APPLICATIONS OF RANDOM MATRICES IN QUANTUM INFORMATION THEORY

ION NECHITA

Abstract. These are notes for five lectures given at the second School of pthe program "Operator Algebras, Groups and Applications to Quantum Information" held in May 2019 at the ICMAT in Mardid.

The goal of this series of lectures is to present some recent results in quantum information theory which make use of random matrices. The main goal is to understand the spectrum of the partially transposed random quntum states, in the limit of large matrix dimension. The mathematical tools needed to do this are Gaussian integraion via the Wick formula, the method of moments, and some basic notions of free probability theory. Along the way, we shall discuss at length the Marchenko-Pastur limit theorem for the Wishart ensemble.

Contents

Date: May 7, 2019.

Introduction

Let us gather here some basic definitions from quantum information theory and set up some notation. A recent excellent monograph containing a lot of relevant material for these lectures is [AS17]. A *quantum state* is a positive semidefinite matrix of unit trace. The set of all quantum states is a convex body denoted by

$$\mathcal{M}_d^{1,+}(\mathbb{C}) := \{\rho \in \mathcal{M}_d(\mathbb{C}) \,:\, \rho \geq 0 \text{ and } \operatorname{Tr}\rho = 1\}.$$

The extremal points of $\mathcal{M}_d^{1,+}(\mathbb{C})$ are the rank one projectors xx^* ($x \in \mathbb{C}^d$, $\|x\| = 1$), and they are called *pure states*.

Of particular interest are states of multiple quantum systems, which are quantum states acting on the *tensor product* of the corresponding Hilbert spaces. Of particular importance are the *separable states*, which in the bipartite case can be described as

$$\mathcal{SEP}_{d_1,d_2} := \operatorname{conv}\{\rho_1 \otimes \rho_2\}_{\rho_i \in \mathcal{M}_{d_i}^{1,+}(\mathbb{C})}.$$

Non-separable states are called *entangled*, and among those, of particular importance is the *maximally entangled state* $d^{-1}\Omega_d \in \mathcal{M}_{d^2}^{1,+}(\mathbb{C})$, where $\Omega_d = \sum_{i=1}^d e_i \otimes e_i$ and $\{e_i\}$ is an orthonormal basis of $\mathbb{C}^d$.

An intermediate set between the set of separable states and the set of all quantum states is the set of *positive partial transpose* states:

$$\mathcal{PPT}_{d_1,d_2} := \{\rho \in \mathcal{M}_{d_1 d_2}^{1,+}(\mathbb{C}) \,:\, \rho^\Gamma := [\mathrm{id}_{d_1} \otimes \mathrm{transp}_{d_2}](\rho) \geq 0\}.$$

The inclusion

$$\mathcal{SEP}_{d_1,d_2} \subseteq \mathcal{PPT}_{d_1,d_2} \tag{0.1}$$

always holds, with equality iff $(d_1, d_2) \in \{(2,2),(2,3),(3,2)\}$, see Figure 1. This fact is a deep result in operator algebra, see [Stø63, Wor76] and [AS17, Section 2.4.5]. One of the main objective of these lectures is to quantify how far is the inclusion (0.1) from being an equality for large dimensions $d_{1,2}$. We shall answer this question in the balanced case ($d_1, d_2 \to \infty$) in Section 3 and in the unbalanced case ($d_1 \to \infty$, d_2 fixed) in Section 5. Sections 1 and 2 contain the random matrix theory pre-requisites, while Section 4 contains the necessary notions and results from Voiculescu's free probability theory needed to state the results in the unbalanced case.

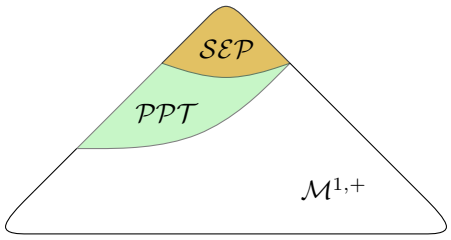


Figure 1. Three convex sets: the set of PPT contains the set of separable quantum states.

1. Gaussian integration. Wick formula

1.1. **Graphical notation for tensors.** In this section, we lay out the foundation for the graphical calculus we shall develop later. We introduce a graphical formalism for representing tensors and tensor contractions that is adapted to quantum information theory. We start at an abstract level, with a purely diagrammatic axiomatization and then we study the Hilbert representations, where graph-theoretic objects shall be associated with concrete elements of Hilbert spaces. This fairly standard graphical notation will be extended in Section 1.3 to random tensors: one can compute

expectation values of diagrams containing Gaussian and/or Haar-unitary boxes by performing a graphical expansion.

1.1.1. *Diagrams, boxes, decorations and wires.* Our starting point is a set $\tilde{S}$ endowed with an involution without fixed point $*$. The set $\tilde{S}$ splits as $S \sqcup S^*$ according to the involution. Elements of $\tilde{S}$ are called *decorations*, and will correspond to vector spaces and their duals.

A *diagram* is a collection of decorated *boxes* and possibly *wires* (or strings) connecting the boxes along their decorations according to rules which we shall specify. In terms of graph theory, a diagram is an unoriented (multi-)graph whose vertices are boxes, and whose edges are strings. Each vertex comes with a (possibly empty) n-tuple of indices (or decorations or labels) in $\tilde{S}^n$. The number n of decorations may depend on the vertex. We say that two diagrams are isomorphic if they are isomorphic as multi-graphs with labeled vertices.

A box is an elementary diagram from which we can construct more elaborate diagrams by putting boxes together and possibly wiring them together. Each box B of a diagram has attached to it a collection of $n(B)$ decorations in $\tilde{S}^{n(B)}$. The union of the decorations attached to a box B is denoted by $S(B) \sqcup S^*(B)$.

Graphically, boxes are represented by rectangles with symbols corresponding to the decorations attached to them (see Figure 2). We take the convention that decorations in S^* are represented by empty (or white) symbols and decorations in S by full (or black) symbols; moreover, we shall depict white decorations on the right hand side of a box, and black decorations on the left (following the standard "right-to-left" matrix multiplication direction). Each decoration is thought as having potentially up to two *attachment points*. An *inner* one (which is attached to the box it belongs to) and an *outer* one, which we shall allow to be attached to a string later on.

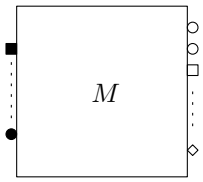


FIGURE 2. A box M

1.1.2. *Constructing new diagrams out of old ones.* Given a family of existing diagrams (e.g. boxes) there exists several ways of creating new diagrams.

(1) One can put diagrams together, i.e. take their disjoint union (when it comes to taking representations in Hilbert spaces, this operation will amount to tensoring). One diagram can be viewed as a box. This amounts to specifying an order between the boxes.
(2) Given a diagram A and a complex number x, one can create a new diagram $A' = xA$.
(3) Given two boxes A, B having the same n-tuples of decorations, one can define $A + B$. This axiom and the previous one (together with evident relations such as $A + A = 2A$ which we don't enumerate in detail) endow the set of identically decorated diagrams with a structure of a complex vector space.
(4) One can add *wires* to an existing diagram (or between two diagrams that have been put together). A wire is allowed between the outer attachment of two decorations only if the decorations have the same shape and different shadings. Such a wire can be created if and only if the two candidate decorations have their outer attachments unoccupied.
(5) There exists an anti-linear *involution* on the diagrams, denoted by $*$. This operation does nothing on the wires. On the boxes, it reverts the shading of the decorations. The involution $*$ is conjugate linear.

1.1.3. *Hilbert structure.* We shall now consider a concrete representation of the diagrams introduced above as tensors in Hilbert spaces. We start by assuming that the set S of full (or black) decorations corresponds to a collection of finite dimensional Hilbert spaces $S = \{V_1, V_2, \cdots\}$. An important fact that will be useful later is that each Hilbert space V_i comes equipped with an orthonormal basis $\{e_1, e_2, \ldots, e_{\dim V_i}\}$. Our aim is to define a $*$-linear map T between the diagrams and tensors in products of Hilbert spaces in the above class and their duals. By duality, white decorations correspond to dual spaces $S^* = \{V_1^*, V_2^*, \cdots\}$. With these conventions, boxes can be seen as tensors whose legs belong to the vector spaces corresponding to its decorations. In a diagram, symbols of the same shape denote isomorphic spaces, but the converse may be false. A particular space V_i (or V_i^*) can appear several times in a box. The reader acquainted with quantum mechanics might think of white shapes as corresponding to "bras" and black shapes corresponding to "kets", but we shall get back to quantum mechanical notions later.

To a box B we therefore associate a tensor

$$T_B \in \left[\bigotimes_{i \in S(B)} V_i\right] \otimes \left[\bigotimes_{j \in S^*(B)} V_j^*\right]. \tag{1.1}$$

Using the canonical duality between tensors and multilinear maps, T_B can also be seen as a function

$$T_B : \bigotimes_{j \in S^*(B)} V_j \to \bigotimes_{i \in S(B)} V_i,$$

We use freely partial duality results, and for example, an element of $V \otimes W^*$ can as well be seen as an element of $\mathcal{L}(W, V)$ or $\mathcal{L}(V^*, W^*)$.

Equation (1.1) defines the map T from the collection of boxes to the collection of vectors in Hilbert spaces obtained by tensoring finitely many copies of $V_i, i \in S(B) \cup S^*(B)$. This map is denoted by

$$T : B \mapsto T_B$$

and we now explain how we can extend it to all diagrams. A *wire* connecting two decorations of the same shape (corresponding to some Hilbert space V) is associated with the identity map (or tensor) $I : V \to V$. Together with our duality axiom, it also corresponds to a canonical tensor contraction (or trace)

$$C : V^* \otimes V \to \mathbb{C}.$$

We denote the set of wires in a diagram $\mathcal{D}$ by $\mathcal{C}(\mathcal{D})$.

