Lecture notes on random matrices

Alice Guionnet

January 2019

In this course we will discuss the law of large numbers and the central limit theorem in random matrix theory.

Contents

1 Method of moments

1.1 Law of Large Numbers (LLN): Wigner's theorem

We consider in this section an $N \times N$ matrix $\mathbf{X}^N$ with real or complex entries such that $\left(\mathbf{X}^N_{ij}, 1 \le i \le j \le N\right)$ are independent and $\mathbf{X}^N$ is self-adjoint; $\mathbf{X}^N_{ij} = \overline{\mathbf{X}}^N_{ji}$. We assume further that

$$\mathbb{E}[\mathbf{X}^N_{ij}] = 0, \lim_{N\to\infty} \max_{1\le i,j\le N} |N\mathbb{E}[|\mathbf{X}^N_{ij}|^2] - 1| = 0. \quad (1)$$

We shall show that, under some finite moments conditions on the entries, the eigenvalues $(\lambda_1,\cdots,\lambda_N)$ of $\mathbf{X}^N$ satisfy the almost sure convergence

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}f(\lambda_i)=\int f(x)d\sigma(x) \tag{2}$$

where f is a bounded continuous function or a polynomial function, σ is the semi-circular law

$$\sigma(dx)=\frac{1}{2\pi}\sqrt{4-x^2}1_{|x|\le 2}dx. \tag{3}$$

We shall prove this convergence for polynomial functions and rely on the fact that for all $k\in\mathbb{N}$, $\int x^k d\sigma(x)$ is null when k is odd and given by the Catalan number

$$C_{k/2}=\frac{\begin{pmatrix} k \\ \frac{k}{2}\end{pmatrix}}{\frac{k}{2}+1} \tag{4}$$

when k is even.

In this section, we use the same notation for complex and for real entries since both cases will be treated at once and yield the same result. The aim of this section is to prove

Theorem 1.1. *[Wigner's theorem [13]] Assume that for all $k\in\mathbb{N}$,*

$$B_k:=\sup_{N\in\mathbb{N}}\sup_{(i,j)\in\{1,\cdots,N\}^2}\mathbb{E}[|\sqrt{N}\mathbf{X}^N_{ij}|^k]<\infty. \tag{5}$$

Then,

$$\lim_{N\to\infty}\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)=\begin{cases}0 & \textit{if } k \textit{ is odd,}\\ C_{\frac{k}{2}} & \textit{otherwise,}\end{cases} \tag{6}$$

where the convergence holds in expectation and almost surely.

The Catalan number C_k will appear here as the number of non-crossing pair partitions of $2k$ elements. Namely, recall that a partition of the (ordered) set $S:=\{1,\cdots,n\}$ is a decomposition

$$\pi=\{V_1,\cdots,V_r\} \tag{7}$$

such that $V_i\cap V_j=\emptyset$ if $i\neq j$ and $\cup V_i=S$. The $V_i, 1\le i\le r$ are called the blocks of the partition and we say that $p\sim_\pi q$ if p,q belong to the same block of the partition π. A partition π of $\{1,\cdots,n\}$ is said to be *crossing* if there exist $1\le p_1<q_1<p_2<q_2\le n$ with

$$p_1\sim_\pi p_2\not\sim_\pi q_1\sim_\pi q_2. \tag{8}$$

It is *non-crossing* otherwise. We leave it as an exercise to the reader to prove that C_k as given in the theorem is exactly the number of non-crossing pair partitions of $\{1,2,\cdots,2k\}$.

Proof. We start the proof by showing the convergence in expectation, for which the strategy is simply to expand the trace over the matrix in terms of its entries. We then use some (easy) combinatorics on trees to find out the main contributing term in this expansion. The almost sure convergence is obtained by estimating the covariance of the considered random variables and applying Borel-Cantelli lemma.

- *Expanding the expectation.*

  Setting $\mathbf{Y}^N = \sqrt{N}\mathbf{X}^N$, we have

$$\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right] = \sum_{i_1,\cdots,i_k=1}^{N} N^{-\frac{k}{2}-1}\mathbb{E}[Y_{i_1i_2}Y_{i_2i_3}\cdots Y_{i_ki_1}] \tag{9}$$

  where $Y_{ij}, 1 \le i,j \le N$, denote the entries of $\mathbf{Y}^N$ (which may eventually depend on N). We denote $\mathbf{i} = (i_1,\cdots,i_k)$ and set

$$P^N(\mathbf{i}) := \mathbb{E}[Y_{i_1i_2}Y_{i_2i_3}\cdots Y_{i_ki_1}]. \tag{10}$$

  Note that P^N depends on N through $\mathbf{Y}^N$. By (5) and Hölder's inequality, $P(\mathbf{i})$ is bounded uniformly by B_k, independently of $\mathbf{i}$ and N. Since the random variables $(Y_{ij}, i \le j)$ are independent and centered, $P(\mathbf{i})$ equals zero unless for any pair (i_p, i_{p+1}), $p \in \{1,\cdots,k\}$, there exists $l \ne p$ such that $(i_p,i_{p+1}) = (i_l,i_{l+1})$ or (i_{l+1},i_l). Here, we used the convention $i_{k+1} = i_1$. To find more precisely which set of indices contributes to the first order in the right hand side of (9), we next provide some combinatorial insight into the sum over the indices.

- *Connected graphs and trees.*

  $V(\mathbf{i}) = \{i_1,\cdots,i_k\}$ will be called the vertices. We identify i_ℓ and i_p iff they are equal. An edge is a pair (i,j) with $i,j \in \{1,\cdots,N\}^2$. At this point, edges are directed in the sense that we distinguish (i,j) from (j,i) when $j \ne i$ and we shall point out later when we consider undirected edges. We denote by $E(\mathbf{i})$ the collection of the k edges $(e_p)_{p=1}^k = (i_p,i_{p+1})_{p=1}^k$.

  We consider the graph $G(\mathbf{i}) = (V(\mathbf{i}),E(\mathbf{i}))$. $G(\mathbf{i})$ is connected by construction. Note that $G(\mathbf{i})$ may contain loops (i.e cycles, for instance edges of type (i,i)) and multiple undirected edges.

  The skeleton $\tilde{G}(\mathbf{i})$ of $G(\mathbf{i})$ is the graph $\tilde{G}(\mathbf{i}) = \left(\tilde{V}(\mathbf{i}),\tilde{E}(\mathbf{i})\right)$ where vertices in $\tilde{V}(\mathbf{i})$ appear only once, edges in $\tilde{E}(\mathbf{i})$ are undirected and appear only once.

  In other words, $\tilde{G}(\mathbf{i})$ is the graph $G(\mathbf{i})$ where multiplicities and orientation have been erased. It is connected, as is $G(\mathbf{i})$.

  We now state and prove a well known inequality concerning undirected connected graphs $G = (V,E)$. If we let, for a discrete finite set A, $|A|$ be the number of its distinct elements, we have the following inequality

$$|V| \le |E| + 1. \tag{11}$$

  Let us prove this inequality and that equality holds only if G is a tree at the same time. This relation is straightforward when $|V| = 1$ and can be proven by induction as follows. Assume $|V| = n$ and consider one vertex v of V. This vertex is contained in l edges of E which we denote $(e_1,\cdots,e_l)$ and with $l \ge 1$ by connectedness. The graph G then decomposes into $(\{v\},\{e_1,\cdots,e_l\})$ and $r \le l$ undirected connected graphs $(G_1,\cdots,G_r)$. We denote $G_j = (V_j,E_j)$, for $j \in \{1,\cdots,r\}$. We have

$$|V| - 1 = \sum_{j=1}^{r}|V_j|, \quad |E| - l = \sum_{j=1}^{r}|E_j|. \tag{12}$$

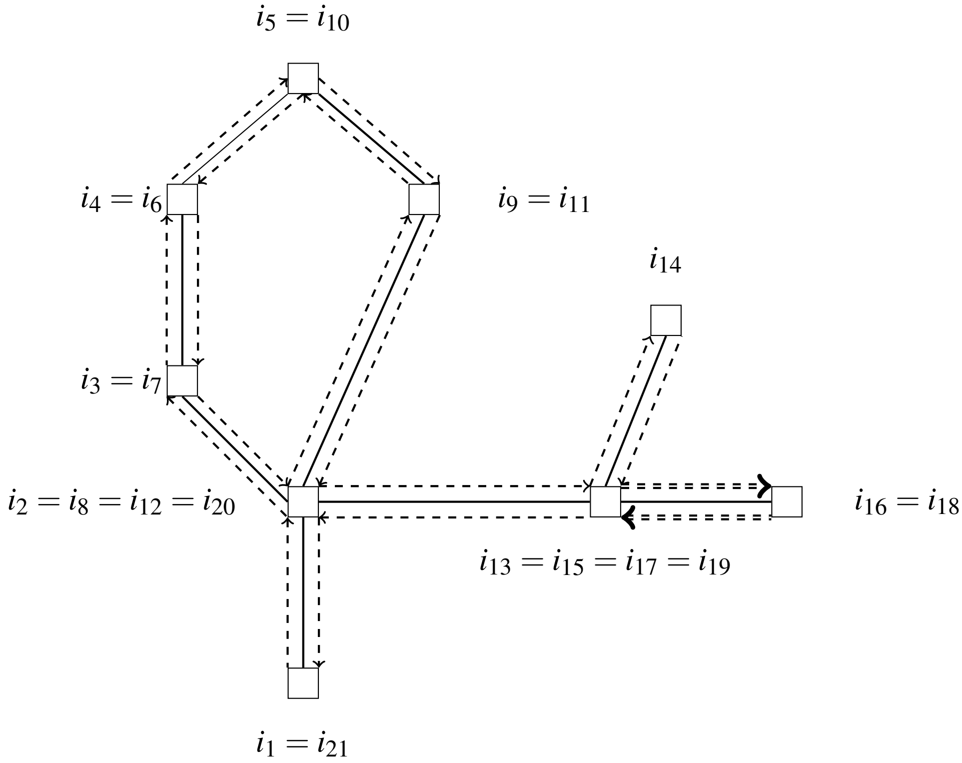

Figure 1: Figure of $G(\mathbf{i})$ (in dash) versus $\tilde{G}(\mathbf{i})$ (in bold), $|\tilde{E}(i)| = 9, |\tilde{V}(\mathbf{i})| = 9$

Applying the induction hypothesis to the graphs $(G_j)_{1\le j\le r}$ gives

$$\begin{aligned} |V|-1 &\le \sum_{i=1}^{r}(|E_j|+1) \\ &= |E|+r-l \le |E| \end{aligned} \tag{13}$$

which proves (11). In the case where $|V| = |E|+1$, we claim that G is a tree, namely G does not have a loop. In fact, for equality to hold, we need to have equalities when performing the previous decomposition of the graph, a decomposition which can be reproduced until all vertices have been considered. If the graph contains a loop, the first time that we erase a vertex of this loop when performing this decomposition, we will create one connected component less than the number of edges we erased and so a strict inequality occurs in the right hand side of (13) (i.e. $r < l$).

- *Convergence in expectation.*

  Since we noticed that $P(\mathbf{i})$ equals zero unless each edge in $E(\mathbf{i})$ is repeated at least twice, we have that

$$|\tilde{E}(\mathbf{i})| \le 2^{-1}|E(\mathbf{i})| = \frac{k}{2}, \tag{14}$$

  and so by (11) applied to the skeleton $\tilde{G}(\mathbf{i})$ we find

$$|\tilde{V}(\mathbf{i})| \le \lfloor \frac{k}{2} \rfloor + 1 \tag{15}$$

where $\lfloor x \rfloor$ is the integer part of x. Thus, since the indices are chosen in $\{1, \cdots, N\}$, there are at most $N^{\lfloor \frac{k}{2} \rfloor + 1}$ indices which contribute to the sum (9) and so we have

$$\left| \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right] \right| \leq B_k N^{\lfloor \frac{k}{2} \rfloor - \frac{k}{2}} \tag{16}$$

where we used (5) and Hölder's inequality. In particular, if k is odd,

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right] = 0. \tag{17}$$

If k is even, the only indices which will contribute to the first order asymptotics in the sum are those such that

$$|\tilde{V}(\mathbf{i})| = \frac{k}{2} + 1, \tag{18}$$

since the other indices will be such that $|\tilde{V}(\mathbf{i})| \leq \frac{k}{2}$ and so will contribute at most by a term $N^{\frac{k}{2}} B_k N^{-\frac{k}{2}-1} = O(N^{-1})$. By the previous considerations, when $|\tilde{V}(\mathbf{i})| = \frac{k}{2} + 1$, we have that

1. $\tilde{G}(\mathbf{i})$ is a tree,
2. $|\tilde{E}(\mathbf{i})| = 2^{-1}|E(\mathbf{i})| = \frac{k}{2}$ and so each edge in $E(\mathbf{i})$ appears exactly twice.

