

Section 3 Invariant Ensembles

Consider the density function of an invariant ensemble as

$$P(H) dH = \frac{1}{Z} \exp\left(-\frac{\beta}{2} N \text{Tr } V(H)\right) dH$$

For GOE / GUE, $V(H) = \frac{1}{2} H^2$. In this case, this model is called a β -ensemble
 $\beta=1$ for GOE; $\beta=2$ for GUE; $\beta=4$ for GSE. symplectic
 \uparrow
 (Gaussian quaternion ensemble)

For general β , it does not correspond to any Wigner ensembles.

① Probability density for β -ensemble

For invariant ensembles, their eigenvectors are uniformly distributed on the unit sphere.

So people mainly focus on their eigenvalues. We will integrate out the eigenvectors in $P(H) dH$ and show the following result.

Theorem 3.1 The joint probability density of the eigenvalues of H is given by

$$(*) P_N(\lambda_1, \lambda_2, \dots, \lambda_N) = \frac{1}{Z_N} \prod_{i < j} |\lambda_i - \lambda_j|^\beta e^{-\frac{\beta}{2} N \sum_{i=1}^N V(\lambda_i)}$$

Rmk 1: Recall that $\Delta(\lambda) = \prod_{i < j} (\lambda_i - \lambda_j)$ is the Vandermonde determinant.

Rmk 2: In the above formula, we can neglect the ordering of the eigenvalues. This can be done by multiplying a $\frac{1}{n!}$ factor when calculating various probabilities.

Rmk 3: We can rewrite (*) as: $P_N(\lambda_1, \lambda_2, \dots, \lambda_N) = \frac{1}{Z_N} \exp\left[-\frac{\beta}{2} \sum_{i=1}^N \log |\lambda_i - \lambda_j|\right]$

Rmk 4: When $|\lambda_i - \lambda_j| = 0$, we have $P_N(\lambda_1, \dots, \lambda_N) = 0$. This indicates a repulsion between different eigenvalues.

Rmk 5: The eigenvalues are strongly correlated. It is useful to think of P_N as a Gibbs measure on N "particles" $\vec{\lambda} = (\lambda_1, \dots, \lambda_N)$:

$$P_N(\vec{\lambda}) d\vec{\lambda} = \exp[-\beta N \mathcal{H}(\vec{\lambda})] \frac{1}{Z_N},$$

$$\mathcal{H}(\vec{\lambda}) := \underbrace{\frac{1}{2} \sum_{i=1}^N V(\lambda_i)}_{\text{confining potential}} - \underbrace{\frac{1}{N} \sum_{i < j} \log |\lambda_i - \lambda_j|}_{\text{logarithmic potential}}$$

[The two terms have the same order " N ".]

A confining potential. A logarithmic potential giving repulsions between particles. This gives an important statistical mechanical system, called the "log-Coulomb gas".

To prove Thm 3.1, we need to show that integrating out the eigenvectors gives the Vandermonde determinant.

Let $H = UDU^*$ be an eigenvalue decomposition of H , where $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ with $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ and U is an orthogonal / ~~unitary~~ unitary matrix. We consider the unitary case. Then U is Haar distributed on $U(N)$.

From the unitary invariance of $\mathbb{P}(H)dH$, we see that after conditioning on D , the eigenmatrix U is drawn from the Haar measure on $U(n)$. In particular, U and D can be taken to be independent:

$$\begin{aligned} \mathbb{P}(H = UDU^* + AH) &= \mathbb{P}(H = (U_{\text{left}} D_0 U_{\text{right}}^* + AH)) \\ \text{for any } V \in U(N) &\quad \mathbb{P}(U \in B, D \in A) = \mathbb{P}(U \in B | D \in A) \mathbb{P}(D \in A) \\ &= \mathbb{P}(U \in B) \mathbb{P}(D \in A). \end{aligned}$$

Proof of Theorem 3.1: Fix a diagonal matrix $D_0 = \text{diag}(\lambda_1^0, \dots, \lambda_N^0)$, $\lambda_1^0 \leq \dots \leq \lambda_N^0$.

Let $\epsilon > 0$ be arbitrarily small. We now compute the probability that H_N lies in an ϵ -ball around D_0 in the H-S norm (Frobenius norm):

$$\mathbb{P}(\|H_N - D_0\|_F \leq \epsilon) \quad \begin{array}{l} \rightarrow \text{The Euclidean norm} \\ \text{on the space of } n \times n \text{ matrices.} \end{array}$$

I. ~~the~~ The simpler direction: The probability density of M_N is $\propto \exp(-\frac{\beta}{2} N \sum_i V(\lambda_i^0))$

The ~~volume~~ volume of the ϵ -ball is $C_N \epsilon^{N^2}$. So

$$(1) \quad \mathbb{P}(\|H_N - D_0\|_F \leq \epsilon) = (C_N + o(1)) \epsilon^{N^2} \exp(-\frac{\beta}{2} N \sum_i V(\lambda_i^0))$$

$o(1) \rightarrow 0$ as $\epsilon \rightarrow 0$ when N is fixed.

II. The harder direction: Let's calculate $\mathbb{P}(\|H_N - D_0\|_F \leq \epsilon)$ in a different way.

Weyl's inequality: Given two hermitian matrices A and B with eigenvalues

$$\lambda_i(A) \leq \dots \leq \lambda_N(A), \quad \lambda_i(B) \leq \dots \leq \lambda_N(B).$$

$$\text{Then, } \forall i, \quad \lambda_i(A) + \lambda_i(B) \leq \lambda_i(A+B) \leq \lambda_i(A) + \lambda_N(B).$$

In particular, if B is a small perturbation with $\|B\| \leq \epsilon$, then

$$\lambda_i(A) - \epsilon \leq \lambda_i(A+B) \leq \lambda_i(A) + \epsilon.$$

Thus, for $H_N = UDU^*$, we have $\|D - D_0\| \leq \epsilon$ when $\|H_N - D_0\|_F \leq \epsilon$.

$$\text{Moreover, } UDU^* = D_0 + O(\epsilon) \Rightarrow D = D_0 + O(\epsilon) \Rightarrow UD - D_0 U^* = O(\epsilon)$$

Now, we make the ansatz

$$D \approx D_0 + \epsilon E, \quad U = \exp(\epsilon S)$$

$\boxed{\text{This diagonal \& unitary (i.e., } U \text{ is a unitary matrix)}}$

where E is a diagonal matrix and S is a skew-adjoint matrix.

$$[\text{Note that } UU^* = \exp(\epsilon S) \exp(\epsilon S^*) - \exp(\epsilon S) \exp(-\epsilon S) = I]$$

Note that $S \mapsto \exp(\epsilon S)$ has a non-degenerate Jacobian in the small ϵ -ball, so the inverse function theorem applies to specify S from U . (Formally, $S = \frac{1}{\epsilon} \log U$.)

WLOG, suppose all λ_i^0 are different (~~other~~ this holds a.s.) -

thus, we make the ansatz: $D = D_0 + \epsilon E$, E is diagonal bounded.

Note, the eigenvalues of D_0 are non-zero and non-degenerate almost surely. Moreover, we have $UDU^* = D_0 + O(\epsilon) \Rightarrow UDU^* = D_0 + O(\epsilon) \Rightarrow UD_0 - D_0 U = O(\epsilon)$. (*)

The only unitary matrices that commute with D_0 are diagonal unitary matrices

$$R = (r_1, r_2, \dots, r_N), \quad \|r_i\|^2 = 1.$$

From (*), we can make the ansatz: $U = \exp(\epsilon S) R$,

where S is a bounded skew-adjoint ($S^* = -S$) matrix with zero diagonal. (The diagonal parts can be included into R .)

We can check that $UU^* = \exp(\epsilon S) \exp(\epsilon S^*) = I$,

$$UD_0 - D_0 U = \exp(\epsilon S) RD_0 - D_0 \exp(\epsilon S) R = RD_0 - D_0 R + O(\epsilon) = O(\epsilon).$$

It is easy to check that $(R, S) \mapsto \exp(\epsilon S) R$ has a non-degenerate Jacobian, so with the inverse function thm, we can uniquely determine R and S from U in the ~~same~~ small ϵ -ball around D_0 .

Now with the ansatz: $D = D_0 + \epsilon E$, $U = \exp(\epsilon S) R$, we get

$$\begin{aligned} H_N &= UDU^* = \exp(\epsilon S) R (D_0 + \epsilon E) R^* \exp(-\epsilon S) \\ &= D_0 + \epsilon E + \epsilon S R D_0 R^* - \epsilon R D_0 R^* S + O(\epsilon^2) \\ &= D_0 + \epsilon E + \epsilon (SD_0 - D_0 S) + O(\epsilon^2). \end{aligned}$$

Thus, $\mathbb{P}(\|H_N - D_0\|_F \leq \epsilon) = \mathbb{P}(\|E + (SD_0 - D_0 S)\| \leq 1 + O(\epsilon))$. (+)

Since U is Haar distributed on $U(N)$, S is locally ^{distributed} as $C_N' \epsilon^{N^2-N}$ times the Lebesgue measure on the space of skew-adjoint matrices with 0 diagonal.

~~On the other hand, E is distributed as $(P_N(\vec{\lambda}^0) + o(1)) \epsilon^N$~~ times the Lebesgue measure on the space of diagonal matrices. Thus we can calculate (+) as:

$$C_N' \epsilon^{N^2} (P_N(\vec{\lambda}^0) + o(1)) \iint_{\|E + SD_0 - D_0 S\|_F \leq 1 + O(\epsilon)} dE ds$$

Consider the map $S \mapsto SD_0 - D_0 S$, where $(SD_0 - D_0 S)_{ij} = (\lambda_j^0 - \lambda_i^0) S_{ij}$.

In other words, the map dilates the (i,j) th entry of S by $\lambda_j^0 - \lambda_i^0$. Hence, the Jacobian of this map is

$$\prod_{i \neq j} |\lambda_j^0 - \lambda_i^0| = \prod_{1 \leq i < j \leq N} |\lambda_j^0 - \lambda_i^0|^2.$$

Applying the change of variable, we get

$$\begin{aligned} \iint_{\|E + SD_0 - D_0 S\|_F \leq 1 + O(\epsilon)} dE ds &= \frac{1}{\prod_{1 \leq i < j \leq N} |\lambda_j^0 - \lambda_i^0|^2} \iint_{\|E + SD_0 - D_0 S\|_F \leq 1 + O(\epsilon)} dE ds \\ &= \frac{C_N'}{\prod_{1 \leq i < j \leq N} |\lambda_j^0 - \lambda_i^0|^2}. \end{aligned}$$

In sum, we have found that

$$P(\|H_N - D_0\|_F \leq \varepsilon) = (C_N + o(1)) \varepsilon^{N^2} \exp\left[-\frac{\beta}{2} N \sum_i V(\lambda_i^\circ)\right]$$

$$= \frac{C_N'''}{\prod_{1 \leq i < j \leq N} |\lambda_j^\circ - \lambda_i^\circ|^2} [P_N(\vec{\lambda}^\circ) + o(1)] \varepsilon^{N^2}$$

$$\Rightarrow P_N(\vec{\lambda}^\circ) = \frac{1}{Z_N} \prod_{1 \leq i < j \leq N} |\lambda_j^\circ - \lambda_i^\circ|^2 \exp\left[-\frac{\beta}{2} N \sum_{i=1}^N V(\lambda_i^\circ)\right]. \quad \square$$

② Mean field approximation of semicircle law

$$P_N(\vec{\lambda}) \quad P_N(\vec{\lambda}^\circ) = \frac{1}{Z_N} \exp[-\beta N H(\vec{\lambda})], \quad H(\vec{\lambda}) = \frac{1}{2} \sum_{i=1}^N V(\lambda_i) - \frac{1}{N} \sum_{i < j} \log |\lambda_i - \lambda_j|.$$

Intuitively, it is plausible that the spectrum should concentrate around the maximum of $P_N(\vec{\lambda})$ as $N \rightarrow +\infty$. It is equivalent to study the minimum of $H(\vec{\lambda})$.

