Strong asymptotic freeness for Wigner and Wishart matrices

Mireille Capitaine, Catherine Donati-Martin

Strong asymptotic freeness for Wigner and Wishart matrices

M. Capitaine[*] and C. Donati-Martin[†]

Abstract

For each n in $\mathbb{N}$, let $X_n = [(X_n)_{jk}]_{j,k=1}^n$ be a random Hermitian matrix such that the n^2 random variables $\sqrt{n}(X_n)_{ii}$, $\sqrt{2n}Re((X_n)_{ij})_{i<j}$, $\sqrt{2n}Im((X_n)_{ij})_{i<j}$ are independent identically distributed with common distribution μ on $\mathbb{R}$. Let $X_n^{(1)}, \ldots, X_n^{(r)}$ be r independent copies of X_n and $(x_1, \ldots x_r)$ be a semicircular system in a $\mathcal{C}^*$-probability space. Assuming that μ is symmetric and satisfies a Poincaré inequality, we show that, almost everywhere, for any non commutative polynomial p in r variables,

$$\lim_{n \longrightarrow +\infty} ||p(X_n^{(1)}, \ldots, X_n^{(r)})|| = ||p(x_1, \ldots x_r)|| \ . \qquad (0.1)$$

We follow the method of [9] and [15] which gave (0.1) in the Gaussian (complex, real or symplectic) case. We also get that (0.1) remains true when the $X_n^{(i)}$ are Wishart matrices while the x_i are Marchenko-Pastur distributed.

Mathematics Subject Classification (2000): 15A52, 46L54, 60F99.
Key words: Random matrices, free probability, asymptotic freeness.

[*]Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex. E-mail: capitain@cict.fr

[†]Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 6, Site Chevaleret, 13 rue Clisson, F-75013 Paris. E-mail: donati@ccr.jussieu.fr

1 Introduction

In the 90's, Voiculescu [18] introduced a random matrix model for a free semi-circular system. He showed that if we take r independent random matrices $(X_n^{(i)})_{i=1,\dots r}$, distributed as $GUE(n, \frac{1}{n})$, then, they are asymptotically free, that is, for every non commutative polynomial p in r variables,

$$\mathbb{E}[\mathrm{tr}_n\, p(X_n^{(1)}, \dots, X_n^{(r)})] \underset{n \longrightarrow \infty}{\longrightarrow} \tau(p(x_1, \dots x_r)) \tag{1.1}$$

where tr_n stands for the normalized trace on $M_n(\mathbb{C})$ and $(x_1, \dots x_r)$ is a free family of semicircular variables in some non commutative probability space $(\mathcal{B}, \tau)$. The result (1.1) holds true for a family of iid Wigner matrices and is proved by Dykema in [6].

In a recent paper, Haagerup and Thorbjørnsen [9] proved a strong version of (1.1), in the GUE case, namely a convergence for the operator norm:

$$\lim_{n \longrightarrow +\infty} ||p(X_n^{(1)}, \dots, X_n^{(r)})|| = ||p(x_1, \dots x_r)|| \ a.s. \tag{1.2}$$

which led to the proof that $Ext(C^*_{red}(F_2))$ is not a group.

Schultz [15] obtained the same result for Gaussian random matrices in the real case (GOE) and in the simplectic case (GSE). Our aim is to extend (1.2) in the case of an independent family of Wigner matrices on one hand and in the case of Wishart matrices on the other hand . Note that the special case $r = 1$ gives the well known convergence of the largest eigenvalue of $X_n^{(1)}$ to the right boundary of the support of x_1 (see [3] for the Wigner case and [7] for the Wishart case; see also [2] and the references therein).

Our approach is very similar to that of [9] and [15]. Therefore, we will recall the main lines of their proofs. First, in proving (1.2), the minoration

$$\liminf_{n \longrightarrow +\infty} ||p(X_n^{(1)}, \dots, X_n^{(r)})|| \geq ||p(x_1, \dots x_r)|| \ a.s.$$

comes rather easily from an a.s. version of (1.1) (obtained in [17] for the GUE case and proved in Section 6 of [15] for the GOE case) (see Lemma 7.2 in [9]). So, the main difficulty is the proof of the reverse inequality:

$$\limsup_{n \longrightarrow +\infty} ||p(X_n^{(1)}, \dots, X_n^{(r)})|| \leq ||p(x_1, \dots x_r)|| \ a.s. \tag{1.3}$$

In the following, we sketch the main steps in the proof of (1.3).

Step 1: A linearisation trick (see [9], Section 2 and Proposition 7.3)
In order to prove (1.3), it is sufficient to prove:

Lemma 1.1 *For all $m \in \mathbb{N}$, all self-adjoint matrices $a_0, \ldots, a_r$*[1] *of size $m \times m$ and all $\epsilon > 0$,*

$$sp(a_0 \otimes 1_n + \sum_{i=1}^{r} a_i \otimes X_n^{(i)}(\omega)) \subset sp(a_0 \otimes 1_{\mathcal{B}} + \sum_{i=1}^{r} a_i \otimes x_i) +]-\epsilon, \epsilon[\quad (1.4)$$

eventually, as $n \longrightarrow \infty$ a.e. in ω. Here, $sp(T)$ denotes the spectrum of the operator T and 1_n the identity matrix.

The analysis of the spectrum of $S_n := a_0 \otimes 1_n + \sum_{i=1}^{r} a_i \otimes X_n^{(i)}$ is done, using the Stieljes transform

$$G_n(\lambda) = \mathbb{E}[(id_m \otimes \mathrm{tr}_n)[(\lambda \otimes 1_n - S_n)^{-1}]], \lambda \in M_m(\mathbb{C}),\ Im(\lambda) \text{ positive definite.} \quad (1.5)$$

The proof of (1.4) requires sharp estimates of the rate of convergence of $G_n(\lambda)$ to $G(\lambda) := (id_m \otimes \tau)[(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}]$ (of order $1/n^2$) where $s = a_0 \otimes 1_{\mathcal{B}} + \sum_{p=1}^{r} a_p \otimes x_p$.

Step 2: In the GUE case, Haagerup and Thorbjørnsen [9] obtains the following estimate

$$||G_n(\lambda) - G(\lambda)|| \leq \frac{C(\lambda)}{n^2}. \quad (1.6)$$

In the GOE, GSE cases, Schultz [15] gets an extra term of order $1/n$, namely

$$||G_n(\lambda) - G(\lambda) - \frac{L(\lambda)}{n}|| \leq \frac{C(\lambda)}{n^2} \quad (1.7)$$

for some functional L.

Step 3 From the previous step, it is shown in section 6 of [9] that

$$\mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))] = (\mathrm{tr}_m \otimes \tau)(\varphi(s)) + O(\frac{1}{n^2}) \quad (1.8)$$

for φ smooth with compact support, and

$$\mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))] = O(\frac{1}{n^2}) \quad (1.9)$$

for φ smooth, constant outside a compact set and such that $supp(\varphi) \cap sp(s) = \emptyset$. In the GOE case (resp. GSE case), Schultz proved in section 5 of [15] that

$$\mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))] = (\mathrm{tr}_m \otimes \tau)(\varphi(s)) + \frac{1}{n}\Lambda(\varphi) + O(\frac{1}{n^2}) \quad (1.10)$$

[1] By a density argument, we can also assume that the matrices a_i are invertible.

where Λ is a distribution with compact support in $sp(s)$ with Stieljes transform

$$f(\lambda) = \mathrm{tr}_m(L(\lambda 1_m)), \lambda \in \mathbb{C} \setminus \mathbb{R}.$$

Therefore, (1.9) still holds for φ with $supp(\varphi) \cap sp(s) = \emptyset$.

Step 4 (1.9), combining with a Gaussian variance estimate, yields (by a standard application of the Borel Cantelli lemma),

$$(\mathrm{tr}_m \otimes \mathrm{tr}_n) 1_F(S_n) = O(n^{-4/3})$$

for $F = \{t \in \mathbb{R}, d(t, sp(s)) \geq \epsilon\}$ which leads to (1.4).

The main difficulties in the generalization of the above to Wigner or Wishart matrices arise in step 2. Indeed, we don't have the gaussian integration by parts' formula anymore. Our approach is inspired by the work of [12] where they use a Taylor expansion (see Lemma 4.1) extending the gausian integration by parts' formula. The remainder of the proof can be completed essentially as in the GOE/GSE case. Hence, in this paper, we shall focus on the obtention of such a master inequality

$$||G_n(\lambda) - G(\lambda) - \frac{1}{n} L(\lambda)|| = O(\frac{1}{n^2})$$

in the case of a family of Hermitian matrices with symmetric iid entries satisfying a Poincaré inequality, as well as in the case of Wishart matrices; we just give some hints when the computations are similar to that of [9], [15].

The paper is organized as follows. In section 2, we introduce notations and preliminaries which will be of basic use later on. In section 3, we describe the proof of (1.8) and (1.10) proved respectively in [9] and [15] in order to make clear the validity of the method in our general framework we state in section 4 (for the Wigner case) and section 5 (for the Wishart case).

2 Notations and preliminaries

This section may contain some definitions already used in the introduction but we choose to gather all the notations in this section for the reader's convenience. To begin with, we introduce some notations on the set of matrices.

- $M_p(\mathbb{C})$ is the set of $p \times p$ matrices with complex entries, $M_p(\mathbb{C})_{sa}$ the subset of self-adjoint elements of $M_p(\mathbb{C})$ and 1_p the identity matrix. In the following, we shall consider two sets of matrices with $p = m$ (m fixed) and $p = n$ with $n \longrightarrow \infty$.

- Tr_p denotes the trace and $\mathrm{tr}_p = \frac{1}{p}\mathrm{Tr}_p$ the normalized trace on $M_p(\mathbb{C})$.

- $||.||$ denotes the operator norm on $M_p(\mathbb{C})$ and $||M||_2 = (\mathrm{Tr}_p(M^*M))^{1/2}$ the Hilbert-Schmidt norm.

- Let $(E_{ij})_{i,j=1}^n$ be the canonical basis of $M_n(\mathbb{C})$ and define a basis of the real vector space of the self-adjoint matrices $M_n(\mathbb{C})_{sa}$ by:

$$\begin{aligned} e_{jj} &= E_{jj}, 1 \leq j \leq n \\ e_{jk} &= \frac{1}{\sqrt{2}}(E_{jk} + E_{kj}), 1 \leq j < k \leq n \\ f_{jk} &= \frac{i}{\sqrt{2}}(E_{jk} - E_{kj}), 1 \leq j < k \leq n \end{aligned}$$

- For a matrix M in $M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$, we denote by

$$M_{ij} := (id_m \otimes \mathrm{Tr}_n)(M(1_m \otimes E_{ji})) \in M_m(\mathbb{C}),\ 1 \leq i, j \leq n$$

and

$$_{\alpha,\beta}M := (\mathrm{Tr}_m \otimes id_n)(M(\hat{E}_{\beta,\alpha} \otimes 1_n)) \in M_n(\mathbb{C}),\ 1 \leq \alpha, \beta \leq m$$

where $(\hat{E}_{\alpha,\beta})$ is the canonical basis of $M_m(\mathbb{C})$.

We now define our matrix model and the random variables of interest.

- $(X_n^{(1)}, \ldots, X_n^{(r)})_{i=1,\ldots r}$ is a set of iid random matrices in $M_n(\mathbb{C})_{sa}$, whose distribution will be specified later (matrices in GUE or GOE in section 3, Wigner matrices in Section 4, Wishart matrices in section 5).

