The PPT square conjecture holds generically for some classes of independent states

Benoît Collins1, Zhi Yin2 and Ping Zhong3

1 Department of Mathematics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan
2 Institute of Advanced Study in Mathematics, Harbin Institute of Technology, Harbin 150006, People’s Republic of China
3 Department of Mathematics and Statistics, University of Wyoming, Dept. 3036, 1000 E. University Ave., Laramie, WY 82071-3036, United States of America

E-mail: collins@math.kyoto-u.ac.jp, hustyinzhi@163.com and pzhong@uwyo.edu

Received 2 July 2018, revised 26 August 2018
Accepted for publication 29 August 2018
Published 14 September 2018

Abstract
Let $|\psi\rangle\langle\psi|$ be a random pure state on $\mathbb{C}^d \otimes \mathbb{C}^s$, where $\psi$ is a random unit vector uniformly distributed on the sphere in $\mathbb{C}^d \otimes \mathbb{C}^s$. Let $\rho_1$ be a random induced state $\rho_1 = \text{Tr}_{\mathbb{C}^s}(|\psi\rangle\langle\psi|)$ whose distribution is $\mu_{d^2,s}$, and let $\rho_2$ be a random induced state following the same distribution $\mu_{d^2,s}$ independent of $\rho_1$. Let $\rho$ be a random state induced by the entanglement swapping of $\rho_1$ and $\rho_2$. We show that the empirical spectrum of $\rho - \mathbb{1}/d^2$ converges almost surely to the Marcenko–Pastur law with parameter $c^2$ as $d \to \infty$ and $s/d \to c$. As an application, we prove that the state $\rho$ is separable generically if $\rho_1$ and $\rho_2$ are PPT entangled.

Keywords: PPT square conjecture, random states, graphical Gaussian calculus

1. Introduction

Quantum entanglement [23] is a resource which can be widely used in quantum information processing. Note that a state is separable if it can be written as a convex combination of product states, and a state is referred to as entangled if it is not separable. One interesting quantum information process is quantum teleportation [5]. In this protocol, Alice can transmit an unknown quantum state to Bob thanks to pre-shared quantum entanglement and classical communication, although it is known that it is impossible to create an ideal copy of an arbitrary quantum state by the no-cloning theorem. Another interesting protocol is entanglement
swapping. It means one can create entanglement between systems which never interact, which is an important technique for quantum teleportation over long distances [32].

In the above two protocols, the maximally entangled state (EPR pair) plays an important role. However, another important type of entangled state exists, called the bound entangled state [22], from which one cannot distill maximally entangled states under local operations and classical communication. For example, states satisfying the positive partial transpose (PPT) property are bound entangled states [22]. Unfortunately, in any deterministic teleportation protocol, the performance of a bound entangled state is not better than a separable (classical) state, as shown in [22]. In view of this, it is natural to wonder what happens to the entanglement swapping protocol with such states.

The so-called PPT square conjecture first appeared in Bäuml’s thesis stated in the following state form [7]: assume that Alice and Charlie share a bound entangled state and that Bob and Charlie also share a bound entangled state; then the state of Alice and Bob, conditioned on any measurement by Charlie, is separable. In other words, this conjecture suggests that the state obtained by the entanglement swapping protocol of PPT entangled states is separable.

There is some evidence to support the PPT square conjecture thus far [7, 8]. In addition, Müller-Hermes recently announced that this conjecture is true for the states on $\mathbb{C}^3 \otimes \mathbb{C}^3$ [17]. However, the main difficulty in studying this conjecture is that we cannot describe the set of all bound entangled states and the conjecture remains open.

Due to the Choi–Jamiołkowski isomorphism [11, 24] between quantum states and quantum channels, there is an equivalent ‘channel’ form of the PPT square conjecture given by Bäuml [7, lemma 14] and Christandl [29]: if $\Phi$ and $\Psi$ are PPT quantum channels, then their composition $\Phi \circ \Psi$ must be entanglement breaking. Recently, Kennedy et al showed that the PPT square conjecture holds asymptotically; namely, they proved that the distance between the iteration of any PPT channel and the set of all entanglement breaking channels goes to zero [25]. This result has been improved by Rahaman et al [28], where they showed that every unital PPT channel becomes entanglement breaking after a finite number of iterations.

In this article, we use two methods widely used in quantum information theory—random matrix theory (RMT) and asymptotic geometric analysis (AGA)—to study the PPT square conjecture. RMT has been heavily used in the non-additivity problem of quantum channels [10, 13, 14, 16, 18–21], and AGA was used to estimate the geometric volume of quantum states with different properties [1, 3, 4, 16, 30, 31].

We outline our approach as follows. We consider two induced states on $\mathbb{C}^d \otimes \mathbb{C}^d$ chosen randomly with distribution $\mu_{d,s}$, where $s$ is the dimension of the environment. Through the work of Aubrun [1] and Aubrun et al [4], we can choose the parameters $s$ and $d$ properly, such that the induced states chosen are PPT entangled generically. In other words, the states are PPT entangled with high probability as $d \to \infty$ and $s/d \to \infty$. Then we study the separability of the state which is obtained from the entanglement swapping protocol of the states. This is done when these two states are chosen independently. Our work is divided into two parts. Firstly, we consider the model of the entanglement swapping protocol of two random induced states and then calculate the moments of the rescaled random matrix model. By using the tools of RMT, we are able to obtain the limits of this model in some asymptotic regime. Secondly, by using AGA and our limit theorem, we show that the state is separable with high probability as $d \to \infty$ and $s/d \to \infty$. In this sense, we prove that the PPT square conjecture holds generically if the states are chosen independently. Moreover, we also consider the random model where the two states are chosen to be the same. However, we can only obtain a weak version of the limit theorem, hence we are not able to use AGA to describe the separability of the induced state directly.
The paper is organized as follows. After this introduction, we collect some relevant results from RMT in section 2. We then introduce our random matrix models and prove some limit theorems in section 3. We apply our results to the PPT square conjecture in section 4 and end the paper with conclusions and some questions.

2. Preliminaries

In this section, we review some relevant results in combinatorics and graphic Gaussian calculus.

2.1. Some combinatorial facts

Let \( I \) be a linearly ordered set of \( p \) elements. We identify it with the set \([p] = \{1, 2, \ldots, p\}\). Denote by \( S_I \) the set of permutations of elements in \( I \). For convenience, we also denote by \( S_p \) the set of permutations of elements in \([p]\). Given a permutation \( \sigma \in S_I \), we denote by \(|\sigma|\) the minimal number of transpositions that multiply to \( \sigma \) and by \( \#\sigma \) the number of cycles of \( \sigma \). We have the following equation

\[
\#\sigma = |I| - |\sigma|.
\]

Let \( d(\sigma, \tau) = |\sigma^{-1}\tau| \), then we have the following triangle inequality:

\[
|\sigma^{-1}\tau| + |\tau^{-1}\pi| \geq |\sigma^{-1}\pi|.
\]

Hence it defines a distance on \( S_I \). We also call \(|\sigma|\) the length of \( \sigma \). If the equality in (2) holds, we say \( \sigma, \tau, \pi \) satisfy the geodesic condition and denote it as \( \sigma = \tau \sim \pi \). We denote the permutations from \( S_p \) which lie on a geodesic from \( \text{id} \) to the full cycle \( \gamma := (1, 2, \cdots, p) \) by

\[
S_{NC}(\gamma) := \{ \pi \in S_p : |\pi| + |\pi^{-1}\gamma| = p - 1 \} \approx \{ \pi \in S_p : \text{id} = \pi - \gamma \}.
\]

For \( \sigma, \pi \in S_{NC}(\gamma) \), we say that \( \sigma \leq \pi \) if \( \sigma \) and \( \pi \) lie on the same geodesic and \( \sigma \) comes before \( \pi \). That is, \( \text{id} = \sigma - \pi - \gamma \) is a geodesic between \( \text{id} \) and \( \gamma \). The set \( S_{NC}(\gamma) \) endowed with ‘\( \leq \)’ becomes a poset. We refer the reader to [27] for more details.

