THE PPT SQUARE CONJECTURE HOLDS GENERICALLY FOR SOME CLASSES OF INDEPENDENT STATES

BENOÎT COLLINS, ZHI YIN, AND PING ZHONG

ABSTRACT. Let $|\psi\rangle\langle\psi|$ be a random pure state on $\mathbb{C}^{d^2} \otimes \mathbb{C}^s$, where ψ is a random unit vector uniformly distributed on the sphere in $\mathbb{C}^{d^2} \otimes \mathbb{C}^s$. Let ρ_1 be a random induced state $\rho_1 = \mathrm{Tr}_{\mathbb{C}^s}(|\psi\rangle\langle\psi|)$ whose distribution is $\mu_{d^2,s}$; and let ρ_2 be a random induced state following the same distribution $\mu_{d^2,s}$ independent from ρ_1. Let ρ be a random state induced by the entanglement swapping of ρ_1 and ρ_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 ρ is separable generically if ρ_1, ρ_2 are PPT entangled.

1. INTRODUCTION

Quantum entanglement [25] 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 called entangled if it is not separable. One interesting quantum information processing is the quantum teleportation [5]. In this protocol, Alice can transmit an unknown quantum state to Bob thanks to preshared 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 the entanglement swapping. It means one can create entanglement between systems which never interact, which is an important technique for quantum teleportation over long distances [35].

In the above two protocols, the maximally entangled state (EPR pair) plays an important role. However, there exists another important type of entangled states, called the bound entangled states [24], from which one can not distill maximally entangled states under local operations and classical communication. For example, states satisfying the positive partial transpose (PPT) property are bound entangled states [24]. Unfortunately, in any deterministic teleportation protocol, the performance of a bound entangled state is not better than a separable (classical) state as shown in [24]. 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 are some evidences to support the PPT square conjecture up to now [7, 8]. In addition, Müller-Hermes announced that this conjecture is true for the states on $\mathbb{C}^3 \otimes \mathbb{C}^3$ [19] recently. However, one main difficulty to study this conjecture is that we can not describe the set of all bound entangled states and the conjecture remains open.

Due to the Choi-Jamiołkowski isomorphism [26, 13] between quantum states and quantum channels, there is an equivalent "channel" form of the PPT square conjecture given by Bäuml [7, Lemma 14] and and Christandl [31]: *if* Φ *and* Ψ *are PPT quantum channels, then their composition* $\Phi \circ \Psi$ *must be entanglement breaking.* Recently, Kennedy-Manor-Paulsen 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 [27]. This result has been improved by Rahaman-Jaques-Paulsen [30], 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. The RMT was heavily used in the non-additivity problem of quantum channels [11, 15, 16, 18, 20, 22, 23], and AGA was used to estimate the geometric volume of quantum states with different properties [1, 4, 18, 32, 34].

We outline our approach as follows. We consider two induced states on $\mathbb{C}^d \otimes \mathbb{C}^d$ chosen randomly with distribution $\mu_{d^2,s}$, where s is the dimension of the environment. By Aubrun [1] and Aubrun-Szarek-Ye's work [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 entanglement swapping protocol of two random induced states and then calculate the moments of the rescaled random matrix model. By using 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 get a weak version of 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 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 σ and by $\#\sigma$ the number of cycles of σ. We have the following equation

$$\#\sigma = |I| - |\sigma|. \tag{1}$$

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

$$|\sigma^{-1}\tau| + |\tau^{-1}\pi| \geq |\sigma^{-1}\pi|. \tag{2}$$

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

$$\begin{aligned} S_{NC}(\gamma) : &= \{\pi \in S_p : |\pi| + |\pi^{-1}\gamma| = p - 1\} \\ &= \{\pi \in S_p : \mathrm{id} - \pi - \gamma\}. \end{aligned}$$

For $\sigma, \pi \in S_{NC}(\gamma)$, we say that $\sigma \leq \pi$ if σ and π lie on the same geodesic and σ comes before π. That is, $\mathrm{id} - \sigma - \pi - \gamma$ is a geodesic between id and γ. The set $S_{NC}(\gamma)$ endowed with "$\leq$" becomes a poset. We refer the reader to [29] 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 π. Given two elements $a, b \in [p]$, we write $a \sim_\pi b$ if a and b belong to the same block of π. A partition π is called crossing if there exist $a_1 < b_1 < a_2 < b_2 \in [p]$ such that $a_1 \sim_\pi a_2 \not\sim_\pi b_1 \sim_\pi b_2$. We call π *non-crossing* partition if π 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 [10] (see also [29, 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 by the following two technical lemmas.