- For a given family $a_0, \ldots a_r$ in $M_m(\mathbb{C})_{sa}$, we define the random variable S_n with values in $M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$ by:

$$S_n = a_0 \otimes 1_n + \sum_{p=1}^{r} a_p \otimes X_n^{(p)} \tag{2.1}$$

and $s \in M_m(\mathbb{C}) \otimes \mathcal{B}$ by

$$s = a_0 \otimes 1_{\mathcal{B}} + \sum_{p=1}^{r} a_p \otimes x_p \tag{2.2}$$

where the $(x_i)_{i=1,\dots r}$ is a free family of self-adjoint operators in a C^* probability space $(\mathcal{B}, \tau)$ with a faithful state τ, whose distribution will be specified in the different cases (semi-circular in sections 3 and 4 or distributed as the Marchenko-Pastur distribution in section 5).

- For any matrix λ in $\mathcal{O}$ where

$$\mathcal{O} := \{\lambda \in M_m(\mathbb{C}) | Im(\lambda) \text{ is positive definite}\},$$

we define the $M_m(\mathbb{C})$ valued rv:

$$H_n(\lambda) = (id_m \otimes \mathrm{tr}_n)[(\lambda \otimes 1_n - S_n)^{-1}], \tag{2.3}$$

$$G_n(\lambda) = \mathbb{E}[H_n(\lambda)] \tag{2.4}$$

and

$$G(\lambda) = (id_m \otimes \tau)[(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}]. \tag{2.5}$$

For $\lambda \in \mathbb{C} \setminus \mathbb{R}$, we also define

$$g_n(\lambda) = \mathrm{tr}_m(G_n(\lambda 1_m))$$

and

$$g(\lambda) = \mathrm{tr}_m(G(\lambda 1_m)).$$

We end this preliminary by recalling some properties of $G(\lambda)$ and of the resolvent $(\lambda \otimes 1_n - S_n)^{-1}$ of the matrix S_n. First, one can easily see that for any λ and λ' in $M_m(\mathbb{C})$ such that $Im(\lambda)$ and $Im(\lambda')$ are positive definite,

$$(\lambda \otimes 1_{\mathcal{B}} - s)^{-1} - (\lambda' \otimes 1_{\mathcal{B}} - s)^{-1} = (\lambda \otimes 1_{\mathcal{B}} - s)^{-1}(\lambda' - \lambda)(\lambda' \otimes 1_{\mathcal{B}} - s)^{-1}. \tag{2.6}$$

Lemma 2.1 *Let λ in $M_m(\mathbb{C})$ such that $Im(\lambda)$ is positive definite. Then*

$$\|(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}\| \leq ||Im(\lambda)^{-1}\| \quad \text{and} \quad ||G(\lambda)|| \leq ||Im(\lambda)^{-1}||. \tag{2.7}$$

Moreover, $G(\lambda)$ is invertible and

$$||G(\lambda)^{-1}|| \leq (||\lambda|| + ||s||)^2 ||Im(\lambda)^{-1}||. \tag{2.8}$$

We refer the reader to section 5 of [9] for a proof of (2.8).

Lemma 2.2 *Let λ in $M_m(\mathbb{C})$ such that $Im(\lambda)$ is positive definite, then*

$$||(\lambda \otimes 1_n - S_n)^{-1}|| \leq ||Im(\lambda)^{-1}||, \tag{2.9}$$

$$\forall 1 \leq k, l \leq n, ||(\lambda \otimes 1_n - S_n)^{-1}_{kl}|| \leq ||Im(\lambda)^{-1}||, \tag{2.10}$$

and for $p \geq 2$,

$$\frac{1}{n}\sum_{k,l=1}^{n} ||(\lambda \otimes 1_n - S_n)^{-1}_{kl}||^p \leq C_m ||Im(\lambda)^{-1}||^p \tag{2.11}$$

where, in the first inequality, $||.||$ denotes the operator norm in $M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$ (in $M_m(\mathbb{C})$ in the others) and C_m a constant depending only on m.

For a Hermitian matrix M, the derivative w.r.t M of the resolvent $R(z) = (z - M)^{-1}$ satisfies:

$$R'_M(z).A = R(z)AR(z) \text{ for all Hermitian matrix } A. \tag{2.12}$$

Sketch of Proof: We just mention the proof of (2.11). From (2.10), it's enough to consider the case $p = 2$.
Let us denote $G^{(n)} = (\lambda \otimes 1_n - S_n)^{-1} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$. Since the operator norm is smaller than the Hilbert-Schmidt norm,

$$\begin{aligned} \frac{1}{n}\sum_{k,l=1}^{n} ||G^{(n)}_{kl}||^2 &\leq \frac{1}{n}\sum_{k,l=1}^{n}\sum_{\alpha,\beta=1}^{m} |_{\alpha,\beta}G^{(n)}_{kl}|^2 \\ &= \frac{1}{n}\mathrm{Tr}_{nm}(G^{(n)}(G^{(n)})^*) \\ &\leq m||G^{(n)}(G^{(n)})^*|| \leq m||Im(\lambda)^{-1}||^2. \end{aligned}$$

where the last inequality follows from (2.9). $\square$

In the sequel, we shall denote by P_k any polynomial of degree k whose coefficients are positive and by C or K any constant; P_k, C or K can depend on the a_l, $l = 1, \ldots, r$, and may vary from line to line.

3 Main ideas in the proofs of (1.8) and (1.10) from [9] and [15]

3.1 Estimate of $\|G_n - G\|$ in [9]

Let us recall the main ideas of [9] in the estimation of $\|G_n(\lambda) - G(\lambda)\|$. In lemma 5.4 of [9], Haagerup and Thorbjørnsen observe in one hand that the matrix-valued Stieljes transform of s satisfies, for any λ in $\mathcal{O}$,

$$\sum_{i=1}^{r} a_i G(\lambda) a_i \;+\; (a_0 - \lambda) + G(\lambda)^{-1} = 0. \tag{3.1}$$

In the other hand, using the Gaussian integration by parts formula, they establish the analogue of (3.1) satisfied by $H_n(\lambda)$ ("Master equation", Lemma 2 [9]):

$$\mathbb{E}\left[\sum_{p=1}^{r} a_p H_n(\lambda) a_p H_n(\lambda) + (a_0 - \lambda) H_n(\lambda) + 1_m\right] = 0. \tag{3.2}$$

Then, using the Gaussian Poincaré inequality to get an estimate of the variance of $H_n(\lambda)$, they deduce from (3.2) the "Master inequality" (Lemma 3 in [9]):

$$\|\sum_{i=1}^{r} a_i G_n(\lambda) a_i G_n(\lambda) \;+\; (a_0 - \lambda) G(\lambda) + 1_m\| \le \frac{C}{n^2}\|Im(\lambda)^{-1}\|^4. \tag{3.3}$$

Moreover, the authors prove that $G_n(\lambda)$ is invertible for any λ in $\mathcal{O}$ and they give an upper bound of the norm of its inverse (Proposition 5.2 [9])

$$\|G_n(\lambda)^{-1}\| \le (\|\lambda\| + K)^2 \|Im(\lambda)^{-1}\|.$$

Hence, they deduce from (3.3) that, for any λ in $\mathcal{O}$,

$$\|a_0 + \sum_{i=1}^{r} a_i G_n(\lambda) a_i \;+\; G_n(\lambda)^{-1} - \lambda\| \le f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|), \tag{3.4}$$

where

$$f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|) = \frac{C}{n^2}(\|\lambda\| + K)^2 \|Im(\lambda)^{-1}\|^5.$$

Further, they set

$$\Lambda_n(\lambda) = a_0 + \sum_{i=1}^{r} a_i G_n(\lambda) a_i \ + \ G_n(\lambda)^{-1}$$

for any λ in $\mathcal{O}$. (3.4) can be rewritten

$$\|\Lambda_n(\lambda) - \lambda\| \leq f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|). \tag{3.5}$$

The authors define

$$\mathcal{O}_n^{'} = \{\lambda \in \mathcal{O}, f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|) < \frac{\epsilon(\lambda)}{2}\}$$

where

$$\epsilon(\lambda) := \frac{1}{\|Im(\lambda)^{-1}\|}.$$

(3.5) implies that, for any λ in $\mathcal{O}_n^{'}$,

$$Im\Lambda_n(\lambda) \geq \frac{1}{2\|Im(\lambda)^{-1}\|} 1_m \tag{3.6}$$

and that in particular $\Lambda_n(\lambda)$ belongs to $\mathcal{O}$ (see Lemma 5.5 [9]). Consequently, applying (3.1), they get that, for any λ in $\mathcal{O}_n^{'}$,

$$a_0 + \sum_{i=1}^{r} a_i G(\Lambda_n(\lambda)) a_i \ + G(\Lambda_n(\lambda))^{-1} = a_0 + \sum_{i=1}^{r} a_i G_n(\lambda) a_i \ + G_n(\lambda)^{-1}. \tag{3.7}$$

In proof of (b) Proposition 5.6, Haagerup and Thorbjørnsen show that (3.7) implies that

$$G_n(\lambda) = G(\Lambda_n(\lambda)) \tag{3.8}$$

for any λ in $\mathcal{O}_n^{''} := \{\lambda \in \mathcal{O}_n^{'}, \epsilon(\lambda) > \sqrt{2\sum_{i=1}^{r} \|a_i\|^2}\}$. Using that $t \mapsto f_n(t^{-1}, t)t^{-1}$ is a continuous strictly decreasing function from $]0; +\infty[$ onto $]0; +\infty[$, they show in proof of (a) Proposition 5.6 [9] that $\mathcal{O}_n^{'}$ is an open connected subset of $M_m(\mathbb{C})$. Thus, by the principle of uniqueness of analytic continuation, (3.8) still holds for any λ in $\mathcal{O}_n^{'}$. Thus, for any λ in $\mathcal{O}_n^{'}$, they get that

$$\begin{aligned}
\|G_n(\lambda) - G(\lambda)\| &\leq \|G(\Lambda_n(\lambda)) - G(\lambda)\| \\
&\leq \|Im(\Lambda_n(\lambda))^{-1}\| \|\lambda - \Lambda_n(\lambda)\| \|Im(\lambda)^{-1}\| \\
&\leq 2 f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|) \|Im(\lambda)^{-1}\|^2
\end{aligned}$$

where the last inequality comes from (3.5), (3.6). Now, if λ belongs to $\mathcal{O}\backslash\mathcal{O}_n'$, they note that

$$\begin{aligned} \|G_n(\lambda) - G(\lambda)\| &\leq 2\|Im(\lambda)^{-1}\| \\ &\leq 4f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|)\|Im(\lambda)^{-1}\|^2 \end{aligned}$$

since

$$\frac{1}{2} \leq f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|)\|Im(\lambda)^{-1}\|.$$

Finally, for any λ in $\mathcal{O}$,

$$\begin{aligned} \|G_n(\lambda) - G(\lambda)\| &\leq 4f_n(\|Im(\lambda)^{-1}\|, \|\lambda\|)\|Im(\lambda)^{-1}\|^2 \\ &= \frac{C}{n^2}(\|\lambda\| + K)^2\|Im(\lambda)^{-1}\|^7. \end{aligned} \quad (3.9)$$

3.2 Estimate of $\|G_n - G - \frac{1}{n}L\|$ in [15]

In the GOE case, a term of order $1/n$ appears in the Master equation so that the estimate of $\|G_n(\lambda) - G(\lambda)\|$ Schultz makes by sticking to the previous proof of [9] is of order $1/n$. Nevertheless, a further study (we will describe in our general framework in section 4) gives her the sharper estimate

$$\|G_n(\lambda) - G(\lambda) - \frac{1}{n}L(\lambda)\| \leq \frac{1}{n^2}(\|\lambda\| + K)^8 P_{13}(\|Im(\lambda)^{-1}\|) \quad (3.10)$$

for any λ such that $Im\lambda$ positive definite or negative definite.

