CHAPTER 10

Random Quantum States

The main goal of this chapter is to prove the following result. Consider a system of N identical particles (e.g., N qubits) in a random pure state. For some $k \leqslant N/2$, let A and B be two subsystems, each consisting of k particles. There exists a threshold function $k_0(N)$ which satisfies $k_0(N) \sim N/5$ as $N \to \infty$ and such that the following holds. *If $k < k_0(N)$, then with high probability the two subsystems A and B share entanglement. Conversely, if $k > k_0(N)$, then with high probability the two subsystems A and B do not share entanglement.*

If the Hilbert space associated to a single particle is $\mathbb{C}^q$ (e.g., $q = 2$ for qubits), the dimension of the system $A \otimes B$ equals q^{2k} and the state ρ describing the $A \otimes B$ subsystem is obtained as a partial trace over an environment of dimension q^{N-2k} (the remaining $N-2k$ particles). If the global system is in a random and uniformly distributed pure state, the state ρ is a random induced state as introduced in Section 6.2.3.4, where its distribution was denoted by $\mu_{q^{2k},q^{N-2k}}$. The central result of the chapter (Theorem 10.12) answers the question whether a random induced state on $\mathbb{C}^d \otimes \mathbb{C}^d$ with distribution $\mu_{d^2,s}$ is separable or entangled. It relies on the volume and mean width estimates from Chapter 9.

Section 10.3 contains results about other thresholds for random induced states: for the PPT vs. non-PPT dichotomy (Theorem 10.17) and for the value of the entanglement of formation being close to maximal or close to minimal (Theorem 10.16).

10.1. Miscellaneous tools

The first sections of this chapter contain an intermediate result (a quantitative central limit theorem) about approximation of random induced states by Gaussian matrices (Proposition 10.6). As a tool, we present some majorization inequalities in Section 10.1.1.

10.1.1. Majorization inequalities. Majorization was introduced in Section 1.3.1. We first state a technical result that ascertains that "flat" vectors (i.e., vectors with a large ℓ_1-norm and small ℓ_∞-norm) majorize many other vectors. Since we need to consider homotheties, it is natural to work in $\mathbb{R}^{n,0}$, the hyperplane of $\mathbb{R}^n$ consisting of vectors whose coordinates add up to 0.

LEMMA 10.1. *Let $x, y \in \mathbb{R}^{n,0}$. Assume that $\|y\|_\infty \leqslant 1$ and $\|y\|_1 \geqslant \alpha n$ for some $\alpha \in (0, 1]$. Then*

$$x \prec (2/\alpha - 1)\|x\|_\infty\, y. \tag{10.1}$$

PROOF OF LEMMA 10.1. By homogeneity, it is enough to verify that the condition $\|x\|_\infty \leqslant 1$ implies $x \prec (2/\alpha - 1)y$. Moreover, it is enough to check this for

x being an extreme point of the set $A \coloneqq \{x \in \mathbb{R}^{n,0} \ : \ \|x\|_\infty \leqslant 1\}$, since the set $\{x \in \mathbb{R}^{n,0} \ : \ x \prec z\}$ is convex for any $z \in \mathbb{R}^{n,0}$.

Extreme points of A are of the following form: $\lfloor n/2 \rfloor$ coordinates are equal to 1 and $\lfloor n/2 \rfloor$ coordinates equal to -1. In the case of odd n there is one remaining coordinate, which is necessarily equal to 0. It is thus enough to verify that if x is of that form, and if y satisfies $\|y\|_\infty \leqslant 1$ and $\|y\|_1 = \alpha n$, then $x \prec (2/\alpha - 1)y$. This is shown by establishing that an average of permutations of y is a multiple of x.

First, average separately the positive and the negative coordinates of y to obtain a vector y' whose coordinates take only two values, one positive and one negative. Since the ℓ_1-norm of the positive and the negative part of y' is equal and amounts to $\alpha n/2$, the support of each part must be at least $\alpha n/2$ and at most $(1 - \alpha/2)n$, and the absolute value of each coordinate at least $\alpha/(2 - \alpha)$.

Assume now that n is even. Next, select a set of $n/2$ equal coordinates (positive or negative, depending on which part has larger support) and average the remaining ones. The obtained vector is a multiple of an extreme point, as needed. If n is odd, select $\lfloor n/2 \rfloor$ equal coordinates (from the dominant sign) and average the remaining ones to produce one zero and $\lfloor n/2 \rfloor$ equal coordinates. The resulting vector is also a multiple of an extreme point. □

A simpler but less precise version of Lemma 10.1 can be obtained without any hypothesis on $\|y\|_\infty$.

LEMMA 10.2. *Let $x, y \in \mathbb{R}^{n,0}$ with $y \neq 0$. Then*

$$x \prec \frac{2n\|x\|_\infty}{\|y\|_1} y. \tag{10.2}$$

PROOF. By homogeneity, we may assume that $\|y\|_\infty = 1$ and the result follows from Lemma 10.1. □

As a consequence, we obtain the fact that if two vectors from $\mathbb{R}^{n,0}$ are flat and close to each other, one is majorized by a small perturbation of the other one.