Lemma 2.2. *Denote by $NC^0(p) = \{\pi \in NC(p) : \pi$ has no singletons$\}$. We then have*

$$\sum_{I \subset [p]} (-1)^{p-|I|} \sum_{\pi \in NC(I)} c^{p-2|I|+2(\#\pi)} = \sum_{\pi \in NC^0(p)} c^{2(\#\pi)-p}. \tag{3}$$

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

$$[\pi'] = \{\pi \in NC(I), \pi = \pi' \cup \{\text{singletons in } I\backslash J\}\}.$$

Noting the fact that for $\pi \in [\pi']$, we have $\#\pi = \#\pi' + |I| - |J|$ and every non-crossing partition $\pi \in NC(I)$ except $\{\{1\}, \{2\}, \cdots, \{p\}\}$ can be decomposed as $\pi = \pi' \cup \{\text{singletons}\}$ by removing singletons from π. We have

$$\begin{aligned}
\text{LHS of (3)} &= \sum_{J\subset[p]} \sum_{J\subset I\subset[p]} (-1)^{p-|I|} \sum_{\substack{\pi\in NC(I)\cap[\pi'] \\ \pi'\in NC^0(J)}} c^{p-2|I|+2\#\pi} \\
&= \sum_{J\subset[p]} \sum_{\pi'\in NC^0(J)} c^{p-2|J|+2\#\pi'} \sum_{J\subset I\subset[p]} (-1)^{p-|I|} \\
&= \sum_{J\subset[p]} \sum_{\pi'\in NC^0(J)} c^{p-2|J|+2\#\pi'} \delta_{|J|,p} = \text{ RHS of (3)},
\end{aligned}$$

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

Lemma 2.3. *Let* $\gamma = (1, 2, \dots, p)$ *be the full cycle of* S_p. *Let* $\delta = \gamma \oplus \gamma \in S_{2p}$ *and* $\beta \in S_{2p}$ *such that* $\beta(i) = \beta(i + p), i = 1, \dots, p$. *Then for any* $\pi \in \{\pi \in S_{2p} : \mathrm{id} - \pi - \delta\}$, π *can be decomposed into* $\pi_1 \oplus \pi_2$, *where* $\mathrm{id} - \pi_1, \pi_2 - \gamma$. *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

$$\pi \leq \{\{1, 2, \dots, p\}, \{p + 1, p + 2, \dots, 2p\}\} \subset NC(2p),$$

thus π can be decomposed into $\pi = \pi_1 \oplus \pi_2$, where $\pi_1 \in NC(p)$ and π_2 is a non-crossing partition of the set $\{p + 1, p + 2, \dots, 2p\}$.

Let V be a cycle of $\beta^{-1}\pi = \beta^{-1}(\pi_1 \oplus \pi_2)$. Since $\beta(\{1, 2, \dots, p\}) = \{p + 1, p + 2, \dots, 2p\}$ and $\{p + 1, p + 2, \dots, 2p\}$ is invariant under $(\beta^{-1}\pi)^2$, we see V must intersect with $\{p + 1, p + 2, \dots, 2p\}$. 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 $\{p + 1, p + 2, \dots, 2p\}$ with $\{1, 2, \dots, p\}$, 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 [15, 16], 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

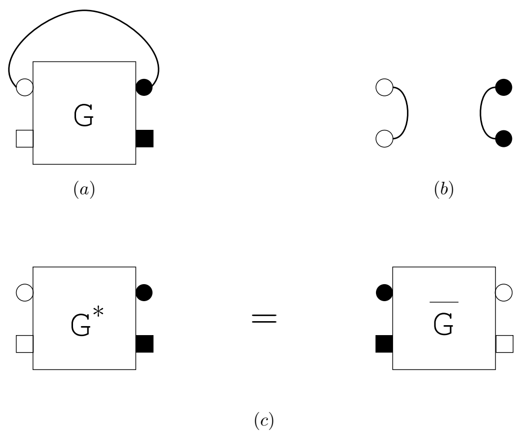

FIGURE 1. diagram (a) for $\mathrm{Tr}_n(G)$, (b) for the non-normalized maximally entangled state, and (c) for the identification between G^* and $\overline{G}$.

decoration can be either filled (black) or empty (white), which corresponds to vector spaces or their dual spaces. And each wire connecting shapes attached to 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 $\mathcal{D}$.

In Figure 1, we give some simple examples of diagrams. In Figure 1 (a), the matrix $G \in M_{nk}$ 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 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 $\overline{G}$ respectively.

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

The operation gives us a new diagram $\mathcal{D}(G)_\pi$, which is called a removal of the original diagram. Using this operation and the Wick formula, one can describe the expectation $\mathbb{E}(\mathcal{D})$ of diagrams $\mathcal{D}$ as the following equation [17, Theorem 3.2] formally:

$$\mathbb{E}(\mathcal{D}(G)) = \sum_{\pi} \mathcal{D}(G)_{\pi}.$$

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

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

where $\#\square$ (resp. $\#\bigcirc$) denotes the number of loops which connects the $\square$ (resp. $\bigcirc$) 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

$$W = \frac{1}{d_1} \mathrm{Tr}_{d_1} \left[(W_1 \otimes W_2)\, P_{d_1} \right], \tag{4}$$

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 $d_1^2 d_2^2 \times d_1^2 d_2^2$ matrix with an operator on $B(H_1) \otimes B(H_2)$. Hence for any $T = \sum_{i,j=1}^{d_1^2} E_{ij} \otimes T_{ij} \in B(H_1) \otimes B(H_2)$, $\mathrm{Tr}_{d_1}(T) = \sum_{j=1}^{d_1^2} T_{jj}$, 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*

$$\mathbb{E}\mathrm{Tr}[W^p] = \sum_{\pi_1, \pi_2 \in S_p} d_2^{\#(\gamma^{-1}\pi_1) + \#(\gamma^{-1}\pi_2)} s^{\#(\pi_1) + \#(\pi_2)} d_1^{\#(\pi_1^{-1}\pi_2) - p}, \tag{5}$$

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

(2) *Case II: If* $W_1 = W_2$, *we have*

$$\mathbb{E}\mathrm{Tr}[W^p] = \sum_{\pi \in S_{2p}} d_2^{\#(\delta^{-1}\pi)} s^{\#(\pi)} d_1^{\#(\beta^{-1}\pi)-p}, \tag{6}$$

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

Proof. The diagrams corresponding to W are given in Figure 2.