Heuristically, we make a mean-field approximation, i.e., the ESD $\frac{1}{N} \sum_i \delta_{\lambda_i}$ can be approximated by a continuous probability measure $p(x) dx$ (\rightarrow this is reasonable for GOE/GUE at least with $P = p_{sc}$.)

Then, we expect that $H(\vec{\lambda})$ is approximately given by

$$H(p) = \frac{1}{2} \int_{\mathbb{R}} V(x) dx$$

$$H(p) = \frac{1}{2} \int_{\mathbb{R}} V(x) p(x) dx + \int_{\mathbb{R}} \int_{\mathbb{R}} p(x) p(y) \log \frac{1}{|x-y|} dx dy.$$

p should minimize this functional. Now, we derive the Euler-Lagrange equation for $H(p)$: consider $p + \delta p$, where p is the minimizer. We should have

$$\frac{\delta H(p)}{\delta p} = 0, \text{ because } H(p + \delta p) \geq H(p) \forall \delta p.$$

$$\text{we have } \delta H(p) = \frac{1}{2} \int_{\mathbb{R}} V(x) \delta p(x) dx + \int_{\mathbb{R}} \int_{\mathbb{R}} (\delta p(x) p(y)) \log \frac{1}{|x-y|} dx dy$$

$$+ \int_{\mathbb{R}} \int_{\mathbb{R}} p(x) (\delta p(y)) \log \frac{1}{|x-y|} dx dy$$

$$\Rightarrow \frac{1}{2} \int_{\mathbb{R}} V(x) \delta p(x) dx + 2 \int_{\mathbb{R}} \left(\int_{\mathbb{R}} p(y) \log \frac{1}{|x-y|} dy \right) \delta p(x) dx = 0$$

Since δp is arbitrary, we ought to have $\frac{1}{2} V(x) + 2 \int_{\mathbb{R}} p(y) \log \frac{1}{|x-y|} dy = 0$ with $\int_{\mathbb{R}} \delta p(x) dx = 0$.

(since $p + \delta p$ should be a prob. density).

This is the E-L equation for p .

Solving the E-L equation, we get a density $p(x)$, which should be the limiting ESD of $\frac{1}{N} \text{Tr } V(H) dH = \frac{1}{N} \exp[-\frac{\beta}{2} N \text{Tr } V(H)] dH$.

Let's solve it for GUE with $V(x) = x^2$. We need to solve

$$\frac{1}{2}x^2 + 2 \int_{\mathbb{R}} p(y) \log |x-y| dy = C \Rightarrow x + 2 \operatorname{PV} \int_{\mathbb{R}} \frac{p(y)}{y-x} dy = 0 \quad (*)$$

Here, $\int_{\mathbb{R}} \frac{p(y)}{y-x} dy$ should be formally understood as a principal value

$$\begin{aligned} \operatorname{P.V.} \int_{\mathbb{R}} \frac{p(y)}{y-x} dy &:= \lim_{\varepsilon \rightarrow 0} \int_{|y-x| \geq \varepsilon} \frac{p(y)}{x-y} dy = \lim_{\eta \rightarrow 0^+} \operatorname{Re} \int_{\mathbb{R}} \frac{p(y)}{y-(x+i\eta)} dy \\ &= \lim_{\eta \rightarrow 0^+} \operatorname{Re} S_p(x+i\eta) =: \operatorname{Re} S_p(x+i0^+) \end{aligned}$$

On the other hand, recall $p(x) = \frac{1}{\pi} \operatorname{Im} S_p(x+i0^+)$.

Then the equation $(*)$ reads : $x + 2 \operatorname{Re} S_p(x+i0^+) = 0$

$$\Rightarrow x \operatorname{Im} S_p(x+i0^+) + 2 [\operatorname{Re} S_p(x+i0^+)] [\operatorname{Im} S_p(x+i0^+)] = 0$$

$$\Rightarrow \operatorname{Im} [x S_p(x+i0^+) + S_p^2(x+i0^+)] = 0$$

Define function $f(z) = z S_p(z) + S_p^2(z)$. The above function means we can make an analytic continuation from $z \in \overline{\mathbb{C}_+} \rightarrow \mathbb{C}$ by making a reflection along \mathbb{R} :

$$f(z) = \overline{f(\bar{z})} \quad \text{for } z \in \mathbb{C}.$$

On the other hand, since $S(z) = \frac{-1+o(1)}{z}$ as $z \rightarrow \infty$, $f \rightarrow -1$ as $z \rightarrow +\infty$.

By Liouville's theorem, f is a constant and

$$f(z) \equiv \lim_{z \rightarrow \infty} [z S_p(z) + S_p^2(z)] = -1$$

This gives the familiar self-consistent equation:

$$S_p^2(z) + z S_p(z) + 1 = 0 \Rightarrow S_p(z) = m_{sc}(z) \Rightarrow p = p_{sc}.$$

The above arguments can be made rigorous for general β ensembles as long as we impose some conditions on V (in particular, need certain condition on its growth at ∞).

③ Bulk eigenvalue distribution of GUE and orthogonal polynomials

Recall that the Vandermonde determinant $\Delta(\vec{\lambda})$ can be expressed (up to a \pm sign) as a determinant of an $n \times n$ matrix

$$M = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

[We take $n \leq N$ in this section.]

Then, $|\Delta_n(\vec{\lambda})|^2$ can be expressed as $\det(MM^*)$, where

$$(MM^*)_{ij} = \sum_{k=0}^{n-1} \lambda_i^k \lambda_j^k, \quad 1 \leq i \leq j \leq n.$$

Through ~~row~~^{column} operations of M , we observe the following fact: if $P_0(\lambda), P_1(\lambda), \dots, P_{n-1}(\lambda)$ are any sequence of polynomials, in which $P_i(x)$ has degree i and the degree i term is exactly x^i , then $\Delta(\lambda_1, \dots, \lambda_n)$ is equal to the det of the matrix $(P_{j-1}(\lambda_i))_{1 \leq i, j \leq n}$. Hence, $|\Delta(\vec{\lambda})|^2$ is equal to the det of the following matrix

$$\left(\sum_{k=0}^{n-1} P_k(\lambda_i) P_k(\lambda_j) \right)_{1 \leq i, j \leq n}$$

Then, we rewrite the prob. density for the eigenvalues of GUE as

We ~~will~~ resort to ~~the original scaling with $|E\lambda_i|^2 = 1$~~ $p(\vec{\lambda}) \propto \det \left(\sum_{k=0}^{n-1} P_k(\lambda_i) e^{-\frac{\lambda_i^2}{4}} P_k(\lambda_j) e^{-\frac{\lambda_j^2}{4}} \right)_{1 \leq i, j \leq n}$.

One particular class of polynomials $\{P_i(\lambda)\}$ is of particular importance to us, that is, the "orthogonal polynomials" with respect to the Gaussian measure.

These are the famous "Hermite polynomials". Define the inner product

$$\langle f, g \rangle := \int_{\mathbb{R}} f(x) g(x) e^{-\frac{x^2}{2}} dx.$$

Let $P_0(x) = 1$. Then we define $P_i(x)$ through the Gram-Schmidt process, i.e., given $P_0(x), P_1(x), \dots, P_{i-1}(x)$ we define $P_i(x)$ (of degree n and has degree n term as ~~as~~) such that $\langle P_n(x), P_i(x) \rangle = 0$ for $0 \leq i \leq n-1$. Furthermore, we can normalize $P_n(x)$ so that $\langle P_n(x), P_n(x) \rangle = 1$. Then we get a sequence of orthonormal polynomials with respect to the Gaussian measure: $\langle P_i, P_j \rangle = \delta_{ij}$.

Then, define the kernel

$$K_n(x, y) := \sum_{k=0}^{n-1} P_k(x) e^{-\frac{x^2}{4}} P_k(y) e^{-\frac{y^2}{4}}.$$

This kernel defines an orthogonal projection Π_{V_n} in $L^2(\mathbb{R})$ to the span of $\{1, x, \dots, x^n\}$. In addition, we have

$$p_n(\vec{\lambda}) \propto \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq n}.$$

Lemma 3.2. For any $k \geq 0$, we have

$$\int_{\mathbb{R}} \det [K_n(\lambda_i, \lambda_j)]_{1 \leq i, j \leq k+1} d\lambda_{k+1} = (n-k) \det (K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k}.$$

~~Rmk:~~ Here, we have neglected the ordering of eigenvalues.

Pf: We observe that

$$\textcircled{1} \quad \int_{\mathbb{R}} K_n(x, x) dx = \sum_{k=0}^{n-1} \int_{\mathbb{R}} P_k(x)^2 e^{-\frac{x^2}{2}} dx = n.$$

$$\textcircled{2} \quad \int_{\mathbb{R}} K_n(x, z) K_n(z, y) dz = \int_{\mathbb{R}} \sum_{k, k'=0}^{n-1} P_k(x) e^{-\frac{x^2}{4}} [P_k(z) e^{-\frac{z^2}{4}} P_{k'}(z) e^{-\frac{z^2}{4}}] P_{k'}(y) e^{-\frac{y^2}{4}} dz \\ = \sum_{k=0}^{n-1} P_k(x) e^{-\frac{x^2}{4}} P_k(y) e^{-\frac{y^2}{4}} = K_n(x, y).$$

When $k=0$, ① completes the proof. Now suppose the result has been proved for $k \geq 0$. Then, we consider the $k+1$ case. We apply the cofactor expansion to the $(k+1)$ -th row of $\det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k+1}$. Then,

$$\begin{aligned} \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k+1} &= \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k} K_n(\lambda_{k+1}, \lambda_{k+1}) \\ &\quad + \sum_{l=0}^k (-1)^{k+1-l} K_n(\lambda_l, \lambda_{k+1}) \det(K_n(\lambda_i, \lambda_j))_{1 \leq i \leq k; 1 \leq j \leq k+1, j \neq l} \\ &\int \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k} K_n(\lambda_{k+1}, \lambda_{k+1}) d\lambda_{k+1} = n \det(\lambda_i, \lambda_j)_{1 \leq i, j \leq k}. \end{aligned}$$

For the sum, we use ② to get that

$$\begin{aligned} &\int K_n(\lambda_l, \lambda_{k+1}) \det(K_n(\lambda_i, \lambda_j))_{1 \leq i \leq k; 1 \leq j \leq k+1, j \neq l} d\lambda_{k+1} \\ &= \det((K_n(\lambda_i, \lambda_j))_{1 \leq i \leq k; 1 \leq j \leq k, j \neq l}, (K_n(\lambda_i, \lambda_l))_{1 \leq i \leq k}) \end{aligned}$$

After row exchange, we get $(-1)^{k-l} \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k}$. \square

Iterating the formula, we get $\int_{\mathbb{R}^n} \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq n} d\lambda_1 \cdots d\lambda_n = n!$.