With this notation, a diagram $\mathcal{D}$ is associated with the tensor T obtained by applying all the contractions ("wires") to the product of tensors represented by the boxes. One is left with a tensor

$$T_{\mathcal{D}} = \left[\prod_{C \in \mathcal{C}(\mathcal{D})} C\right] \left(\bigotimes_{B \text{ box of } \mathcal{D}} T_B\right).$$

This is well defined (provided that one specifies one total order on the boxes): the order of the factors in the product does not matter, since wires act on different spaces. For a box B, we denote by $FS(B) \subset S(B)$ the subset of black decorations which have no wires attached (we call such a decoration *free*). $FS^*(B)$ is defined in the same manner for white decorations (dual spaces). With this notation, the tensor $T_{\mathcal{D}}$ associated to a diagram $\mathcal{D}$ can be seen in two ways: as an element of a Hilbert space

$$T_{\mathcal{D}} \in \left[\bigotimes_{j \in \bigcup_B FS^*(B)} V_j^*\right] \otimes \left[\bigotimes_{i \in \bigcup_B FS(B)} V_i\right],$$

or, equivalently, as a linear map

$$T_{\mathcal{D}} : \bigotimes_{j \in \bigcup_B FS^*(B)} V_j \to \bigotimes_{i \in \bigcup_B FS(B)} V_i.$$

We need two further axioms to ensure that we are indeed dealing with acceptable Hilbert representations.

(1) A diagram such that all outer attachments of its decorations are occupied by wires corresponds canonically to an element in $\mathbb{C}$. In addition, a trivial box with a given decoration of type i closed on itself by a wire into a loop takes a value in $\mathbb{N}$. This value is called the dimension of V_i.
(2) Given a diagram $\mathcal{D}$, if it is canonically paired to its dual $\mathcal{D}^*$ by strings, the result lies in $\mathbb{R}^+$.

1.1.4. *Special diagrams.* To make our calculus useful, we need to introduce a few special diagrams (equivalently, boxes) satisfying some specific axioms.

(1) **The trivial box.** A wire connecting two identically shaped decorations of different shading corresponds to the identity map $I : V \to V$. We shall call this box the *trivial* or the identity box.

FIGURE 3. Trivial box, corresponding to the identity matrix.

It satisfies the following identity axiom:

FIGURE 4. Trivial axiom: $I = I \cdot I$.

(2) **Bras and kets.** The simplest boxes one can consider are vectors and linear forms. Following the quantum mechanics "bra" and "ket" vocabulary, vectors, or (1, 0)-tensors have no white decorations and only one black decoration, whereas linear forms (or (0,1)-tensors) have one white label and no black labels. We represent in Figure 5 a ket $x \in V$ and a bra $\varphi \in V^*$.

FIGURE 5. A ket (left) and bra (right).

(3) **The Bell state.** Since each space $V \in S$ comes equipped with a particular fixed basis $\{e_i\}_{i=1}^{\dim V}$, we can define the *bra Bell state* as the tensor (it is in fact a linear form)

$$\Omega^* = \sum_{i=1}^{\dim V} e_i^* \otimes e_i^*,$$

and its ket counterpart (which is a vector in $V \otimes V$)

$$\Omega = \sum_{i=1}^{\dim V} e_i \otimes e_i.$$

This notation is needed in the sense that Bell states are not canonical and are not well defined from the sole data of V. They rely on some additional real structure of the vector

space V which can be encoded by the data of an explicit basis. Bell states are represented in Figure 6(a). They satisfy the graphical axiom in Figure 6(b). Bell states play a central role in our formalism; we shall see later that they allow us to define the *transposition* of a box and even to consider wires connecting identical decorations.

FIGURE 6. Bell states (ket and bra, left) and axiom (right).

(4) **Unitary boxes.** Boxes associated to unitary matrices U satisfy the graphical axiom depicted in Figure 7 which corresponds to the identities $UU^* = U^*U = I$.

FIGURE 7. Axioms for unitary matrices.

1.1.5. *Examples.* Let us now look at some simple diagrams which illustrate this formalism.

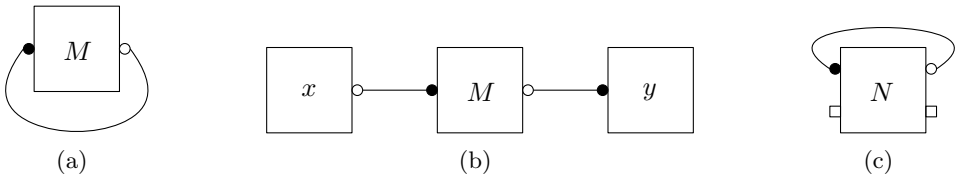


FIGURE 8. Some simple diagrams: a trace, a scalar product, and a partial trace.

Suppose that each diagram in Figure 8 comes equipped with two vector spaces V_1 and V_2 which we shall represent respectively by circle and square shaped symbols. In the first diagram, M is a tensor (or a matrix, depending on which point of view we adopt) $M \in V_1^* \otimes V_1$, and the wire applies the contraction $V_1^* \otimes V_1 \to \mathbb{C}$ to M. The result of the diagram $\mathcal{D}_a$ is thus $T_{\mathcal{D}_a} = \mathrm{Tr}(M) \in \mathbb{C}$. In the second diagram, again there are no free decorations, hence the result is the complex number $T_{\mathcal{D}_b} = \langle y, Mx \rangle$. Finally, in the third example, N is a $(2,2)$ tensor or a linear map $N \in \mathcal{L}(V_1 \otimes V_2, V_1 \otimes V_2)$. When one applies to the tensor N the contraction of the couple (V_1, V_1^*), the result is the partial trace of N over the space V_1: $T_{\mathcal{D}_c} = \mathrm{Tr}_{V_1}(N) \in \mathcal{L}(V_2, V_2)$.

Bell states allow us to introduce the *transposition* operation for a tensor (or a box) as follows. We define, as usual, transposition for a matrix M (or a tensor $M \in V^* \otimes V$) and we extend it in a trivial way to more general situations. Graphically, the box corresponding to the transposed tensor M^t is defined in Figure 9(a): the box M has its "legs" transposed; the wires are connecting decorations of the same color, so this is a non-canonical, basis dependent, operation. Note however that this operation is different from the involution $*$ applied to the same box. Bell states allow for wires connecting identical shaped symbols of the same color, as in Figure 9(b). Such non-canonical tensor contractions ($V \otimes V \to \mathbb{C}$ or $V^* \otimes V^* \to \mathbb{C}$) are shorthand graphical notations for the corresponding diagram containing a Bell state, and we shall use them quite often in what follows.

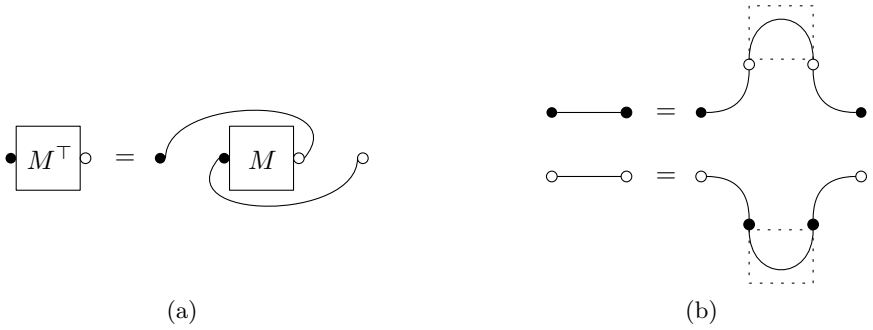


FIGURE 9. The transposition operation (left) and Bell states allowing for wires between same-color decorations (right).

Also, for reasons which shall be clear later, we shall sometimes make substitution $\overline{M} = (M^*)^\top$. Finally, by grouping two Bell states together, one obtains the (non-canonical) tensor ω (Figure 10), called "the maximal entangled state". It corresponds to the tensor

$$\omega = \Omega\Omega^* = \sum_{i=1}^{\dim V} \sum_{j=1}^{\dim V} e_i \otimes e_i \otimes e_j^* \otimes e_j^* \in V \otimes V \otimes V^* \otimes V^*.$$

The reader with background in quantum information will notice that the maximally entangled state we just defined is *not normalized* in order to be a density matrix. The reader with background in planar algebra theory will recognize a multiple of the Jones projection.

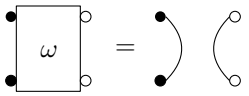


FIGURE 10. The (un-normalized) maximal entangled state ω.

The graphical formalism we have just defined allows us to perform some operations directly on the level of diagrams, bypassing sometimes cumbersome algebraic manipulations involving summing over families of indices. As a first example of this philosophy, we prove graphically in Figure 11 that the partial trace of the (un-normalized) maximally entangled state is the identity matrix.

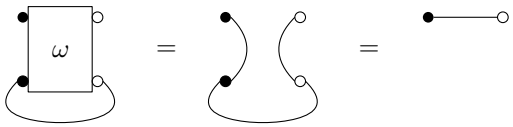


FIGURE 11. The partial trace of the (un-normalized) maximal entangled state ω is the identity operator.

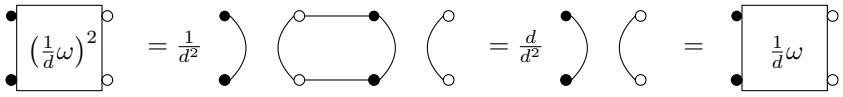


FIGURE 12. The normalized maximal entangled state is a projection. Note that the loop in the second diagram is equal to the dimension of the vector space, d here.

1.1.6. *Comments on other existing graphical calculi.* The above formalism is the one that seemed the most compatible with Weingarten calculus. Here, we comment about already existing graphical formalisms, in the hope that this section will serve as a dictionary for the reader acquainted to one of the calculi below.

