On extremal eigenvalues that lays outside of the support of the spectral distribution

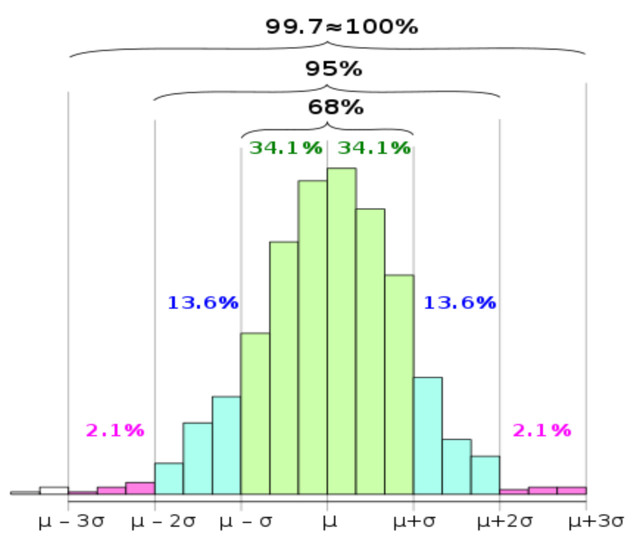
Table of Contents

Opening Words

This document is a little heavy in mathematics. Doing research usually follows the three steps: **concept clarification**, **criterion setting** and **system building.** I'll try to do the work according to this guidance and try to use the least mathematics for the topic.

There are four keywords in the title: **extremal eigenvalue**, **lay outside**, **support**, **spectral distribution**. It is not sufficient to only tell you that it belongs to the area of random matrix theory(RMT) and the document is to study extremal eigenvalues' fluctuation. Better to pay attention to the words "**random matrix**" and "**extremal eigenvalue**", since they are the origin and destination of the document, respectively.

For a preliminary understanding, it is better to see the graph illustrated. Just consider the common way in statistics to describe a data distribution, e.g. in univariate Gaussian distribution, let μ and σ be the mean value and standard deviation. It is known that the data within one standard deviation of the mean account for about 68% of the set, while within two standard deviations account for about 95%. And within three standard deviations namely, in the domain $[\mu \mp 3\sigma]$ which includes around 99.7% of the data, suppose we can see the domain as the main part of the data distribution(since it is nearly 100%), then the so-called extremal data outside it can be seen as outliers.



*The figure is from https://en.wikipedia.org/wiki/68%E2%80%9395%E2%80%9399.7_rule

Similarly, but to be more complex, the study object in this document is also some extremal data but with more randomness.

1 Basics and Notations

Before going into formal content, the necessary mathematics are listed here.
There are lots of beauty and profound mathematical properties, but only the needed results are put here, without so much in detailed explanation.
Not all the terms are explicitly used or mentioned, but they are foundational to the results in next parts and listed here

According to convention, the N-dimensional real number field is denoted by $\mathbb{R}^n$; N-dimensional complex number field, $\mathbb{C}^n$

Deterministic Matrices: usually in which the elements are deterministic functions or certain numbers. The elements of matrices are also called entries.Here the identity matrix is denoted by I

Eigenvalues and Eigenvectors
For a square matrix A, a pair of scalar and non-zero vector (λ,v)satisfying

$$A\mathrm{v} = \lambda\mathrm{v}$$

is called an eigenvalue–eigenvector pair.
The eigenvalue of a square matrix is called the spectrum of a matrix.

Transpose: a matrix operator, which flips a matrix over its diagonal; that is, it switches the row and column indices of the matrix A by producing another matrix, often denoted by A^T

Trace:, denoted by Tr, which is the sum of diagonal entries or eigenvalues, (trace invariance by similarity transformation)

Symmetric matrix: the square matrix that is equal to its transpose
Hermitian matrix: the complex square matrix that is equal to its own conjugate transpose, for the reason, its diagonal entries are real.
The two are both diagonalizable with real eigenvalues.

Singular Values
For a matrix A, if it is not diagonalizable, there is an alternative decomposition(SVD,singular value decomposition) as

$$A = VSU^T$$

where S is a non-negative diagonal matrix, whose elements are called the singular values of A, and U,V are two real, orthogonal matrices.

A is said to be **positive semi-definite** or **non-negative-definite** $\Leftrightarrow \mathrm{x}^T A\mathrm{x} \geq 0, \forall \mathrm{x} \in \mathbb{R}^n$

Dirac delta function: also known as the unit impulse symbol, is a generalized function or distribution over the real numbers x, whose value is zero everywhere except at a base point λ , and whose integral over the entire real line is equal to one; denoted by $\delta_\lambda(x)$

almost surely, denoted by a.s., for convergence, it means the measure of divergent points is zero.

convergent weakly:A sequence $X_1, X_2 ...$, of real-valued random variables is said to converge in distribution, or converge weakly, or converge in law to a random variable X if $\lim_{n\to\infty} F_n(x) = F(x), \forall x \in \mathbb{R}$ at which F is continuous. Here F_n and F are the cumulative distribution functions of random variables X_n and X, respectively.

Support: the points of non-zero probability for a probability distribution, denoted by $supp(\mathbb{P})$

2 Introdunction to Random Matrices

Random matrices: some or all the entries are random variables, also called random ensembles

Dyson index: denoted by $\beta = 1,2,4$,which is referred to the random matrices with real entries, complex entries and quaternionic entries,respectively.

2.1 Spectral Distribution

For a Hermitian matrix H of size $N \times N$, we denote by $\{\lambda_1 \geq \ldots \geq \lambda_N\}$ the set of eigenvalues of H, ranked in decreasing order,.

We define and denote by $\rho_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(H)}$ the **empirical spectral distribution(ESD)**, which is also called the sample eigenvalue density.

