Annales Henri Poincaré

Asymptotics of Random Density Matrices

Ion Nechita

Abstract. We investigate random density matrices obtained by partial tracing larger random pure states. We show that there is a strong connection between these random density matrices and the Wishart ensemble of random matrix theory. We provide asymptotic results on the behavior of the eigenvalues of random density matrices, including convergence of the empirical spectral measure. We also study the largest eigenvalue (almost sure convergence and fluctuations).

1. Introduction

Physicists and computer scientists working with finite size quantum systems are often interested in properties of *typical* states, such as entanglement, entropy, etc. In order to estimate such quantities, one has to endow the set of states (pure or mixed) with a certain probability measure and compute averages with respect to this measure. It has been known for a certain while now that there exists an "uniform" (in a way which will precised later) measure on the set $\mathcal{E}_n$ of pure states of size n. However, the situation is less simple when dealing with density matrices: there is no widely accepted candidate for a "canonical" measure on the set $\mathcal{D}_n$ of all density matrices of size n.

One may find in the literature two classes of probabilities on $\mathcal{D}_n$:

- the *induced measures*, where random density matrices are obtained by partial tracing a larger random pure state,
- the *metric measures*, where the measure is the volume element associated to a particular distance on $\mathcal{D}_n$.

Depending on the physical nature of the model, one may choose different measures from one class or the other. In this work we study the measures of the first class.

The induced measures were introduced by Braunstein [2] and studied later by Hall [5], Życzkowski and Sommers [13,15]. In the first part of this work we provide a rigorous construction of these measures. In the second part, we give some new

explicit and recurrence formulas for the moments and we study the asymptotic behavior of the spectrum of such random density matrices. Our approach is based on the connection with the well-known theory of *Wishart* random matrices.

Our paper is organized as follows. In Section 2 we recall the construction of the induced measures, adding mathematical rigor to the existing literature. Section 3 is devoted to recalling some results on the Wishart ensemble and making explicit the connection with random density matrices. We deduce the distribution of the eigenvalues and we study the moments. In Section 4 we study two models of large random density matrices, providing results on the behavior of the spectrum. A discussion of the results as well as ideas for generalizing our work are presented at the end of the paper.

2. From pure states to density matrices

We start by introducing and motivating the model of random density matrices that we consider.

As explained in the Introduction, one would like to endow the set of density matrices on a complex Hilbert space $\mathcal{H}$ with a "natural" probability measure. It turns out that there is no straightforward way of doing this, so one has to make some additional hypothesis in order for a probability measure to stand out as the most natural one. Our approach here is based on the definition of a density matrix as it is usually understood in the theory of open quantum systems. We consider that the system described by the density matrix is coupled to an environment and that the compound system is in a random pure state. More precisely, we shall make two assumptions:

(A1) The system $\mathcal{H}$ is *coupled* to an environment $\mathcal{K}$ and the compound system $\mathcal{H} \otimes \mathcal{K}$ is in a pure state $|\psi\rangle$.

(A2) The pure state $|\psi\rangle$ is *uniformly* distributed on the set of all pure states on $\mathcal{H} \otimes \mathcal{K}$.

The first assumption is motivated by a large class of models considered in physics or quantum information theory. The general framework is provided by a system $\mathcal{H} \otimes \mathcal{K}$ in a pure state, isolated from its environment. Suppose that one has access only to the sub-system $\mathcal{H}$. This may happen for several different reasons: $\mathcal{K}$ may be not accessible (e.g., $\mathcal{H}$ and $\mathcal{K}$ are in distant galaxies) or $\mathcal{K}$ may be too complicated to study (a heat bath or a noisy channel, for example). In these situations, it is natural to make the assumption (A1). Let us turn now to the second assumption. If one has no *a priori* information on the systems $\mathcal{H}$ and $\mathcal{K}$, it makes sense to suppose (A2). Moreover, it turns out that there exists an unique uniform probability measure on the set of pure states of given size, so we shall consider *uniform* pure states on the compound system.

However, there are situations when one of the two hypotheses (A1) or (A2) is not physically motivated. For instance, when one has no knowledge of an environment coupled to the system $\mathcal{K}$, there is no reason to suppose (A1). Instead, one

should use other probability measures, such as the *Bures measure* (see the discussion in [13]). On the other hand, even if (A1) corresponds to the physical reality, one may have extra information on the system $\mathcal{H}$ or $\mathcal{K}$ (or both). For example, it may be that the state of the environment $\mathcal{K}$ has a particular form; thus, it makes no sense to assume (A2) and our model would not be adapted to such situations.

In the next section, motivated by the assumption (A2), we shall construct the uniform measure on the set of pure states of given size. Then, by partial tracing, we shall provide the probability which verifies the assumptions (A1) and (A2).

2.1. The canonical probability measure on the pure states

In quantum mechanics, a pure state is described by a norm one vector in a n-dimensional complex vector space $\mathcal{H}$. The phase of pure states is not determined, i.e.,

$$|e^{i\theta}\psi\rangle = |\psi\rangle \quad \forall\, \theta \in \mathbb{R}\,. \tag{1}$$

In order to make this definition rigorous, we introduce the following equivalence relation on $\mathcal{H} \smallsetminus \{0\}$:

$$x \sim y \Leftrightarrow \exists\, \lambda \in \mathbb{C}^* \quad such\ that \quad x = \lambda y\,. \tag{2}$$

Definition 1. A pure state $|\psi\rangle$ is an element of the quotient space $(\mathcal{H} \smallsetminus \{0\})/\sim$. We denote by $\mathcal{E}_n$ the set of pure states of size n.

As all complex Hilbert spaces are isomorphic to $\mathbb{C}^n$, the set $\mathcal{E}_n$ is the set of rays in $\mathbb{C}^n$. We endow $\mathcal{E}_n$ with the associated quotient topology and the Borel σ-field. We now turn to the construction of the uniform probability measure on $\mathcal{E}_n$.

As stated in the assumption (A2), the probability on $\mathcal{E}_n$ should be the most *uniform* one, as there is no *a priori* information on the state $|\psi\rangle$. In particular, as there is no preferred basis of $\mathcal{H}$, the uniform measure should be invariant by changes of bases. In our framework ($\mathcal{H}$ is a complex Hilbert space), changes of bases are provided by unitary applications. As a consequence, we ask that the uniform probability measure should be unitarily invariant.