We can explore $G(\mathbf{i})$ by following the path P of edges $i_1 \to i_2 \to i_3 \cdots \to i_k \to i_1$. Since $\tilde{G}(\mathbf{i})$ is a tree, $G(\mathbf{i})$ appears as a fat tree where each edge of $\tilde{G}(\mathbf{i})$ is repeated exactly twice. We then see that each pair of directed edges corresponding to the same undirected edge in $\tilde{E}(\mathbf{i})$ is of the form $\{(i_p, i_{p+1}), (i_{p+1}, i_p)\}$ (since otherwise the path of edges has to form a loop to return to i_0). Therefore, for these indices, $\lim_N P^N(\mathbf{i}) = \lim_N E[|\sqrt{N}X_{ij}^N|^2]^{\frac{k}{2}} = 1$ does not depend on $\mathbf{i}$.

Finally, observe that $G(\mathbf{i})$ gives a pair partition of the edges of the path P (since each undirected edge has to appear exactly twice) and that this partition is non crossing (as can be seen by unfolding the path, keeping track of the pairing between edges by drawing an arc between paired edges). Moreover, this pair partition is the same for $\mathbf{i}$ and $\mathbf{j}$ iff the graphs $G(\mathbf{i})$ and $G(\mathbf{j})$ are isomorphic (that is they corresponding to a different labeling of the vertices of the same rooted graph). For a given graph G, there corresponds approximately $N^{|V|}$ possible labeling of its vertices corresponding to isomorphic graphs $G(\mathbf{i})$. Hence, only the graphs with maximum numbers of vertices, that is trees, will contribute to the leading order, that is $|V| = \frac{k}{2} + 1$. Therefore we have proved that

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right] = \sharp\{\text{ non-crossing pair partitions of } k \text{ edges }\}. \tag{19}$$

- *Almost sure convergence.* To prove the almost sure convergence, we estimate the variance and then use Borel Cantelli's lemma. The variance is given by

$$\begin{aligned}\operatorname{Var}((\mathbf{X}^N)^k) &:= \mathbb{E}\left[\frac{1}{N^2}\left(\operatorname{Tr}\left((\mathbf{X}^N)^k\right)\right)^2\right] - \mathbb{E}\left[\frac{1}{N}\operatorname{Tr}\left((\mathbf{X}^N)^k\right)\right]^2 \\ &= \frac{1}{N^{2+k}} \sum_{\substack{i_1,\dots,i_k=1 \\ i'_1,\dots,i'_k=1}}^{N} [P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]\end{aligned} \tag{20}$$

with

$$P(\mathbf{i},\mathbf{i}') := \mathbb{E}[Y_{i_1i_2}Y_{i_2i_3}\cdots Y_{i_ki_1}Y_{i'_1i'_2}\cdots Y_{i'_ki'_1}]. \tag{21}$$

We denote $G(\mathbf{i},\mathbf{i}')$ the graph with vertices $V(\mathbf{i},\mathbf{i}') = \{i_1,\cdots,i_k,i'_1,\cdots,i'_k\}$ and edges $E(\mathbf{i},\mathbf{i}') = \{(i_p,i_{p+1})_{1\le p\le k},(i'_p,i'_{p+1})_{1\le p\le k}\}$. For $\mathbf{i},\mathbf{i}'$ to contribute to the sum, $G(\mathbf{i},\mathbf{i}')$ must be connected. Indeed, if $E(\mathbf{i})\cap E(\mathbf{i}') = \emptyset$, $P(\mathbf{i},\mathbf{i}') = P(\mathbf{i})P(\mathbf{i}')$. Moreover, as before, each edge must appear at least twice to give a non zero contribution so that $|\tilde{E}(\mathbf{i},\mathbf{i}')| \le k$. Therefore, we are in the same situation as before, and if $\tilde{G}(\mathbf{i},\mathbf{i}') = (\tilde{V}(\mathbf{i},\mathbf{i}'),\tilde{E}(\mathbf{i},\mathbf{i}'))$ denotes the skeleton of $G(\mathbf{i},\mathbf{i}')$, we have the relation

$$|\tilde{V}(\mathbf{i},\mathbf{i}')| \le |\tilde{E}(\mathbf{i},\mathbf{i}')| + 1 \le k+1. \tag{22}$$

This already shows that the variance is at most of order N^{-1} (since $P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')$ is bounded uniformly, independently of $(\mathbf{i},\mathbf{i}')$ and N), but we need a slightly better bound to prove the almost sure convergence. To improve our bound let us show that the case where $|\tilde{V}(\mathbf{i},\mathbf{i}')| = |\tilde{E}(\mathbf{i},\mathbf{i}')| + 1 = k+1$ can not occur. In this case, we have seen that $\tilde{G}(\mathbf{i},\mathbf{i}')$ must be a tree since then equality holds in (22). Also, $|\tilde{E}(\mathbf{i},\mathbf{i}')| = k$ implies that each edge appears with multiplicity exactly equal to 2. For any contributing set of indices $\mathbf{i},\mathbf{i}'$, $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i})$ and $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i}')$ must share at least one edge (i.e one edge must appear with multiplicity one in each subgraph) since otherwise $P(\mathbf{i},\mathbf{i}') = P(\mathbf{i})P(\mathbf{i}')$. This is a contradiction. Indeed, if we explore $\tilde{G}(\mathbf{i},\mathbf{i}')$ by following the path $i_1 \to i_2 \to \cdots \to i_1$, we see that each (non-oriented) visited edge appears twice or this path makes a loop. The first case is impossible since $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i})$ and $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i}')$ share one edge and each edge of $\tilde{G}(\mathbf{i},\mathbf{i}')$ has multiplicity 2, and the second case is also impossible since $\tilde{G}(\mathbf{i},\mathbf{i}')$ is a tree. Therefore, we conclude that for all contributing indices,

$$|\tilde{V}(\mathbf{i},\mathbf{i}')| \le k \tag{23}$$

which implies

$$\operatorname{Var}((\mathbf{X}^N)^k) \le p_k N^{-2} \tag{24}$$

with p_k a constant independent of N. Applying Chebychev's inequality gives for any $\delta > 0$

$$\mathbb{P}\left(\left|\frac{1}{N}\operatorname{Tr}\left((\mathbf{X}^N)^k\right) - \mathbb{E}\left[\frac{1}{N}\operatorname{Tr}\left((\mathbf{X}^N)^k\right)\right]\right| > \delta\right) \le \frac{p_k}{\delta^2 N^2}, \tag{25}$$

and so Borel-Cantelli's lemma implies

$$\lim_{N\to\infty}\left|\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)-\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]\right|=0 \quad a.s. \tag{26}$$

The proof of the theorem is complete.

□

Exercise 1.2. *take $\{\mathbf{X}^{N,\ell}, 1\le\ell\le m\}$ be m independent Wigner matrices such that*

$$\mathbb{E}[\mathbf{X}^{N,\ell}_{ij}]=0,\forall 1\le i,j\le N, 1\le\ell\le m, \quad \lim_{N\to\infty} max_{1\le i,j\le N}|N\mathbb{E}[|\mathbf{X}^{N,\ell}_{ij}|^2]-1|=0$$

Assume that for all $k\in\mathbb{N}$,

$$B_k:=\sup_{1\le\ell\le m}\sup_{N\in\mathbb{N}}\sup_{ij\in\{1,\cdots,N\}^2}\mathbb{E}[|\sqrt{N}\mathbf{X}^{N,\ell}_{ij}|^k]<\infty. \tag{27}$$

Then, for any $\ell_j\in\{1,\cdots,m\}, 1\le j\le k$,

$$\lim_{N\to\infty}\frac{1}{N}tr\left(\mathbf{X}^{N,\ell_1}\mathbf{X}^{N,\ell_2}\cdots\mathbf{X}^{N,\ell_k}\right)=\sigma^m(X_{\ell_1}\cdots X_{\ell_k})$$

where the convergence holds in expectation and almost surely. $\sigma^m(X_{\ell_1}\cdots X_{\ell_k})$ is the number $|NP(X_{\ell_1}\cdots X_{\ell_k})|$ of non-crossing pair partitions of labelled points $S(X_{\ell_1}\cdots X_{\ell_k})$ given by $(1,\ell_1),(2,\ell_2),\ldots,(k,\ell_k)$ so that every block contains points with the same label.

Additional question: Extend σ^m by linearity to polynomials. Show that $\sigma^m(1)=1$ and for all polynomials $P_1,\ldots,P_k$ all $i_k\in\{1,\ldots,m\}$ so that $i_{k+1}\neq i_k$

$$\sigma^m\left((P_1(X_{i_1})-\sigma^m(P_1))\cdots(P_k(X_{i_k})-\sigma^m(P_k))\right)=0$$

We say σ^m is the law of m free variables.

Exercise 1.3. *Take for $L\in\mathbb{N}$, $\mathbf{X}^{N,L}$ the $N\times N$ self-adjoint matrix such that $\mathbf{X}^{N,L}_{ij}=(2L)^{-\frac{1}{2}}1_{|i-j|\le L}X_{ij}$ with $(X_{ij}, 1\le i\le j\le N)$ independent centered real random variables having all moments finite and $E[X_{ij}^2]=1$. The purpose of this exercise is to show that for all $k\in\mathbb{N}$,*

$$\lim_{L\to\infty}\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}((\mathbf{X}^{N,L})^k)\right]=C_{k/2} \tag{28}$$

with C_x null if x is not an integer. Moreover, if $L(N)\in\mathbb{N}$ is a sequence going to infinity with N so that $L(N)/N$ goes to zero, prove that

$$\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}((\mathbf{X}^{N,L(N)})^k)\right]=C_{k/2}. \tag{29}$$

If $L(N)=\lfloor\alpha N\rfloor$, one can also prove the convergence of the moments of $\mathbf{X}^{N,L(N)}$. Show that this limit can not be given by the Catalan numbers $C_{k/2}$ by considering the case $k=2$.

Hint: Show that for $k\ge 2$

$$\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}((\mathbf{X}^{N,L})^k)\right]=(2L)^{-k/2}\sum_{\substack{|i_2-\lfloor\frac{N}{2}\rfloor|\le L,\\ |i_{p+1}-i_p|\le L, p\ge 2}}\mathbb{E}[X_{\lfloor\frac{N}{2}\rfloor i_2}\cdots X_{i_k\lfloor\frac{N}{2}\rfloor}]+O(N^{-1}). \tag{30}$$

Then prove that the contributing indices to the above sum correspond to the case where $G(0,i_2,\cdot,i_k)$ is a tree with $k/2$ vertices and show that being given a tree there are approximately $(2L)^{\frac{k}{2}}$ possible choices of indices $i_2,\cdots,i_k$.

Exercise 1.4. *Show that we have the same result if instead of* (5) *we assume*

$$\limsup_{N\in\mathbb{N}} N \sup_{(i,j)\in\{1,\cdots,N\}^2} \mathbb{E}[|\mathbf{X}_{ij}^N|^k] = 0 \qquad \forall k>2. \tag{31}$$

Exercise 1.5. *Take $X^{N,M}$ to be a $N\times M$ matrix with independent centered entries such that $N\mathbb{E}[|X_{i,j}|^2]=1$ and satisfying* (5). *Assume N/M goes to one. Show that*

$$\lim_{N\to\infty}\mathbb{E}[\frac{1}{N}\mathrm{Tr}((X^{N,M}(X^{N,M})^*)^k)] = \int x^{2k}d\sigma(x)$$

is the Catalan number.

Exercise 1.6. *Take $X^{N,M}$ to be a $N\times M$ matrix with independent centered entries such that $N\mathbb{E}[|X_{i,j}|^2]=1$ and satisfying* (5). *Assume N/M goes to c. Show that*

$$\lim_{N\to\infty}\mathbb{E}[\frac{1}{N}\mathrm{Tr}((X^{N,M}(X^{N,M})^*)^k)] = \sum c^\ell m(\ell,k)$$

where $m(\ell,k)$ is an integer number counting the number of non-crossing pair partitions of $2k$ points such that if we color in a bipartite way the faces of the non-crossing partition in such a way that the first face is black, then it has ℓ black faces.