Our calculus is mainly inspired by Bob Coecke's *Kindergarten Quantum Mechanics* [Coe10]. However we choose not to orient the strings; rather, we separate with color (black/white) the vector spaces and their duals, therefore there is only one possible pairing. A common feature of the two calculi is the central place occupied in the formalisms by Bell states.

V.F.R. Jones's theory of planar algebras [Jon99] is also connected to our graphical calculus. One of our diagrammatic axioms is the existence of a Bell state. This is very closely related to the axioms of Temperley-Lieb algebras and the diagrammatic for a Jones projection. Most of our calculus could take place in Jones' *bipartite graph planar algebra*.

However it is not clear whether planarity plays an important role in our calculus. More generally, one could view our calculus as fitting in the frame of traced monoidal (or tensor) categories. Here, our objects are our elementary family of Hilbert spaces, their duals and all their finite Hilbert tensor products. The monoidal structure corresponds graphically consists in copying two diagrams side to side, and amounts to taking tensor product of the Hilbert spaces. The trace corresponds to the conditional expectations obtained by our wiring procedure. We refer for example to [JSV96].

1.2. **Gaussian integration.** The Gaussian (or normal) distribution is arguably the most important probability distribution in mathematics and in science, due to the *Central Limit Theorem*: properly normalized sums of independent, identically distributed (i.i.d.) random variables converge to a Gaussian distribution.

In the real case, a *Gaussian distribution* of mean m and variance σ^2 has the following density with respect to the Lebesgue measure $\mathrm{d}x$:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m^2)}{2\sigma^2}\right);$$

if X is such a random variable, we write $X \sim \mathcal{N}(m, \sigma^2)$, see Figure 13, left panel, for some examples. One can consider multi-dimensional Gaussian distributions, characterized by a vector $m \in \mathbb{R}^n$ and a positive definite *covariance matrix* $\Sigma \in \mathcal{M}_n(\mathbb{R})$. The density of such a random vector $\mathbb{R}^n \ni X \sim \mathcal{N}(m, \Sigma)$ reads

$$\frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}\langle x-m, \Sigma^{-1}(x-m)\rangle\right).$$

Of importance in what follows is the multi-variate *standard* (i.e. zero mean, identity variance) complex case, where a random variable $Z \in \mathbb{C}^n$ is said to have a standard normal distribution if it has density

$$\frac{1}{\pi^n} \exp(-\|z\|^2/2).$$

In particular, a scalar standard complex random variable Z has independent real and imaginary parts, both having distribution $\mathcal{N}(0, 1/2)$, see Figure 13, right panel. More general complex Gaussian vectors $Z \in \mathbb{C}^d$ are described by a complex vector m and a positive definite complex covariance matrix Σ:

$$\begin{aligned} \forall i, \qquad & \mathbb{E}Z_i = m_i \\ \forall i, j, \qquad & \mathbb{E}[\bar{Z}_i Z_j] = \Sigma_{ij}. \end{aligned}$$

We now address the question of computing integrals with respect to a Gaussian distribution. The combinatorial method described below is known in the literature as *Wick's formula*, or Isserlis'

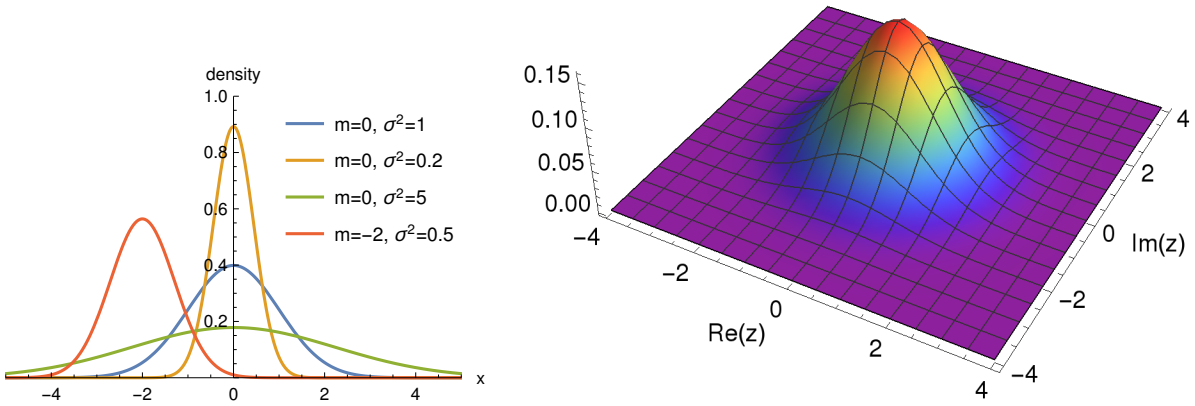
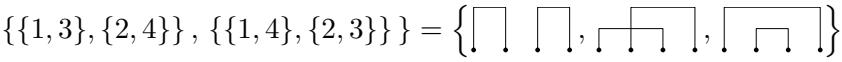


FIGURE 13. Gaussian distributions, in the real case (left) and in the complex case (right).

formula [Iss18]. We denote below by $P_2(k)$ the set of pair-partitions of the set $[k] = \{1, 2, \ldots, k\}$; for example,

$$P_2(4) = \{\, \{\{1,2\},\{3,4\}\}\,,\ \{\{1,3\},\{2,4\}\}\,,\ \{\{1,4\},\{2,3\}\}\,\} =$$



Theorem 1.1. *Let Z be a (complex) Gaussian n-variate random vector with zero mean. Then*

$$\mathbb{E}[Z_{i_1} Z_{i_2} \cdots Z_{i_k}] = \sum_{\pi \in P_2(k)} \prod_{\{s,t\} \in \pi} \mathbb{E}[Z_{i_s} Z_{i_t}].$$

Proof. We shal prove the statement in the real case and leave the complex setting to the reader. First, note that if k is odd, both sides are zero: the LHS is zero by the invariance of a centered Gaussian distribution over a global sign change, while the RHS is zero since there are no pair partitions of $[k]$; we assume thus $k = 2r$. The proof strategy follows [Wit85]. The first ingredient of the proof is a *Laplace transform*: for all $\lambda \in \mathbb{R}^n$, we have

$$\mathbb{E} \exp\langle \lambda, Z \rangle = \exp\left(\frac{\langle \lambda, \Sigma \lambda \rangle}{2} \right).$$

We leave the proof of the claim above as an exercise, it follows from a linear change of variables in the Gaussian integral. Taking partial derivatives with respect to the λ_{i_s} variables and evaluating at $\lambda = 0$, we obtain

$$\mathbb{E}[Z_{i_1} Z_{i_2} \cdots Z_{i_k}] = \frac{\partial^k}{\partial \lambda_{i_1} \cdots \partial \lambda_{i_k}} \bigg|_{\lambda = 0} \exp\left(\frac{\langle \lambda, \Sigma \lambda \rangle}{2} \right).$$

To evaluate the derivative on the RHS, we use *Faà di Bruno's formula for the chain rule*, see [Har06, Proposition 1 and equation (4)]:

$$\frac{\partial^k}{\partial x_1 \cdots \partial x_k} f(y) = \sum_{\pi \in P(k)} f^{(\#\pi)}(y) \prod_{B \in \pi} \frac{\partial^{|B|} y}{\prod_{j \in B} \partial x_j},$$

where the sum is over *all* partitions π of $[k]$, and $\#\pi$ denotes the number of blocks of a given partition π (see Section 2.2 for the combinatorics of (non-crossing) partitions). In our situation, f above is the exponential function, so $f^{(\#\pi)} = f = \exp$, while y is a quadratic function of the x variables, so only *pair partitions* survive. □

It is a remarkable property of the Gaussian distribution that all the moments of the distribution can be computed using only the covariance. For example, for a centered Gaussian vector Z having covariance matrix Σ, we have

$$\mathbb{E}[Z_1 Z_2 Z_3 Z_4] = \Sigma_{12}\Sigma_{34} + \Sigma_{13}\Sigma_{24} + \Sigma_{14}\Sigma_{23}.$$

Since the number of pair partitions of $[2n]$ is

$$|P_2(2n)| = (2n-1)(2n-3)\cdots 5 \cdot 3 \cdot 1 =: (2n)!!,$$

we have the following corollary.

Corollary 1.2. *If X is a real standard Gaussian random variable, we have, for all $n \geq 0$:*

$$\mathbb{E}[X^{2n}] = (2n)!!$$
$$\mathbb{E}[X^{2n+1}] = 0.$$

For Z a complex standard Gaussian, he have

$$\mathbb{E}[\bar{Z}^m Z^n] = \delta_{m,n}(m+n)!!.$$

1.3. **Graphical Wick formula.** We shall now recast the Wick formula above in the graphical formalism described previously. Consider a diagram which contains a new special box G corresponding to a *Gaussian random matrix* (i.e. the entries of the matrix are i.i.d. standard complex Gaussian random variables). We shall compute the expected value of a random diagram with respect to the Gaussian probability measure; as we shall see, this operation will consist of *expanding* the diagram, by erasing the Gaussian boxes and replacing them with wires.

To start, consider $\mathcal{D}$ a diagram which contains, amongst other constant tensors, boxes corresponding to independent Gaussian random matrices of *covariance one* (identity). One can deal with more general Gaussian matrices by multiplying the standard ones with constant matrices. Note that a box can appear several times, adjoints of boxes are allowed and the diagram may be disconnected. Also, Gaussian matrices need not be square.

The expectation value of such a random diagram $\mathcal{D}$ can be computed by a *removal* procedure as in the unitary case. Without loss of generality, we assume that we do not have in our diagram adjoints of Gaussian matrices, but instead their complex conjugate box. This assumption allows for a more straightforward use of the Wick formula from Theorem 1.1. We can assume that $\mathcal{D}$ contains only one type of random Gaussian box G; other independent random Gaussian matrices are assumed constant at this stage as they can be removed in the same manner afterwards.

