Phase transitions for random states and a semicircle law for the partial transpose

Guillaume Aubrun,[1,*] Stanisław J. Szarek,[2,3,†] and Deping Ye[4,‡]

[1]*Institut Camille Jordan, Université Claude Bernard Lyon 1, 69622 Villeurbanne CEDEX, France*
[2]*Case Western Reserve University, Cleveland, Ohio 44106-7058, USA*
[3]*Institut de Mathématiques de Jussieu, Université Pierre et Marie Curie, 75005 Paris, France*
[4]*Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, Canada A1C 5S7*

For a system of N identical particles in a random pure state, there is a threshold $k_0 = k_0(N) \sim N/5$ such that two subsystems of k particles each typically share entanglement if $k > k_0$, and typically do not share entanglement if $k < k_0$. By "random" we mean here uniformly distributed on the sphere of the corresponding Hilbert space. The analogous phase transition for the positive partial transpose (PPT) property can be described even more precisely. For example, for N qubits the two subsystems of size k are typically in a PPT state if $k < k_1 := N/4 - 1/2$ and typically in a non-PPT state if $k > k_1$. Since, for a given state of the entire system, the induced state of a subsystem is given by the partial trace, the above facts can be rephrased as properties of random induced states. An important step in the analysis depends on identifying the asymptotic spectral density of the partial transposes of such random induced states, a result which is interesting in its own right.

I. INTRODUCTION

If all that we know about a quantum system is its dimension n (the number of levels) and that it is well isolated from the environment, a reasonable model—or at least a reasonable first guess—for the state of the system is a unit vector selected at random from the sphere of an n-dimensional complex Hilbert space $\mathcal{H}$. If the system interacts with some part of the environment, represented by an ancilla space $\mathcal{H}_a$, the quantum formalism suggests as a model the so-called (random) *induced state*, obtained after partial tracing, over $\mathcal{H}_a$, a random pure state on the space $\mathcal{H} \otimes \mathcal{H}_a$. The same description applies if we are primarily interested in a *subsystem* of an isolated system, the setup that is addressed in the abstract.

The above is just one example of how a random paradigm arises naturally in the quantum context. In the last few years probabilistic considerations have become a very fruitful approach in quantum information theory, the highlights being the fundamental paper [1] by Hayden, Leung, and Winter and, more recently, Hastings's proof that suitably chosen random channels provide a counterexample to the additivity conjecture for classical capacity of quantum channels [2].

Although random states have been considered for many years, their properties (e.g., *are they typically entangled?*) remained elusive. In this Rapid Communication we describe a reasonably general way to handle such questions. Of course, the induced state ρ being random, we cannot expect to be able to tell what ρ is. However, we may be able to infer some properties of ρ if they are *generic* (that is, occur with probability close to 1). As it turns out, for many natural properties, a phenomenon of phase transition takes place (at least when $\dim \mathcal{H}$ is sufficiently large): the generic behavior of ρ "flips" to the opposite one when $s := \dim \mathcal{H}_a$ changes

from being a little smaller than certain threshold dimension s_0 to being larger than s_0.

For simplicity, we will focus on the *random induced states* mentioned at the beginning of the Introduction. This leads (see Refs. [3,4]) to a natural family of probability measures on $\mathcal{D}(\mathcal{H})$, the set of states on $\mathcal{H}$, where s, the dimension of the ancilla space, is a parameter. For specificity, consider $\mathcal{H} = \mathbf{C}^d \otimes \mathbf{C}^d$ and let us concentrate on two properties: *entanglement* and *positive partial transpose* (PPT). This choice is based, first, on the importance of these concepts and, second, on the differences in their respective mathematical features, which allow to present the diverse techniques needed to handle the problems.

Concerning the importance aspect, we note that detecting and exploiting entanglement—originally discovered in the 1930s [5]—is a central problem in quantum information and quantum computation, at least since Shor's work [6] on integer factoring. Next, the positive partial transpose is the simplest test for entanglement (Peres-Horodecki PPT criterion; see Refs. [7,8]) and is at the center of the important *distillability conjecture* [9], a positive answer to which would give a physical and operational meaning to the PPT property. On the other hand, from the computational complexity point of view, verifying the PPT property is easy (just check whether the partial transpose of the state ρ is positive semidefinite), while deciding whether ρ is entangled is a computationally intractable problem (NP-hard) [10].