On the other hand, neglecting the ordering eigenvalues, if we extend $p_n(\lambda)$ from the simplex $\{\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n\}$ to \mathbb{R}^n , its integral is $n!$. Since $\det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq n}$ is symmetric in $\lambda_1, \dots, \lambda_n$, we thus get the Gaudin-Mehta formula

$$p_n(\lambda_1, \dots, \lambda_n) = \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq n} \text{ on } \{\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n\}.$$

Moreover, the above lemma 3.2 shows that $p_k(\lambda_1, \dots, \lambda_k) = \det(K_n(\lambda_i, \lambda_j))_{1 \leq i, j \leq k}$ is the k -point correlation function:

$$\begin{aligned} &\int_{\mathbb{R}^k} p_k(\lambda_1, \dots, \lambda_k) F(\lambda_1, \dots, \lambda_k) d\lambda_1 \cdots d\lambda_k \\ &= \mathbb{E} \sum_{\substack{1 \leq i_1, \dots, i_k \leq n \\ \text{distinct}}} F(\lambda_{i_1}, \dots, \lambda_{i_k}), \quad \text{supported on } \{x_1 \leq x_2 \leq \dots \leq x_k\} \end{aligned}$$

When $k=1$, we get $\mathbb{E} \mu_{H_n}^{(dx)} = \frac{1}{n} K_n(x, x) dx$, on

$$\mathbb{E} \mu_{\sqrt{n} H_n} = \frac{1}{\sqrt{n}} K_n(\sqrt{n}x, \sqrt{n}x) dx.$$

When $k=2$, we get $p_2(\lambda_1, \lambda_2) = \det \begin{pmatrix} K(\lambda_1, \lambda_1) & K(\lambda_1, \lambda_2) \\ K(\lambda_2, \lambda_1) & K(\lambda_2, \lambda_2) \end{pmatrix}$.

Recall that $P_i(x)$ is orthogonal to x^j for $0 \leq j < i$. This implies that $xP_i(x)$ is orthogonal to x^j for $0 \leq j < i-1$. On the other hand, $xP_i(x)$ has degree $i+1$, so $xP_i(x)$ must lie in the span of $\{P_{i-1}, P_i, P_{i+1}\}$. This gives that

$$xP_i = \cancel{aP_{i-1}} + bP_i + cP_{i+1}, \text{ where } c \neq 0.$$

Rearrange, we can write that

$$P_{i+1}(x) = (a_i x + b_i)P_i - c_i P_{i-1}, \quad a_i \neq 0.$$

Note $\int_{\mathbb{R}} xP_i(x)P_{i+1}(x)e^{-\frac{x^2}{2}} dx = \frac{1}{a_i}, \quad \int_{\mathbb{R}} xP_i(x)P_{i-1}(x)dx = \frac{c_i}{a_i}.$

$$\Rightarrow \frac{1}{a_{i-1}} = \frac{c_i}{a_i} \Rightarrow c_i = \frac{a_i}{a_{i-1}}, \text{ with } a_{-1} = \infty.$$

Consider $P_{i+1}(x)P_i(y) - P_i(x)P_{i+1}(y) = a_i(x-y)P_i(x)P_i(y) - \frac{a_i}{a_{i-1}}(P_{i-1}(x)P_i(y) - P_{i-1}(y)P_i(x)).$

$$\Rightarrow P_i(x)P_i(y) = \frac{P_{i+1}(x)P_i(y) - P_i(x)P_{i+1}(y)}{a_i(x-y)} - \frac{P_i(x)P_{i-1}(y) - P_{i-1}(x)P_i(y)}{a_{i-1}(x-y)}.$$

Summing over them gives the Christoffel-Darboux formula:

$$K_n(x,y) = \sum_{i=0}^{n-1} P_i(x)P_i(y) e^{-\frac{x^2+y^2}{4}} = \frac{P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)}{a_{n-1}(x-y)} e^{-\frac{x^2+y^2}{4}}.$$

To understand the asymptotic behavior of $K_n(x,y)$ as $n \rightarrow \infty$, need to understand the behavior $P_n(x)$ as $n \rightarrow \infty$.

Note that the above discussion is very general and not restricted to Hermite polynomials.

It only requires a system of orthogonal polynomials.

Hermite polynomials are defined as

$$H_k(x) := (-1)^k e^{\frac{x^2}{2}} \frac{d^k}{dx^k} e^{-\frac{x^2}{2}} \quad \text{with leading coefficient 1.}$$

Hermite functions are defined as

$$\Psi_k(x) = \frac{H_k(x)}{(2\pi)^{1/4} \sqrt{k!}} e^{-\frac{x^2}{4}}, \quad \int \Psi_k(x) \Psi_l(x) dx = \delta_{kl}.$$

Hermite functions are eigenfunctions

of the Hamiltonian for the Harmonic oscillator in quantum mechanics. They have the celebrated asymptotics:

$$\Psi_{2m}(x) = \frac{(-1)^m}{n^{1/4} \sqrt{\pi}} \cos(\sqrt{n}x) + \tilde{o}(n^{-\frac{1}{4}}),$$

$$\Psi_{2m+1}(x) = \frac{(-1)^m}{n^{1/4} \sqrt{\pi}} \sin(\sqrt{n}x) + \tilde{o}(n^{-\frac{1}{4}}), \quad \text{as } n \rightarrow \infty \text{ and } |2m-n|=O(1).$$

They are uniform over $|x| \leq Cn^{-\frac{1}{2}}$. In addition, $a_{n-1} \sim \frac{1}{\sqrt{n}}$.

$$\text{Hence, we get } K_n(x, y) \approx \sqrt{n} \frac{\psi_n(x)\psi_{n-1}(y) - \psi_{n-1}(x)\psi_n(y)}{x-y}$$

$$\approx \frac{1}{\pi} \frac{\sin(\sqrt{n}x)\cos(\sqrt{n}y) - \sin(\sqrt{n}y)\cos(\sqrt{n}x)}{x-y}$$

$$= \frac{1}{\pi} \frac{\sin(\sqrt{n}(x-y))}{x-y}$$

We consider the scaling $\frac{y}{\sqrt{n}p_{sc}(0)}$, $p_{sc}(0) = \frac{1}{\pi}$. Then

$$K_n\left(\frac{y_1}{\sqrt{n}p_{sc}(0)}, \frac{y_2}{\sqrt{n}p_{sc}(0)}\right) \approx \sqrt{n}p_{sc}(0) \frac{\sin[\pi(y_1 - y_2)]}{\pi(y_1 - y_2)}.$$

This gives the Dyson sine kernel around 0. In general, we have

$$\frac{1}{\sqrt{n}p_{sc}(x)} K_n\left(\sqrt{n}x + \frac{y_1}{\sqrt{n}p_{sc}(x)}, \sqrt{n}x + \frac{y_2}{\sqrt{n}p_{sc}(x)}\right) \rightarrow S(y_1, -y_2) := \frac{\sin(\pi(y_1 - y_2))}{\pi(y_1 - y_2)} \text{ for } x \in [-2, 2].$$

We resort to the scaling $H_n \rightarrow \frac{1}{\sqrt{n}}H_n$, where the typical eigengap is of order $\frac{1}{n}$. Then,

$$\frac{1}{[p_{sc}(E)]^n} P_k\left(E + \frac{x_1}{np_{sc}(E)}, \dots, E + \frac{x_k}{np_{sc}(E)}\right) \rightarrow \det(S(x_i - x_j))_{1 \leq i, j \leq n}.$$

This means, for any $E \in [-2, 2]$ and bounded F ,

$$\begin{aligned} & \mathbb{E} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} F(n p_{sc}(E)(\lambda_{i_1} - E), \dots, n p_{sc}(E)(\lambda_{i_k} - E)) \\ & \rightarrow \int_{y_1 \leq y_2 \leq \dots \leq y_k} F(y_1, y_2, \dots, y_k) \det(S(y_i - y_j))_{1 \leq i, j \leq k} dy_1 \dots dy_k. \end{aligned}$$

In particular, using the 2-point correlation function, one can derive that "Wigner surmise": given a bulk energy E , let i s.t. $i = N(\int_{-2}^E p_{sc}(x) dx + o(1))$. Then

$N p_{sc}(E)(\lambda_{i+1} - \lambda_i)$ converges in law to $p(x) = \sqrt{\frac{32x^2}{\pi^2}} e^{-4x^2/\pi}$ in the Hermitian case.

Later, we will show the "bulk universality", i.e., the Dyson sine kernel and the Wigner surmise also occur for general Wigner matrices.

At edge, a different scaling has to be used. Using the definition of Hermite functions, we can check the following identity:

$$\psi'_N(x) = -\frac{x}{2} \psi_N(x) + \sqrt{N} \psi_{N-1}(x).$$

With this identity, we can check that

$$K_N(x, y) = \sqrt{N} \frac{\psi_N(x)\psi_{N-1}(y) - \psi_N(y)\psi_{N-1}(x)}{x-y} = \frac{\psi_N(x)\psi'_N(y) - \psi_N(y)\psi'_N(x)}{x-y} - \frac{1}{2} \psi_N(x)\psi_N(y).$$

The Plancherel-Rotach edge asymptotics for ψ_N gives that

$N^{\frac{1}{12}} \gamma_N (2\sqrt{N} + \frac{u}{N^{1/6}}) \rightarrow A(u)$, where $A(u)$ is the Airy function:

$$A(u) = \frac{1}{\pi} \int_0^\infty \cos(\frac{1}{3}t^3 + ut) dt$$

(It is called Airy function because it is a solution to the Airy equation $y'' - xy = 0$.)

Then, we see that

$$\frac{1}{N^{1/6}} K_N (2\sqrt{N} + \frac{u}{N^{1/6}}, 2\sqrt{N} + \frac{v}{N^{1/6}}) \rightarrow A(u, v),$$

$$A(u, v) := \frac{A(u) A'(v) - A'(u) A(v)}{u-v} \text{ the Airy kernel.}$$

Resort to the scaling $H_N \rightarrow \frac{1}{\sqrt{N}} H_N$, we have

$$N^{\frac{k}{3}} P_k \left(2 + \frac{d_1}{N^{2/3}}, 2 + \frac{d_2}{N^{2/3}}, \dots, 2 + \frac{d_k}{N^{2/3}} \right) \rightarrow \det(A(d_i, d_j))_{1 \leq i, j \leq k}.$$

When $k=2$, we have

$$\sum_{j \neq k} \mathbb{E} F(N^{2/3}(\lambda_j - 2), N^{2/3}(\lambda_k - 2)) \rightarrow \int_{\mathbb{R}^2} da_1 da_2 F(a_1, a_2) \det(A(d_i, d_j))_{1 \leq i, j \leq 2}.$$

A much harder problem is to derive the limiting distribution of $N^{2/3}(\lambda_N - 2)$.