PROPOSITION 10.3. *Let $x, y \in \mathbb{R}^{n,0}$. Assume that $\|x - y\|_\infty \leqslant \varepsilon$ and $\|y\|_1 \geqslant \alpha n$ for some $\alpha > 0$. Then*

$$x \prec \left(1 + \frac{2\varepsilon}{\alpha}\right) y.$$

PROOF. We use the following elementary property of majorization: if $x_1 \prec \lambda_1 y$ and $x_2 \prec \lambda_2 y$ for some positive λ_1, λ_2, then $x_1 + x_2 \prec (\lambda_1 + \lambda_2)y$. We apply this fact with $x_1 = y$, $\lambda_1 = 1$ and $x_2 = x - y$. Lemma 10.2 shows that we can choose $\lambda_2 = 2\varepsilon/\alpha$, and the Proposition follows. □

EXERCISE 10.1. Provide an alternative proof of Lemma 10.2 by using directly the definition of majorization.

10.1.2. Spectra and norms of unitarily invariant random matrices. A lot of information about a self-adjoint matrix can be retrieved from its spectrum; for example, all unitarily invariant norms can be computed if one knows the eigenvalues (see Section 1.3.2). In contrast, computing the values of other norms or gauges (e.g., the gauge associated to the set of separable states) usually requires some knowledge about the eigenvectors.

However, if the matrix is random and if its distribution is unitarily invariant, it is possible to circumvent this difficulty. Heuristically, the principle we are going

to establish and use is as follows: *if A and B are two unitarily invariant random matrices with similar spectra, then, for any norm or gauge $\|\cdot\|$, the typical values of $\|A\|$ and of $\|B\|$ are comparable.*

It is convenient to work in the hyperplane $\mathsf{M}_n^{\mathrm{sa},0}$ of self-adjoint complex $n \times n$ matrices with trace zero. One says that a $\mathsf{M}_n^{\mathrm{sa},0}$-valued random variable A is *unitarily invariant* if, for any $U \in \mathsf{U}(n)$, the random matrices A and $UAU^\dagger$ have the same distribution. Recall also that μ_{SC} is the standard semicircular distribution, that $\mu_{\mathrm{sp}}(A)$ is the empirical spectral distribution of a self-adjoint matrix A, and that d_∞ denotes the ∞-Wasserstein distance. All these concepts were introduced in Section 6.2.

Proposition 10.4. *Let A and B be two $\mathsf{M}_n^{\mathrm{sa},0}$-valued random variables which are unitarily invariant and satisfy the following conditions*

$$\mathbf{P}(d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{SC}}) \leqslant \varepsilon) \geqslant 1 - p \quad \text{and} \quad \mathbf{E}\, d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{SC}}) \leqslant \varepsilon \tag{10.3}$$

for some $\varepsilon, p \in (0,1)$, and similarly for B. Then, for any convex body $K \subset \mathsf{M}_n^{\mathrm{sa},0}$ containing the origin in its interior,

$$\frac{1-p}{1+C\varepsilon}\, \mathbf{E}\, \|A\|_K \leqslant \mathbf{E}\, \|B\|_K \leqslant \frac{1+C\varepsilon}{1-p}\, \mathbf{E}\, \|A\|_K$$

for some absolute constant C.

Proof of Proposition 10.4. Note that possible relations between A and B (such as independence) are irrelevant in the present situation. Consider the following function on $\mathbb{R}^{n,0}$ (recall that $\mathbb{R}^{n,0}$ denotes the hyperplane of vectors of sum zero in $\mathbb{R}^n$)

$$\phi(x) = \mathbf{E}\, \|U \operatorname{Diag}(x) U^\dagger\|_K,$$

where $U \in \mathsf{U}(n)$ denotes a Haar-distributed random unitary matrix (independent of everything else) and $\operatorname{Diag}(x)$ is the diagonal matrix whose ii-th entry is x_i. Unitary invariance implies that

$$\mathbf{E}\, \|A\|_K = \mathbf{E}\, \phi(\operatorname{spec}(A)) \tag{10.4}$$

and similarly for B (see Exercise 10.2). Let E be the event $\{d_\infty(\mu_{\mathrm{sp}}(B), \mu_{\mathrm{SC}}) \leqslant \varepsilon\}$. Assume for the moment that E holds, we have then (see Exercise 6.25)

$$\begin{aligned} \|B\|_1 &= n \int |x|\, \mathrm{d}\mu_{\mathrm{sp}}(B)(x) \geqslant n \int_{-2}^{2} (|x| - \varepsilon)^+\, \mathrm{d}\mu_{\mathrm{SC}}(x) \\ &\geqslant n \int_{-2}^{2} (|x| - 1)^+\, \mathrm{d}\mu_{\mathrm{SC}}(x) = \alpha n, \end{aligned}$$