We call \( \pi = \{V_1, \ldots, V_r\} \) a partition of the set \([p]\) if the sets \( V_i \) (\( i = 1, \ldots, r \)) are pairwise disjoint, non-empty subsets of \([p]\) such that \( V_1 \cup \cdots \cup V_r = [p] \). We use \( \#\pi \) to denote the number of blocks of \( \pi \). Given two elements \( a, b \in [p] \), we write \( a \sim b \) if \( a \) and \( b \) belong to the same block of \( \pi \). A partition \( \pi \) is called a crossing if \( a_1 < b_1 < a_2 < b_2 \in [p] \) exist such that \( a_1 \sim a_2 \not\sim b_1 \sim b_2 \). We call \( \pi \) a non-crossing partition if \( \pi \) is not crossing. Denote by \( NC(p) \) the set of all non-crossing partitions of \([p]\). We also denote by \( NC(I) \) the set of all non-crossing partitions of the linearly ordered set \( I \).

A partition can be naturally identified with a permutation. We will use the following identification of non-crossing partitions due to Biane [9] (see also [27, lecture 23]).

**Lemma 2.1.** Let \( \gamma = (1, 2, \ldots, p) \) be the full cycle of \( S_p \). There is a bijection between \( NC(p) \) and the set \( S_{NC}(\gamma) \) which preserves the poset structure.

We end this subsection with the following two technical lemmas.
Lemma 2.2. Denote by $\text{NC}^0(p) = \{ \pi \in \text{NC}(p) : \pi \text{ has no singletons} \}$. We then have
\[
\sum_{J \subset [p]} (-1)^{|J|-|p|} \sum_{\pi \in \text{NC}(J)} c^{p-2|J|+2(\#\pi)} = \sum_{\pi \in \text{NC}(p)} c^{2(\#\pi)-p}.
\]

Proof. It suffices to prove that only the terms with $J = [p]$ and $\pi \in \text{NC}[p]$ without singletons in the above sum survive. Given a subset $J \subset I$ and a non-crossing partition $\pi' \in \text{NC}^0(J)$ having no singletons, we denote by $[\pi']$ the set of non-crossing partitions in $\text{NC}(I)$ which extend $\pi'$ by adding singletons:
\[
[\pi'] = \{ \pi \in \text{NC}(I), \pi = \pi' \cup \{ \text{singletons in } I \setminus J \} \}.
\]

Note the fact that for $\pi \in [\pi']$, we have $\#\pi = \#\pi' + |I| - |J|$ and every non-crossing partition $\pi \in \text{NC}(I)$ except $\{1, 2, \ldots, |p|\}$ can be decomposed as $\pi = \pi' \cup \{ \text{singletons} \}$ by removing singletons from $\pi$. We have
\[
\text{LHS of (3)} = \sum_{J \subset [p]} \sum_{J' \subset [p]} (-1)^{|J|-|p|} \sum_{\pi' \in \text{NC}(J')} c^{p-2|J|+2\#\pi'} \sum_{\pi \in \text{NC}(J)} (-1)^{|J|} c^{p-2|J|+2\#\pi} \delta_{|J|, p} = \text{RHS of (3)},
\]

where we have used the fact that $\sum_{J \subset [p]} (-1)^{|J|-|p|} = 0$ if $|J| \neq p$.

Lemma 2.3. Let $\gamma = (1, 2, \ldots, p)$ be the full cycle of $S_p$. Let $\delta = \gamma \oplus \gamma \in S_{2p}$ and $\beta \in S_{2p}$ such that $\beta(i) = \beta(i+p), i = 1, \ldots, p$. Then for any $\pi \in \{ \pi \in S_{2p} : \text{id} - \pi = \delta \}$, $\pi$ can be decomposed into $\pi_1 \oplus \pi_2$, where $\text{id} = \pi_1, \pi_2 = \gamma$. Moreover, we have
\[
\#(\beta^{-1}\pi) = \#(\pi_1\pi_2).
\]

Proof. We shall use the identification between non-crossing partitions and geodesic permutations (see lemma 2.1). The non-crossing partition
\[
\pi \subset \{1, 2, \ldots, p\}, \{ p+1, p+2, \ldots, 2p \} \subset \text{NC}(2p),
\]
thus $\pi$ can be decomposed into $\pi = \pi_1 \oplus \pi_2$, where $\pi_1 \in \text{NC}(p)$ and $\pi_2$ is a non-crossing partition of the set $\{ p+1, p+2, \ldots, 2p \}$.

Let $V$ be a cycle of $\beta^{-1}\pi = \beta^{-1}(\pi_1 \oplus \pi_2)$. Since $\beta(\{1, 2, \ldots, p\}) = \{ p+1, p+2, \ldots, 2p \}$ and $\{ p+1, p+2, \ldots, 2p \}$ is invariant under $(\beta^{-1}\pi)^2$, we see $V$ must intersect with $\{ p+1, p+2, \ldots, 2p \}$. Now let $1 \leq a \leq p$, we have $\beta^{-1}(\pi_1 \oplus \pi_2)(p+a) = \beta^{-1}(p+\pi_2(a)) = \pi_2(a)$ and $\beta^{-1}(\pi_1 \oplus \pi_2)(\pi_2(a)) = \beta^{-1}(\pi_1(\pi_2(a))) = p + \pi_1(\pi_2(a))$. Hence, by identifying $\{ p+1, p+2, \ldots, 2p \}$ with $\{1, 2, \ldots, p\}$, the action of $(\beta^{-1}\pi)^2$ is the same as the action of $\pi_1\pi_2$. In particular, we have $\#(\beta^{-1}\pi) = \#(\pi_1\pi_2)$.
2.2. Graphical Gaussian calculus

In [13, 14], the first named author and Nechita introduced a graphical formulation of the Weingarten calculus, which is very useful to evaluate moments of the output of the quantum channel they are interested in. Let us briefly review the main ideas and refer the reader to the original article for details.

A diagram is a collection of boxes with certain decorations and possibly wires which connect the boxes along their decorations according to some rules. Each decoration can be either filled (black) or empty (white), which corresponds to vector spaces or their dual spaces, and each wire connecting the shapes attached to the boxes corresponds to a tensor of a vector space with its dual, which produces a partial trace operation. A diagram consisting of boxes and wires is denoted by $D$.

In figure 1, we give some simple examples of diagrams. In figure 1(a), the matrix $G \in \mathbb{M}_{nk}$ is represented as a box with two decorations—— the round one stands for the $n$-dimensional Hilbert space and the square one stands for the $k$-dimensional part. The wire connecting the round decorations stands for tracing over the $n$-dimensional part, where we identify $\mathbb{C}^{nk} \cong \mathbb{C}^n \otimes \mathbb{C}^k$.

After taking partial trace, figure 1(a) ends with a matrix in $\mathbb{M}_k$. The diagram in figure 1(b) represents the (non-normalized) maximally entangled state and figure 1(c) shows the equivalence of two diagrams corresponding to $G^*$ and $G$ respectively.

Let us now very briefly describe how to compute the expectation values of diagrams containing boxes of $G$ and $G^*$, where $G$ is a matrix whose entries are independent random variables with the same distribution $\mathcal{N}(0, 1)$. We label each box in the diagram and fill each box with a matrix $G$, or its relative $G^*$. To emphasize that $G$ is involved in the diagram, we use $D(G)$ to represent such diagrams. Given a permutation $\pi$, a wire will be added to connect a decoration labeled white (respectively black) of the box $G$ having index $i$, with the same decoration labeled black (respectively white) of the box $G^*$ having an index $\pi(i)$. We must add enough wires so that every decoration labeled white must be paired with the same decoration labeled black.