Tracy and Widom (1993, 1994) show that it converges to a limiting distribution, which is now referred to as the Tracy-Widom law: $F_\beta(x)$, $\beta = 1, 2, 4$.

In particular,

$$F_2(s) = \exp\left(-\int_s^\infty (x-s) q^2(x) dx\right), \text{ where } q(x) \text{ is the solution to Painlevé equation II: } q''(s) = sq(s) + 2q(s)^3, \\ q(s) \sim A(s) \text{ as } s \rightarrow \infty.$$

"Edge universality" also holds, i.e., $N^{2/3}(\lambda_N - 2)$ also converges to the TW law as $N \rightarrow \infty$ for general Wigner matrices.

Section 4 Rigidity of eigenvalues

Thm 4.1 (Local semicircle law) Let H_N be a real Wigner matrix with independent entries (up to $H = H^*$) with mean 0 and variance $\frac{1}{N}$. Its entries have finite moments up to any order, i.e., $\forall k \in \mathbb{N}, \exists C_k > 0$ so that

$$\max_{i,j} |E| \sqrt{N} |h_{ij}|^k \leq C_k.$$

Then, uniformly in $|E| \leq \frac{\varepsilon^{-1}}{N}$ and $\eta \geq N^{-1+\varepsilon}$ for a small constant $\varepsilon > 0$, we have:

(1) (Entrywise local law)

$$\max_{i,j} |G_{ij}(z) - \delta_{ij} m_{sc}(z)| \prec \gamma_F(z), \quad \gamma_F(z) := \sqrt{\frac{\text{Im } m_{sc}(z)}{N\eta}} + \frac{1}{N\eta}.$$

(2) (Averaged local law) $|m_N(z) - m_{sc}(z)| \prec \frac{1}{N\eta}, \quad m_N = \frac{1}{N} \text{Tr } G(z).$

Outside of the spectrum (i.e., $[-2, 2]$), we have the stronger averaged local law:

(3) (Averaged local law)

$$|m_N(z) - m_{sc}(z)| \prec \frac{1}{N(K_E + \eta)} + \frac{1}{(N\eta)^2 \sqrt{K_E + \eta}}$$

uniformly in $\{z : 20 \leq |E| \leq \frac{\varepsilon^{-1}}{N}, N^{-1+\varepsilon} \leq \eta \leq 10, N\eta \sqrt{K_E + \eta} \geq N^\varepsilon\}$ for any small const $\varepsilon > 0$. Here K_E is defined as the distance to the edges: $K_E := \min\{|E-2|, |E+2|\}$.

We will prove this theorem later. Before that, we consider some of its applications.

In this section, we use it to establish the "rigidity of eigenvalues".

Thm 4.2: Given the local laws, we have that

$$|\lambda_j - \sigma_j| \propto (j, N-j+1)^{-\frac{1}{3}} N^{-\frac{2}{3}} \quad \forall 1 \leq j \leq N,$$

where σ_j is defined as

$$\int_{-2}^{\sigma_j} p_{sc}(x) dx = \frac{j}{N}. \quad (\text{Recall that } \int_{-\infty}^{\sigma_j} d\mu_{H_N}(x) = \frac{j}{N}.)$$

Rmk: Note that this is much stronger than the semicircle law we proved before. Previously, we have only shown that the Kolmogorov distance between μ_{H_N} and μ_{sc} is $O(1)$ with prob. $1-O(1)$.

Thm 4.2 indeed implies that

$$\sup_x \left| \int_{-\infty}^x p_{sc}(t) dt - \int_{-\infty}^x d\mu_{H_N}(x) \right| \prec \frac{1}{N}.$$

This indicates a strong correlation between eigenvalues. For independent point process, the typical order of fluctuation should be $N^{-\frac{1}{2}}$.

- ① We first show an upper/lower bound on the largest/smallest eigenvalue.

Prop: Given the local law, for any constants $\varepsilon, D > 0$, we have that

$$\mathbb{P}(\lambda_N > 2 + N^{-\frac{2}{3}} + \varepsilon) \leq N^{-D} \text{ for large enough } N.$$

Pf: Let $\eta = N^{-\frac{2}{3}}$ and choose $E = 2 + K$ for some $K \geq N^{-\frac{2}{3}} + \varepsilon \gg N^{\frac{1}{2}}\eta$. From the local law, $|m_N(z) - m_{sc}(z)| \propto \frac{1}{N(K_E + \eta)} + \frac{1}{(N\eta)^2 \sqrt{K_E + \eta}}$, we get

$$|\operatorname{Im} m_N(z) - \operatorname{Im} m_{sc}(z)| \leq \frac{N^{\frac{1}{2}\varepsilon}}{NK} \ll \frac{1}{N\eta} \text{ w.h.p.}$$

On the other hand, if there is an eigenvalue λ with $|\lambda - E| \leq \eta$, we have

$$\operatorname{Im} m_N(z) = \frac{1}{N} \sum_K \frac{\eta}{(K-E)^2 + \eta^2} \geq \frac{1}{2N\eta}.$$

For $m_{sc}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}$, we can check that $\operatorname{Im} m_{sc}(z) \approx \frac{\eta}{\sqrt{K+\eta}} \ll \frac{1}{N\eta}$. (see the lemma below)

Hence, $\operatorname{Im} m_N(z) \leq \operatorname{Im} m_{sc}(z) + |\operatorname{Im} m_N(z) - \operatorname{Im} m_{sc}(z)| \ll \frac{1}{N\eta}$ w.h.p.

We get a contradiction. \square

Lemma (Properties of $m_{sc}(z)$) $m_{sc}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}$ is the unique solution to the self-consistent equation $m(z) + \frac{1}{m(z)} + z = 0$ with $\operatorname{Im} m(z) > 0$ for $\operatorname{Im} z > 0$. Furthermore, we have that:

$$(i) |m_{sc}(z)| = |m_{sc}(z) + z|^{-1} \leq 1.$$

(ii) For any constant $C > 0$, there exists a constant $\varepsilon > 0$ such that for $E \in [-C, C]$ and $\eta \in [0, \varepsilon]$, we have

$$\varepsilon \leq |m_{sc}(z)| \leq 1 - \varepsilon.$$

Furthermore, $|m_{sc}(E)| = 1$ for $E \in [-2, 2]$.

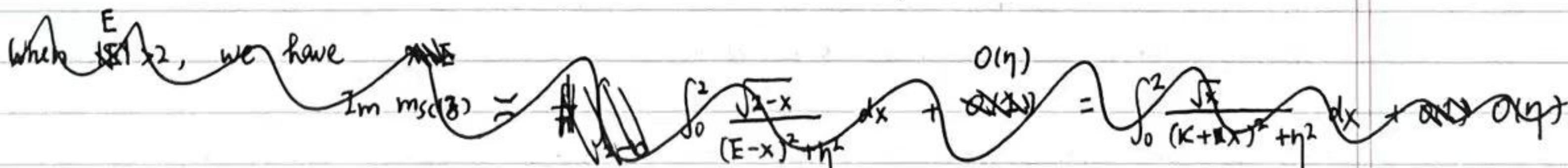
$$(iii) |1 - m_{sc}^2(z)| \approx \sqrt{K_E + \eta}, \quad K_E := |E| - 2.$$

(iv)

$$\operatorname{Im} m_{sc}(z) \approx \begin{cases} \sqrt{K_E + \eta} & \text{if } |E| \leq 2 \\ \frac{\eta}{\sqrt{K_E + \eta}} & \text{if } |E| > 2 \end{cases}.$$

Proof involves a direct calculation. The estimate (iv) is more general, and, unlike other estimates, does not pertain to the semicircle law. For $\delta = E + i\eta$,

$$\operatorname{Im} m_{sc}(z) = \operatorname{Im} \int_{-2}^2 \frac{f_{sc}(x) dx}{x - (E + i\eta)} = \int_{-2}^2 \frac{\eta}{(x - E)^2 + \eta^2} f_{sc}(x) dx.$$



* When $E \geq 2$, we have $\operatorname{Im} m_{sc}(z) \sim \eta \int_{2-C}^2 \frac{\sqrt{2-x}}{(E-x)^2 + \eta^2} dx + O(\eta)$

$$= \eta \int_0^C \frac{\sqrt{x}}{(x+\kappa)^2 + \eta^2} dx + O(\eta),$$

where $\int_0^C \frac{\sqrt{x}}{(x+\kappa)^2 + \eta^2} dx \sim \int_0^C \frac{\sqrt{x}}{(x+\kappa+\eta)^2} dx \sim \int_0^{\frac{C}{\kappa+\eta}} \frac{1}{\sqrt{\kappa+\eta}} \frac{\sqrt{y}}{y^2+1} dy \sim \frac{1}{\sqrt{\kappa+\eta}}$.

* When $E \in [0, 2]$, we have $\operatorname{Im} m_{sc}(z) \sim \underbrace{\eta \int_{E-\kappa}^2 \frac{\sqrt{2-x}}{(E-x)^2 + \eta^2} dx}_{I} + \underbrace{\eta \int_{-2}^{E-\kappa} \frac{\sqrt{2-x}}{(E-x)^2 + \eta^2} dx}_{II}$.

For the second term, $II = \eta \int_{(1+C)\kappa}^4 \frac{\sqrt{y}}{(y-\kappa)^2 + \eta^2} dy \sim \eta \int_{C\kappa}^{4-\kappa} \frac{\sqrt{y}}{y^2 + \eta^2} dy \sim \sqrt{\eta} \int_{C\kappa/\eta}^{(4-\kappa)/\eta} \frac{\sqrt{y}}{y^2 + 1} dy \sim \sqrt{\eta} \wedge \frac{\eta}{\sqrt{\kappa}}$.

For the first term, $I = \eta \int_0^{(1+C)\kappa} \frac{\sqrt{y}}{(y-\kappa)^2 + \eta^2} dy = \eta \int_{-\kappa}^{C\kappa} \frac{\sqrt{y+\kappa}}{y^2 + \eta^2} dy = \sqrt{\eta} \int_{-\kappa/\eta}^{C\kappa/\eta} \frac{\sqrt{y+\kappa/\eta}}{y^2 + 1} dy$.

Note that $I \sim \sqrt{\eta}$ if $\kappa \leq \eta$.

On the other hand, if $\kappa \geq \eta$, we have $I = \sqrt{\eta} \int_{-\kappa/\eta}^{C\kappa/\eta} \frac{\sqrt{y+\kappa/\eta}}{y^2 + 1} dy + O\left(\frac{\eta}{\sqrt{\kappa}}\right) \sim \sqrt{\kappa} \int_{-\kappa/\eta}^{C\kappa/\eta} \frac{1}{y^2 + 1} dy + O\left(\frac{\eta}{\sqrt{\kappa}}\right) \sim \sqrt{\kappa}$.