$\alpha \approx 0.16$ being a numerical constant. Applying Proposition 10.3 to the vectors $\operatorname{spec}(A)$ and $\operatorname{spec}(B)$, we conclude that (with $C = 2/\alpha$)

$$\operatorname{spec}(A) \prec (1 + C d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{sp}}(B)))\operatorname{spec}(B).$$

Since ϕ is convex and permutationally invariant, it follows that

$$\phi(\operatorname{spec}(A)) \leqslant (1 + C d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{sp}}(B)))\phi(\operatorname{spec}(B)).$$

Using the fact that $d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{sp}}(B)) \leqslant \varepsilon + d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{SC}})$ and taking expectation over A yields

$$\mathbf{E}\, \phi(\operatorname{spec}(A)) \leqslant (1 + 2C\varepsilon)\phi(\operatorname{spec}(B)).$$

Recall that the above inequality is true conditionally on E. Consequently,

$$\mathbf{E}\,\phi(\mathrm{spec}(B)) \geqslant \mathbf{E}\,\phi(\mathrm{spec}(B))\mathbf{1}_E \geqslant (1+2C\varepsilon)^{-1}\mathbf{P}(E)\,\mathbf{E}\,\phi(\mathrm{spec}(A)).$$

In view of (10.4) and since $\mathbf{P}(E) \geqslant 1-p$ by hypothesis, this shows that

$$\mathbf{E}\,\|A\|_K \leqslant \frac{1+2C\varepsilon}{1-p}\,\mathbf{E}\,\|B\|_K.$$

The other inequality follows by symmetry. □

If ε is large (2 or larger), the hypothesis $d_\infty(\mu_{\mathrm{sp}}(A), \mu_{\mathrm{SC}}) \leqslant \varepsilon$ does not prevent A from being identically zero. However, an isomorphic version of Proposition 10.4 can be similarly obtained under the hypothesis that the spectra of A and B are reasonably flat.

PROPOSITION 10.5 (see Exercise 10.3). *Let A and B be two $\mathsf{M}_n^{\mathrm{sa},0}$-valued random variables which are unitarily invariant. Assume that*

$$\mathbf{P}(\|A\|_1 \geqslant c_1 n) \geqslant 1-p \quad \text{and} \quad \mathbf{E}\,\|A\|_\infty \leqslant C_2, \tag{10.5}$$

and similarly for B. Then, for any convex body $K \subset \mathsf{M}_n^{\mathrm{sa},0}$ containing the origin in the interior,

$$C^{-1}\,\mathbf{E}\,\|A\|_K \leqslant \mathbf{E}\,\|B\|_K \leqslant C\,\mathbf{E}\,\|A\|_K$$

with $C = (1-p)^{-1}(2C_2/c_1)$.

EXERCISE 10.2 (Retrieving unitarily invariant distributions from the spectrum). Let A be a $\mathsf{M}_n^{\mathrm{sa},0}$-valued random variable which is unitarily invariant. Recall that $\mathrm{Diag}(\mathrm{spec}(A))$ is the diagonal matrix whose diagonal entries are the eigenvalues of A arranged in the non-increasing order. Let $U \in \mathsf{U}(n)$ be a Haar-distributed random unitary matrix independent of A. Show that the random matrix $U\,\mathrm{Diag}(\mathrm{spec}(A))U^\dagger$ has the same distribution as A.

EXERCISE 10.3 (All flat unitarily invariant distributions look alike). Prove Proposition 10.5.

10.1.3. Gaussian approximation to induced states. We are going to investigate typical properties of random induced states, in the large dimension regime. Their spectral properties where discussed in Section 6.2.3, and are described either by the Marčenko–Pastur distribution (when s is proportional to n) or by the semi-circular distribution (when $s \gg n$).

However, we are also interested in properties that cannot be inferred from the spectrum (the main example being separability vs. entanglement on a bipartite system). In this context, it is useful to compare induced states with their Gaussian approximation. Indeed, the Gaussian model allows to connect with tools from convex geometry, such as the mean width.

It is convenient to work in the hyperplane $\mathsf{M}_n^{\mathrm{sa},0}$ and to consider the shifted operators $\rho - \mathrm{I}/n$, which we compare with a GUE_0 random matrix (see Section 6.2.2). The following proposition compares the expected value of any norm (or gauge) computed for both models.

PROPOSITION 10.6. *Given integers n, s, denote by $\rho_{n,s}$ a random induced state on $\mathbb{C}^n$ with distribution $\mu_{n,s}$, and by G_n an $n \times n$ GUE_0 random matrix. Let $C_{n,s}$ be*

the smallest constant such that the following holds: for any convex body $K \subset \mathsf{M}_n^{\mathrm{sa},0}$ containing 0 in the interior,

$$C_{n,s}^{-1}\,\mathbf{E}\left\|\frac{G_n}{n\sqrt{s}}\right\|_K \leqslant \mathbf{E}\left\|\rho_{n,s}-\frac{\mathrm{I}}{n}\right\|_K \leqslant C_{n,s}\,\mathbf{E}\left\|\frac{G_n}{n\sqrt{s}}\right\|_K. \tag{10.6}$$

Then
(i) For any sequences (n_k) and (s_k) such that $\lim_{k\to\infty} n_k = \lim_{k\to\infty} s_k/n_k = \infty$, we have $\lim_{k\to\infty} C_{n_k,s_k} = 1$.
(ii) For any $a > 0$, we have $\sup\{C_{n,s} \ : \ s \geqslant an\} < \infty$.

