that the $(k, \sigma(k))$ entry of A is a_k and the $(k, \tau(k))$ entry of B is b_k. Therefore, the $(\sigma(k), k)$ entry of $A^\dagger$ is a_k^*, and the $(\tau(k), k)$ entry of $B^\dagger$ is b_k^*. They imply that there are exactly two entries in each row of H_2, which can be expressed with a_i, b_j and their conjugates, for $i, j \in \{1, \cdots, d\}$. This implies that the two elements $a_{\sigma^{-1}(i)}^* b_j$, which is the $\big(d(i-1) + j, d(\sigma^{-1}(i) - 1) + \tau(j)\big)$ entry of $X^\dagger X$, and $a_i b_{\tau^{-1}(j)}^*$, which is the $\big(d(i-1) + j, d(\sigma(i) - 1) + \tau^{-1}(j)\big)$ entry of $X^\dagger X$, are in the same the $(d(i-1) + j)$-th row of $X^\dagger X$ and that both are non-diagonal entries of $X^\dagger X$. Further, the diagonal entry of $X^\dagger X$ in this row is $(\left|a_{\sigma^{-1}(i)}\right|^2 + \left|b_{\tau^{-1}(j)}\right|^2)$. Recall that $\sigma(i) \neq i, \tau(i) \neq i, \forall i \in \{1, \cdots, d\}$. Equation (5) implies that the largest eigenvalue λ_1 of $X^\dagger X$ satisfies:

$$\lambda_1 \quad \leq \quad \max_{i,j} \big(\left|a_{\sigma^{-1}(i)}\right|^2 + \left|b_{\tau^{-1}(j)}\right|^2 + \left|a_{\sigma^{-1}(i)}\right| \left|b_j\right| + |a_i| \left|b_{\tau^{-1}(j)}\right| \big). \tag{44}$$

Applying the basic inequality, we obtain $\left|a_p\right| \left|b_q\right| + |a_s| \, |b_t| \leq (\left|a_p\right|^2 + \left|b_q\right|^2 + |a_s|^2 + |b_t|^2)/2$. Since $p \neq s$ and $q \neq t$, the constraint (42) implies that $|a_i|^2 + \left|b_j\right|^2 \leq \frac{1}{d}$ and $\left|a_p\right| \left|b_q\right| + |a_s| \, |b_t| \leq \frac{1}{2d}$. Hence, we have $\lambda_1 \leq \frac{3}{2d}$, and thus, $\lambda_1 + \lambda_2 \leq \frac{3}{d}$. This implies Conjecture 1 holds for $d \geq 6$. Next, we will prove that Conjecture 1 holds for $d = 5$.

Let us recall (43). The inequality (7) implies that the second largest eigenvalue of $X^\dagger X$ satisfies $\lambda_2(X^\dagger X) \leq \lambda_2(H_1) + \lambda_1(H_2)$, where $\lambda_2(H_1)$ means the second largest eigenvalue of H_1 and $\lambda_1(H_2)$ means the largest eigenvalue of H_2. Since H_1 is diagonal, $\lambda_2(H_1)$ is the second largest diagonal element of H_1. Applying Lemma 4 to H_2, one can show that $\lambda_1(H_2)$ should not be greater than the largest root of these quadratic polynomials $\lambda^2 = \sum_{\substack{k=1 \\ k \neq k_1}}^{n} \left|X^\dagger X_{(k_1,k)}\right| \sum_{\substack{k=1 \\ k \neq k_2}}^{n} \left|X^\dagger X_{(k_2,k)}\right|$ for $k_1 \neq k_2$. Suppose

$k_1 = d(i-1)+j$ and $k_2 = d(p-1)+q$ with $(i,j) \neq (p,q)$. Hence, we can bound $\lambda_1(X^\dagger X) + \lambda_2(X^\dagger X)$ as follows.

$$(\lambda_1+\lambda_2) \leq \max_{(i,j)\neq(p,q)} \left(\left|a_{\sigma^{-1}(i)}\right|^2 + \left|b_{\tau^{-1}(j)}\right|^2 + \left|a_{\sigma^{-1}(p)}\right|^2 + \left|b_{\tau^{-1}(q)}\right|^2 + 2\sqrt{c} \right), \tag{45}$$

where $c = \left(\left|a_{\sigma^{-1}(i)}\right| \left|b_j\right| + |a_i| \left|b_{\tau^{-1}(j)}\right| \right) \left(\left|a_{\sigma^{-1}(p)}\right| \left|b_q\right| + |a_p| \left|b_{\tau^{-1}(q)}\right| \right)$.