In sum, we get $I \sim \sqrt{\kappa+\eta}$.

② The Helffer - Sjöstrand formula

The H-S formula provides a way to control the distance between two measures through bounding the difference between their Stieltjes transforms.

Lemma 4.3: Let $f \in C_c^1(\mathbb{R})$, and $\chi \in C_c^\infty(\mathbb{R})$ be a smooth cutoff function with support in $[-1, 1]$ and with $\chi(y) \equiv 1$ for $|y| \leq 1/2$. Then, we have that

$$f(x) = \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{iy f''(x) \chi(y) + i(f(x) + iy f'(x)) \chi'(y)}{\lambda - (x+iy)} dx dy.$$

Rmk: The H-S formula can be regarded as an extension of the Cauchy integral formula:

$$f(x) = \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{z - x} dz, \quad C_\lambda \text{ encircles } x.$$

Pf: Define $\tilde{f}(x+iy) := (f(x) + iy f'(x)) \chi(y)$.

For $\partial_{\bar{z}} = (\partial_x + i\partial_y)/2$, we claim that

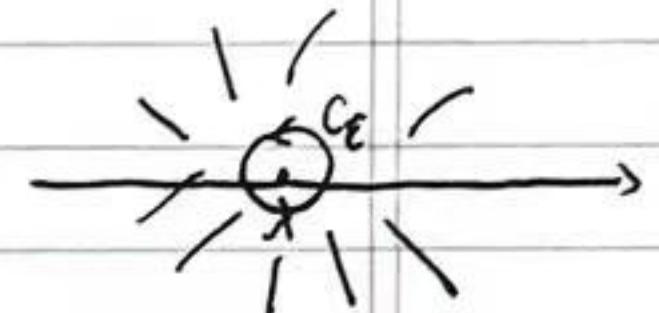
$$(*) \quad f(\lambda) = \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{\partial_{\bar{z}} \tilde{f}(x+iy)}{\lambda - (x+iy)} dx dy.$$

Expanding $\partial_{\bar{z}} \tilde{f}(x+iy)$, we conclude the proof. The proof of (*) is a simple application of the Green's theorem. With $\partial_{\bar{z}}(\lambda - z)^{-1} = 0$, we can write that

$$\begin{aligned} \frac{1}{\pi} \int \frac{\partial_{\bar{z}}(\tilde{f}(x+iy))}{\lambda - (x+iy)} dx dy &= \frac{1}{\pi} \int \partial_{\bar{z}} g(x+iy) dx dy, \text{ where } g(x+iy) = \frac{\tilde{f}(x+iy)}{\lambda - (x+iy)}. \\ &= \frac{1}{2\pi} \int [\partial_x g(x+iy) + i\partial_y g(x+iy)] dx dy. \end{aligned}$$

From the compact support condition of g and Green's theorem, we get

$$\frac{1}{2\pi} \int_{|x+iy-\lambda| \geq \varepsilon} [\partial_x g(x+iy) + i\partial_y g(x+iy)] dx dy$$



$$= \frac{1}{2\pi} \oint_{C_\varepsilon} [ig(x,y) dx - g(x,y) dy] dz = \frac{1}{2\pi i} \oint_{C_\varepsilon} \frac{\tilde{f}(x+iy)}{x+iy-\lambda} dz$$

$$\rightarrow \tilde{f}(\lambda) = f(\lambda) \text{ as } \varepsilon \rightarrow 0.$$

Further, the integral $\int_{|x+iy-\lambda| \leq \varepsilon} [\partial_x g(x+iy) + i\partial_y g(x+iy)] dx dy \rightarrow 0$ as $\varepsilon \rightarrow 0$. This gives (*). \square

Remark: The formula (*) also holds for the choice $\tilde{f}(x+iy) = f(x)g(y)$. But, we need to add $i y f'(x) g(y)$ to make $\partial_{\bar{z}} \tilde{f}(x+iy) = O(|y|)$ for small $|y|$, which will be used in the following proof.

③ The convergence Speed of the ESD

Define the empirical distribution

$$F_N(E) := \int_{-\infty}^E P_{H_N}(x) dx = \frac{1}{N} \#\{k : \lambda_k \leq E\}.$$

Define the distribution function of the semicircle density

$$F_{SC}(E) := \int_{-\infty}^E P_{SC}(x) dx.$$

Given

Theorem 4.4 Under the averaged local law $|\lambda_N(z) - m_0(z)| \prec \frac{1}{N^\eta}$ uniformly in $|z| \leq \varepsilon^{-1}$ and $|y| \geq N^{-1+\varepsilon}$ for any small constant $\varepsilon > 0$, we have that

$$\sup_{E \in \mathbb{R}} |F_N(E) - F_{SC}(E)| \prec \frac{1}{N}.$$

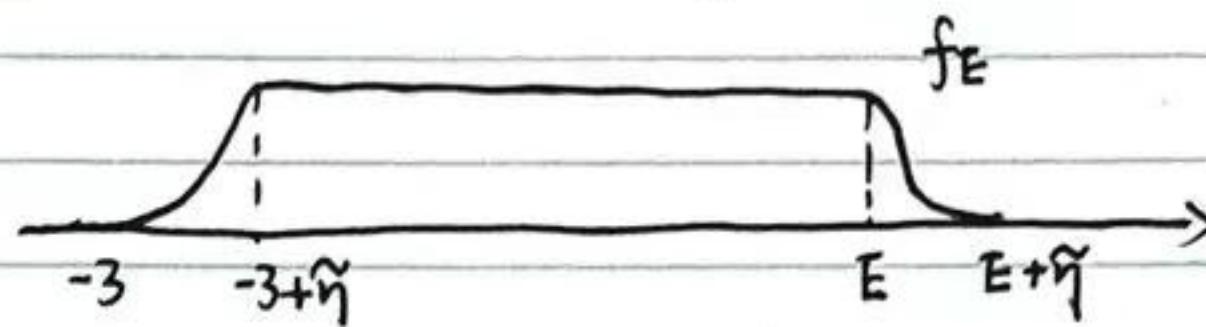
Pf: We already know that H_N has no eigenvalue outside $[-2 - N^{-\frac{2}{3} + \varepsilon}, 2 + N^{-\frac{2}{3} + \varepsilon}]$. So outside this interval, $F_N(E) - F_{SC}(E) = 0$ w.h.p.

We only need to control $F_N(E) - F_{SC}(E)$ for $E \in [-2 - N^{-\frac{2}{3} + \varepsilon}, 2 + N^{-\frac{2}{3} + \varepsilon}]$.

We pick $\tilde{\eta} = N^{-1+\varepsilon}$ for a small constant $\varepsilon > 0$. We know that

$$|m_N(z) - m_{SC}(z)| = \left| \int \frac{dF_N(x)}{x-z} - \int \frac{dF_{SC}(x)}{x-z} \right| \leq \frac{1}{N\tilde{\eta}} \quad \text{for } E \in (-3, 3) \text{ and } z \geq \tilde{\eta}.$$

Given $E \in [-2 - N^{-\frac{2}{3} + \varepsilon}, 2 + N^{-\frac{2}{3} + \varepsilon}]$, we define a smooth cutoff function $f_E(x)$ such that $\text{supp}(f_E) \subseteq [-3, E + \tilde{\eta}]$, $f_E^{(0)} = 1$ for $x \in [-3 + \tilde{\eta}, E]$, and $\|f_E'\|_\infty \leq C\tilde{\eta}^{-1}$, $\|f_E''\|_\infty \leq C\tilde{\eta}^{-2}$, and $f_E \uparrow$ on $[-3, -3 + \tilde{\eta}]$, $f_E \downarrow$ on $[E, E + \tilde{\eta}]$.



Note that $\int f_E(x) p_{H_N}(x) dx$ and $\int f_E(x) p_{SC}(x) dx$ approximates $F_N(E)$ and $F_{SC}(E)$.

We now bound

$$\begin{aligned} \int f_E(x) [p_{H_N}(x) - p_{SC}(x)] dx &\stackrel{H-S}{=} \frac{1}{\pi} \int_{\mathbb{R}^2} \int_{\mathbb{R}} \frac{i y f_E''(x) \chi(y) + i(f_E'(x) + iy f_E'(x)) \chi'(y)}{\lambda - (x+iy)} \\ &\quad [p_{H_N}(\lambda) - p_{SC}(\lambda)] d\lambda dx dy \\ &= \frac{1}{\pi} \int_{\mathbb{R}^2} [iy f_E''(x) \chi(y) + i(f_E'(x) + iy f_E'(x)) \chi'(y)] (m_N(z) - m_{SC}(z)) dx dy \\ &= -\frac{1}{\pi} \int_{\mathbb{R}^2} [y f_E''(x) \chi(y) + f_E(x) \chi'(y)] \text{Im}(m_N(z) - m_{SC}(z)) dx dy \\ &\quad - \frac{1}{\pi} \int_{\mathbb{R}^2} y f_E'(x) \chi'(y) \text{Re}(m_N(z) - m_{SC}(z)) dx dy = I + II + III. \end{aligned}$$

Note $\chi'(y) \neq 0$ for $|y| \leq 1$. Furthermore, $f_E'(x)$ is of order $O(\frac{1}{\tilde{\eta}})$ on two intervals of size $\tilde{\eta}$.

For $1/2 \leq |y| \leq 1$, we have $|\text{Re}(m_N(z) - m_{SC}(z))| \leq \frac{1}{N}$. Hence, we have

$$|III| \leq \frac{1}{N} \times \frac{1}{\tilde{\eta}} \times \tilde{\eta} = \frac{1}{N}.$$

Similarly, we have $|II| \leq \frac{1}{N}$. For I, by symmetry, we only need to consider the upper complex plane with $y \geq 0$. We consider two regimes: $0 < y \leq \tilde{\eta}$ and $y > \tilde{\eta}$.