3.3 From Step 2 to Step 3

¿From the previous estimates (3.9) and (3.10), Haagerup, Thorbjørnsen and Schultz immediately get that, for any λ in $\mathbb{C} \setminus \mathbb{R}$,

$$|r_n(\lambda)| \leq \frac{1}{n^2}(\|\lambda\| + K)^\alpha P_k(|Im(\lambda)^{-1}|) \quad (3.11)$$

where

-in the GUE case [9]

$$r_n(\lambda) = g_n(\lambda) - g(\lambda)\ ,\ \alpha = 2\ \ k = 7.$$

-in the GOE case [15]

$$r_n(\lambda) = g_n(\lambda) - g(\lambda) - \frac{1}{n} tr_m(L(\lambda 1_m)) \ , \ \alpha = 8 \ \ k = 13.$$

Since S_n and s are selfadjoint, by the spectral theory, there exist unique probability measures μ_n and μ on $\mathbb{R}$ such that

$$\int \varphi d\mu_n = \mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))]$$

$$\int \varphi d\mu = (\mathrm{tr}_m \otimes \tau)(\varphi(s)).$$

g_n and g are the Stieljes transforms of μ_n and μ. Moreover, in Lemma 5.5 in [15], Schultz proves by using a characterisation theorem of Tillmann that $l(\lambda) := tr_m(L(\lambda 1_m))$ is the Stieljes transform of a distribution Λ with compact support in $sp(s)$. Hence, using the inverse Stieljes tranform, Haagerup, Thorbjornsen and Schultz get respectively that, for any φ in $\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$,
- in [9]

$$\int \varphi d\mu_n - \int \varphi d\mu = -\frac{1}{\pi} \lim_{y \to 0^+} Im \int_{\mathbb{R}} \varphi(x) r_n(x+iy) dx. \tag{3.12}$$

- in [15]

$$\int \varphi d\mu_n - \int \varphi d\mu - \frac{\Lambda(\varphi)}{n} = -\frac{1}{\pi} \lim_{y \to 0^+} Im \int_{\mathbb{R}} \varphi(x) r_n(x+iy) dx. \tag{3.13}$$

Hence, the remainder of the two proofs (in [9] and [15]) deals with the estimation of

$$\limsup_{y \to 0^+} | \int_{\mathbb{R}} \varphi(x) h(x+iy) dx|$$

where h is an analytic function on $\mathbb{C} \setminus \mathbb{R}$ which satisfies

$$|h(\lambda)| \leq (\|\lambda\| + K)^\alpha P_k(|Im(\lambda)^{-1}|). \tag{3.14}$$

In [9] section 6, Haagerup and Thorbjørnsen introduce a very clever family of functions $\{I_p(\lambda), p \geq 1\}$ defined by

$$I_p(\lambda) = \frac{1}{(p-1)!} \int_0^{+\infty} h(\lambda + t) t^{p-1} \exp(-t) dt.$$

They note that

$$I_1(\lambda) - I_1^{'}(\lambda) = h(\lambda)$$
$$I_p(\lambda) - I_p^{'}(\lambda) = I_{p-1}(\lambda),\ p \geq 2,$$

so that for any φ in $\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R})$ and $y > 0$,

$$\int_{\mathbb{R}} \varphi(x) h(x+iy) dx = \int_{\mathbb{R}} (1+D)^p \varphi(x) I_p(x+iy) dx.$$

Now, they choose $p = k+1$ where k is the degree of the polynomial in the right hand side of (3.14) (that is $p = 8$ in [9] and $p = 14$ in [15]) and estimate $I_{k+1}(\lambda)$ for $Im\lambda > 0$. Using (3.14), it is not difficult to see that

$$\lim_{r \to +\infty} \int_{[r, r+ir]} \frac{1}{k!} h(\lambda + z) z^k \exp(-z) dz = 0.$$

Thus, by Cauchy's integral theorem, the authors get

$$\begin{aligned} I_{k+1}(\lambda) &= \lim_{r \to +\infty} \int_{[0, r+ir]} \frac{1}{k!} h(\lambda + z) z^k \exp(-z) dz \\ &= \int_0^{+\infty} \frac{1}{k!} h(\lambda + (1+i)t)(1+i)^{k+1} t^k \exp(-(1+i)t) dt. \end{aligned}$$

Plugging in (3.14), one gets for any λ such that $Im\lambda > 0$,

$$\begin{aligned} |I_{k+1}(\lambda)| &\leq \frac{2^{\frac{k+1}{2}}}{k!} \int_0^{+\infty} (|\lambda| + \sqrt{2}t + K)^\alpha P_k(|Im\lambda + t|^{-1}) t^k \exp(-t) dt \\ &\leq \frac{2^{\frac{k+1}{2}}}{k!} \int_0^{+\infty} (|\lambda| + \sqrt{2}t + K)^\alpha P_k(t^{-1}) t^k \exp(-t) dt \\ &\leq \int_0^{+\infty} (|\lambda| + \sqrt{2}t + K)^\alpha Q(t) \exp(-t) dt \end{aligned}$$

where $Q(t) = \frac{2^{\frac{k+1}{2}}}{k!} P_k(t^{-1}) t^k$ is a polynomial.
It follows by dominated convergence

$$\limsup_{y \to 0^+} |\int_{\mathbb{R}} \varphi(x) h(x+iy) dx| \leq \int_{\mathbb{R}} \int_0^{+\infty} |(1+D)^p \varphi(x)| (|x| + \sqrt{2}t + K)^\alpha Q(t) \exp(-t) dt dx < +\infty.$$

Dealing with $h(\lambda) = n^2 r_n(\lambda)$ one gets

$$\limsup_{y \to 0^+} | \int_{\mathbb{R}} \varphi(x) r_n(x+iy) dx | \leq \frac{C}{n^2}. \tag{3.15}$$

Combining (3.15) with respectively (3.12) and (3.13) , one gets respectively (1.8) and (1.10).

4 The iid case

We consider a Hermitian matrix $X_n = [(X_n)_{jk}]_{j,k=1}^n$ of size n for which the n^2 rv $((X_n)_{ii})$, $(\sqrt{2}Re((X_n)_{ij})_{i<j}$, $(\sqrt{2}Im((X_n)_{ij})_{i<j}$ are independent identically distributed with common distribution $\mu/\sqrt{n}$ where μ is a symmetric distribution with variance 1 on $\mathbb{R}$ which satisfies a Poincaré inequality (see section 4.2). We call X_n a Wigner matrix with distribution μ. Let $X_n^{(1)}, \ldots, X_n^{(r)}$ be r independent copies of X_n. We present our main technical tool (see [12]):

Lemma 4.1 *Let ξ be a real-valued rv such that $\mathbb{E}(|\xi|^{p+2}) < \infty$. Let ϕ be a function from $\mathbb{R}$ to $\mathbb{C}$ such that the first $p+1$ derivatives are continuous and bounded. Then,*

$$\mathbb{E}(\xi\phi(\xi)) = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!} \mathbb{E}(\phi^{(a)}(\xi)) + \epsilon \tag{4.1}$$

where κ_a are the cumulants of ξ, $|\epsilon| \leq C \sup_t |\phi^{(p+1)}(t)| \mathbb{E}(|\xi|^{p+2})$, C depends on p only.

In the following, we shall apply this identity with a function $\phi(\xi)$ given by the Stieljes transform of a random matrix. It follows from the Lemma 2.2 and (2.12) above that the conditions of Lemma 4.1 (bounded derivatives) are fulfilled.

4.1 The master equation

Note that since μ satisfies a Poincaré inequality, we have $\int |x|^q d\mu(x) < +\infty$ for any q in $\mathbb{N}$ (see Corollary 3.2 and Proposition 1.10 in [13]). Note also that, since μ is symmetric, any odd cumulant of μ vanishes.

Theorem 4.1 *With the previous notations,*

$$\mathbb{E}\left[\sum_{i=1}^{r} a_i H_n(\lambda) a_i H_n(\lambda) \ + \ (a_0-\lambda)H_n(\lambda)+1_m\right] + \frac{1}{n}R_n(\lambda)+\epsilon_n = 0 \quad (4.2)$$

where $\|\epsilon_n\| \leq \frac{P_6(\|Im(\lambda)^{-1}\|)}{n^2}$ *and* $R_n(\lambda)$ *denotes the quantity*

$$\frac{\kappa_4}{2}\mathbb{E}\left[\sum_{p=1}^{r}\frac{1}{n^2}\sum_{k,l=1}^{n} a_p(\lambda\otimes 1_n - S_n)^{-1}_{kk} a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll} a_p(\lambda\otimes 1_n - S_n)^{-1}_{kk} a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll}\right]$$

where κ_4 *is the fourth cumulant of the distribution* μ*. Note that*

$$\|R_n(\lambda)\| \leq P_4(\|Im(\lambda)^{-1}\|). \quad (4.3)$$

Proof: We shall apply formula (4.1) to the $\mathcal{M}_m(\mathbb{C})$-valued function $\phi(\xi) = (\lambda\otimes 1_n - S_n)^{-1}_{ij}$ for $1 \leq i,j \leq n$ and ξ is one of the variable $(X_n^{(p)})_{kk}$, $\sqrt{2}Re((X_n^{(p)})_{kl})$, $\sqrt{2}Im((X_n^{(p)})_{kl})$ for $1 \leq k < l \leq n$ and $p \leq r$.
We notice that

$$\begin{aligned}
\frac{\partial\phi}{\partial Re((X_n^{(p)})_{kk})} &= \phi'_{X_n^{(p)}}.e_{kk}, \ 1\leq k\leq n\\
\frac{\partial\phi}{\partial\sqrt{2}Re((X_n^{(p)})_{kl})} &= \phi'_{X_n^{(p)}}.e_{kl}, \ 1\leq k<l\leq n\\
\frac{\partial\phi}{\partial\sqrt{2}Im((X_n^{(p)})_{kl})} &= \phi'_{X_n^{(p)}}.f_{kl}, \ 1\leq k<l\leq n\\
\sqrt{2}Re((X_n^{(p)})_{kl}) &= \mathrm{Tr}_n(X_n^{(p)}e_{kl}), \ 1\leq k<l\leq n\\
\sqrt{2}Im((X_n^{(p)})_{kl}) &= \mathrm{Tr}_n(X_n^{(p)}f_{kl}), \ 1\leq k<l\leq n\\
(X_n^{(p)})_{kk} &= \mathrm{Tr}_n(X_n^{(p)}e_{kk}), \ 1\leq k\leq n.
\end{aligned}$$

Let $1\leq p\leq r$, $1\leq k\leq l\leq n$ be fixed. For simplicity, we write ϕ', ϕ'', ϕ''' for the first derivatives of ϕ with respect to $\sqrt{2}Re((X_n^{(p)})_{kl})$. Then, according to (2.12),

$$\begin{aligned}
\phi' &= \left[(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}\right]_{ij}\\
\phi'' &= 2\left[(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}\right]_{ij}\\
\phi''' &= 6\left[(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}\right.\\
&\qquad \left. a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}\right]_{ij}
\end{aligned}$$

Writing (4.1) in this setting gives

$$\mathbb{E}[\mathrm{Tr}_n(X_n^{(p)}e_{kl})(\lambda\otimes 1_n - S_n)^{-1}_{ij}] = \frac{1}{n}\mathbb{E}[\phi'] + \frac{\kappa_4}{6n^2}\mathbb{E}[\phi'''] + O(n^{-3}) \qquad (4.4)$$

where the $O(n^{-3})$ means the norm of this term is smaller than $\frac{C\|a_p\|^5\|Im(\lambda)^{-1}\|^6}{n^3}$. Multiplying by n gives the equation, denoted by $A_{ij}^{kl}(p)$:

$$n\mathbb{E}[\mathrm{Tr}_n(X_n^{(p)}e_{kl})(\lambda\otimes 1_n - S_n)^{-1}_{ij}] = \mathbb{E}[\phi'] + \frac{\kappa_4}{6n}\mathbb{E}[\phi'''] + O(n^{-2}) \qquad (4.5)$$

with the analogous equations with f_{pq} (denoted by $B_{ij}^{kl}(p)$) and e_{pp}.