The **Stieltjes transform** of ρ_N is denoted for $z \in \mathbb{C} \setminus \mathbb{R}$ by

$$g_{\rho_N}(z) = \int_R \frac{1}{z-x} d\rho_N(x) = \frac{1}{N} Tr(zI - H)^{-1}$$

, $z \notin \{\lambda_i : 1 \leq i \leq N\}$

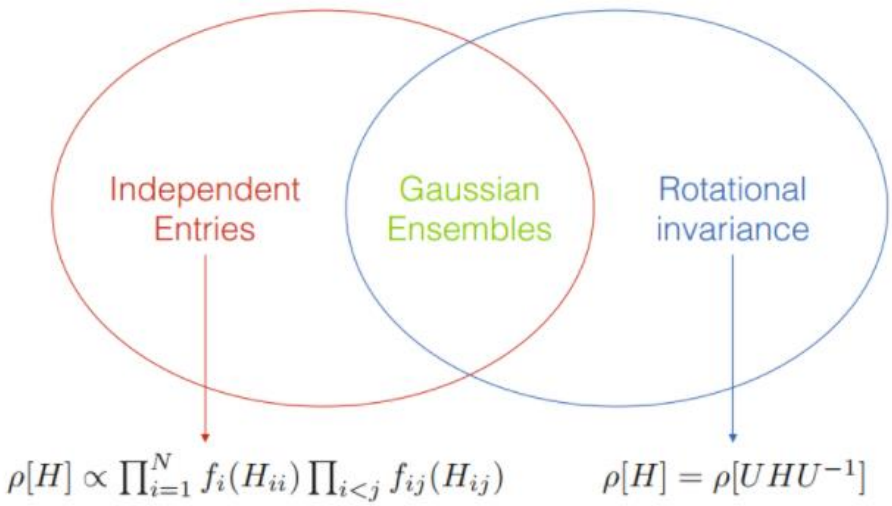
If $\rho_N(x)$ converges weakly to some probability measure $\rho(x)$ as $N \to \infty$, $\rho(x)$ is called the **limiting spectral distribution (LSD)** of H , accordingly, the Stieltjes transform of $\rho(x)$ is

$$g_\rho(z) = \int_{supp\{\rho\}} \frac{1}{z-x} d\rho(x)$$

Both ESD and LSD are mean value, which are used to represent the average effect of eigenvalues.
*Stieltjes transform is not explicitly used in document, but it is an important foundation for the results in next parts,so it is still put here.

2.2 Classification

Here we resort to a simple way to classify random matrices, which is from section3.2 in [1]



*The figure is from ref[1], $\prod_i^N f_i\,(H_{ii}) \prod_{i<j} f_i\,(H_{ij})$ is the joint distribution of a Wigner matrix

For the mathematically oriented readers, looking for more formal classifications of random matrix models, it is recommended to read the mini-review [47] and references therein.

a) Wigner matrices

Usually, the matrices have independent entries and some symmetric properties, including real symmetric matrices, complex Hermitian matrices..., e.g:

- adjacency matrices of random graphs
- Lévy matrices(independent power-law entries)
- power-law banded matrices
- ...

b) Rotational invariance matrices

The matrices, through a similarity transformation, the spectral distribution is invariant. As for similarity transformation, it includes :

- unitary transformation,
- biunitary transformation,

- ...

For simplicity, this property of rotational invariance means that any two matrices that are related via a similarity transformation occur in the ensemble with the same probability, and it indicates that the eigenvectors are not that important, as we can rotate our matrices as freely as we wish, and still leave their statistical weight unchanged. e.g.:

- Wishart-Laguerre ensembles(i.e., Laguerre unitary ensemble, LUE),which is the famous 'Wishart matrices'
- Jacobiclassical ensembles, the so-called "weakly-confined" ensembles
- ...

c)*Gaussian Ensembles

The intersection of a) and b) is only Gaussian Ensembles, which was proved in [46] Due to the fine properties, Gaussian ensembles are often used as a good starting point for study. There are just the three:

- GOE: Gaussian Orthogonal Ensemble, real entries
- GUE: Gaussian Unitary Ensemble, complex entries
- GSE: Gaussian Symplectic Ensemble, quaternionic entries

d) Others

Others are exisitent, for instance, biorthogonal ensembles, which are non-invariant and with non-independent entries...

3 Extremal Eigenvalues' Fluctuation in Classical Models

In this part, we start from the two widely-studied and classical random matrix models, namely, wigner matrices and sample covariance matrices. The two corresponds to a) and b) in the classification part.

3.1 Classical Wigner matrices

The Wigner matrices here are real symmetric or complex Hermitian matrices whose entries are independent up to the symmetry condition, with the form:

$$W_N = \frac{1}{\sqrt{N}} X_N$$

where N is sample size, $(X_N)_{ii}$,$\sqrt{2}Re\big((X_N)_{ij}\big)_{i<j}$, $\sqrt{2}Im\big((X_N)_{ij}\big)_{i<j}$ are i.i.d with distribution τ with variance σ^2 and $mean$ 0; if $\tau = \mathcal{N}(0, \sigma^2)$, W_N^G is G.U.E.-matrix.

3.1.1 Asymptotics

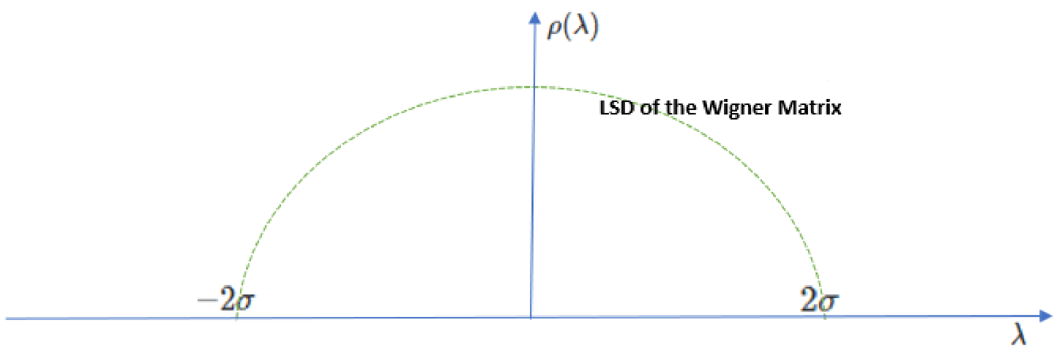
Here are the results on the asymptotic behaviors of the spectral measure[5,6] and extremal eigenvalues[7]
Limiting Spectral Distribution:

$$\mu_{W_N} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}(W_N) \to \mu_{sc}(\lambda), a.s. N \to +\infty$$

where the spectral density:

$$\rho(\lambda) = \frac{d\mu_{sc}(\lambda)}{d\lambda} = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - \lambda^2}1[\lambda_-, \lambda_+]$$

here the left edge $\lambda_- = -2\sigma$,and the right dege$\lambda_+ = +2\sigma$



It can be seen that the LSD of the Wigner matrix is a semicircle with the radius of 2σ
It is worth mentioning that the above result can be called **bulk universality**[65],which concerns the local statistics of eigenvalues in the interior of the spectrum,we can refer to[18] for further discussion

Extremal eigenvalues: If $\int x^4 d\tau(x) < \infty$, then

$$\lambda_{max}(W_N) \to \lambda_+ \ and \ \lambda_{min}(W_N) \to \lambda_- , a.s. N \to +\infty$$

We can see when $N \to \infty$,the spectral density of the Wigner matrix weakly converges to a semi-circle shape distribution with the bulk of $[-2\sigma, 2\sigma]$, and meanwhile, the largest eigenvalue(λ_{min}) and the smallest eigenvalue(λ_{max}) weakly converge to the right edge(2σ) and left edge(-2σ), respectively.