The operation gives us a new diagram $D(G)_\pi$, which is called a removal of the original diagram. Using this operation and the Wick formula, one can describe the expectation $E(D)$ of diagrams $D$ as the following equation [15, theorem 3.2] formally:

$$E(D(G)) = \sum_\pi D(G)_\pi.$$

In order to calculate $D(G)_\pi$, we have to count the contributions for each $\pi$ of every decoration. For instance, suppose there are two kinds of decorations, say $\square$ and $\circ$ and their corresponding dimensions are $n$ and $k$ respectively. Thus for each $\pi$,

$$D(G)_\pi = n^{\#\square} k^{\#\circ},$$

where $\#\square$ (respectively $\#\circ$) denotes the number of loops which connects the $\square$ (respectively $\circ$) decorations.

In this paper, our random matrix models are related to the Wishart matrix $W \in W(n, s)$, which is an $n \times n$ random matrix of the form $W = GG^*$, where $G$ is a $n \times s$ random matrix whose entries are independent identically distributed random variables with the standard complex Gaussian distribution. We will calculate the moments of certain random matrix models using the graphical calculus.
3. Entanglement swapping process of two Wishart matrices

Let $W_1, W_2 \in W(d_1 d_2, s)$ be two Wishart matrices with the same parameters $(d_1 d_2, s)$; we would like to analyze the spectrum distribution of $d_2^2 \times d_2^2$ matrix $W$ which is obtained by the entanglement swapping process of $W_1$ and $W_2$ given by

$$W = \frac{1}{d_1} \Tr_{d_1} \left[ (W_1 \otimes W_2) P_{d_1} \right],$$

where $\Tr_{d_1}$ is the partial trace and $P_{d_1}$ is the Bell projection on the two parties with dimension $d_1$.

More precisely, let $H_i = \mathbb{C}^{d_i} \otimes \mathbb{C}^{d_i}$ and we identify a $d_2^2 \times d_2^2$ matrix with an operator on $B(H_1) \otimes B(H_2)$. Hence for any $T = \sum_{i,j=1}^{d_1} E_{ij} \otimes T_{ij} \in B(H_1) \otimes B(H_2)$, $\Tr_{d_1}(T) = \sum_{i,j=1}^{d_1} T_{jj}$, where $E_{ij}$ denotes the elementary matrix with 1 at the $(i,j)$-entry and 0 at other entries. Using Dirac notation, the Bell vector is $|\phi\rangle = \frac{1}{\sqrt{d_1}} (|1\rangle \otimes |1\rangle + \cdots |d_1\rangle \otimes |d_1\rangle)$, and the Bell projection is $P_{d_1} = d_1 |\phi\rangle \langle \phi| = \sum_{i,j=1}^{d_1} |i\rangle \langle j| \otimes |i\rangle \langle j|$, where $\{|1\rangle, \cdots, |d_1\rangle\}$ is an orthonormal basis of $\mathbb{C}^{d_1}$.

3.1. Moment formulas of the random matrix $W$

**Proposition 3.1.** The moments of $W$ are given by the following formula.

(1) Case I: If $W_1$ and $W_2$ are chosen independently, we have

$$\text{ETr}[W^p] = \sum_{\pi_1, \pi_2 \in S_p} d_1^{\#(\pi_1) + \#(\pi_2)} d_2^{\#(\gamma^{-1} \pi_1) + \#(\pi_2)} s^{\#(\pi_1) + \#(\pi_2) - p},$$

where $\gamma = (1, 2, \ldots, p)$ is the full cycle of $S_p$. 

Figure 1. Diagram (a) for $\Tr_n(G)$, (b) for the non-normalized maximally entangled state, and (c) for the identification between $G^*$ and $G$. 

J. Phys. A: Math. Theor. 51 (2018) 425301
B Collins et al

\[ \begin{array}{c}
\text{(a)} \\
\text{G} \\
\text{(b)} \\
\text{G^*} \\
= \\
\text{G} \\
\end{array} \]
(2) Case II: If \( W_1 = W_2 \), we have

\[
\mathbb{E} \text{Tr}[W^p] = \sum_{\pi \in S_{2p}} \frac{d_2^2\#(\gamma^{-1}\pi)\#(\pi)d_1^2\#(\beta^{-1}\pi)-p}{d_1},
\]

where \( \delta = \gamma \oplus \gamma \in S_{2p}, \gamma = (1, 2, \ldots, p) \) and \( \beta \in S_{2p} \) is defined as \( \beta(i) = \beta(i+p) \).

**Proof.** The diagrams corresponding to \( W \) are given in figure 2.

For case I, by using the graphic Gaussian calculus, we see that the \( p \)th moment of \( W \) is given by the formula

\[
\mathbb{E} \text{Tr}[W^p] = \sum_{\pi_1, \pi_2 \in S_p} \mathcal{D}(G_1, G_2)_{\pi_1, \pi_2}.
\]

To calculate the \( \mathcal{D}(G_1, G_2)_{\pi_1, \pi_2} \), we label the \( G_1 \) (respectively \( G_2 \)) and the \( \overline{G}_1 \) (respectively \( \overline{G}_2 \)) boxes with \( 1, \ldots, p \). A removal \( (\pi_1, \pi_2) \in S_p \times S_p \) of the boxes \( G_1, G_2 \) and \( \overline{G}_1, \overline{G}_2 \) connects the decorations in the following way:

1. the white (respectively black) decorations of the \( i \)th \( G_1 \) block are paired with the white (respectively black) decorations of the \( \pi_1(i) \)th \( \overline{G}_1 \) block;
2. the white (respectively black) decorations of the \( i \)th \( G_2 \) block are paired with the white (respectively black) decorations of the \( \pi_2(i) \)th \( \overline{G}_2 \) block.

We can now compute the contributions for each pairing \( (\pi_1, \pi_2) \) as follows:

1. white ‘\( \circ \)’-loops: \( d_2^2\#(\gamma^{-1}\pi_1)\#(\gamma^{-1}\pi_2) \);
2. white ‘\( \square \)’-loops: \( d_1^2\#(\pi_1^{-1}\pi_2) \);
3. black ‘\( \blacklozenge \)’-loops: \( s\#(\pi_1) + \#(\pi_2) \);
4. normalization factors \( d_1^{-1} \) from the Bell projection \( P_{d_1}; d_1^{-p} \).

Hence

\[
\mathcal{D}(G_1, G_2)_{\pi_1, \pi_2} = d_2^2\#(\gamma^{-1}\pi_1)\#(\gamma^{-1}\pi_2)\#(\pi_1)\#(\pi_2)d_1^2\#(\pi_1^{-1}\pi_2)-p,
\]

which completes the proof of case I.
For case II, we label the $G$ and the $\overline{G}$ boxes in the following manner: $1^T, \ldots, p^T$ for the $G$ (respectively $\overline{G}$) boxes that are on the top of the diagram and $1^B, \ldots, p^B$ for the $G$ (respectively $\overline{G}$) boxes that are on the bottom of the diagram. We shall rename the labels as \{1^T, \ldots, p^T, 1^B, \ldots, p^B\} \simeq \{1, \ldots, 2p\}. With this notation, the two fixed permutations $\delta$ and $\beta \in S_{2p}$ introduced in the main text are as follows: for all $i$,

$$\delta(i^T) = (i + 1)^T, \delta(i^B) = (i + 1)^B, \text{ and } \beta(i^T) = i^B, \beta(i^B) = i^T. $$

Now we have the following formula for the $p$th moment of $W$ given by

$$\mathbb{E}\text{Tr}[W^p] = \sum_{\pi \in S_{2p}} D(G)_{\pi}. $$

A removal $\pi \in S_{2p}$ of the boxes $G$ and $\overline{G}$ connects the decorations in the following way: the white (respectively black) decorations of the $i$th $G$ block are paired with the white (respectively black) decorations of the $\pi(i)$th $G$ block.