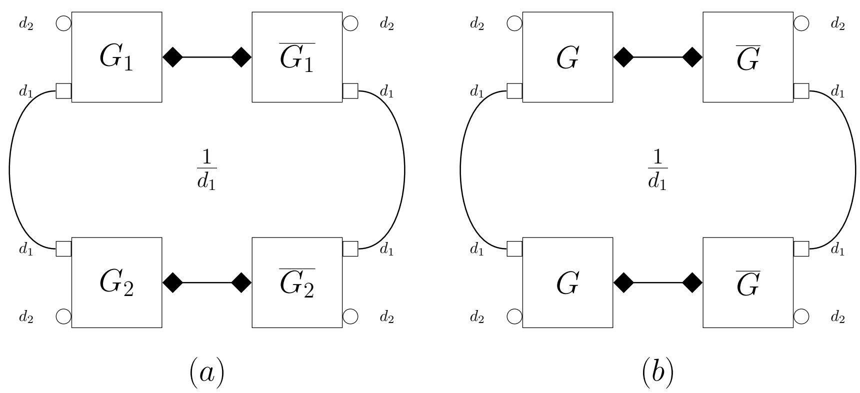

FIGURE 2. diagram (a) for the case I, (b) for the case II.

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

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

To calculate the $\mathcal{D}(G_1, G_2)_{\pi_1,\pi_2}$, we label the G_1 (resp. G_2) and the $\overline{G_1}$ (resp. $\overline{G_2}$) 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 (resp. black) decorations of the i-th G_1 block are paired with the white (resp. black) decorations of the $\pi_1(i)$-th $\overline{G_1}$ block;
(2) the white (resp. black) decorations of the i-th G_2 block are paired with the white (resp. black) decorations of the $\pi_2(i)$-th $\overline{G_2}$ block.

We can now compute the contributions for each pairing (π_1, π_2) as follows,

(1) white "○"-loops: $d_2^{\#(\gamma^{-1}\pi_1)+\#(\gamma^{-1}\pi_2)}$;
(2) white "□"-loops: $d_1^{\#(\pi_1^{-1}\pi_2)}$;
(3) black "◆"-loops: $s^{\#\pi_1+\#\pi_2}$;
(4) normalization factors d_1^{-1} from the Bell projection P_{d_1}: d_1^{-p}.

Hence

$$\mathcal{D}(G_1, G_2)_{\pi_1,\pi_2} = d_2^{\#(\gamma^{-1}\pi_1)+\#(\gamma^{-1}\pi_2)} s^{\#(\pi_1)+\#(\pi_2)} d_1^{\#(\pi_1^{-1}\pi_2)-p},$$

which completes the proof of the case I.

For Case II, we label the G and the $\overline{G}$ boxes in the following manner: $1^T, \dots, p^T$ for the G (resp. $\overline{G}$) boxes that are on the top of the diagram and $1^B, \dots, p^B$ for the G (resp. $\overline{G}$) boxes that are on the bottom of the diagram. We shall rename the labels as $\{1^T, \dots, p^T, 1^B, \dots, p^B\} \simeq \{1, \dots, 2p\}$. With this notation, the two fixed permutations δ and $\beta \in S_{2p}$ introduced in the main text are the following: 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 p-th moment of W given by

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

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

On the other hand, the contributions for each pairing π are given by

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

Hence

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

which finishes the proof. □

3.2. **Moments of 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:

$$Z = d_2 s \left(\frac{W}{d_2^2 s^2} - \frac{\mathbb{1}}{d_2^2} \right). \tag{7}$$

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 ($d_1, d_2 \to \infty$ and $s/d_2 \to c$) are given by*

$$\lim_{\substack{d_1, d_2 \to \infty \\ s/d_2 \to c}} \frac{1}{d_2^2} \mathbb{E}\mathrm{Tr}[Z^p] = \sum_{\pi \in NC^0(p)} c^{2(\#\pi) - p}. \tag{8}$$

Proof. By binomial identity, in both cases we have

$$m_p(Z) := \frac{1}{d_2^2} \mathbb{E}\mathrm{Tr}[Z^p] = \frac{1}{d_2^2} \sum_{I \subset [p]} \left(-\frac{s}{d_2} \right)^{|I^c|} \left(\frac{1}{d_2 s} \right)^{|I|} \mathrm{Tr}[W^{|I|}].$$