In the special case when $n := \dim \mathcal{H}$ equals $s = \dim \mathcal{H}_a$, partial tracing over the ancilla space $\mathcal{H}_a$ leads to the uniform distribution on $\mathcal{D}(\mathbf{C}^n)$ (i.e., uniform with respect to the Lebesgue measure, or Hilbert-Schmidt volume, denoted by vol). More generally, when $s \geqslant n$, the corresponding probability measure $\mu_{n,s}$ has a density with respect to the Lebesgue measure on $\mathcal{D}(\mathbf{C}^n)$, which has a simple form [3]

$$\frac{d\mu_{n,s}}{d\text{vol}}(\rho) = \frac{1}{Z_{n,s}}(\det \rho)^{s-n}, \tag{1}$$

*aubrun@math.univ-lyon1.fr
†szarek@cwru.edu
‡deping.ye@mun.ca

where $Z_{n,s}$ is a normalization factor. Note that Eq. (1) defines the measure $\mu_{n,s}$ (in particular) for every real $s \geqslant n$, while the partial trace construction makes sense only for integer values of s. If $s < n$, the measure $\mu_{n,s}$ is concentrated on the boundary of $\mathcal{D}(\mathbf{C}^n)$, but still can be described analytically. Another way to implement these measures is to start from the *complex* Wishart-Laguerre matrices $W_{n,s}$ ($n \times n$, with s degrees of freedom) [11], a classical ensemble in statistics and mathematical physics, and to normalize them to have trace 1.

In spite of the explicitness of the formula (1), it is not easy to find—even approximately, and even for $s = n = d^2$—the probability that a random induced state has PPT or is entangled. This is because these traits are not encoded in a simple way in the spectral properties of ρ. It was shown in Ref. [12]—via methods of high-dimensional probability—that the proportion of states (measured in the sense of $\mu_{n,n}$, i.e., the Legesgue measure) that are unentangled, or *separable*, is extremely small in large dimensions. This was extended to the case when $s = \dim \mathcal{H}_a$ is slightly larger than $n = d^2$ in [13], while, on the other hand, it was proved in Ref. [1] that random induced states on $\mathbf{C}^d \otimes \mathbf{C}^d$ are typically separable when s is proportional to $n^2 = d^4$. The paper [12] also established that unentangled states are extremely rare even among PPT states (again, when $s = n = d^2$). However, even such a simple question as *Does the proportion of the PPT states among all states go to* 0 *as the dimension increases?*, originally asked in Ref. [14], has not been rigorously addressed prior to the work that we describe in this Rapid Communication. The results we summarize go a long way in filling the gaps in understanding the phenomena in question (see Refs. [15,16] for details and references). We show that the threshold between entanglement and separability occurs when s is roughly of order $n^{3/2} = d^3$, and that the threshold between NPT (i.e., non-PPT) and PPT is when $s \sim 4n = 4d^2$.

The heuristics behind the consequences stated in the Abstract is now as follows. If we have a system of N particles (with D levels each) which is in a random pure state, and two subsystems of k particles each, then the "joint state" of the subsystems is modeled by a random induced state on $\mathbf{C}^d \otimes \mathbf{C}^d$ with $d = D^k$ and $s = D^{N-2k}$. In particular, the relation $k = N/5$, or $N = 5k$, corresponds exactly to $s = d^3$. A similar argument applies to the PPT property.

Another consequence of the results is that, for a large range of parameters, when the ancilla dimension s is roughly between $4d^2$ and d^3, a generic random state is both PPT and entangled. Such states are *bound entangled*, or *undistillable* [9] and, in spite of being entangled, are useless for purposes such as teleportation or superdense coding (cf. Ref. [17]). However, since for small systems ($2 \otimes 2$ and $2 \otimes 3$; see Refs. [18,19]) PPT is equivalent to separability, one is tempted to think that bound entangled states are an anomaly, and that the PPT property remains a good proxy for separability in higher dimensions. Our results imply that this heuristic becomes misleading for large systems and that PPT and separability are quantitatively very different properties.