A removal of the diagram $\mathcal{D}$ is a pairing between *Gaussian boxes* G and their conjugates $\bar{G}$. The set of removals is denoted by $\mathrm{Rem}_G(\mathcal{D})$ and it may be empty: if the number of G boxes is different from the number of $\bar{G}$ boxes, then $\mathrm{Rem}_G(\mathcal{D}) = \emptyset$ (since no pairing between matrices and their conjugates can exist). Otherwise, a removal r can identified with a permutation $\alpha \in \mathcal{S}_p$, where p is the number of G and $\bar{G}$ boxes. In the Gaussian/Wick calculus, one pairs conjugate boxes: white and black decorations are paired in an identical manner, hence only one permutation is needed to encode the removal.

To each removal r associated to a permutation $\alpha \in \mathcal{S}_p$ corresponds a removed diagram $\mathcal{D}_r$ constructed as follows. One starts by erasing the boxes G and $\bar{G}$, but keeps the decorations attached to these boxes. Then, the decorations (white *and* black) of the i-th G box are paired with the decorations of the $\alpha(i)$-th $\bar{G}$ box in a coherent manner, see Figure 14.

The graphical reformulation of the Wick formula from Theorem 1.1 becomes the following theorem, which we state without proof.

Theorem 1.3. *The following holds true:*

$$\mathbb{E}_G[\mathcal{D}] = \sum_{r \in \mathrm{Rem}_G(\mathcal{D})} \mathcal{D}_r.$$

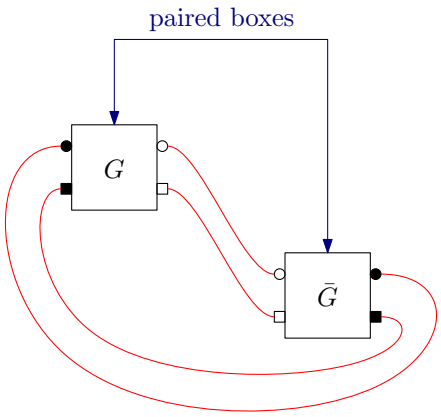


FIGURE 14. Pairing of boxes in the Gaussian case

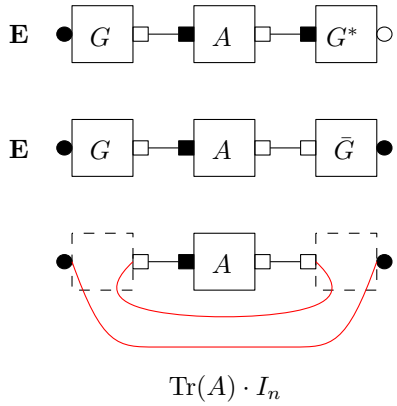


FIGURE 15. Applying Theorem 1.3 to compute $\mathbb{E}[GAG^*]$.

In Figure 15, we present an example of application of the theorem above. We consider, on the first row, the diagram corresponding to $\mathbb{E}[GAG^*]$, where $G \in \mathcal{M}_{n\times k}(\mathbb{C})$ is a $n \times k$ Gaussian matrix, and $A \in \mathcal{M}_k(\mathbb{C})$ is a square, deterministic matrix. The first row contains the diagram $\mathcal{D}$ associated to the algebraic expression. In the second row, we rewrite the same diagram, replacing G^* by $\bar{G}^\top$, in order to be able to apply Theorem 1.3. The third row contains the result of the application: we erase the $G/\bar{G}$ boxed and we add the wires corresponding to the permutation $(1) \in \mathcal{S}_1$ (in red). We recognize the diagrams for the identity matrix and for the trace of A: $\mathbb{E}[GAG^*] = \mathrm{Tr}(A)I_n$.

2. The Wishart ensemble

The birth of random matrix theory can be traced to statistics and physics. Wishart introduced the distribution that bears his name in the 1920's [Wis28], in order to explain the discrepancy between the eigenvalues of a measured covariance matrix, and an expected covariance matrix. Later, Wigner was studying nuclear physics when he introduced [Wig55] the semi-circle distribution. Since then, random matrix theory has played a role in many fields of mathematics and science, including operator algebras [VDN92], combinatorics, complex analysis, theoretical physics and telecommunication theory, just to cite a few. Quantum information theory is definitely one of the most recent of fields of application; for more on this, we direct the interested reader to the recent review [CN16]. The classical reference for random matrix theory is Mehta's book [Meh04]; more recent monographs written in a mathematical language are [AGZ10, MS17].

In quantum information theory, randomness is built in, by the axioms of quantum mechanics. Since quantum states are modeled by (unit trace, positive semidefinite) matrices, it is clear that the two fields intersect. However, we can see two more reasons for the use of random matrices in quantum information. First, we would like to understand the *typical properties* of quantum states and channels, relative to tasks ans paradigms in quantum information theory. Very early, properties such as the average entanglement of quantum states were studied [Pag93], and several probability distribution over the set of quantum states were introduced [ŻS01]. Second, it turns out that some problems – in particular the minimum output entropy additivity problem, which we discuss at length here – did not have an obvious non-random answer, therefore it became not only natural, but also important, to consider random quantum objects.

One paper which popularized the use of random techniques in quantum information was [HLSW04]. This work pointed out that some well-established techniques in the mathematics of random matrices – measure concentration in this case – could be of use in quantum information.

2.1. **Wishart matrices and their limit distribution.** Historically the first ensemble of random matrices having been studied is the Wishart ensemble [Wis28], see [BS10, Chapter 3] or [AGZ10, Section 2.1] for a modern presentation.

Definition 2.1. *Let $G \in \mathcal{M}_{d\times s}(\mathbb{C})$ be a random matrix with complex, standard, i.i.d. Gaussian entries. The distribution of the positive-semidefinite matrix $W = GG^* \in \mathcal{M}_d(\mathbb{C})$ is called a* Wishart distribution *of parameters (d, s) and is denoted by $\mathcal{W}_{d,s}$.*

The study of the asymptotic behavior of Wishart random matrices is due to Marčenko and Pastur [MP67], while the stronger convergence results have been proved by analytic tools such as determinantal point processes; one can also recover the stronger forms of the theorem as direct consequences of the much more general results [Mal12]. Since we aim at giving complete proofs of our results, we state it here in a rather week form: the convergence in moments.

Definition 2.2. *A sequence of random matrices X_d is said to* converge in moments *to a probability distribution ν if for all positive integers p, we have*

$$\lim_{d\to\infty} \mathbb{E}\int t^p \mathrm{d}\mu_{X_d} = \mathbb{E}\frac{1}{d}\operatorname{Tr}(X_d^p) = \int t^p \mathrm{d}\nu,$$

where μ_{X_d} is the empirical eigenvalue distribution *of X_d*

$$\mu_{X_d} = \frac{1}{d}\sum_{i=1}^{d} \delta_{\lambda_i(X_d)}.$$

Theorem 2.3. *Consider a sequence s_d of positive integers which behaves as $s_d \sim cd$ as $d \to \infty$, for some constant $c \in (0, \infty)$. Let W_d be a sequence of positive-semidefinite random matrices such that W_d is distributed according to $\mathcal{W}_{d,s_d}$. Then, the sequence W_d converges in moments to the* Marčenko-Pastur distribution *π_c given by*

$$\pi_c = \max(1-c, 0)\delta_0 + \frac{\sqrt{(b-x)(x-a)}}{2\pi x}\,\mathbf{1}_{(a,b)}(x)\,\mathrm{d}x, \tag{2.1}$$

where $a = (1-\sqrt{c})^2$ and $b = (1+\sqrt{c})^2$.

The Marčenko-Pastur distribution π_c is sometimes called the *free Poisson distribution*, see [NS06, Proposition 12.11]. We plotted in Figure 16 its density in the cases $c = 1$ and $c = 4$.

Remark 2.4. *The Dirac mass appearing in (2.1) is due to the fact that if $c < 1$, the matrix W_d is rank deficient. Since $cd < d$, a fraction $1 - c$ of the eigenvalues of W_d are null, yielding the Dirac mass at zero.*

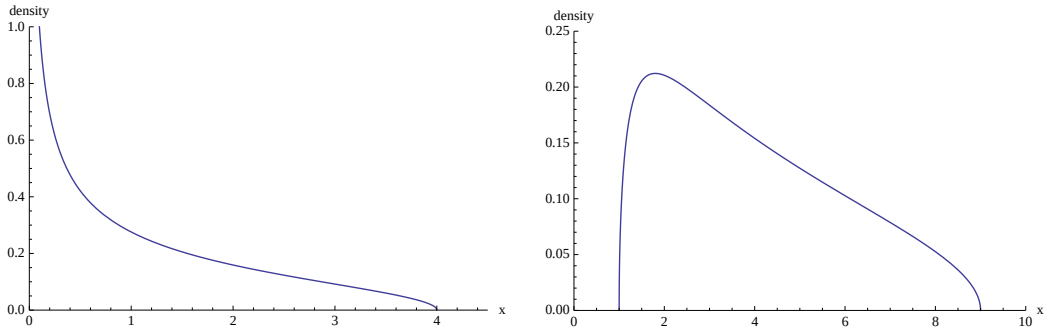


FIGURE 16. The density of the Marčenko-Pastur distributions π_1 (left) and π_4 (right).

We postpone the proof of Theorem 2.3 to Section 2.3.

We end this section by the statement of the so-called Carleman condition, which ensures that a sequence of moments defines a unique probability measure.