1.2 CLT for Wigner matrices

In the previous section, we proved Wigner's theorem by evaluating $\int x^p dL_{\mathbf{X}^N}(x)$ for $p\in\mathbb{N}$ (in the sequel, $dL_{\mathbf{X}^N}$ will denote the empirical measure associated with the matrix model $\mathbf{X}^N$). We shall push this computation one step further here and prove a central limit theorem. Namely, setting

$$\int x^k d\bar{L}_{\mathbf{X}^N}(x) := \mathbb{E}\left[\int x^k dL_{\mathbf{X}^N}(x)\right], \tag{32}$$

we shall prove that

$$M_k^N := N\left(\int x^k dL_{\mathbf{X}^N}(x) - \int x^k d\bar{L}_{\mathbf{X}^N}(x)\right) = \sum_{i=1}^N \left(\lambda_i^k - \mathbb{E}[\lambda_i^k]\right) \tag{33}$$

converges in law to a centered Gaussian variable. We will be rather sketchy here, we refer to [2] for a complete and clear treatment and [1] for a simplified exposition of the full proof of the theorem we state below. To simplify, we assume here that $\mathbf{X}^N$ is a Wigner matrix with

$$X_{ij}^N = \frac{B_{ij}}{\sqrt{N}}, \tag{34}$$

where $(B_{ij}, 1 \le i \le j \le N)$ are independent real equidistributed random variables. Moreover we assume here that their marginal distribution μ has all moments finite (in particular (5) is satisfied) and satisfies

$$\int x d\mu(x) = 0 \text{ and } \int x^2 d\mu(x) = 1. \tag{35}$$

We shall show why the following statement holds.

Theorem 1.7. *Let*

$$\begin{aligned}\sigma_k^2 = k^2 \left[C_{\frac{k-1}{2}}\right]^2 + \frac{k^2}{2}\left[C_{\frac{k}{2}}\right]^2 \left[\int x^4 d\mu(x) - 1\right] \\ + \sum_{r=3}^{\infty} \frac{2k^2}{r} \left(\sum_{\substack{k_i \ge 0 \\ 2\sum_{i=1}^r k_i = k-r}} \prod_{i=1}^{r} C_{k_i} \right)^2\end{aligned} \tag{36}$$

In this formula, C_x equals zero if x is not an integer and otherwise is equal to the Catalan number.

Then, M_k^N converges in moments to the centered Gaussian variable with variance σ_k^2, i.e., for all $l \in \mathbb{N}$,

$$\lim_{N\to\infty} \mathbb{E}\left[(M_k^N)^l\right] = \frac{1}{\sqrt{2\pi}\sigma_k} \int x^l e^{-\frac{x^2}{2\sigma_k^2}} dx. \tag{37}$$

Remark 1.8. *Unlike the standard central limit theorem for independent variables, the variance here depends on $\int x^4 d\mu(x)$.*

Outline of the proof.

- *We first prove that the statement is true when $l = 2$.* (It is clearly true for $l = 1$ since X_k^N is centered.) We thus want to show

$$\sigma_k^2 = \lim_{N\to\infty} \mathbb{E}\left[(M_k^N)^2\right]. \tag{38}$$

Below (22), we proved that $\mathbb{E}\left[(X_k^N)^2\right]$ is bounded, uniformly in N. Furthermore, we can write

$$\mathbb{E}\left[(M_k^N)^2\right] = \frac{1}{N^k} \sum_{\mathbf{i},\mathbf{i}'} [P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')] \tag{39}$$

where the sum over $\mathbf{i},\mathbf{i}'$ will hold on graphs $\tilde{G}(\mathbf{i},\mathbf{i}') = (\tilde{V}(\mathbf{i},\mathbf{i}'), \tilde{E}(\mathbf{i},\mathbf{i}'))$ so that

$$|\tilde{V}(\mathbf{i},\mathbf{i}')| \le k, \quad |\tilde{E}(\mathbf{i},\mathbf{i}')| \le k. \tag{40}$$

Since $[P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]$ is uniformly bounded, the only contributing graphs to the leading order will be those for which $|\tilde{V}(\mathbf{i},\mathbf{i}')| = k$. Then, since we always have $|\tilde{V}(\mathbf{i},\mathbf{i}')| \le |\tilde{E}(\mathbf{i},\mathbf{i}')| + 1$, we have two cases:

- $|\tilde{E}(\mathbf{i},\mathbf{i}')| = k - 1$ in which case the skeleton $\tilde{G}(\mathbf{i},\mathbf{i}')$ will again be a tree but with one edge less than the total number possible; this means that one edge appears with multiplicity four and belongs to $\tilde{E}(\mathbf{i}) \cap \tilde{E}(\mathbf{i}')$, the other edges appearing with multiplicity 2. Hence, the graphs of $\tilde{E}(\mathbf{i})$ and $\tilde{E}(\mathbf{i}')$ are both trees (which implies that k is even); there are $C^2_{\frac{k}{2}}$ such trees, and they are glued by a common edge, to choose among $\frac{k}{2}$ edges in each of the tree. Finally, there are two possible choices to glue the two trees according to the orientation. Thus, there are

$$2\left(\frac{k}{2}\right)^2 C^2_{\frac{k}{2}} = \left(\frac{k^2}{2}\right) C^2_{\frac{k}{2}} \tag{41}$$

such graphs and then

$$P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}') = \int x^4 d\mu(x) - 1. \tag{42}$$

We hence obtain the contribution $(\frac{k^2}{2})C^2_{\frac{k}{2}}(\int x^4 d\mu(x) - 1)$ to the variance.

• $|\tilde{E}(\mathbf{i},\mathbf{i}')| = k$. In this case, the graph is not a tree anymore and because $|\tilde{E}(\mathbf{i},\mathbf{i}')| - |\tilde{V}(\mathbf{i},\mathbf{i}')| = 1$, it contains exactly one cycle. This can be seen either by closer inspection of the arguments given after (11) or by using the formula which relates the genus of a graph and its number of vertices, faces and edges;

$$\sharp\text{vertices} + \sharp\text{faces} - \sharp\text{edges} = 2 - 2g \leq 2 \tag{43}$$

The faces are defined by following the boundary of the graph; each of these boundaries are exactly one cycle of the graph except one (since a graph has always one boundary) and therefore

$$\sharp\text{faces} = 1 + \sharp\text{cycles}. \tag{44}$$

So we get, for a connected graph with skeleton $(\tilde{V}, \tilde{E})$,

$$|\tilde{V}| \leq |\tilde{E}| + 1 - \sharp\text{cycles} \tag{45}$$

In our case, $\sharp$vertices $= \sharp$edges $= k$ and $\sharp$cycles ≥ 1 (since the graph is not a tree), which implies that $\sharp$cycles $= 1$. This implies also that both graphs have to share this cycle and have the same edges. The cycle can have length one, corresponding to a loop at a vertex: We then create the rest of the graph by constructing a tree with $k - 1/2$ edges, rooted next to the loop. We have $(C_{k-1/2})^2$ such possibilities. We then must choose two roots on each of these graphs to be the starting edge of the two explorating paths: we have k^2 possibilities. The case where the cycle has length two amounts to the previous case that we already considered (with an edge with multiplicity 4). For cycle made of $r \geq 3$ edges, the counting is similar: we have possibly trees at each vertices of the cycle, k^2 choices of the roots and 2 possible choice to glue the cycles according to orientation or not. We then have to divide by r because of the symmetry around the cycle. This yields

$$2k^2 \left(\sum_{2\sum k_i = k-r} \prod C_{k_i} \right)^2 / r.$$

Adding the number of such graphs completes the proof of the convergence of $\mathbb{E}\left[(M_k^N)^2\right]$ to σ_k^2 (see [2] for more details).

- *Convergence to the Gaussian law.*

  We next show that M_k^N is asymptotically Gaussian. This amounts to prove that $\lim_{N\to\infty}\mathbb{E}[(M_k^N)^{2l+1}] = 0$ whereas

$$\lim_{N\to\infty}\mathbb{E}[(M_k^N)^{2l}] = \sharp\{\text{number of pair partitions of } 2l \text{ elements}\} \times \sigma_k^{2l}. \tag{46}$$

Again, we shall expand the expectation in terms of graphs and write for $l \in \mathbb{N}$,

$$\mathbb{E}[(M_k^N)^l] = \frac{1}{N^{\frac{kl}{2}}} \sum_{\mathbf{i}_1,\cdots,\mathbf{i}_l} P(\mathbf{i}^1,\cdots,\mathbf{i}^l) \tag{47}$$

with $P(\mathbf{i}^1,\cdots,\mathbf{i}^l)$ given by

$$\begin{aligned}\mathbb{E}\Big[\Big(B_{i_1^1 i_2^1}\cdots B_{i_k^1 i_1^1} &- \mathbb{E}[B_{i_1^1 i_2^1}\cdots B_{i_k^1 i_1^1}]\Big)\cdot \\ &\cdots\Big(B_{i_1^l i_2^l}\cdots B_{i_k^l i_1^l} - \mathbb{E}[B_{i_1^l i_2^l}\cdots B_{i_k^l i_1^l}]\Big)\Big]\end{aligned} \tag{48}$$

We denote by $G(\mathbf{i}^1,\cdots,\mathbf{i}^l) = (V(\mathbf{i}^1,\cdots,\mathbf{i}^l), E(\mathbf{i}^1,\cdots,\mathbf{i}^l))$ the corresponding graph; $V(\mathbf{i}^1,\cdots,\mathbf{i}^l) = \{i_n^j, 1\le j\le l, 1\le n\le k\}$ and $E(\mathbf{i}^1,\cdots,\mathbf{i}^l) = \{(i_n^j, i_{n+1}^j), 1\le j\le l, 1\le n\le k\}$ with the convention $i_{k+1}^j = i_1^j$. As before, $P(\mathbf{i}^1,\cdots,\mathbf{i}^l)$ equals zero unless each edge appears with multiplicity 2 at least. Also, because of the centering, it vanishes if there exists a $j\in\{1,\cdots,l\}$ so that $E(\mathbf{i}^j)$ does not intersect $E(\mathbf{i}^1,\cdots,\mathbf{i}^{j-1},\mathbf{i}^{j+1},\cdots,\mathbf{i}^l)$. Consequently, we must have $c \le \lfloor l/2 \rfloor$. Let us decompose $G(\mathbf{i}^1,\cdots,\mathbf{i}^l)$ into its connected components $(G_1,\cdots,G_c)$. We claim that

$$|V(\mathbf{i}^1,\cdots,\mathbf{i}^l)| \le \frac{(k-1)l}{2} + c. \tag{49}$$

To prove this bound let $(V(\mathbf{i}^1,\cdots,\mathbf{i}^l), E')$ be a spanning forest of $G(\mathbf{i}^1,\cdots,\mathbf{i}^l)$. We choose E' so that each edge has multiplicity two (E' can be obtained by removing some edges from the original graph). We claim that we can choose E' so that for each j

$$|E(\mathbf{i}^j)\cap E'| \le k-1\,.$$

Indeed, if we have a spanning forest and a j such that $|E(\mathbf{i}^j)\cap E'| = k$ this means that $G(\mathbf{i}^j)$ is a tree and each edge has multiplicity two. But we also know that there exists k such that $E(\mathbf{i}^j)\cap E(\mathbf{i}^k)| \neq \emptyset$ and let e be in this intersection. We let $\tilde{E}'$ be obtained by removing one edge e from $E(\mathbf{i}^j)\cap E'$ but keeping all other edges from E'. Clearly, $\tilde{E}'$ is still a spanning forest with the same properties and we can continue like that for all j, hence obtaining the desired property.

This allows to conclude as then $|E'| \le (k-1)l$ and so the number of edges in the spanning forest $(V(\mathbf{i}^1,\cdots,\mathbf{i}^l), E'')$ obtained by keeping only simple edges in E' is such that $|E''| \le (k-1)l/2$. Finally, we have

$$|V(\mathbf{i}^1,\cdots,\mathbf{i}^l)| \le |E''| + c \le \frac{(k-1)l}{2} + c\,.$$

We deduce that indices corresponding to graphs with c connected compnents have weight $N^{\frac{(k-1)l}{2}+c-\frac{kl}{2}}$. Thus, to get a first order contribution we must have l even and $c=\frac{l}{2}$. In that case, we write the pairing $(s_j,r_j)_{1\le j\le c}$ so that $(G(\mathbf{i}_{s_j}),G(\mathbf{i}_{r_j}))_{1\le j\le c}$ are connected for all $1\le j\le c$ (with the convention $s_j<r_j$). By independence of the entries, we have

$$P(\mathbf{i}_1,\cdots,\mathbf{i}_l)=\prod_{j=1}^{c}P(\mathbf{i}_{s_j},\mathbf{i}_{r_j}) \tag{50}$$

and so we have proved that

$$\begin{aligned} N^{-kl}\sum_{\mathbf{i}_1,\cdots,\mathbf{i}_l}P(\mathbf{i}_1,\cdots,\mathbf{i}_l) &= \sum_{\substack{s_1<\cdots<s_c\\ r_1<\cdots r_c\\ r_j>s_j}}\left(N^{-k}\sum_{\mathbf{i}_1,\mathbf{i}_2}P(\mathbf{i}_1,\mathbf{i}_2)\right)^c+o(1)\\ &=\sigma_k^{2c}\sum_{\substack{s_1<\cdots<s_c\\ r_1<\cdots r_c\\ r_j>s_j}}1+o(1) \end{aligned} \tag{51}$$

which proves the claim since

$$\frac{1}{\sqrt{2\pi}}\int x^{2c}e^{-\frac{x^2}{2}}dx=\sum_{\substack{s_1<\cdots<s_c\\ r_1<\cdots r_c\\ r_j>s_j}}1=(2c-1)(2c-3)(2c-5)\cdots 1. \tag{52}$$

This completes the proof of the moments convergence. □

2 Method based on Dyson-Schwinger equations

In this section, we show how to derive topological expansions from Dyson-Schwinger equations for the simplest model : the GUE. The Gaussian Unitary Ensemble is the sequence of $N\times N$ hermitian matrices $X_N, N\ge 0$ such that $(X_N(ij))_{i\le j}$ are independent centered Gaussian variables with variance $1/N$ that are complex outside of the diagonal (with independent real and imaginary parts). Then, we shall discuss the following expansion, true for all integer k

$$\mathbb{E}[\frac{1}{N}\mathrm{Tr}(X_N^k)]=\sum_{g\ge 0}\frac{1}{N^{2g}}M_g(k)\,.$$

This expansion is called a topological expansion because $M_g(k)$ is the number of maps of genus g which can be build by matching the edges of a vertex with k labelled half-edges. We remind here that a map is a connected graph properly embedded into a surface (i.e so that edges do not cross). Its genus is the smallest genus of a surface so that this can be done. This identity is well known [14] and was the basis of several breakthroughs in enumerative geometry [9, 11]. It can be proven by expanding the trace into products of Gaussian entries and using Wick calculus to compute these moments. In this section, we show how to derive it by using Dyson-Schwinger equations.