While, as we postulated, random induced states form the most natural family of probability measures on $\mathcal{D}(\mathbf{C}^n)$, our methods are fairly robust and allow handling of other random models. For example, another popular way to construct random states is to consider mixtures of pure states. Our analysis

applies to this model as well: if $\nu_{n,s}$ is the distribution of $\frac{1}{s}\sum_{i=1}^{s} |\psi_i\rangle\langle\psi_i|$, where (ψ_i) are independent uniform pure states, then all the results presented for the measures $\mu_{n,s}$ remain valid *mutatis mutandis* for $\nu_{n,s}$.

II. RESULTS

We recapitulate the setting: $n = \dim \mathcal{H}$, ψ is a (random) unit vector uniformly distributed on the sphere of $\mathcal{H} \otimes \mathbf{C}^s$, and $\rho = \mathrm{tr}_{\mathbf{C}^s} |\psi\rangle\langle\psi|$ is a random state on $\mathcal{H}$ whose distribution is denoted by $\mu_{n,s}$. Further, we assume that $n = d^2 > 1$ and $\mathcal{H} = \mathbf{C}^d \otimes \mathbf{C}^d$; states on $\mathcal{H}$ will be considered entangled, PPT, etc., with respect to this particular splitting. For definiteness, the partial transpose Γ will be the transposition in the second factor, i.e., defined (by its action on product states) via $(\tau_1 \otimes \tau_2)^\Gamma = \tau_1 \otimes \tau_2^T$. The first result describes the phase transition between generic entanglement and generic separability.

Theorem 1 [16]. There exist effectively calculable constants $C, c > 0$ and a threshold function $s_0 = s_0(d)$ satisfying

$$cd^3 \leqslant s_0(d) \leqslant Cd^3 \ln^2 d, \qquad (2)$$

such that if ρ is a random state on $\mathbf{C}^d \otimes \mathbf{C}^d$ distributed according to the measure $\mu_{d^2,s}$ and if $\varepsilon > 0$, then
  (i) for $s \leqslant (1 - \varepsilon)s_0(d)$ we have
    $\mathbf{P}(\rho \text{ is separable}) \leqslant 2\exp[-c(\varepsilon)d^3]$ and
  (ii) for $s \geqslant (1 + \varepsilon)s_0(d)$ we have
    $\mathbf{P}(\rho \text{ is entangled}) \leqslant 2\exp[-c(\varepsilon)s]$,
where $c(\varepsilon) > 0$ depends only on ε.   ∎

Let us mention that our methods extend also to the multipartite setting and to "unbalanced" systems such as $\mathbf{C}^{d_1} \otimes \mathbf{C}^{d_2}$, $d_1 \neq d_2$–see Ref. [16] for precise statements.

The idea behind the proof of Theorem 1, based on tools from high-dimensional convexity, is quite general and can be used to estimate thresholds for other properties of random induced states (beyond separability), provided the set of states with given property is a convex subset $\mathcal{K} \subset \mathcal{D}(\mathcal{H})$ and has some minimal invariance properties. However, in the special case of PPT we have a more precise result.

Theorem 2 [15]. Let ρ be a random state on $\mathbf{C}^d \otimes \mathbf{C}^d$ distributed according to $\mu_{d^2,s}$. Set $s_1(d) = 4d^2$ and let $\varepsilon > 0$. Then
  (i) for $s \leqslant (1 - \varepsilon)s_1(d)$ we have
    $\mathbf{P}(\rho \text{ is PPT}) \leqslant 2\exp[-c(\varepsilon)d^2]$ and
  (ii) for $s \geqslant (1 + \varepsilon)s_1(d)$ we have
    $\mathbf{P}(\rho \text{ is non-PPT}) \leqslant 2\exp[-c(\varepsilon)s]$,
where $c(\varepsilon) > 0$ depends only on ε.   ∎

As noted in Ref. [15], it is likely that the sharp estimate in (i) is of order $\exp(-c(\varepsilon)d^4)$; this conjecture leads to interesting large deviation problems for matrices ρ^Γ.