On the other hand, the contributions for each pairing $\pi$ are given by

1. white $\circ$-loops: $d_2^{\#(\delta^{-1}\pi)}$;
2. white $\square$-loops: $d_2^{\#(\beta^{-1}\pi)}$;
3. black $\bullet$-loops: $s^{\#\pi}$;
4. normalization factors $d_1^{-1}$ from the Bell projection $P_{d_1}$: $d_1^{-p}$.

Hence

$$D(G)_{\pi} = d_2^{\#(\delta^{-1}\pi)} s^{\#\pi} d_1^{\#(\beta^{-1}\pi) - p}, $$

which finishes the proof.

3.2. Moments of the rescaled matrix of $W$ in the asymptotic regime $d_1, d_2 \to \infty, s/d_2 \to c$

We introduce the following rescaled matrix:

$$Z = d_2 s \left( \frac{W}{d_2^2} - \frac{\mathbb{I}}{d_2} \right). $$

Then we have the following theorem.

**Theorem 3.2.** Let $c > 0$ be a constant; then in cases I and II, the moments of $Z$ under the asymptotic regime $(d_1, d_2 \to \infty\text{ and } s/d_2 \to c)$ are given by

$$\lim_{d_1, d_2 \to \infty, s/d_2 \to c} \frac{1}{d_2^2} \mathbb{E}\text{Tr}[Z^p] = \sum_{\pi \in \text{NC}(p)} c^{2(\#\pi) - p}. $$

**Proof.** By binomial identity, in both cases we have

$$m_p(Z) := \frac{1}{d_2^2} \mathbb{E}\text{Tr}[Z^p] = \frac{1}{d_2^2} \sum_{I \subset [p]} \left( -\frac{s}{d_2} \right)^{|I|} \left( \frac{1}{d_2} \right)^{|I|} \text{Tr}[W^{|I|}]. $$
Case I: Let \( s = cd_2 \), we have

\[
m_p(Z) = \frac{1}{d_2^s} \sum_{|\pi| \leq |p|} \left( -\frac{s}{d_2^s} \right)^{|\pi|} \left( \frac{1}{d_2^s} \right)^{|p|} \sum_{\pi_1, \pi_2 \in S_l} d_2^\#(\gamma_1^{-1}\pi_1) + \#(\gamma_2^{-1}\pi_2) + \#(\pi_2) \cdot d_4^\#(\pi_1^{-1}\pi_2) - |l|,
\]

\[
= \sum_{\pi \in S_l} (-1)^{p-|l|} \sum_{\pi_1, \pi_2 \in S_l} d_2^\#(\pi_1, \pi_2) \cdot c^{\#(\pi_1)} - |\pi_1| + p \cdot d_4^\#(\pi_1^{-1}\pi_2) - |l|,
\]

where \( f_l(\pi_1, \pi_2) = 2|l| - |\gamma_1^{-1}\pi_1| - |\gamma_2^{-1}\pi_2| - |\pi_1| - |\pi_2| - 2 \). Note that the power of \( d_2 \) becomes

\[
f_l(\pi_1, \pi_2) \leq 2(|l| - 1 - |\gamma_1|) = 0,
\]

where we used the fact that \( |\gamma_1| = |l| - 1 \). Therefore the power of \( d_2 \) terms converge to zero as \( d_2 \to \infty \), except in the cases that satisfy the following geodesic condition \( id - \pi_1, \pi_2 - \gamma_1 \). Moreover, for \( \pi_1, \pi_2 \in S_l \), \( \#(\pi_1^{-1}\pi_2) \leq |l| \) and the equality holds if and only if \( \pi_1 = \pi_2 \). Hence, when \( d_1, d_2 \to \infty, s/d_2 \to c \), the only terms in the moments \( m_p(Z) \) which survive are those for which \( id - \pi_1 = \pi_2 - \gamma_1 \) and \( |\pi_1| + |\pi_2| = p \). That is,

\[
\lim_{d_1, d_2 \to \infty} m_p(Z) = \sum_{|\pi| \leq |p|} (-1)^{p-|l|} \sum_{id - \pi = \gamma_1} c^{p-2|\pi|}
\]

\[
= \sum_{|\pi| \leq |p|} (-1)^{p-|l|} \sum_{\pi \in NC(l)} c^{2\#(\pi) - 2|l| + p}
\]

\[
= \sum_{\pi \in NC_p} c^{2\#(\pi) - p}, \tag{9}
\]

where we used lemma 2.2. This completes the proof for case I.

Case II: \( W_1 = W_2 \), by applying proposition 3.1, we have

\[
m_p(Z) = \frac{1}{d_2^s} \sum_{|\pi| \leq |p|} \left( -\frac{s}{d_2^s} \right)^{|\pi|} \left( \frac{1}{d_2^s} \right)^{|p|} \sum_{\pi \in S_{id}, \delta} d_2^\#(\delta_1^{-1}\pi) + \#(\pi) \cdot d_4^\#(\delta_1^{-1}\pi)
\]

\[
= \sum_{|\pi| \leq |p|} (-1)^{p-|l|} \sum_{\pi \in S_{id}, \delta} d_2^{2|l| - |\delta_1^{-1}\pi| - |\pi| - 2} c^{-|\pi| + p} d_4^\#(\delta_1^{-1}\pi) - |l|, \tag{10}
\]

where we have set \( s = cd_2 \). The power of \( d_2 \) is given by

\[
2|l| - |\delta_1^{-1}\pi| - |\pi| - 2 \leq 2|l| - |\delta_1| - 2 = 0,
\]

where we have used the fact that \( |\delta_1| = 2|l| - 2 \). Therefore, when \( d_2 \to \infty, s/d_2 \to c \), the only terms in the moments \( m_p(Z) \) which survive are those for which \( id - \pi - \delta_1 \), and we have

\[
\lim_{d_1, d_2 \to \infty} m_p(Z) = \sum_{|\pi| \leq |p|} (-1)^{p-|l|} \sum_{id - \pi - \delta_1} c^{p-|\pi|} d_4^\#(\delta_1^{-1}\pi) - |l|.
\]

By lemma 2.3, \( \pi \) can be decomposed into the direct sum of \( \pi_1 \oplus \pi_2 \) such that \( id - \pi_1, \pi_2 - \gamma_1 \), and \( \#(\delta_1^{-1}\pi) = \#(\pi_1 \pi_2) \). Hence we have
\[ \lim_{s/d_2 \to \infty} m_p(Z) = \sum_{I \subseteq [p]} (-1)^{|I|} \sum_{\pi \in NC(I)} c^{p-|\pi_1|-|\pi_2|} \]

where we used the identification between the set \( \{ \pi \in S_I : \text{id} - \pi - \gamma \} \) and \( NC(I) \).

Recall that the density of the Marcenko–Pastur distribution with parameter \( c \) is given by

\[ \mu_{MP}(x) = \sqrt{4c - (x-1-c)^2} \frac{1}{2\pi x} \mathbb{1}_{[(\sqrt{c} - 1, \sqrt{c}+1)]}(x) dx. \]

The density function of the standard semicircular distribution is

\[ \mu_{SC}(x) = \frac{1}{2\pi} \sqrt{4-x^2} \mathbb{1}_{[-2,2]}(x) dx. \]

We have the following result similar to [16, corollaries 2.4 and 2.5]. For completeness, we provide the proof. \( \square \)

**Corollary 3.3.** As \( d_1, d_2 \to \infty, s/d_2 \to c \), the random matrix \( Z \) converges in moments to a centered Marchenko–Pastur distribution of parameter \( c^2 \) (rescaled by \( c \)).

The random matrix \( Z \) converges in moments to the standard semicircular distribution when \( d_1, d_2 \to \infty \) and \( s/d_2 \to \infty \).