Suppose $|a_i|$ and $|b_j|$ are in the decreasing order. In order to maximize the rhs of (45), we can assume that $x_1 = |a_1|$, $x_2 = |a_2|$, $x_3 = |b_1|$ and $x_4 = |b_2|$ and other $|a_i|$'s and $|b_j|$'s are zero. Then, our problem can be transformed into an optimization task as follows.

$$\begin{aligned} \max \quad & f(x_1,x_2,x_3,x_4) = 2x_1^2 + x_3^2 + x_4^2 + 2\sqrt{(x_1x_3+x_2x_4)(x_1x_4+x_2x_3)} \\ s.t. \quad & x_i > 0, \quad i = 1,2,3,4, \\ & \sum_{i=1}^{4} x_i^2 = \frac{1}{d}, \\ & x_1 - x_2 \geq 0, \quad x_3 - x_4 \geq 0, \\ & x_1^2 - x_2^2 \geq x_3^2 - x_4^2. \end{aligned} \tag{46}$$

Since $\sum_{i=1}^{4} x_i^2 = \frac{1}{d}$, we let $x_1 = \sqrt{\frac{1}{d}}(\cos a \cos b)$, $x_2 = \sqrt{\frac{1}{d}}(\cos a \sin b)$, $x_3 = \sqrt{\frac{1}{d}}(\sin a \cos c)$ and $x_4 = \sqrt{\frac{1}{d}}(\sin a \sin c)$. Then, we obtain $f = \frac{1}{4d}\big(4 + \cos 2(a-b) + 2\cos 2b + \cos 2(a+b) + 2\sqrt{2}\sqrt{\sin^2 2a(\sin 2b + \sin 2c)}\big)$, which implies f obtains its maximum only when $\sin 2c = 1$. Furthermore, $\sin 2c = 1$ implies $x_3 = x_4$. Substituting $x_3 = x_4$ and $\sum_{i=1}^{4} x_i^2 = \frac{1}{d}$ into $f(x_1,x_2,x_3,x_4) = 2x_1^2 + x_3^2 + x_4^2 + 2\sqrt{(x_1x_3+x_2x_4)(x_1x_4+x_2x_3)}$, f actually is a two-variable function, that is $f(x_1,x_2) = \frac{1}{d} + x_1^2 - x_2^2 + 2(x_1+x_2)\sqrt{\frac{1}{2d} - \frac{1}{2}(x_1^2+x_2^2)}$. Let $d = 5$. Based on the optimization task expressed by Equation (46), the following task is specifically to show:

$$\begin{aligned} \max \quad & f(x_1,x_2) = \frac{1}{5} + x_1^2 - x_2^2 + 2(x_1+x_2)\sqrt{\frac{1}{10} - \frac{1}{2}(x_1^2+x_2^2)} \leq \frac{1}{2} \\ s.t. \quad & 0 < x_2 \leq x_1, \quad x_1^2 + x_2^2 < \frac{1}{5}. \end{aligned} \tag{47}$$

We first consider the boundary case, i.e., $x_1 = x_2$. In such a case, we have $f_1(x_1) := f(x_1,x_1) = \frac{1}{5} + 4x_1\sqrt{\frac{1}{10} - x_1^2} \leq \frac{2}{5}$. This implies that the maximum over the boundary is upper bounded by $\frac{2}{5}$. Then, let us consider the non-boundary case. It is known that the maximum in this case occurs when both partial derivatives of $f(x_1,x_2)$ equal zero. By computing, one can show the two partial derivatives of $f(x_1,x_2)$ as follows.