2.1 Combinatorics versus analysis

In order to calculate the electromagnetic momentum of an electron, Feynman used diagrams and Schwinger used Green's functions. Dyson unified these two approaches thanks to Dyson-Schwinger equations. On one hand they can be thought as equations for the generating functions of the graphs that are enumerated, on the other they can be seen as equations for the invariance of the underlying measure. A baby version of this idea is the combinatorial versus the analytical characterization of the Gaussian law $\mathcal{N}(0,1)$. Let X be a random variable with law $\mathcal{N}(0,1)$. On one hand it is the unique law with moments given by the number of matchings :

$$\mathbb{E}[X^n] = \#\{\text{pair partitions of } n \text{ points}\} =: P_n \,. \tag{53}$$

On the other hand, it is also defined uniquely by the integration by parts formula

$$\mathbb{E}[Xf(X)] \;=\; \mathbb{E}\left[f'(X)\right] \tag{54}$$

$$\tag{55}$$

for all smooth functions f going to infinity at most polynomially. If one applies the latter to $f(x) = x^n$ one gets

$$m_{n+1} := \mathbb{E}\left[X^{n+1}\right] \;=\; \mathbb{E}\left[nX^{n-1}\right] = nm_{n-1}\,.$$

This last equality is the induction relation for the number P_{n+1} of pair partitions of $n+1$ points by thinking of the n ways to pair the first point. Since $P_0 = m_0 = 1$ and $P_1 = m_1 = 0$, we conclude that $P_n = m_n$ for all n. Hence, the integration by parts formula and the combinatorial interpretation of moments are equivalent.

2.2 GUE : combinatorics versus analysis

When instead of considering a Gaussian variable we consider a matrix with Gaussian entries, namely the GUE, it turns out that moments are as well described both by integration by parts equations and combinatorics. In fact moments of GUE matrices can be seen as generating functions for the enumeration of interesting graphs, namely maps, which are sorted by their genus. We shall describe the full expansion, the so-called topological expansion, at the end of this section. In this section, we discuss the large dimension expansion of moments of the GUE up to order $1/N^2$ as well as central limit theorems for these moments, and characterize these asymptotics both in terms of equations similar to the previous integration by parts, and by the enumeration of combinatorial objects.

Let us be more precise. A matrix $X = (X_{ij})_{1\le i,j\le N}$ from the GUE is the random $N\times N$ Hermitian matrix so that for $k<j$, $X_{kj} = X^{\mathbb{R}}_{kj} + iX^{i\mathbb{R}}_{kj}$, with two independent real centered Gaussian variables with covariance $1/2N$ (denoted later $\mathcal{N}\left(0,\frac{1}{2N}\right)$) variables $X^{\mathbb{R}}_{kj}, X^{i\mathbb{R}}_{kj}$) and for $k\in\{1,\dots,N\}$, $X_{kk}\sim\mathcal{N}\left(0,\frac{1}{N}\right)$. then, we shall prove that

$$\mathbb{E}[\frac{1}{N}\mathrm{tr}(X^k)] = M_0(k) + \frac{1}{N^2}M_1(k) + o(\frac{1}{N^2}) \tag{56}$$

where

- $M_0(k) = C_{k/2}$ denotes the Catalan number : it vanishes if k is odd and is the number of non-crossing pair partitions of $2k$ (ordered) points, that is pair partitions so that any two blocks (a,b) and (c,d) is such that $a < b < c < d$ or $a < c < d < b$. C_k can also be seen to be the number of rooted trees embedded into the plane and k edges, that is trees with a distinguished edge and equipped with an exploration path of the vertices $v_1 \to v_2 \to \cdots \to v_{2k}$ of length $2k$ so that (v_1, v_2) is the root and each edge is visited twice (once in each direction). C_k can also be seen as the number of planar maps build over one vertex with valence k : namely take a vertex with valence k, draw it on the plane as a point with k half-edges. Choose a root, that is one of these half-edges. Then the set of half-edges is in bijection with k ordered points (as we drew them on the plane which is oriented). A matching of the half-edges is equivalent to a pairing of these points. Hence, we have a bijection between the graphs build over one vertex of valence k by matching the end-points of the half-edges and the pair partitions of k ordered points. The pairing is non-crossing iff the matching gives a planar graph, that is a graph that is properly embedded into the plane (recall that an embedding of a graph in a surface is proper iff the edges of the graph do not cross on the surface). Hence, $M_0(k)$ can also be interpreted as the number of planar graphs build over a rooted vertex with valence k. Recall that the genus g of a graph (that is the minimal genus of a surface in which it can be properly embedded) is given by Euler formula :

  $$2 - 2g = \#Vertices + \#Faces - \#Edges\,,$$

  where the faces are defined as the pieces of the surface in which the graph is embedded which are separated by the edges of the graph. If the surface as minimal genus, these faces are homeomorphic to discs.

- $M_1(k)$ is the number of graphs of genus one build over a rooted vertex with valence k. Equivalently, it is the number of rooted trees with $k/2$ edges and exactly one cycle.

Moreover, we shall prove that for any $k_1, \ldots, k_p$ $(\mathrm{tr}(X^{k_j}) - \mathbb{E}[\mathrm{tr}(X^{k_j})])_{1 \le j \le p}$ converges in moments towards a centered Gaussian vector with covariance

$$M_0(k,\ell) = \lim_{N\to\infty} \mathbb{E}\left[(\mathrm{tr}(X^k) - \mathbb{E}[\mathrm{tr}(X^k)])(\mathrm{tr}(X^\ell) - \mathbb{E}[\mathrm{tr}(X^\ell)])\right].$$

$M_0(k,\ell)$ is the number of connected planar rooted graphs build over a vertex with valence k and one with valence ℓ. Here, both vertices have labelled half-edges and two graphs are counted as equal only if they correspond to matching half-edges with the same labels (and this despite of symmetries). Equivalently $M_0(k,\ell)$ is the number of rooted trees with $(k+\ell)/2$ edges and an exploration path with $k+\ell$ steps such that k consecutive steps are colored and at least an edge is explored both by a colored and a non-colored step of the exploration path.

Recall here that convergence in moments means that all mixed moments converge to the same mixed moments of the Gaussian vector with covariance M. We shall use that the moments of a centered Gaussian vector are given by Wick formula :

$$m(k_1,\ldots,k_p) = \mathbb{E}[\prod_{i=1}^{p} X_{k_i}] = \sum_{\pi} \prod_{\text{blocks } (a,b) \text{ of } \pi} M(k_a,k_b)$$

which is in fact equivalent to the induction formula we will rely on :

$$m(k_1,\dots,k_p)=\sum_{i=2}^{p}M(k_1,k_i)m(k_2,\dots,k_{i-1},k_{i+1},\dots,k_p)\,.$$

Convergence in moments towards a Gaussian vector implies of course the standard weak convergence as convergence in moments implies that the second moments of $Z_N:=(\mathrm{tr}(X^{k_j})-\mathbb{E}[\mathrm{tr}(X^{k_j})])_{1\le j\le p}$ are uniformly bounded, hence the law of Z_N is tight. Moreover, any limit point has the same moments than the Gaussian vector. Since these moments do not blow too fast, there is a unique such limit point, and hence the law of Z_N converges towards the law of the Gaussian vector with covariance M. We will discuss at the end of this section how to generalize the central limit theorem to differentiable test functions, that is show that $Z_N(f)=\mathrm{tr}f(X)-\mathbb{E}[\mathrm{tr}f(X)]$ converges towards a centered Gaussian variable for any bounded differentiable function. This requires more subtle uniform estimates on the covariance of $Z_N(f)$ for which we will use Poincaré's inequality.

The asymptotic expansion (56) as well as the central limit theorem can be derived using combinatorial arguments and Wick calculus to compute Gaussian moments. This can also be obtained from the Dyson-Schwinger (DS) equation, which we do below.

2.2.1 Dyson-Schwinger Equations

Let :

$$Y_k:=\mathrm{tr}X^k-\mathbb{E}\mathrm{tr}X^k$$

We wish to compute for all integer numbers $k_1,\dots,k_p$ the correlators :

$$\mathbb{E}\left[\mathrm{tr}X^{k_1}\prod_{i=2}^{p}Y_{k_i}\right].$$

By integration by parts, one gets the following Dyson-Schwinger equations

Lemma 2.1. *For any integer numbers $k_1,\dots,k_p$, we have*

$$\mathbb{E}\left[trX^{k_1}\prod_{i=2}^{p}Y_{k_i}\right]=\mathbb{E}\left[\frac{1}{N}\sum_{\ell=0}^{k_1-2}trX^{\ell}trX^{k_1-2-p}\prod_{i=2}^{p}Y_{k_i}\right]$$

$$+\mathbb{E}\left[\sum_{i=2}^{p}\frac{k_i}{N}trX^{k_1+k_i-2}\prod_{j=2,j\neq i}^{p}Y_{k_j}\right]\tag{57}$$

Proof. Indeed, we have

$$\begin{aligned}\mathbb{E}\left[\mathrm{tr}X^{k_1}\prod_{i=2}^{p}Y_{k_i}\right]&=\sum_{i,j=1}^{N}\mathbb{E}\left[X_{ij}(X^{k_1-1})_{ji}\prod_{i=2}^{p}Y_{k_i}\right]\\&=\frac{1}{N}\sum_{i,j=1}^{N}\mathbb{E}\left[\partial_{X_{ji}}\left((X^{k_1-1})_{ji}\prod_{i=2}^{p}Y_{k_i}\right)\right]\end{aligned}$$

where we noticed that since the entries are Gaussian independent complex variables, for any smooth test function f,

$$\mathbb{E}[X_{ij}f(X_{k\ell},k\le\ell)] = \frac{1}{N}\mathbb{E}[\partial_{X_{ji}}f(X_{k\ell},k\le\ell)]\,. \tag{58}$$

But, for any $i,j,k,\ell\in\{1,\dots,N\}$ and $r\in\mathbb{N}$

$$\partial_{X_{ji}}(X^r)_{k\ell} = \sum_{s=0}^{r-1}(X^s)_{kj}(X^{r-s-1})_{i\ell}$$

where $(X^0)_{ij} = 1_{i=j}$. As a consequence

$$\partial_{X_{ji}}(Y_r) = rX^{r-1}_{ij}\,.$$

The Dyson-Schwinger equations follow readily. □

Exercise 2.2. *Show that*

1. *If X is a GUE matrix,* (58) *holds. Deduce* (2.1).
2. *take X to be a GOE matrix, that is a symmetric matrix with real independent Gaussian entries $N_{\mathbb{R}}(0,\frac{1}{N})$ above the diagonal, and $N_{\mathbb{R}}(0,\frac{2}{N})$ on the diagonal. Show that*

$$\mathbb{E}[X_{ij}f(X_{k\ell},k\le\ell)] = \frac{1}{N}\mathbb{E}[\partial_{X_{ji}}f(X_{k\ell},k\le\ell)] + \frac{1}{N}\mathbb{E}[\partial_{X_{ij}}f(X_{k\ell},k\le\ell)]\,.$$

   Deduce that a formula analogous to (2.1) *holds provided we have an additional term $N^{-1}\mathbb{E}\left[k_1 trX^{k_1}\prod_{i=2}^{p}Y_{k_i}\right]$.*

2.2.2 Dyson-Schwinger equation implies genus expansion

We will show that the DS equation (2.1) can be used to show that :

$$\mathbb{E}\left[\frac{1}{N}\mathrm{tr}X^k\right] = M_0(k) + \frac{1}{N^2}M_1(k) + o(\frac{1}{N^2})$$

Next orders can be derived similarly. Let :

$$m_k^N := \mathbb{E}\left[\frac{1}{N}\mathrm{tr}X^k\right]$$

By the DS equation (with no Y terms), we have that :

$$m_k^N = \mathbb{E}\left[\sum_{\ell=0}^{k-2}\frac{1}{N}\mathrm{tr}X^\ell\frac{1}{N}\mathrm{tr}X^{k-\ell-2}\right]\,. \tag{59}$$

We now assume that we have the self-averaging property that for all $\ell\in\mathbb{N}$:

$$\mathbb{E}\left[\left(\frac{1}{N}\mathrm{tr}X^\ell - \mathbb{E}\left[\frac{1}{N}\mathrm{tr}X^\ell\right]\right)^2\right] = o(1)$$

as $N \to \infty$ as well as the boundedness property $\sup_N m_\ell^N < \infty$. We will show both properties are true in Lemma 2.3. If this is true, then the above expansion (59) gives us :

$$m_k^N = \sum_{\ell=0}^{k-2} m_\ell^N m_{k-\ell-2}^N + o(1)$$

As $\{m_\ell^N, \ell \leq k\}$ are uniformly bounded, they are tight and so any limit point $\{m_\ell, \ell \leq k\}$ satisfies

$$m_k = \sum_{\ell=0}^{k-2} m_\ell m_{k-\ell-2}, m_0 = 1, m_1 = 0.$$

This equation has clearly a unique solution.