Recall how we can obtain the master equation in the gaussian case (GUE case) from (4.5) which reads in this case:

$$n\mathbb{E}[\mathrm{Tr}_n(X_n^{(p)}e_{kl})(\lambda\otimes 1_n - S_n)^{-1}_{ij}] = \mathbb{E}\left[(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes e_{kl}(\lambda\otimes 1_n - S_n)^{-1}\right]_{ij}. \qquad (4.6)$$

By a linear combination with the analogous equation with f_{kl}, we have:

$$n\mathbb{E}[\mathrm{Tr}_n(X_n^{(p)}E_{kl})(\lambda\otimes 1_n - S_n)^{-1}_{ij}] = \mathbb{E}\left[(\lambda\otimes 1_n - S_n)^{-1}a_p\otimes E_{kl}(\lambda\otimes 1_n - S_n)^{-1}\right]_{ij}$$

for all $1\leq k,l\leq n$.
Now, take in the above formula $i=k$, $j=l$ and consider $\frac{1}{n^2}\sum_{k,l}$, we then obtain:

$$\frac{1}{n}\sum_{k,l}\mathbb{E}[(X_n^{(p)})_{lk}(\lambda\otimes 1_n - S_n)^{-1}_{kl}] = \frac{1}{n^2}\sum_{k,l}\mathbb{E}\left[(\lambda\otimes 1_n - S_n)^{-1}_{kk}a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll}\right] \qquad (4.7)$$

that is

$$\mathbb{E}[id_m\otimes \mathrm{tr}_n((1_m\otimes X_n^{(p)})(\lambda\otimes 1_n - S_n)^{-1})] = \mathbb{E}\left[H_n(\lambda)a_pH_n(\lambda)\right]. \qquad (4.8)$$

Now, from the above equation,

$$\begin{aligned}\mathbb{E}\left[\sum_{p=1}^r a_pH_n(\lambda)a_pH_n(\lambda)\right] &= \sum_{p=1}^r \mathbb{E}[id_m\otimes\mathrm{tr}_n((a_p\otimes 1_n)(1_m\otimes X_n^{(p)})(\lambda\otimes 1_n - S_n)^{-1})]\\ &= \sum_{p=1}^r \mathbb{E}[id_m\otimes\mathrm{tr}_n((a_p\otimes X_n^{(p)})(\lambda\otimes 1_n - S_n)^{-1})]\\ &= \mathbb{E}[id_m\otimes\mathrm{tr}_n(S_n - a_0\otimes 1_n)(\lambda\otimes 1_n - S_n)^{-1})]\\ &= -1_m + (\lambda - a_0)\mathbb{E}[H_n(\lambda)]\end{aligned}$$

implying the master formula in the GUE case:

$$\mathbb{E}\left[\sum_{p=1}^{r} a_p H_n(\lambda) a_p H_n(\lambda) + (a_0 - \lambda) H_n(\lambda) + 1_m\right] = 0.$$

Keeping in mind these computations, we now study the terms coming from third derivatives.
We thus consider $A(p) = \frac{1}{n^2}\sum_{k,l} A_{kl}^{kl}(p)$ (resp. $B(p)$) and study all the contributions of the different terms.

Study of the third derivative
Writing as before the terms appearing in ϕ''', we can see that all the terms except one contains at least two G_{kl} and then, according to Lemma 2.2, these terms will give a contribution in $O(n^{-2})$ in $A(p)$. The only term to be considered is:

$$\frac{1}{n^2}\sum_{k,l}\mathbb{E}((\lambda\otimes 1_n - S_n)^{-1}_{kk} a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll} a_p(\lambda\otimes 1_n - S_n)^{-1}_{kk} a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll}).$$

Now, using the same linear combination giving (4.7) in the GUE case, we obtain that the corrective term of order $1/n$ appearing in the iid case is:

$$\frac{1}{n}\{\frac{\kappa_4}{2}\mathbb{E}\left[\frac{1}{n^2}\sum_{k,l=1}^{n}(\lambda\otimes 1_n - S_n)^{-1}_{kk} a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll} a_p(\lambda\otimes 1_n - S_n)^{-1}_{kk}\right.$$
$$\left. a_p(\lambda\otimes 1_n - S_n)^{-1}_{ll}\right]\}.$$

The proof of the Theorem is complete. □

4.2 Variance estimate

We assume that μ satisfies a Poincaré inequality: there exists a positive constant C such that for any $\mathcal{C}^1$ function $f : \mathbb{R} \to \mathbb{C}$ such that f and f' are in $L^2(\mu)$,

$$\mathbf{V}(f) \leq C \int |f'|^2 d\mu,$$

with $\mathbf{V}(f) = \mathbb{E}(|f - \mathbb{E}(f)|^2)$. We refer the reader to [4] for a characterization of the measures on $\mathbb{R}$ which satisfy a Poincaré inequality (see also [1]). For example, $\mu(dx) = \exp(-|x|^\alpha)dx$ with $\alpha \geq 1$ satisfies the Poincaré inequality.

For any matrices $A_1, \ldots, A_r$, define $\|(A_1, \ldots, A_r)\|_e^2 := \sum_{j=1}^r \|A_j\|_2^2$. Let $\Psi : (M_n(\mathbb{C})_{sa})^r \to \mathbb{R}^{rn^2}$ be the canonical isomorphism introduced in Remark 3.4 in [9].

Lemma 4.2 *For any function $f : \mathbb{R}^{rn^2} \to \mathbb{C}$ be a $\mathcal{C}^1$-function such that f and the gradient $\nabla(f)$ are both polynomially bounded,*

$$\mathbf{V}\big[f \circ \Psi(X_n^{(1)}, \ldots, X_n^{(r)})\big] \leq \frac{C}{n}\mathbb{E}\{\|\nabla \left[f \circ \Psi(X^{(1)}, \ldots, X^{(r)})\right] \|_e^2\}. \tag{4.9}$$

Proof $\mu^{(n)} := \mu/\sqrt{n}$ satisfies the Poincaré inequality

$$\int |g - \int g d\mu^{(n)}|^2 d\mu^{(n)} \leq \frac{C}{n} \int |g'|^2 d\mu^{(n)}.$$

(4.9) readily follows by the tensorisation property of the Poincaré inequality.

4.3 Master inequality

We follow the lines of the proof of Theorem 4.5 in [9]. Using the master equality (4.2), we easily get

$$\|\textstyle\sum_{i=1}^r a_i G_n(\lambda) a_i G_n(\lambda) \; + \; (a_0 - \lambda) G_n(\lambda) + 1_m + \frac{1}{n} R_n(\lambda) + \epsilon_n\|$$

$$\leq \| \sum_{i=1}^r a_i^2 \| \mathbb{E}\{\|H_n(\lambda) - \mathbb{E}(H_n(\lambda))\|^2\}.$$

Thanks to (4.9), the following of the proof of Theorem 4.5 in [9] still holds and we similarly get

$$\mathbb{E}\{\|H_n(\lambda) - E(H_n(\lambda))\|^2\} \leq \frac{Cm^3}{n^2} \| \sum_{i=1}^r a_i^2 \| \|(Im(\lambda))^{-1}\|^4$$

and therefore

$$\| \sum_{i=1}^r a_i G_n(\lambda) a_i G_n(\lambda) \; + \; (a_0 - \lambda) G_n(\lambda) + 1_m + \frac{1}{n} R_n(\lambda)\| \leq \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n^2}. \tag{4.10}$$

4.4 Estimation of $\|G_n - G\|$

In the Gaussian case, Haagerup and Thorbjørnsen in [9] and Schultz in [15] prove that $G_n(\lambda)$ is invertible for any λ such that $Im\lambda$ is positive definite. In our more general case, we are going to use the master inequality (4.10) in order to prove that, for any λ in some subset of $\mathcal{O}$, $G_n(\lambda)$ is invertible and to get an upper bound of $\|G_n(\lambda)^{-1}\|$.
Set

$$B_n(\lambda) = \sum_{i=1}^{r} a_i G_n(\lambda) a_i \; + \; (a_0 - \lambda).$$

Now, from the master inequality (4.10) and (4.3), we get

$$\|\sum_{i=1}^{r} a_i G_n(\lambda) a_i G_n(\lambda) \; + \; (a_0 - \lambda) G_n(\lambda) + 1_m\| \leq \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n}. \quad (4.11)$$

that is

$$\|B_n(\lambda) G_n(\lambda) + 1_m\| \leq \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n}.$$

Hence, for any λ such that $\frac{P_6(\|(Im(\lambda))^{-1}\|)}{n} < \frac{1}{2}$, $B_n(\lambda)G_n(\lambda)$ is invertible with

$$\|[B_n(\lambda)G_n(\lambda)]^{-1}\| \leq 2.$$

Thus, for such a λ, $G_n(\lambda)$ is also obviously invertible with

$$\|G_n(\lambda)^{-1}\| \leq 2\|B_n(\lambda)\| \leq 2(\|a_0\| + \|\lambda\| + \sum_{i=1}^{r} \|a_i\|^2 \|(Im(\lambda))^{-1}\|) \quad (4.12)$$

Now, from the inequality (4.11) and using (4.12), we get readily that for any λ in $\mathcal{O}$ such that $\frac{P_6(\|(Im(\lambda))^{-1}\|)}{n} < \frac{1}{2}$,

$$\|\sum_{i=1}^{r} a_i G_n(\lambda) a_i + (a_0 - \lambda) + G_n(\lambda)^{-1}\| \leq \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n} 2(\|a_0\| + \|\lambda\| + \sum_{i=1}^{r} \|a_i\|^2 \|(Im(\lambda))^{-1}\|). \quad (4.13)$$

Define

$$\mathcal{O}_n' = \left\{ \lambda \in \mathcal{O}, \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n} 2(\|a_0\| + \|\lambda\| + \sum_{i=1}^{r} \|a_i\|^2 \|(Im(\lambda))^{-1}\|) < \frac{1}{2\|(Im(\lambda))^{-1}\|} \right\}.$$