$$\begin{aligned} \frac{\partial f(x_1,x_2)}{\partial x_1} &= 2x_1 - \frac{x_1(x_1+x_2)}{\sqrt{\frac{1}{10} - \frac{x_1^2+x_2^2}{2}}} + 2\sqrt{\frac{1}{10} - \frac{x_1^2+x_2^2}{2}}, \\ \frac{\partial f(x_1,x_2)}{\partial x_2} &= -2x_2 - \frac{x_2(x_1+x_2)}{\sqrt{\frac{1}{10} - \frac{x_1^2+x_2^2}{2}}} + 2\sqrt{\frac{1}{10} - \frac{x_1^2+x_2^2}{2}}. \end{aligned} \tag{48}$$

Set the two partial derivatives in Equation (48) equal to zero. One can solve the root as $x_1' = \sqrt{\frac{3}{40} + \frac{1}{10\sqrt{2}}}$, $x_2' = \frac{1}{20}\big(3\sqrt{10(3+2\sqrt{2})} - 4\sqrt{5(3+2\sqrt{2})}\big)$. One can further obtain $f(x_1',x_2') \approx 0.483$, which is strictly less than $\frac{1}{2}$ by computing. Therefore, combining the above two cases, the inequality in Equation (47) holds.

Till now, all dimensions $d \geq 5$ have been studied. This completes the proof. $\square$

Combining Lemma 2 and Theorem 3, we have that Conjecture 1 holds when $X \in \mathcal{N}_d \cup \mathcal{P}_3$ with $d \geq 5$.

3.4. A Sufficient Condition for Conjecture 1

The localization of eigenvalues has been a hot topic of matrix analysis for years. Several results show that eigenvalues can be bounded by traces. In this subsection, we use the recent results on the bounds of eigenvalues to formulate a sufficient condition for Conjecture 1.

Lemma 9. *([22], Theorem 2.1.) Let $A \in \mathbb{H}^{n\times n}$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then, for any $1 \leq k \leq n$,*

$$\left|\lambda_k - \frac{\mathrm{Tr}A}{n}\right| \leq \sqrt{\frac{n-1}{n}\left(\|A\|^2 - \frac{\mathrm{Tr}^2 A}{n}\right)}. \tag{49}$$

Here, $\|A\| := (\mathrm{Tr}A^\dagger A)^{1/2} = (\sum_{j=1}^n \lambda_j^2)^{1/2}$, i.e., the l_2-norm.

Then, we have the following result by using Lemma 9.

Lemma 10. *Conjecture 1 holds for X satisfying $\left\|X^\dagger X\right\|^2 \leq \frac{1}{10}$ when $d = 4$.*

Proof of Lemma 10. Let $d = 4$ in Conjecture 1. Since $X^\dagger X$ is Hermitian, we have $\left\|X^\dagger X\right\|^2 = \mathrm{Tr}(X^\dagger X)^2 = \sum_{j=1}^n \lambda_j^2$, where λ_j's are eigenvalues of $X^\dagger X$. Recall that $\mathrm{Tr}X^\dagger X = 1$. Therefore, $\left\|X^\dagger X\right\|^2$ is lower bounded by $\frac{1}{16}$. Applying Lemma 9 to $X^\dagger X$, the largest eigenvalue of $X^\dagger X$ satisfies:

$$\lambda_1 \leq \frac{1}{16} + \sqrt{\frac{15}{16}}\left(\left\|X^\dagger X\right\|^2 - \frac{1}{16}\right)^{1/2}. \tag{50}$$

One can show that $\left\|X^\dagger X\right\|^2 \leq \frac{1}{10}$ is equivalent to $\frac{1}{16} + \sqrt{\frac{15}{16}}\left(\left\|X^\dagger X\right\|^2 - \frac{1}{16}\right)^{1/2} \leq \frac{1}{4}$. This implies that $\left\|X^\dagger X\right\|^2 \leq \frac{1}{10}$ suffices to obtain $\lambda_1 + \lambda_2 \leq \frac{1}{2}$. This completes the proof. $\square$

4. Discussion

It is known that all entangled two-qubit states are distillable; but for $3 \otimes 3$ or $2 \otimes 4$ systems, there exist bound entangled states: entangled states that cannot be distilled [9]. Furthermore, 2×2 PPT (positive partial transpose) states, which are defined in Section 1, are separable, while 3×3 PPT states may be entangled. This indicates that 3×3 non-distillable states exist. View the two sets of 2×2 PPT states and 3×3 PPT states as two balls, respectively. One can find the difference between the two balls geometrically.