Definition 2. A measure ν_n on $\mathcal{E}_n$ is *unitarily invariant* if

$$\nu_n(UA) = \nu_n(A)\,,$$

for all unitary $U \in \mathcal{U}(n)$ and for all Borel subset $A \subset \mathcal{E}_n$.

It turns out that this condition is strong enough to completely specify a measure on $\mathcal{E}_n$, i.e., there is an unique unitarily invariant probability measure on $\mathcal{E}_n$. This follows from a well-known result in probability theory regarding group actions (see [7]). Let us recall it here.

Let G be a topological group acting on a topological space X. We call its action *transitive* if for all $x, y \in X$, there exists $g \in G$ such that $y = g \cdot x$ and *proper* if for all $g \in G$, the application $X \ni x \mapsto g \cdot x$ is proper, i.e., the pre-image of a compact set is compact. We then have the following

Theorem 1 ([7]). *Let G be a topological group which acts transitively and properly on a topological space X. Suppose that both G and X are locally compact and separable. Then there exists an unique (up to a constant) measure ν on X which is G-invariant.*

In order apply this result to our situation, we consider the action of the unitary group $\mathcal{U}(n)$ on the set $\mathcal{E}_n$ by left multiplication. We obtain the following proposition.

Proposition 1. *The action of $\mathcal{U}(n)$ on $\mathcal{E}_n$ is transitive and proper and thus there exists an* unique *unitarily invariant probability measure ν_n on $\mathcal{E}_n$.*

Proof. First of all, notice that this action is well defined: the class $|U\psi\rangle$ does not depend on ψ, but only on the class $|\psi\rangle$; we say that the multiplication by an unitary is a *class application*. In order to show that the action is transitive, consider two classes $|\psi\rangle$ and $|\varphi\rangle$ and an unitary $U \in \mathcal{U}(n)$ such that $U\psi = \varphi$ (such an unitary always exists). It follows then that $U|\psi\rangle = |\varphi\rangle$. Finally, the action is compact, as the set $\mathcal{E}_n$ is compact and the multiplication applications are continuous. Thus, the action verifies the hypothesis of Theorem 1, and as a consequence there is an unique unitarily invariant measure on $\mathcal{E}_n$. Moreover, given the compacity of $\mathcal{E}_n$ we can choose the measure of unit mass, which concludes the proof of the proposition. □

Existence and unicity being settled, one would like to dispose of more concrete descriptions on the distribution ν_n. It turns out that there are two ways of doing that.

First of all, let us recall the definition of a complex Gaussian random variable. Let X and Y be two independent real Gaussian random variables of mean 0 and variance $1/2$. Then $Z = X + iY$ is said to have a complex Gaussian distribution of mean 0 and variance 1. We denote by $\mathcal{N}_{\mathbb{C}}(0, 1)$ the law of Z. A complex vector $(Z_1, \ldots, Z_n)$ is said to have distribution $\mathcal{N}_{\mathbb{C}}^n(0, \mathrm{I}_n)$ if the random variables $Z_1, \ldots, Z_n$ are independent and have distribution $\mathcal{N}_{\mathbb{C}}(0, 1)$.

Consider now a complex Gaussian vector $X \sim \mathcal{N}_{\mathbb{C}}^n(0, \mathrm{I}_n)$ and the projection application

$$\Pi : \mathbb{C}^n \approx \mathcal{H} \to \mathcal{E}_n \tag{3}$$

$$x \mapsto |x\rangle . \tag{4}$$

It is well-known in probability theory that the law of X is unitarily invariant in $\mathbb{C}^n$. This property remains valid for the projection $\Pi(X)$ and thus the law of $|X\rangle$ is unitarily invariant on $\mathcal{E}_n$. As there is only one unitarily invariant distribution on $\mathcal{E}_n$, we have $|X\rangle \sim \nu_n$.

We can also obtain the law ν_n from another well-known probability distribution, the *Haar measure* on $\mathcal{U}(n)$. In order to do this, consider a Haar-distributed unitary matrix U. Obviously, the distribution of U is unitarily invariant; the same will hold true for the first column Y of U and for its class $|Y\rangle$. Thus $|Y\rangle$ has distribution ν_n. We sum up these results in the following

Proposition 2.   *1. Let X be a random complex vector of law $\mathcal{N}_{\mathbb{C}}^n(0, \mathrm{I}_n)$. Then the class $|X\rangle$ of X is distributed along ν_n.*

   2. Let U be a random unitary matrix distributed along the Haar measure on $\mathcal{U}(n)$ and let Y be the first column of U. Then the class $|Y\rangle$ has distribution ν_n.

2.2. The induced measure on density matrices

In this section we effectively construct the *induced measures* on density matrices that will be studied in the rest of the article. As stated in the Introduction, the induced measure of parameters n and k is obtained as follows:

- Consider a product space $\mathcal{H} \otimes \mathcal{K}$ of two complex Hilbert spaces $\mathcal{H}$ (of dimension n) and $\mathcal{K}$ – the environment – of dimension k. This is the global space *system + environment.*
- Take an uniform random pure state $|\psi\rangle$ on $\mathcal{H} \otimes \mathcal{K}$ (see the assumption (A2)).
- Consider the (random) pure density matrix $|\psi\rangle\langle\psi|$ corresponding to the pure state $|\psi\rangle$.
- Take $\rho = \mathrm{Tr}_{\mathcal{K}}(|\psi\rangle\langle\psi|)$, the partial trace of $|\psi\rangle\langle\psi|$ with respect to the environment $\mathcal{K}$. The law of the random variable ρ is the desired probability measure, which we shall note $\mu_{n,k}$.

As in our formalism $|\psi\rangle$ is an equivalence class, we shall define the pure density matrix $|\psi\rangle\langle\psi|$ by:

$$|\psi\rangle\langle\psi| = \frac{\psi \cdot \psi^*}{\mathrm{Tr}(\psi \cdot \psi^*)} \in \mathcal{M}_{nk}(\mathbb{C}). \tag{5}$$

Clearly, $\psi \mapsto |\psi\rangle\langle\psi|$ is a class function (it does not depend on the representant ψ chosen, but only on the class $|\psi\rangle$), so $|\psi\rangle\langle\psi|$ is well-defined. The normalizing factor $\mathrm{Tr}(\psi \cdot \psi^*)$ appears because we want the matrix $|\psi\rangle\langle\psi|$ to be trace one; this could have been avoided by considering a norm one vector ψ, since $\mathrm{Tr}(\psi \cdot \psi^*) = \|\psi\|^2$.