Case I: Let $s = cd_2$, we have

$$\begin{aligned} m_p(Z) &= \frac{1}{d_2^2} \sum_{I \subset [p]} \left(-\frac{s}{d_2}\right)^{|I^c|} \left(\frac{1}{d_2 s}\right)^{|I|} \cdot \\ &\cdot \sum_{\pi_1, \pi_2 \in S_I} d_2^{\#(\gamma_I^{-1}\pi_1) + \#(\gamma_I^{-1}\pi_1)} s^{\#(\pi_1)+\#(\pi_2)} d_1^{\#(\pi_1^{-1}\pi_2) - |I|} \\ &= \sum_{I \subset [p]} (-1)^{p-|I|} \sum_{\pi_1, \pi_2 \in S_I} d_2^{f_I(\pi_1,\pi_2)} c^{-|\pi_1| - |\pi_2| + p} d_1^{\#(\pi_1^{-1}\pi_2) - |I|}, \end{aligned}$$

where $f_I(\pi_1, \pi_2) = 2|I| - |\gamma_I^{-1}\pi_1| - |\gamma_I^{-1}\pi_2| - |\pi_1| - |\pi_2| - 2$. Note that the power of d_2 becomes

$$f_I(\pi_1, \pi_2) \leq 2(|I| - 1 - |\gamma_I|) = 0,$$

where we used the fact that $|\gamma_I| = |I| - 1$. 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 $id - \pi_1, \pi_2 - \gamma_I$. Moreover, for $\pi_1, \pi_2 \in S_I$, $\#(\pi_1^{-1}\pi_2) \leq |I|$ 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 $id - \pi_1 = \pi_2 - \gamma_I$ and $|\pi_1| + |\pi_2| = p$. That is,

$$\begin{aligned} \lim_{\substack{d_1, d_2 \to \infty \\ s/d_2 \to c}} m_p(Z) &= \sum_{I \subset [p]} (-1)^{p-|I|} \sum_{id - \pi - \gamma_I} c^{p - 2|\pi|} \\ &= \sum_{I \subset [p]} (-1)^{p-|I|} \sum_{\pi \in NC(I)} c^{2\#\pi - 2|I| + p} \\ &= \sum_{\pi \in NC^0(p)} c^{2(\#\pi) - p}, \end{aligned} \tag{9}$$

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

Case II: $W_1 = W_2$, by applying Proposition 3.1, we have

$$\begin{aligned} m_p(Z) &= \sum_{I \subset [p]} \frac{1}{d_2^2} \left(-\frac{s}{d_2}\right)^{|I^c|} \left(\frac{1}{d_2 s}\right)^{|I|} \sum_{\pi \in S_{I \cup I}} d_2^{\#(\delta_I^{-1}\pi)} s^{\#(\pi)} d_1^{\#(\beta_I^{-1}\pi)} \\ &= \sum_{I \subset [p]} (-1)^{p-|I|} \sum_{\pi \in S_{I \cup I}} d_2^{2|I| - |\delta_I^{-1}\pi| - |\pi| - 2} c^{-|\pi| + p} d_1^{\#(\beta_I^{-1}\pi) - |I|}, \end{aligned} \tag{10}$$

where we have set $s = cd_2$. The power of d_2 is given by

$$2|I| - |\delta_I^{-1}\pi| - |\pi| - 2 \leq 2|I| - |\delta_I| - 2 = 0,$$

where we have used the fact that $|\delta_I| = 2|I| - 2$. 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 $id -$

$\pi - \delta_I$, and we have

$$\lim_{\substack{d_2\to\infty \\ s/d_2\to c}} m_p(Z) = \sum_{I\subset[p]} (-1)^{p-|I|} \sum_{id-\pi-\delta_I} c^{p-|\pi|} d_1^{\#(\beta_I^{-1}\pi)-|I|}.$$

By Lemma 2.3, π can be decomposed into direct sum of $\pi_1 \oplus \pi_2$ such that $\mathrm{id} - \pi_1, \pi_2 - \gamma_I$, and $\#(\beta_I^{-1}\pi) = \#(\pi_1\pi_2)$. Hence we have

$$\begin{aligned}
\lim_{\substack{d_1,d_2\to\infty \\ s/d_2\to c}} m_p(Z) &= \sum_{I\subset[p]} (-1)^{p-|I|} \sum_{id-\pi_1=\pi_2^{-1}-\gamma_I} c^{p-|\pi_1|-|\pi_2|} \\
&= \sum_{I\subset[p]} (-1)^{p-|I|} \sum_{id-\pi_1=\pi_2-\gamma_I} c^{p-|\pi_1|-|\pi_2|} \\
&= \sum_{I\subset[p]} (-1)^{p-|I|} \sum_{id-\pi-\gamma_I} c^{p-2|\pi|} \\
&= \sum_{I\subset[p]} (-1)^{p-|I|} \sum_{\pi\in NC(I)} c^{2\#\pi-2|I|+p} \\
&= \sum_{\pi\in NC^0(p)} c^{2(\#\pi)-p},
\end{aligned}$$

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

Recall that the density of the Marcenko-Pastur distribution with parameter c is given by

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

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 [18, Corollary 2.4, 2.5]. For completeness, we provide the proof.

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 paramemter c^2 are equal to c^2, except the first one, which is zero (see [29, 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 recalling 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. □

3.3. **Almost surely 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 the following result.

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*

$$\frac{1}{d^2}\mathrm{Tr}[Z^p] \to \mathbb{E}\left(\frac{1}{d^2}\mathrm{Tr}[Z^p]\right),$$

almost surely as $d \to \infty, s/d \to c$.