Remark 10.7. We emphasize that the quantity $\mathbf{E}\,\|G_n\|_K$ appearing in (10.6) is exactly the Gaussian mean width of the polar set K°. Indeed, the standard Gaussian vector in the space $\mathsf{M}_n^{\mathrm{sa},0}$ (equipped with the Hilbert–Schmidt scalar product, as always) is exactly a GUE_0 random matrix. In view of (4.32), we could have equivalently formulated Proposition 10.6 using the usual mean width: if $\tilde{C}_{n,s}$ denotes the smallest constant such that the inequalities

$$\tilde{C}_{n,s}^{-1}\frac{w(K^\circ)}{\sqrt{s}} \leqslant \mathbf{E}\left\|\rho_{n,s}-\frac{\mathrm{I}}{n}\right\|_K \leqslant \tilde{C}_{n,s}\frac{w(K^\circ)}{\sqrt{s}}, \tag{10.7}$$

are true for every convex body containing 0 in the interior, then the conclusions of Proposition 10.6 hold for $\tilde{C}_{n,s}$ instead of $C_{n,s}$.

Proof. It is easy to check that (10.6) holds for some $C_{n,s} < +\infty$ if n and s are fixed (see Exercise 10.4). Moreover, we know from Theorem 6.35(i) that, for every fixed n,

$$\sup\{C_{n,s} \ : \ s \in \mathbb{N}\} < +\infty. \tag{10.8}$$

(i) Assume that $n = n_k$ and $s = s_k$, with n_k and s_k/n_k both tending to infinity, and denote $A_k = \sqrt{ns}(\rho_{n,s} - \mathrm{I}/n)$ and $B_k = G_n/\sqrt{n}$. Consider the random variables $X_k = d_\infty(\mu_{\mathrm{sp}}(A_k), \mu_{\mathrm{SC}})$ and $Y_k = d_\infty(\mu_{\mathrm{sp}}(B_k), \mu_{\mathrm{SC}})$. We know from Theorem 6.23 and Theorem 6.35(iii) that X_k and Y_k converge to zero in probability. We also claim that $\lim \mathbf{E}\,X_k = \lim \mathbf{E}\,Y_k = 0$; this follows from the fact that $X_k \leqslant 2 + \|A_k\|$, $Y_k \leqslant 2 + \|B_k\|$ and from Proposition 6.24 and Proposition 6.33. Part (i) follows now from Proposition 10.4.
(ii) Let A_k and B_k be as before, but now we only assume that $s_k \geqslant an_k$ for some $a > 0$. We argue by contradiction: suppose that C_{n_k,s_k} tends to infinity. We know from (10.8) that the sequence (n_k) cannot be bounded, so we may assume $\lim_k n_k = +\infty$. Similarly, using part (i), we may assume that s_k/n_k is bounded, and therefore (by passing to a subsequence) that $\lim s_k/n_k = \lambda \in [a, \infty)$. We know from Theorem 6.35(ii) and Theorem 6.23 that $\mu_{\mathrm{sp}}(A_k)$ and $\mu_{\mathrm{sp}}(B_k)$ converge in probability towards a nontrivial deterministic limit, and therefore satisfy the hypotheses of Proposition 10.5 for some constants p, c_1, C_2. □

Exercise 10.4. Let X and Y two $\mathbb{R}^n$-valued random vectors with the property that, for any $\theta \in S^{n-1}$, we have $0 < \mathbf{E}\,|\langle X, \theta\rangle| < +\infty$ and $0 < \mathbf{E}\,|\langle Y, \theta\rangle| < +\infty$. Show that there exists a constant C (depending on n, X, Y) such that, for any convex body K containing the origin in the interior, we have $\mathbf{E}\,\|X\|_K \leqslant C\,\mathbf{E}\,\|Y\|_K$.

10.1.4. Concentration for gauges of induced states. We present a concentration result valid for any gauge evaluated on random induced states.