Since $t \mapsto \frac{P_6(t^{-1})}{n} 2(\|a_0\| + t + \sum_{i=1}^r \|a_i\|^2 t^{-1})t^{-1}$ is a continuous strictly decreasing function from $]0, +\infty[$ onto $]0, +\infty[$, one can prove that $\mathcal{O}_n'$ is an open connected subset of $M_m(\mathbb{C})$ by following the proof of (a) Proposition 5.6 in [9]. Note that, using the inequality

$$\frac{1}{\|(Im(\lambda))^{-1}\|} \leq \|\lambda\|,$$

one immediately gets that for any λ in $\mathcal{O}_n'$,

$$\|(Im(\lambda))^{-1}\|(\|a_0\| + \|\lambda\| + \sum_{i=1}^r \|a_i\|^2 \|(Im(\lambda))^{-1}\|) \geq 1$$

and thus that

$$\frac{P_6(\|(Im(\lambda))^{-1}\|)}{n} \leq \frac{1}{4}.$$

Consequently, for any λ in $\mathcal{O}_n'$, $G_n(\lambda)$ is invertible and (4.13) holds. Defining for any λ in $\mathcal{O}_n'$,

$$\Lambda_n(\lambda) = a_0 + \sum_{i=1}^r a_i G_n(\lambda) a_i + G_n(\lambda)^{-1}$$

and sticking to the proof of [9] described in section II. 2.1, we get that, for any λ in $\mathcal{O}$,

$$\begin{aligned}
\|G_n(\lambda) - G(\lambda)\| &\leq 4\frac{P_6(\|(Im(\lambda))^{-1}\|)}{n} 2(\|a_0\| + \|\lambda\| + \sum_{i=1}^r \|a_i\|^2 \|(Im(\lambda))^{-1}\|)\|(Im(\lambda))^{-1}\|^2 \\
&\leq (\|\lambda\| + K)\frac{P_9(\|(Im(\lambda))^{-1}\|)}{n} \qquad (4.14)
\end{aligned}$$

Note that, in the following we will use (4.13) in the simplest form:

$$\forall \lambda \in \mathcal{O}_n', \; \|\sum_{i=1}^r a_i G_n(\lambda) a_i + (a_0 - \lambda) + G_n(\lambda)^{-1}\| \leq (\|\lambda\| + K)\frac{P_7(\|(Im(\lambda))^{-1}\|)}{n}.$$

4.5 Convergence of $R_n(\lambda)$

Let x_i, $i = 1, \ldots, r$ be self-adjoint operators in a C^* probability space $(\mathcal{B}, \tau)$. We assume that the x_i are free and identically semi-circular distributed with mean 0 and variance 1. Then, G satisfies (3.1).

Proposition 4.1 *Let a be a matrix in $M_m(\mathbb{C})$. Then,*

$$\mathbb{E}[\frac{1}{n}\sum_{k=1}^{n}(\lambda\otimes 1_n - S_n)^{-1}_{kk}\ a\ (\lambda\otimes 1_n - S_n)^{-1}_{kk}] = GaG + O\left(\frac{P_{10}(\|(Im(\lambda))^{-1}\|)}{n}(\|\lambda\| + K)\right) \quad (4.15)$$

Proof: We start from the resolvent identity:

$$\lambda(\lambda\otimes 1_n - S_n)^{-1}_{kk} = 1_m + \sum_{l=1}^{n}(S_n)_{kl}(\lambda\otimes 1_n - S_n)^{-1}_{lk}.$$

We write $G^{(n)}(\lambda) = (\lambda\otimes 1_n - S_n)^{-1}$ and $D_a^{(n)}(\lambda) = \frac{1}{n}\sum_{k=1}^{n} G^{(n)}(\lambda)_{kk}\ a\ G^{(n)}(\lambda)_{kk}$. From the above identity,

$$\begin{aligned}
&\lambda\ \frac{1}{n}\sum_{k=1}^{n} G^{(n)}(\lambda)_{kk}\ a\ G^{(n)}(\lambda)_{kk} \\
&\quad = a\ \frac{1}{n}\sum_{k=1}^{n} G^{(n)}(\lambda)_{kk} + \frac{1}{n}\sum_{k,l=1}^{n}(S_n)_{kl}G^{(n)}(\lambda)_{lk}\ a\ G^{(n)}(\lambda)_{kk} \\
&\quad = aH_n(\lambda) + a_0 D_a^{(n)}(\lambda) + \frac{1}{n}\sum_{p=1}^{r} a_p \sum_{k,l=1}^{n}(X_n^{(p)})_{kl}G^{(n)}(\lambda)_{lk}\ a\ G^{(n)}(\lambda)_{kk}
\end{aligned}$$

We take the expectation and we use the integration by part formula (4.1) for the last term:

$$\begin{aligned}
\mathbb{E}((X_n^{(p)})_{kl}\Phi(X_n^{(1)},\ldots,X_n^{(r)})] &= \mathbb{E}(\mathrm{Tr}_n(X_n^{(p)}E_{lk})\Phi(X_n^{(1)},\ldots,X_n^{(r)})] \\
&= \frac{1}{n}\mathbb{E}[\Phi'_p(X_n^{(1)},\ldots,X_n^{(r)}).E_{lk}] + O\left(\frac{P_4(\|(Im(\lambda))^{-1}\|)}{n^2}\right)
\end{aligned}$$

with $\Phi(X_n^{(1)},\ldots,X_n^{(r)}) = G^{(n)}(\lambda)_{lk}\ a\ G^{(n)}(\lambda)_{kk}$. Then,

$$\begin{aligned}
\mathbb{E}[(X_n^{(p)})_{kl}G^{(n)}(\lambda)_{lk}aG^{(n)}(\lambda)_{kk}] &= \frac{1}{n}\mathbb{E}[G^{(n)}(\lambda)_{ll}a_pG^{(n)}(\lambda)_{kk}aG^{(n)}(\lambda)_{kk}] \\
&\quad + \frac{1}{n}\mathbb{E}[G^{(n)}(\lambda)_{lk}aG^{(n)}(\lambda)_{kl}a_pG^{(n)}(\lambda)_{lk}] \\
&\quad + O\left(\frac{P_4(\|(Im(\lambda))^{-1}\|)}{n^2}\right)
\end{aligned}$$

Thus, we obtain from the resolvent identity,

$$(\lambda - a_0)\mathbb{E}(D_a^{(n)}(\lambda)) = a\mathbb{E}(H_n(\lambda)) +$$
$$\sum_{p=1}^{r} a_p \sum_{k,l=1}^{n} \frac{1}{n^2}\mathbb{E}[G^{(n)}(\lambda)_{ll} a_p G^{(n)}(\lambda)_{kk} a G^{(n)}(\lambda)_{kk} +$$
$$G^{(n)}(\lambda)_{lk} a G^{(n)}(\lambda)_{kl} a_p G^{(n)}(\lambda)_{lk}] + O\left(\frac{P_4(\|(Im(\lambda))^{-1}\|)}{n}\right).$$

¿From Lemma 2.2,

$$\frac{1}{n^2}\sum_{k,l=1}^{n} \mathbb{E}[G^{(n)}(\lambda)_{lk} a G^{(n)}(\lambda)_{kl} a_p G^{(n)}(\lambda)_{lk}] = O\left(\frac{P_3(\|(Im(\lambda))^{-1}\|)}{n}\right),$$

thus,

$$(\lambda - a_0)\mathbb{E}(D_a^{(n)}(\lambda)) = a\mathbb{E}(H_n(\lambda)) + \sum_{p=1}^{r} a_p \mathbb{E}[H_n(\lambda) a_p D^{(n)}(\lambda)] + O\left(\frac{P_4(\|(Im(\lambda))^{-1}\|)}{n}\right)$$

From the estimate of the variance of H_n, we have:

$$\mathbb{E}[H_n(\lambda) a_p D_a^{(n)}(\lambda)] = \mathbb{E}[H_n(\lambda)] a_p \mathbb{E}[D_a^{(n)}(\lambda)] + O\left(\frac{P_4(\|(Im(\lambda))^{-1}\|)}{n}\right).$$

Then, using also the estimation of $\|G_n(\lambda) - G(\lambda)\|$ we get

$$(\lambda - a_0 - \sum_{p=1}^{r} a_p G(\lambda) a_p)\mathbb{E}(D_a^{(n)}(\lambda)) = aG + O\left(\frac{P_9(\|(Im(\lambda))^{-1}\|)}{n}(\|\lambda\| + K)\right).$$

Using (3.1) and (2.9) we finally get

$$\mathbb{E}(D_a^{(n)}(\lambda)) = GaG + O\left(\frac{P_{10}(\|(Im(\lambda))^{-1}\|)}{n}(\|\lambda\| + K)\right).$$

□

¿From the above proposition, we obtain:

Proposition 4.2 *$R_n(\lambda)$ defined in Theorem 4.1 converges as n tends to infinity to*

$$R(\lambda) = \frac{\kappa_4}{2}\sum_{p=1}^{r} a_p G(\lambda) a_p G(\lambda) a_p G(\lambda) a_p G(\lambda).$$

More precisely,

$$\|R_n(\lambda) - R(\lambda)\| \leq (\|\lambda\| + K)^2 \frac{P_{20}(\|(Im(\lambda))^{-1}\|)}{n}. \tag{4.16}$$

Proof: It's enough to prove the convergence of each coordinate of the $m \times m$ matrix $R_n(\lambda)$. This will actually follow from the convergence of terms of the form:

$$\mathbb{E}\left[{}_{\alpha,\beta}\left(n^{-1} \sum_{k=1}^{n} G_{kk}^{(n)} a G_{kk}^{(n)} \right) {}_{\gamma,\delta}\left(n^{-1} \sum_{k=1}^{n} G_{kk}^{(n)} b G_{kk}^{(n)} \right) \right] \tag{4.17}$$

for a et b elements of the canonical basis in $\mathcal{M}_m(\mathbb{C})$. Since, applying Lemma 4.2, we have

$$\mathbb{E}(\|D_a^{(n)}(\lambda) - \mathbb{E}(D_a^{(n)}(\lambda))\|^2) \leq \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n},$$

the above quantity (4.17) is of the same order as:

$$\mathbb{E}\left[{}_{\alpha,\beta}\left(n^{-1} \sum_{k=1}^{n} G_{kk}^{(n)} a G_{kk}^{(n)} \right) \right] \mathbb{E}\left[{}_{\gamma,\delta}\left(n^{-1} \sum_{k=1}^{n} G_{kk}^{(n)} b G_{kk}^{(n)} \right) \right]$$

According to Proposition 4.1, this last quantity converges towards ${}_{\alpha,\beta}(GaG)\ {}_{\gamma,\delta}(GbG)$. Thus, the convergence of R_n to R follows with the estimation (4.16). $\square$

We define

$$L(\lambda) = (id_m \otimes \tau)[(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}(R(\lambda)G(\lambda)^{-1} \otimes 1_{\mathcal{B}})(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}].$$

4.6 Estimation of $\|G(\lambda) - G_n(\lambda) + \frac{1}{n}L(\lambda)\|$

Following 4.24 in [15], one gets for any λ in $\mathcal{O}_n^{'}$,

$$\begin{aligned}
&\|G(\lambda) - G_n(\lambda) + \tfrac{1}{n}L(\lambda)\| \\
\leq\ & \|(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}\| \|(\Lambda_n(\lambda) \otimes 1_{\mathcal{B}} - s)^{-1}\| \|\Lambda_n(\lambda) - \lambda + \frac{1}{n}R(\lambda)G(\lambda)^{-1}\| \\
& + \frac{1}{n}\|(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}\| \|R(\lambda)G(\lambda)^{-1}\| \|(\lambda \otimes 1_{\mathcal{B}} - s)^{-1} - (\Lambda_n(\lambda) \otimes 1_{\mathcal{B}} - s)^{-1}\| \\
\leq\ & \|(Im(\lambda))^{-1}\| \|(Im(\Lambda_n(\lambda)))^{-1}\| \|\Lambda_n(\lambda) - \lambda + \frac{1}{n}R(\lambda)G(\lambda)^{-1}\| \\
& + \frac{C}{n}\|(Im(\lambda))^{-1}\|^7 \|(Im(\Lambda_n(\lambda)))^{-1}\| \|\Lambda_n(\lambda) - \lambda\| (\|\lambda\| + K)^2
\end{aligned}$$

where we made use of the estimates (2.7), (2.8), (2.6) and the upper bound

$$\|R(\lambda)\| \leq C\|(Im(\lambda))^{-1}\|^4.$$

Now, for any λ in $\mathcal{O}_n'$,

$$\|(Im(\Lambda_n(\lambda)))^{-1}\| \leq 2\|(Im(\lambda))^{-1}\|$$

and

$$\|\Lambda_n(\lambda) - \lambda\| \leq \frac{P_6(\|(Im(\lambda))^{-1}\|)}{n}(\|\lambda\| + K).$$

