The largest eigenvalue of finite rank deformation of Wigner matrices

M. Capitaine, C. Donati-Martin and D. Féral

1 The model

$$M_N = X_N + A_N := \frac{1}{\sqrt{N}} W_N + A_N$$

where W_N is a $N \times N$ Hermitian matrix such that $(W_N)_{ii}$, $\sqrt{2}Re((W_N)_{ij})_{i<j}$, $\sqrt{2}Im((W_N)_{ij})_{i<j}$ are iid with common distribution μ. μ is assumed to be symmetric with variance σ^2 and it satisfies a Poincaré inequality.
A_N is a deterministic, Hermitian matrix.

Example: $\mu = N(0;\sigma^2)$, $X_N \sim GUE(N, \frac{\sigma^2}{N})$.

2 Some known result in the non deformed case ($A_N = 0$)

- Convergence of the spectral measure $\mu_{X_n} := \frac{1}{N}\sum_i \delta_{\lambda_i(X_N)}$ to the Wigner distribution $\mu_{sc} = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}1_{[-2\sigma,2\sigma]}$.

- Convergence a.s. of $\lambda_{max}(X_N)$ to 2σ (the right end-point of the support of the limiting distribution)
- Fluctuations (Tracy-Widom, Soshnikov)

  $$\sigma^{-1} N^{2/3} \left(\lambda_{max}(X_N) - 2\sigma \right) \xrightarrow{\mathcal{L}} \text{ T-W distribution } F_2$$

  where the distribution F_2 can be expressed with the Fredholm determinant of an operator associated to the Airy kernel.

3 The deformation

A_N Hermitian of finite rank r (independent of N) with eigenvalues θ_i of multiplicity k_i; $\theta_1 > \theta_2 > \ldots > \theta_J$.
Convergence of the spectral measure to the semicircular distribution μ_{sc}.
What about the extremal eigenvalues?

1) The Gaussian case (Péché)
Ex: θ_1 with multiplicity 1.

1) si $0 \leq \theta_1 < \sigma$, $\quad \sigma^{-1} N^{2/3} \left(\lambda_{max}(M_N) - 2\sigma\right) \xrightarrow{\mathcal{L}} F_2$

2) si $\theta_1 = \sigma$, $\quad \sigma^{-1} N^{2/3} \left(\lambda_{max}(M_N) - 2\sigma\right) \xrightarrow{\mathcal{L}} F_3$

3) si $\theta_1 > \sigma$, $\quad N^{1/2} \left(\lambda_{max}(M_N) - \rho_{\theta_1}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_{\theta_1}^2)$
with $\rho_{\theta_1} = \theta_1 + \frac{\sigma^2}{\theta_1} > 2\sigma$.

2) The non Gaussian case for a particular A_N (Féral-Péché)
A_N is the deformation defined by $(A_N)_{ij} = \frac{\theta}{N}$, so that $r = 1$ and $\theta_1 = \theta$.
Same TCL as in the Gaussian case, universality of the fluctuations (independent of μ, the distribution of the entries).

3) The non Gaussian case, A_N general

Theorem 1 *a.s. behaviour of the spectrum of M_N. Let $J_{+\sigma}$ (resp. $J_{-\sigma}$) be the number of j's such that $\theta_j > \sigma$ (resp. $\theta_j < -\sigma$).*

(a) $\forall 1 \leq j \leq J_{+\sigma},\ \forall 1 \leq i \leq k_j,$

$$\lambda_{k_1+\cdots+k_{j-1}+i}(M_N) \longrightarrow \rho_{\theta_j} \quad a.s.,$$

(b) $\lambda_{k_1+\cdots+k_{J_{+\sigma}}+1}(M_N) \longrightarrow 2\sigma \quad a.s.,$

(c) $\lambda_{k_1+\cdots+k_{J-J_{-\sigma}}}(M_N) \longrightarrow -2\sigma \quad a.s.,$

(d) $\forall j \geq J - J_{-\sigma} + 1, \forall 1 \leq i \leq k_j,$

$$\lambda_{k_1+\cdots+k_{j-1}+i}(M_N) \longrightarrow \rho_{\theta_j} \quad a.s.$$

Remark: Same result as in the sample covariance matrices (Bai-Silverstein, Baik-Silverstein)

Theorem 2 *Fluctuations*
Let $A_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ and assume that $\theta > \sigma$. Then

$$\sqrt{N}\Big(\lambda_{max}(M_N) - \rho_\theta\Big) \xrightarrow{\mathcal{L}} (1 - \frac{\sigma^2}{\theta^2})\Big\{\mu * \mathcal{N}(0, v_\theta)\Big\}.$$

where $v_\theta = v(\theta, \sigma^2, \int x^4 d\mu(x))$.