We now turn to the third step of the above construction and recall that the partial trace is the unique application $\mathrm{Tr}_{\mathcal{K}} : \mathcal{M}_{nk}(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ such that

$$\mathrm{Tr}\left((A \otimes \mathrm{I}_{\mathcal{K}})B\right) = \mathrm{Tr}\left(A \, \mathrm{Tr}_{\mathcal{K}}(B)\right) \quad \forall A \in \mathcal{M}_n(\mathbb{C}), \quad B \in \mathcal{M}_{nk}(\mathbb{C}). \tag{6}$$

Its expression for elementary matrices ($a_1, a_2 \in \mathcal{H}, b_1, b_2 \in \mathcal{K}$) is

$$\mathrm{Tr}_{\mathcal{K}}\left[(a_1 \otimes b_1) \cdot (a_2 \otimes b_2)^*\right] = \langle b_2, b_1\rangle \cdot a_1 a_2^*. \tag{7}$$

We have now all the elements needed for the definition of the induced measures:

Definition 3. The induced measure of parameters n and k is the distribution $\mu_{n,k}$ of the random density matrix

$$\rho = \mathrm{Tr}_{\mathcal{K}}(|\psi\rangle\langle\psi|), \tag{8}$$

where $|\psi\rangle$ is an uniform pure state on $\mathcal{H} \otimes \mathcal{K}$ of distribution ν_{nk}.

In order to get a better understanding of the measure $\mu_{n,k}$, we write ψ in an orthonormal basis $\{e_i \otimes f_j; 1 \leq i \leq n, 1 \leq j \leq k\}$ of $\mathcal{H} \otimes \mathcal{K}$:

$$\psi = \sum_{i=1}^n \sum_{j=1}^k \psi_{ij}\, e_i \otimes f_j \,. \tag{9}$$

Thus the matrix $|\psi\rangle\langle\psi|$ has coordinates (in the same basis):

$$|\psi\rangle\langle\psi|_{ij;i'j'} = \frac{\psi_{ij}\,\overline{\psi_{i'j'}}}{\sum_{\alpha=1}^n \sum_{\beta=1}^k |\psi_{\alpha\beta}|^2} \,. \tag{10}$$

After taking the partial trace, we obtain

$$\rho_{ii'} = \frac{\sum_{j=1}^k \psi_{ij}\,\overline{\psi_{i'j}}}{\sum_{\alpha=1}^n \sum_{\beta=1}^k |\psi_{\alpha\beta}|^2} \,. \tag{11}$$

Now, if we arrange the coordinates ψ_{ij} of ψ in a $n \times k$ matrix X such that $X(i,j) = \psi_{ij}$, we have

$$\rho = \frac{X \cdot X^*}{\mathrm{Tr}(X \cdot X^*)} \,. \tag{12}$$

Several important remarks should be made at this point. First of all, consider $U \in \mathcal{U}(n)$ and the density matrix ρ' obtained by replacing ψ with $(U \otimes \mathrm{I}_{\mathcal{K}})\psi$:

$$\rho' = \mathrm{Tr}_{\mathcal{K}}\left(|(U \otimes \mathrm{I}_{\mathcal{K}})\psi\rangle\langle(U \otimes \mathrm{I}_{\mathcal{K}})\psi|\right) . \tag{13}$$

By the properties of the partial trace, we have that $\rho' = U\rho U^*$. But recall that the law of $|\psi\rangle$ is unitarily invariant; it is thus invariant by $U \otimes \mathrm{I}_{\mathcal{K}}$ (which is an element of $\mathcal{U}(nk)$). Hence the law $\mu_{n,k}$ is invariant by unitary conjugation. Being positive, and thus self-adjoint, density matrices diagonalize:

$$\rho = VDV^* , \tag{14}$$

with V an unitary and D a diagonal matrix with positive entries. The unitary invariance of $\mu_{n,k}$ corresponds to the fact that V is distributed along the Haar measure on $\mathcal{U}(n)$. Remains, of course, the question of the distribution of D, the diagonal matrix of eigenvalues, which will be answered in Section 3.2 (see Proposition 4).

Another important question concerns the law of the matrix X. Recall that the coordinates of X are those of ψ, rearranged in a $n \times k$ matrix. Since the pure state $|\psi\rangle$ is distributed along the uniform measure ν_{nk}, we know, by the second point of Proposition 2, that we can take for ψ a complex Gaussian vector in $\mathbb{C}^{nk}$. Thus, the elements of X are independent, complex Gaussian random variables.

Lemma 1. *Let X be a $n \times k$ complex matrix such that the entries are independent identically distributed (i.i.d.) $\mathcal{N}_{\mathbb{C}}(0,1)$ random variables. Then, the matrix*

$$\rho = \frac{X \cdot X^*}{\mathrm{Tr}(X \cdot X^*)} \tag{15}$$

has distribution $\mu_{n,k}$.

This lemma motivates the study of matrices of type $W = X \cdot X^*$, which will be taken up in the next section.

3. Wishart matrices. Results at fixed size

3.1. The Wishart ensemble

This section is devoted to introducing the Wishart ensemble of random matrices. Introduced in the 1930's to study covariance matrices in statistics, Wishart matrices have found many applications, both theoretical (random matrix theory) and practical: principal component analysis, engineering, etc. Let us start by recalling the definition of the Wishart ensemble:

Definition 4. Let X be a $n \times k$ complex matrix such that the entries are i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$ random variables. The $n \times n$ matrix $W = X \cdot X^*$ is called a *Wishart random matrix* of parameters n and k.

In virtue of Lemma 1, there is a strong connection between the distribution of Wishart matrices and the random density matrices we study. More precisely, if W is a Wishart matrix, then

$$\rho = \frac{W}{\operatorname{Tr} W} \tag{16}$$

has distribution $\mu_{n,k}$.