Thus,

$$\begin{aligned}\|G(\lambda) - G_n(\lambda) + \frac{1}{n}L(\lambda)\| &\leq 2\|(Im(\lambda))^{-1}\|^2\|\Lambda_n(\lambda) - \lambda + \frac{1}{n}R(\lambda)G(\lambda)^{-1}\| \\ &\quad + \frac{P_{14}(\|(Im(\lambda))^{-1}\|)}{n^2}(\|\lambda\| + K)^3.\end{aligned}$$

Now, for any λ in $\mathcal{O}_n'$,

$$\begin{aligned}\|\Lambda_n(\lambda) - \lambda + \frac{1}{n}R(\lambda)G(\lambda)^{-1}\| &\leq \|\Lambda_n(\lambda) - \lambda + \frac{1}{n}R_n(\lambda)G_n(\lambda)^{-1}\| \\ &\quad + \frac{1}{n}\|R_n(\lambda)(G_n(\lambda)^{-1} - G(\lambda)^{-1})\| \\ &\quad + \frac{1}{n}\|(R_n(\lambda) - R(\lambda))G(\lambda)^{-1})\| \\ &\leq \frac{P_7(\|(Im(\lambda))^{-1}\|)}{n^2}(\|\lambda\| + K) \\ &\quad + \frac{P_4(\|(Im(\lambda))^{-1}\|)}{n}\|G_n(\lambda)^{-1} - G(\lambda)^{-1}\| \\ &\quad + \frac{1}{n}\|R_n(\lambda) - R(\lambda)\|(\|\lambda\| + K)^2\|(Im(\lambda))^{-1}\|\end{aligned}$$

where we used (4.10), (4.12), (4.3) and (2.8). Moreover, one easily gets

$$\begin{aligned}\|G_n(\lambda)^{-1} - G(\lambda)^{-1}\| &= \|G_n(\lambda)^{-1}(G(\lambda) - G_n(\lambda))G(\lambda)^{-1}\| \\ &\leq \|G_n(\lambda)^{-1}\|\|G(\lambda) - G_n(\lambda))\|\|G(\lambda)^{-1}\|.\end{aligned}$$

Consequently, using the estimate (4.14) of $\|G_n(\lambda) - G(\lambda)\|$ together with (4.12) and (2.8), we get

$$\|G_n(\lambda)^{-1} - G(\lambda)^{-1}\| \leq (\|\lambda\| + K)^4\frac{P_{11}(\|(Im(\lambda))^{-1}\|)}{n}.$$

We conclude that

$$\begin{aligned}\|G(\lambda)-G_n(\lambda)+\frac{1}{n}L(\lambda)\| &\leq (\|\lambda\|+K)^4\frac{P_{17}(\|(Im(\lambda))^{-1}\|)}{n^2}\\ &+\frac{2}{n}\|R_n(\lambda)-R(\lambda)\|(\|\lambda\|+K)^2\|(Im(\lambda))^{-1}\|^3.\end{aligned}$$

Using (4.16), we can conclude that, for any λ in $\mathcal{O}'_n$,

$$\|G(\lambda)-G_n(\lambda)+\frac{1}{n}L(\lambda)\| \leq (\|\lambda\|+K)^4\frac{P_{23}(\|(Im(\lambda))^{-1}\|)}{n^2}.$$

Now, for λ in $\mathcal{O}\setminus\mathcal{O}'_n$,

$$\begin{aligned}1 &\leq 4\frac{P_6(\|(Im(\lambda))^{-1}\|)}{n}(\|a_0\|+\|\lambda\|+\|\sum_{i=1}^r a_i^2\|\|(Im(\lambda))^{-1}\|)\|(Im(\lambda))^{-1}\|\}\\ &\leq (\|\lambda\|+K)\frac{P_8(\|(Im(\lambda))^{-1}\|)}{n}.\end{aligned}$$

We get

$$\begin{aligned}\|G(\lambda)-G_n(\lambda)+\frac{1}{n}L(\lambda)\| &\leq \|G(\lambda)-G_n(\lambda)\|+\frac{1}{n}\|L(\lambda)\|\\ &\leq (\|\lambda\|+K)\frac{P_8(\|(Im(\lambda))^{-1}\|)}{n}\\ &\quad\times\left[(\|\lambda\|+K)\frac{P_9(\|(Im(\lambda))^{-1}\|)}{n}+\frac{1}{n}\|(Im(\lambda))^{-1}\|^7(\|\lambda\|+K)^2\right]\\ &\leq (\|\lambda\|+K)^3\frac{P_{17}(\|(Im(\lambda))^{-1}\|)}{n^2}.\end{aligned}$$

Thus, one can easily see that one can choose K and P_{23} such that for any λ in $\mathcal{O}$,

$$\|G(\lambda)-G_n(\lambda)+\frac{1}{n}L(\lambda)\| \leq (\|\lambda\|+K)^4\frac{P_{23}(\|(Im(\lambda))^{-1}\|)}{n^2}. \qquad (4.18)$$

Note that, since under our hypothesises, S_n and $-S_n$ are identically distributed, the arguments of [15] to prove her theorem 4.5 still hold. Thus, (4.18) is also valid for any λ such that $Im\lambda$ is negative definite.

4.7 Spectrum of S_n

• From step 2 to step 3

Sticking to the proof of Lemma 5.5 of [15], we get that,

$$l(\lambda) := tr_m(L(\lambda 1_m), \lambda \in \mathbb{C} \setminus \mathbb{R},$$

is the Stieljes transform of a distribution Λ with compact support in $sp(s)$. Hence, the proof described in section 3.2 still holds (with $\alpha = 4$ and $k = 23$); thus we can state that for any smooth function φ with compact support

$$\mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))] = (\mathrm{tr}_m \otimes \tau)(\varphi(s)) + \frac{1}{n}\Lambda(\varphi) + O(\frac{1}{n^2}). \qquad (4.19)$$

Moreover, following the proof of Lemma 5.6 in [15], one can show that $\Lambda(1) = 0$ and deduce that, for φ smooth, constant outside a compact set and such that $supp(\varphi) \cap sp(s) = \emptyset$,

$$\mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))] = O(\frac{1}{n^2}).$$

• Step 4

The proof of step 4 is exactly the same as in [9] so that we have proved that, for any $\varepsilon > 0$ and almost surely

$$Spect(S_n) \subset Spect(s) + (-\varepsilon, \varepsilon)$$

when n goes to infinity. Note that this result implies that

$$sup_n \|X_n^{(p)}\| < +\infty \quad a.e.$$

4.8 The main theorem

4.8.1 First inequality

By the same arguments of [9] in Proposition 7.3, we can deduce the following inequality from the above inclusion of the spectrum of S_n.

Proposition 4.3 *Almost everywhere, for all polynomials p in r non commuting variables,*

$$\limsup_{n \to +\infty} \|p(X_n^{(1)}, \ldots, X_n^{(r)}\| \leq \|p(x_1, \ldots, x_r)\|.$$

4.8.2 Second inequality

Proposition 4.4 *Almost everywhere, for all polynomials p in r non commuting variables,*

$$\liminf_{n\to+\infty} \|p(X_n^{(1)}, \dots, X_n^{(r)}\| \geq \|p(x_1, \dots, x_r)\|.$$

¿From [9], Proof of Lemma 7.2, it is clear that this proposition follows from the almost sure asymptotic freeness of the $X_n^{(i)}$ together with the property that

$$a.e \sup_n \|X_n^{(i)}\| < +\infty.$$

The proof of the first point follows the proof of Theorem 6.2 in [15]; nevertheless, we modify the proof of Lemma 6.5 in [15] to get the analogue in our context without needing such a result as Lemma 6.4 in [15].

Lemma 4.3 *Let d be in $\mathbb{N}^*$, $i_1, \dots, i_d$ be in $\{1, \dots, r\}$ and n be in $\mathbb{N}^*$. Define $f : M_n(\mathbb{C})^r \to \mathbb{C}$ by*

$$f(v_1, \dots, v_r) = \mathrm{tr}_n(v_{i_1} \dots v_{i_d}).$$

Then, there is a constant $C > 0$ such that

$$\mathbf{V} f(X_n^{(1)}, \dots, X_n^{(r)}) \leq \frac{C}{n^2}.$$

Proof: Applying Poincaré Inequality (4.9), we get

$$\mathbf{V} f(X_n^{(1)}, \dots, X_n^{(r)}) \leq \frac{C}{n} \mathbb{E}\{\|\nabla f(X_n^{(1)}, \dots, X_n^{(r)})\|_e^2\}.$$

Now, let $v = (v_1, \dots, v_r)$ be in $M_n(\mathbb{C})^r$ and $w = (w_1, \dots, w_r)$ be in $M_n(\mathbb{C})^r_{sa}$ with $\|w\|_e = 1$. By the Cauchy Schwartz inequality

$$\begin{aligned}
|\frac{d}{dt}_{t=0} f(v + tw)| &= \frac{1}{n} |\,\mathrm{Tr}_n(w_{i_1} v_{i_2} v_{i_3} \dots v_{i_d}) + \mathrm{Tr}_n(v_{i_1} w_{i_2} v_{i_3} \dots v_{i_d}) + \dots \\
&\qquad\qquad \dots + \mathrm{Tr}_n(v_{i_1} v_{i_2} \dots v_{i_{d-1}} w_{i_d})| \\
&\leq \frac{1}{n} \left\{ \|w_{i_1}\|_2 \|v_{i_2} v_{i_3} \dots v_{i_d}\|_2 + \dots + \|w_{i_d}\|_2 \|v_{i_1} v_{i_2} \dots v_{i_{d-1}}\|_2 \right\} \\
&\leq \frac{1}{n} \left\{ \|v_{i_2} v_{i_3} \dots v_{i_d}\|_2 + \|v_{i_1} v_{i_3} \dots v_{i_d}\|_2 + \dots + \|v_{i_1} v_{i_2} \dots v_{i_{d-1}}\|_2 \right\}
\end{aligned}$$

Thus,

$$\begin{aligned}\|\nabla f(X_n^{(1)},\dots,X_n^{(r)})\|_e^2 &\le \frac{1}{n^2}\Big[\big\{\mathrm{Tr}_n(X_n^{(i_2)}X_n^{(i_3)}\dots X_n^{(i_d)}X_n^{(i_d)}\dots X_n^{(i_2)})\big\}^{\frac12}\\ &\qquad+\dots+\big\{\mathrm{Tr}_n(X_n^{(i_1)}\dots X_n^{(i_{d-1})}X_n^{(i_{d-1})}\dots X_n^{(i_1)})\big\}^{\frac12}\Big]^2\\ &\le \frac{C}{n^2}\big\{\mathrm{Tr}_n(X_n^{(i_2)}X_n^{(i_3)}\dots X_n^{(i_d)}X_n^{(i_d)}\dots X_n^{(i_2)})\\ &\qquad+\dots+\mathrm{Tr}_n(X_n^{(i_1)}\dots X_n^{(i_{d-1})}X_n^{(i_{d-1})}\dots X_n^{(i_1)})\big\}\end{aligned}$$

for some constant C depending on d, and we get that

$$\begin{aligned}\mathbb{E}\left(\|\nabla f(X_n^{(1)},\dots,X_n^{(r)})\|_e^2\right) &\le \frac{C}{n}\big\{\mathbb{E}\left(\mathrm{tr}_n(X_n^{(i_2)}X_n^{(i_3)}\dots X_n^{(i_d)}X_n^{(i_d)}\dots X_n^{(i_2)})\right)\\ &\qquad+\dots+\mathbb{E}\left(\mathrm{tr}_n(X_n^{(i_1)}\dots X_n^{(i_{d-1})}X_n^{(i_{d-1})}\dots X_n^{(i_1)})\right)\big\}.\end{aligned}$$

Each term inside the brackets of the left hand side is uniformly bounded in n since it converges as n tends to infinity according to the result of asymptotic freeness in mean of Dykema in [6]. The result follows. □

Lemma 4.3 yields the almost sure asymptotic freeness of the $X_n^{(i)}$ using the Borel Cantelli lemma.