$\longrightarrow$ Non universality of the fluctuations.

4 Elements of Proof of Theorem 1

Step 1 Prove that a.s.

$$Spect(M_N) \subset K_\sigma(\theta_1, \ldots \theta_J) + [-\epsilon, +\epsilon] \qquad (1)$$

for N large, where $K_\sigma(\theta_1, \cdots, \theta_J) :=$

$$\left\{\rho_{\theta_J}; \cdots ; \rho_{\theta_{J-J_{-\sigma}+1}}\right\} \cup [-2\sigma; 2\sigma] \cup \left\{\rho_{\theta_{J_{+\sigma}}}; \cdots ; \rho_{\theta_1}\right\}.$$

Tool: The Stieltjes transform: for $z \in \mathbb{C}\backslash\mathbb{R}$, define $g_N(z) = \mathrm{tr}_N(G_N(z))$ where $G_N(z) = (zI_N - M_N)^{-1}$ is the resolvent of M_N. We set $h_N(z) = \mathbb{E}[g_N(z)]$.

$$g_N(z) = \int \frac{1}{z-x} d\mu_{M_N}(x);\ h_\sigma(z) = \int \frac{1}{z-x} d\mu_{sc}(x).$$

Aim: Obtain a precise estimate

$$h_\sigma(z) - h_N(z) + \frac{1}{N} L_\sigma(z) = O(\frac{1}{N^2}) \qquad (2)$$

where L_σ is the Stieltjes transform of a distribution η with compact support in K_σ.
With the help of the inverse Stieltjes transform,

$$\mathbb{E}[\mathrm{tr}_N(\varphi(M_N))] = \int \varphi(x) d\mu_{sc}(x) + \frac{1}{N} \int \varphi(x) d\eta(x) + O(\frac{1}{N^2}),$$

for φ smooth with compact support;
and some variance estimates, we deduce from (2)

$$\mathrm{tr}_N\, 1_{{}^c K_\sigma^\varepsilon(\theta_1, \cdots, \theta_J)}(M_N) = O(1/N^{\frac{4}{3}})\ a.s.$$

and therefore the inclusion of the spectrum (1).

Proof of (2):

1) The Gaussian Case:

- The Gaussian integration by parts formula:
  $\phi : \mathbb{R} \to \mathbb{C}$, ξ *standard Gaussian*

$$\mathbb{E}(\xi\phi(\xi)) = \mathbb{E}(\phi^{'}(\xi)).$$

  $\Phi : \mathcal{H}_n(\mathbb{C}) \to \mathbb{C}$, $H \in \mathcal{H}_n(\mathbb{C})$,

$$\frac{N}{\sigma^2}\mathbb{E}[\mathrm{Tr}(X_N H)\Phi(X_N)] = \mathbb{E}[\Phi^{'}(X_N) \cdot H]$$

Apply it for $\Phi(X_N) = [(zI_N - X_N - A_N)^{-1}]_{kl} = G_N(z)_{kl}$ and $H = E_{kl}$; then sum over k and l.

$$\to \sigma^2\mathbb{E}[g_N^2(z)] - z\mathbb{E}[g_N(z)] + 1 + \frac{1}{N}\mathbb{E}[\mathrm{Tr}(G_N(z)A_N)] = 0$$

$$\to \sigma^2 h_N^2(z) - zh_N(z) + 1 + \frac{1}{N}\mathbb{E}[\mathrm{Tr}(G_N(z)A_N)] = O(\frac{1}{N^2})$$

Recall that $\sigma^2 h_\sigma^2(z) - zh_\sigma(z) + 1 = 0$.

Estimate for $\mathbb{E}[\mathrm{Tr}(G_N(z)A_N)]$:

$A_N = U^*\Lambda U$ where Λ is a diagonal matrix with entries $\lambda_i \neq 0$ for $i \leq r$, $\lambda_i = 0$, $i > r$. We can show using

- The Gaussian integration by parts formula
- Some variance estimates
- $h_N(z) = h_\sigma(z) + O(\frac{1}{N})$

the estimate

$$\mathbb{E}[\mathrm{Tr}(G_N(z)A_N)] = \sum_{i=1}^{r} \frac{\lambda_i}{z - \lambda_i - \sigma^2 h_\sigma(z)} + O(\frac{1}{N})$$

Set

$$R_G^{A_N}(z) = \sum_{i=1}^{r} \frac{\lambda_i}{z - \lambda_i - \sigma^2 h_\sigma(z)} = \sum_{\theta_i \neq 0} k_i \frac{\theta_i}{z - \theta_i - \sigma^2 h_\sigma(z)}.$$

Then,

$$\sigma^2 h_N^2(z) - z h_N(z) + 1 + \frac{1}{N} R_G^{A_N}(z) = O(\frac{1}{N^2})$$

leading to

$$h_N(z) - h_\sigma(z) + \frac{1}{N} L(z) = O(\frac{1}{N^2})$$

where $L(z) = h_\sigma^{-1}(z)\mathbb{E}[(z - sc)^{-2}] R_G^{A_N}(z)$.