Proposition 2.5. *Let μ be a probability measure on $\mathbb{R}$ having finite moments*

$$m_n = \int t^p \mathrm{d}\mu(t)$$

which satisfy

$$\sum_{n=1}^{\infty} m_{2n}^{-1/(2n)} = +\infty.$$

Then, μ is the only measure on $\mathbb{R}$ having the sequence (m_n) as moments.

2.2. **Non-crossing partitions and permutations.** For a permutation $\sigma \in \mathcal{S}_p$, denote by $\#\sigma$ the number of its cycles, including the trivial ones (fixed points). Denote also by $|\sigma|$ its *length*, i.e. the minimum number of transposition which multiply to σ. It is well known that for all permutations $\sigma \in \mathcal{S}_p$,

$$\#\sigma + |\sigma| = p.$$

The set of *non-crossing partitions* will play a crucial role in what follows. Recall that a partition π of $[p] := \{1, 2, \ldots, p\}$ is called non-crossing if there are now quadruples (a, b, c, d) such that a, b (resp. c, d) belong to the same block of π, and $a < c < b < d$; see Figure 17 for some examples. The are supremum and infimum operations on $NC(p)$, which turn it into a lattice, see [NS06, Lecture 9]. The number of elements in the set $NC(p)$ is the *Catalan number*

$$\mathrm{Cat}_p = \frac{1}{p+1}\binom{2p}{p}.$$

These numbers satisfy the recurrence relation

$$\mathrm{Cat}_p = \sum_{i=1}^{p} \mathrm{Cat}_{i-1}\mathrm{Cat}_{p-i},$$

and thus their generating series is given by

$$M(z) = \sum_{p=0}^{\infty} \mathrm{Cat}_p z^p = \frac{1 - \sqrt{1-4z}}{2z}.$$

We collect now a some properties of the distance function over the symmetric group, which allow us to bijectively identify a subset of $\mathcal{S}_p$ with $NC(p)$. This result can be traced back to [Bia97].

FIGURE 17. A non-crossing partition $\{\{1,2,4\},\{3\}\}$ (left) vs. a crossing one $\{\{1,3\},\{2,4\}\}$ (right).

Lemma 2.6. *The function $d(\sigma,\tau) = |\sigma^{-1}\tau|$ is an integer valued distance on $\mathcal{S}_p$. Besides, it has the following properties:*

- *the diameter of $\mathcal{S}_p$ is $p-1$;*
- *$d(\cdot,\cdot)$ is left and right translation invariant;*
- *for three permutations $\sigma_1,\sigma_2,\tau \in \mathcal{S}_p$, the quantity $d(\tau,\sigma_1)+d(\tau,\sigma_2)$ has the same parity as $d(\sigma_1,\sigma_2)$;*
- *the set of* geodesic *points (elements which saturate the triangular inequality) between the identity permutation* id *and some permutation $\sigma \in \mathcal{S}_p$ is in bijection with the set of non-crossing partitions smaller than π, where the partition π encodes the cycle structure of σ. Moreover, the preceding bijection preserves the lattice structure.*

2.3. **Proof of the Marcenko-Pastur theorem.** Proof of the Marčenko-Pastur theorem We have now all the elements to present a short and elegant proof of Theorem 2.3.

Proof of Theorem 2.3. The proof will consist of three independent steps: computing the moments, at fixed d, of the random matrix W_d, letting $d \to \infty$ and computing the limiting moments, and finally identifying the probability measure having precisely these moments.

Step 1. Moment formula

We are interested, for any fixed dimensions d, s, in computing the p-th moment of the random matrix $W_d = GG^*$, where G is a $d \times s$ matrix with i.i.d. complex standard Gaussian random entries. To do this, we consider the diagram $\mathcal{D}$ corresponding to the random variable $\mathrm{Tr}(W_d^p)$. This diagram contains p pairs $(G,\bar{G})$ of Gaussian boxes, which are connected as in Figures 18 and 19. More precisely, the label corresponding to $\mathbb{C}^d$ which is attached to the i-th G-box is connected to the corresponding label attached to the $(i-1)$-th $\bar{G}$-box. On the other hand, the label corresponding to $\mathbb{C}^s$ which is attached to the i-th G-box is connected to the corresponding label attached to the i-th $\bar{G}$-box. Using the graphical Wick formula from Theorem 1.3, we have

$$\mathbb{E}\,\mathrm{Tr}(W_d^p) = \mathbb{E}\mathcal{D} = \sum_{\alpha\in\mathcal{S}_p} \mathcal{D}_\alpha,$$

where $\mathcal{D}_\alpha$ is the removal diagram obtained by deleting the $G/\bar{G}$ boxed and connecting the labels according to the permutation α. It is clear that each diagram $\mathcal{D}_\alpha$ consists only of loops of two types: ones coming from round labels corresponding to $\mathbb{C}^d$ spaces, and others coming from square labels corresponding to $\mathbb{C}^s$ spaces. The number of loops of each type is the number of cycles in the permutation $\beta^{-1}\alpha$, where β encodes the initial wiring of the labels of each type; see Figures 18 and 19 for some examples. In conclusion, we have

$$\mathbb{E}\,\mathrm{Tr}(W_d^p) = \sum_{\alpha\in\mathcal{S}_p} d^{\#(\gamma^{-1}\alpha)} s^{\#\alpha}. \tag{2.2}$$

In the formula above, $\#(\cdot)$ is the number of cycles function, and γ is the full cycle permtuation

$$\gamma = (p\,(p-1)\cdots 3\,2\,1) \in \mathcal{S}_p.$$

Step 2. Asymptotic moments

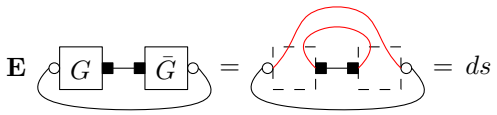


FIGURE 18. The first moment of a Wishart matrix using the graphical Wick calculus from Theorem 1.3. Round labels correspond to $\mathbb{C}^d$, while square labels correspond to $\mathbb{C}^s$.

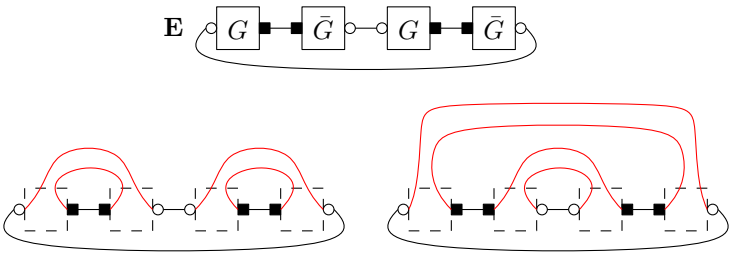


FIGURE 19. The second moment of a Wishart matrix using the graphical Wick calculus. On the top row, the diagram for $\mathbb{E}\,\mathrm{Tr}(W_d^2)$. On the bottom row, the two diagrams corresponding to the permutations $\mathrm{id} = (1)(2)$, on the left, and (12), on the right. Their values are respectively ds^2 and $d^2 s$.

Let us now consider the asymptotic regime we are interested in, $d \to \infty$ and $s \sim cd$, for some fixed parameter $c \in (0, \infty)$. Since the terms in (2.2) are all positive, we have

$$\mathbb{E}\,\mathrm{Tr}(W_d^p) \sim \sum_{\alpha \in \mathcal{S}_p} c^{\#\alpha} d^{\#(\gamma^{-1}\alpha)+\#\alpha}.$$

The dominating terms in the sum above are those maximizing the quantity $\#(\gamma^{-1}\alpha) + \#\alpha$ over the symmetric group. Using the properties of the distance function $|\cdot|$ on permutations from Lemma 2.6, we have

$$\#(\gamma^{-1}\alpha) + \#\alpha = 2p - (|\alpha| + |\gamma^{-1}\alpha|) \leq 2p - |\gamma| = p + 1,$$

where equality is attained iff α is a *geodesic* permutation (it saturates the triangle inequality $|\mathrm{id}^{-1}\alpha| + |\alpha^{-1}\gamma| \geq |\mathrm{id}^{-1}\gamma|$). We conclude that

$$\mathbb{E}\,\mathrm{Tr}(W_d^p) \sim d^{p+1} \sum_{\sigma \in NC(p)} c^{\#\sigma}.$$

Notice that considering only the dominating terms from the sum (2.2), indexed over all permutations, selects the ones for which the permutations are non-crossing partitions.

Step 3. The Marčenko-Pastur distribution

We are going to treat here the case $c = 1$; the general case is similar. We can rewrite the asymptotic moment formula as

$$\lim_{d\to\infty} \mathbb{E}\frac{1}{d}\,\mathrm{Tr}\left[(d^{-1}W_d)^p\right] = \mathrm{Cat_p}.$$

We claim that the unique probability measure μ having the Catalan numbers as moments is the one from (2.1):

$$\pi_1 = \frac{\sqrt{x(4-x)}}{2\pi x}\,\mathbf{1}_{(0,4)}(x)\,dx.$$

To show this, recall that the generating function of the Catalan number must be the moment generating function of μ:

$$M_\mu(z) = \sum_{p=0}^{\infty} z^p \int t^p \mathrm{d}\mu = \frac{1-\sqrt{1-4z}}{2z},$$

where the relation above holds formally (as a power series in z), and analytically, in a small neighborhood of 0. The *Cauchy transform* of μ reads now

$$G_\mu(z) = \int \frac{1}{z-t}\mathrm{d}\mu(t) = z^{-1}M_\mu(z^{-1}) = \frac{1-\sqrt{1-4z^{-1}}}{2},$$

which holds now on a neighborhood of the infinity in the complex plane. One recovers the density of μ via the *Stieltjes inversion formula*, which says that if we denote by

$$h_\varepsilon(t) := -\frac{1}{\pi}\Im G_\mu(t+i\varepsilon),$$

then

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \lim_{\varepsilon\to 0} h_\varepsilon(t).$$

In our case, we recover $\mu = \pi_1$.