PROPOSITION 10.8. *Let $s \geqslant n$, let $K \subset \mathrm{D}(\mathbb{C}^n)$ be a convex body with inradius r, and let ρ be a random state with distribution $\mu_{n,s}$. Let M be the median of $\|\rho - \mathrm{I}/n\|_{K_0}$, with $K_0 = K - \mathrm{I}/n$. Then, for every $\eta > 0$,*

$$\mathbf{P}\left(\left|\left\|\rho - \frac{\mathrm{I}}{n}\right\|_{K_0} - M\right| \geqslant \eta\right) \leqslant \exp(-s) + 2\exp(-n^2 s r^2 \eta^2/72).$$

PROOF OF PROPOSITION 10.8. We know that ρ has the same distribution as $AA^\dagger$, where A is an $n \times s$ matrix uniformly distributed on the Hilbert–Schmidt sphere S_{HS}. Consider the function $f : S_{\mathrm{HS}} \to \mathbb{R}$ defined by

$$f(A) = \left\|AA^\dagger - \frac{\mathrm{I}}{n}\right\|_{K_0}. \tag{10.9}$$

For every $t > 0$, denote by Ω_t the subset $\Omega_t = \{A \in S_{\mathrm{HS}} \ : \ \|A\|_\infty \leqslant t\}$. The function f is the composition of several operations:
(a) the map $A \mapsto \|A\|_{K_0}$, which is $1/r$-Lipschitz with respect to the Hilbert–Schmidt norm.
(b) the map $A \mapsto A - \mathrm{I}/n$, which is an isometry for the Hilbert–Schmidt norm,
(c) the map $A \mapsto AA^\dagger$, which is $2t$-Lipschitz on Ω_t (see Lemma 8.22).
It follows that the Lipschitz constant of the restriction of f to Ω_t is bounded by $2t/r$. We now apply the local version of Lévy's lemma (Corollary 5.35) and obtain that, for every $\eta > 0$,

$$\mathbf{P}(|f - M| \geqslant \eta) \leqslant \mathbf{P}(S_{\mathrm{HS}} \backslash \Omega_t) + 2\exp(-n s r^2 \eta^2/8t^2).$$

If we choose $t = 3/\sqrt{n}$, then $\mathbf{P}(S_{\mathrm{HS}} \backslash \Omega_t) \leqslant \exp(-s)$ (apply Proposition 6.36 with $\varepsilon = \sqrt{s/n}$) and the result follows. □

REMARK 10.9. Taking $t = 1$ in the argument above, one obtains that the global Lipschitz constant of f is bounded by $2/r$. This implies (see Proposition 5.29) that any two central values for f differ by at most $C/(r\sqrt{ns})$.

10.2. Separability of random states

Assume now that we work in a bipartite Hilbert space, and for simplicity consider the case of $\mathbb{C}^d \otimes \mathbb{C}^d$ where both parties play a symmetric role. Throughout this section we write Sep for $\mathrm{Sep}(\mathbb{C}^d \otimes \mathbb{C}^d)$ and consider random induced states on $\mathbb{C}^d \otimes \mathbb{C}^d$ with distribution $\mu_{d^2,s}$.

10.2.1. Almost sure entanglement for low-dimensional environments. Since the maximally mixed state lies in the interior of the set of separable states, and since the measures $\mu_{d^2,s}$ converge weakly towards the Dirac mass at the maximally mixed state (see Section 6.2.3.4), it follows that $\mu_{d^2,s}(\mathrm{Sep})$ tends to 1 when s tends to infinity (d being fixed). Conversely, the following result shows that random induced states are entangled with probability one when $s \leqslant (d-1)^2$.

PROPOSITION 10.10. *Let d, s be integers with $s \leqslant (d-1)^2$. Then $\mu_{d^2,s}(\mathrm{Sep}) = 0$.*

PROOF. Let $S \subset \mathbb{C}^d \otimes \mathbb{C}^d$ be the range of ρ. The random subspace S is Haar-distributed on the Grassmann manifold $\mathsf{Gr}(s, \mathbb{C}^d \otimes \mathbb{C}^d)$. We use the following simple fact which is an immediate consequence of the definition of separability: if

ρ is separable, then S is spanned by product vectors. The Proposition now follows from Theorem 8.1: when $s \leqslant (d-1)^2$, S almost surely contains no nonzero product vector. □

PROBLEM 10.11. *For which values of d, s do we have $\mu_{d^2,s}(\mathrm{Sep}) = 0$?*

EXERCISE 10.5. Let d, s be integers with $s \geqslant d^2$. Show that $0 < \mu_{d^2,s}(\mathrm{Sep}) < 1$.

EXERCISE 10.6. Let d, s be integers such that $\mu_{d^2,s}(\mathrm{Sep}) > 0$. Show that $\mu_{d^2,t}(\mathrm{Sep}) > 0$ for every $t \geqslant s$. (Cf. Problem 10.14.)

10.2.2. The threshold theorem. From the two extreme cases, $s \leqslant (d-1)^2$ and $s = \infty$, we may infer that induced states are more likely to be separable when the environment has larger dimension. As it turns out, a phase transition takes place (at least when d is sufficiently large): the generic behavior of ρ "flips" to the opposite one when s changes from being a little smaller than a certain threshold dimension s_0 to being larger than s_0. More precisely, we have the following theorem.