Quantum communications need pure entangled states, which can be distilled from not completely entangled states. The distillability of bipartite states is connected with the possibility of obtaining asymptotically pure maximally-entangled states by LOCC from many copies of a given state. Such maximally-entangled states can be then used for transmitting qubits by means of teleportation. This is why entanglement distillation plays a fundamental role in quantum information.

Definition 1 shows that if Conjecture 1 with $d = 4$ were true, it is still possible that Werner states might be n-distillable with some integer $n > 2$. However it is widely believed that Werner states with $\alpha = -1/2$ may be undistillable [1,2,9,23,24].

To sum up, in this paper, we introduced a main open problem in quantum information, i.e., the distillability problem. One way to attack the problem is to solve the equivalent mathematical problem Conjecture 1. It is known that Conjecture 1 holds for the set of normal matrices with $d \geq 4$. We carried out the first step to prove Conjecture 1 for non-normal matrices and, thus, the distillability problem.

Author Contributions: Funding acquisition, L.C. Investigation, Y.S. and L.C. Supervision, L.C. Writing, original draft, Y.S. and L.C. Writing, review and editing, Y.S. and L.C. All authors contributed equally to this paper. All authors have read and approved the final manuscript.

Funding: Yi Shen and Lin Chen were supported by the Beijing Natural Science Foundation (4173076), the National Natural Science Foundation of China (Grant No. 11501024) and the Fundamental Research Funds for the Central Universities (Grant Nos. KG12040501, ZG226S18C1 and ZG216S1810).

Acknowledgments: The authors are grateful to the editors and anonymous reviewers for their helpful comments and constructive suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. The Brief Proof of Lemma 2 and Some Related New Findings

Proof of Lemma 2. It is easy to verify that the operator X in (1) is normal iff operators A and B are normal. Since X is normal, we can replace singular values with moduli of eigenvalues, which means:

$$\lambda_{ij} = a_i + b_j \tag{A1}$$

where a_i and b_j are eigenvalues of A and B, respectively, and λ_{ij} are eigenvalues of X. We then have:

$$\begin{aligned}\sup_{X\in\mathcal{N}_d}(\sigma_1^2+\sigma_2^2) &= \sup_{X\in\mathcal{N}_d}(|\lambda_1|^2+|\lambda_2|^2) \\ &= \sup_{X\in\mathcal{N}_d}\max_{\substack{i,j,k,l\in\{1,\cdots,d\},\\(i,j)\neq(k,l)}}(|a_i+b_j|^2+|a_k+b_l|^2),\end{aligned}\tag{A2}$$

where λ_1 and λ_2 are two eigenvalues of X with the largest moduli. From (2), parameters a_i's and b_j's should satisfy:

$$\sum_{i=1}^d a_i = \mathrm{Tr}A = 0, \quad \sum_{j=1}^d b_j = \mathrm{Tr}B = 0, \tag{A3}$$

$$\sum_{i=1}^d |a_i|^2 + \sum_{j=1}^d |b_j|^2 = \mathrm{Tr}A^\dagger A + \mathrm{Tr}B^\dagger B = \frac{1}{d}. \tag{A4}$$

The proof of [9] (Theorem 1) shows that (A2) is further equivalent to the following term.

$$\sup_{X\in\mathcal{N}_d}(\sigma_1^2+\sigma_2^2) = \sup_{X\in\mathcal{N}_d}\max\{|a_1+b_1|^2+|a_2+b_2|^2, \quad |a_1+b_1|^2+|a_1+b_2|^2\}. \tag{A5}$$