**Proof.** It is known that all the free cumulants of a centered Marchenko–Pastur (free Poisson) distribution with parameter \( c \) are equal to \( c^2 \), except the first one, which is zero (see [27, lecture 12]). Hence, we read from (8) and the moment-free cumulant formula that the distribution determined by the moment series (8) is the centered Marchenko–Pastur distribution of parameter \( c^2 \) recall by the factor \( c \).

When \( d_1, d_2 \to \infty, s/d_2 \to \infty \), the moment of the limit is nonzero only when \( p \) is even and \( \#\pi = p/2 \). In this case, the set \( \{ \pi \in NC^0(2n) : \#\pi = n \} \) is the same as the non crossing pair partitions of \([2n]\), whose cardinality is the \( n \)th Catalan number. Hence, the limit distribution is the semicircular distribution. \( \square \)

### 3.3. Almost sure convergence of \( Z \) in the asymptotic regime \( d_1 = d_2 \to \infty, s/d_2 \to c \)

The main result in this subsection is as follows.

**Theorem 3.4.** Let \( d_1 = d_2 = d \). For the random matrix \( Z \) defined in (7), if \( W_1, W_2 \) are chosen independently, we have
\[
\frac{1}{d^2} \text{Tr}[Z^p] \to \mathbb{E} \left( \frac{1}{d^2} \text{Tr}[Z^p] \right),
\]
almost surely as \( d \to \infty, s/d \to c. \)

**Proof.**

We have

\[
\text{Tr}[Z^p] = \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \left( \frac{s}{d} \right)^{p-k} \text{Tr} \left( \frac{W}{ds} \right)^k.
\]

Hence, it is equivalent to show

\[
\frac{1}{d^2} \text{Tr} \left( \left( \frac{W}{ds} \right)^p \right) \to \mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left( \left( \frac{W}{ds} \right)^p \right) \right) \text{ almost surely.}
\]

By using the Chebyshev inequality and the Borel–Cantelli lemma, it is sufficient to show the following inequality:

\[
\sum_{d=1}^{\infty} \mathbb{E} \left( \left( \frac{1}{d^2} \text{Tr} \left( \left( \frac{W}{ds} \right)^p \right) \right)^2 \right) < \infty.
\]

When \( d_1 = d_2 = d \) and \( s/d = c \), we write (5) in case (I) of theorem 3.1 as

\[
\frac{1}{d^2} \mathbb{E} \left( \text{Tr} \left( \left( \frac{W}{ds} \right)^p \right) \right) = \sum_{\pi_1, \pi_2 \in S_p} d^f(\pi_1, \pi_2) e^{-|\pi_1| - |\pi_2| + p},
\]

where

\[
f(\pi_1, \pi_2) = 2p - |\gamma^{-1}\pi_1| - |\gamma^{-1}\pi_2| - |\pi_1| - |\pi_2| - 2 + \#(\pi_1^{-1}\pi_2) - p = g(\pi_1, \pi_2) + \#(\pi_1^{-1}\pi_2) - p \leq 0,
\]

and the function \( g \) is defined as

\[
g(\pi_1, \pi_2) = 2p - |\gamma^{-1}\pi_1| - |\gamma^{-1}\pi_2| - |\pi_1| - |\pi_2| - 2
\leq p - |\gamma^{-1}\pi_2| - |\pi_2| - 1 \leq 0.
\]

We hence deduce that \( f(\pi_1, \pi_2) = 0 \) if and only if \( \text{id} - \pi_1 = \pi_2 - \gamma \). Note that \( |\alpha^{-1}\beta| + |\beta^{-1}\pi| \) has the same parity as \( |\alpha^{-1}\pi| \) for any permutation \( \alpha, \beta, \pi \). Hence all possible values of the function \( g(\pi_1, \pi_2) \) are \(-2k, k \in \mathbb{N}\). Therefore, we have

\[
\frac{1}{d^2} \mathbb{E} \left( \text{Tr} \left( \left( \frac{W}{ds} \right)^p \right) \right) = \sum_{id = \pi_1, \pi_2 - \gamma} e^{-|\pi_1| - |\pi_2| + p} + O \left( \frac{1}{d^2} \right).
\]

Hence we have

\[
\left( \mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left( \left( \frac{W}{ds} \right)^p \right) \right) \right)^2 = \sum_{id = \pi_1, \pi_2 - \gamma} e^{-|\pi_1| - |\pi_2| - |\pi_1| - |\pi_2| + 2p} + O \left( \frac{1}{d^2} \right).
\]

(12)
The term $\mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left[ \left( \frac{W}{ds} \right)^p \right] \right)^2$ is more involved to estimate and one needs to introduce the permutation $\tilde{\gamma} = \gamma \oplus \gamma \in S_{2p}$. By graphical Gaussian calculus, we have

$$\mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left[ \left( \frac{W}{ds} \right)^p \right] \right)^2 = \sum_{\pi_1, \pi_2 \in S_{2p}} d^p(\pi_1, \pi_2) c^{-|\pi_1| - |\pi_2| + 2p},$$

where $f(\pi_1, \pi_2) = 4p - |\tilde{\gamma}^{-1}\pi_1| - |\tilde{\gamma}^{-1}\pi_2| - |\pi_1| - |\pi_2| - 4 + \#(\pi_1^{-1}\pi_2) - 2p$. This can be done similarly to the proof of proposition 3.1, and we leave the details to the reader.

One can easily show that

$$\tilde{g}(\pi_1, \pi_2) := 4p - |\tilde{\gamma}^{-1}\pi_1| - |\tilde{\gamma}^{-1}\pi_2| - |\pi_1| - |\pi_2| - 4 \leq 4p - 4 + 2|\tilde{\gamma}| \leq 0.$$ 

The inequality will be saturated when $id - \pi_1, \pi_2 - \tilde{\gamma}$. Note that $|\tilde{\gamma}^{-1}\pi_1| + |\pi_1|$ and $|\tilde{\gamma}^{-1}\pi_2| + |\pi_2|$ has the same parity as $|\tilde{\gamma}| = 2(p - 1)$. Hence, we have

$$\mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left[ \left( \frac{W}{ds} \right)^p \right] \right)^2 = \sum_{id - \pi_1, \pi_2 - \tilde{\gamma}} c^{-|\pi_1| - |\pi_2| + 2p} + O \left( \frac{1}{d^2} \right).$$

Moreover, if $id - \pi_1, \pi_2 - \tilde{\gamma}$, the partitions $\pi_1$ and $\pi_2$ can be decomposed into

$$\pi_1 = \pi_1^{(1)} \oplus \pi_1^{(2)},$$

$$\pi_2 = \pi_2^{(1)} \oplus \pi_2^{(2)},$$

where $\pi_i^{(1)} \in S_p$ and $\pi_i^{(2)} \in S_p$, $i = 1, 2$. We then have

$$\mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left[ \left( \frac{W}{ds} \right)^p \right] \right)^2 = \sum_{\pi_1^{(1)} \oplus \pi_1^{(2)} = \pi_1, \pi_2^{(1)} \oplus \pi_2^{(2)} = \pi_2} c^{-|\pi_1^{(1)}| - |\pi_1^{(2)}| - |\pi_2^{(1)}| - |\pi_2^{(2)}| + 2p} + O \left( \frac{1}{d^2} \right).$$

Combining equations (12) and (13), we have

$$\mathbb{E} \left( \frac{1}{d^2} \text{Tr} \left[ \left( \frac{W}{ds} \right)^p \right] \right)^2 - \left( \mathbb{E} \left( \frac{1}{d^2} \left[ \left( \frac{W}{ds} \right)^p \right] \right) \right)^2 = O \left( \frac{1}{d^2} \right),$$

which completes our proof. \qed

**Corollary 3.5.** Let $d_1 = d_2 = d$. If $W_1, W_2$ are chosen independently, then the distribution of the random matrix $Z$ defined in (7) converges to the centered Marchenko–Pastur distribution of parameter $c^2$ (rescaled by the factor $c$) almost surely as $d \to \infty, s/d \to c$.