We shall give a list of results on Wishart matrices that will be used later in the study of random density matrices. As the results are rather classical in random matrix theory, we will not supply proofs, but only references to the original papers.

We start with a result on the eigenvalues of a Wishart matrix. Being of the form $W = X \cdot X^*$, Wishart matrices are positive and thus they admit n non-negative eigenvalues $\lambda_1, \ldots, \lambda_n$. The next proposition provides the distribution of the random vector $(\lambda_1, \ldots, \lambda_n)$ (see [8]).

Proposition 3. *Let W be a random $n \times n$ Wishart matrix with parameters n and k. Then the distribution of the eigenvalues $(\lambda_1, \ldots, \lambda_n)$ has a density with respect to the Lebesgue measure on $\mathbb{R}_+^n$ which is given by*

$$\Phi_{n,k}^{\mathcal{W}}(\lambda_1, \ldots, \lambda_n) = C_{n,k}^{\mathcal{W}} \exp\left(-\sum_{i=1}^{n} \lambda_i\right) \prod_{i=1}^{n} \lambda_i^{k-n} \Delta(\lambda)^2, \tag{17}$$

where $C_{n,k}^{\mathcal{W}}$ is the constant $\left[\prod_{j=0}^{n-1} \Gamma(n+1-j)\Gamma(k-j)\right]^{-1}$ and

$$\Delta(\lambda) = \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j). \tag{18}$$

When studying large random matrices, one important question is to what resembles the spectrum of a random matrix in the limit $n \to \infty$? In order to

answer such a question, one introduces the *empirical spectral measure*

$$L_n(W) = \frac{1}{n}\sum_{i=1}^{n}\delta_{\lambda_i},\tag{19}$$

which is a random probability measure (it depends on W, which is random). It turns out that, almost surely, the random measures $L_n(W)$ converge to a deterministic probability measure, the Marchenko–Pastur distribution.

Definition 5. For $c \in]0,\infty[$, we denote by μ_c the *Marchenko–Pastur* probability measure given by the equation

$$\mu_c = \max\{1-c,0\}\delta_0 + \frac{\sqrt{(x-a)(b-x)}}{2\pi x}\mathbf{1}_{[a,b]}(x)dx,\tag{20}$$

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

The result is contained in the following theorem (see [4]).

Theorem 2. *Assume that $c \in]0,\infty[$, and let $(k(n))_n$ be a sequence of integers such that $\lim_{n\to\infty} k(n)/n = c$. Consider a sequence of random matrices $(W_n)_n$ such that for all n, W_n is a Wishart matrix of parameters n and $k(n)$. Define the renormalized empirical eigenvalue distribution of W_n by*

$$L_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{n^{-1}\lambda_i(W_n)},$$

where $\lambda_1(W_n),\ldots,\lambda_n(W_n)$ are the eigenvalues of W_n. Then, almost surely, the sequence $(L_n)_n$ converges weakly to the Marchenko–Pastur distribution μ_c.

Another object of interest in random matrix theory is the largest eigenvalue of a large random matrix. The next result shows that in the Wishart case, it converges almost surely to the right edge of the support of the Marchenko–Pastur distribution; similarly to the Central Limit Theorem, the nature of the fluctuations is known (see [1] and [6]).

Theorem 3. *Assume that $c \in]0,\infty[$, and let $(k(n))_n$ be a sequence of integers such that $\lim_{n\to\infty} k(n)/n = c$. Consider a sequence of random matrices $(W_n)_n$ such that for all n, W_n is a Wishart matrix of parameters n and $k(n)$, and let $\lambda_{max}(W_n)$ be the largest eigenvalue of W_n. Then, almost surely,*

$$\lim_{n\to\infty}\frac{1}{n}\lambda_{max}(W_n) = (\sqrt{c}+1)^2.\tag{21}$$

Moreover, the following limit holds in distribution

$$\lim_{n\to\infty}\frac{\lambda_{max}(W_n) - n(\sqrt{c}+1)^2}{n^{1/3}(1+\sqrt{c})(1+1/\sqrt{c})^{1/3}} = \mathcal{W}_2.\tag{22}$$

Here, $\mathcal{W}_2$ is the Tracy–Widom law of order 2; as even the definition of this probability distribution is well beyond the scope of this work, we encourage the reader to look it up in [14], the original paper of Tracy and Widom.

3.2. The spectrum of a density matrix

Recall from Section 2.2 that when considering the diagonalization of a random density matrix

$$\rho = VDV^*,\qquad(23)$$

the unitary matrix V is distributed along the Haar measure on the unitary group $\mathcal{U}(n)$. In this section we compute the distribution of the diagonal matrix D, i.e., the spectrum of a density matrix with distribution $\mu_{n,k}$.

Here, as well as in the next section, the parameters n and k will be fixed, and we shall assume that $k \geq n$. If $n > k$, by a property of the partial trace application, the matrix ρ will have $n - k$ null eigenvalues and k eigenvalues identical to those of the density matrix

$$\sigma = \mathrm{Tr}_{\mathcal{H}}(|\psi\rangle\langle\psi|)\,.$$

In consequence, the study of the spectrum of ρ is equivalent to the study of the spectrum of σ. Moreover, the size of σ's environment (n) is larger than the dimension of σ itself (k), and we can apply the first case. In conclusion, whenever n is larger than k, we interchange n and k, and we append $n - k$ null eigenvalues to the spectrum of ρ.

Recall that if W is a Wishart matrix of parameters n and k, then $\rho = W/\mathrm{Tr}\,W$ has distribution $\mu_{n,k}$. It follows that if $(\lambda_1, \ldots, \lambda_n)$ are the eigenvalues of W and $(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)$ are those of ρ, then we have

$$\tilde{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}\,,\quad \forall 1 \leq i \leq n\,.\qquad(24)$$

As the trace of a density matrix equals one, the (random) vector $(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)$ is confined in the $(n-1)$-dimensional probability simplex $\Sigma_{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\}$. Note that $\tilde{\lambda}_n$ is a function of $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{n-1}$, so we will show that the distribution of $(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{n-1})$ admits a density w.r.t. the Lebesgue measure on Σ_{n-1}.