Question:

- L Stieltjes transform of a distribution ?
- Support of this distribution?

$\longleftrightarrow$ Analyticity of L (+ conditions); set of singular points.

If $|\theta_i| > \sigma$, $x \in \mathbb{R}\backslash[-2\sigma, 2\sigma]$,

$$x - \theta_i - \sigma^2 h_\sigma(x) = 0 \Longleftrightarrow x = \theta_i + \frac{\sigma^2}{\theta_i} := \rho_{\theta_i}.$$

2) The non Gaussian case

GIP replaced by: (Khorunzhy, Khoruzhenko, Pastur)

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

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

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

Apply to $\xi = Re((X_N)_{ij}), Im((X_N)_{ij}), (X_N)_{ii}$, the odd cumulants vanish (μ symmetric). One must consider the third derivative of $\Phi = (G_N(z))_{kl}$.

One obtains:

$$\sigma^2 h_N^2(z) - z h_N(z) + 1 + \frac{1}{N} R(z) = O(\frac{1}{N^2})$$

where $R(z) = R_G^{A_N}(z) + \kappa_4 R^0_{\Phi'''}(z)$.

$R^0_{\Phi'''}(z)$ is analytic on $\mathbb{C} \backslash [-2\sigma, 2\sigma]$.

Step 2: A.s. convergence of the first extremal eigenvalues of M_N.

Lemma 2 *(Weyl) Let B and C be two $N \times N$ Hermitian matrices. For any pair of integers j, k such that $1 \leq j, k \leq N$ and $j + k \leq N + 1$, we have*

$$\lambda_{j+k-1}(B + C) \leq \lambda_j(B) + \lambda_k(C).$$

For any pair of integers j, k such that $1 \leq j, k \leq N$ and $j + k \geq N + 1$, we have

$$\lambda_j(B) + \lambda_k(C) \leq \lambda_{j+k-N}(B + C).$$

5 Sketch of Proof of Theorem 2

$$\lambda_1(M_N) = \theta + \frac{(W_N)_{11}}{\sqrt{N}} + \check{M}_{\cdot 1}^* \widehat{G}(\lambda_1(M_N))\check{M}_{\cdot 1}$$

where $\widehat{G}$ is the resolvent associated to the Wigner matrix of size $N-1$ obtained from M_N by removing the first row and column;

$$\check{M}_{\cdot 1} = {}^t\left((M_N)_{21}, \ldots, (M_N)_{N1}\right) = \frac{1}{\sqrt{N}}\, {}^t\left((W_N)_{21}, \ldots, (W_N)_{N1}\right).$$

Use the resolvent equation:

$$\widehat{G}(\lambda_1(M_N)) - \widehat{G}(\rho_\theta) = -(\lambda_1(M_N) - \rho_\theta)\widehat{G}(\rho_\theta)\widehat{G}(\lambda_1(M_N))$$

and

Theorem 3 *(Bai-Yao, Baik-Silverstein) Let* $B = (b_{ij})$ *be a* $N \times N$ *random Hermitian matrix and* $Y_N = (y_1, \ldots, y_N)$ *be an independent vector of size* N *which contains i.i.d standardized entries with bounded fourth moment and such that* $\mathbb{E}(y_1^2) = 0$ *if* y_1 *is complex. Assume that*

(i) there exists a constant $a > 0$ *(not depending on* N*) such that* $||B|| \leq a$,

(ii) $\frac{1}{N}\mathrm{Tr}B^2$ *converges in probability to a number* a_2,

(iii) $\frac{1}{N}\sum_{i=1}^N b_{ii}^2$ *converges in probability to a number* a_1^2.

Then the random variable $\frac{1}{\sqrt{N}}(Y_N^* B Y_N - \mathrm{Tr}B)$ *converges in distribution to a Gaussian variable with mean zero and variance*

$$(\mathbb{E}|y_1|^4 - 1 - t/2)a_1^2 + (t/2)a_2$$

where $t = 4$ *when* Y_1 *is real and is* 2 *when* y_1 *is complex.*