**Remark.**

1. Note that the almost sure convergence of $Z$ to the semicircular distribution also follows from the more general result by Bai and Yin [6], which claims that the almost sure convergence holds for any fixed $d_1$ under the asymptotic regime $d_2 \to \infty, s/d_2 \to \infty.$
The method of using the graphical Gaussian calculus to evaluate moments provides an alternative and self-contained proof.

(2) The above result does not hold for case II, because if we let \( d_1 = d_2 = d \), the power of \( d \) terms in equation (10) is as follows: 
\[ 2|I| - |\delta T^{-1} \pi| - |\pi| - 2 + \#(\delta T^{-1} \pi) - |I| \leq \#(\delta T^{-1} \pi) - |I| \leq |I| \neq 0, \]
where \( \pi \in S_{d,d} \). Hence by using our moments technique, we do not know whether the convergence in moments holds or not.

(3) The random matrix model used in this section should be compared with the model studied in [16]. Though these two models are different, their limit distributions are the same.

4. Applications to the PPT square conjecture

4.1. The PPT square conjecture

Let us first recall some notations used in quantum information theory. Consider \( C^n = C^{d_1} \otimes C^{d_2} \) with dimension \( n = d_1 d_2 \). A quantum state \( \rho \) on \( C^n \) is a positive operator with \( \text{Tr}(\rho) = 1 \). A state is called separable if it can be written as a linear combination of product states. A state \( \rho \) that is not separable is called entangled. A state \( \rho \) on \( C^n \) is called PPT (positive partial transpose) if \( \rho^T = (\text{Id} \otimes T)(\rho) \) is a positive operator, where \( T \) is the transpose operator on \( M_{d_2}(C) \).

By definition, we see that the partial transpose of a separable state is always positive. The PPT property is relatively easier to check and is a useful criterion to study entanglement. We refer the reader to [23] for further information about entangled states.

Let \( \rho_1, \rho_2 \) be quantum states on \( C^{d_1} \otimes C^{d_2} \). The typical entanglement swapping protocol can be represented by
\[
\rho = \frac{\text{Tr}_{d_1}[\rho_1 \otimes \rho_2 P_{d_1}]}{\text{Normalization factor}},
\]
where \( P_{d_1} \) is the Bell projection on \( C^{d_1} \otimes C^{d_1} \).

To illustrate the action of entanglement swapping, we look at the case when \( d_1 = d_2 = d \) and \( \rho_1, \rho_2 \) are \( d \)-dimensional maximally entangled states. Then \( \rho \) is also a \( d \)-dimensional maximally entangled state. This can be easily seen via the graphical language in figure 3.

The PPT square conjecture suggests that if \( \rho_1 \) and \( \rho_2 \) are bound entangled states, then \( \rho \) defined in (14) is separable. Since the bound entangled state with negative partial transpose is still a mystery, in practice, we will focus on the PPT bound entangled states [22].

4.2. Random induced state and PPT entanglement threshold

We denote by \( \mu_{n,s} \) the distribution of the induced state \( \text{Tr}_s[\phi] \langle \phi | \), where \( \phi \) is uniformly distributed on the unit sphere in \( C^n \otimes C^s \). A random state \( \rho_{n,s} \) on \( C^n \) with distribution \( \mu_{n,s} \) is called a random induced state.

The induced states can be described by a random matrix model. Let \( W = GG^* \in W(n,s) \) be a Wishart matrix with parameter \((n,s)\). It is known that \( W/\text{Tr}[W] \) is a random state with distribution \( \mu_{n,s} \) (see [26] for example). Moreover, \( W/\text{Tr}[W] \) and \( \text{Tr}[W] \) are independent [12, 26]. In addition, \( \text{Tr}[W] \) is strongly concentrated around \( ns \). Hence in this sense we can write
\[
\rho_{n,s} = \frac{W}{\text{Tr}[W]} \approx \frac{W}{ns} \text{ for sufficiently large } n,s.
\]
With the Wishart matrix model, it is possible to estimate the spectrum of $\rho_{n,s}$ in the asymptotic regime $s, n \to \infty$. By using techniques in AGA, Aubrun [1] and Aubrun et al [4] found the following PPT and entanglement threshold of $\rho_{n,s}$ respectively. Note here that the partial transpose and separability of $\rho_{n,s}$ are with respect to the bipartition $C^n = C^d \otimes C^d$, thus $n = d^2$.

**Proposition 4.1 ([1]).**

For a given $\epsilon > 0$, we have

1. If $s \leq (4 - \epsilon)d^2$, the probability that $\rho_{n,s}$ is PPT exponentially decays to 0 as $s \to \infty$.
2. If $s \geq (4 + \epsilon)d^2$, the probability that $\rho_{n,s}$ is PPT exponentially decays to 1 as $s \to \infty$.

**Proposition 4.2 ([4]).**

For a given $\epsilon > 0$, there exist constants $C_1, C_2$ and a function $s_0 = s_0(d)$ such that

$$C_1 d^3 \leq s_0 \leq C_2 d^3 \log^2(d)$$

and

1. If $s < (1 - \epsilon)s_0$, the probability that $\rho_{n,s}$ is separable exponentially decays to 0 as $s \to \infty$.
2. If $s > (1 + \epsilon)s_0$, the probability that $\rho_{n,s}$ is separable exponentially decays to 1 as $s \to \infty$.

Denote by $M_{n,0}^{sa}$ the set of all $n \times n$ self-adjoint matrices with trace 0. Let $K$ be a convex body in $M_{n,0}^{sa}$, with $\| \cdot \|_K$ the gauge function defined by $\| x \|_K = \inf \{ t \geq 0, x \in tK \}$.

Following [4], define a gauge $\phi_K$ on $\mathbb{R}^{n,0}$ by

$$\phi_K(x) = \int_{U(n)} \| U \text{Diag}(x) U^* \|_K dU.$$ 

There are two crucial facts in the work of Aubrun et al [4].

1. $\mathbb{E} \phi_K(sp(\rho_{n,s} - \mathbb{I}/n)) = \mathbb{E} \| \rho_{n,s} - \mathbb{I}/n \|_K$, where $sp(\rho_{n,s} - \mathbb{I}/n)$ is the spectrum vector of $(\rho_{n,s} - \mathbb{I}/n)$ in $\mathbb{R}^{n,0}$. This is due to the Haar unitary invariance of $(\rho_{n,s} - \mathbb{I}/n)$.
2. When $n$ and $s/n$ tend to infinity, the empirical spectral distribution of $\sqrt{s} (\rho_{n,s} - \mathbb{I}/n)$ converges to $\mu_{SC}$, in probability, with respect to the $\infty$-Wasserstein distance. This is because as $n$ and $s/n$ tend to infinity, $\sqrt{s} (\rho_{n,s} - \mathbb{I}/n)$ almost surely converges to $\mu_{SC}$.
With the above two facts, the gauge of \((\rho_{n,s} - \mathbb{I}/n)\) and \(G_n\) are comparable in the asymptotic regime \(n \to \infty, s/n \to \infty\), where \(G_n\) is a GUE ensemble in \(M_n^{\mathbb{C}}\). More precisely, by [4, proposition 3.1], we have
\[
\mathbb{E} \left\| \rho_{n,s} - \mathbb{I} \right\|_K \approx \mathbb{E} \frac{1}{n \sqrt{s}} \| G_n \|_K, \text{ as } n \to \infty, s/n \to \infty.
\]
The symbol ‘\(\approx\)’ means that the limit of the ratio of the left hand side and the right hand side equals one as \(n \to \infty\) and \(s/n \to \infty\).

Combining the above two propositions, we see that if we chose the parameter properly such that \(4d^2 < s < s_0\), the random state \(\rho_{n,s}\) would be generically PPT entangled. In other words, \(\rho_{n,s}\) is PPT entangled with high probability as \(n \to \infty\).

