Spectrum of deformed random matrices and free probability

Mireille Capitaine, C Donati-Martin

Spectrum of deformed random matrices and free probability

M. Capitaine*, C. Donati-Martin†

Abstract

The aim of this paper is to show how free probability theory sheds light on spectral properties of deformed matricial models and provides a unified understanding of various asymptotic phenomena such as spectral measure description, localization and fluctuations of extremal eigenvalues, eigenvectors behaviour.

1 Introduction

The sum or product of large independent random matrices are models of interest in applications such as mathematical finance, wireless communication, statistical learning (see for example Johnstone [50]). Indeed, in these contexts, one matrix is seen as the signal with significant parameters and the other one as the noise. The general question is to know whether the observation of signal plus noise can give access to the significant parameters.

Pastur [61] and Marchenko-Pastur [60] investigated the limiting spectral distributions of respectively additive deformation of Wigner matrices or multiplicative deformation of Wishart matrices and by the way provide a description of the global asymptotic behavior of the spectrum. Concerning the behavior of largest eigenvalues, the pionner works goes back to Furedi-Komlos [45] where it is shown that for a non centered Wigner matrix, the largest eigenvalue separates from the bulk of eigenvalues and has Gaussian fluctuations. In 2005, Baik, Ben Arous and Péché [12] proved a phase transition phenomenon for the behavior of the largest eigenvalue of the so-called non white Wishart matrices (that is multiplicative perturbations of a Wishart matrix). The multiplicative perturbation they considered is a finite rank perturbation of the identity matrix. The authors pointed out a threshold (depending on the covariance of the sample vectors) giving two different behaviors : either the largest eigenvalue sticks to the bulk and fluctuates according to the Tracy Widom distribution, either the largest eigenvalue separates from the bulk and has Gaussian fluctuations. The phenomenon described is called the BBP transition. Later, a phase transition phenomenon was proved for more general finite rank deformations of both matrices of iid type or unitarily invariant models [13, 64, 25, 26, 58] and for general spiked sample covariance matrices [10]. Some papers investigated eigenvectors of finite rank deformations

*CNRS, Institut de Mathématiques de Toulouse, Université Paul Sabatier, 31062 Toulouse, France

†Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay 78035 Versailles, France

[63, 25, 26].

It is well known since Voiculescu's works [76] that free probability sheds light on the global asymptotic spectrum of the sum or product of large independent random matrices. In the last five years, several articles have highlighted the role of the free analytic subordination functions (introduced by Voiculescu [77, 78], Biane [29]) in the analysis of both global and local behaviour of eigenvalues and eigenspaces of polynomials in asymptotically free random matrices. Thus, the eigenvalue distribution of any polynomial in asymptotically free random matrices can be found via its Cauchy-Stieltjes transform through a direct computation of the subordination functions associated to a free additive convolution of matrix-valued free random variables. We refer to [22] and [20] for the analysis of selfadjoint polynomials, and to [21] for the determination of Brown measure for non-selfadjoint polynomials. More surprisingly, the analytic subordination property for free convolutions turned out to be fundamental in the study of outliers of spiked deformed models (eigenvalues of the deformed matrix converging outside the support of the limiting spectral distribution) and allows the study to be performed on arbitrary classical deformed models, while explaining the first order asymptotic in the BBP phenomenon [37, 19, 32, 33, 70]. The same subordination functions provide (via their derivatives) the limiting behaviour of the eigenvectors associated to the outliers ([25, 26, 32, 19]). Very recently, the local laws of sums of independent unitarily invariant random matrices is studied in ([14]) using subordination.

In this paper, we investigate three classical deformed models. It turns out that various properties of the spectrum, i.e. convergence (or fluctuations) of eigenvalues and eigenvectors, of these models can be presented in an unified way, using the framework of free probability, free convolutions and free subordination functions. Therefore, this paper will focus on free probability and take another look of various results in the literature from this point of view without being exhaustive.

In the rest of this introduction, we describe the three models of deformed random matrices: additive deformation of a Wigner or unitarily invariant matrix, multiplicative deformation of a Wishart type or unitarily invariant matrix and information plus noise type model. We then present the results on the limiting spectral distributions and the seminal works on the convergence of extreme eigenvalues, for finite rank perturbations.

Section 2 is devoted to free probability : we first provide some background on the theory and then focus on free convolutions and subordinations properties which will be central in the study of general deformed models. In Section 3, we give a free probabilistic interpretation of the limiting spectral distribution (LSD) of Section 1 and describe the support of the LSD, in terms of subordinations functions. Section 4 is devoted to the behavior of outliers of general spikes models. The location of these outliers as well as the norm of the projection of corresponding eigenvectors onto the spikes eigengenspaces of the perturbation are described by means of the subordinations functions. In the last section, we describe some results on fluctuations of outliers and eigenvalues at soft edges of the LSD. The first results was obtained for Gaussian models with finite rank deformations. For general deformations, one can get universal result by considering mobile edges of the support of the Girko's type deterministic equivalent measure obtained by replacing the LSD of A_N by its empirical spectral measure in the free convolution describing the limiting LSD of the deformed model. We also provide a free probabilistic interpretation of the appearance of different rates or different asymptotic distributions for fluctuations as uncovered by Johansson [49] and Lee and Schnelli [54].

Notations:
For a Hermitian matrix H of size $N \times N$, we denote by $\{\lambda_1(H) \geq \lambda_2(H) \geq \ldots \geq \lambda_N(H)\}$ the set of eigenvalues of H, ranked in decreasing order, and by $\mu_H = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(H)}$ its spectral measure. The Stieltjes transform of μ_H is defined for $z \in \mathbb{C}\backslash\mathbb{R}$ by

$$g_{\mu_H}(z) = \int_{\mathbb{R}} \frac{1}{z-x} \mathrm{d}\mu_H(x) = \frac{1}{N}\operatorname{Tr}(zI - H)^{-1}.$$

If μ_H converges weakly to some probability measure $\mathbb{P}$ as N goes to infinity, $\mathbb{P}$ is called the limiting spectral measure (LSD) of H and supp($\mathbb{P}$) denotes its support.