3.1.2 All eigenvalues fluctuate in their definitional domains

*The discussion here is actually not only limited to the case of Wigner matrices.

We can just think about it easily. In the process of $N \to \infty$,the edges of the bulk of the spectral density of the Wigner matrix is moving on the way to the limits, the smallest and largest eigenvalues are also moving on the way to their limits, as said previously, the spectral distribution is merely spectral expectation, so there must exist deviations to it from other eigenvalues' moving behaviors, in other words, the distances between the moving edges and other eigenvalues are a dynamic process, which can be described through probability distribution.

Let's relax the process to a simpler case and fix the edges as the edges of the limiting spectral density(we can also call the kind of edge as the classical edge), namely in the Wigner case, they are-2σ and 2σ . Theoretically, any eigenvalue is probable to occur in any place of its domain of definition. In fact, It has been observed that some eigenvalues are probable to walk beyond the one edge or two edges. That is to say, outside the bulk, the zone where the eigenvalues should not be existent in terms of spectral distribution, the probability of eigenvalues is non-zero, although it can be small. Thus, there exist the transition regions between the edges and some eigenvalues.

With regard to the largest and smallest eigenvalues, after randomly moving in transition regions, they finally converge to some points, which in the Wigner case are just the right and left edges, respectively. Intuitively, it is quite natural to consider the distance between an extremal eigenvalue and a near edge. In a general sense, the results on the distance between an edge and some kind of extremal eigenvalue have been investigated quite sufficiently.

Concretely, the fluctuating local density of eigenvalues can be computed from the number of eigenvalues within a distance, for concrete derivations, we can refer to page-75 in[2].

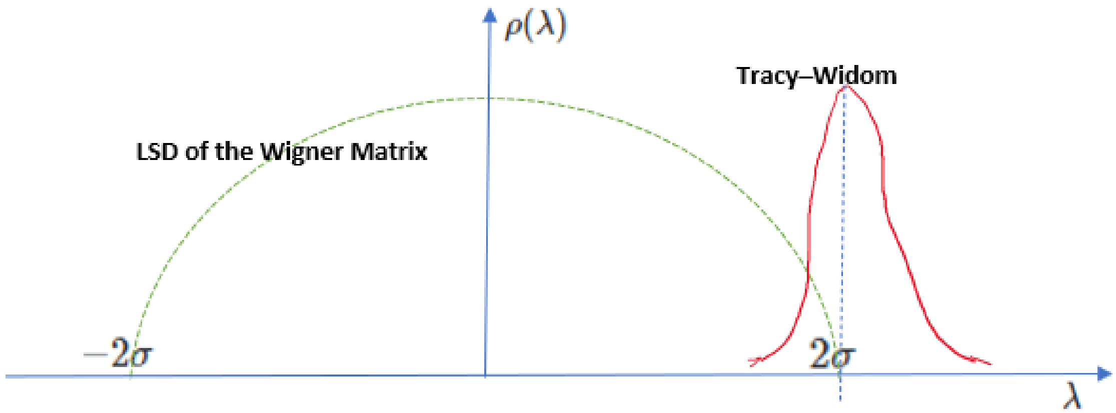
3.1.3 Edge universality and Tracy–Widom law

Starting from the largest eigenvalue
The first beautiful result was achieved by Tracy and Widom on the distribution of the largest eigenvalue of gaussian wigner matrices.

This result can be formally stated : the rescaled distribution of $\lambda_{max} - \lambda_+$ converges, as $N \to \infty$, follows the Tracy–Widom distribution, usually denoted by TW_β:

$$\mathbb{P}\left(N^{2/3}(\lambda_{max} - \lambda_+) \leq s\right) = TW_{\beta \in (1,2)}(s)$$

here for the case,$\lambda_+ = 2\sigma$.
For intuitive understanding, see the graph below.



*The curve of Tracy-Widom distribution in this and later graphs is only **schematic plot**, which is not in its actual size.

Simply speaking for the Wigner case, the largest eigenvalue λ_{max} does not fluctuate very far away from the right classical edge λ_{+}. Take for example N = 1000,and λ_{max} is within $1000^{-2/3} = 0.01$ away from $\lambda_{+} = 2\sigma$. This indicates that the width of the region around λ_{+} within which one expects to observe the largest eigenvalue of a Wigner matrix goes down as $N^{-2/3}$.

For the **density of Tracy–Widom density**, we have known that $tw_{\beta\in(1,2)}(s) = TW'_{\beta\in(1,2)}(s)$, w.r.t the left and right far tails:

$$tw_{\beta\in(1,2)}(s) \propto -s^{3/2}, as\ N \to \infty;\ \ tw_{\beta\in(1,2)}(s) \propto -|s|^{3}, as\ N \to -\infty$$

It is noted that the left tail is much thinner than the right tail, it shows pushing the largest eigenvalue inside the bulk is more difficult than pulling one eigenvalue away from λ_{+}. Using this analogy and the formalism in section 5.4.2 in [2], the large deviation regime of the Tracy–Widom problem (i.e. $\lambda_{max} - \lambda_{+} = O(1)$) can be obtained. Note that the result is exponentially small in N as the $s^{3/2}$ behavior for $s \to \infty$ combines with $N^{2/3}$ to give a linear dependence in N

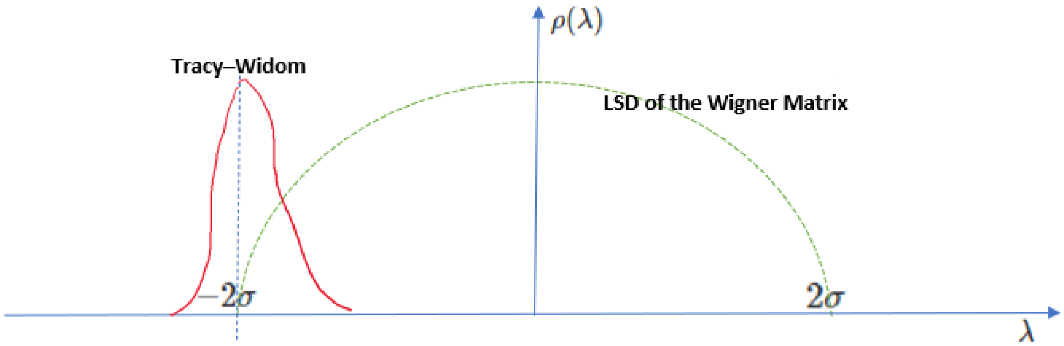
As discussed, the distribution of the largest eigenvalue follows the law of Tracy–Widom, which exhibits a type of universality, it's called the **edge universality**.