4.3. The PPT square conjecture generically holds when the states are chosen independently

Let \(\rho_1\) and \(\rho_2\) be two random induced states with distribution \(\mu_{d_1,d_2,s}\). Namely, we can write
\[
\rho_i = \frac{W_i}{\text{Tr}[W_i]} \quad i = 1, 2,
\]
where \(W_i \in W(d_1 d_2, s), i = 1, 2\). Recall that the state that is induced by the ‘entanglement swapping’ protocol is the following:
\[
\rho = \frac{\text{Tr}_{d_1} [\rho_1 \otimes \rho_2 P_{d_1}]}{\text{Normalization factor}} = \frac{\frac{1}{d_1} \text{Tr}_{d_1} [W_1 \otimes W_2 P_{d_1}]}{\text{Normalization factor}}
\]
where \(W\) is defined in (4).

**Lemma 4.3.** If \(\rho_1\) and \(\rho_2\) are two random induced states with distribution \(\mu_{d_1,d_2,s}\), then the random state \(\rho\) on \(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}\) converges in moments to \(\mu_{d_1,d_2,s}\) as \(d_1 \to \infty\).

**Proof.** For \(i = 1, 2\), we let \(W_i = G_i G_i^*\), where \(G_i\)’s entries are \(\{g_{j,k}^{(i)}\}, j = 1, \ldots, d_i, k = 1, \ldots, d_2, t = 1, \ldots, s\), \(i = 1, 2\) and each \(g_{j,k}^{(i)}\) is the standard normal distribution. After a simple calculation we can write
\[
W = GG^*.
\]
where \(G\) is a \(d_1^2 \times s^2\) matrix with the following entries \(\{g_{k,k',t,t'} = 1, \ldots, d_1, t, t' = 1, \ldots, s\}\):
\[
g_{k,k',t,t'} = \frac{1}{\sqrt{d_1^2}} \sum_{j=1}^{d_1} g_{j,k,1}^{(1)} g_{j,k',1}^{(2)}.
\]
(16)

If we suppose that \(\rho_1\) and \(\rho_2\) are independent, i.e. \(g_{j,k}^{(i)}\) are independent standard normal random variables, then due to the classical central limit theorem, the distribution of \(g_{k,k',t,t'}\) converges in moments to a standard normal distribution as \(d_1 \to \infty\) for every \(k, k', t, t'\). Hence \(W \in W(d_1^2, s^2)\) when \(d_1 = \infty\).

By letting \(p = 1\) in the equation (5), we have \(\mathbb{E}(\text{Tr}[W]) = d_1^2 s^2\). Moreover, \(W/\text{Tr}[W]\) and \(\text{Tr}[W]\) are independent as \(d_1 \to \infty\) (this is due to lemma 4.3 and the discussion in section 4.2).

In addition, \(W\) is strongly concentrated around \(d_1^2 s^2\) for sufficiently large \(d_1, d_2\) and \(s\). Hence
we can formally use $W/(d_2^2 s^2)$ to replace $\rho = W/(\text{Tr}[W])$ for sufficiently large $d_1, d_2$ and $s$ (see [2, proposition 6.34] and the discussion after that). In addition, we can further treat $W$ following the same distribution as the Wishart matrix $W(d_2^2, s^2)$ for large $d_1$ for our purpose.

**Remark.** The above lemma and discussion provide a direct explanation that the results in section 3.2 also hold for our model $W$ defined in (4) when $W_1, W_2$ are chosen independently. To make the argument rigorous, we need the almost sure convergence developed in section 3.3.

In the rest of this section, we assume $d_1 = d_2 = d$, and denote $n = d^2$. Suppose $\rho_1$ and $\rho_2$ are two independent random states with distribution $\mu_{dE}$, which are generically PPT entangled, then the state $\rho$ in equation (15) is generically separable. Our idea is to adapt the arguments in [4]. To this end, according to the lemma 4.3 and the discussion, we can approximately write $\rho = W/ns^2$ for sufficiently large $d$ and $s$, where $W \in W(n, s^2)$. So instead of $\rho$, we can consider $W/ns^2$ in this asymptotic picture $(d, s \to \infty)$. Recall that there are two important ingredients in [4]. The fact that $\rho = W/ns^2$ enables us to compare the gauge of the spectrum vector of $\rho - \mathbb{I}/n$ and itself is due to the Haar unitary invariance of $W$. On the other hand, for the second ingredient, we have to use our theorem 3.4, by which we can compare the gauge $\rho - \mathbb{I}/n$ and the GUE ensemble $G_n$. Then by using the concentration of the measure technique (see for instance [4, section 2.2]), we can show the separability of $\rho$.

In conclusion, we can roughly say that the distribution of $\rho$ is $\mu_{n,s^2}$, and the parameter $s^2$ makes $\rho$ generically separable. The following is our final theorem.

**Theorem 4.4.** Let $\rho_1$ and $\rho_2$ be two independent random states with distribution $\mu_{dE}$, which are generically PPT entangled, then the state $\rho$ in equation (15) is generically separable.

**Proof.** By lemma 4.3 and the discussion above, we can formally write

$$\rho = \frac{W}{ns^2}, \text{ as } d, s \to \infty,$$

where $W \in W(d^2, s^2)$. Denote by $S$ the set of all separable quantum states on $\mathbb{C}^d \otimes \mathbb{C}^d$, and let $S_0 = S - \mathbb{I}/n$. Obviously, $S_0$ is a convex set of $M_n^{sa,0}$. Hence for the convex body $S_0$ in $M_n^{sa,0}$, we have

$$\mathbb{E} \phi_{S_0}(sp(\rho - \mathbb{I}/n)) = \mathbb{E} \left\| \rho - \frac{\mathbb{I}}{n} \right\|_{S_0}, \text{ as } d \to \infty.$$

Here let us mention that in [4], the above equation holds for arbitrary $d$. However, the condition $d \to \infty$ is necessary in our paper, since $\rho$ is Haar unitary invariant only if $d = d_1 \to \infty$, which is a necessary condition for the equation. On the other hand, combining with the theorem 3.4, $ds(\rho - \mathbb{I}/n)$ almost surely converges to $\mu_{SC}$ as $d \to \infty, s/d \to \infty$. Similar to [4, proposition 3.1] we have

$$\mathbb{E} \left\| \rho - \frac{\mathbb{I}}{n} \right\|_{S_0} \approx \mathbb{E} \frac{1}{ns} \left\| G_n \right\|_{S_0}, \text{ as } d \to \infty, s/d \to \infty,$$

where $G_n$ is a GUE ensemble in $M_n^{sa,0}$. The symbol ‘$\approx$’ means the terms of the left (right) hand side are bounded by each other (up to constants $\varepsilon_{d, s}, c_{d, s},$ and $c_{d, s} \xrightarrow{asd \to \infty, s/d \to \infty}$).
Recall that the quantum state $\rho$ on $\mathbb{C}^d \otimes \mathbb{C}^d$ is separable if and only if 
\[ \|\rho\|_{S_0} \leq 1 \iff \left\| \rho - \frac{1}{n} I \right\|_{S_0} \leq 1. \]

Moreover, we have, by [4, section 4],
\[ \mathbb{E} \left\| \frac{G_n}{n} \right\|_{S_0} = O(d^2 \log d). \]

Hence
\[ \mathbb{E} \left\| \rho - \frac{1}{n} I \right\|_{S_0} = O(d^2 \log d). \]

By the concentration of the measure technique, for any $t > 0$ we have
\[ \mathbb{P} \left( \left\| \rho - \frac{1}{n} I \right\|_{S_0} > \mathbb{E} \left\| \rho - \frac{1}{n} I \right\|_{S_0} + t \right) \leq e^{-2t^2/n}, \]

where we have used the fact that $\rho \rightarrow \|\rho\|_{S_0}$ is a $2n$-Lipschitz function (see [4, lemma 3.4]) on the real sphere $S^{2n-1}$. If $s^2 > s_0 = d^3 \log d$, then
\[ \mathbb{P} \left( \left\| \rho - \frac{1}{n} I \right\|_{S_0} > 1 + t \right) \leq e^{-2t^2/n} \leq e^{-dt^2}. \]

Hence the probability that $\rho$ is entangled decays exponentially to 0 as $d \to \infty$. However, since $\rho_1$ and $\rho_2$ are PPT entangled, the required parameters should satisfy $s > 4d^2$. Thus $s^2 > 16d^4 > s_0$ for sufficiently large $d$, which implies that $\rho$ is generically separable. \hfill \Box

**Remark.** The above discussion on the concentration of measure techniques is an adaption of Szarek’s lecture ‘Geometric Functional Analysis and QIT’ in the trimester program of the Centre Emile Borel ‘Analysis in Quantum Information Theory’.