1.1 Classical Random matricial models and deformations

We first review classical models in random matrix theory and the pionner works on the behavior of their spectrum.

1.1.1 Wigner matrices

Wigner matrices are real symmetric or complex Hermitian random matrices whose entries are independent (up to the symmetry condition). They were introduced by Wigner in the fifties, in connection with nuclear physics. In this paper, we will consider Hermitian Wigner matrices of the following form :

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

where $(X_N)_{ii}$, $\sqrt{2}\Re e((X_N)_{ij})_{i<j}$, $\sqrt{2}\Im m((X_N)_{ij})_{i<j}$ are i.i.d, with distribution τ with variance σ^2 and mean zero.
If $\tau = \mathcal{N}(0,\sigma^2)$, $W_N =: W_N^G$ is a G.U.E.-matrix. The first result concerns the behavior of the spectral measure.

Theorem 1 ([80, 81])**.**

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

where

$$\frac{d\mu_{sc}}{dx}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\, 1_{[-2\sigma, 2\sigma]}(x)$$

The second result establishes the behavior of the extremal eigenvalues.

Theorem 2 ([11])**.** *If* $\int x^4 \mathrm{d}\tau(x) < +\infty$, *then*

$$\lambda_1(W_N) \to 2\sigma \text{ and } \lambda_N(W_N) \to -2\sigma \text{ a.s when } N \to +\infty.$$

1.1.2 Sample covariance matrices

Random matrices first appeared in mathematical statistics in the 1930s with the works of Hsu, Wishart and others. They considered sample covariance matrices of the form:

$$S_N = \frac{1}{p} X_N X_N^* \tag{1}$$

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

The behavior of the spectral measure for large size ($p = p(N)$ tends to ∞ as N tends to ∞) was handled in the seminal work of Marchenko-Pastur.

Theorem 3 ([60])**.** *If* $c_N := \frac{N}{p(N)} \to c \in]0; 1]$ *when* $N \to \infty$,

$$\mu_{S_N} \to \mu_{\mathrm{MP}} \quad \text{a.s when} \quad N \to +\infty$$

where

$$\mu_{\mathrm{MP}}(dx) = \frac{1}{2\pi c x} \sqrt{(b-x)(x-a)}\, 1_{[a,b]}(x) \mathrm{d}x$$

where $a = (1-\sqrt{c})^2$, $b = (1+\sqrt{c})^2$.

For extremal eigenvalues of sample covariance matrices, we have :

Theorem 4 ([46, 9, 82])**.** *If* $\int x^4 \mathrm{d}\tau(x) < +\infty$,

$$\lambda_1(S_N) \to (1+\sqrt{c})^2 \text{ a.s when } N \to +\infty,$$

$$\lambda_N(S_N) \to (1-\sqrt{c})^2 \text{ a.s when } N \to +\infty.$$

1.1.3 Deformations

Motivated by statistics, wireless communications or imaging one may consider deformations of these models.
Let A_N be a deterministic matrix. One may wonder how the spectrum of a classical random model is impacted by the following perturbations.

i) Additive perturbation of a Wigner matrix : A_N is a $N \times N$ Hermitian matrix and W_N a Wigner matrix (see section 1.1.1),

$$M_N = W_N + A_N.$$

ii) Multiplicative perturbation: A_N is non negative Hermitian $N \times N$ matrix and S_N a sample covariance matrix defined as in section 1.1.2,

$$M_N = A_N^{\frac{1}{2}} S_N A_N^{\frac{1}{2}}.$$

iii) Information-Plus-Noise type model: A_N, X_N are rectangular $N \times p$ matrices with X_N a random matrix with i.i.d. entries defined as in section 1.1.2, σ is some positive real number.

$$M_N = (\sigma \frac{X_N}{\sqrt{p}} + A_N)(\sigma \frac{X_N}{\sqrt{p}} + A_N)^*.$$

In the litterature, these three kinds of deformations have been also considered for isotropic models, 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.

1.2 Convergence of spectral measures

This section is devoted to the study of the limiting spectral measure of the deformed matrix M_N for the different models i) to iii).

1.2.1 Finite rank deformations

When A_N is a finite rank deformation of the null matrix in case i) and iii) (resp. of the identity matrix in case ii)), the limiting spectral measure is not affected by the deformation.

Proposition 1. *We assume that* $\mathrm{rank}(A_N) = r$, *r fixed, independent of N in case i) and iii), (resp.* $\mathrm{rank}(A_N - I) = r$ *in case ii)). Then, when N goes to infinity,*

- *In case i), $\mu_{M_N} \to \mu_{sc}$,*
- *In case ii) and iii), $\mu_{M_N} \to \mu_{\mathrm{MP}}$.*

This follows from the rank inequalities (see [7, Appendix A.6]) where for a matrix A, F^A denotes the cumulative distribution function of the spectral measure μ_A and the norm is the supremum norm:
- Let A and B two $N \times N$ Hermitian matrices. Then,

$$\|F^A - F^B\| \leq \frac{1}{N} \mathrm{rank(A-B)},$$

- Let A and B two $N \times p$ matrices. Then,

$$\|F^{AA^*} - F^{BB^*}\| \leq \frac{1}{N} \mathrm{rank(A-B)}.$$

.