The uniqueness clame comes from the fact that π_1 is compactly supported, hence it satisfies the Carleman condition from Proposition 2.5. □

3. Random quantum states and their partial transposition

In this lecture, we turn to the main goal of the series, that is the study of the partial transposition of random quantum states. Having discussed all the presrequisites from random matrix theory in the previous lecture, we first specify what we mean when we talk about *random quantum states*, and then move on to study the partial transposition of this model of random matrices.

3.1. Random quantum states: the induced measures.

Endowing the set of quantum states with natural probability measures is an important task, at least for two reasons. On the one hand, it allows to study and understand the properties of *typical quantum states*, providing qualitative and quantitative answers to questions such as: is a generic quantum state entangled or separable? or what is the probability that the PPT (positive partial transpose) entanglement criterion fails, i.e. what is the probability that a random state is PPT entangled? Answering these questions, under reasonable notions of randomness, is the main topic of this series of lectures. On the other hand, random constructions have proven to be very successful in quantum information theory (we cite the resolution in the negative of the additivity conjecture of the minimum output entropy of quantum channels by Hastings [Has09] as an example); hence, different ensembles of quantum states are a valuable source of examples (and counter-examples) for the theory.

In the case of random pure quantum states, there is a clear candidate for a probability distribution: *the Lebesgue measure on the unit sphere of the corresponding complex Hilbert space*. Given the rotation invariance of the (complex) Gaussian distribution on $\mathbb{C}^d$, a random pure d-dimmensional quantum state can be defined simply by renormalizing a standard complex Gaussian random vector $g \in \mathbb{C}^d$:

$$|\psi\rangle = \frac{g}{\|g\|}.$$

Note that in the definition above, the random variables ψ and $\|g\|$ are *independent*: this is again a consequence of the rotational invariance of the standard Gaussian distribution, and can be seen as a generalization of the scalar (1D) situation, where the modulus $|z|$ of a complex Gaussian $z = |z|e^{i\varphi}$ is independend of its angle (or phase) φ.

In the mixed state case, there is no unique candidate for a probability measure on the convex body $\mathcal{M}_d^{1,+}$. Of course, the set of states inherits the Lebesgue measure of its ambient space, and

one can normalize it to have unit mass. We shall see however that the Lebesgue measure is just a special case of a one parameter family of probability distributions which are very natural, both from a mathematical and from a physical viewpoint. Before going into details, let us briefly mention here another distribution on the set of states which has received a lot of attention, and which is motivated by considerations from statistics, the *Bures measure* [Hal98, SZ03, OSŻ10].

Let us introduced the family of *induced measures* starting from a physical perspective. Assume that the system of interest (modelled by the Hilbert space $\mathbb{C}^d$) is coupled to a s-dimensionnal environment $\mathbb{C}^s$ and that the joint system is in a pure state $|\psi\rangle$, which is distributed uniformly on the unit sphere of the product Hilbert space $\mathbb{C}^d \otimes \mathbb{C}^s \cong \mathbb{C}^{ds}$. The reduced density matrix $\rho = \mathrm{Tr}_s\, |\psi\rangle\langle\psi|$ is a random mixed quantum state, and the *induced measure of parameters* d, s is the distribution of this random matrix. Note that ρ is a $d \times d$ random matrix, the parameter s appearing in the expression of its denisty. One can compute the probability distribution of this random matrix [ŻS01, ŻPNC11]

$$\mathrm{d}\mathbb{P}(\rho) = C_{d,s} \det \rho^{s-d} \mathbf{1}_{\rho \geq 0,\, \mathrm{Tr}\, \rho = 1} \mathrm{dLeb}(\rho), \tag{3.1}$$

where $C_{d,s}$ is a normalizing constant and Leb is the Lebesgue measure on the set of $d \times d$ hermitian matrices. In particular, it is a remarkable fact [ŻS03] that, for $s = d$ (i.e. the size of the envoronment is equal to the size of the system of interest), one recovers a uniform density, thus the Lebesgue measure (or the Hilbert-Schmidt measure) on the set of density matrices. Integrating out the Haar-distributed eigenvectors from (3.1), one obtains the probability density of the spectrum $(\lambda_1, \ldots, \lambda_d)$ of ρ, with respect to the Lebesgue measure on the probability simplex $\Delta_{d-1} := \{x \in \mathbb{R}^d \,:\, x_i \geq 0 \text{ and } \sum_i x_i = 1\}$:

$$\mathrm{d}\mathbb{P}(\lambda_1, \ldots, \lambda_d) = C'_{d,s} \prod_{i=1}^{d} \lambda_i^{s-d} \prod_{1 \leq i < j \leq d} (\lambda_i - \lambda_j)^2 \mathbf{1}_{\lambda_i \geq 0, \sum_i \lambda_i = 1} \mathrm{dLeb}(\lambda),$$

it is a remarkable fact that random quantum states following the induced distribution of parameters (d, s) can also be obtained as normalized Wishart matrix of the same parameters, see also [Nec07, ŻPNC11]

$$\rho = \frac{W}{\mathrm{Tr}\, W} = \frac{GG^*}{\mathrm{Tr}(GG^*)},$$

where G is a $d \times s$ random matrix with i.i.d. standard complex Gaussian entries. To establish this equivalence, one uses the independence of the random variables ρ and $\mathrm{Tr}\, W$ appearing above, see [Nec07, Proposition 4 and Corollary 1].

In what follows, we shall not work with normalized quantum states, but with Wishart matrices. This choice is motivated by the fact that the PPT criterion is about positivity, and normalization does not play any role in it. We shall thus work with the cone of positive semidefinite matrices (and the Wishart matrices) instead of working with quantum states. Abusing notation, we define the three cones of interest below:

$$\begin{aligned}
\mathcal{SEP}_{d,n} &= \{A \in \mathcal{M}_{dn} \,:\, A = \sum_i B_i \otimes C_i, \text{ where } B_i, C_i \geq 0\} \\
&\subseteq \\
\mathcal{PPT}_{d,n} &= \{A \in \mathcal{M}_{dn} \,:\, A^\Gamma = [\mathrm{id}_d \otimes \mathrm{transp}_n](A) \geq 0\} \\
&\subseteq \\
\mathcal{PSD}_{dn} &= \{A \in \mathcal{M}_{dn} \,:\, A \geq 0\}.
\end{aligned}$$

3.2. **The partial transpose of Wishart random matrices.** We study here the asymptotical eigenvalue distribution of the partial transposition of random quantum states. The question whether a given mixed quantum state is separable or entangled has been proven to be an NP-hard

one [Gur03]. To circumvent this worst-case intractability, *entanglement criteria* are used. These are efficiently computable conditions which are necessary for separability; in other words, an entanglement criterion is a (usually convex) super-set $\mathcal{X}_d$ of the set of separable states, for which the membership problem is efficiently solvable (see [AS15] for the number of such criteria needed to obtain a good approximation of the set of separable states). As in the previous section, from a probabilistic point of view, estimating the probability that a random quantum state (sampled from the induced ensemble) is an element of $\mathcal{X}_d$ is central.

In what follows we shall tackle this problem for one entanglement criterion in the framework of *thresholds*. Given a family $G_d \subseteq \mathcal{PSD}_d$ of convex cones, a pair of functions (s'_d, s''_d) is called a threshold for the family G_d if the following two properties are satisfied:

(1) If W_d is a sequence of Wishart random matrices of parameters (d, s_d) with $s_d \geq s''_d$, then

$$\lim_{d\to\infty} \mathbb{P}[W_d \in G_d] = 1.$$

(2) If W_d is a sequence of Wishart random matrices of parameters (d, s_d) with $s_d \leq s'_d$, then

$$\lim_{d\to\infty} \mathbb{P}[W_d \in G_d] = 0.$$

Let us start with the most used example, the *positive partial transpose* criterion (PPT). The PPT criterion has been introduced by Peres in [Per96]: if a positive semidefinite matrix $A \in \mathcal{M}_d \otimes \mathcal{M}_n$ is separable, then

$$A^\Gamma := [\mathrm{id} \otimes \mathrm{transp}](A) \geq 0.$$

Note that the positivity of A^Γ is equivalent to the positivity of $A^\intercal = [\mathrm{transp} \otimes \mathrm{id}](A)$, so it does not matter on which tensor factor the transpose application acts. We denote by $\mathcal{PPT}_{d,n}$ the PPT cone

$$\mathcal{PPT}_{d,n} := \{A \in \mathcal{M}_{dn} \,:\, A^\Gamma \geq 0\} \supseteq \mathcal{SEP}_{d,n}.$$

This necessary condition for separability has been shown to be also sufficient for qubit-qubit and qubit-qutrit systems ($dn \leq 6$) in [HHH96]; the result was a simple consequence of the fact that all the positive application from $\mathcal{M}_2$ to $\mathcal{M}_{2,3}$ are decomposable. These non trivial facts are due to Woronowocz [Wor76]. The PPT criterion for random quantum states has first been studied numerically in [ŽPBC07]. The analytic results in the following proposition are from [Aub12] (in the balanced case) and from [BN13] (in the unbalanced case); see also [FŚ13] for some improvements in the balanced case and the relation to meanders.

Proposition 3.1. *Consider a sequence $W_d \in \mathcal{M}_{dn_d}$ of random Wishart matrices of parameters (dn_d, cdn_d), where n_d is a function of d and c is a positive constant.*

In the balanced regime $n_d = d$, the (properly rescaled) empirical eigenvalue distribution of the matrices W_d^Γ converges to a semicircular measure $\mu_{SC(1,1/c)}$ of mean 1 and variance $1/c$, see (4.1). In particular, the threshold for the sets $\mathcal{PPT}_{d,d}$ $(d \to \infty)$ is $c_0 = 4$.