On the other hand, let $M_0(k)$ be the number of maps of genus 0 with one vertex with valence k. These satisfy the Catalan recurrence :

$$M_0(k) = \sum_{\ell=0}^{k-2} M_0(\ell) M_0(k-\ell-2)$$

This recurrence is shown by a Catalan-like recursion argument, which goes by considering the matching of the first half edge with the ℓth half-edge, dividing each map of genus 0 into two sub-maps (both still of genus 0) of size ℓ and $k-\ell-2$, for $\ell \in \{0, \dots, k-2\}$.

Since m and M_0 both satisfy the same recurrence (and $M_0(0) = m_0^N = 1, M_0(1) = m_1^N = 0$), we deduce that $m = M_0$ and therefore we proved by induction (assuming the self-averaging works) that :

$$m_k^N = M_0(k) + o(1) \text{ as } N \to \infty$$

It remains to prove the self-averaging and boundedness properties.

Lemma 2.3. *There exists finite constants D_k and E_k, $k \in \mathbb{N}$, independent of N, so that for integer number ℓ, every integer numbers $k_1, \dots, k_\ell$ then :*

$$\text{a)} \qquad c^N(k_1, \dots, k_p) := \mathbb{E}\left[\prod_{i=1}^{\ell} Y_{k_i}\right] \textit{ satisfies } |c^N(k_1, \dots, k_p)| \leq D_{\sum k_i}$$

and

$$\text{b)} \qquad m_{k_1}^N := \mathbb{E}\left[\frac{1}{N} trX^{k_1}\right] \textit{ satisfies } |m_{k_1}^N| \leq E_{k_1}.$$

Proof. The proof is by induction on $k = \sum k_i$. It is clearly true for $k = 0, 1$ where $E_0 = 1, E_1 = 0$ and $D_k = 0$. Suppose the induction hypothesis holds for $k-1$. To see that *b)* holds, by the DS equation, we first observe that :

$$\begin{aligned} \mathbb{E}\left[\frac{1}{N}\mathrm{tr}X^k\right] &= \mathbb{E}\left[\sum_{\ell=0}^{k-2} \frac{1}{N}\mathrm{tr}X^\ell \frac{1}{N}\mathrm{tr}X^{k-\ell-2}\right] \\ &= \sum_{\ell=0}^{k-2} (m_\ell^N m_{k-\ell-2}^N + \frac{1}{N^2} c^N(\ell, k-\ell-2)) \end{aligned}$$

Hence, by the induction hypothesis we deduce that

$$\left|\mathbb{E}\left[\frac{1}{N}\mathrm{tr}X^k\right]\right| \le \sum_{\ell=0}^{k-2}(E_\ell E_{k-2-\ell}+D_{k-2}) := E_k\,.$$

To see that $a)$ holds, we use the DS equation as follows

$$\begin{aligned}
\mathbb{E}\left[Y_{k_1}\prod_{j=2}^{p}Y_{k_j}\right] &= \mathbb{E}\left[\mathrm{tr}X_{k_1}\prod_{j=2}^{p}Y_{k_j}\right]-\mathbb{E}\left[\mathrm{tr}X_{k_1}\right]\mathbb{E}\left[\prod_{j=2}^{p}Y_{k_j}\right]\\
&= \frac{1}{N}\mathbb{E}\left[\sum_{\ell=0}^{k-2}\mathrm{tr}X^{\ell}\mathrm{tr}X^{k_1-\ell-2}\prod_{j=2}^{p}Y_{k_j}\right]\\
&\quad +\mathbb{E}\left[\sum_{i=2}^{p}\frac{k_i}{N}\mathrm{tr}X^{k_1+k_i-2}\prod_{j=2,j\neq i}^{p}Y_{k_j}\right]\\
&\quad -\mathbb{E}\left[\frac{1}{N}\sum_{\ell=0}^{k-2}\mathrm{tr}X^{\ell}\mathrm{tr}X^{k_1-\ell-2}\right]\mathbb{E}\left[\prod_{j=2}^{p}Y_{k_j}\right]\,.
\end{aligned}$$

We next substract the last term to the first and observe that

$$\begin{aligned}
&\mathrm{tr}X^{\ell}\mathrm{tr}X^{k_1-\ell-2}-\mathbb{E}[\mathrm{tr}X^{\ell}\mathrm{tr}X^{k_1-\ell-2}]\\
&\qquad = NY_\ell m^N_{k_1-2-\ell}+NY_{k_1-2-\ell}m^N_\ell+Y_\ell Y_{k_1-2-\ell}-c^N(\ell,k_1-2-\ell)
\end{aligned}$$

to deduce

$$\begin{aligned}
\mathbb{E}\left[Y_{k_1}\prod_{j=2}^{p}Y_{k_j}\right] &= 2\sum_{\ell=0}^{k_1-2}m^N_\ell c^N(k_1-2-\ell,k_2,\dots,k_p)\\
&+\sum_{i=2}^{p}k_i m^N_{k_1+k_i-2}c^N(k_2,..,k_{i-1},k_{i+1},.,k_p)\\
&-\frac{1}{N}\sum_{\ell=0}^{k_1-2}[c^N(\ell,k_1-2-\ell)c^N(k_2,\dots,k_p)-c^N(\ell,k_1-2-\ell,k_2,\dots,k_p)]\\
&+\frac{1}{N}\sum_{i=2}^{p}k_i c^N(k_1+k_i-2,k_2,..,k_{i-1},k_{i+1},.,k_p) \qquad (60)
\end{aligned}$$

which is bounded uniformly by our induction hypothesis. □

As a consequence, we deduce

Corollary 2.4. *For all $k\in\mathbb{N}$, $\frac{1}{N}\mathrm{Tr}(X^k)$ converges almost surely towards $M_0(k)$.*

Proof. Indeed by Borel Cantelli Lemma it is enough to notice that it follows from the summability of

$$P\Big(|\mathrm{Tr}(X^k-\mathbb{E}\Big(\mathrm{Tr}(X^k)\Big)|\ge N\varepsilon\Big)\le\frac{c^N(k,k)}{\varepsilon^2N^2}\le\frac{D_{2k}}{\varepsilon^2N^2}.$$

□

2.3 Central limit theorem

The above self averaging properties prove that $m_k^N = M_0(k) + o(1)$. To get the next order correction we analyze the **limiting covariance** $c^N(k,\ell)$. We will show that

Lemma 2.5. *For all $k,\ell \in \mathbb{N}$, $c^N(k,\ell)$ converges as N goes to infinity towards the unique solution $M_0(k,\ell)$ of the equation*

$$M_0(k,\ell) \;=\; 2\sum_{p=0}^{\ell-2} M_0(p)M_0(k-2-p,\ell) + \ell M_0(k+\ell-2)$$

so that $M_0(k,\ell) = 0$ if $k+\ell \le 1$.

As a consequence we will show that

Corollary 2.6. *$N^2(m_k^N - M_0(k)) = m_k^1 + o(1)$ where the numbers $(m_k^1)_{k\ge 0}$ are defined recursively by :*

$$m_k^1 = 2\sum_{\ell=0}^{k-2} m_\ell^1 M_0\,(k-\ell-2) + \sum_{\ell=0}^{k-2} M_0(\ell,k-\ell-2)$$

Proof. (Of Lemma 2.5) Observe that $c^N(k,\ell)$ converges for $K = k+\ell \le 1$ (as it vanishes uniformly). Assume you have proven convergence towards $M_0(k,\ell)$ up to K. Take $k_1 + k_2 = K+1$ and use (60) with $p=1$ to deduce that $c^N(k_1,k_2)$ satisfies

$$c^N(k_1,k_2) = 2\sum_{\ell=0}^{k_1-2} m_\ell^N c^N(k_1-\ell-2,k_2) + k_2 m_{k_1+k_2-2}^N + \frac{1}{N}\sum c^N(\ell,k_1-\ell-2,k_2)\,.$$

Lemma 2.3 implies that the last term is at most of order $1/N$ and hence we deduce by our induction hypothesis that $c(k_1,k_2)$ converges towards $M_0(k_1,k_2)$ which is given by the induction relation

$$M_0(k_1,k_2) \;=\; 2\sum_{\ell=0}^{k_1} M_0(\ell)M_0(k_1-2-\ell,k_2) + k_2 M_0(k_1+k_2-2)\,.$$

Moreover clearly $M_0(k_1,k_2) = 0$ if $k_1+k_2 \le 1$. There is a unique solution to this equation. □

Exercise 2.7. *Show by induction that*

$$M_0(k,\ell) = \#\{\textit{planar maps with 1 vertex of degree } \ell \textit{ and one vertex of degree } k\}$$

Proof. (of Corollary 2.6) Again we prove the result by induction over k. It is fine for $k = 0,1$ where $c_k^1 = 0$. By (60) with $p=0$ we have :

$$\begin{aligned} N^2(m_k^N - M_0(k)) \;&=\; 2\sum M_0(\ell)N^2\left(m_{k-\ell-2}^N - M_0(k-2-\ell)\right) \\ &\quad + \sum N^2\left(m_\ell^N - M_0(\ell)\right)(m_{k-\ell-2} - M_0(k-2-\ell)) \\ &\quad + \sum c^N(\ell,k-\ell-2) \end{aligned}$$

from which the result follows by taking the large N limit on the right hand side. □

Exercise 2.8. *Show that $c_k^1 = m_1(k)$ is the number of planar maps with genus 1 build on a vertex of valence k.(The proof goes again by showing that $m_1(k)$ satisfies the same type of recurrence relations as c_k^1 by considering the matching of the root : either it cuts the map of genus 1 into a map of genus 1 and a map of genus 0, or there remains a (connected) planar maps.)*

Theorem 2.9. *For any polynomial function $P = \sum \lambda_k x^k$, $Z_N(P) = trP - \mathbb{E}[trP]$ converges in moments towards a centered Gaussian variable $Z(P)$ with covariance given by*

$$\mathbb{E}[Z(P)\overline{Z}(P)] = \sum \lambda_k \overline{\lambda}_{k'} M_0(k,k') .$$

Proof. It is enough to prove the convergence of the moments of the Y_k's. Let

$$c^N(k_1,\dots,k_p) = \mathbb{E}\left[Y_{k_1}\cdots Y_{k_p}\right] .$$

Then $c^N(k_1,\dots,k_p)$ converges to $G(k_1,\dots k_p)$ given by :

$$G(k_1,\dots,k_p) = \sum_{i=2}^{k} M_0(k_1,k_i) G(k_2,\dots,\hat{k}_i,\dots,k_p) \tag{61}$$

where ˆ is the absentee hat.