1.2.2 General deformations

We now assume that for some probability ν, the spectrum of the deformation A_N satisfies

$$\mu_{A_N} \underset{N\to+\infty}{\longrightarrow} \nu \qquad \text{for cases i) and ii)}$$
$$\mu_{A_N A_N^*} \underset{N\to+\infty}{\longrightarrow} \nu \qquad \text{for case iii).}$$

The following theorem is a review of the pionner results concerning the limiting spectral distribution, in the three deformed models. The spectral distribution is characterized via an equation satisfied by its Stieltjes transform. In Section 3, we will give a free probabilistic interpretation of these distributions.

Theorem 5. *Convergence of the spectral measure* μ_{M_N}

i) Deformed Wigner matrices ([61], [3] Theorem 5.4.5)

$$\mu_{M_N} \underset{N\to+\infty}{\longrightarrow} \mu_1 \text{ weakly}$$

with

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_1}(z) = \int \frac{1}{z - \sigma^2 g_{\mu_1}(z) - t} \mathrm{d}\nu(t). \tag{2}$$

ii) Sample covariance matrices ([60, 71])

$$\mu_{M_N} \underset{N\to+\infty}{\longrightarrow} \mu_2 \text{ weakly}$$

with

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_2}(z) = \int \frac{1}{z - t(1 - c + czg_{\mu_2}(z))} \mathrm{d}\nu(t). \tag{3}$$

iii) Information-Plus-Noise type matrices ([42])

$$\mu_{M_N} \underset{N\to+\infty}{\longrightarrow} \mu_3 \text{ weakly}$$

with

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_3}(z) = \int \frac{1}{(1 - c\sigma^2 g_{\mu_3}(z))z - \frac{t}{1 - c\sigma^2 g_{\mu_3}(z)} - \sigma^2(1 - c)} \mathrm{d}\nu(t). \tag{4}$$

Remark 1. *The limiting measures μ_1, μ_2, μ_3 are deterministic. Note that they are not always explicit. They are* **universal** *(do not depend on the distribution of the entries of X_N) and only depend on A_N through the limiting spectral measure ν.*

Remark 2. *If $\nu = \delta_0$ in* (2)*, we recover the equation satisfied by the Stieltjes transform of the semicircular distribution. The same is true for the Marchenko-Pastur distribution in* (3) *with $\nu = \delta_1$ and in* (4) *with $\nu = \delta_0$.*

Remark 3. *Such functional equations for Stieltjes transforms have been obtained for deformations of unitarily invariant models by [62] and [73].*

1.3 Convergence of extreme eigenvalues

We now present the seminal works on the behavior of the largest (or smallest) eigenvalues of classical models with finite rank deformation. As we have seen above, the limiting behavior of the spectral measure is not modified by a deformation A_N of finite rank (or such that $I_N - A_N$ is of finite rank in the multiplicative case). This is no longer true for the extremal eigenvalues. The following results deal with finite rank perturbations and Gaussian type or unitarily invariant models : we give the precise statement for the convergence of the largest eigenvalue in the Gaussian case and refer to the original papers in the unitarily invariant case.

1.3.1 Multiplicative deformations

A complete study (convergence, fluctuations) of the behavior of the largest eigenvalue of a deformation of a Gaussian sample covariance matrix was considered in a paper of Baik-Ben Arous and Péché where they exhibit a striking phase transition phenomenon for the largest eigenvalue, according to the value of the spiked eigenvalues of the deformation. They considered the following sample covariance matrix :

$$M_N = A_N^{1/2} S_N^G A_N^{1/2}$$

where S_N^G is a L.U.E. matrix as defined in Section 1.1.2 and the perturbation A_N[1] is given by

$$A_N = \text{diag}\ (\underbrace{1, \ldots, 1}_{N-r \text{ times}}, \pi_1, \ldots, \pi_r)$$

where r is fixed, independent of N and $\pi_1 \geq \ldots \geq \pi_r > 0$ are fixed, independent of N, such that for all $i \in \{1, \ldots, r\}$, $\pi_i \neq 1$. The π_i's are called the spikes of A_N.

Theorem 6. *(BBP phase transition)[12, 13]*
Let $\omega_c = 1 + \sqrt{c}$,

- *If $\pi_1 > \omega_c$, a.s when $N \to +\infty$*

$$\lambda_1(M_N) \to \pi_1 \left(1 + \frac{c}{(\pi_1 - 1)}\right) > (1 + \sqrt{c})^2.$$

Therefore the largest eigenvalue of M_N is an "outlier" since it converges outside the support of the limiting empirical spectral distribution and then does not stick to the bulk.

- *If $\pi_1 \leq \omega_c$, a.s when $N \to +\infty$*

$$\lambda_1(M_N) \to (1 + \sqrt{c})^2.$$

[1] In the Gaussian case, we may assume, without loss of generality, that the perturbation is diagonal

The same phenomenon of phase transition was established by Benaych-Georges and Nadakuditi [25] in the case of a deformation of an unitarily invariant model of the form :

$$M_N = A_N^{1/2} U_N B_N U_N^* A_N^{1/2}, \tag{5}$$

where

- B_N is a deterministic $N \times N$ Hermitian non negative definite matrix such that:
- $\mu_{B_N} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(B_N)}$ weakly converges to some probability measure μ with compact support $[a, b]$.
- the smallest and largest eigenvalues of B_N converge to a and b.
- U_N is a random $N \times N$ unitary matrix distributed according to Haar measure.
- $A_N - I_N$ is a deterministic Hermitian matrix having r non-zero eigenvalues $\theta_1 \geq \ldots \geq \theta_r > -1$ r, $\theta_i, i = 1, \ldots r$, fixed independent of N .

1.3.2 Additive deformations

Let us consider $M_N = W_N^G + A_N$ where W_N^G is a G.U.E. matrix as defined in Section 1.1.1 and A_N is defined by

$$A_N = \text{diag}\,(\underbrace{0, \ldots, 0}_{N-r \text{ times}}, \theta_1, \ldots, \theta_r)$$

for some fixed r, independent of N, and some fixed $\theta_1 \geq \cdots \geq \theta_r$, independent of N.
An analog of the BBP phase transition phenomenon for this model was obtained by Péché [64].