On $0 < y \leq \tilde{\eta}$, we have $|\text{Im } m_{SC}(z)| \leq 1$ and $|\text{Im } m_N(z)| \leq C\tilde{\eta}/y$ w.h.p. Using the inequality

$$y |\text{Im } m_N(E+iy)| \leq \tilde{\eta} |\text{Im } m_N(E+i\tilde{\eta})| \text{ for } y \leq \tilde{\eta}.$$

Furthermore, $f_E''(x)$ is of order $O(\tilde{\eta}^{-2})$ on two intervals of size $\tilde{\eta}$. Hence,

$$\begin{aligned} &\int_{x \in [-3, E + \tilde{\eta}]} \int_{0 < y \leq \tilde{\eta}} |y f_E''(x) \chi(y)| |\text{Im}(m_N(z) - m_{SC}(z))| dx dy \\ &\leq C\tilde{\eta} \int_{x \in [-3, -3 + \tilde{\eta}] \cup [E, E + \tilde{\eta}]} \int_{0 < y \leq \tilde{\eta}} |f_E''(x)| dx dy \leq C\tilde{\eta} \cdot \frac{1}{\tilde{\eta}^2} \cdot \tilde{\eta}^2 = O(\tilde{\eta}). \end{aligned}$$

For the regime $\tilde{\eta} < y \leq 1$, we use integration by parts in x to get

$$\int dx \int_{\tilde{\eta} < y \leq 1} dy y f_E''(x) \chi(y) \operatorname{Im}(m_N(z) - m_{sc}(z)) dx dy$$

$$= - \int dx \int_{\tilde{\eta} < y \leq 1} dy y f_E'(x) \chi(y) \partial_x \operatorname{Im}(m_N(z) - m_{sc}(z)) dx dy$$

Cauchy-Riemann condition

$$= \int dx \int_{\tilde{\eta} < y \leq 1} dy y f_E'(x) \chi(y) \partial_y \operatorname{Im}(m_N(z) - m_{sc}(z)) dx dy$$

$$= - \int dx \tilde{\eta} f_E'(x) \operatorname{Im}(m_N(x+i\tilde{\eta}) - m_{sc}(x+i\tilde{\eta})) dx$$

$$- \int dx \int_{\tilde{\eta} < y \leq 1} dy [x(y) + y \chi'(y)] f_E'(x) \operatorname{Im}(m_N(z) - m_{sc}(z)) dx dy$$

$$= J_1 + J_2 + J_3.$$

Using the averaged local law, $|m_N(x+i\tilde{\eta}) - m_{sc}(x+i\tilde{\eta})| < \frac{1}{N\tilde{\eta}}$ and that $f_E'(x)$ is of order $O(\tilde{\eta}^{-1})$ on two intervals of size $\tilde{\eta}$, we get that

$$|J_1| \leq C \frac{1}{N\tilde{\eta}} \cdot \tilde{\eta} \cdot \frac{1}{\tilde{\eta}} \cdot \tilde{\eta} = O\left(\frac{1}{N}\right).$$

J_3 can be bounded in the same way as III: $|J_3| \leq \frac{1}{N} \times \frac{1}{\tilde{\eta}} \times \tilde{\eta} = \frac{1}{N}$.

Finally, J_2 can be bounded as

$$|J_2| \leq \frac{1}{\tilde{\eta}} \times \tilde{\eta} \times \int_{\tilde{\eta} < y \leq 1} x(y) \frac{1}{Ny} dy \leq \frac{1}{N} \log \frac{1}{\tilde{\eta}} \leq \frac{\log N}{N} \leq \frac{1}{N}.$$

To summarize, we have shown that

$$\int f_E'(x) [\rho_{H_N}(x) - \rho_{sc}(x)] dx \leq \frac{1}{N}.$$

Our goal is to bound $\int 1_{(-\infty, E]}(x) [\rho_{H_N}(x) - \rho_{sc}(x)] dx$. We bound

$$\begin{aligned} & \int |f_E(x) - 1_{(-\infty, E]}(x)| \rho_{H_N}(x) dx \text{ and } \underbrace{\int |f_E(x) - 1_{(-\infty, E]}(x)| \frac{\rho_{sc}(x)}{\rho_{H_N}(x)} dx}_{\text{separately.}} \\ & \leq \int_E^{E+\tilde{\eta}} \rho_{H_N}(x) dx \text{ w.h.p. because} \\ & \quad \text{there is no eigenvalue of } H_N \text{ inside } (-\infty, 3]. \quad \left| \begin{array}{l} \leq C\tilde{\eta} \text{ because } \rho_{sc}(x) = 0 \text{ on } (-\infty, 3] \text{ and,} \\ \rho_{sc}(x) \leq C, |f_E(x) - 1_{(-\infty, E]}(x)| \text{ is non-zero on} \\ \text{two intervals of size } \tilde{\eta}. \end{array} \right. \end{aligned}$$

Moreover, $\int_E^{E+\tilde{\eta}} \rho_{H_N}(x) dx \leq \int [f_{E+\tilde{\eta}}(x) - f_{E-\tilde{\eta}}(x)] \rho_{H_N}(x) dx$

$$\leq \frac{1}{N} + \int [f_{E+\tilde{\eta}}(x) - f_{E-\tilde{\eta}}(x)] \rho_{sc}(x) dx \leq C\tilde{\eta}.$$

In sum, we obtain that $|F_N(E) - F_{sc}(E)| = \left| \int_{(-\infty, E]} 1_{(-\infty, E]}(x) (\rho_N(x) - \rho_{sc}(x)) dx \right| \prec \frac{1}{N}$.
 The above arguments are uniformly in E because the local law is uniform in E . \square

④ Rigidity of eigenvalues (Proof of Theorem 4.1)

Recall that

$$\frac{j}{N} = \int_{-\infty}^{\lambda_j} \rho_N(x) dx, \quad \frac{j}{N} = \int_{-\infty}^{\gamma_j} \rho_{sc}(x) dx. \quad (*)$$

We already know that $\int_{-\infty}^E [\rho_N(x) - \rho_{sc}(x)] dx$ is small. We want to turn this estimate into bounds of $|\lambda_j - \gamma_j|$.

By symmetry, we consider $j \leq N/2$. We aim to show that

$$|\lambda_j - \gamma_j| \prec j^{-1/3} N^{-2/3} \text{ uniformly in } 1 \leq j \leq \frac{N}{2}.$$

From (*), we get $0 = \int_{-\infty}^{\lambda_j} \rho_N(x) dx - \int_{-\infty}^{\gamma_j} \rho_{sc}(x) dx = \int_{-\infty}^{\lambda_j} (\rho_N(x) - \rho_{sc}(x)) dx - \int_{\lambda_j}^{\gamma_j} \rho_{sc}(x) dx$.

Hence, $\left| \int_{\lambda_j}^{\gamma_j} \rho_{sc}(x) dx \right| \leq |F_N(\lambda_j) - F_{sc}(\lambda_j)| \prec \frac{1}{N}$.

Note that $\rho_{sc}(x) \sim \sqrt{(x+2)(2-x)} \sim \sqrt{k_x}$ for $x \in [-2, 1]$. Then, we can derive that $F_{sc}(x) \sim k_x^{3/2}$ and $\rho_{sc}(x) \sim \sqrt{k_x} \sim F_{sc}(x)^{1/3}$, $x \in [-2, 1]$. This implies that,

$$\gamma_j + 2 \sim \left(\frac{j}{N}\right)^{2/3}, \quad \rho_{sc}(\gamma_j) \sim \left(\frac{j}{N}\right)^{1/3}, \quad 1 \leq j \leq \frac{N}{2}.$$

If we know that $\rho_{sc}(\gamma_j)$ and $f_{sc}(\lambda_j)$ are comparable, then we have

$$\left[\frac{1}{N} \succ \left| \int_{\lambda_j}^{\gamma_j} \rho_{sc}(x) dx \right| \gtrsim \rho_{sc}(\gamma_j) |\gamma_j - \lambda_j| \Rightarrow |\gamma_j - \lambda_j| \prec \frac{1}{N} \frac{1}{\rho_{sc}(\gamma_j)} = j^{-\frac{1}{3}} N^{-\frac{2}{3}}. \right]$$

To give a rigorous proof, we first consider $j \geq N^{\frac{\varepsilon}{2}}$. Since $F_{sc}(\gamma_j) = \frac{j}{N} \geq N^{-1+\frac{\varepsilon}{2}}$ and $F_{sc}(\gamma_j) = F_N(\lambda_j)$, we have

$$\left| F_{sc}(\lambda_j) - F_{sc}(\gamma_j) \right| \prec \frac{1}{N} \leq N^{-\frac{\varepsilon}{2}} F_{sc}(\gamma_j).$$

Hence, $|F_{sc}(\lambda_j) - F_{sc}(\gamma_j)| \leq N^{-\frac{\varepsilon}{2}} F_{sc}(\gamma_j)$ w.h.p. This shows that $F_{sc}(\lambda_j) \sim F_{sc}(\gamma_j)$, implying that $\rho_{sc}(\lambda_j) \sim \rho_{sc}(\gamma_j)$. This completes the proof.

For indices $j \leq N^{\frac{\varepsilon}{2}}$, we use the trivial bound: $\lambda_1 \leq \lambda_j \leq \lambda_{N^{\frac{\varepsilon}{2}}}$.

But we already know that $\lambda_1 \geq -2 - N^{-2/3+\varepsilon/2}$ w.h.p. and $\lambda_{N^{\frac{\varepsilon}{2}}} \leq -2 + N^{\frac{\varepsilon}{6}} \left(\frac{N^{\varepsilon/2}}{N}\right)^{2/3}$

In addition, we have $|\gamma_j + 2| \leq C \left(\frac{N^{\varepsilon/2}}{N}\right)^{2/3} = C N^{-\frac{2}{3} + \frac{\varepsilon}{3}}$. $= -2 + N^{\frac{\varepsilon}{2}}$ w.h.p.
 $N^{-\frac{2}{3} + \varepsilon/2}$

Hence, $|\lambda_j - \gamma_j| \leq C N^{-\frac{2}{3} + \frac{\varepsilon}{2}} \leq N^{-\frac{2}{3} + \varepsilon}$ w.h.p.

Section 5 Some applications of local semicircle law

① Delocalization of eigenvectors

Thm 5.1 (Delocalization of eigenvectors) Under the setting of local semicircle law, we have $\max_{1 \leq i \leq N} \|u_i\|_\infty \propto N^{-1/2}$.

Pf.: We have the entrywise local law $\max_{i,j} |G_{ij}(z) - m_{sc}(z) \delta_{ij}| \prec \underline{\gamma}(z)$ uniformly in $z \in [-3, 3]$ and $N^{-1+\varepsilon} \leq \eta \leq 1$, where $\underline{\gamma}(z) := \sqrt{\frac{\text{Im } m_{sc}(z)}{N\eta}} + \frac{1}{N\eta} \ll 1$.

We have

$$\text{Im } G_{ii}(z) = \left\| \sum_{k=1}^N \frac{u_k(i)}{(\lambda_k - z)^2 + \eta^2} \right\|^2, \quad z = E + i\eta.$$

Taking $E = \lambda_k$ and $\eta = N^{-1+\varepsilon}$, by rigidity of eigenvalues, we know that $|\lambda_k| \leq 3$ w.h.p. Hence, $|G_{ii}(z)| \leq |\text{Im } m_{sc}(z)| + |\text{Im } G_{ii}(z) - \text{Im } m_{sc}(z)| \leq 2$ w.h.p. Hence, w.h.p.,

$$2 \geq \text{Im } G_{ii}(\lambda_k + iN^{-1+\varepsilon}) \geq \left\| \frac{\eta u_k(i)}{\eta^2} \right\|^2 = \frac{|u_k(i)|^2}{\eta} \Rightarrow |u_k(i)|^2 \leq 2\eta \text{ w.h.p.}$$

Since E is arbitrary, we get

$$\|u_k\|_\infty \propto N^{-1/2} \text{ by taking a union bound over } i. \quad \square$$

② Deformed Wigner matrix

Consider a rank-1 deformation of H : $\tilde{H} = H + d\vec{v}\vec{v}^*$, where \vec{v} is a unit vector. Without loss of generality, we take $d \geq 0$. We want to study the limiting behavior of the largest eigenvalue and its eigenvector. In particular, we want to know when there will be an outlier outside the semicircle support $[-2, 2]$.