In conclusion,

Theorem 4.2 *Let $X_n^{(1)},\dots,X_n^{(r)}$ be independent Wigner matrices associated to a symmetric distribution μ which satisfies a Poincaré inequality. Let $(x_1,\dots,x_r)$ be a semicircular system. Then, almost everywhere, for all polynomials p in r non commuting variables*

$$\lim_{n\to+\infty}\|p(X_n^{(1)},\dots,X_n^{(r)})\| = \|p(x_1,\dots,x_r)\|.$$

5 The Wishart case

We consider a $n\times n$ Hermitian matrix Y, distributed as a Wishart matrix of parameter $p(n)\ge n$ and variance $\frac{1}{n}$ that is with density w.r.t the Lebesgue measure dM on $\mathcal{M}_{sa}(\mathbb{C})$:

$$C_{n,p}1_{(M\ge 0)}(\det(M))^{p(n)-n}\exp(-n\,\mathrm{tr}(M)).$$

We assume that $\frac{p(n)}{n} \underset{n\to\infty}{\longrightarrow} \alpha$ for some $\alpha \geq 1$. More precisely, according to Dirichlet theorem ([16], Lemme 14.1), there exists subsequences $p(n)$ and $q(n)$ of integers tending to ∞ such that:

$$|\frac{p(n)}{q(n)} - \alpha| \leq \frac{1}{q(n)^2}.$$

So, we shall consider a matrix Y of size $q(n)$ and parameter $p(n)$. For simplicity, we shall denote the subsequence $q(n)$ by n and therefore, we will assume in this section that:

$$|\frac{p(n)}{n} - \alpha| \leq \frac{1}{n^2}. \tag{5.1}$$

It is well know that the spectral measure of Y converges to the so called Marchenko-Pastur distribution μ_α [14]:

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

5.1 Differentiation formula for the Wishart ensemble

Lemma 5.1 *Let Φ a C^1 function on $\mathcal{M}_{sa}(\mathbb{C})$ with $\Phi(0) = 0$, then:*

$$\mathbb{E}\left[\Phi'(Y).H\right] - n\mathbb{E}\left[\Phi(Y)\operatorname{Tr}_n(H)\right] + (p(n)-n)\mathbb{E}\left[\Phi(Y)\operatorname{Tr}_n(Y^{-1}H)\right] = 0 \tag{5.2}$$

for all hermitian matrix H, or by linearity for $H = E_{jk}$, $1 \leq j,k \leq n$.

Proof: Since the Lebesgue measure is invariant by translation,

$$\mathbb{E}[\Phi(Y)] = \int \Phi(M+\epsilon X)\exp(-n\operatorname{Tr}_n(M+\epsilon X))(\det(M+\epsilon X))^{p(n)-n} 1_{(M+\epsilon X\geq 0)}dM.$$

Now, by differentiation with respect to ϵ and taking $\epsilon = 0$, we obtain (5.2) using $\partial(\det M)^s = s(\det M)^s M^{-1}$.

5.2 The master equation

Let $(X_n^{(1)}, \ldots, X_n^{(r)})_{i=1,\ldots r}$ be r independent copies of the random matrix Y. We shall apply (5.2) with

$$\Phi(X_n^{(l)}) = \left[(1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1}\right]_{jk} \in \mathcal{M}_m(\mathbb{C})$$

and $H = E_{jk}$. Then,

$$\Phi'(X_n^{(l)}).E_{jk} = (\lambda \otimes 1_n - S_n)^{-1}_{kk} + [(1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1}]_{jj} a_l [(\lambda \otimes 1_n - S_n)^{-1}]_{kk}$$

and

$$\Phi(X_n^{(l)}) \operatorname{Tr}_n((X_n^{(l)})^{-1} E_{jk}) = (X_n^{(l)})^{-1}_{kj} \left[(1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1}\right]_{jk}.$$

The sum over j of the terms in the above equation gives:

$$(\lambda \otimes 1_n - S_n)^{-1}_{kk}.$$

Now, if we sum the identities obtained by (5.2) over j, k, and dividing by n^2, we obtain:

$$\begin{aligned} &\mathbb{E}[id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})] \\ &\quad + \mathbb{E}[id \otimes \operatorname{tr}_n((1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1})\ a_l\ id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})] \\ &\quad - \mathbb{E}[id \otimes \operatorname{tr}_n((1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1})] \\ &\quad + (\frac{p(n)}{n} - 1)\mathbb{E}[id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})] = 0 \end{aligned} \tag{5.3}$$

which can be written as:

$$\mathbb{E}\left[id \otimes \operatorname{tr}_n((1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1})(1_m - a_l\ id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})\right]$$

$$= \frac{p(n)}{n}\mathbb{E}[id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})]. \tag{5.4}$$

Proposition 5.1 *1. For $\lambda \in \mathcal{O}$,*

$$|\mathbb{E}\left[id \otimes \operatorname{tr}_n((1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1})(1_m - a_l\ id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})\right] -$$
$$\mathbb{E}\left[id \otimes \operatorname{tr}_n((1_m \otimes X_n^{(l)})(\lambda \otimes 1_n - S_n)^{-1})\right] \mathbb{E}\left[(1_m - a_l\ id \otimes \operatorname{tr}_n((\lambda \otimes 1_n - S_n)^{-1})\right]|$$
$$\leq \frac{P_4(||(Im(\lambda))^{-1}||)}{n^2}.$$

2. For $\lambda \in \mathcal{O}$,

$$||G_n(\lambda)^{-1}|| \leq (||\lambda|| + K)^2 ||(Im(\lambda))^{-1}||$$

3. If a_l is invertible and $\lambda \in \mathcal{O}$, then $(1_m - a_l G_n(\lambda))$ is invertible and

$$||(1_m - a_l G_n(\lambda))^{-1}|| \leq ||a_l^{-1}||(||\lambda|| + K)^2 ||(Im(\lambda))^{-1}||.$$

If $||(Im(\lambda))^{-1}|| < \frac{1}{2||a_l||}$, then $(1_m - a_l G_n(\lambda))$ is invertible and

$$||(1_m - a_l G_n(\lambda))^{-1}|| \leq 2. \tag{5.5}$$

Sketch of Proof:
1. The variance estimate follows from the Gaussian Poincaré inequality since we can write $Y = \frac{1}{n} X^* X$ for a rectangular Gaussian matrix X. We proceed as in [9, Section 4]. We need some estimate on the maximal eigenvalue of Y, i.e. $\mathbb{E}[\lambda_{max}]$ and $\mathbb{E}[\lambda_{max}^3]$ are bounded, independently of n. This can be proved, as in Lemma 5.1 of [9], using previous results in [8].
2. The proof is the same as Proposition 5.2 in [9].
3. If a_l is invertible,

$$(1 - a_l G_n(\lambda)) = a_l (a_l^{-1} - G_n(\lambda)) := a_l T.$$

Now, the matrix T satisfies, $Im(T) = -Im(G_n(\lambda))$ and thus is positive definite (see the proof of Proposition 5.2 in [9]). Its inverse T^{-1} satisfies:

$$||T^{-1}|| \leq ||Im(T)^{-1}|| = ||Im(G_n(\lambda))^{-1}|| \leq (||\lambda|| + K)^2 ||(Im(\lambda))^{-1}||.$$

(5.5) follows from the majoration $||G_n(\lambda)|| \leq ||(Im(\lambda))^{-1}||$. $\square$

From (5.4) and Proposition 5.1, we obtain:

$$\Bigg|\Bigg| \sum_{l=1}^{r} a_l \mathbb{E}[id \otimes \mathrm{tr}_n((1_m \otimes Y^{(l)})(\lambda \otimes 1_n - S_n)^{-1})] - \frac{p(n)}{n} \sum_{l=1}^{r} a_l \mathbb{E}[H_n(\lambda)](1_m - a_l \mathbb{E}[H_n(\lambda)])^{-1} \Bigg|\Bigg| \leq (||\lambda|| + K)^2 \frac{P_5(||(Im(\lambda))^{-1}||)}{n^2}$$

The first line of the above equation equals:

$$\mathbb{E}[id \otimes \mathrm{tr}_n(S_n - a_0 \otimes 1_n)(\lambda \otimes 1_n - S_n)^{-1})] = -1_m + (\lambda - a_0)\mathbb{E}[H_n(\lambda)].$$

We have thus obtain the master inequality:

$$||-a_0 G_n(\lambda) + \lambda G_n(\lambda) - \frac{p(n)}{n} \sum_{l=1}^{r} a_l G_n(\lambda)(1_m - a_l G_n(\lambda))^{-1} - 1_m|| \leq (||\lambda|| + K)^2 \frac{P_5(||(Im(\lambda))^{-1}||)}{n^2} \tag{5.6}$$

or since the matrices $(1_m - a_l G_n(\lambda))^{-1}$ and $a_l G_n(\lambda)$ commute,

$$||-a_0 G_n(\lambda)+\lambda G_n(\lambda)-\frac{p(n)}{n}\sum_{l=1}^{r}(1_m-a_l G_n(\lambda))^{-1}a_l G_n(\lambda)\ -1_m|| \leq (||\lambda||+K)^2\frac{P_5(||(Im(\lambda))^{-1}||)}{n^2} \tag{5.7}$$

5.3 Estimation of $||G_n(\lambda) - G(\lambda)||$

Let $x_i, i \leq r$ be a free family of self adjoint variables in a C^*-probability space $(\mathcal{B}, \tau)$, with Marchenko-Pastur distribution μ_α, with parameter α.

$$s = a_0 \otimes 1_{\mathcal{B}} + \sum_{i=1}^{r} a_i \otimes x_i$$

Using the known expression of the R transform of the distribution of x_i (see [5], [10, Example 3.3.5][2]):

$$R_x(z) = \alpha(1-z)^{-1}, z \in \mathbb{C}\backslash\mathbb{R},$$

we can show the following

Lemma 5.2 *G satisfies the following equation: for $\lambda \in \mathcal{O}$,*

$$a_0 + \alpha\sum_{i=1}^{r}(1_m - a_i G(\lambda))^{-1}a_i + G(\lambda)^{-1} = \lambda. \tag{5.8}$$

Sketch of Proof: From the definition of the R transformation with amalgation over $M_m(\mathbb{C})$, we can show that:

$$R_{a\otimes x}(\lambda) = R_x(a\lambda)a,\ a \in M_m(\mathbb{C})_{sa}, \lambda \in \mathcal{O}$$

and then, by freeness asumption

$$R_{\sum_{i=1}^r a_i \otimes x_i}(\lambda) = \sum_{i=1}^{r} R_{x_i}(a_i\lambda)a_i.$$

(5.8) follows, using the relation between R and G. $\square$.