Theorem 7.
- *If $\theta_1 \leq \sigma$, then $\lambda_1(M_N) \underset{N\to+\infty}{\longrightarrow} 2\sigma$ a.s..*
- *If $\theta_1 > \sigma$, then $\lambda_1(M_N) \underset{N\to+\infty}{\longrightarrow} \rho_{\theta_1}$ a.s. with $\rho_{\theta_1} := \theta_1 + \frac{\sigma^2}{\theta_1} > 2\sigma$.*

A similar result has been obtained by Benaych-Georges and Nadakuditi [25] for

$$M_N = U_N B_N U_N^* + A_N,$$

where U_N, B_N satisfy the same hypothesis as in (5) and A_N is a finite rank perturbation of the null matrix.

1.3.3 Information plus noise type matrices

Such a phase transition phenomenon was established by Loubaton and Vallet in [58] for the singular values of a finite rank deformation of a Ginibre ensemble.
Let X_N be a $N \times p$ rectangular matrix as defined in Section 1.1.2 with iid complex Gaussian entries, and A_N be a finite rank perturbation of the null matrix with non zero eigenvalues θ_i.

Theorem 8. *([58]) Let* $M_N = (\sigma \frac{X_N}{\sqrt{p}} + A_N)(\sigma \frac{X_N}{\sqrt{p}} + A_N)^*$*. Then, as* $N \to +\infty$ *and* $N/p \to c \in]0;1]$ *, almost surely,*

$$\lambda_i(M_N) \quad \longrightarrow \begin{cases} \frac{(\sigma^2+\theta_i)(\sigma^2 c+\theta_i)}{\theta_i} & \text{if } \theta_i > \sigma^2\sqrt{c}, \\ \sigma^2(1+\sqrt{c})^2 & \text{otherwise.} \end{cases}$$

Again, this result was extended in [26] to the case $M_N = (V_N + A_N)(V_N + A_N)^*$ where V_N is a biunitarily invariant matrix such that the empirical spectral measure of $V_N V_N^*$ converges to a deterministic compactly supported measure μ with convergence of the largest (resp. smallest) eigenvalue of $V_N V_N^*$ to the right (resp. left) end of the support of μ.

1.4 Eigenvectors associated to outliers

Now, one can wonder in the spiked deformed models, when some eigenvalues separate from the bulk, how the corresponding eigenvectors of the deformed model project onto those of the perturbation. There are some pionneering results concerning finite rank perturbations: [63] in the real Gaussian sample covariance matrix setting, and [25, 26] dealing with finite rank additive or multiplicative perturbations of unitarily invariant matrices. Here is the result for eigenvector projection corresponding to the largest outlier for finite rank additive perturbation.

Theorem 9. *Let* $M_N = U_N B_N U_N^* + A_N$ *where* U_N *is a Haar unitary matrix,* B_N *satisfies the same hypothesis as in (5) and* A_N *has all but finitely many non zero eigenvalues* $\theta_1 > \ldots > \theta_J$*. Then, if* $\theta_1 > 1/\lim_{z \downarrow b} g_\mu(z)$*, almost surely,*

$$\lambda_1(M_N) \underset{N\to+\infty}{\longrightarrow} g_\mu^{-1}(1/\theta_1) := \rho_{\theta_1}$$

and for any $i = 1, \ldots, J$*, if* ξ *is a unit eigenvector associated to* $\lambda_1(M_N)$*,*

$$\|P_{\mathrm{Ker}(\theta_i I - A)}\xi\|^2 \underset{N\to+\infty}{\longrightarrow} -\frac{\delta_{i1}}{\theta_1^2 g_\mu'(\rho_{\theta_1})}.$$

It turns out that the results of the above Section may be interpreted in terms of free probability theory and this very analysis allows to extend the results of Subsections 1.3 and 1.4 to non-finite rank deformations. Therefore, in the next section, we start by recalling some background in free probability theory for the reader's convenience.

2 Free probability theory

We refer to [79] for an introduction to free probability theory.

A non-commutative probability space is a unital algebra $\mathcal{A}$ over $\mathbb{C}$, endowed with a linear functional $\phi : A \to \mathbb{C}$ such that $\phi(1) = 1$. Elements of $\mathcal{A}$ are called non-commutative random variables.

If $(a_i)_{i=1,\ldots,q}$ is a family of non-commutative random variables in $(\mathcal{A},\phi)$, the distribution $\mu_{(a_i)_{i=1,\ldots,q}}$ of $(a_i)_{i=1,\ldots,q}$ is the linear functional on the algebra $\mathbb{C}\langle X_i | i=1,\ldots,q\rangle$ of polynomials in the non commutating variables $(X_i)_{i=1,\ldots,q}$ given by

$$\mu_{(a_i)_{i=1,\ldots,q}}(P) = \phi\left(P\left(\mu_{(a_i)_{i=1,\ldots,q}}\right)\right).$$

If $\mathcal{A}$ is a C^*-algebra endowed with a state ϕ, then for any selfadjoint element a in $\mathcal{A}$, there exists a measure ν_a on $\mathbb{R}$ such that, for every polynomial P, we have

$$\mu_a(P) = \int P(t)\mathrm{d}\nu_a(t).$$

Then we identify μ_a and ν_a.