The proof of Theorem 2 (except for the exponential estimates on the probabilities, which require a unified approach common to both theorems) depends on methods of random matrix theory and, specifically, on the following result that identifies *asymptotic spectral density* of the partial transpose of random induced states, and which is of independent interest.

If A is a Hermitian matrix, we will denote by $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ the largest and the smallest eigenvalues of A. If $a \in \mathbf{R}$ and $\sigma > 0$, the *semicircular distribution* $\mu_{\mathrm{SC}(a,\sigma^2)}$ is the probability measure with support $[a - 2\sigma, a + 2\sigma]$
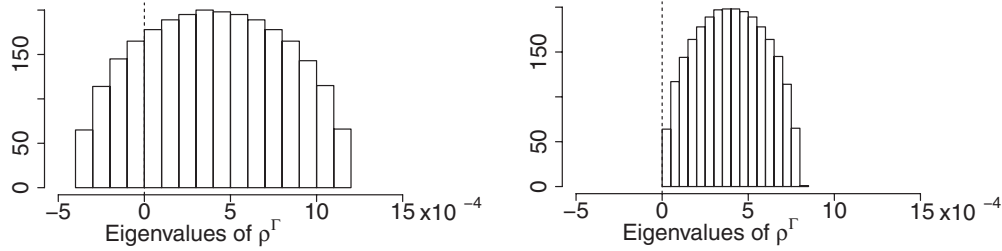
FIG. 1. Histogram showing distribution of the eigenvalues of ρ^Γ, where ρ is a random state on $\mathbf{C}^{50} \otimes \mathbf{C}^{50}$ chosen according to the distribution $\mu_{2500,2500}$ ($\alpha = 1$, left) or $\mu_{2500,10000}$ ($\alpha = 4$, right). In both cases the median eigenvalue is about $\frac{1}{2500} = 4 \times 10^{-4}$.

and density $(2\pi\sigma^2)^{-1}\sqrt{4\sigma^2 - (x-a)^2}$. We then have the following theorem.

Theorem 3 [15]. Given $\alpha > 0$, let ρ_d be a random mixed state on $\mathbf{C}^d \otimes \mathbf{C}^d$ distributed according to $\mu_{d^2, \lfloor \alpha d^2 \rfloor}$. Then, as d tends to $+\infty$, the eigenvalue distributions of ρ_d^Γ approaches the deterministic measure $\mu_{\mathrm{SC}(1,1/\alpha)}$ in the following sense: for any interval $I \subset \mathbf{R}$, the proportion of eigenvalues of ρ_d^Γ inside the rescaled interval $\frac{1}{d^2}I$ converges (in probability) toward $\mu_{\mathrm{SC}(1,1/\alpha)}(I)$.

Moreover, we also have convergence of the extreme eigenvalues $\lambda_{\max}(d^2\rho_d^\Gamma)$ and $\lambda_{\min}(d^2\rho_d^\Gamma)$ to respectively $1 + 2/\sqrt{\alpha}$ and $1 - 2/\sqrt{\alpha}$, the endpoints of the support of $\mu_{\mathrm{SC}(1,1/\alpha)}$. ∎

It is easy to numerically "check" the conclusion of Theorem 3 (this was first noticed in Ref. [20]). For example, Fig. 1 shows sample distributions of eigenvalues of a partially transposed random state on $\mathbf{C}^{50} \otimes \mathbf{C}^{50}$, when $\alpha = 1$ and $\alpha = 4$ (sample size 1 in each case).

Because of the link between random induced states and the Wishart ensemble $W_{n,s}$, Theorem 3 holds also for that ensemble (real or complex, although it is the complex setting that is most relevant to the quantum theory); in that case the rescaling factor d^2 is not needed. We emphasize that this is rather unexpected since the asymptotic spectral density of the Wishart ensemble *itself* is given by the Marchenko-Pastur distribution [21].