In fact,The Tracy–Widom law holds to more general Wigner matrices[8-13], **even in the case** that the symmetric assumption is partially removed[14,15]. The results can be found:

- Edge universality for Wigner matrices whose entries have vanishing third moments was proved in [42];
- Edge universality without moment matching was proved in [43] for Wigner matrices and in [44, 45] for generalized Wigner matrices;

- A necessary and sufficient condition on the entries' distributions for the edge universality to hold was given in [16].
- The fluctuations of the largest eigenvalue of a real symmetric or complex Hermitian Wigner matrix of size N converge to the Tracy–Widom laws has been proved at the rate of $O\left(N^{-1/3+\omega}\right), \forall\omega > 0$, as $N \to \infty$ in[48].

Then what about the smallest eigenvalue?
The distribution of the smallest eigenvalue λ_{min} around the lower edge λ_{-} also follows the Tracy–Widom law. So is a similar graph:



As to the tails, still the lift is thinner than the right, but it means putting the smallest eigenvalue inside the bulk is easier than dragging outside.

For any ordered eigenvalue
For more **general cases** ,in [20], the authors proved that after proper rescaling, the 1st, 2nd, 3rd, etc. eigenvalues of Wigner matrices weakly converge to Tracy–Widom distribution. This is to say,the distance between a classical edge(the left edge or the right edge) and any ordered eigenvalue follows the Tracy–Widom distribution as N$\to\infty$. Obviously, the case of extremal eigenvalues are included in.

To recap, in the Wigner matrices, the extremal eigenvalues weakly converge to the bulk edges of the limiting spectral distribution, and the corresponding fluctuations follow the Tracy–Widom distribution.

3.2 Classical sample covariance matrices

After introducing Wigner matrices, with a similar introducing structure, let's go into a kind of invariance matrices--- sample covariance matrices

3.2.1 Similarity and complexity

Although **MOST** of the laws on the fluctuation of eigenvalue hold in the part for Wigner matrices still hold for the sample covariance matrices here, there are some more complex conditions(mainly since dependent entries) and important exceptions of interest.

3.2.2 Asymptotics

When the sample covariance matrix is assumed as

$$S_N = \frac{1}{p(N)} X_N X_N^*$$

where X_N is a complex random matrix with i.i.d. entries. It is assumed that X_N is a $N \times p(N)$ and $p(N) \leq N$ complex random matrix, $Re\big((X_N)_{ij}\big), Im\big((X_N)_{ij}\big), i = 1, \ldots, N, j = 1, \ldots, p(N)$ are i.i.d,following distribution τ with variance $\frac{1}{2}$ and mean 0. Note that the spectra of $\frac{1}{p(N)} X_N X_N^*$ and $\frac{1}{p(N)} X_N^* X_N$ differ by$|p(N) - N|$ zero eigenvalues. If τ is Gaussian, S_N=:S_N^G is an L.U.E matrix,i.e.the Wishart matrix

From the definition, we can see, the sample covariance matrices are positive semidefinite.

The behavior of the spectral measure as $N \to \infty$ follows Marcenko–Pastur distribution

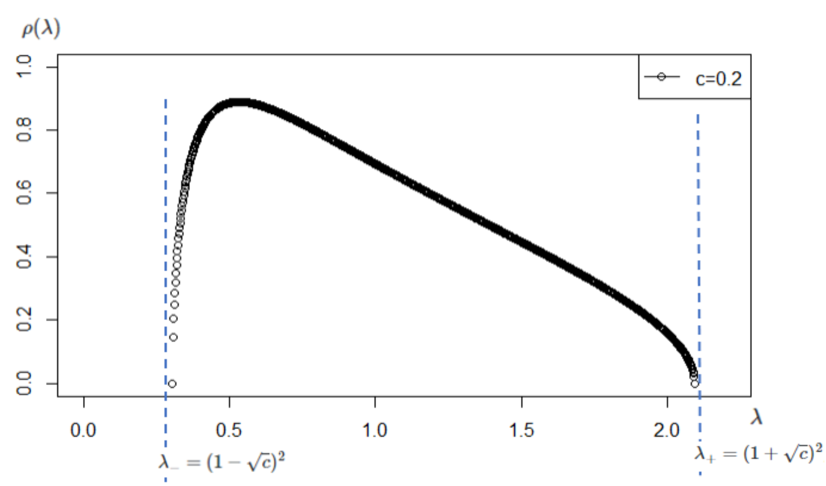
Limiting Spectral Distribution ([21]): If $C_N := \frac{N}{p(N)} \to c \in (0,1]$ as $N \to \infty$, then

$$\mu_{SC} \to \mu_{MP}, as\ N \to \infty$$

where the spectral density

$$\rho(\lambda) = \frac{d\mu_{SC}}{d\lambda} = \frac{1}{2\pi c\lambda} \sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)} 1_{[\lambda_-, \lambda_+]}$$

where $\lambda_- = \left(1 - \sqrt{c}\right)^2, \lambda_+ = \left(1 + \sqrt{c}\right)^2$,c is called aspect ratio. The graph is an



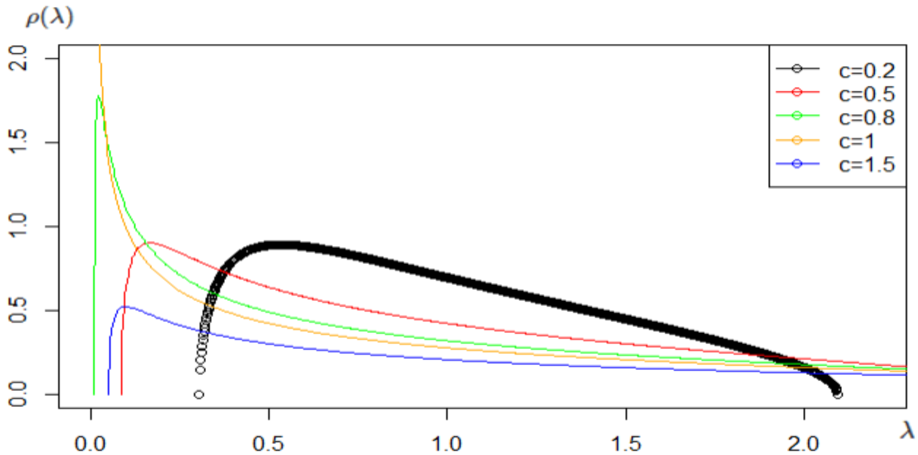
example with $c = 0.2$.