THEOREM 10.12. *Define a function $s_0(d)$ as $s_0(d) = w(\mathrm{Sep}(\mathbb{C}^d \otimes \mathbb{C}^d)^\circ)^2$. This function satisfies*

$$cd^3 \leqslant s_0(d) \leqslant Cd^3 \log^2 d \tag{10.10}$$

for some constants c, C and is the threshold between separability and entanglement in the following sense. If ρ is a random state on $\mathbb{C}^d \otimes \mathbb{C}^d$ induced by the environment $\mathbb{C}^s$, then, for any $\varepsilon > 0$,

(i) if $s \leqslant (1-\varepsilon)s_0(d)$, we have

$$\mathbf{P}(\rho \text{ is entangled}) \geqslant 1 - 2\exp(-c(\varepsilon)d^3), \tag{10.11}$$

(ii) if $s \geqslant (1+\varepsilon)s_0(d)$, we have

$$\mathbf{P}(\rho \text{ is separable}) \geqslant 1 - 2\exp(-c(\varepsilon)s), \tag{10.12}$$

where $c(\varepsilon)$ is a constant depending only on ε.

As a corollary, we recover the result mentioned in the preamble of the chapter: given N identical particles in a generic pure state, if we assign k of them to Alice and k of them to Bob, their shared state suddenly jumps from typically entangled to typically separable when k crosses a certain threshold value $k_N \sim N/5$. We state the result for qubits only, but both the statement and the proof easily generalize to D-level particles for $D > 2$.

COROLLARY 10.13 (see Exercise 10.8). *Given an integer N, there is $k_N \sim N/5$ with the following property. For some integer $k \leqslant N/2$, decompose $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$ as $\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{E}$ with $\mathcal{A} = \mathcal{B} = (\mathbb{C}^2)^{\otimes k}$ and $\mathcal{E} = (\mathbb{C}^2)^{\otimes(N-2k)}$, and consider a unit vector $\psi \in \mathcal{H}$ chosen uniformly at random. Let $\rho = \mathrm{Tr}_{\mathcal{E}} |\psi\rangle\langle\psi|$ be the induced state on $\mathcal{A} \otimes \mathcal{B}$. Then*

(1) for $k < k_N$, $\mathbf{P}(\rho \text{ is entangled}) \geqslant 1 - 2\exp(-\alpha^N)$,
(2) for $k > k_N$, $\mathbf{P}(\rho \text{ is separable}) \geqslant 1 - 2\exp(-\alpha^N)$,

where $\alpha > 1$ is a constant independent of N.

PROOF OF THEOREM 10.12. The inequalities (10.10) are a direct consequence of Theorem 9.6.

We next present a detailed proof of part (ii). Let $\rho_{d^2,s}$ be a random state with distribution $\mu_{d^2,s}$. Denote $\mathrm{Sep}_0 = \mathrm{Sep} - \mathrm{I}/d^2$. Consider also the function $f(\rho) = \left\|\rho - \frac{\mathrm{I}}{d^2}\right\|_{\mathrm{Sep}_0}$ and the quantity $E_{d,s} := \mathbf{E}\, f(\rho_{d^2,s})$.

Fix $\varepsilon > 0$, and let s, d be such that $s \geqslant (1+\varepsilon)s_0(d)$. Appealing to Proposition 10.6 (in the version given in Remark 10.7), we obtain

$$E_{d,s} \leqslant \tilde{C}_{n,s} \frac{w(K^\circ)}{\sqrt{s}} \leqslant \frac{\tilde{C}_{n,s}}{\sqrt{1+\varepsilon}}, \tag{10.13}$$

where $\tilde{C}_{n,s}$ is the constant appearing in (10.7). The constants $\tilde{C}_{n,s}$ tend to 1 as d and s tend to infinity under the constraint $s \geqslant (1+\varepsilon)s_0(d)$.

Let $M_{d,s}$ be the median of $f(\rho_{d^2,s})$. We know from Proposition 10.8 (the inradius of Sep being $\Theta(1/d^2)$, see Table 9.1) that

$$\mathbf{P}\left(f(\rho_{d^2,s}) > M_{d,s} + \eta\right) \leqslant \exp(-s) + 2\exp(-cs\eta^2). \tag{10.14}$$

Remark 10.9 implies that $|M_{d,s} - E_{d,s}| \leqslant Cd/\sqrt{s}$. It follows then from (10.13) that there is an $\eta > 0$ (depending only on ε) with the property that $M_{d,s} + \eta \leqslant 1$ for all d large enough and $s \geqslant (1+\varepsilon)s_0(d)$. The inequality (10.12) follows now from (10.14) and from the obvious remark that a state ρ is entangled if and only if $f(\rho) > 1$. Small values of d can be taken into account by adjusting the constants if necessary. Note that the argument yields *a priori* a bound $C' \exp(-c'(\varepsilon)s)$, possibly with $C' > 2$, but the bound (10.12) follows then with $c(\varepsilon) = c'(\varepsilon)/\log_2 C'$.

The proof of part (i) goes along similar lines, particularly if we do not care about the exact power of d appearing in the exponent of the probability bound in (10.11); this is because Proposition 10.8 yields an estimate parallel to (10.14) for $\mathbf{P}\left(f(\rho_{d^2,s}) < M_{d,s} - \eta\right)$. There are some fine points which emerge when s is relatively small, but they can be handled using inequalities from Exercise 10.7; see [**ASY14**] for details. See also Remark 10.15. □

The fine points in the proof of part (i) of Theorem 10.12 would disappear if the answer to the following natural problem was positive (cf. Exercise 10.6).