Proposition 4. *The distribution of the (unordered) eigenvalues* $\tilde{\lambda}_1(\rho), \ldots, \tilde{\lambda}_{n-1}(\rho)$ *has a density with respect to the Lebesgue measure on* Σ_{n-1} *given by*

$$\Phi_{n,k}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{n-1}) = C_{n,k} \prod_{i=1}^n (\tilde{\lambda}_i)^{k-n} \Delta(\tilde{\lambda})^2\,,\qquad(25)$$

where

$$C_{n,k} = \frac{\Gamma(nk)}{\prod_{j=0}^{n-1} \Gamma(n+1-j)\Gamma(k-j)}\,.\qquad(26)$$

Remark 1. In the formula (25), there are only $n-1$ variables; $\tilde{\lambda}_n$ is not a variable, but merely the notation $\tilde{\lambda}_n = 1 - (\tilde{\lambda}_1 + \cdots + \tilde{\lambda}_{n-1})$.

Proof. Let us start from the Wishart distribution of eigenvalues and consider the change of variables

$$(\lambda_1, \ldots, \lambda_n) \mapsto (\lambda_1, \ldots, \lambda_{n-1}, S) \mapsto \tag{27}$$

$$\mapsto (\lambda_1/S, \ldots, \lambda_{n-1}/S, S) = (\tilde\lambda_1, \ldots, \tilde\lambda_{n-1}, S), \tag{28}$$

where $S = \sum_{i=1}^n \lambda_i$ is the sum of the Wishart eigenvalues. The Jacobian of this transformation equals $1/S^{n-1}$, and we get

$$\Phi_{n,k}^{(\tilde\lambda, S)}(\tilde\lambda_1, \ldots, \tilde\lambda_{n-1}, S) = C_{n,k}^{\mathcal{W}} \exp(-S) \prod_{i=1}^n (S\tilde\lambda_i)^{k-n} \Delta(S\tilde\lambda)^2 S^{n-1}. \tag{29}$$

We get now to the crucial point of the proof. Clearly, the above density factorizes as

$$\Phi_{n,k}^{(\tilde\lambda, S)}(\tilde\lambda_1, \ldots, \tilde\lambda_{n-1}, S) = C_{n,k}^{\mathcal{W}} \times \left[\prod_{i=1}^n \tilde\lambda_i^{k-n} \Delta(\tilde\lambda)^2 \right] \times \left[S^{nk-1} \exp(-S) \right]. \tag{30}$$

Hence, the normalized eigenvalues $(\tilde\lambda_1, \ldots, \tilde\lambda_{n-1})$ and the sum of the Wishart eigenvalues S are *independent* random variables.

In order to compute the distribution of $(\tilde\lambda_1, \ldots, \tilde\lambda_{n-1})$, it suffices to take the marginal with respect to S; we get

$$\Phi_{n,k}(\tilde\lambda_1, \ldots, \tilde\lambda_{n-1}) = C_{n,k} \prod_{i=1}^n \tilde\lambda_i^{k-n} \Delta(\tilde\lambda)^2, \tag{31}$$

where

$$C_{n,k} = C_{n,k}^{\mathcal{W}} \cdot \int_0^\infty S^{nk-1} e^{-S} dS = \Gamma(nk) C_{n,k}^{\mathcal{W}} \tag{32}$$

$$= \frac{\Gamma(nk)}{\prod_{j=0}^{n-1} \Gamma(n+1-j)\Gamma(k-j)}. \tag{33}$$

$\square$

As a byproduct of the proof, we also obtain the following characterization of the induced measure.

Corollary 1. *The law of a random density matrix ρ of parameters n and k is the law of a Wishart matrix W of the same parameters* conditioned *by* $\operatorname{Tr} W = 1$.

Proof. From the formula (30) we see that the normalized eigenvalues and the trace of a Wishart matrix are independent random variables. Thus, taking the marginal with respect to the trace is equivalent to conditioning on the event $(\operatorname{Tr} W = 1)$. Note however that $(\operatorname{Tr} W = 1)$ has zero probability. $\square$

In Figure 1 we have plotted the density functions for $n = 2$ and several values of k using the analytic formula (25). For $n = 3$ we have randomly generated random density matrices and plotted the probability simplex Σ_2 along with the points corresponding to the spectra (Figure 2). We notice that for large values of k
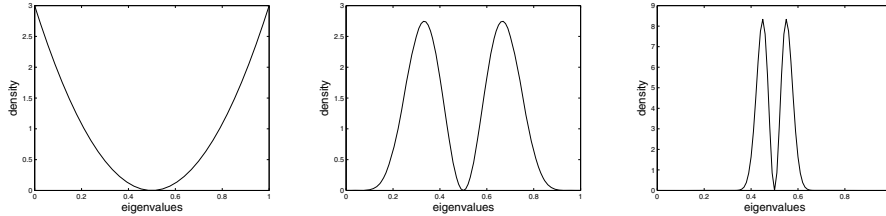
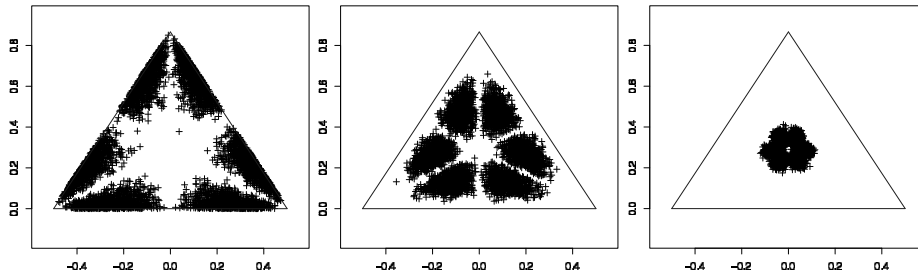FIGURE 1. Theoretical eigenvalue distribution for $(n = 2, k = 2)$, $(n = 2, k = 10)$ and $(n = 2, k = 50)$.

FIGURE 2. Empirical eigenvalue distribution (5000 matrices) for $(n = 3, k = 3)$, $(n = 3, k = 10)$ and $(n = 3, k = 100)$.

(the size of the environment), the spectrum concentrates to the middle point in Σ_{n-1}. This is a general phenomenon and it will be studied in Section 4.1.