This type of moment convergence is equivalent to a Wick formula and is enough to prove (by the moment method) that $Y_{k_1},\dots,Y_{k_p}$ are jointly Gaussian. Again, we will prove this by induction by using the DS equations. Now assume that (61) holds for any $k_1,\dots,k_p$ such that $\sum_{i=1}^{p} k_i \le k$. (induction hypothesis) We use (60). Notice by the a priori bound on correlators of Lemma 2.3(a) that the terms with a $1/N$ are negligible in the right hand side and m_k^N is close to $M_0(k)$, yielding

$$\begin{aligned}\mathbb{E}\left[Y_{k_1}\prod_{j=2}^{p} Y_{k_j}\right] &= 2\sum_{\ell=0}^{k_1-2} M_0(\ell) c^N(k_1-2-\ell,k_2,\dots,k_p) \\ &+ \sum_{i=2}^{p} k_i M_0(k_1+k_i-2) c^N(k_2,..,k_{i-1},k_{i+1},.,k_p) + O(\frac{1}{N})\end{aligned}$$

By using the induction hypothesis, this gives rise to :

$$\begin{aligned}\mathbb{E}\left[\prod_{i=1}^{p} Y_{k_i}\right] &= 2\sum M_0(\ell) G(k_1-\ell-2,k_2,\dots,k_p) \\ &\quad + \sum k_i M_0(k_i+k_j-2) G(k_2,\dots,\hat{k}_i,\dots k_p) + o(1)\end{aligned}$$

It follows that

$$G(k_1,\dots,k_p) = 2\sum M_0(\ell) G(k_1-\ell-2,k_2,\dots,k_p) + \sum k_i M_0(k_i+k_j-2) G(k_2,\dots,\hat{k}_i,\dots k_p) .$$

But using the induction hypothesis, we get

$$G(k_1,\dots,k_p) = \sum_{i=2}^{p} (2\sum M_0(\ell) M(k_1-\ell-2,k_i) + k_i M_0(k_i+k_j-2)) G(k_2,\dots,\hat{k}_i,\dots k_p)$$

which yields the claim since

$$M_0(k_1,k_i) = 2\sum M_0(\ell)M(k_1-\ell-2,k_i) + k_iM_0(k_1+k_i-2)\,.$$

□

2.4 GUE topological expansion

The "topological expansion" reads

$$\mathbb{E}\left[\frac{1}{N}\mathrm{tr}\left[X^k\right]\right] = \sum_{g\geq 0}\frac{1}{N^{2g}}M_g(k)$$

where $M_g(k)$ is the number of rooted maps of genus g build over a vertex of degree k. Here, a "map" is a connected graph properly embedded in a surface and a "root" is a distinguished oriented edge. A map is assigned a genus, given by the smallest genus of a surface in which it can be properly embedded. This complete expansion (not that the above series is in fact finite) can be derived as well either by Wick calculus or by Dyson-Schwinger equations : we leave it as an exercise to the reader. We will see later that cumulants of traces of moments of the GUE are related with the enumeration of maps with several vertices.

Exercise 2.10. *Show that Theorem 1.7 implies that M_k^N converges weakly to the centered Gaussian variable with variance σ_k^2. Hint: control tails to approximate bounded continuous functions by polynomials.*

Exercise 2.11. *Show that if $\mathbf{X}_1^N,\dots,\mathbf{X}_k^N$ are independent Wigner matrices*

$$tr(X_{i_1}\cdots X_{i_k}) - \mathbb{E}[tr(X_{i_1}\cdots X_{i_k})]$$

converges in law towards a centered Gaussian variable.

2.5 Generalization to Beta-ensembles

The approach by Dyson-Schwinger expansion can be generalized to invariant ensembles such as the so-called Beta ensembles and even to discrete Beta-ensembles given by the distribution of uniform tilings [5]. In this section we simply give heuristics to show how such a generalization can be made possible. The distribution of Beta-ensembles is the probability measure on $\mathbb{R}^N$ given by

$$dP_N^{\beta,V}(\lambda_1,\dots,\lambda_N) = \frac{1}{Z_N^{\beta,V}}\Delta(\lambda)^\beta e^{-N\beta\sum V(\lambda_i)}\prod_{i=1}^N \delta\lambda_i$$

where $\Delta(\lambda) = \prod_{i<j}|\lambda_i-\lambda_j|$.

Remark 2.12. *In the case $V(X) = \frac{1}{2}x^2$ and $\beta = 2$, $P_N^{2,x^2/4}$ is exactly the law of the eigenvalues for a matrix taken in the GUE as we were considering in the previous chapter (the case $\beta = 1$ corresponds to GOE and $\beta = 4$ to GSE). This is left as a (complicated) exercise, see e.g. [1].*

In this case we can also write the Dyson-Schwinger equations by a simple integration by parts. For instance we have for any bouded continuous function f

$$\mathbb{E}[\frac{1}{N}\sum V'(\lambda_i)f(\lambda_i)] = \mathbb{E}[\frac{1}{2N^2}\sum_{i,j=1}^{N}\frac{f(\lambda_i)-f(\lambda_j)}{\lambda_i-\lambda_j}] + (\frac{1}{\beta}-\frac{1}{2})\mathbb{E}[\frac{1}{N}\sum_{i=1}^{N} f'(\lambda_i)].$$

Again, this is an equation on the moments of the empirical measure

$$\hat{\mu}_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}$$

However, it is not closed in general, even if we look at the higher order Dyson-Schwinger equations. For instance even if we assume that we have self-averaging (namely the covariance goes to zero), the limit point of $\hat{\mu}_N$ (for the weak topology) will satisfy:

$$\int fV'd\mu = \frac{\beta}{2}\int\int\frac{f(x)-f(y)}{x-y}d\mu(x)d\mu(y) \tag{62}$$

for all C^1 function f. This equation has in general several solutions, for instance when V has two big wells like $A(x^2-1)^2$ for A big enough. However it has a unique solution for instance when V is convex, and in fact the previous results can be readily generalized to the case where V is convex, and a small perturbation of the quadratic potential. We restrict ourselves to $\beta = 2$ to simplify.

Theorem 2.13. *Take $\beta = 2$. Assume $V(x) = \frac{1}{2}x^2 + \sum_{i=1}^{p} t_i x^i$ is such that $V''(x) \geq c$ for all x.then there exists $\varepsilon(c) > 0$ such that if $max|t_i| \leq \varepsilon(c)$, for all k,*

$$\lim_{N\to\infty}\mathbb{E}[\int x^k d\hat{\mu}_N(x)] = \sum_{k_1,\ldots,k_p\geq 0}\prod(-t_i)^{k_i}k_i!M(k,k_1,\ldots,k_p)$$

where $M(k,k_1,\ldots,k_p)$ is the number of planar maps build over one vertex of valence k and k_i vertices of valence i, $1 \leq i \leq p$. Moreover, $(N(\int x^k d\hat{\mu}_N(x) - \mathbb{E}[\int x^k d\hat{\mu}_N(x)]))_{k\leq K}$ converge towards a centered Gaussian vector with covariance

$$C(k,k') = \sum_{k_1,\ldots,k_p\geq 0}\prod(-t_i)^{k_i}k_i!M(k,k',k_1,\ldots,k_p)$$

This type of results holds in a much greater generality, for instance in the multi-matrix case where only such perturbative results hold [7, 8]. The main point to prove the law of large number is to notice that

- Because V is strictly convex, by BrascampLieb inequalities the λ_i stay bounded by some $M(c) < \infty$ with exponentially large probability. By concentration inequalities the covariance is going to zero like $1/N^2$
- The probability measures $\mathbb{E}[\hat{\mu}^N]$ are therefore tight and there limit points satisfy (62). But taking $f(x) = x^k$ we can check that there exists a unique solution since if we had two μ and μ' then

$$\Delta_k = |\mu(x^k) - \mu'(x^k)| \leq 2\sum_{r\leq k-2}\Delta(r)M(c)^{\ell-2-r} + \sum|t_i|i\Delta(k+i-1)$$

Taking $\gamma < 1/M(c)$ we deduce that

$$\Delta(\gamma) = \sum \gamma^k \Delta(k) \le (\frac{2\gamma^2}{1-\gamma M(c)} + \sum |t_i| \gamma^{-i+1}) \Delta(\gamma)$$

The conclusion follows when the RHS is smaller than 1.

To prove the central limit theorem, the idea is quite similarly to see that eventhough the equations on the moments are not closed, they are approximately closed. We refer the reader to [8].

In fact, the law of large numbers and the CLT hold in much greater generality [10, 6]:

Theorem 2.14. • *Assume that* $\liminf_{|x|\to\infty} \frac{V(x)}{\ln(|x|)} > 1$ *(i.e. $V(x)$ goes to infinity fast enough to dominate the log term at infinity) and V is continuous. Then there exitsts a probability measure μ_V such that for all bounded continuous function f*

$$\lim_{N\to\infty} \mathbb{E}[\int f(x) d\hat{\mu}^N(x)] = \int f(x) d\mu_V(x) .$$

- *Assume V C^{35}. Assume that μ_V has a connected support and moreover that it vanishes at the boundary points like a square root (and is strictly positive otherwise). Then for f sufficiently smooth, $N(\int f(x)d\hat{\mu}^N(x) - \mathbb{E}[\int f(x)d\hat{\mu}^N(x)])$ converges towards a Gaussian variable.*

As pointed out by Peter Forrester, the law of large numbers is a consequence of the large deviation principle which roughly states that

$$\begin{aligned} \frac{dP_N^{\beta,V}}{d\lambda} &= \frac{1}{Z_N^{\beta,V}} \exp\left\{ \frac{1}{2}\beta \sum_{i\neq j} \ln|\lambda_i - \lambda_j| - \beta N \sum V(\lambda_i) \right\} \\ "=" & \frac{1}{Z_N^{\beta,V}} \exp\left\{ -\beta N^2 \mathbf{E}(\hat{\mu}_N) \right\} \end{aligned}$$

where $\mathbf{E}$ is the energy

$$\mathbf{E}(\mu) = \int \int [\frac{1}{2}V(x) + \frac{1}{2}V(y) - \frac{1}{2}\ln|x-y|] d\mu(x) d\mu(y)$$

One can check that there exists a unique minimizer to $\mathbf{E}$ from which the convergence follows. Let us stress the ideas to get the CLT. We first rewrite the Dyson-Schwinger equations by linearizing $\hat{\mu}_N$ around its limit as follows:

Lemma 2.15. *Let $f_i : \mathbb{R} \to \mathbb{R}$ be C_b^1 functions, $0 \le i \le K$. Let $M_N = N(\hat{\mu}_N - \mu_V)$. Then,*

$$\begin{aligned} \mathbb{E}[M_N(\Xi f_0) \prod_{i=1}^K N\hat{\mu}_N(f_i)] &= (\frac{1}{\beta} - \frac{1}{2}) \mathbb{E}[\hat{\mu}_N(f_0') \prod_{i=1}^K N\hat{\mu}_N(f_i)] \\ &+ \frac{1}{\beta} \sum_{\ell=1}^K \mathbb{E}[\hat{\mu}_N(f_0 f_\ell') \prod_{i\neq \ell} N\hat{\mu}_N(f_i)] \\ &+ \frac{1}{2N} \mathbb{E}[\int \frac{f_0(x) - f_0(y)}{x-y} dM_N(x) dM_N(y) \prod_{i=1}^p N\hat{\mu}_N(f_i)] \end{aligned}$$

where

$$\Xi f(x) = V'(x)f(x) - \int \frac{f(x)-f(y)}{x-y} d\mu_V(y)\,.$$

Ξ *will be called the master operator.*

The above equations are again a direct consequence of integration by parts. Then, to solve these equations, it is enough to show that moments of M_N are of order one and to invert Ξ. The later can be done exactly when μ_V has a connected support and vanishes like a square root (these are Tricomi airfol equations [12]). The former is a conseqeunce of concentration of measure and Dyson-Schwinger equations. Let us assume we could prove it. Then, we see that we can solve recursively Dyson-Schwinger equation. We first have

$$\mathbb{E}[(\hat{\mu}^N - \mu_V)(f_0)] = \frac{1}{N}(\frac{1}{\beta} - \frac{1}{2})\mathbb{E}[\hat{\mu}_N((\Xi^{-1} f_0)')] + O(\frac{1}{N^2})$$

we can then deduce the limit of the covariance from the first equation with $K = 1$

$$\lim_{N\to\infty} \mathbb{E}[(M_N(f_0)(N(\hat{\mu}_N(f_1) - \mathbb{E}[\hat{\mu}^N(f_1]))] = \mu_V(\Xi^{-1} f_0 f_1')\,.$$

We can continue like this to get the large N expansion of moments of $\hat{\mu}^N$, and in particular the CLT. We send the interested reader to my notes of a course I gave in Columbia (see my webpage at the top of the Articles) for more details and more uses of Dyson-Schwinger equations.

3 Heavy tails matrices

3.1 Law of large numbers

For heavy tails matrices, the spectral measure still converges but we have a different limit than the semi-circle law. Here are the models we would like to study :

Definition 3.1 (Models of symmetric heavy tailed matrices with i.i.d. sub-diagonal entries)**.**
Let $A = (a_{i,j})_{i,j=1\le N}$ be a random symmetric matrix with i.i.d. sub-diagonal entries $(a_{i,j})_{i\le j}$.