Problem 10.14 (As environment increases, entanglement decreases). *Fix an integer $d \geqslant 2$. Is it true that the function $s \mapsto \mu_{d^2,s}(\mathrm{Sep})$ is non-decreasing?*

Remark 10.15. An alternative and simpler argument to prove part (i) of Theorem 10.12 is sketched in Exercise 10.9. That argument also has the advantage that it produces explicitly an entanglement witness certifying that the induced state is entangled. However, the argument works only in the range $s \leqslant cd^3$ for some constant $c > 0$; while this does not cover the entire range, it handles the case of relatively small s that does not readily follow from Proposition 10.8.

Exercise 10.7 (Partial results on monotonicity of entanglement). Set $\pi_{d,s} := \mu_{d^2,s}(\mathrm{Sep}(\mathbb{C}^d \otimes \mathbb{C}^d))$.
(i) Show that the function $d \mapsto \pi_{d,s}$ is non-increasing for any integer $s \geqslant 1$.
(ii) Show the inequality $\pi_{2d,s} \leqslant \pi_{d,4s}$.

Exercise 10.8 (Proof of the $N/5$ threshold result). Prove Corollary 10.13 by combining Theorem 10.12 (applied with $\varepsilon = 1/2$) and Exercise 10.7.

Exercise 10.9 (The induced state is its own witness). Let ρ be a random state on $\mathbb{C}^d \otimes \mathbb{C}^d$ with distribution $\mu_{d^2,s}$, and $W = \rho - \mathrm{I}/d^2$.

(i) Show that $\mathrm{Tr}(W\rho)$ is of order $1/s$ with high probability.
(ii) Show that for any unit vector $x \in \mathbb{C}^d \otimes \mathbb{C}^d$ and $0 < \eta < 1$, we have

$$\mathbf{P}\Big(|\langle x|W|x\rangle| > \frac{\eta}{d^2}\Big) \leqslant C\exp(-cs\eta^2).$$

(iii) Conclude that with high probability, $\sup\{\mathrm{Tr}(\sigma W) \ : \ \sigma \in \mathrm{Sep}\} \leqslant Cd^{-3/2}s^{-1/2}$.
(iv) Conclude that in the regime $s \leqslant cd^3$, with high probability, W witnesses the fact that ρ is entangled.

10.3. Other thresholds

10.3.1. Entanglement of formation. Theorem 10.12 settles the "entanglement vs. separability" dichotomy for random induced states. In the generic entanglement regime, we could be more precise and ask about quantitative estimates: *how strongly* is a random state entangled?

To address the above question we need a method to quantify the amount of entanglement present in a quantum state. The approach from the preceding section allows to use the value of the gauge $\|\rho - \mathrm{I}/d^2\|_{\mathrm{Sep}_0}$ as a measure of the strength of entanglement. In this section we will work with invariants that are more "native" to quantum information theory.

For a pure state ψ, the entropy of entanglement $E(\psi)$ was introduced in (8.1). A possible way to extend this definition to mixed states is to use a "convex roof" construction. For a state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$, define its *entanglement of formation* $E_F(\rho)$ as

$$E_F(\rho) = \inf\Big\{\sum p_i E(\psi_i) \ : \ \rho = \sum p_i |\psi_i\rangle\langle\psi_i|\Big\}, \tag{10.15}$$

the infimum being taken over all decompositions of ρ as convex combinations of pure states. Equivalently, the entanglement of formation is the smallest convex function which coincides with the entropy of entanglement on pure states.

Entanglement of pure states was studied in Chapter 8. In particular, for a random pure state ψ (which corresponds to the case $s = 1$), we typically have $E_F(|\psi\rangle\langle\psi|) = E(\psi) = \log d - \frac{1}{2} + o(1)$; see Lemma 8.13. Here is a statement describing a "behavior shift" which takes place as s increases.

THEOREM 10.16 (Entanglement of formation for random induced states). *Let ρ be a random state on $\mathbb{C}^d \otimes \mathbb{C}^d$ with distribution $\mu_{d^2,s}$.*

(1) If $s \leqslant cd^2/\log^2 d$, then with high probability $E_F(\rho) \geqslant \log(d) - 1$.
(2) If $0 < \varepsilon < 1$ and $s \geqslant C\varepsilon^{-2} d^2 \log^2 d$, then with high probability $E_F(\rho) \leqslant \varepsilon$.

PROOF. Assume $s \leqslant d^2$. If S denotes the range of ρ, then S is a random Haar-distributed s-dimensional subspace of $\mathbb{C}^2 \otimes \mathbb{C}^2$. We use the following relaxation

$$E_F(\rho) \geqslant \inf\{E(\psi) \ : \ \psi \in S\}.$$

We then conclude using Theorem 8.15 that, with high probability, $E_F(\rho) \geqslant \log(d) - 1$ provided $s \leqslant cd^2/\log^2 d$.