Proof. We have

$$\mathrm{Tr}[Z^p] = \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \left(\frac{s}{d}\right)^{p-k} \mathrm{Tr}\left[\left(\frac{W}{ds}\right)^k\right]. \tag{11}$$

Hence, it is equivalent to show

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

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

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

When $d_1 = d_2 = d$ and $s/d = c$, we write (5) in Case (I) of Theorem 3.1 as

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

where

$$\begin{aligned} f(\pi_1,\pi_2) &= 2p - |\gamma^{-1}\pi_1| - |\gamma^{-1}\pi_2| - |\pi_1| - |\pi_2| - 2 + \#(\pi_1^{-1}\pi_2) - p \\ &= g(\pi_1,\pi_2) + \#(\pi_1^{-1}\pi_2) - p \le 0, \end{aligned}$$

and the function g is defined as

$$\begin{aligned} g(\pi_1,\pi_2) &= 2p - |\gamma^{-1}\pi_1| - |\gamma^{-1}\pi_2| - |\pi_1| - |\pi_2| - 2 \\ &\le p - |\gamma^{-1}\pi_2| - |\pi_2| - 1 \le 0. \end{aligned}$$

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

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

Hence we have

$$\left(\mathbb{E}\left(\frac{1}{d^2}\mathrm{Tr}\left[\left(\frac{W}{ds}\right)^p\right]\right)\right)^2 = \sum_{\substack{id-\pi_1,\pi_2-\gamma\\ id-\pi_1',\pi_2'-\gamma}} c^{-|\pi_1|-|\pi_2|-|\pi_1'|-|\pi_2'|+2p} + O\left(\frac{1}{d^2}\right). \tag{12}$$

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

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

where $\bar{f}(\pi_1,\pi_2) = 4p - |\bar{\gamma}^{-1}\pi_1| - |\bar{\gamma}^{-1}\pi_2| - |\pi_1| - |\pi_2| - 4 + \#(\pi_1^{-1}\pi_2) - 2p$. 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

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

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

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

Moreover, if $id - \pi_1, \pi_2 - \bar{\gamma}$, the partitions π_1 and π_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

$$\mathbb{E}\left(\frac{1}{d^2}\mathrm{Tr}\left[\left(\frac{W}{ds}\right)^p\right]\right)^2 = \sum_{\substack{id-\pi_1^{(1)},\pi_1^{(2)}-\gamma\\ id-\pi_2^{(1)},\pi_2^{(2)}-\gamma}} c^{-|\pi_1^{(1)}|-|\pi_1^{(2)}|-|\pi_2^{(1)}|-|\pi_2^{(2)}|+2p} + O\left(\frac{1}{d^2}\right). \tag{13}$$

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

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

which completes our proof. □

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 Marcheko-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 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 the case II, because if we let $d_1 = d_2 = d$, the power of d terms in equation (10) is the following: $2|I| - |\delta_I^{-1}\pi| - |\pi| - 2 + \#(\beta_I^{-1}\pi) - |I| \leq \#(\beta_I^{-1}\pi) - |I| \leq |I| \neq 0$, where $\pi \in S_{I \sqcup I}$. 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 [18]. 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 $\mathbb{C}^n = \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ with dimension $n = d_1 d_2$. A quantum state ρ on $\mathbb{C}^n$ is a positive operator with $\mathrm{Tr}(\rho) = 1$. A state is called *separable* if it can be written as a linear combination of product states. A state ρ that is not separable is called *entangled*. A state ρ on $\mathbb{C}^n$ is called PPT (positive partial transpose) if $\rho^\Gamma = (\mathrm{Id} \otimes T)(\rho)$ is a positive operator, where T is the transpose operator on $M_{d_2}(\mathbb{C})$. 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 [25] for some more information about entangled states.

Let ρ_1, ρ_2 be quantum states on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$, the typical entanglement swapping protocol can be represented by

$$\rho = \frac{\mathrm{Tr}_{d_1}[\rho_1 \otimes \rho_2 \ P_{d_1}]}{\text{Normalization factor}}, \tag{14}$$

where P_{d_1} is the Bell projection on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_1}$.

To illustrate the action of entanglement swapping, we look at the case when $d_1 = d_2 = d$ and ρ_1, ρ_2 are d-dimensional maximally entangled states. Then ρ is also a d-dimensional maximally entangled states. This can be easily seen via the graphical language in Figure 3.

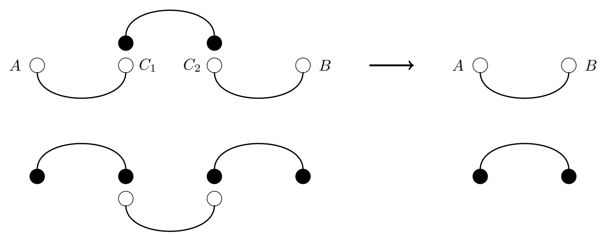

FIGURE 3. Diagram picture for the entanglement swapping of maximally entangled state.

The PPT square conjecture suggests that if ρ_1 and ρ_2 are bound entangled states then ρ 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 [24].

4.2. **Random induced state and PPT entanglement threshold.** We denote by $\mu_{n,s}$ the distribution of the induced state $\mathrm{Tr}_s|\phi\rangle\langle\phi|$, where ϕ is uniformly distributed on the unit sphere in $\mathbb{C}^n \otimes \mathbb{C}^s$. A random state $\rho_{n,s}$ on $\mathbb{C}^n$ 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 $W/\mathrm{Tr}[W]$ is a random state with distribution $\mu_{n,s}$ (see [28] for example). Moreover, $W/\mathrm{Tr}[W]$ and $\mathrm{Tr}[W]$ are independent [28, 14]. In addition, $\mathrm{Tr}[W]$ is strongly concentrated around ns. Hence in this sense we can write

$$\rho_{n,s} = \frac{W}{\mathrm{Tr}[W]} \approx \frac{W}{ns} \text{ for sufficiently large } n, s.$$

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-Szarek-Ye [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 $\mathbb{C}^n = \mathbb{C}^d \otimes \mathbb{C}^d$, thus $n = d^2$.