3.3. Moments

The aim of this section is to provide formulas for the moments of order q of a random density matrix of distribution $\mu_{n,k}$. In order to do that, we shall introduce the some notation: $\mathbb{E}_{n,k}[\cdot]$ will denote the expectation with respect to the law $\mu_{n,k}$ and $\mathbb{E}_{n,k}^{\mathcal{W}}[\cdot]$ the expectation with respect to the law of Wishart matrices with parameters n and k. We will use the corresponding result on the Wishart ensemble and derive explicit formulas, as well as recurrence relations. The following proposition provides a bridge between the moments of a density matrix and those of a Wishart matrix with the same parameters.

Proposition 5. *Let $\mathbb{E}_{n,k}[\mathrm{Tr}(\rho^q)]$ be the moment of a random density matrix of parameters n and k and let $\mathbb{E}_{n,k}^{\mathcal{W}}[\mathrm{Tr}(W^q)]$ be the moment of a Wishart matrix having the same parameters. Then,*

$$\mathbb{E}_{n,k}\big[\mathrm{Tr}(\rho^q)\big] = \frac{\mathbb{E}_{n,k}^{\mathcal{W}}\big[\mathrm{Tr}(W^q)\big]}{nk(nk+1)\cdots(nk+q-1)}. \tag{34}$$

Proof. By using the same technique as in the proof of Proposition 4, we get

$$\mathbb{E}_{n,k}^{\mathcal{W}}\big[\operatorname{Tr}(W^q)\big] = \mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^q)\big]\frac{\Gamma(nk+q)}{\Gamma(nk)}\,, \tag{35}$$

which is the same as equation (34).   □

We can find in the literature different explicit and recurrence formulas for $\mathbb{E}_{n,k}^{\mathcal{W}}[\operatorname{Tr}(W^q)]$. From the one in [4], we get

$$\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^q)\big] = \frac{\Gamma(nk)}{\Gamma(nk+q)}\sum_{j=1}^{q}(-1)^{j-1}\frac{[k+q-j]_q[n+q-j]_q}{(q-j)!(j-1)!}\,, \tag{36}$$

where $[a]_q = a(a-1)\cdots(a-q+1)$. The recurrence formula (see [4])

$$\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^q)\big] = \frac{(2q-1)(n+k)}{(nk+q-1)(q+1)}\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^{q-1})\big] \tag{37}$$

$$+ \frac{(q-2)\big((q-1)^2-(k-n)^2\big)}{(nk+q-1)(nk+q-2)(q+1)}\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^{q-2})\big] \tag{38}$$

allows us to easily compute some averages:

$$\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^2)\big] = \frac{n+k}{nk+1}\,, \tag{39}$$

$$\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^3)\big] = \frac{n^2+3nk+k^2+1}{(nk+1)(nk+2)}\,, \tag{40}$$

$$\mathbb{E}_{n,k}\big[\operatorname{Tr}(\rho^4)\big] = \frac{n^3+6n^2k+6nk^2+k^3+5n+5k}{(nk+1)(nk+2)(nk+3)}\,, \quad etc. \tag{41}$$

These formulas are consistent with the ones of [13] and [15].

4. Asymptotics

The last part of this paper is devoted to the study of random density matrices corresponding to *large systems*. We shall consider two models, both motivated physically:

1. In the first model, the size of the density matrix n is constant and the size of the environment k tends to infinity. Such a situation arises typically when one studies a small system (a qubit, a pair of qubits, etc.) coupled to a much larger environment. We show that in the limit $k \to \infty$, density matrices distributed along $\mu_{n,k}$ converge to the maximally mixed (or chaotic) state I/n.
2. In the second model, both n and k tend to infinity and $k/n \to c > 0$. This model describes a large system coupled to a large environment with constant ratio of size $(\dim\mathcal{K}/\dim\mathcal{H} \approx c)$. In this case we show that the spectral measures of density matrices of law $\mu_{n,k}$ converge to a deterministic measure known in random matrix theory as the *Marchenko–Pastur distribution* (see

Definition 5). We also study the convergence and the fluctuations of the largest eigenvalue of random density matrices.

4.1. The first model

Consider the density function of $\mu_{n,k}$ with n fixed and $k \to \infty$:

$$\Phi_{n,k}(\lambda_1, \ldots, \lambda_{n-1}) = C_{n,k} \prod_{i=1}^{n} \lambda_i^{k-n} \Delta(\lambda)^2 \,. \tag{42}$$

As n is fixed, the Vandermonde factor $\Delta(\lambda)$ is constant. The other factor, properly normalized in order to get a probability density, is the Dirichlet measure of parameter $\alpha = k - n + 1$:

$$\Phi_{n,k}^{'}(\lambda_1, \ldots, \lambda_{n-1}) = C_{n,k}^{'} \prod_{i=1}^{n} \lambda_i^{\alpha-1} \,. \tag{43}$$

The next result is well-known in probability theory. We shall sketch its proof for the sake of completeness.

Theorem 4. *The Dirichlet measure converges weakly as $\alpha \to \infty$ to the Dirac measure $\delta_{(1/n,\ldots,1/n)}$*

Proof. The idea behind the proof is to show that the variance of a Dirichlet-distributed random variable converges to 0 as its parameter converges to infinity. Let X be such a random variable. X has a density with respect to the Lebesgue measure on the probability simplex given by:

$$f(x_1, \ldots, x_n) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} \prod_{i=1}^{n} x_i^{\alpha-1} \,.$$

It is easy to compute

$$\mathbb{E}\left[\left\|X - \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)\right\|^2\right] = n\mathbb{E}\left[x_1^2 - \frac{2x_1}{n} + \frac{1}{n^2}\right] = \frac{\alpha+1}{n\alpha+1} - \frac{1}{n} \to 0 \,. \tag{44}$$

$\square$

As the maximally mixed state I/n is the unique state having spectrum $\{1/n, \ldots, 1/n\}$, we get:

Corollary 2. *Density matrices of the first model converge almost surely to the maximally mixed (or chaotic) state I/n.*

Remark 2. The same result can be obtained by an entropic argument. It turns out that the mean von Neumann entropy $S(\rho) = -\operatorname{Tr}(\rho \log \rho)$ can be computed for a random density matrix distributed along $\mu_{n,k}$:

$$\mathbb{E}_{n,k}\big[S(\rho)\big] = \sum_{i=k+1}^{nk} \frac{1}{i} - \frac{n-1}{2k} \,.$$

This formula has been conjectured by Page [9] and has been subsequently proved (see [11, 12]) using various methods. Let us explain how it implies the corollary. First, fix n and let k grow to infinity, as in our model. The mean entropy is easily seen to converge to $\log n$. This turns out to be the maximum von Neumann entropy for a system with n degrees of freedom. It is attained at the state I/n, the unique state of maximum uncertainty.