For the second part, denote by a the smallest eigenvalue of ρ and consider the convex combination

$$\rho = (\rho - a\,\mathrm{I}) + a\,\mathrm{I} = (1 - d^2 a)\sigma + d^2 a \frac{\mathrm{I}}{d^2}$$

for some state σ. Using the convexity of E_F and the obvious facts that $E_F(\sigma) \leqslant \log d$ and $E_F(\mathrm{I}/d^2) = 0$, we obtain $E_F(\rho) \leqslant (1 - d^2 a)\log d$. However, we know from Proposition 6.36 (or Exercise 6.43) that $a \geqslant \frac{1}{d^2} - \frac{C}{d\sqrt{s}}$ with large probability. It follows that as long as $s \geqslant C^2\varepsilon^{-2} d^2 \log^2 d$, then

$$E_F(\rho) \leqslant \frac{Cd\log(d)}{\sqrt{s}} \leqslant \varepsilon. \qquad \square$$

EXERCISE 10.10. Check that $E_F(\rho) = 0$ if and only if ρ is separable.

10.3.2. Threshold for PPT. The machinery developed in this chapter can be applied to any property instead of separability and allows to reduce the estimation of threshold dimensions to the estimation of a geometric quantity (the mean width for the polar set).

One natural example is the PPT property. Since $\mathrm{PPT} = \mathrm{D} \cap \Gamma(\mathrm{D})$, where Γ is the partial transpose, it follows easily (arguing as in the first part of the proof of Proposition 9.8) that $w\big(\mathrm{PPT}_0^\circ\big) \leqslant 2w(\mathrm{D}_0^\circ) \simeq d$. The threshold s_1 appearing in this approach satisfies then

$$s_1(d) = w(\mathrm{PPT}_0^\circ)^2 = \Theta(d^2).$$

However, we know that the spectrum of large-dimensional partially transposed random states is described by a non-centered semicircular distribution (see Theorem 6.30). A more precise estimation of the threshold follows (note that the distribution $SC(\lambda, \lambda)$ appearing in Theorem 6.30 has support $[\lambda - 2\sqrt{\lambda}, \lambda + 2\sqrt{\lambda}]$, which is included in $[0, +\infty)$ if and only if $\lambda \geqslant 4$).

Theorem 10.17 (Threshold for the PPT property). *Define $s_1(d) = 4d^2$. Let ρ be a random state on $\mathbb{C}^d \otimes \mathbb{C}^d$ with distribution $\mu_{d^2,s}$. Then*

(i) if $s \leqslant (1-\varepsilon)s_1(d)$, we have

$$\mathbf{P}(\rho \text{ is PPT}) \leqslant 2\exp(-c(\varepsilon)d^2),$$

(ii) if $s \geqslant (1+\varepsilon)s_1(d)$, we have

$$\mathbf{P}(\rho \text{ is PPT}) \geqslant 1 - 2\exp(-c(\varepsilon)s).$$

Here $c(\varepsilon)$ is a constant depending only on ε.

The comparison between Theorems 10.12, 10.16 and 10.17 is instructive: if s is sufficiently larger than d^2, but sufficiently smaller than d^3, random states are typically PPT and entangled (in particular they cannot be distilled, see Chapter 12), but have an amount of entanglement extremely small when measured via the entanglement of formation.

Exercise 10.11. Explain the presence of expressions of the form $\Omega_\varepsilon(d^2)$ and $\Omega_\varepsilon(s)$ in the exponents in Theorem 10.17.

Notes and Remarks

Theorem 10.12, as well as the preliminary results from Section 10.1, are from [**ASY14**]. A high-level non-technical overview can be found in [**ASY12**]. In particular, the existence of a separability threshold around the value $s = d^3$ was proved in [**ASY14**]; previously only the cases $s \leqslant d^2$ or $s \geqslant d^4$ were covered (see e.g [**HLW06**]).

The answer to Problem 10.11 is known for qubits: we have $\mu_{4,2}(\mathrm{Sep}(\mathbb{C}^2 \otimes \mathbb{C}^2)) = 0$ and $\mu_{4,3}(\mathrm{Sep}(\mathbb{C}^2 \otimes \mathbb{C}^2)) > 0$. As explained in section 7.1 of [**ASY14**], this follows from results of [**RW09**] and [**SBŻ06**], respectively.

The entanglement of formation is only one of the many possible ways to quantify entanglement of mixed states. However, other measures are harder to manipulate. For a survey of the subject of entanglement measures see [**PV07**].

The threshold for the entanglement of formation (Theorem 10.16) is essentially from [**HLW06**], and the threshold for the PPT property (Theorem 10.17) is from [**Aub12**] (see also [**ASY12**]).

Other thresholds functions have been computed or estimated: for the realignment criterion [**AN12**], for the k-extendibility property [**Lan16**], and for still other properties [**CNY12, JLN14, JLN15**] (including the absolute PPT property and the reduction criterion).