[2]We warm the reader that the R transform defined in this book differs by a factor z from the Voiculescu R transform we used here

Theorem 5.1 *For any $\lambda \in \mathcal{O}$,*

$$||G(\lambda) - G_n(\lambda)|| \leq (||\lambda|| + K)^4 \frac{P_8(||(Im(\lambda))^{-1}||)}{n^2}. \tag{5.9}$$

Proof: We can proceed as in the proof of Theorem 5.7 in [9]. We just mention the different steps:

Step 1: Define $\Lambda_n(\lambda) = a_0 + G_n(\lambda)^{-1} + \alpha \sum_l (1_m - a_l G_n(\lambda))^{-1} a_l$. From the master inequality (5.7), Proposition 5.1 and (5.1), we can show that:

$$||\lambda - \Lambda_n(\lambda)|| \leq (||\lambda|| + K)^4 \frac{P_6(||(Im(\lambda))^{-1}||)}{n^2}.$$

Then, for $\lambda \in \mathcal{O}'_n$ of the form

$$\mathcal{O}'_n = \left\{ \lambda \in \mathcal{O}, (||\lambda|| + K)^4 \frac{P_6(||(Im(\lambda))^{-1}||)}{n^2} < \frac{1}{2} \frac{1}{||(Im(\lambda))^{-1}||} \right\},$$

we have $Im(\Lambda_n(\lambda)) \geq \frac{1}{2||(Im(\lambda))^{-1}||}$ and in particular $\Lambda_n(\lambda) \in \mathcal{O}$.

Step 2: For $\lambda \in \mathcal{O}'_n$, we can consider $G(\Lambda_n(\lambda))$ and we have, from the identity (5.8):

$$a_0 + G_n(\lambda)^{-1} + \alpha \sum_l (1_m - a_l G_n(\lambda))^{-1} a_l = a_0 + G(\Lambda_n(\lambda))^{-1} + \alpha \sum_l (1_m - a_l G(\Lambda_n(\lambda)))^{-1} a_l. \tag{5.10}$$

Lemma 5.3 *(see [9], Propostion 5.6) For $\lambda \in \mathcal{O}'_n$,*

$$G(\Lambda_n(\lambda)) = G_n(\lambda). \tag{5.11}$$

Proof: As in [9], it's enough to prove (5.11) for $\lambda \in \mathcal{O}''_n$, a non empty subset of the connected subset $\mathcal{O}'_n$. Put $x = G_n(\lambda)$ and $y = G(\Lambda_n(\lambda))$, then, from (5.10),

$$a_0 + x^{-1} + \alpha \sum_l (1_m - a_l x)^{-1} a_l = a_0 + y^{-1} + \alpha \sum_l (1_m - a_l y)^{-1} a_l$$

so that

$$y + \alpha \sum_l x(1_m - a_l x)^{-1} a_l \, y = x + \alpha \sum_l x(1_m - a_l y)^{-1} a_l \, y,$$

Thus,

$$\begin{aligned} y - x &= \alpha \sum_l x[(1_m - a_l y)^{-1} - (1_m - a_l x)^{-1}] a_l\ y \\ &= \alpha \sum_l x(1_m - a_l y)^{-1}[(1 - a_l x) - (1 - a_l y)](1_m - a_l x)^{-1} a_l\ y \\ &= \alpha \sum_l x(1_m - a_l y)^{-1} a_l (y - x)(1_m - a_l x)^{-1} a_l\ y \end{aligned}$$

In particular, we have,

$$\|y - x\| \leq \left(\alpha \|x\| \|y\| \sum_l \|(1_m - a_l y)^{-1}\| \|(1_m - a_l x)^{-1}\| \|a_l\|^2 \right) \|y - x\| \tag{5.12}$$

Now, we have

$$\|x\| = \|G_n(\lambda)\| \leq ||(Im(\lambda))^{-1}||$$

and

$$\|y\| = \|G(\Lambda_n(\lambda))\| \leq (||(Im(\Lambda_n(\lambda)))^{-1}|| \leq 2||(Im(\lambda))^{-1}||$$

for $\lambda \in \mathcal{O}'_n$ (see Step 1).
Moreover, from Proposition 5.1, for $||(Im(\lambda))^{-1}||$ small enough,

$$\|(1_m - a_l x)^{-1}\| \leq 2;\ \|(1_m - a_l y)^{-1}\| \leq 2.$$

Set

$$\mathcal{O}''_n = \{\lambda \in \mathcal{O}'_n; \alpha \|G_n(\lambda)\| \|G(\Lambda_n(\lambda))\| \sum_l \|(1_m - a_l G(\Lambda_n(\lambda)))^{-1}\| \|(1_m - a_l G_n(\lambda))^{-1}\| \|a_l\|^2 < 1\},$$

Then, from (5.12), for $\lambda \in \mathcal{O}''_n$, $G(\Lambda_n(\lambda)) = G_n(\lambda)$. Now, it is easy to see, from the above estimates, that $\lambda = it1_m \in \mathcal{O}''_n$ for t large enough, so $\mathcal{O}''_n$ is a non empty set. $\square$

Step 3: The estimation of $G(\Lambda_n(\lambda)) - G(\lambda)$ is obtained as in Subsection 3.1 (considering the two cases $\lambda \in \mathcal{O}'_n$ and $\lambda \in \mathcal{O} \backslash \mathcal{O}'_n$). $\square$

5.4 The spectrum of S_n

¿From Theorem 5.1 and the proof described in Section 3.2 (see also Section 6 in [9]), we can prove that for φ smooth, constant outside a compact set and such that $supp(\varphi) \cap sp(s) = \emptyset$

$$\mathbb{E}[(\mathrm{tr}_m \otimes \mathrm{tr}_n)(\varphi(S_n))] = O(\frac{1}{n^2}).$$

from which we deduce that, for any $\varepsilon > 0$ and almost surely

$$Spect(S_n) \subset Spect(s) + (-\varepsilon, \varepsilon)$$

when n goes to infinity.

5.5 The main theorem

We can now prove:

Theorem 5.2 *There exists a set N of probability 0 such that for all non commutative polynomial p in r variables, and all $\omega \in \Omega \backslash N$,*

$$\lim_{n \longrightarrow \infty} ||p(X_n^{(1)}(\omega), \ldots, X_n^{(r)}(\omega))|| = ||p(x_1, \ldots x_r)||. \tag{5.13}$$

Proof: The inequality

$$\limsup_{n \longrightarrow \infty} ||p(X_n^{(1)}, \ldots, X_n^{(r)})|| \leq ||p(x_1, \ldots x_r)|| \ a.s. \tag{5.14}$$

follows from the above inclusion of the spectrum of S_n and the arguments developed in [9], Section 7. The reverse inequality

$$\liminf_{n \longrightarrow \infty} ||p(X_n^{(1)}, \ldots, X_n^{(r)})|| \geq ||p(x_1, \ldots x_r)|| \ a.s. \tag{5.15}$$

follows, as in Lemma 7.2 in [9], from the a.s. asymptotic freeness of the $(X_n^{(i)})_{i=1,\ldots,r}$ and $\sup_n \|X_n^{(i)}\| < \infty$ a.s.. The first point was proved by Hiai and Petz (see [10], [11]) and the second point follows from (5.14). $\square$

Remark: If we only assume the convergence of $\frac{p(n)}{n}$ to α with $|\frac{p(n)}{n} - \alpha| \leq \frac{1}{n}$, then an extra term appears in the estimation of $G - G_n$ at order n^{-2}, namely:

$$||G(\lambda) - G_n(\lambda) + (\alpha - \frac{p(n)}{n})L(\lambda)|| \leq \frac{C(\lambda)}{n^2}$$

with

$$L(\lambda) = (id_m \otimes \tau)[(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}(R(\lambda) \otimes 1_{\mathcal{B}})(\lambda \otimes 1_{\mathcal{B}} - s)^{-1}]$$

and $R(\lambda) = \sum_l (1_m - a_l G(\lambda))^{-1} a_l$.
As in Schultz [15] and in the iid case (see Section 4), this term gives rise to a distribution with compact support in $sp(s)$ and the conclusion remains true.

References

[1] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., Scheffer, G. *Sur les inégalités de Sobolev logarithmiques.* (French) [Logarithmic Sobolev inequalities] Panoramas et Synthèses [Panoramas and Syntheses], 10. Société Mathématique de France, Paris, 2000.

[2] Bai, Z. D.: *Methodology in Spectral analysis of large dimensional random matrices. A review.* Statistica Sinica **9** (1999), pp 611-677.

[3] Bai, Z.D., Yin, Y. Q.: *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix.* Ann. of Proba. **16** (1988), pp 1729-1741.

[4] Bobkov, S., Gotze.: *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.* J. Funct. Anal. **163** (1999), pp 1-28.

[5] Capitaine, M. and Casalis, M.:*Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices.* Indiana Univ. Math. J. **53** (2004), pp 397–431.

[6] Dykema, K.: *On Certain Free Product Factors via an Extended Matrix Model.* J. Funct. Anal. **112** (1993), pp 31-60.

[7] Geman S.: *A limit theorem for the norm of random matrices.* Annals of Probability 8 (1980) pp 252-261.

[8] Haagerup, U. and Thorbjørnsen, S.: *Random matrices with complex Gaussian entries.* Expo. Math. **21** (2003), pp 293–337.

[9] Haagerup, U. and Thorbjørnsen, S.: *A new application of random matrices:* $Ext(C^*_{red}(F_2))$ *is not a group.* To appear in Ann. Math.

[10] Hiai, F. and Petz, D.: "The semicircle law, free random variables and entropy." Mathematical Surveys and Monographs, 77. American Mathematical Society, Providence, RI, 2000.

[11] Hiai, F. and Petz, D.: *Asymptotic freeness almost everywhere for random matrices.* Acta Sci. Math. **66** (2000) pp 809–834.

[12] Khorunzhy, A., Khoruzhenko, B., Pastur,L.: *Asymptotic properties of large random matrices with independent entries*, J. Math. Phys. **37** (1996), pp 5033-5060.

[13] Ledoux, L.: "The Concentration of Measure Phenomenon." Mathematical Surveys and Monographs, Volume 89, A.M.S, 2001.

[14] Marchenko, V. and Pastur, L.: *The distribution of eigenvalues in a certain sets of random matrices.* Math. Sb. **72** (1967), pp 507-536.

[15] Schultz, H.: *Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases.* To appear in *Prob. Th. Rel. Fields.*

[16] Tenenbaum, G.: "Introduction à la théorie analytique et probabiliste des nombres." Institut Elie Cartan, Nancy, 1990.

[17] Thorbjørnsen, S.:*Mixed moments of Voiculescu's Gaussian Random matrices.* J. Funct. Anal. **176** (2000), pp 213-246.

[18] Voiculescu, D.: *Limit laws for random matrices and free products.* Invent. Math. **104** (1991), pp 201-220.