In[41],for some well-designed sample covariance matrices, the authors find a certain threshold c_+ such that when $c > c_+$,the limiting spectral distribution also exhibits convex decay at the right edge of the spectrum

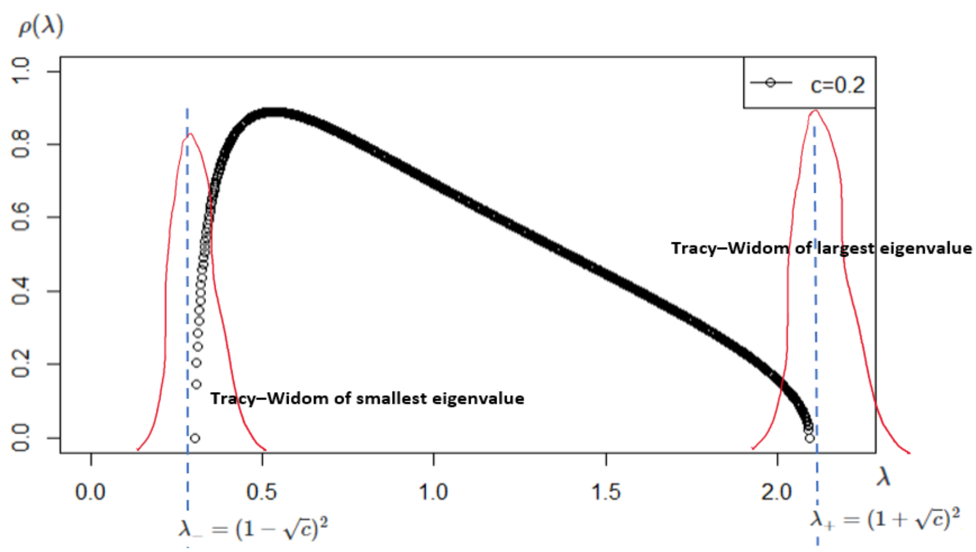
Extremal eigenvalues([2,22-24]): If $\int x^4 d\tau(x) < \infty$, then

$$\lambda_{max}(S_N) \to \lambda_+ \ and \ \lambda_{min}(S_N) \to \lambda_- \ , a.s. N \to +\infty$$

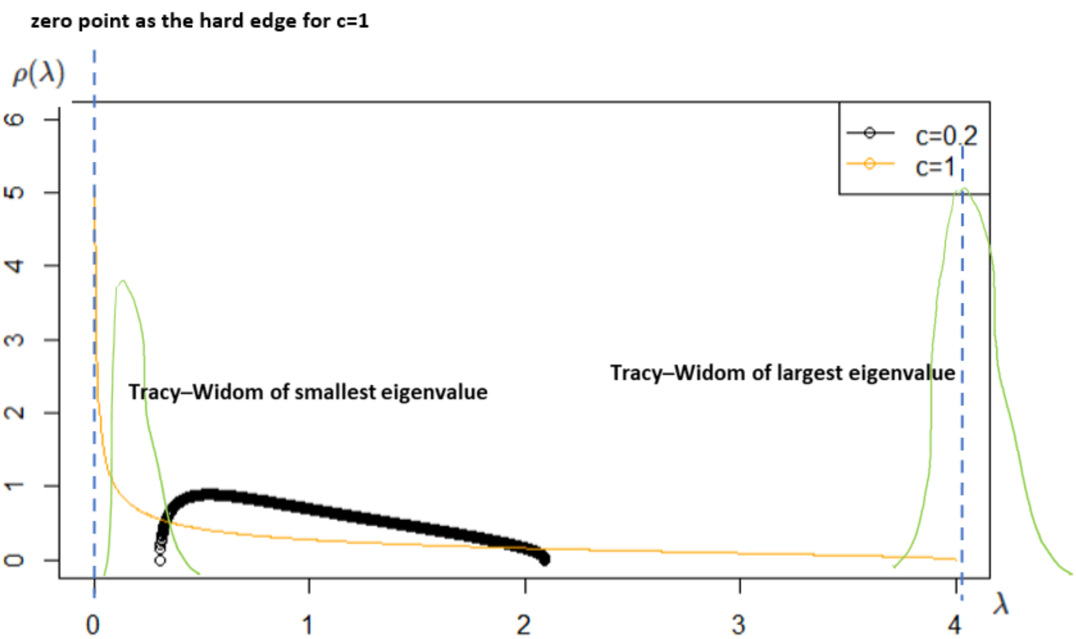
For the two results, if we extend to more general conditions on entries, or let c be $(0, \infty]$,or choose other assumptions, some formalized results are also achieved, for further readings on the asymptotics of empirical spectral measure, we refer to[25]; more details, on the asymptotics of extremal eigenvalues of sample covariance matrices, are well-studied are in[26-29]. The graph below are the examples with more c values.



3.2.3 Fluctuation in definitional domain



Like that of Wigner matrices, all the eigenvalues of sample covariance matrices fluctuate and follows the Tracy–Widom law(details in next part) , but under some condition, some eigenvalues only fluctuate inside the bulk, the corresponding edge is called **hard edge**,e.g. the fluctuation from the smallest eigenvalue to the left edge. Take the Wishart matrices as the example, when $c = 1$, the lower edge $\lambda_- = 0$, this is a “hard” edge since all eigenvalues of the empirical matrix must be non-negative. In contrast, the Wigner semicircle edges are “soft”.



3.2.4 Edge universality and Tracy–Widom law

Like wigner case, for the largest eigenvalue of Wishart matrices, as $N \to \infty$, the rescaled distribution of $\lambda_{max} - \lambda_+$ converges to the Tracy–Widom distribution, with the similar form

$$\mathbb{P}\left(N^{2/3}(\lambda_{max} - \lambda_+) \leq \gamma s\right) = TW_{\beta\in(1,2)}(s)$$

here for the case,$\lambda_+ = \left(1+\sqrt{c}\right)^2$, $\gamma = \sqrt{c}\lambda_+^{2/3}$.

In fact, this law on the largest eigenvalue is not restricted to the Wishart ensemble and it is universal for general sample covariance matrices, even removing the condition on c (see [30]). As to the Tracy–Widom law on the smallest eigenvalue, it still holds when $c \neq 1$([31]).