A family of unital subalgebras $(\mathcal{A}_i)_{i=1,\ldots q}$ in $(\mathcal{A},\phi)$ is freely independent if for every $p \geq 1$, for every $(a_1,\ldots,a_p)$ such that, for every k in $\{1,\ldots,p\}$, $\phi(a_k) = 0$ and a_k is in $\mathcal{A}_{i(k)}$ for some $i(k)$ in $\{1,\ldots,q\}$ with $i(k) \neq i(k+1)$, then $\phi(a_1,\ldots,a_p) = 0$. Random variables are free in $(\mathcal{A},\phi)$ if the subalgebras they generate with 1 are freely independent.

For each n in $\mathbb{N}\setminus\{0\}$, let $(a_i^n)_{i=1,\ldots,q}$ be a family of noncommutative random variables in a noncommutative probability space $(\mathcal{A}_n,\phi_n)$. The sequence of joint distributions $\mu_{(a_i^n)_{i=1,\ldots,q}}$ converges as n tends to $+\infty$, if there exists a distribution μ such that $\mu_{(a_i^n)_{i=1,\ldots,q}}(P)$ converges to $\mu(P)$ as n tends to $+\infty$ for every P in $\mathbb{C}\langle X_i | i=1,\ldots,q\rangle$. μ is called the limit distribution of $(a_i^n)_{i=1,\ldots,q}$. If $(a_i)_{i=1,\ldots,q}$ is a family of noncommutative random variables with distribution μ, we also say that $(a_i^n)_{i=1,\ldots,q}$ converges towards $(a_i)_{i=1,\ldots,q}$. A family of noncommutative random variables $(a_i^n)_{i=1,\ldots,q}$ is said to be asymptotically free as n tends to ∞ if it has a limit distribution μ and if $(X_1,\ldots,X_q)$ are free in $(\mathbb{C}\langle X_i | i=1,\ldots,q\rangle,\mu)$.

Additive and multiplicative free convolutions arise as natural analogues of classical convolutions in the context of free probability theory. For two Borel probability measures μ and ν on the real line, one defines the free additive convolution $\mu \boxplus \nu$ as the distribution of $a+b$, where a and b are free self-adjoint random variables with distributions μ and ν, respectively. Similarly, if both μ,ν are supported on $[0,+\infty)$, their free multiplicative convolution $\mu \boxtimes \nu$ is the distribution of the product ab, where, as before, a and b are free positive random variables with distributions μ and ν, respectively. The product ab of two free positive random variables is usually not positive, but it has the same moments as the positive random variables $a^{1/2}ba^{1/2}$ and $b^{1/2}ab^{1/2}$. We refer to [74, 75, 59, 27] for the definitions and main properties of free convolutions. In the following sections, we briefly recall the analytic approach developed in [74, 75] to calculate the additive and multiplicative free convolutions of compactly supported measures and the analytical definition of the rectangular free convolution introduced by F. Benaych-Georges in [23]. Finally, we present the fundamental analytic subordination properties [77, 29, 78, 17] of these three convolutions.

2.1 Free convolution

2.1.1 Free additive convolution

The Stieltjes transform of a compactly supported probability measure μ is conformal in the neighborhood of ∞, and its functional inverse g_μ^{-1} is meromorphic at zero with principal part $1/z$. The R-transform [74] of μ is the convergent power series defined by

$$R_\mu(z) = g_\mu^{-1}(z) - \frac{1}{z}.$$

The free additive convolution of two compactly supported probability measures μ and ν is another compactly supported probability measure characterized by the identity

$$R_{\mu \boxplus \nu} = R_\mu + R_\nu$$

satisfied by these convergent power series.

2.1.2 Multiplicative free convolution on $[0, +\infty)$

Recall the definition of the moment-generating function of a Borel probability measure μ on $[0, +\infty)$:

$$\psi_\mu(z) = \int_{[0,+\infty)} \frac{zt}{1 - zt} \, \mathrm{d}\mu(t), \quad z \in \mathbb{C} \setminus \left\{ z \in \mathbb{C} \colon \frac{1}{z} \in \operatorname{supp}(\mu) \right\}.$$

This function is related to the Cauchy-Stieltjes transform of μ via the relation

$$\psi_\mu(z) = \frac{1}{z} g_\mu\left(\frac{1}{z}\right) - 1.$$

Recall also the so-called eta transform

$$\eta_\mu(z) = \frac{\psi_\mu(z)}{1 + \psi_\mu(z)}.$$

The Σ-transform [75, 27] of a compactly supported Borel probability measure $\mu \neq \delta_0$ is the convergent power series defined by

$$\Sigma_\mu(z) = \frac{\eta_\mu^{-1}(z)}{z},$$

where η_μ^{-1} is the inverse of η_μ relative to composition. The free multiplicative convolution of two compactly supported probability measures $\mu \neq \delta_0 \neq \nu$ is another compactly supported probability measure characterized by the identity

$$\Sigma_{\mu \boxtimes \nu}(z) = \Sigma_\mu(z) \Sigma_\nu(z)$$

in a neighbourhood of 0.

2.1.3 Rectangular free convolution

Let c be in $]0;1]$. Let τ be a probability measure on $\mathbb{R}^+$. Define for z in $\mathbb{C}\setminus[0;+\infty[$,

$$M_\tau(z)=\int_{\mathbb{R}^+}\frac{t^2z}{1-t^2z}\mathrm{d}\tau(t),\quad H_\tau^{(c)}(z):=z\,(cM_\tau(z)+1)\,(M_\tau(z)+1),$$
$$\text{and}\quad T^{(c)}(z)=(cz+1)(z+1).$$

The transform C_τ [23] defined as follows is called the rectangular R-transform:

$$C_\tau^{(c)}(z)=T^{(c)^{-1}}\left(\frac{z}{H_\tau^{(c)^{-1}}(z)}\right),\quad \text{for}\quad z\quad \text{small enough.}$$

The rectangular free convolution with ratio c of two probability measures μ and ν on $\mathbb{R}^+$ is the unique probability measure on $\mathbb{R}^+$ whose rectangular R-transform is the sum of the rectangular R-transforms of μ and ν, and it is denoted by $\mu\boxplus_c\nu$. Then, we have for z small enough,

$$C_{\mu\boxplus_c\nu}(z)=C_\mu(z)+C_\nu(z).$$

2.2 Analytic subordinations

The analytic subordination phenomenon for free convolutions was first noted by Voiculescu in [77] for free additive convolution of compactly supported probability measures. Biane [29] extended the result to free additive convolutions of arbitrary probability measures on $\mathbb{R}$, and also found a subordination result for multiplicative free convolution. A new proof was given later, using a fixed point theorem for analytic self-maps of the upper half-plane [18]. Note that such a subordination property allows to give a new definition of free additive convolution [39]. Finally, S. Belinschi, F. Benaych-Georges, and A. Guionnet [17] established such a phenomenon for the rectangular free convolution.