5. Conclusion

In this paper, we studied the random matrix which is obtained by the ‘entanglement swapping’ protocol of two Wishart matrices. By using Gaussian graphical calculus, we are able to obtain some limit theorems of our random models, where the limit distribution is the Marcenko–Pastur law (respectively semi-circle law) under proper asymptotic regime. An interesting application is that we have proved that the PPT square conjecture holds generically if we independently choose the states.

Some people believe that the PPT square conjecture might not hold true thanks, for example, to heuristic volume considerations of the relative volume of PPT states inside all states, compared to the relative volume of separable states. In addition, thanks to the existence of famous precedents, such as random counterexamples to the MOE additivity problem [14, 20],

---

4Szarek’s lecture notes for a talk given in the trimester ‘Analysis in Quantum Information Theory’ which is based on the book [2] and the paper [4].
we think that it is natural to hope for a counterexample to the PPT square conjecture built with random techniques. However, our investigations so far tend rather to serve as evidence that the PPT square conjecture might be true (at least, very often, in a natural probabilistic sense). We considered many random models, including non-independent models, but they seem to yield similar conclusions, so we focused on a specific natural model where both maps are independent.

This suggests that in order to exhibit a counterexample, there must exist a correlation between the chosen states, and that these correlations have to be of a type that remains to be uncovered, but a priori not of the same flavor as those successfully used in the broad area of probabilistic quantum information theory.

Acknowledgments

BC is supported by JSPS Kakenhi 17K18734, 15KK0162 and 17H04823, and ANR-14-CE25-0003. ZY is supported by NSFC No. 11771106 and No. 11431011. PZ is partially supported by NSFC (No. 11501423, No. 11671308) and NSERC research grant. PZ also wishes to thank IASM of the Harbin Institute of Technology for their warm hospitality in summer 2016. The work was conducted when PZ was a postdoctoral fellow at the University of Waterloo. The authors are grateful to the IHP for holding the trimester ‘Analysis in Quantum Information Theory’ during fall 2017, which provided a fruitful environment to meet and work on the project of this paper.

ORCID iDs

Zhi Yin https://orcid.org/0000-0002-9972-8979

References

[1] Aubrun G 2012 Partial transposition of random states and non-centered semicircular distributions Random Matrices: Theory Appl. 1 1250001
[2] Aubrun G and Szarek S J 2017 Alice and Bob Meet Banach: the Interface of Asymptotic Geometric Analysis and Quantum Information Theory (Mathematical Surveys and Monographs vol 223) (Providence, RI: American Mathematical Society)
[3] Aubrun G, Szarek S J and Werner E 2011 Hastings additivity counterexample via Dvoretzky’s theorem Commun. Math. Phys. 305 85–97
[4] Aubrun G, Szarek S J and Ye D P 2014 Entanglement thresholds for random induced states Commun. Pure Appl. Math. 67 129–71
[5] Bennett C H, Brassard G, Crepeau C, Jozsa R, Peres A and Wootters W K 1993 Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels Phys. Rev. Lett. 70 1895–9
[6] Bai Z D and Yin Y Q 1988 Convergence to the semicircle law Ann. Probab. 16 863–75
[7] Bäuml S 2010 On bound key and the use of bound entanglement Diploma Thesis Ludwig Maximilians Universität München
[8] Bäuml S, Christandl M, Horodecki K and Winter A 2015 Limitations on quantum key repeaters Nat. Commun. 6 6908
[9] Biane P 1997 Some properties of crossings and partitions Discrete Math. 175 41–53
[10] Brandao F and Horodecki M 2010 On Hastings’ counterexamples to the minimum output entropy additivity conjecture Open Syst. Inf. Dyn. 17 31
[11] Choi M D 1975 Completely positive linear maps on complex matrices Linear Algebr. Appl. 10 285
[12] Collins B 2005 Product of random projections, Jacobi ensembles and universality problems arising from free probability Probab. Theory Relat. Fields **133** 315–44
[13] Collins B and Nechita I 2010 Random quantum channels I: graphical calculus and the Bell state phenomenon Commun. Math. Phys. **297** 345–70
[14] Collins B and Nechita I 2011 Random quantum channels II: entanglement of random subspaces, Rényi entropy estimates and additivity problems Adv. Math. **226** 1181–201
[15] Collins B and Nechita I 2011 Gaussianization and eigenvalue statistics for random quantum channels (III) Ann. Appl. Probab. **21** 1136–79
[16] Collins B, Nechita I and Ye D P 2012 The absolute positive partial transpose property for random induced states Random Matrices: Theory Appl. **01** 1250002
[17] Christandl M, Müller-Hermes A and Wolf M M 2018 When do composed maps become entanglement breaking? (arXiv:1807.01266)
[18] Fukuda M and Nechita I 2015 Additivity rates and PPT property for random quantum channels Ann. Math. Blaise Pascal **22** 1–72
[19] Grudka A, Horodecki M and Pankowski Ł 2010 Constructive counterexamples to the additivity of the minimum output Rényi entropy of quantum channels for all $p > 2$ J. Phys. A: Math. Theor. **43** 425304
[20] Hastings M B 2009 Superadditivity of communication capacity using entangled inputs Nat. Phys. **5** 255
[21] Hayden P and Winter A 2008 Counterexamples to the maximal $p$-norm multiplicativity conjecture for all $p > 1$ Commun. Math. Phys. **284** 263
[22] Horodecki M, Horodecki P and Horodecki R 1998 Mixed-state entanglement and distillation: is there a bound entanglement in nature Phys. Rev. Lett. **80** 5239
[23] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2007 Quantum entanglement Rev. Mod. Phys. **81** 865–942
[24] Jamiołkowski A 1972 Linear transformations which preserve trace and positive semidefiniteness of operators Rep. Math. Phys. **3** 275
[25] Kennedy M, Manor N A and Paulsen V I 2017 Composition of PPT maps (arXiv:1710.08475)
[26] Nechita I 2007 Asymptotics of random density matrices Ann. Henri Poincaré **8** 1521–38
[27] Nica A and Speicher R 2006 *Lectures on the Combinatorics of Free Probability* (LMS Lecture Note Series vol 335) (Cambridge: Cambridge University Press)
[28] Rahaman M, Jaques S and Paulsen V I 2018 Eventually entanglement breaking maps (arXiv: 1801.05542)
[29] Ruskai M, Junge M, Kribs D, Hayden P and Winter A 2012 Banff Int. Research Station Workshop: Operator Structures in Quantum Information Theory www.birs.ca/workshops/2012/12w5084/report12w5084.pdf (Accessed: 31 September 2018)
[30] Sommers H-J and Zyczkowski K 2003 Bures volume of the set of mixed quantum states J. Phys. A: Math. Theor. **37** 8457
[31] Zyczkowski K and Sommers H-J 2001 Induced measures in the space of mixed quantum states J. Phys. A: Math. Gen. **34** 7111–25
[32] Żukowski M, Zeilinger A, Horne M A and Ekert A K 1993 ‘Event-ready-detectors’ Bell experiment via entanglement swapping Phys. Rev. Lett. **71** 4287–90