1. *We say that A is a* **Lévy matrix** *of parameter α in $]0,2[$ when $A = X/a_N$ where the entries x_{ij} of X have absolute values in the domain of attraction of an α-stable distribution, more precisely for all $u \ge 0$*

$$\varphi\left(|x_{ij}| \ge u\right) = \frac{L(u)}{u^\alpha} \tag{63}$$

with a slowly varying function L, and

$$a_N = \inf\{u : P\left(|x_{ij}| \ge u\right) \le \frac{1}{N}\}$$

($a_N = \tilde{L}(N)N^{1/\alpha}$, with $\tilde{L}(\cdot)$ a slowly varying function).

2. *We say that A is a* **Wigner matrix with exploding moments** *with parameter* $(C_k)_{k\geq 1}$ *whenever the entries of A are centered, and for any* $k \geq 1$

$$\lim_{N\to\infty} N\mathbb{E}\left[(a_{ij})^{2k}\right] = C_k, \tag{64}$$

with $(C_{k+1})_{k\geq 0}$ *the sequence of moments of a unique measure m.*

A particular case of matrices with exploding moments is the case of the adjacency matrix of an Erdös-Rényi graph, i.e. of a matrix A such that $A_{ij} = 1$ with probability p/N and 0 with probability $1 - p/N$. It is an exploding moments Wigner matrix, with $C_k = p$ for all $k \geq 1$ (and $m = p\delta_1$). But in all these heavy tails random matrices, there will be typically only finitely many rentries of order one, and many very small or vanishing entries, which is very different from the light tails random matrices which have all entries small (but not vanishing).

The main assumption we will make on the matrices we shall consider is a bit more general than these two types of models and reads as follows.

Assumption 3.2. *Let* μ_N *be the law of* $a_{ij}, i \leq j$. *Assume that uniformly on t in compacts of* $\mathbb{C}^-$

$$\lim_{N\to\infty} N \int (e^{-itx^2} - 1) d\mu_N(x) = \Phi(t)$$

with Φ *such that there exists g on* $\mathbb{R}^+$ *bounded by* Cy^κ *for some* $\kappa > -1$ *such that for* $t \in \mathbb{C}^-$,

$$\Phi(t) = \int_0^\infty g(y) e^{\frac{iy}{t}} dy. \tag{65}$$

Furthermore assume that X with law μ_N *can be decomposed into the law of* $A + B$ *where*

$$P(A \neq 0) \ll N^{-1} \qquad \mathbb{E}[B^2] \ll N^{-1/2}$$

Note that these hypotheses are fulfiled by our examples.

-In the case of Lévy matrices, $\Phi(\lambda) = -\sigma(i\lambda)^{\alpha/2}$ and the expression

$$-\sigma(i\lambda)^{\alpha/2} = \int_{y=0}^{+\infty} C_\alpha y^{\frac{\alpha}{2}-1} e^{i\frac{y}{\lambda}} \mathrm{d}y$$

shows the existence of g satisfying (65) : $g(y) = C_\alpha y^{\frac{\alpha}{2}-1}$. The last point is satisfied with for some $a \in (0, 1/2(2-\alpha))$

$$A = 1_{|x_{ij}| > N^a a_N} \frac{x_{ij}}{a_N}, \quad B = 1_{|x_{ij}| \leq N^a a_N} \frac{x_{ij}}{a_N}.$$

- In the case of Wigner matrices with exploding moment, one first needs to use the following formula, for $\xi \in \mathbb{C}$ with positive real part :

$$1 - e^{-\xi} = \int_0^{+\infty} \frac{J_1(2\sqrt{t})}{\sqrt{t}} e^{-t/\xi} \mathrm{d}t, \tag{66}$$

where J_1 the Bessel function of the first kind defined by $J_1(s) = \frac{s}{2} \sum_{k\geq 0} \frac{(-s^2/4)^k}{k!(k+1)!}$.

It follows that

$$N(\varphi_N(\lambda)-1) = N\mathbb{E}(e^{-i\lambda a^2}-1) = -N\mathbb{E}\int_0^{+\infty} \frac{J_1(2\sqrt{t})}{\sqrt{t}} e^{-\frac{t}{i\lambda a^2}} \mathrm{d}t = \int_0^{+\infty} g_N(y) e^{i\frac{y}{\lambda}} \mathrm{d}y$$

with

$$g_N(y) := -N\frac{\mathbb{E}[|a|J_1(2\sqrt{y}|a|)]}{\sqrt{y}} = -N\mathbb{E}[a^2\frac{J_1(2\sqrt{ya^2})}{\sqrt{ya^2}}] = \int f_y(x)\mathrm{d}m_N(x)$$

for $f_y(x) := -\frac{J_1(2\sqrt{xy})}{\sqrt{xy}}$ and m_N the measure with kth moment given by $N\mathbb{E}[(a_{ij})^{2k}]$. As m_N converges weakly to m and f_y is continuous and bounded, we have

$$g_N(y) \to -\int \frac{J_1(2\sqrt{xy})}{\sqrt{xy}} \mathrm{d}m(x) =: g(y)\,.$$

In the context of heavy tails matrices, we can not use moments anymore. We shall instead rely on Stieljes transform given for $z \in \mathbb{C}\backslash\mathbb{R}$ by

$$G_N(z) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{z-\lambda_i}$$

where the λ_i are the eigenvalues of A.

Theorem 3.3. *Under Assumption 3.2, G_N converges almost surely towards G given , for $z \in \mathbb{C}^+$, by*

$$G(z) = i\int_0^{\infty} e^{itz} e^{\rho_z(t)} dt$$

where $\rho_z : \mathbb{R}^+ \to \{x+iy; x \le 0\}$ is the unique solution analytic in $z \in \mathbb{C}^+$ of the fixed point equation

$$\rho_z(t) = \int_0^{\infty} g(y) e^{\frac{iy}{t}z + \rho_z(\frac{y}{t})} dy$$

A key observation is the following concentration of measure result:

Lemma 3.4. *Let $\|f\|_{TV}$ be the total variation norm,*

$$\|f\|_{TV} = \sup_{x_1<\cdots<x_n} \sum_{i=2}^{n} |f(x_i) - f(x_{i-1})|$$

Then, for any $\delta > 0$ and any function f with finite total variation norm so that $E[|\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i)|] < \infty$,

$$P\left(\left|\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i)]\right| \ge \delta\|f\|_{TV}\right) \le 2e^{-\frac{N\delta^2}{8}}$$

Remark 3.5. *Note that the above speed is not optimal for laws μ,ν which have sufficiently fast decaying tails, in which case $\sum_{i=1}^{N} f(\lambda_i) - \mathbb{E}[\sum_{i=1}^{N} f(\lambda_i)]$ is of order one. However it is the optimal rate for instance for heavy tails matrices where the central limit theorem holds for $N^{-1/2}(\sum_{i=1}^{N} f(\lambda_i) - \mathbb{E}[\sum_{i=1}^{N} f(\lambda_i)])$.*

Remark 3.6. *Note that we only required independence of the vectors, rather than the entries.*

Proof of Lemma 3.4. Let us first recall Azuma-Hoeffding's inequality

Lemma 3.7. *(Azuma-Hoeffding's inequality) Suppose $M_k, k \geq 0$ is a martingale for the filtration $\mathcal{F}_k$ and $|M_k - M_{k-1}| \leq c_k$. Then for all $t \geq 0$*

$$P(M_n - M_0 \geq t) \leq \exp\{-\frac{t^2}{2\sum_{k=1}^n c_k^2}\}.$$

We finally prove Lemma 3.4 for a continuously differentiable function f, the generalization to all functions with finite variation norm then holds by density. We then have $\|f\|_{TV} = \int |f'(x)|dx$. We apply Azuma-Hoeffding 's inequality to

$$M_k = \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i)|\mathcal{F}_k]$$

where $\mathcal{F}_k$ is the filtration generated by $\{X_N(i,j), 1 \leq i \leq j \leq k\}$ for Wigner matrices. M_k is a martingale obviously and

$$M_N - M_0 = \frac{1}{N}\sum_{i=1}^N f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i)].$$

Therefore we need to bound for each $k \in \{1, \cdots, N\}$

$$M_k - M_{k-1} \quad = \quad \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i) - \frac{1}{N}\sum_{i=1}^N f(\tilde{\lambda}_i)|\mathcal{F}_k].$$

where in the above expectation λ_i and $\tilde{\lambda}_i$ are the eigenvalues of the $N \times N$ matrix X_N and Z_N respectively, where Z_N has the same entries than X_N except for the kth vector where we take independent copies. Hence the eigenvalues λ and $\tilde{\lambda}$ are the eigenvalues of two operators which differ at most by a rank one perturbation. This implies that their spectral measures are close by the following lemma :

Lemma 3.8. *let X, Y be two $N \times N$ Hermitian matrices so that $Y - X$ has rank one. Let $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_N$ (resp. $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \cdots \leq \tilde{\lambda}_N$) be the ordered eigenvalues of X and Y respectively. Then, for any C^1 function g on the real line*

$$\left|\sum_{i=1}^N g(\lambda_i) - \sum_{i=1}^N g(\tilde{\lambda}_i)\right| \leq 2\|g\|_{TV}.$$

Proof. Since the two matrices differ only by a rank one matrix, the eigenvalues λ_i and $\tilde{\lambda}_i$ are interlaced by Weyl interlacing property, see e.g [1, Theorem A.7] :

$$\tilde{\lambda}_{i-1} \leq \lambda_i \leq \tilde{\lambda}_{i+1}.$$

If g is increasing we deduce that

$$\sum_{i=1}^{N-2} g(\tilde{\lambda}_i) \le \sum_{i=2}^{N-1} g(\lambda_i) \le \sum_{i=3}^{N} g(\tilde{\lambda}_i)$$

which implies

$$|\sum_{i=1}^{N} g(\lambda_i) - \sum_{i=1}^{N} g(\tilde{\lambda}_i)| \le 2\|g\|_\infty \tag{67}$$

Decomposing $f(x)-f(0)$ as the difference of two increasing functions

$$f(x)-f(0) = \int_0^x f'(y)1_{f'(y)\ge 0}dy - \int_0^x (-f')(y)1_{f'(y)<0}dy$$

proves the claim since

$$\|f\|_{TV} \ge \|\int_0^x f'(y)1_{f'(y)\ge 0}dy\|_\infty + \|\int_0^x (-f')(y)1_{f'(y)<0}dy\|_\infty$$

□

Proof of Theorem 3.3 Because of Lemma 3.4, it is enough to prove that G_N converges in L^1. To do so we shall use will study the following Schur complement formula.

Lemma 3.9. *Let X be a symmetric matrix, and let X_i denote the i-th column of X with the entry $X(i,i)$ removed (i.e., X_i is an $N-1$-dimensional vector). Let $X^{(i)}$ denote the matrix obtained by erasing the i-th column and row from X. Then, for every $z \in \mathbb{C}\setminus\mathbb{R}$,*

$$(X-zI)^{-1}(i,i) = \frac{1}{X(i,i)-z-X_i^*(X^{(i)}-zI_{N-1})^{-1}X_i}. \tag{68}$$

Proof of Lemma 3.9 Note first that from Cramer's rule,

$$(X-zI_N)^{-1}(i,i) = \frac{\det(X^{(i)}-zI_{N-1})}{\det(X-zI)}. \tag{69}$$

Write next

$$X-zI_N = \begin{pmatrix} X^{(N)}-zI_{N-1} & X_N \\ X_N^* & X(N,N)-z \end{pmatrix},$$

and use the matrix identity

$$\begin{aligned} \det\begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \det\left(\begin{pmatrix} A & 0 \\ C & D-CA^{-1}B \end{pmatrix}\begin{pmatrix} 1 & A^{-1}B \\ 0 & 1 \end{pmatrix}\right) \\ &= \det A \det(D-CA^{-1}B) \end{aligned} \tag{70}$$

with $A = X^{(N)}-zI_{N-1}$, $B = X_N$, $C = X_N^*$ and $D = X(N,N)-z$ to conclude that

$$\begin{aligned} &\det(X-zI_N) = \\ &\det(X^{(N)}-zI_{N-1})\det\left[X(N,N)-z-X_N^*(X^{(N)}-zI_{N-1})^{-1}X_N\right]. \end{aligned}$$

The last formula holds in the same manner with $X^{(i)}, X_i$ and $X(i,i)$ replacing $X^{(N)}, X_N$ and $X(N,N)$ respectively. Substituting in (69) completes the proof of Lemma 3.9. □The main idea in the proof is that the convergence and fluctuations of the term $X_i^T(X^{(i)} - zI_{N-1})^{-1}X_i$ in terms of G_N will provide the convergence and fluctuations of $G_N(z)$. To this end let us write

$$\begin{aligned} X_i^T(X_N^{(i)} - zI)^{-1}X_i &= \sum_{j\neq k} X_{ij}X_{ik}(X_N^{(i)} - zI)^{-1}_{jk} + \sum_j |X_{ij}|^2(X^{(i)} - zI)^{-1}_{jj} \\ &=: O(z) + D(z) \end{aligned}$$