III. MATHEMATICS BEHIND THE RESULTS

Although Theorems 2 and 3 have similar statements, the tools used in their proofs are very different, which parallels the differences in the computational complexity of PPT vs that of entanglement. However, combining all the tools is often necessary to obtain the full strength of the results.

We first describe the proof of Theorem 1, which is of geometric nature and where the concept of *mean width* plays a central role. To present it, let us introduce basic concepts associated to a convex body $K \subset \mathbf{R}^m$ containing the origin in the interior (see Ref. [22] for more background). The *gauge* of K is the function $\|\cdot\|_K$ defined for $x \in \mathbf{R}^m$ by

$$\|x\|_K = \inf\{t \geqslant 0 : x \in tK\}.$$

The *polar* (or dual) body of K is defined as

$$K^\circ = \{y \in \mathbf{R}^m : \langle x, y \rangle \leqslant 1 \ \forall x \in K\}.$$

If u is a vector from the unit sphere S^{m-1}, the *support function* of K in the direction u is $h_K(u) := \max_{x \in K} \langle x, u \rangle = $

$\|u\|_{K^\circ}$. Note that $h_K(u) + h_K(-u)$ is the distance between the two hyperplanes tangent to K and normal to u. The mean width of K is then defined as

$$w(K) := \int_{S^{m-1}} h_K(u)\, d\sigma(u) = \int_{S^{m-1}} \|u\|_{K^\circ} d\sigma(u),$$

where $d\sigma$ is the normalized spherical measure on S^{m-1}.

In our setting, the relevant convex body is $K = \mathcal{S}^\circ$, where $\mathcal{S}$ is the set of mixed separable states on $\mathcal{H} = \mathbf{C}^d \otimes \mathbf{C}^d$. The ambient space $\mathbf{R}^m$ is the space of self-adjoint trace 1 operators on $\mathcal{H}$ (hence $m = d^4 - 1$), where the maximally mixed state plays the role of the origin. Since $K^\circ = (\mathcal{S}^\circ)^\circ = \mathcal{S}$ (the bipolar theorem) and since separability of ρ is equivalent to $\|\rho\|_{\mathcal{S}} \leqslant 1$, the crucial question is whether $w(\mathcal{S}^\circ)$ is smaller or larger than 1. An analysis of this question leads to the following value of the threshold function appearing in Theorem 1:

$$s_0(d) = w(\mathcal{S}^\circ)^2.$$

Assertions (i) and (ii) can then be derived from concentration of measure, and the heart of the proof is showing Eq. (2), especially the upper bound.

Determining the threshold $s_0(d)$ requires finding the typical value of the gauge associated to $\mathcal{S}$, computing which, as we mentioned, is a hard problem. We take an indirect route and find the order of magnitude of the threshold using the machinery of high-dimensional geometry, especially the so-called MM^*-estimate.

The MM^*-estimate (see Refs. [22,23]) is a general theorem which relates the mean width of a convex body and the mean width of its polar. While the abstract formulation may require an affine change of coordinates, in the present situation, because of the symmetries of $\mathcal{S}$ (invariance under local unitary conjugations), we can deduce via simple representation theory the inequalities

$$1 \leqslant w(\mathcal{S})w(\mathcal{S}^\circ) \leqslant C \ln d,$$

where $C > 0$ is a universal constant. Since $w(\mathcal{S})$ can be estimated by standard techniques of high-dimensional probability [12], this allows to establish the order of magnitude of $w(\mathcal{S}^\circ)$ [and hence of $s_0(d)$] up to polylog factors.