Proposition 4.1. [1] *For given $\epsilon > 0$, we have*

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

Proposition 4.2. [4] *For 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 $\rho_{d^2,s}$ is separable exponentially decays to 0 as $s \to \infty$.*

(2) *If* $s > (1+\epsilon)s_0$*, the probability that* $\rho_{d^2,s}$ *is separable exponentially decays to* 1 *as* $s \to \infty$.

Denote by $M_n^{sa,0}$ the set of all $n \times n$ self-adjoint matrices with trace 0. Let K be a convex body in $M_n^{sa,0}$, with $\|\cdot\|_K$ the gauge function defined by $\|x\|_K = \inf\{t \geq 0,\ x \in tK\}$. Following [4], define a gauge ϕ_K on $\mathbb{R}^{n,0}$ by

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

There are two crucial facts in Aubrun-Szarek-Ye's work [4].

(1) $\mathbb{E}\ \phi_K(sp(\rho_{n,s} - \mathbb{1}/n)) = \mathbb{E}\|\rho_{n,s} - \mathbb{1}/n\|_K$, where $sp(\rho_{n,s} - \mathbb{1}/n)$ is the spectrum vector of $(\rho_{n,s} - \mathbb{1}/n)$ in $\mathbb{R}^{n,0}$. This is due to the Haar unitary invariance of $(\rho_{n,s} - \mathbb{1}/n)$.
(2) When n and s/n tend to infinity, the empirical spectral distribution of $\sqrt{ns}(\rho_{n,s} - \mathbb{1}/n)$ converges to μ_{SC}, in probability, with respect to the ∞-Wasserstein distance. This is because as n and s/n tend to infinity, $\sqrt{ns}(\rho_{n,s} - \mathbb{1}/n)$ almost surely converges to μ_{SC}.

With the above two facts, the gauge of $(\rho_{n,s} - \mathbb{1}/n)$ and G_n are comparable in the asymptotic regime $n \to \infty, s/n \to \infty$, where G_n is a GUE ensemble in $M_n^{sa,0}$. More precisely, by [4, Proposition 3.1], we have

$$\mathbb{E}\ \left\|\rho_{n,s} - \frac{\mathbb{1}}{n}\right\|_K \approx \mathbb{E}\ \frac{1}{n\sqrt{s}}\ \|G_n\|_K\,, \text{as } n \to \infty, s/n \to \infty.$$

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 ρ_1 and ρ_2 be two random induced states with distribution $\mu_{d_1 d_2,s}$. Namely, we can write

$$\rho_i = \frac{W_i}{\mathrm{Tr}[W_i]}, i = 1, 2,$$

where $W_i \in W(d_1 d_2, s), i = 1, 2$. Recall the state that is induced by the "entanglement swapping" protocol is the following:

$$\begin{aligned}\rho = \frac{\mathrm{Tr}_{d_1}\left[\rho_1 \otimes \rho_2\ P_{d_1}\right]}{\text{Normalization factor}} &= \frac{\frac{1}{d_1}\mathrm{Tr}_{d_1}\left[W_1 \otimes W_2\ P_{d_1}\right]}{\text{Normalization factor}} \\ = \frac{W}{\mathrm{Tr}[W]},\end{aligned} \tag{15}$$

where W is defined in (4).

Lemma 4.3. *If* ρ_1 *and* ρ_2 *are two random induced states with distribution* $\mu_{d_1 d_2,s}$*, then the random state* ρ *on* $\mathbb{C}^{d_2} \otimes \mathbb{C}^{d_2}$ *converges in moments to* $\mu_{d_2^2,s^2}$ *as* $d_1 \to \infty$.

Proof. For $i = 1, 2$, we let $W_i = G_i G_i^*$, where G_i's entries are $\{g_{jk,t}^{(i)}, j = 1, \ldots, d_1, k = 1, \ldots, d_2, t = 1, \ldots, s\}, i = 1, 2$ and each $g_{jk,t}^{(i)}$ is the standard normal distribution. After simple calculation we can write

$$W = GG^*,$$

where G is a $d_2^2 \times s^2$ matrix with the following entries $\{g_{kk',tt'}, k, k' = 1, \ldots, d_2, t, t' = 1, \ldots, s\}$:

$$g_{kk',tt'} = \frac{1}{\sqrt{d_1}} \sum_{j=1}^{d_1} g_{jk,t}^{(1)} g_{jk',t'}^{(2)}. \tag{16}$$

If we suppose ρ_1 and ρ_2 are independent, i.e, $g_{jk,t}^{(i)}$ are independent standard normal random variables, then due to the classical central limit theorem, the distribution of $g_{kk',tt'}$ converges in moments to a standard normal distribution as $d_1 \to \infty$ for every k, k', t, t'. Hence $W \in \mathcal{W}(d_2^2, s^2)$ when $d_1 = \infty$. □