2.2.1 Free additive subordination property

Let us define the reciprocal Cauchy-Stieltjes transform $J_\mu(z)=1/g_\mu(z)$, which is an analytic self-map of the upper half-plane. Given Borel probability measures μ and ν on $\mathbb{R}$, there exist two unique analytic functions $\omega_1,\omega_2\colon\mathbb{C}^+\to\mathbb{C}^+$ such that

1. $\lim_{y\to+\infty}\omega_j(iy)/iy=1$, $j=1,2$;
2. $$\omega_1(z)+\omega_2(z)-z=J_\mu(\omega_1(z))=J_\nu(\omega_2(z))=J_{\mu\boxplus\nu}(z),\quad z\in\mathbb{C}^+.\tag{6}$$
3. In particular (see [18]), for any $z\in\mathbb{C}^+\cup\mathbb{R}$ so that ω_1 is analytic at z, $\omega_1(z)$ is the attracting fixed point of the self-map of $\mathbb{C}^+$ defined by
$$w\mapsto J_\nu(J_\mu(w)-w+z)-(J_\mu(w)-w).$$
A similar statement, with μ,ν interchanged, holds for ω_2.

In particular, according to (6), we have for any $z \in \mathbb{C}^+$,

$$g_{\mu \boxplus \nu}(z) = g_\mu(\omega_1(z)) = g_\nu(\omega_2(z)). \tag{7}$$

2.2.2 Multiplicative subordination property

Given Borel probability measures μ, ν on $[0, +\infty)$, there exist two unique analytic functions $F_1, F_2 \colon \mathbb{C} \setminus [0, +\infty) \to \mathbb{C} \setminus [0, +\infty)$ so that

1. $\pi > \arg F_j(z) \geq \arg z$ for $z \in \mathbb{C}^+$ and $j = 1, 2$;
2. $$\frac{F_1(z)F_2(z)}{z} = \eta_\mu(F_1(z)) = \eta_\nu(\omega_2(z)) = \eta_{\mu \boxtimes \nu}(z), \quad z \in \mathbb{C} \setminus [0, +\infty). \tag{8}$$
3. In particular (see [18]), for any $z \in \mathbb{C}^+ \cup \mathbb{R}$ so that F_1 is analytic at z, the point $h_1(z) := F_1(z)/z$ is the attracting fixed point of the self-map of $\mathbb{C} \setminus [0, +\infty)$ defined by
$$w \mapsto \frac{w}{\eta_\mu(zw)} \eta_\nu \left(\frac{\eta_\mu(zw)}{w} \right).$$
A similar statement, with μ, ν interchanged, holds for ω_2.

In particular (8) yields

$$\psi_{\mu \boxtimes \nu}(z) = \psi_\mu(F_1(z)) = \psi_\nu(F_2(z)). \tag{9}$$

2.2.3 Rectangular free subordination property

Let c be in $]0; 1]$. Let μ and ν be two probability measures on $\mathbb{R}^+$. Assume that the rectangular R-transform $C_\mu^{(c)}$ of μ extends analytically to $\mathbb{C} \backslash \mathbb{R}^+$; this happens for example if μ is $\boxplus_c$ infinitely divisible. Then there exist two unique meromorphic functions Ω_1, Ω_2 on $\mathbb{C} \setminus \mathbb{R}^+$ so that

$$H_\mu^{(c)}(\Omega_1(z)) = H_\nu^{(c)}(\Omega_2(z)) = H_{\mu \boxplus_c \nu}^{(c)}(z),$$

$\Omega_j(\overline{z}) = \overline{\Omega_j(z)}$ and $\lim_{x \uparrow 0} \Omega_j(x) = 0$, $j \in \{1; 2\}$.

The fonctions ω_i, F_i and Ω_i in the three previous subsections are called subordination functions.

2.3 Asymptotic freeness of independent random matrices

Free probability theory and random matrix theory are closely related. Indeed the purely algebraic concept of free relation of noncommutative random variables can be also modeled by random matrix ensembles if the matrix size goes to infinity. Let $(\Omega, \mathcal{F}, P)$ be a classical probability space and for every $N \geq 1$, let us denote by $\mathcal{M}_N$ the algebra of complex $N \times N$ matrices. Let $\mathcal{A}_N$ be the algebra of $N \times N$ random matrices $(\Omega, \mathcal{F}, P) \to \mathcal{M}_N$. Define

$$\Phi_N : \begin{cases} \mathcal{A}_N \to \mathbb{C} \\ A \to \frac{1}{N} \operatorname{Tr} A \end{cases}$$

For each $N \geq 1$, $(\mathcal{A}_N, \Phi_N)$ is a non-commutative probability space and we will consider random matrices in this non-commutative context.
In his pioneering work [76], Voiculescu shows that independent Gaussian Wigner matrices converge in distribution as their size goes to infinity to free semi-circular variables. The following result from Corollary 5.4.11 [3] extends Voiculescu's seminal observation.