4.2. The second model

In the second model, both the size of the density matrix and the size of the environment tend to infinity. In order to use the results on the Wishart ensemble (Theorems 2 and 3), we need appropriate results on the behavior of the trace S of a Wishart matrix.

Lemma 2. *Assume that $c \in]0, \infty[$, and let $(k(n))_n$ be a sequence of integers such that $\lim_{n \to \infty} k(n)/n = c$. Consider a sequence of random matrices $(W_n)_n$ such that for all n, W_n is a Wishart matrix of parameters n and $k(n)$. Let $S_n = \operatorname{Tr} W_n$ be the trace of W_n. Then*

$$\frac{S_n}{nk(n)} \to 1 \quad \text{almost surely} \tag{45}$$

and

$$\frac{S_n - nk(n)}{\sqrt{nk(n)}} \Rightarrow \mathcal{N}(0, 1)\,, \tag{46}$$

where "$\Rightarrow$" denotes the convergence in distribution.

Proof. Recall that $W_n = X_n \cdot X_n^*$, when X_n is a $n \times k(n)$ matrix with i.i.d. complex Gaussian entries. We have

$$S_n = \sum_{i=1}^{n} \sum_{j=1}^{k(n)} |X_{ij}|^2 = \sum_{i=1}^{n} \sum_{j=1}^{k(n)} \left(\operatorname{Re}(X_{ij})^2 + \operatorname{Im}(X_{ij})^2 \right). \tag{47}$$

The random variables $\{\operatorname{Re}(X_{ij}), \operatorname{Im}(X_{ij})\}_{ij}$ are i.i.d. with distribution $\mathcal{N}(0, 1/2)$ and thus, by the law of large numbers, we have, almost surely,

$$\lim_{n \to \infty} \frac{S_n}{2nk(n)} = \frac{1}{2}\,, \tag{48}$$

completing the proof of the first result. The second result follows from the Central Limit Theorem. $\qquad \square$

We can now state and prove the analogue of Theorem 2 for random density matrices:

Theorem 5. *Assume that $c \in]0, \infty[$, and let $(k(n))_n$ be a sequence of integers such that $\lim_{n \to \infty} k(n)/n = c$. Consider a sequence of random density matrices $(\rho_n)_n$ such that for all n, ρ_n has distribution $\mu_{n,k(n)}$. Define the renormalized empirical distribution of ρ_n by*

$$L_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{cn\lambda_i(\rho_n)}\,, \tag{49}$$

where $\lambda_1(\rho_n), \cdots, \lambda_n(\rho_n)$ are the eigenvalues of ρ_n. Then, almost surely, the sequence $(L_n)_n$ converges weakly to the Marchenko–Pastur distribution μ_c.

Proof. We know (Theorem 2) that the empirical distribution of eigenvalues for the Wishart ensemble

$$L_n^{\mathcal{W}} = \frac{1}{n} \sum_{i=1}^{n} \delta_{n^{-1}\lambda_i(W_n)}, \qquad (50)$$

converges almost surely to the Marchenko–Pastur distribution of parameter c. Recall that the eigenvalues of the density matrix $\rho_n = W_n / \operatorname{Tr}(W_n)$ are those of W_n divided by the trace S_n of W_n. We have thus the following formula for the empirical spectral measure of ρ:

$$L_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{cn\lambda_i(W_n)/S_n} = \frac{1}{n} \sum_{i=1}^{n} \delta_{n^{-1}\lambda_i(W_n) \cdot \frac{cn^2}{S_n}}. \qquad (51)$$

The last equation is the same as equation (50) with the Dirac measures perturbed by a factor of cn^2/S_n which converges, almost surely, to 1 (by the preceding lemma). We are now going to show that such a perturbation does not change the limit in distribution. In order to achieve this, recall that when the limit measure is compactly supported, the convergence in distribution is equivalent to the convergence of moments. If we compute the q-th moment of the measures $L_n^{\mathcal{W}}$ and L_n, we find:

$$\langle x^q, L_n^{\mathcal{W}} \rangle = \frac{1}{n} \sum_{i=1}^{n} \left(n^{-1}\lambda_i(W_n) \right)^q, \qquad (52)$$

and, respectively,

$$\langle x^q, L_n \rangle = \frac{1}{n} \sum_{i=1}^{n} \left(n^{-1}\lambda_i(W_n) \right)^q \cdot \left(\frac{cn^2}{S_n} \right)^q. \qquad (53)$$

These expressions have the same limit as $n \to \infty$ for all q, and thus L_n converges to the Marchenko–Pastur distribution. $\qquad \square$

In Figure 3, we have plotted for several values of c and large n and k a histogram of the spectrum for *one* density matrix and the theoretical density of the Marchenko–Pastur distribution (see Remark 3). We can see that the empirical histogram matches closely the theoretical curve for rather mild values of n (here $n = 1000$).

We now turn to the study of the largest eigenvalue of random density matrices. As before, we use the corresponding result on the Wishart ensemble (Theorem 3) and the control over the trace (Lemma 2):

Theorem 6. *Assume that $c \in]0, \infty[$, and let $(k(n))_n$ be a sequence of integers such that $\lim_{n \to \infty} k(n)/n = c$. Consider a sequence of random matrices $(\rho_n)_n$ such that for all n, ρ_n has distribution $\mu_{n,k(n)}$, and let $\lambda_{max}(\rho_n)$ be the largest eigenvalue of ρ_n. Then, almost surely,*

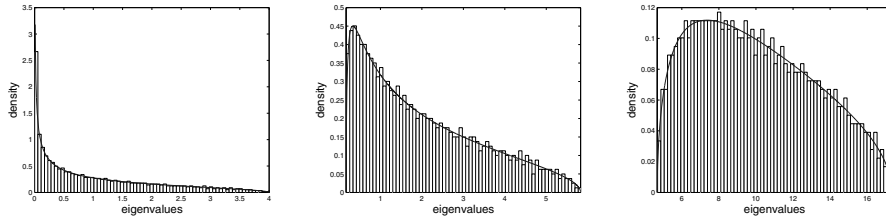$$\lim_{n \to \infty} cn\lambda_{max}(\rho_n) = (\sqrt{c} + 1)^2. \qquad (54)$$

FIGURE 3. Empirical and limit measures for $(n = 1000, k = 1000)$, $(n = 1000, k = 2000)$ and $(n = 1000, k = 10000)$.