By letting $p = 1$ in the equation (5), we have $\mathbb{E}(\mathrm{Tr}[W]) = d_2^2 s^2$. Moreover, $W/\mathrm{Tr}[W]$ and $\mathrm{Tr}[W]$ 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 $d_2^2 s^2$ for sufficiently large d_1, d_2 and s. Hence we can formally use $W/(d_2^2 s^2)$ to replace $\rho = W/(\mathrm{Tr}[W])$ 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 $\mathcal{W}(d_2^2, s^2)$ 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 ρ_1 and ρ_2 are two independent random states with distribution $\mu_{d^2,s}$ which are generically PPT entangled, then the state ρ 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 $W \in \mathcal{W}(n, s^2)$. So instead of ρ, we can consider W/ns^2 in this asymptotic picture $(d, s \to \infty)$. Recall 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 $\rho - \mathbb{1}/n$ and itself, this 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 $\rho - \mathbb{1}/n$ and the GUE ensemble G_n. Then by using the concentration of measure technique (see for instance [4, Section 2.2]), we can show the separability of ρ.

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

Theorem 4.4. *Let ρ_1 and ρ_2 be two independent random states with distribution $\mu_{d^2,s}$ which are generically PPT entangled, then the state ρ 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 $\mathcal{S}$ the set of all separable quantum states on $\mathbb{C}^d \otimes \mathbb{C}^d$, and let $\mathcal{S}_0 = \mathcal{S} - \mathbb{1}/n$. Obviously, $\mathcal{S}_0$ is a convex set of $M_n^{sa,0}$. Hence for the convex body $\mathcal{S}_0$ in $M_n^{sa,0}$, we have

$$\mathbb{E}\ \phi_{\mathcal{S}_0}(sp(\rho - \mathbb{1}/n)) = \mathbb{E}\ \left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0}, \text{as } d \to \infty.$$

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 ρ 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, $ds(\rho - \mathbb{1}/n)$ almost surely converges to μ_{SC} as $d \to \infty, s/d \to \infty$. Similar to [4, Proposition 3.1] we have

$$\mathbb{E}\ \left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0} \approx \mathbb{E}\ \frac{1}{ns} \|G_n\|_{\mathcal{S}_0}, \text{as } d \to \infty, s/d \to \infty,$$

where G_n is a GUE ensemble in $M_n^{sa,0}$. The symbol "$\approx$" means the terms of left (right) hand side are bounded by each other (up to constants $c_{d,s}, C_{d,s}$, and $c_{d,s}, C_{d,s} \to 1 \text{as} d \to \infty, s/d \to \infty$).

Recall that the quantum state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$ is separable if and only if [4]

$$\|\rho\|_{\mathcal{S}_0} \le 1 \iff \left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0} \le 1.$$

Moreover, we have, by [4, Section 4],

$$\mathbb{E} \left\| \frac{G_n}{n} \right\|_{\mathcal{S}_0} = O(d^{\frac{3}{2}} \log d).$$

Hence

$$\mathbb{E}\ \left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0} = \frac{O(d^{\frac{3}{2}} \log d)}{s}.$$

By the concentration of measure technique, for any $t > 0$ we have

$$\mathbb{P}\left(\left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0} > \mathbb{E}\ \left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0} + t \right) \le e^{-s^2 t^2/n},$$

where we have used the fact that $\rho \to \|\rho\|_{\mathcal{S}_0}$ is a $2n$-Lipschitz function (see [4, Lemma 3.4]) on the real sphere S^{2ns^2-1}. If $s^2 > s_0 = d^3 \log^2 d$, then

$$\mathbb{P}\left(\left\| \rho - \frac{\mathbb{1}}{n} \right\|_{\mathcal{S}_0} > 1 + t \right) \le e^{-s^2 t^2/n} \le e^{-dt^2}.$$

Hence the probability that ρ is entangled decays exponentially to 0 as $d \to \infty$. However, since ρ_1 and ρ_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 ρ is generically separable. □

Remark. The above discussion on concentration of measure techniques is an adaption of S. J. Szarek's lecture "Geometric Functional Analysis and QIT" [33] 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 are the Marcenko-Pastur law (resp. semi-circle law) under proper asymptotic regime. An interesting application is that we have proved the PPT square conjecture holds generically if we independently chose 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 [16, 22], we think that it remains natural to hope for 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 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: B.C. is supported by JSPS Kakenhi 17K18734, 15KK0162 and 17H04823, and ANR-14-CE25-0003. Z.Y. is supported by NSFC No. 11771106 and No. 11431011. P.Z. is partially supported by NSFC (No. 11501423, No. 11671308) and NSERC research grant. P.Z. also wants to thank IASM of the Harbin Institute of Technology for warm hospitality in summer 2016. The work was done when P.Z. 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 the fall 2017, that provided a fruitful environment to meet and work on the project of this paper.

References

[1] G. Aubrun, Partial transposition of random states and non-centered semicircular distributions, *Random Matrices: Theory Appl.* **1**, 1250001 (2012).

[2] G. Aubrun and S. J. Szarek, *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*, Mathematical Surveys and Monographs Volume: 223, American Mathematical Society, (2017).

[3] G. Aubrun, S. J. Szarek, E. Werner, Hastingss additivity counterexample via Dvoretzky's theorem, *Comm. Math. Phys.* **305** (1), 85-97 (2011).