The results can be found:

- Edge universality for sample covariance matrices when$\sqrt{N}X_N)_{ij}$ have vanishing third moments was proved in [32] under the condition $c \neq 1$;
- Edge universality without moment matching for sample covariance matrices was proved in [33] when $c \neq 1$;

- A necessary and sufficient condition on the entries' distributions for the edge universality to hold was given in [34].
- When $c = 1$ (i.e. $N = p(N)$), the smallest eigenvalue exhibits a different asymptotic behavior, which is referred to as the **hard edge** case, and the corresponding universality in general sample covariance matrices was studied in[35,36].
- When $c > 0$,the convergent rate to its Tracy–Widom limit has been improved to $N^{-1/3}$ in[37].

To recap, compared to Wigner matrices, since the increased complexity in elements, relation, structure and interaction in sample covariance matrices, the asymptotics of extremal values presents more difficulties for study. At least, it is known that in a general sense, the convergence and fluctuation of extremal values have achieved the comparable results to Wigner matrices

3.3 Summary:Universality at soft edges

With regard to the universality of the Tracy-Widom fluctuations at soft edges, it can extend to quite general deformed Wigner matrices or sample covariance matrices without outliers. The involved methods pursue a Green function comparison strategy ([40, 49, 50]) or make use of anisotropic local laws [51].

Two keywords can be found: **'deformed'** and **'outlier'**, so we go to the next section.

4 Outliers in Deformed Models

Here we draw a line on outliers.
For a square random matrix with size $N \times N$, as $N \to \infty$,if an eigenvalue converge to outside the bulk and all edges, the eigenvalue can be seen as an outlier, although it can fluctuate inside the bulk.

Later in the deformed models, the clear cause and mechanism for outlier's occurrence can be found.

4.1 Common deformations

Deformed models here means the perturbation or distortion of some classical ensembles(such as Wigner matrices, sample covariance matrices...), here we introduce three:

- i) Additive perturbation, $M_N = W_N + A_N$,
  - W_N is a Wigner matrix,
  - A_N is an $N \times N$ Hermitian matrix
- ii) Multiplicative perturbation, $M_N = A_N^{\frac{1}{2}} S_N A_N^{\frac{1}{2}}$
  - S_N is a sample covariance matrix,
  - A_N is a $N \times N$ non negative Hermitian matrix

- iii) Information-Plus-Noise, $M_N = \left(\sigma \frac{X_N}{\sqrt{p}} + A_N\right)\left(\sigma \frac{X_N}{\sqrt{p}} + A_N\right)^*$
  - A_N is a rectangular $N \times p(N)$ matrix,
  - X_N is a rectangular $N \times p(N)$ matrix,with i.i.d. entries
  - σ is some positive real numbers
- Other random perturbations

Usually,M_N is called **perturbation** matrix;A_N, **deformation** matrix.

Additionally, these three kinds of deformations have been also considered for isotropic models, if replacing in i) and ii) W_N, S_N by UBU^* with U Haar distributed and B deterministic, and in iii) X_N by a random matrix whose distribution is biunitarily invariant.

Obviously, the complexity to study deformed models is increased quite a lot in the deformed models. On the basis of original methods and by the empowerment of **free convolution and analytic subordination** from free probability theory(see more in[52,1,2]), lots of important results are attained elegantly.

4.2 Asymptotics

Setting different conditions on A_N can render different results, the usual settings are on rank, the fixed number of eigenvalues, the fixed multiplicity of some independent eigenvalues; these fixed eigenvalues are called **spikes**.

When A_N is finite rank, LSD of M_N is still the same as that of W_N or S_N. But to an extremal eigenvalue of M_N, it **does not** hold always. There is a critical threshold, if a corresponding eigenvalue in A_N goes beyond the threshold, the extremal eigenvalue does not stick to the bulk of M_N and become an outlier, which is said to be generated by a spiked eigenvalue of A_N. In fact, this law can be extended to any eigenvalue of M_N.

In a general sense, the relevant criterion for a spiked eigenvalue of A_N to generate an outlier in the spectrum of the deformed model is to belong to some set related to the subordination functions, which can provide the definite supports of bulk and outliers, respectively.

For the asymptotic behavior of the deformed model, the location of outliers, the property of the corresponding eigenvectors, in [0], the authors give a unified presentation on the results:

$M_N = A_N + Y_N$ $\mu_{A_N} \to_{N\to+\infty} \nu$ $\mu_{Y_N} \to_{N\to+\infty} \mu$ $\theta \in$ Spect(A_N) θ multiplicity k_i $\theta \notin$ supp(ν)	$M_N = A_N^{1/2} Y_N A_N^{1/2}$ $\mu_{A_N A_N^*} \to_{N\to+\infty} \nu$ $\mu_{Y_N} \to_{N\to+\infty} \mu$ $\theta \in$ Spect(A_N) θ multiplicity k_i $\theta > 0, \theta \notin$ supp(ν)	$M_N = (A_N + Y_N)(A_N + Y_N)^*$ $\mu_{A_N A_N^*} \to_{N\to+\infty} \nu$ $\mu_{Y_N Y_N^*} \to_{N\to+\infty} \mu$ $\sqrt{\mu}$ or $\sqrt{\nu}$ $\boxplus_c$ infinitely divisible $\theta \in$ Spect$(A_N A_N^*)$ θ multiplicity k_i $\theta > 0, \theta \notin$ supp(ν)
$\mu_{M_N} \to_{N\to+\infty} \mu \boxplus \nu$	$\mu_{M_N} \to_{N\to+\infty} \mu \boxtimes \nu$	$\mu_{M_N} \to_{N\to+\infty} (\sqrt{\mu} \boxplus_c \sqrt{\nu})^2$
$g_\tau(z) = \int_{\mathbb{R}} \frac{d\tau(x)}{z-x}$	$\Psi_\tau(z) = \frac{1}{z} g_\tau(\frac{1}{z}) - 1$	$H^{(c)}_{\sqrt{\tau}} = \frac{c}{z} g_\tau(\frac{1}{z})^2 + (1-c) g_\tau(\frac{1}{z})$
$g_{\mu\boxplus\nu}(z) = g_\nu(\omega_{\mu,\nu}(z))$	$\Psi_{\mu\boxtimes\nu}(z) = \Psi_\nu(F_{\mu,\nu}(z))$	$H^{(c)}_{\sqrt{\mu}\boxplus_c\sqrt{\nu}}(z) = H^{(c)}_{\sqrt{\nu}}(\Omega_{\mu,\nu}(z))$
k_i outliers of M_N in the neighborhood of each ρ s.t $\omega_{\mu,\nu}(\rho) = \theta$	k_i outliers of M_N in the neighborhood of each ρ s.t $\frac{1}{F_{\mu,\nu}(1/\rho)} = \theta$	k_i outliers of M_N in the neighborhood of each ρ s.t $\frac{1}{\Omega_{\mu,\nu}(1/\rho)} = \theta$
ξ eigenvector of M_N associated to an outlier in the neighborhood of ρ s.t $\omega_{\mu,\nu}(\rho) = \theta$ $\lVert P_{\mathrm{Ker}(\theta I - A)}\xi\rVert^2 \to_{N\to+\infty} \frac{1}{\omega'_{\mu,\nu}(\rho)}$	ξ eigenvector of M_N associated to an outlier in the neighborhood of ρ s.t $\frac{1}{F_{\mu,\nu}(1/\rho)} = \theta$ $\lVert P_{\mathrm{Ker}(\theta I - A)}\xi\rVert^2 \to_{N\to+\infty} \frac{\rho F_{\mu,\nu}(1/\rho)}{F'_{\mu,\nu}(1/\rho)}$	ξ eigenvector of M_N associated to an outlier in the neighborhood of ρ s.t $\frac{1}{\Omega_{\mu,\nu}(1/\rho)} = \theta$ $\lVert P_{\mathrm{Ker}(\theta I - A)}\xi\rVert^2 \to_{N\to+\infty} \frac{\rho^2 g_{(\sqrt{\mu}\boxplus_c\sqrt{\nu})^2}(\rho)}{\theta^2 g_\nu(\theta)\Omega'_{\mu,\nu}(1/\rho)}$