Theorem 10. *Let $\{D_N(i)\}_{1 \leq i \leq p}$ be a sequence of uniformly bounded real diagonal matrices with empirical measure of diagonal elements converging to μ_i, $i = 1, \ldots, p$ respectively. Let $\{U_N(i)\}_{1 \leq i \leq p}$ be independent unitary matrices following the Haar measure, independent from $\{D_N(i)\}_{1 \leq i \leq p}$.*

- *The noncommutative variables $\{U_N(i) D_N(i) U_N(i)^*\}_{1 \leq i \leq p}$ in the noncommutative probability space $(\mathcal{A}_N, \Phi_N)$ are almost surely asymptotically free, the law of the marginals being given by the μ_i's.*
- *The spectral distribution of $D_N(1) + U_N(2) D_N(2) U_N(2)^*$ converges weakly almost surely to $\mu_1 \boxplus \mu_2$ goes to infinity.*
- *Assume that $D_N(1)$ and $D_N(2)$ are nonnegative. Then, the empirical spectral measure of $(D_N(1))^{\frac{1}{2}} U_N(2) D_N(2) U_N(2)^* (D_N(1))^{\frac{1}{2}}$ converges weakly almost surely to $\mu_1 \boxtimes \mu_2$ as N goes to infinity.*

Thus, if μ is the eigenvalue distribution of a large selfadjoint random matrix A and ν is the eigenvalue distribution of a large selfadjoint random matrix B then, when A and B are in generic position, $\mu \boxplus \nu$ is nearly the eigenvalue distribution of $A + B$. Similarly, dealing with nonnegative matrices $\mu \boxtimes \nu$ is nearly the eigenvalue distribution of $A^{\frac{1}{2}} B A^{\frac{1}{2}}$.

Similarly, for independent rectangular $n \times N$ random matrices A and B such that $n/N \to c \in]0; 1]$, when A and B are in generic position, F. Benaych-Georges [23] proved that rectangular free convolution with ratio c provides a good understanding of the asymptotic global behaviour of the singular values of $A + B$.

Theorem 11. *Let A and B be independent rectangular $N \times p$ random matrices such that A or B is invariant, in law, under multiplication, on the right and on the left, by any unitary matrix. Assume that there exists two laws μ and ν such that, for the weak convergence in probability, we have $\displaystyle\frac{1}{N} \sum_{s \text{ sing. val. of } A_N} \delta_s \to \mu$, $\displaystyle\frac{1}{N} \sum_{s \text{ sing. val. of } B_N} \delta_s \to \nu$ as n and N goes to infinity with*

$N/p \to c \in]0;1]$. *Then*

$$\frac{1}{N} \sum_{s \text{ sing. val. of } A_N+B_N} \delta_s \underset{N\to+\infty}{\longrightarrow} \mu \boxplus_c \nu,$$

for the weak convergence in probability.

3 Study of LSD of deformed ensembles through free probability theory

In this section, we take an other look of the LSD described in Section 1.2, in the light of free probability theory introduced in the previous Section. We also characterize the limiting supports in terms of free subordination functions. This will prove to be fundamental to understand the outliers phenomenon for spiked models in Section 4. We finish with an analysis of the different behaviors of the density at edges of the support of free additive, multiplicative, rectangular convolutions with semi-circular, Marchenko-Pastur and the square-root of Marchenko-Pastur distributions respectively. This provides the rate for fluctuations of eigenvalues at edges as it will discussed in Section 5.

3.1 Free probabilistic interpretation of LSD

As noticed in Remark 1, the limiting spectral distributions of the deformed models investigated in Section 1.1.3 are **universal** in the sense that they do not depend on the distribution of the entries of X_N. Therefore, choosing Gaussian entries and applying Theorem 10 and Theorem 11, we readily get the following free probabilistic interpretation of the limiting measures as well as of the equations satisfied by the limiting Stieltjes transforms in Theorem 5.

- Deformed Wigner matrices

$$\mu_{M_N} \underset{N\to+\infty}{\longrightarrow} \mu_1 \text{ weakly}, \quad \mu_1 = \mu_{sc} \boxplus \nu$$

- Sample covariance matrices

$$\mu_{M_N} \underset{N\to+\infty}{\longrightarrow} \mu_2 \text{ weakly}, \quad \mu_2 = \mu_{\mathrm{MP}} \boxtimes \nu$$

- Information-Plus-Noise type matrices

$$\mu_{M_N} \underset{N\to+\infty}{\longrightarrow} \mu_3 \text{ weakly}, \quad \mu_3 = (\sqrt{\mu_{\mathrm{MP}}} \boxplus_c \sqrt{\nu})^2.$$

The equations (2), (3) and (4) satisfied by the limiting Stieltjes transforms correspond to free subordination properties and exhibit the subordination functions $\omega_{\mu_{sc},\nu}$ with respect to the semi-circular distribution μ_{sc} for the free additive convolution, $F_{\mu_{\mathrm{MP}},\nu}$ with respect to the Marchenko-Pastur distribution μ_{MP} for the free multiplicative convolution, and $\Omega_{\mu_{\mathrm{MP}},\nu}$ with respect to the pushforward of the Marchenko-Pastur distribution by the square root function $\sqrt{\mu_{\mathrm{MP}}}$ for the rectangular free convolution.