[4] G. Aubrun, S. J. Szarek, and D. P. Ye, Entanglement thresholds for random induced states, *Comm. Pure Appl. Math.* **67** (1), 129-171 (2014).
[5] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.* **70** 1895-1899 (1993).
[6] Z. D. Bai and Y. Q. Yin, Convergence to the Semicircle Law, *Ann. Probab.* **16** 863-875 (1988).
[7] S. Bäuml, *On Bound Key And The Use Of Bound Entanglement*, Diploma thesis, Ludwig Maximilians Universität München, (2010).
[8] S. Bäuml, M. Christandl, K. Horodecki, A. Winter, Limitations on quantum key repeaters, *Nat. Comm.* **6**, 6908 (2015).
[9] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.* **70**, 1895 (1993).
[10] P. Biane, Some properties of crossings and partitions, *Discrete Mathematics*, **175** (1), 41-53 (1997).
[11] F. Brandao, M. Horodecki, On Hastings' counterexamples to the minimum output entropy additivity conjecture, *Open Syst. Inf. Dyn.* **17**, 31 (2010).
[12] D. Bouwmeester, J. W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, Experimental quantum teleportation , *Nat.* **390** (6660), 575-579 (1997).
[13] M. D. Choi, Completely positive linear maps on complex matrices, *Linear Alg. Appl.* **10**, 285 (1975).
[14] B. Collins, Product of random projections, Jacobi ensembles and universality problems arising from free probability, *Probab. Theory Related Fields* **133** (3), 315-344 (2005).
[15] B. Collins and I. Nechita, Random quantum channels I: Graphical calculus and the Bell state phenomenon, *Comm. Math. Phys.* **297**, 345-370 (2010).
[16] B. Collins and I. Nechita, Random quantum channels II: Entanglement of random subspaces, Rényi entropy estimates and additivity problems, *Adv. Math.* **226**, 1181-1201 (2011).
[17] B. Collins and I. Nechita, Gaussianization and eigenvalue statistics for random quantum channels (III), *Ann. Appl. Prob.* **21** No. 3, 1136-1179 (2011).
[18] B. Collins, I. Nechita, and D. P. Ye, The absolute positive partial transpose property for random induced states, *Random Matrices: Theory Appl.* **01** 1250002 (2012).
[19] M. Christandl, A. Müller-Hermes, M. M. Wolf, When do composed maps become entanglement breaking? *arXiv.* 1807.01266.
[20] M. Fukuda and I. Nechita, Additivity rates and PPT property for random quantum channels, *Annales mathématiques Blaise Pascal.* **22** (1), 1-72 (2015).
[21] A. Grudka, M. Horodecki and Ł. Pankowski, 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 (2010).
[22] M. B. Hastings, Superadditivity of communication capacity using entangled inputs, *Nat. Phys.* **5**, 255 (2009).
[23] P. Hayden and A. Winter, Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$, *Comm. Math. Phys.* **284**, 263 (2008).
[24] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-state entanglement and distillation: Is there a bound entanglement in nature, *Phys. Rev. Lett.* **80**, 5239 (1998).
[25] R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81** (2), 865-942 (2007).
[26] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, *Rep. Math. Phys.* **3**, 275 (1972).
[27] M. Kennedy, N. A. Manor, and V. I. Paulsen, Composition OF PPT Maps, arXiv:1710.08475.
[28] I. Nechita, Asymptotics of random density matrices, *Ann. Henri Poincare.* **8**, 1521-1538 (2007).
[29] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, LMS Lecture Note Series, Vol 335, Cambridge University Press, (2006).

[30] M. Rahaman, S. Jaques, and V. I. Paulsen, Eventually Entanglement Breaking Maps, arXiv:1801.05542.
[31] M. Ruskai, M. Junge, D. Kribs, P. Hayden, and A. Winter, *BanffInternational Research Station workshop: Operator Structures in Quantum Information Theory* (2012). Available at https://www.birs.ca/workshops/2012/12w5084/report12w5084.pdf.
[32] H-J Sommers and K. Zyczkowski, Bures volume of the set of mixed quantum states, *J. Phys. A: Math. Theor.* **37** (35), 8457 (2003).
[33] Szarek's lecturenotes for a talk given in the trimester "Analysis in Quantum Information Theory" which is based on the book [2] and the paper [4].
[34] K. Zyczkowski and H-J Sommers, Induced measures in the space of mixed quantum states, *J. Phys. A: Math. Theor.* **34** (35), 7111-25 (2001).
[35] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, "Event-ready-detectors" Bell experiment via entanglement swapping, *Phys. Rev. Lett.* **71** (26), 4287-4290 (1993).

DEPARTMENT OF MATHEMATICS, GRADUATE SCHOOL OF SCIENCE, KYOTO UNIVERSITY, KYOTO 606-8502, JAPAN
E-mail address: collins@math.kyoto-u.ac.jp

INSTITUTE OF ADVANCED STUDY IN MATHEMATICS, HARBIN INSTITUTE OF TECHNOLOGY, HARBIN 150006, CHINA
E-mail address: hustyinzhi@163.com

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF WYOMING, DEPT. 3036, 1000 E. UNIVERSITY AVE., LARAMIE, WY 82071-3036, USA
E-mail address: pzhong@uwyo.edu