Then, we have that the following two inequalities hold:

$$|a_1+b_1|^2+|a_2+b_2|^2 \leq \tfrac{1}{2}, \tag{A6}$$
$$|a_1+b_1|^2+|a_1+b_2|^2 \leq \tfrac{1}{2}, \tag{A7}$$

under the constraints (A3) and (A4). The first inequality follows from the identity:

$$|x+y|^2 = 2(|x|^2+|y|^2) - |x-y|^2 \leq 2(|x|^2+|y|^2) \tag{A8}$$

which implies:

$$\begin{aligned}|a_1+b_1|^2+|a_2+b_2|^2 &\leq 2(|a_1|^2+|b_1|^2+|a_2|^2+|b_2|^2) \\ &\leq \frac{2}{d} = \frac{1}{2}.\end{aligned}\tag{A9}$$

The second inequality follows from the following Proposition A1 proven by [9] (Proposition 6). This completes the proof. □

Proposition A1. *Suppose $\overrightarrow{a}$ and $\overrightarrow{b}$ are $d \geq 3$-dimensional vectors with complex elements $\tilde{a}_i$ and $\tilde{b}_i$ satisfying the constraints:*

$$\sum_{i=1}^{d} \tilde{a}_i = \sum_{i=1}^{d} \tilde{b}_i = 0, \qquad \sum_{i=1}^{d} |\tilde{a}_i|^2 + \sum_{i=1}^{d} |\tilde{b}_i|^2 = \frac{1}{d}. \tag{A10}$$

Then, the following equality:

$$\max_{\overrightarrow{a}, \overrightarrow{b}} (|\tilde{a_1} + \tilde{b_1}|^2 + |\tilde{a_1} + \tilde{b_2}|^2) = \frac{3d-4}{d^2} \tag{A11}$$

holds.

Following the ideas in the proof of [9] (Proposition 6), one can similarly obtain the corollary as follows.

Corollary A1. *Suppose $k \in [0, \frac{1}{d}]$, and $\overrightarrow{a}$ and $\overrightarrow{b}$ are $d \geq 3$-dimensional vectors with complex elements $\tilde{a}_i$ and $\tilde{b}_i$ satisfying the constraints:*

$$\sum_{i=1}^{d} \tilde{a}_i = \sum_{i=1}^{d} \tilde{b}_i = 0, \qquad \sum_{i=1}^{d} |\tilde{a}_i|^2 + \sum_{i=1}^{d} |\tilde{b}_i|^2 = \frac{1}{d} - k. \tag{A12}$$

Then, the following equality:

$$\max_{\overrightarrow{a}, \overrightarrow{b}} (|\tilde{a_1} + \tilde{b_1}|^2 + |\tilde{a_1} + \tilde{b_2}|^2) = \frac{(3d-4)(1-dk)}{d^2} \tag{A13}$$

holds.

The maximum can be reached when:

$$\begin{aligned} \tilde{a_1} &= \frac{d-1}{d}\sqrt{\frac{2(1-dk)}{3d-4}}, & \tilde{a_i} &= -\frac{1}{d}\sqrt{\frac{2(1-dk)}{3d-4}}, (i > 1), \\ \tilde{b_1} = \tilde{b_2} &= \frac{d-2}{d}\sqrt{\frac{1-dk}{2(3d-4)}}, & \tilde{b_i} &= -\frac{2}{d}\sqrt{\frac{1-dk}{2(3d-4)}}, (i > 2). \end{aligned} \tag{A14}$$

Next, we show that Conjecture 1 no longer holds when $d = 3$. Therefore, we claim that $d \geq 4$ is essential to Conjecture 1.

Lemma A1. *Let $\mathcal{N}_d$ be a subset of normal operators X in (1) satisfying the constraints (2). Then, for $d = 3$, we have:*

$$\frac{5}{9} \leq \sup_{X \in \mathcal{N}_d} (\sigma_1^2 + \sigma_2^2) \leq \frac{2}{3}, \tag{A15}$$

where σ_1 and σ_2 are the two largest singular values of operator X.