However, it is impossible to give a comprehensive interpretation on the table in the document, since there are lots of theories, especially free probability theory, and technical assumptions to introduce, and the notations in the table are quite different**. If possible later, I'll write a new document on the part**.

Still, an intuitive explanation is here, and please allow me skip some conceptual explanations:

- The 1st row shows the basic set of each deformed model
- The 2nd row gives the LSD of each deformed model
- The 3rd and 4th are two different integral transformation(based on Stieltjes transform ,free convolution and analytic subordination) for each deformed model
- The 5th gives the sets of outliers' occurrence of each deformed model
- The last row indicates the limiting projection of the eigenvectors associated to outliers' kernel space.
- Concerning the limiting projection of the eigenvectors associated to outliers of Information-Plus-Noise type models, has been proved only in the iid case for diagonal perturbation A_N and in the isotropic case for nite rank perturbation A_N[0]

4.3 Non-universal fluctuations of outliers

The phenomenon is existent for the fluctuations of the outliers : the limiting distribution can depend on the distribution of the **entries** (thus, the dependence is a kind of non- universality), according to the localization/delocalization of the eigenvectors of A_N. There are some discussions under different model assumptions:

- Gaussian case, see[58,59], on random perturbations, see [60,61]
- Wigner case, see[54,55,56,57]
- Sample covariance case, see[53,62,63]

5 References

[0] M Capitaine, C Donati-Martin,Spectrum of deformed random matrices and free probability,2016

[1]Giacomo Livan, Marcel Novaes, Pierpaolo Vivo, Introduction to Random Matrices Theory and Practice,2018

[2]Jean-Philippe Bouchaud and Marc Potters, A First Course in Random Matrix Theory (for Physicists, Engineers and Data Scientists) 1st Edition, 2020

[3] Gernot Akemann,Jinho Baik, and Philippe Di Francesc, Oxford Handbook of Random Matrix Theory,2015

[5] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. Ann. of
Math. (2) 67, (1958), 325-327

[6] Y.Q. Yin, Z.D. Bai and P.R. Krishnaiah. On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. Probab. Theory Related Fields 78 (4):509-521,1988.

[7] Z. D. Bai and Y. Q. Yin. Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. Ann. Probab., 16(4), (1988), 1729-1741.

[8] Sinai, Y., Soshnikov, A.: A Refinement of Wigner's Semicircle Law in a Neighborhood of the Spectrum Edge, Functional Anal. and Appl. 32, 114–131 (1998).

[9] Soshnikov, A.: Universality at the Edge of the Spectrum in Wigner Random Matrices, Commun. Math. Phys. 207,697-733 (1999).

[10] Tao, T., Vu, V.: Random Matrices: Universality of Local Eigenvalue Statistics up to the Edge, Commun. Math. Phys.298, 549-572 (2010).

[11] Erd˝os, L, Yau, H.-T., Yin, J.: Rigidity of eigenvalues of generalized Wigner matrices, Adv. Math. 229(3), 1435-1515(2012).

[12] Alt, J., Erd˝os, L, Kr¨uger, T, Schr¨oder, D.: Correlated random matrices: band rigidity and edge universality, Ann. Probab. 48(2), 963-1001 (2020).

[13] Bourgade, P., Erd˝os, L., Yau, H.-T.:Edge universality of beta ensembles, Commun. Math. Phys. 332.1 261-353 (2014).

[14] P´ech´e, S., Soshnikov, A.: On the Lower Bound of the Spectral Norm of Symmetric Random Matrices with Independent Entries, Electron. Commun. Probab. 13, 280–290 (2008).

[15] P´ech´e, S., Soshnikov, A.: Wigner Random Matrices with Non-Symmetrically Distributed Entries, J. Stat. Phys. 129,
857–884 (2007).

[16] Lee, J. O., Yin, J.: A Necessary and Sufficient Condition for Edge Universality of Wigner Matrices, Duke Math. J. 163(1), 117-173, (2014).

[17] Erd˝os, L., Yau, H.-T.: Universality of Local Spectral Statistics of Random Matrices, Bull. Amer. Math. Soc. 49,377-414 (2012).

[18] Erd˝os, L., Yau, H.-T.: A Comment on the Wigner-Dyson-Mehta Bulk Universality Conjecture for Wigner Matrices,Electron. J. Probab. 17, 1-5 (2012).