Moreover,

$$\lim_{n \to \infty} \frac{n^{2/3}\left[cn\lambda_{max}(\rho_n) - (\sqrt{c} + 1)^2\right]}{(1 + \sqrt{c})(1 + 1/\sqrt{c})^{1/3}} = \mathcal{W}_2 \quad \text{in distribution}. \tag{55}$$

Proof. By the first part of Theorem 3, the (normalized) largest eigenvalue $\frac{1}{n}\lambda_{max}(W_n)$ of a Wishart matrix converges almost surely to $(\sqrt{c}+1)^2$. Obviously, we have

$$\lambda_{max}(\rho_n) = \frac{\lambda_{max}(W_n)}{S_n}, \tag{56}$$

and, by Lemma 2, $S_n/(cn^2)$ converges (almost surely) to 1. Finally, we obtain formula (54).

For the second part of the theorem, what we need to do, normalizations apart, is to show that the trace of a Wishart matrix fluctuates less than the largest eigenvalue. For the Wishart case, we have

$$\lambda_{max}(W_n) = n\left(\sqrt{c} + 1\right)^2 + n^{1/3}\left(1 + \sqrt{c}\right)\left(1 + 1/\sqrt{c}\right)^{1/3}\left(\mathcal{W}_2 + \mathrm{o}(1)\right), \tag{57}$$

and

$$S_n = nk(n) + \sqrt{nk(n)}\left(\mathcal{N} + \mathrm{o}(1)\right). \tag{58}$$

Again, $\lambda_{max}(\rho_n) = \lambda_{max}(W_n)/S_n$ and after simplifications, one obtains the desired formula (55). $\qquad\square$

Remark 3. Note that Theorem 5 and the first part of Theorem 6 deal with *almost sure* convergences. This means that when considering sequences of random density matrices of increasing size, the respective convergences will fail only on a set of null measure. This is to be compared with typicality results for random density matrices obtained recently in [3, 10] by concentration of measure techniques. These results give bounds (at fixed matrix size) on the probability that a random matrix is far from its expected value, while our results deal with the more subtle convergence of rescaled quantities, such as the spectral distribution or the largest eigenvalue.

5. Conclusions

We investigated random density matrices distributed along the so-called *induced measures*. After introducing them as partial traces of larger random pure states, we provided some explicit and recurrence relations for the moments of such density matrices. Using results on Wishart matrices, we then considered large density matrices. In a first model, a fixed size system was coupled to a very large environment; we showed that an uniform pure state on the compound system corresponds to the maximally mixed (or chaotic) density matrix on the fixed-size system. In parallel with Wishart matrices, we studied the regime $\dim \mathcal{K} / \dim \mathcal{H} \to c$. We obtained the almost sure convergence of the empirical spectral measure and of the largest eigenvalue, as well as the fluctuations of the largest eigenvalue. Results from random matrix theory were easily adapted for density matrices. Other important quantities, such as correlation functions, require a more detailed analysis, and this will be the subject of further work. Also, it may be interesting to study such asymptotics for other probability measures on density matrices, such as the Bures measure.

Acknowledgements

References

[1] Z. Bai, *Methodologies in spectral analysis of large-dimensional random matrices, a review*, Statist. Sinica **9** (1999), no. 3, pp. 611–677.

[2] S. Braunstein, *Geometry of quantum inference*, Phys. Lett. A **219** (1996), no. 3–4, pp. 129–174.

[3] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi, *Canonical typicality*, Phys. Rev. Lett. **96** (2006), no. 5, pp. 050403.

[4] U. Haagerup and S. Thorbjørnsen, *Random matrices with complex Gaussian entries*, Expo. Math. **21** (2003), no. 4, pp. 293–335.

[5] M. Hall, *Random quantum correlations and density operator distributions*, Phys. Lett. A **242** (1998), no. 3, pp. 123–129.

[6] I. Johnstone, *On the distribution of the largest eigenvalue in principal components analysis*, Ann. Statist. **29** (2001), no. 2, pp. 295–327.

[7] O. Kallenberg, *Foundations of modern probability*, Springer-Verlag, 2002.

[8] M. Mehta, *Random matrices*, Elsevier/Academic Press, 2004.

[9] D. Page, *Average entropy of a subsystem*, Phys. Rev. Lett. **71** (1993), no. 9, pp. 1291–1294.

[10] S. Popescu, A. J. Short and A. Winter, *The foundations of statistical mechanics from entanglement: Individual states vs. averages*, `arXiv:quant-ph/0511225v3`.

[11] J. Sánchez-Ruiz, *Simple proof of Page's conjecture on the average entropy of a subsystem*, Phys. Rev. E **52** (1995), no. 5, pp. 5653–5655.

[12] S. Sen, *Average entropy of a quantum subsystem*, Phys. Rev. Lett. **77** (1996).

[13] H.-J. Sommers and K. Życzkowski, *Statistical properties of random density matrices*, J. Phys. A **37** (2004), no. 35, pp. 8457–8466.

[14] C. Tracy and H. Widom, *On orthogonal and symplectic matrix ensembles*, Comm. Math. Phys. **177** (1996), no. 3, pp. 727–754.

[15] K. Życzkowski and H.-J. Sommers, *Induced measures in the space of mixed quantum states*, J. Phys. A **34** (2001), no. 35, pp. 7111–7125.

Ion Nechita
Université Lyon 1
CNRS, UMR 5208
Institut Camille Jordan
Bâtiment du Doyen Jean Braconnier
43, bd du 11 novembre 1918
F-69622 Villeurbanne Cedex
France
e-mail: `nechita@math.univ-lyon1.fr`