Proof of Lemma A1. According to (A8), we have:

$$|a_1 + b_1|^2 + |a_2 + b_2|^2 \leq 2(|a_1|^2 + |b_1|^2 + |a_2|^2 + |b_2|^2) \leq \frac{2}{d} = \frac{2}{3}. \tag{A16}$$

Proposition A1 implies that:

$$\max(|a_1 + b_1|^2 + |a_1 + b_2|^2) = \frac{3d-4}{d^2} = \frac{5}{9}. \tag{A17}$$

Due to the relation (A5), we obtain (A15) for $d = 3$. $\square$

References

1. DiVincenzo, D.P.; Shor, P.W.; Smolin, J.A.; Terhal, B.M.; Thapliyal, A.V. Evidence for bound entangled states with negative partial transpose. *Phys. Rev. A* **2000**, *61*, 062312.
2. Dür, W.; Cirac, J.I.; Lewenstein, M.; Bruß, D. Distillability and partial transposition in bipartite systems. *Phys. Rev. A* **2000**, *61*, 062313.
3. Terhal, B.M. A family of indecomposable positive linear maps based on entangled quantum states. *Linear Algebra Its Appl.* **2000**, *323*, 61–73.
4. Poon, E. Preservers of maximally entangled states. *Linear Algebra Its Appl.* **2015**, *468*, 122–144.
5. Cariello, D. A gap for PPT entanglement. *Linear Algebra Its Appl.* **2016**, *529*, 89–114.
6. Hou, J.; Qi, X. Linear maps preserving separability of pure states. *Linear Algebra Its Appl.* **2013**, *439*, 1245–1257.
7. Alfsen, E.; Shultz, F. Finding decompositions of a class of separable states. *Linear Algebra Its Appl.* **2012**, *437*, 2613–2629.
8. Woerdeman, H.J. The separability problem and normal completions. *Linear Algebra Its Appl.* **2004**, *376*, 85–95.
9. Pankowski, L.; Piani, M.; Horodecki, M.; Horodecki, P. A few steps more towards npt bound entanglement. *IEEE. Trans. Inf. Theory* **2010**, *56*, 1–19.
10. Vianna, R.O.; Doherty, A.C. Distillability of Werner states using entanglement witnesses and robust semidefinite programs. *Phys. Rev. A* **2006**, *74*, 052306.
11. Bandyopadhyay, S.; Roychowdhury, V. Classes of *n*-copy undistillable quantum states with negative partial transposition. *Phys. Rev. A* **2003**, *68*, 022319.
12. Chen, L.; Chen, Y.-X. Rank-three bipartite entangled states are distillable. *Phys. Rev. A* **2008**, *78*, 022318.
13. Horodecki, M.; Horodecki, P.; Horodecki, R. Inseparable two spin- $\frac{1}{2}$ density matrices can be distilled to a singlet form. *Phys. Rev. Lett. Phys. Rev. A* **1997**, *78*, 574–577.
14. Rains, E.M. Bound on distillable entanglement. *Phys. Rev. A* **1999**, *60*, 179.
15. Chen, L.; Dokovic, D.Z. Distillability and PPT entanglement of low-rank quantum states. *J. Phys. A* **2011**, *44*, 285303.
16. Chen, L.; Dokovic, D.Z. Distillability of non-positive-partial-transpose bipartite quantum states of rank four. *Phys. Rev. A* **2016**, *94*, 052318.
17. Peres, A. Separability criterion for density matrices. *Phys. Rev. Lett.* **1996**, *77*, 1413.
18. Horodecki, M.; Horodecki, P.; Horodecki, R. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A* **1996**, *223*, 1–8.
19. Werner, R.F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **1989**, *40*, 4277–4281.
20. Bernstein, D. *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed.; Princeton University Press: Princeton, NJ, USA, 2009.
21. Horn, R.; Johnson, C. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 1985.
22. Huang, T.; Wang, L. Improving eigenvalues of complex matrices using traces. *Linear Algebra Its Appl.* **2007**, *426*, 841–854.
23. Hiroshima, T. Bound entangled states with non-positive partial transpose exist. *arXiv* **2008**, arXiv:0802.0329
24. Hiroshima, T. A problem of existence of bound entangled states with non-positive partial transpose and the hilbert's 17th problem. *arXiv* **2012**, arXiv:1007.1830.