[19] Oxford Handbook of Random Matrix Theory,2015

[20]Alexander Soshnikov,Universality at the edge of the spectrum in Wigner

random matrices,2003.1
[21] L. A. Pastur. On the spectrum of random matrices. Theoretical and Mathematical Physics, Volume 10, Issue 1, (1972), 67-74.
[22] Z. D. Bai, J. W. Silverstein and Y.Q. Yin. A note on the limit of the largest eigenvalue of a large-dimensional sample covariance matrix. J. Multivariate Anal., 26(2), (1988), 166-168.
[23] S. Geman. A limit theorem for the norm of random matrices. Ann. Probab., 8 (2), (1980) 252-261.
[24] Y.Q. Yin, Z.D. Bai and P.R. Krishnaiah. On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. Probab. Theory Relared Fields 78 (4):509-521, 1988.
[25] Marchenko, V. A., Pastur, L. A.: Distribution of eigenvalues for some sets of random matrices, Matematicheskii Sbornik 114(4), 507-536 (1967).
[26] Geman, S.: A limit theorem for the norm of random matrices, Ann. Probab. 252-261 (1980).
[27] Silverstein, J. W.: On the weak limit of the largest eigenvalue of a large dimensional sample covariance matrix, Journal of Multivariate Analysis 30(2), 307-311 (1989).
[28] Yin, Y. Q., Bai, Z. D., Krishnaiah, P. R.: On the limit of the largest eigenvalue of the large dimensional sample covariance matrix, Probab. Theory Relat. Fields 78(4), 509-521 (1988).
[29]Bai, Z. D,Yin, Y. Q.,Limit of the smallest eigenvalue of a large dimensional sample covariance matrix, the anals of probability,1993,1275-1294
[30] Peche, S.: Universality results for the largest eigenvalues of some sample covariance matrix ensembles, Probab. Theory Relat. Fields 143(3), 481-516 (2009).
[31] Feldheim, O. N., Sodin, S.: A universality result for the smallest eigenvalues of certain sample covariance matrices, Geometric And Functional Analysis 20(1), 88-123 (2010).
[32] Wang, K.: Random covariance matrices: Universality of local statistics of eigenvalues up to the edge, Random Matrices: Theory and Applications 1(01), 1150005 (2012).
[33] Pillai, N., Yin, J.: Universality of covariance matrices, Ann. Appl. Probab. 24(3), 935-1001 (2014).
[34] Ding, X., Yang, F.: A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices, Ann. Appl. Probab. 28(3), 1679-1738 (2018).
[35] Ben Arous, G., P´ech´e, S.: Universality of local eigenvalue statistics for some sample covariance matrices, Comm. Pure Appl. Math. 58, 1-42 (2005).
[36] Tao, T., Vu, V.: Random matrices: The distribution of the smallest singular values, Geometric And Functional Analysis 20(1), 260-297 (2010).
[37]Kevin Schnelli, Yuanyuan Xu, Convergence rate to the Tracy--Widom laws for the largest eigenvalue of sample covariance matrices,2021.8
[38]Sandrine Peche, Universality results for largest eigenvalues of some sample covariance matrix ensembles,2007.5
[39]Alexander Soshnikov, A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices,2001.4

[40]Ji Oon Lee, Kevin Schnelli, Tracy-Widom Distribution for the Largest Eigenvalue of Real Sample Covariance Matrices with General Population,2014.9
[41]Jinwoong Kwak, Ji Oon Lee, Jaewhi Park, Extremal eigenvalues of sample covariance matrices with general population,2020.9
[42] Tao, T., Vu, V.: Random Matrices: Universality of Local Eigenvalue Statistics up to the Edge, Commun. Math. Phys. 298, 549-572 (2010).
[43] Erd˝os, L, Yau, H.-T., Yin, J.: Rigidity of eigenvalues of generalized Wigner matrices, Adv. Math. 229(3), 1435-1515 (2012).
[44] Alt, J., Erdos, L, Kruger, T, Schröder, D.: Correlated random matrices: band rigidity and edge universality, Ann. Probab. 48(2), 963-1001 (2020).
[45] Bourgade, P., Erdos, L., Yau, H.-T.:Edge universality of beta ensembles, Commun. Math. Phys. 332.1 261-353 (2014).
[46] C.E. Porter and N. Rosenzweig, Annals of the Acad. Sci. Fennicae, Serie A VI Physica 44, 1 (1960), reprinted in C.E. Porter, Statistical Theories of Spectra: fluctuations (Academic Press,New York, 1965)
[47] M. Zirnbauer, Symmetry classes in random matrix theory (2004)
---[48] Kevin Schnelli, Yuanyuan Xu, Convergence rate to the Tracy-Widom laws for the largest eigenvalue of Wigner matrices,2021.2
[49] Z. Bao, G. Pan, W. Zhou. Universality for the Largest Eigenvalue of Sample Covariance Matrices with General Population. Ann. Statist. 43.1, (2015), 382-421.
[50] J. O. Lee and K. Schnelli. Edge Universality for Deformed Wigner Matrices. (2014), arXiv:1407.8015.
[51] A. Knowles and J. Yin. Anisotropic local laws for random atrices(2014),arXiv:1410.3516
[52] Xiang-Gen Xia, A Simple Introduction to Free Probability Theory and Its Application to Random Matrices,2019.2
[53] Z. Bai and J. Yao. On sample eigenvalues in a generalized spiked population model. J. Multivariate Anal. 106 (2012), 167-177.
[54] M. Capitaine, C. Donati-Martin, D. Feral : Central limit theorems for eigenvalues of de-formations of Wigner matrices. Ann. Inst. Henri Poincare Probab. Stat. 48 no. 1, (2012), 107-133.
[55] A. Knowles and J. Yin. The outliers of a deformed Wigner matrix. The Annals of Probability Vol. 42, No. 5, (2014), 1980-2031.
[56] A. Pizzo, D. Renfrew, A. Soshnikov. On nite rank deformations of Wigner matrices. Ann. Inst. Henri Poincare Probab. Stat. 49, no. 1, (2013), 64-94
[57] D. Renfrew, A. Soshnikov. On nite rank deformations of Wigner matrices II. Delocalized perturbations Random Matrices Theory Appl. 2 (2013), no. 1, 1250015, 36 pp.
[58] S. Peche. The largest eigenvalue of small rank perturbations of Hermitian random matrices. Probab. Theory Related Fields 134, no. 1, (2006), 127-173
[59] T. Shcherbina. On universality of local edge regime for the deformed Gaussian unitary ensemble. Journal of Statistical Physics 143(3), (2011), 455-481.
[60] K. Johansson. From Gumbel to Tracy-Widom. Probability Theory and Related Fields. 138, (1), (2007), 75-112.
[61] J. O. Lee and K. Schnelli. Extremal eigenvalues and eigenvectors of deformed Wigner ma-trices, Probability Theory and Related Fields pp 1-77 First online: 06

January 2015
[62] J. Baik, G. Ben Arous and S. Peche. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. Ann. Probab. 33 no. 5, (2005), 1643-1697.
[63] W. Hachem, A. Hardy and J. Najim. Large Complex Correlated Wishart Matrices: Fluc-tuations and Asymptotic Independence at the Edges, to appear in Annals of Probability (2015).
[64] Transition from Tracy-Widom to Gaussian fluctuations of extremal eigenvalues of sparse Erdős-Rényi graphs
[65] Ji Oon Lee, Jun Yin,2012.1, A Necessary and Sufficient Condition for Edge Universality of Wigner matrices