As indicated earlier, the same scheme yields estimates for the thresholds corresponding to other properties including the PPT, but it *does not* allow us to recover the precise order $4d^2$ appearing in Theorem 2. However, the latter result (except for quantitative estimates on the probabilities, which require further work, again based on the concentration of measure) follows readily from Theorem 3, which describes

very precisely the spectrum of the partial transpose of a random induced state: the PPT condition is equivalent to $\lambda_{\min}(\rho_d^\Gamma) \geqslant 0$, which is generic if $1 - 2/\sqrt{\alpha} > 0$; similarly, $\lambda_{\min}(\rho_d^\Gamma) < 0$ is generic if $1 - 2/\sqrt{\alpha} < 0$—hence $\alpha = \frac{s}{d^2} = 4$ is the critical value. In turn, to show Theorem 3 we use the moment method, a standard technique from random matrix theory. The idea is to identify the asymptotic spectral density by computing its moments. This leads to problems in asymptotic combinatorics: the moments of semicircular distributions are given by the Catalan numbers, corresponding to the dominant combinatorial terms, while the statements about convergence of extreme eigenvalues are proved by refining the calculations and carefully estimating contributions of lower-order combinatorial terms.

IV. CONCLUSIONS

We established that random induced states on $\mathcal{H} = \mathbf{C}^d \otimes \mathbf{C}^d$ exhibit a phase-transition phenomenon with respect to the dimension s of the ancilla space. We exemplified the phenomenon on two properties: positive partial transpose, for which the threshold value of s is $4d^2$, and entanglement, for which the threshold is d^3 (up to a polylog factor). This allows us to determine whether two subsystems of an isolated system typically share (or typically do not share) entanglement when knowing only the sizes of those subsystems, and similarly for the PPT property. In fact, we provide a "black box" approach, which applies to many natural properties of quantum states. Our results motivate further study of the geometry of sets of quantum states, and that of large deviation behavior of some random matrix ensembles related to quantum information theory.

We expect the probabilistic methods to continue to play a major role in quantum theory. Indeed, the latter field usually involves high-dimensional objects; for example, the quantum analog of a byte [a state on $(\mathbf{C}^2)^{\otimes 8}$–a qubyte, one may say] "lives" in a space of dimension $2^{16} - 1 = 65\,535$. While this makes numerical schemes mostly impractical (the well-known *curse of dimensionality*), randomness is boosted by the presence of many free parameters (one may call this phenomenon *the blessing of dimensionality*). The current level of understanding of these aspects of the theory is arguably comparable to that of combinatorics in the 1950s, when the power of the probabilistic method [24] began to be appreciated and, subsequently, the study of random graphs became an intensive area of research.

[1] P. Hayden, D. Leung, and A. Winter, Commun. Math. Phys. **265**, 95 (2006).

[2] M. B. Hastings, Nat. Phys. **5**, 255 (2009).

[3] K. Życzkowski and H.-J. Sommers, J. Phys. A **34**, 7111 (2001).

[4] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, England, 2006).

[5] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[6] P. W. Shor, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, edited by S. Goldwasser (IEEE Comput. Soc. Press, Los Alamitos, CA, 1994), pp. 124-134.

[7] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[8] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[9] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).

[10] L. Gurvits, in *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, (electronic) (ACM, New York, 2003), pp. 10–19.

[11] J. Wishart, Biometrika **20A**, 32 (1928).

[12] G. Aubrun and S. J. Szarek, Phys. Rev. A **73**, 022109 (2006).

[13] D. Ye, J. Phys. A **43**, 315301 (2010).

[14] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).

[15] G. Aubrun, Random Matrices: Theory and Applications (to appear), e-print arXiv:1011.0275.

[16] G. Aubrun, S. Szarek, and D. Ye, e-print arXiv:1106.2264.

[17] P. W. Shor, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **86**, 2681 (2001).

[18] E. Størmer, Acta Math. **110**, 233 (1963).

[19] S. L. Woronowicz, Rep. Math. Phys. **10**, 165 (1976).

[20] M. Žnidarič, T. Prosen, G. Benenti, and G. Casati, J. Phys. A **40**, 13787 (2007).

[21] V. A. Marchenko and L. A. Pastur, Mat. Sb. (N.S.) **72**(114), 507 (1967).

[22] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge Tracts in Mathematics No. 94 (Cambridge University Press, Cambridge, England, 1989).

[23] T. Figiel and N. Tomczak-Jaegermann, Israel J. Math. **33**, 155 (1979).

[24] P. Erdős and A. Rényi, Publ. Math. Debrecen **6**, 290 (1959).