In the unbalanced regime $n_d = n$ fixed, the (properly rescaled) empirical eigenvalue distribution of the matrices $d^{-1}W_d^\Gamma$ converges to a free difference of free Poisson distributions (see Section 4 for the definitions)

$$\pi_{cn(n+1)/2} \boxminus \pi_{cn(n-1)/2}.$$

In particular, the threshold for the sets $\mathcal{PPT}_{d,n}$ (n fixed, $d \to \infty$) is

$$c_0 = 2 + 2\sqrt{1 - \frac{1}{n^2}}.$$

Proof. We are going to sketch the proof of the convergence result in the unbalanced case; for the balanced case, see [Aub12] and for the threshold in the unbalanced case, see [BN13, Section 6].

Using again the graphical Wick formula, one can find the following expression for the (unnormalized) moments of W_d^Γ:

$$\mathbb{E}\,\mathrm{Tr}[({W_d}^\Gamma)^p] = \sum_{\alpha\in\mathcal{S}_p} s^{\#\alpha} d^{\#(\gamma^{-1}\alpha)} n^{\#(\gamma\alpha)}.$$

Using the fact that, for every noncrossing partition $\sigma \in NC(p)$, denoting by $e(\sigma)$ the number of blocks of even size of σ, we have $1 + \mathrm{e}(\sigma) = \#(\sigma\gamma)$, we arrive at the formula

$$\begin{aligned}\mathbb{E}(dn)^{-1}\,\mathrm{Tr}[(d^{-1}{W_d}^\Gamma)^p] &\sim \sum_{\sigma\in NC(p)} n^{\#\sigma+e(\sigma)} c^{\#\sigma}\\ &\sim \sum_{\sigma\in NC(p)}\prod_{b\in\sigma} cn^{1+\mathbf{1}_{|b| \text{ is even}}}\\ &\sim \sum_{\sigma\in NC(p)}\prod_{b\in\sigma}\left(\frac{cn(n+1)}{2}+\frac{cn(n-1)}{2}(-1)^{|b|}\right).\end{aligned}$$

We can now identify the free difference of free Poisson operators using the free cumulant approach of [NS06]: the free cumulant of order p of the limiting measure is

$$\frac{cn(n+1)}{2}+\frac{cn(n-1)}{2}(-1)^{|b|}.$$

□

Remark 3.2. *The computation of the limiting distribution of in the unbalanced case performed above was done using the method of moments. A more general approach, allowing to answer the same question for general maps and general matrix distributions, was provided in* [ANV16] *using operator valued free probability theory.*

Remark 3.3. *The value of the threshold in the theorem above has a practical significance: if one considers a random pure quantum state on $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^n \otimes \mathbb{C}^{cdn}$, takes the partial trace on the third subsystem, and the partial transposition on the second subsystem, then the resulting matrix is positive semidefinite if $c > c_0$, and has negative eigenvalues if $c < c_0$, with large probability as n is fixed and $d \to \infty$.*

4. An introduction to free probability theory

4.1. **Non-commutative probability spaces. Freeness.** We have studied random matrices in the previous lecture by their moments: the only properties of the ambient probability space we have used were the fact that the random variables have an algebra structure, and the existence of the expectation functional. We abstract out these notions in the following definition [NS06, Lecture 1].

Definition 4.1. *A* non-commutative probability space *is an algebra $\mathcal{A}$ with unit endowed with a tracial state φ. An element of $\mathcal{A}$ is called a (non-commutative) random variable.*

In these lectures we have already encountered the non-commutative probability space of *random matrices* $(\mathcal{M}_d(L^{\infty-}(\Omega,\mathbb{P})), \mathbb{E}[d^{-1}\,\mathrm{Tr}(\cdot)])$, where we use the standard notation $L^{\infty-}(\Omega,\mathbb{P}) = \cap_{p\geq 1} L^p(\Omega,\mathbb{P})$; the $L^{\infty-}$ space contains all random variables with moments of all orders. We shall encounter another example in Section 4.2.

In classical probability theory, the notion of *independence* of random variables plays a very important role; in particular, it allows to compute the joint distribution of independent random variables in terms of the marginal distributions (i.e. the distributions of the individual random variables). The notion of freeness is a non-commutative alternative to classical independence.

Definition 4.2. *Let $\mathcal{A}_1, \ldots, \mathcal{A}_k$ be subalgebras of $\mathcal{A}$ having the same unit as $\mathcal{A}$. They are said to be* free *if for all $a_i \in \mathcal{A}_{j_i}$ $(i = 1, \ldots, k)$ such that $\varphi(a_i) = 0$, one has*

$$\varphi(a_1 \cdots a_k) = 0$$

as soon as $j_1 \neq j_2$, $j_2 \neq j_3, \ldots, j_{k-1} \neq j_k$. Collections $S_1, S_2, \ldots$ of random variables are said to be free if the unital subalgebras they generate are free.

Let $(a_1, \ldots, a_k)$ be a k-tuple of selfadjoint random variables and let $\mathbb{C}\langle X_1, \ldots, X_k\rangle$ be the free $*$-algebra of non commutative polynomials on $\mathbb{C}$ generated by the k indeterminates $X_1, \ldots, X_k$. The *joint distribution* of the family $\{a_i\}_{i=1}^k$ is the linear form

$$\begin{aligned}\mu_{(a_1,\ldots,a_k)} : \mathbb{C}\langle X_1, \ldots, X_k\rangle &\to \mathbb{C} \\ P &\mapsto \varphi(P(a_1, \ldots, a_k)).\end{aligned}$$

In the case of a single, self-adjoint random variable x, if the moments of x coincide with those of a compactly supported probability measure μ, i.e.

$$\forall p \geq 1, \qquad \varphi(x^p) = \int t^p d\mu(t),$$

we say that x has distribution μ. The most important distribution in free probability theory is the semicircular distribution

$$\mu_{SC(0,1)} = \frac{\sqrt{4 - x^2}}{2\pi} \mathbf{1}_{[-2,2]}(x) dx,$$

which is, for reasons we will not get into, the free world equivalent of the Gaussian distribution in classical probability (see [NS06, Lecture 8] for the details). A random variable x having distribution $\mu_{SC(0,1)}$ has the Catalan number for moments:

$$\varphi(x^p) = \begin{cases} \mathrm{Cat}_p := \frac{1}{p+1}\binom{2p}{p}, & \text{if } p \text{ is even} \\ 0, & \text{if } p \text{ is odd.} \end{cases}$$

More generally, if x has distribution $\mu_{SC(0,1)}$, we say that $y = \sigma x + m$ has distribution

$$\mu_{SC(m,\sigma^2)} = \frac{\sqrt{4\sigma^2 - (x-m)^2}}{2\pi\sigma^2} \mathbf{1}_{[m-2\sigma, m+2\sigma]}(x) \mathrm{d}x. \tag{4.1}$$

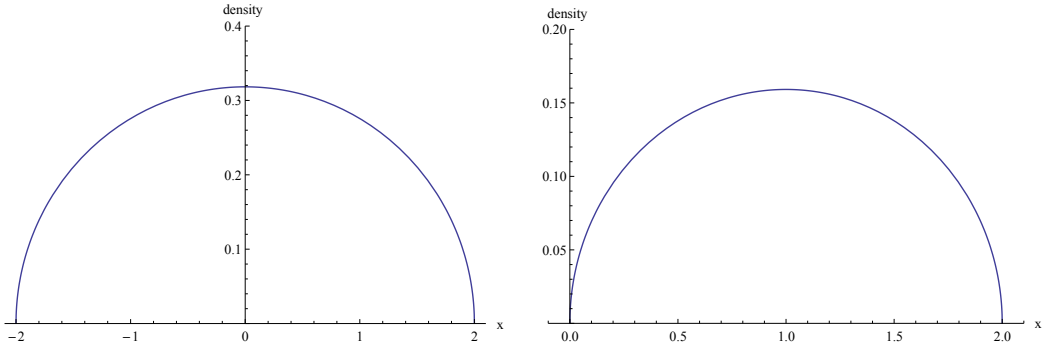


FIGURE 20. The density of the semicircular distributions $\mu_{SC(0,1)}$ (left) and $\mu_{SC(1,1/4)}$ (right).

Remark 4.3. *If the non-commutative random variable x has (standard) semicircular distribution, then x^2 has a free Poisson (or Marchenko-Pastur distribution) of parameter $c = 1$.*

Given a k-tuple $(a_1, \ldots, a_k)$ of free random variables such that the distribution of a_i is μ_{a_i}, the joint distribution $\mu_{(a_1,\ldots,a_k)}$ is uniquely determined by the μ_{a_i}'s. A family $(a_1^n, \ldots, a_k^n)_n$ of k-tuples of random variables is said to *converge in distribution* towards $(a_1, \ldots, a_k)$ iff for all $P \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$, $\mu_{(a_1^n,\ldots,a_k^n)}(P)$ converges towards $\mu_{(a_1,\ldots,a_k)}(P)$ as $n \to \infty$. Sequences of random variables $(a_1^n)_n, \ldots, (a_k^n)_n$ are called *asymptotically free* as $n \to \infty$ iff the k-tuple $(a_1^n, \ldots, a_k^n)_n$ converges in distribution towards a family of free random variables.

Given two free random variables $a, b \in \mathcal{A}$, the distribution μ_{a+b} is uniquely determined by μ_a and μ_b. The free additive convolution of μ_a and μ_b is defined by $\mu_a \boxplus \mu_b = \mu_{a+b}$. When $x = x^* \in \mathcal{A}$, we identify μ_x with the spectral measure of x with respect to τ. The operation $\boxplus$ induces a binary operation on the set of probability measures on $\mathbb{R}$. Similarly, we write $\mu_a \boxminus \mu_b = \mu_{a-b}$.