We first observe that the off diagonal terms $O(z)$ will always be negligible

Lemma 3.10. *Under the assumption 3.2, for all $\varepsilon > 0$, for any matrix C such that $N^{-1}tr(CC^T)$ is bounded independently of N*

$$\lim_{N\to\infty} \mathbb{P}(|\sum_{j\neq k} X_{ij}X_{ik}C_{jk}| \geq \delta) = 0\,.$$

Proof. recall that we can write $X_{ij} = A_{ij} + B_{ij}$ where $P(A_{ij} \neq 0) \ll N^{-1}$ and $\mathbb{E}[B_{ij}^2] \ll N^{-1/2}$ Hence

$$\mathbb{P}(|\sum_{j\neq k} X_{ij}X_{ik}C_{jk}| \geq \delta) \leq \mathbb{P}(|\sum_{j\neq k} B_{ij}B_{ik}C_{jk}| \geq \delta) + o(1)$$

We then apply Chebychev inequality and the independence of the B's to deduce by Chebychev's inequality that

$$\mathbb{P}(|\sum_{j\neq k} B_{ij}B_{ik}C_{jk}| \geq \delta) \leq \frac{1}{\delta^2}\sum_{j\neq k} \mathbb{E}[B^2]^2 C_{jk}^2 = \frac{N\mathbb{E}[B^2]^2}{\delta^2}\frac{1}{N}\mathrm{tr}(C^2)$$

which goes to zero as N goes to infinity. □

The next difficulty is that the diagonal elements of the resolvent are not approximately deterministic anymore, but rather behave like independent random variables. Indeed $\sum_j |X_{ij}|^2(zI - X^{(i)})^{-1}_{jj}$ remains random in the limit $N \to \infty$ as can be seen if we compute for instance its Fourier transform for $t \in \mathbb{R}^+$:

$$\begin{aligned} \mathbb{E}[e^{-it\sum_j |X_{ij}|^2(zI - X^{(i)})^{-1}_{jj}}] &= \mathbb{E}\left[\prod_{j=1}^N \mathbb{E}_{X_i}[e^{-it|X_{ij}|^2(-X^{(i)}+zI)^{-1}_{jj}}]\right] \\ &= \mathbb{E}\left[\prod_{j=1}^N (1 + \frac{1}{N}\Phi(t(z - X_N^{(i)})^{-1}_{jj}) + o(\frac{1}{N}))\right] \\ &= \mathbb{E}\left[e^{\frac{1}{N}\sum_{j=1}^N \Phi(t(z - X_N^{(i)})^{-1}_{jj}) + o(1)}\right] \end{aligned}$$

where in the second line we used Assumption 3.2. To prove the convergence of G_N it is thus natural to study the order parameter

$$\rho_z^N(t) := \frac{1}{N}\sum_{j=1}^N \Phi(t(z - X_N)^{-1}_{jj})\,.$$

First, we can show that $\rho_z^N(t)$ self-averages as in Lemma 3.4 thanks to the fact that Φ is smooth on $\mathbb{C}^-$. Indeed, we can use also Azuma-Hoefding inequality and the same martingale decomposition to reduce the problem to bound uniformly

$$\sum_{i=1}^N \Phi((z-X^{(j)})_{ii}^{-1}) - \sum_{i=1}^N \Phi((z-X^{(j+1)})_{ii}^{-1})$$

where $X^{(j)} - X^{(j-1)}$ has rank one. But, following Lemma C.3 in [6], we notice that if X, Y are two Hermitian matrices so that $X - Y$ has rank one, then for any function f with finite total variation norm, we have

$$\left|\sum_{i=1}^N f((z-X)_{ii}^{-1}) - \sum_{i=1}^N f((z-Y)_{ii}^{-1})\right| \le \|f\|_{TV} \sum_{i=1}^N \left|\left((z-X)^{-1} - (z-Y)^{-1}\right)_{ii}\right|$$

But $M := (z-X)^{-1} - (z-Y)^{-1}$ has rank one and is uniformly bounded by $2/|\Im z|$, hence $M = \pm\|M\| ee^*$ for some unit vector v. It follows that

$$\left|\sum_{i=1}^N f((z-X)_{ii}^{-1}) - \sum_{i=1}^N f((z-Y)_{ii}^{-1})\right| \le \|f\|_{TV} \frac{2}{|\Im z|} \sum_{i=1}^N |v_i|^2 = \|f\|_{TV} \frac{2}{|\Im z|},$$

which is the desired bound.

To get an equation for the order parameter ρ_z^N, notice that by Schur complement formula and symmetry that

$$\mathbb{E}[\rho_z^N(t)] = \mathbb{E}\left[\Phi\left(t\left(z - X(1,1) + X_1^T(zI - X^{(1)})^{-1}X_1\right)^{-1}\right)\right].$$

Again, we may neglect the off diagonal terms as in Lemma 3.10 (note that the A_i's may be assumed to vanish with high probability and then the L^2 norm argument holds), and therefore deduce that since Φ is smooth by continuous by assumption

$$\mathbb{E}[\rho_z^N(t)] = \mathbb{E}\left[\Phi\left(t\left(z - \sum_{k=2}^N |X_{ik}|^2 (zI - X^{(1)})_{kk}^{-1}\right)^{-1}\right)\right] + o(1).$$

Therefore we deduce from Assumption (65) that

$$\begin{aligned}
\mathbb{E}[\rho_z^N(t)] &= \int_0^\infty g(y)\mathbb{E}[e^{i\frac{y}{t}(z - X_{ii} - \sum |X_{ij}|^2 (z-X^i)_{jj}^{-1})}]dy + o(1) \\
&= \int_0^\infty g(y) e^{i\frac{y}{t}z} \prod_{j=1}^N (1 + \frac{1}{N}\Phi(\frac{y}{t}(z - X^i)_{jj}^{-1}))]dy + o(1) \\
&= \int_0^\infty g(y) e^{i\frac{y}{t}z} e^{\rho_z^N(y/t)} dy + o(1) \qquad (71)
\end{aligned}$$

It is not hard to see that ρ_z^N is sequentially tight as a continuous function on $\mathbb{R}^+$, for instance by Arzela-Ascoli theorem. By (71), any limit point ρ_z satisfies

$$\rho_z(t) = \int_0^\infty g(y) e^{i\frac{y}{t}z} e^{\rho_z(y/t)} dy.$$

We also note that by definition, ρ_z^N takes its values in $\{x+iy, x \le 0\}$. We claim that there exists at most one solution with values with non positive real part for $\Im z$ big enough. Indeed if we had two such solutions ρ and $\tilde\rho$, and we denote $\Delta(t) = |\rho_z(t) - \tilde\rho_z(t)|$ their difference, then as $|g(y)| \le C y^\kappa, \kappa > -1$

$$\Delta(t) \le C \int_0^\infty y^\kappa \Delta(y/t) \wedge 1 e^{-\Im z y/t} dy = C t^{\kappa+1} \int y^\kappa \Delta(y) \wedge 1 e^{-\Im z y} dy$$

where we used that $\rho, \tilde\rho$ have non positive real parts. Integrating under $t^\kappa e^{-\Im z t} dt$ on both sides yields

$$I := \int y^\kappa \Delta(y) e^{-\Im z y} dy \le C \int t^{2\kappa+1} e^{-\Im z t} dt \times \int y^\kappa \Delta(y) \wedge 1 e^{-\Im z y} dy$$

Since $\int y^\kappa \Delta(y) \wedge 1 e^{-\Im z y} dy$ is finite and smaller than I, we deduce for $\Im z$ large enough so that

$$C \int t^{2\kappa+1} e^{-\Im z t} dt < 1$$

that $I = 0$. But we claim that for each given t, $N \to \rho_z^N(t)$ is analytic away from the real axis. Indeed, Φ is analytic on $\{x+iy, y<0\}$ by (65) and $(z-X)^{-1}_{ii}$ is analytic on $\Im z > 0$, with image in $\{x+iy, y<0\}$. We have also seen it is uniformly bounded. Hence any limit point must be analytic on $\{\Im z > 0\}$ by Montel's theorem. We conclude that $\rho_z(t)$ is uniquely determined by its values for $\Im z$ large and therefore uniquely defined by our equation. To conclude, $\rho_z^N(t)$ converges almost surely and in L^1 towards $\rho_z(t)$.

This characterizes also the limit of G_N. Indeed, by concentration inequalities we have almost surely that

$$\begin{aligned}
G_N(z) &= \mathbb{E}[G_N(z)] + o(1) \\
&= \mathbb{E}[\frac{1}{z - \sum |X_{ji}|^2 (z - X^{(1)})^{-1}_{jj}}] + o(1) \\
&= i \int_0^\infty dt \mathbb{E}\left[e^{it(z - \sum |X_{ji}|^2 (z - X^{(1)})^{-1}_{jj})} \right] + o(1) \\
&= i \int_0^\infty dt e^{itz + \rho_z(t)} + o(1)
\end{aligned}$$

As G_N is tight on $\Omega_\varepsilon = \{\Im z \ge \varepsilon\}$ for all $\varepsilon > 0$ by Arzela-Ascoli theorem, and its limit points are analytic by Montel's theorem (as G_N is uniformly bounded on Ω_ε), this implies the convergence of G_N on $\mathbb{C}^+$ to the unique analytic function on $\mathbb{C}^+$ given by the above formula for $\Im z$ large enough.

3.2 CLT

In the case of heavy tails matrices the CLT holds in the usual scale:

Theorem 3.11. *Under the assumptions of Theorem 3.3 and if we assume additionally that we can write*

$$\Phi(x+y) = \int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{\frac{it}{x} + \frac{is}{y}} d\tau(t,s)$$

with a measure $d\tau(t,s) = \delta_{t=0} d\mu(s) + \delta_{s=0} d\mu(t) + f(t,s) ds dt$ *with* $|f(t,s)| \le Ct^{\kappa} + Cs^{\kappa}$, $\kappa > -2$, *then for all* $z_1, \ldots, z_p \in \mathbb{C} \backslash \mathbb{R}$,

$$\left(\mathbb{Z}_N(z_i) := \frac{1}{\sqrt{N}} \left(\mathrm{Tr}((z_i - X)^{-1}) - \mathbb{E}[\mathrm{Tr}((z_i - X)^{-1})] \right), 1 \le i \le p \right)$$

converges in law towards a centered Gaussian vector.

As an exercise, you can check that the new assumption is verified for Erdös-Renyi matrices and for Lévy except for the bound on f (some argument then needs to be adapted). Notice that the scaling of the fluctuations is $N^{-\frac{1}{2}}$ as for independent variables, but differently from light tails matrices. The interested reader can read

References

[1] G. Anderson, A. Guionnet, O. Zeitouni *An Introduction to Random Matrices*. Cambridge studies in advanced mathematics, 118 (2009).

[2] G. Anderson, O. Zeitouni *A CLT for a band matrix model*, Probab. Theory Rel. Fields, 2005, **134**, 283–338

[3] Benaych Georges, Florent and Guionnet, Alice and Male, Camille, *Central limit theorem for heavy tailed random matrices*, Comm. Math. Phys. **329** 2014,641-686

[4] Bordenave, Charles and Guionnet, Alice, *Localization and delocalization of eigenvectors for heavy-tailed random matrices*, Probab. Theory Related Fields,**157** 2013, 885–953.

[5] Borodin, Alexei and Gorin, Vadim and Guionnet, Alice, *Gaussian asymptotics of discrete* β*-ensembles*, Publications mathématiques de l'IHÉS, 2016, 1–78.

[6] Borot, G. and Guionnet, A.; Asymptotic expansion of beta matrix models in the one-cut regime. *Comm. Math. Phys.*, **317**, 447-483 (2013)

[7] A. Guionnet, E. Maurel-Segala *Combinatorial aspects of matrix models*, Alea, 2006, **1**, 241–279. *Second order asymptotics for matrix models*, Ann. Probab., **35**, 2007, 2160–2212.

[8] A. Guionnet, E. Maurel-Segala *Second order asymptotics for matrix models*, Ann. Probab., **35**, 2007, 2160–2212.

[9] J. Harer and D. Zagier, *The Euler characteristic of the moduli space of curves* Invent. Math.,**85** 1986, 457–485.

[10] K. Johansson *On the fluctuations of eigenvalues of random Hermitian matrices.* Duke Math. J. **91** 1998, 151–204.

[11] Kontsevich, Maxim, *Vassiliev's knot invariants*, Gelfand Seminar, Adv. Soviet Math., **16**, 1993, 137–150

[12] F.G. Tricomi, *Integral equations*,Pure Appl. Math., London (1957)

[13] E.P. Wigner *On the distribution of the roots of certain symmetric matrices*, Annals Math. , **67**, 1958, 325–327.

[14] A. Zvonkin, *Matrix integrals and maps enumeration: an accessible introduction*, Math. Comput. Modeling, **26** 1997, 281–304.