Theorem 5.2 Under the setting of local semicircle law, we have that:

(i) If $d > 1$, then

$$|\lambda_N - \theta(d)| \prec N^{-\frac{1}{2}}, \quad \text{where } \theta(d) := d + \frac{1}{d}.$$

(ii) If $d < 1$, then

$$|\lambda_N - 2| \prec N^{-\frac{2}{3}}.$$

Pf.: Consider the eigenvalue equation $\det(\tilde{H} - \lambda I) = 0$

$$\Rightarrow \det(H - \lambda + d\vec{v}\vec{v}^*) = \det(G(\lambda)) \det(1 + d\vec{v}\vec{v}^* G(\lambda))$$

$$= \det(G(\lambda)) \det(dI) \det(d^{-1} + \vec{v}^* G(\lambda) \vec{v}).$$

[We used $\det(1+AB) = \det(1+BA)$.]

Hence, if λ is not an eigenvalue of H , we must have

$$d^{-1} + \vec{v}^* G(\lambda) \vec{v} = 0.$$

To solve this equation, we need an estimate on $\vec{v}^* G(\lambda) \vec{v}$ for $\lambda \notin [-2, 2]$. This is provided by the following isotropic local law.

Theorem 5.3 (Isotropic local semicircle law) For any deterministic unit vectors $\vec{u}, \vec{v} \in \mathbb{R}^N$, we have that

$$|\vec{u}^* G(\beta) \vec{v} - m_{sc}(\beta) \vec{u}^* \vec{v}| \propto \sqrt{\beta}$$

uniformly in $|\beta| \leq \varepsilon^{-1}$ and $\eta \geq N^{-1+\varepsilon}$. In addition, outside $[-2, 2]$, we have

that $|\vec{u}^* G(\beta) \vec{v} - m_{sc}(\beta) \vec{u}^* \vec{v}| \propto \frac{1}{\sqrt{n} (\kappa + \eta)^{1/4}}$

uniformly in $|\beta| \geq 2 + N^{-\frac{2}{3} + \varepsilon}$ and $\eta \geq 0$ (in particular, it holds for real β).

As a consequence of this isotropic local law, we get the "isotropic localization" of the eigenvectors using

$$|\vec{v}^* \vec{u}_k|^2 \leq \eta \operatorname{Im} [\vec{v}^* G(\lambda_k + i\eta) \vec{v}] .$$

Corollary For any deterministic unit vector $\vec{v} \in \mathbb{R}^n$, we have that

$$\max_{1 \leq i \leq N} |\vec{v}^* u_i| \propto N^{-1/2} .$$

Define: $h_1(\beta) := m_{sc}(\beta) + d^{-1}$, $h_2(\beta) := \vec{v}^* G(\beta) \vec{v} + d^{-1}$.

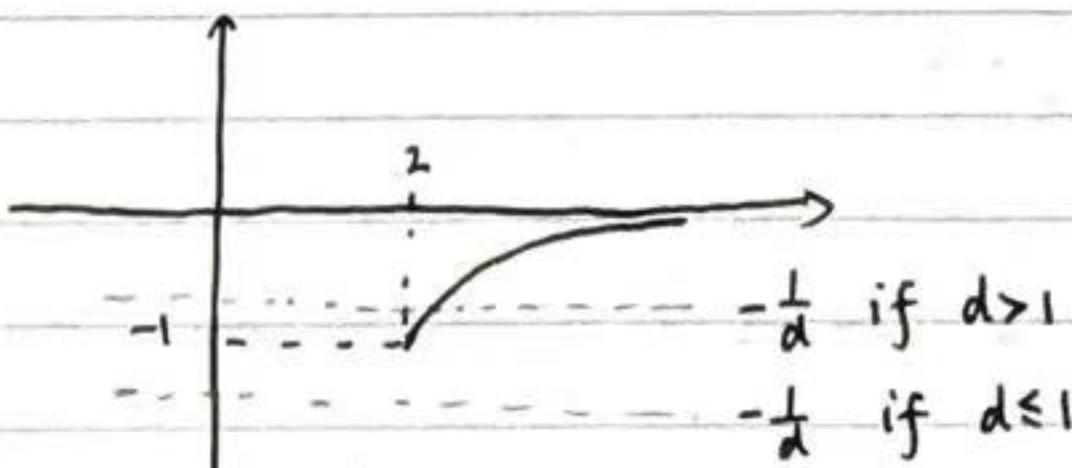
The function $\frac{h_2(\beta)}{h_1(\beta)}$ is analytic on the right half complex plane $\{\beta: \operatorname{Re} \beta \geq 2 + N^{-\frac{2}{3} + \varepsilon}\}$.

Furthermore, w.h.p., H has no eigenvalue inside S_ε , so $h_2(\beta)$ is also analytic on S_ε .

~~$m_{sc}(\beta) + d^{-1} = 0 \Rightarrow m_{sc}(\beta) = -d^{-1} \Rightarrow \beta = d + \frac{1}{d}$ (using $\beta + m_{sc}(\beta) + \frac{1}{m_{sc}(\beta)} = 0$).~~

Moreover

Now, we solve $m_{sc}(x) = -d^{-1}$. Note that $m_{sc}(x) = -1$ at $x=2$ and $m_{sc}(x)$ is monotonically increasing on $[2, +\infty)$ because $m_{sc}'(x) = \int_x^2 \frac{P_{sc}(t)}{(t-x)^2} dt > 0$.



Hence, the equation has a unique solution on $(2, +\infty)$ iff $d > 1$. Moreover, the solution is given by

$$x = \theta(d) = d + \frac{1}{d} \text{ by using } \beta + m_{sc}(\beta) + \frac{1}{m_{sc}(\beta)} = 0 .$$

(i) Suppose $d > 1$. Taking ε small enough such that $\{\beta: |\beta - \theta(d)| = \frac{1}{N^{1-\frac{1}{2}+\varepsilon}} \subseteq S_\varepsilon\}$, then this circle encloses one zero of $h_1(\beta)$. ④

Using the isotropic local law, we get that w.h.p., $|h_2(z) - h_1(z)| = |\vec{v}^* G(z) \vec{v} - m_{sc}(z)| \leq N^{-\frac{1}{2} + \frac{\epsilon}{2}}$.

Also, $|h_1(z)| \geq N^{-\frac{1}{2} + \epsilon}$ for $|z - \theta(d)| = N^{-\frac{1}{2} + \epsilon}$. $\Rightarrow |h_1(z)| \geq |h_2(z) - h_1(z)|$.

With the Rouché's theorem, $h_2(z)$ has a zero inside $\{z - \theta(d)| = N^{-\frac{1}{2} + \epsilon}\}$ w.h.p.

Hence, $\exists \lambda_i$ of \tilde{H} such that

$$|\lambda_i - \theta(d)| \leq N^{-\frac{1}{2} + \epsilon} \text{ w.h.p.}$$

By Cauchy interlacing, this λ_i must be λ_N .

(ii) Suppose $d \leq 1$. We consider $h_2(x) = \vec{v}^* G(x) \vec{v} + d^{-1}$ for $x \in I_\epsilon = \{x: x \geq 2 + N^{-\frac{2}{3} + \epsilon}\}$.

Using the isotropic local law, we get that

$$|h_1(x) - h_2(x)| \ll N^{-\frac{1}{3} - \frac{\epsilon}{4}} \text{ uniformly in } x \in I_\epsilon.$$

On the other hand, using the monotonicity of $h_1(x)$ on $[2, +\infty)$, we get that

$$h_1(x) \geq h_1(2 + N^{-\frac{2}{3} + \epsilon}) + d^{-1}$$

$$\geq -1 + N^{-\frac{1}{3} + \frac{\epsilon}{2}} + O(N^{-\frac{2}{3} + \epsilon}) + d^{-1} \geq N^{-\frac{1}{3} + \frac{\epsilon}{2}} + O(N^{-\frac{2}{3} + \epsilon}).$$

In sum, we get that w.h.p., $h_2(x) \geq h_1(x) - |h_1(x) - h_2(x)| \geq \frac{1}{2} N^{-\frac{1}{3} + \frac{\epsilon}{2}}$ w.h.p.

for $x \in I_\epsilon$. Hence, $h_2(x)$ has no solution on I_ϵ , which gives that

$$\lambda_N(\tilde{H}) \leq 2 + N^{-\frac{2}{3} + \epsilon} \text{ w.h.p.}$$

On the other hand, by Cauchy interlacing, $\lambda_N(\tilde{H}) \geq \lambda_N(H) \geq 2 - N^{-\frac{2}{3} + \epsilon}$ w.h.p.

by rigidity of $\lambda_N(H)$. \square

Theorem 5.4 Under the setting of local semicircle law, we have that:

(i) If $d > 1$, then

$$\min_{\lambda \in \mathbb{C} \setminus \{-1\}} |\lambda \vec{v}^* \vec{u}_N - (1 - \frac{1}{d^2})^{1/2}| \ll N^{-\frac{1}{2}},$$

where \vec{u}_i denotes the eigenvector corresponding to λ_i . Furthermore, $\forall \vec{u} \perp \vec{v}$, $\|\vec{u}\| = 1$,

$$|\vec{u}^* \vec{u}_N| \ll N^{-\frac{1}{2}}.$$

(ii) If $d < 1$, then

$$|\vec{u}^* \vec{u}_N| \ll N^{-\frac{1}{2}} \text{ for any fixed unit vector } \vec{u} \in \mathbb{R}^N.$$

(Here, \vec{u} can be \vec{v} . This estimate shows that \vec{u}_N is isotropically delocalized.)

Proof: (i) By Theorem 5.2, we have that $|\lambda_N - \theta_N| \ll N^{-\frac{1}{2}}$. Hence, let γ_N be a contour $\{z: |z - \theta(d)| = \frac{1}{2}|\theta(d) - 2|\}$, we know γ_N only encloses $\theta(d)$ but no other eigenvalue of \tilde{H} . Using the Cauchy integral formula, we get

$$-\frac{1}{2\pi i} \oint_{\gamma_N} \frac{1}{\tilde{H} - z} dz = -\frac{1}{2\pi i} \oint_{\gamma_N} \sum_k \frac{\vec{u}_k \vec{u}_k^*}{\lambda_k - z} dz = \vec{u}_N \vec{u}_N^*$$

Hence, for any unit vector $\vec{u} \in \mathbb{R}^N$, we have that

$$|\vec{u}^* \vec{u}_N|^2 = -\frac{1}{2\pi i} \oint_G \vec{u}^* \frac{1}{H-\lambda} \vec{u} d\lambda, \quad \text{where } (\tilde{H}-\lambda)^{-1} = (H-\lambda + d\vec{v} \vec{v}^*)^{-1} \\ = (G(\lambda)^{-1} + d\vec{v} \vec{v}^*)^{-1}.$$

Woodbury matrix identity $(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)VA^{-1}$.