4.2. **The full Fock space, free semicircular random variables.** We discuss now a more abstract non-commutative probability space, in which freeness appears naturally.

Definition 4.4. *Let H be a complex Hilbert space. The* full Fock space *over H is defined to be*

$$\mathcal{F}(H) = \bigoplus_{n=0}^{\infty} H^{\otimes n} = \mathbb{C}\Omega \oplus \bigoplus_{n=1}^{\infty} H^{\otimes n}.$$

The bounded operators on $\mathcal{F}(H)$, together with the vacuum state

$$\tau(X) = \langle \Omega, X\Omega \rangle$$

form a non-commutative probability space. We also define, for a vector $f \in H$, the creation and annihilation operators $\ell(h)$ and $\ell(h)^*$, defined as follows:

$$\ell(f)\Omega = f$$
$$\ell(f) f_1 \otimes \cdots \otimes f_n = f \otimes f_1 \otimes \cdots \otimes f_n$$

and

$$\ell(f)^*\Omega = 0$$
$$\ell(f)^* f_1 = \langle f, f_1 \rangle \Omega$$
$$\ell(f)^* f_1 \otimes \cdots \otimes f_n = \langle f, f_1 \rangle f_2 \otimes \cdots \otimes f_n.$$

The following theorem is taken from [NS06, Section 7], where it is proven in a more general form.

Theorem 4.5. *Let $f, g \in H$ be two orthogonal vectors. Then the non-commutative random variables $x = \ell(f) + \ell(f)^*$ and $y = \ell(g) + \ell(g)^*$ are semicircular and free.*

Proof. Let us first show that both x and y have semicircular distributions; moreover, without loss of generality, let us assume that $\|f\| = 1$, and task to show that x has $\mu_{SC(0,1)}$ distribution.

To do this, fix some moment order p, and consider $\tau(x^p)$:

$$\tau(x^p) = \sum_{w:[p]\to\{1,*\}} \langle \Omega, \ell(f)^{w(p)} \ell(f)^{w(p-1)} \cdots \ell(f)^{w(2)} \ell(f)^{w(1)} \Omega \rangle.$$

For each choice of the function w, the scalar product above is either 0 or 1; we have thus to count how many choices of w give 1. It is clear that a function w gives 1 iff $p = 2q$ is even, and the lattice path induced by w is a *Dyck* path. Recall that a Dyck path is a path in the lattice $\mathbb{Z}^2$, starting at $(0, 0)$, ending at $(p = 2q, 0)$, having $(1, \pm 1)$ steps, and, importantly, staying above the x-axis at all times; see Figure 21 for an example. The number of such paths is given by the Catalan numbers, and the first part of the proof is complete.

Let us now show that x and y are free. Let us first identify which elements in the algebra generated by $\{1, \ell(f)\}$ are traceless. It is easy enough to see that, after some cancellations of the form $\ell(f)^*\ell(f) = \|f\|^2$, the only such elements are of the form

$$\ell(f) \cdots \ell(f)\ell(f)^* \cdots \ell(f)^*,$$

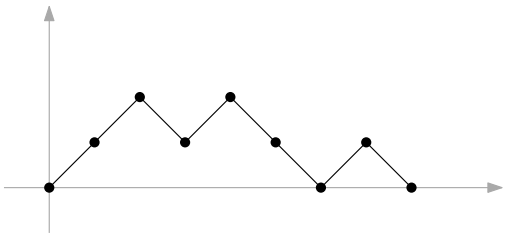


FIGURE 21. A Dyck path.

where the product above is non empty. The conclusion follows by considering arbitrary alternating products of the above type for f and g, and by noting that whenever $\ell(f)^*\ell(g)$ appears, the end result is zero; hence, the $*$-algebras generated by $\ell(f)$ and $\ell(g)$ are free. The conclusion follows. □

5. Partial transposition of random quantum states: the unbalanced case

5.1. Other entanglement criteria.

5.2. Conclusion: $\mathcal{SEP}$ vs. $\mathcal{PPT}$.

References

[AGZ10] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. Cambridge University Press, 2010. 11, 12

[ANV16] Octavio Arizmendi, Ion Nechita, and Carlos Vargas. On the asymptotic distribution of block-modified random matrices. *Journal of Mathematical Physics*, 57(1):015216, 2016. 19

[AS15] Guillaume Aubrun and Stanislaw Szarek. Dvoretzky's theorem and the complexity of entanglement detection. *arXiv preprint arXiv:1510.00578*, 2015. 18

[AS17] Guillaume Aubrun and Stanisław J Szarek. *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*, volume 223. American Mathematical Soc., 2017. 2

[Aub12] Guillaume Aubrun. Partial transposition of random states and non-centered semicircular distributions. *Random Matrices: Theory and Applications*, 1(02):1250001, 2012. 18

[Bia97] Philippe Biane. Some properties of crossings and partitions. *Discrete Mathematics*, 175(1):41–53, 1997. 13

[BN13] Teodor Banica and Ion Nechita. Asymptotic eigenvalue distributions of block-transposed Wishart matrices. *Journal of Theoretical Probability*, 26(3):855–869, 2013. 18

[BS10] Zhidong Bai and Jack W Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010. 12

[CN16] Benoit Collins and Ion Nechita. Random matrix techniques in quantum information theory. *Journal of Mathematical Physics*, 57(1), 2016. 11

[Coe10] Bob Coecke. Quantum picturalism. *Contemporary physics*, 51(1):59–83, 2010. 8

[FŚ13] Motohisa Fukuda and Piotr Śniady. Partial transpose of random quantum states: Exact formulas and meanders. *Journal of Mathematical Physics*, 54(4):042202, 2013. 18

[Gur03] Leonid Gurvits. Classical deterministic complexity of edmonds' problem and quantum entanglement. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 10–19. ACM, 2003. 18

[Hal98] Michael JW Hall. Random quantum correlations and density operator distributions. *Physics Letters A*, 242(3):123–129, 1998. 17

[Har06] Michael Hardy. Combinatorics of partial derivatives. *the electronic journal of combinatorics*, 13(1):1, 2006. 9

[Has09] Matthew B Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255–257, 2009. 16

[HHH96] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1):1–8, 1996. 18

[HLSW04] Patrick Hayden, Debbie Leung, Peter W Shor, and Andreas Winter. Randomizing quantum states: Constructions and applications. *Communications in Mathematical Physics*, 250(2):371–391, 2004. 12

[Iss18] Leon Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918. 9

[Jon99] Vaughan FR Jones. Planar algebras, i. *arXiv preprint math/9909027*, 1999. 8

[JSV96] André Joyal, Ross Street, and Dominic Verity. Traced monoidal categories. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 119, pages 447–468. Cambridge University Press, 1996. 8

[Mal12] Camille Male. The norm of polynomials in large random and deterministic matrices. *Probability Theory and Related Fields*, 154(3-4):477–532, 2012. 12

[Meh04] Madan Lal Mehta. *Random matrices*, volume 142. Academic press, 2004. 11

[MP67] Vladimir A Marčenko and Leonid Andreevich Pastur. Distribution of eigenvalues for some sets of random matrices. *Sbornik: Mathematics*, 1(4):457–483, 1967. 12

[MS17] James A Mingo and Roland Speicher. *Free probability and random matrices*, volume 35. Springer, 2017. 11

[Nec07] Ion Nechita. Asymptotics of random density matrices. *Annales Henri Poincaré*, 8(8):1521–1538, 2007. 17

[NS06] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006. 12, 13, 19, 20, 21

[OSŻ10] V Osipov, Hans-Juergen Sommers, and K Życzkowski. Random bures mixed states and the distribution of their purity. *Journal of Physics A: Mathematical and Theoretical*, 43(5):055302, 2010. 17

[Pag93] Don N Page. Average entropy of a subsystem. *Physical review letters*, 71(9):1291, 1993. 12

[Per96] Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413, 1996. 18

[Stø63] Erling Størmer. Positive linear maps of operator algebras. *Acta Mathematica*, 110(1):233–278, 1963. 2

[SZ03] Hans-Jürgen Sommers and Karol Zyczkowski. Bures volume of the set of mixed quantum states. *Journal of Physics A: Mathematical and General*, 36(39):10083, 2003. 17

[VDN92] Dan V Voiculescu, Ken J Dykema, and Alexandru Nica. *Free random variables*. 1. American Mathematical Soc., 1992. 11

[Wig55] Eugene P Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, pages 548–564, 1955. 11

[Wis28] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928. 11, 12

[Wit85] CS Withers. The moments of the multivariate normal. *Bulletin of the Australian Mathematical Society*, 32(1):103–107, 1985. 9

[Wor76] Stanisław Lech Woronowicz. Positive maps of low dimensional matrix algebras. *Reports on Mathematical Physics*, 10(2):165–183, 1976. 2, 18

[ŽPBC07] Marko Žnidarič, Tomaž Prosen, Giuliano Benenti, and Giulio Casati. Detecting entanglement of random states with an entanglement witness. *Journal of Physics A: Mathematical and Theoretical*, 40(45):13787, 2007. 18

[ŻPNC11] Karol Życzkowski, Karol A Penson, Ion Nechita, and Benoit Collins. Generating random density matrices. *Journal of Mathematical Physics*, 52(6):062201, 2011. 17

[ŻS01] Karol Życzkowski and Hans-Jürgen Sommers. Induced measures in the space of mixed quantum states. *Journal of Physics A: Mathematical and General*, 34(35):7111, 2001. 12, 17

[ŻS03] Karol Życzkowski and Hans-Jürgen Sommers. Hilbert-schmidt volume of the set of mixed quantum states. *Journal of Physics A: Mathematical and General*, 36(39):10115, 2003. 17

CNRS, Laboratoire de Physique Théorique, IRSAMC, Université de Toulouse, UPS, F-31062 Toulouse, France

Email address: nechita@irsamc.ups-tlse.fr