Then, we can write that $(\tilde{H}-\lambda)^{-1} = G(\lambda) - G(\lambda)\vec{v}(d^{-1} + \vec{v}^*G(\lambda)\vec{v})^{-1}\vec{v}^*G(\lambda)$.

Thus, we get that

$$|\vec{u}^* \vec{u}_N|^2 = -\frac{1}{2\pi i} \oint_G [\vec{u}^* G(\lambda) \vec{u} - \vec{u}^* G(\lambda) \vec{v} \frac{1}{d^{-1} + \vec{v}^* G(\lambda) \vec{v}} \vec{v}^* G(\lambda) \vec{u}] d\lambda \\ = +\frac{1}{2\pi i} \oint_G (\vec{u}^* G(\lambda) \vec{v})(\vec{v}^* G(\lambda) \vec{u}) \frac{1}{d^{-1} + \vec{v}^* G(\lambda) \vec{v}} d\lambda.$$

Taking $\vec{u} = \vec{v}$, we get that

$$|\vec{v}^* \vec{u}_N|^2 = \Theta \frac{1}{2\pi i} \oint_G (\vec{v}^* G(\lambda) \vec{v})^2 \frac{1}{d^{-1} + \vec{v}^* G(\lambda) \vec{v}} d\lambda.$$

Note that the integrand has a simple pole at $\lambda = \sigma_N$ as we have shown before, where $\vec{v}^* G(\sigma_N) \vec{v} = -d^{-1}$. Hence,

$$|\vec{v}^* \vec{u}_N|^2 = \Theta \frac{1}{2\pi i} (d^{-2}) \oint_G \frac{1}{d^{-1} + \vec{v}^* G(\lambda) \vec{v}} d\lambda \quad (\text{by the residue theorem}) \\ = \Theta \frac{1}{2\pi i} d^{-2} \oint_G \frac{1}{d^{-1} + m_{sc}(\lambda) + O(N^{-1/2})} d\lambda \quad (\text{by the isotropic local law}) \\ = \Theta \frac{1}{2\pi i} d^{-2} \oint_G \frac{1}{d^{-1} + m_{sc}(\lambda)} d\lambda + O(N^{-1/2})$$

Note that the integrand has a simple pole at $\lambda = \theta(d)$. Using the residue theorem, we get that

$$+\frac{1}{2\pi i} \oint_G \frac{1}{d^{-1} + m_{sc}(\lambda)} d\lambda = \Theta \frac{1}{m_{sc}(\theta(d))} = \Theta f'(m_{sc}(\theta(d))) = -1 + d^2,$$

$$f(m) = -m - \frac{1}{m} \quad (\text{recall } \lambda + m + \frac{1}{m} = 0)$$

where $f(m) = f(m + \frac{1}{m})$ is the inverse function of m_{sc} . Thus, we have

$$|\vec{v}^* \vec{u}_N|^2 = 1 - \frac{1}{d^2} + O(N^{-1/2}).$$

Taking square root, we get $\min_{\alpha \in \{+1, -1\}} |\alpha \vec{v}^* \vec{u}_N - (1 - \frac{1}{d^2})^{1/2}| \propto N^{-1/2}$.

On the other hand, if $\vec{u} \perp \vec{v}$, then $\vec{u}^* G(\lambda) \vec{v} = (\vec{u}^* \vec{v}) \cdot m_{sc}(\lambda) + O(N^{-1/2}) = O(N^{-1/2})$.

Using this fact, we get

$$|\vec{u}^* \vec{v}_N|^2 \leq \Theta(N^{-1/2})^N \left| \oint_G \frac{1}{d^{-1} + \vec{v}^* G(\lambda) \vec{v}} d\lambda \right| \leq N^{-1}.$$

$$\Rightarrow |\vec{u}^* \vec{v}_N| \propto N^{-1/2}.$$

(ii) By Theorem 5.2, we have that $|\lambda_N - 2| \prec N^{-2/3}$.

Using the spectral decomposition, we obtain that

$$|\vec{u}^* \vec{v}_N|^2 \propto \eta \operatorname{Im} [\vec{u}^* \tilde{G}(\lambda_N + i\eta) \vec{u}], \quad \tilde{G}(z) := (\tilde{H} - z)^{-1}.$$

Using the Woodbury matrix identity, for $z = \lambda_N + i\eta$, we have

$$\vec{u}^* \tilde{G}(z) \vec{u} = \vec{u}^* G(z) \vec{u} - \vec{u}^* G(z) \vec{v} \frac{1}{d^{-1} + \vec{v}^* G(z) \vec{v}} \vec{v}^* G(z) \vec{u}$$

By the isotropic local law for $G(z)$, we get that $\vec{u}^* G(z) \vec{u}$, $\vec{u}^* G(z) \vec{v}$ and $\vec{v}^* G(z) \vec{u}$ are of order 1 with high probability. Furthermore, by the anisotropic local law,

$$|d^{-1} + \vec{v}^* G(z) \vec{v}| = |d^{-1} + m_{sc}(z) + O(N^{1/2})| \gtrsim 1 \text{ w.h.p.}$$

In sum, we get that $|\vec{u}^* \tilde{G}(z) \vec{u}| = O(1)$ w.h.p. for any $\eta \geq N^{-1+\epsilon}$. Hence,

$$|\vec{u}^* \vec{v}_N|^2 \propto N^{-1+\epsilon} \forall \text{ small constant } \epsilon > 0. \quad \square$$

Rmk: Theorem 5.3 and Theorem 5.4 give the "LLN" (or first order asymptotics) of the largest eigenvalue λ_N and its eigenvector \vec{u}_N . We can also derive their "CLT" (second order asymptotics). For the largest eigenvalue λ_N , we know that:

(1) If $d < 1$, then $N^{2/3}(\lambda_N - 2) \rightarrow \text{TW}$ in law.

(2) If $d > 1$ and the entries of H are GOE / GUE, then $N^{1/2}(\lambda_N - \theta(d))$ converges to a centered Gaussian.

As d crosses 1, a famous BBP transition occurs for λ_N (Baik - Ben Arous - Peche). However, if H is non-Gaussian, (2) may not hold: there may be some non-Gaussian part that depends on the third and fourth moments of the H entries. Here, we only consider the simpler Gaussian case. We will show that:

Theorem 5.5: Given a GOE / GUE H , if $d > 1$, we have that

$$N^{1/2}(\lambda_N - \theta(d)) \rightarrow N(0, \sigma^2(d)), \text{ in law,}$$

$$\text{where } \theta(d) = \frac{2}{\beta} \frac{|d|^{1/2} - 1}{d^4}, \quad \text{where } \beta = 1 \text{ for GOE and } \beta = 2 \text{ for GUE.}$$

Proof: We only show the proof for GOE. By rotational invariance of H , it suffices to choose $\vec{v} = \vec{e}_1$, i.e.,

$$\tilde{H} = H + \begin{pmatrix} 1 & & \\ & 0 & \\ & & \ddots \\ & & & 0 \end{pmatrix}.$$

We have seen that λ_N satisfies the equation

$$\vec{v}^* G(\lambda_N) \vec{v} = G_{11}(\lambda_N) = d^{-1} \Rightarrow m_{sc}(\lambda_N) + [G_{11}(\lambda_N) - m_{sc}(\lambda_N)] = d^{-1}. \quad (*)$$

By Thm 5.2, we know that $|\lambda_N - \theta(d)| \prec N^{-1/2}$. Hence, for large enough N ,

$\lambda_N > 2$ and $\lambda_N - 2 \geq 1$ w.h.p. Then, the isotropic local law gives that $|G_{11}(\lambda_N) - m_{SC}(\lambda_N)| \ll N^{-1/2}$.

Now, we perform a Taylor expansion of $m_{SC}(\lambda_N)$ around $m_{SC}(\theta(d)) = -d^{-1}$:

$$m_{SC}(\lambda_N) + [G_{11}(\lambda_N) - m_{SC}(\lambda_N)] = m_{SC}(\theta(d))$$

$$\Rightarrow m_{SC}'(\theta(d)) \cdot [\lambda_N - \theta(d)] + O(N^{-1}) = -[G_{11}(\lambda_N) - m_{SC}(\lambda_N)] \\ = -[G_{11}(\theta(d)) - m_{SC}(\theta(d))] + [\Delta G_{11}(\theta(d))] \\ - \Delta G_{11}(\lambda_N) \quad (*)$$

We denote $\Delta G_{11} := G_{11} - M_{SC}$.

To bound $\Delta G_{11}(\theta(d)) - \Delta G_{11}(\lambda_N)$, we need to control $G_{11}(z) - m_{SC}(z)$ for z between $\theta(d)$ and λ_N . The sharp way to do this is using Cauchy integral formula again:

$$G_{11}'(z) - m_{SC}'(z) = \frac{1}{2\pi i} \oint_{C_z} \frac{1}{\zeta - z} \Delta G_{11}(\zeta) d\zeta, \quad C_z := \{\zeta : |\zeta - z| = \varepsilon\} \text{ for a small enough constant } \varepsilon > 0.$$

Then, using the isotropic local law $|\Delta G_{11}(z)| \ll N^{-1/2}$, we get that

$$|G_{11}'(z) - m_{SC}'(z)| \ll N^{-1/2}.$$

Hence, we have

$$|\Delta G_{11}(\lambda_N) - \Delta G_{11}(\theta(d))| \ll N^{-1/2} |\lambda_N - \theta(d)| \ll N^{-1}.$$

Then, equation $(*)$ becomes

$$m_{SC}'(\theta(d)) \cdot [\lambda_N - \theta(d)] = -[G_{11}(\theta(d)) - m_{SC}(\theta(d))] + O(N^{-1}). \quad (*)$$

Previously, we have calculated that

$$m_{SC}'(\theta(d)) = \frac{1}{f'(m_{SC}(\theta(d)))} = \frac{1}{f'(-d^{-1})} = \frac{1}{d^2 - 1}.$$

Thus, from $(*)$, we get that

$$\sqrt{N}(\lambda_N - \theta(d)) = (d^2 - 1) \left\{ -\sqrt{N} [G_{11}(\theta(d)) - m_{SC}(\theta(d))] \right\} + O(N^{-1/2}).$$

Hence, we only need to establish the CLT of

$$Y(d) := -\sqrt{N} [G_{11}(\theta(d)) - m_{SC}(\theta(d))].$$

There are many ways to prove the CLT of $Y(d)$. Some popular ones include:

calculating the moments / characteristic function of Y , and compare them with those of a centered Gaussian. Here, we use Stein's method, that is, we show that for any ~~bounded~~ smooth function $f \in C_c^\infty(\mathbb{R})$,

$$\mathbb{E}[Y f(Y)] = S^2 \mathbb{E}[f'(Y)] + o(1).$$

(This without $o(1)$, this is the Gaussian integration by parts formula with S^2 giving the asymptotic variance of Y .)