- Deformed Wigner matrices

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_1}(z) = \int \frac{1}{z - \sigma^2 g_{\mu_1}(z) - t} \mathrm{d}\nu(t) = g_\nu(\omega_{\mu_{sc},\nu}(z)).$$

$$\omega_{\mu_{sc},\nu}(z) = z - \sigma^2 g_{\mu_1}(z).$$

- Sample covariance matrices

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_2}(z) = \int \frac{1}{z - t(1 - c + czg_{\mu_2}(z))} \mathrm{d}\nu(t).$$

$$\to \qquad \psi_{\mu_1}\left(\frac{1}{z}\right) = \psi_\nu(F_{\mu_{\mathrm{MP}},\nu}\left(\frac{1}{z}\right))$$

$$\psi_\tau(z) = \int \frac{tz}{1 - tz} d\tau(t) = \frac{1}{z} g_\tau(\frac{1}{z}) - 1,$$

$$F_{\mu_{\mathrm{MP}},\nu}(z) = z - cz + cg_{\mu_1}(\frac{1}{z}).$$

- Information-Plus-Noise type matrices

$$\mu_3 = (\sqrt{\mu_{\mathrm{MP}}} \boxtimes_c \sqrt{\nu})^2$$

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_3}(z) = \int \frac{1}{(1 - c\sigma^2 g_{\mu_3}(z))z - \frac{t}{1 - c\sigma^2 g_{\mu_3}(z)} - \sigma^2(1 - c)} \mathrm{d}\nu(t).$$

$$\to \quad H^{(c)}_{\sqrt{\mu_3}}\left(\frac{1}{z}\right) = H^{(c)}_{\sqrt{\nu}}\left(\Omega_{\mu_{\mathrm{MP}},\nu}\left(\frac{1}{z}\right)\right)$$

$$H^{(c)}_{\sqrt{\tau}}(z) = \frac{c}{z} g_\tau(\frac{1}{z})^2 + (1 - c) g_\tau(\frac{1}{z}),$$

$$\Omega_{\mu_{\mathrm{MP}},\nu}(z) = \frac{1}{\frac{1}{z}(1 - c\sigma^2 g_{\mu_3}(\frac{1}{z}))^2 - (1 - c)\sigma^2(1 - c\sigma^2 g_{\mu_3}(\frac{1}{z}))}.$$

3.2 Limiting supports of LSD

For each deformed model introduced in Section 1.1.3 involving i.i.d entries, several authors studied the limiting support [28, 40, 43, 57, 58, 33].

It turns out that in each case there is a one to one correspondance involving the subordination functions between the complement of the support of the limiting spectral measure and some set in the complement of the limiting support of the deformation, as follows.

- Deformed Wigner $\mu_1 = \mu_{sc} \boxplus \nu$

$$\mathbb{R} \setminus \mathrm{supp}(\mu_1) \underset{\phi_1}{\overset{\varphi_1}{\rightleftarrows}} \mathcal{O}_1 \subset \mathbb{R} \setminus \mathrm{supp}(\nu),$$

$$\mathcal{O}_1 := \{u \in \mathbb{R} \setminus \mathrm{supp}(\nu), \phi_1'(u) > 0\}$$

$$u \in \mathbb{R} \setminus \mathrm{supp}(\nu), \quad \phi_1(u\ ) = u + \sigma^2 g_\nu(u).$$

$$x \in \mathbb{R} \setminus \mathrm{supp}(\mu_1), \quad \varphi_1(x\ ) = x - \sigma^2 g_{\mu_1}(x).$$

- Sample covariance matrices $\mu_2 = \mu_{\mathrm{MP}} \boxtimes \nu$

$$\mathbb{R} \setminus \{\mathrm{supp}(\mu_2)\} \underset{\phi_2}{\overset{\varphi_2}{\rightleftarrows}} \mathcal{O}_2 \subset \mathbb{R} \setminus \{\mathrm{supp}(\nu)\}, \tag{10}$$

$$\mathcal{O}_2 = \left\{u \in^c \{\mathrm{supp}(\nu)\}, \, \phi_2'(u) > 0\right\}$$

$$u \in \mathbb{R} \setminus \mathrm{supp}(\nu), \quad \phi_2(u) = u + cu \int \frac{t}{u-t} \mathrm{d}\nu(t).$$

$$x \in \mathbb{R} \setminus \mathrm{supp}(\mu_2), \quad \varphi_2(x) = \begin{cases} \frac{x}{(1-c)+cxg_{\mu_2}(x)} & \text{if } \ c < 1 \\ \frac{1}{g_{\mu_2}(x)} & \text{if } \ c = 1. \end{cases}$$

Note that φ_2 is well defined on $\mathbb{R} \setminus \mathrm{supp}(\mu_2)$ since its denominator never vanishes according to Lemma 6.1 in [7].

- Information-Plus-Noise type model $\mu_3 = (\sqrt{\mu_{\mathrm{MP}}} \boxtimes_c \sqrt{\nu})^2$

$$\mathbb{R} \setminus \mathrm{supp}(\mu_3) \underset{\phi_3}{\overset{\varphi_3}{\rightleftarrows}} \mathcal{O}_3 \subset \mathbb{R} \setminus \mathrm{supp}(\nu),$$

$$\mathcal{O}_3 = \left\{ u \in \mathbb{R} \setminus \mathrm{supp}(\nu), \phi_3'(u) > 0, g_\nu(u) > -\frac{1}{\sigma^2 c} \right\}.$$

$$u \in \mathbb{R} \setminus \mathrm{supp}(\nu), \quad \phi_3(u) = u(1 + c\sigma^2 g_\nu(u))^2 + \sigma^2(1-c)(1 + c\sigma^2 g_\nu(u)) \tag{11}$$

$$x \in \mathbb{R} \setminus \mathrm{supp}(\mu_3), \quad \varphi_3(x) = x(1 - c\sigma^2 g_{\mu_3}(x))^2 - (1-c)\sigma^2(1 - c\sigma^2 g_{\mu_3}(x))$$

Note that φ_1 corresponds to the extension of $\omega_{\mu_{sc},\nu}$ on $\mathbb{R} \setminus \mathrm{supp}(\mu_1)$ and for $i = 2, 3$, φ_i coincides on $\mathbb{R} \setminus \{\mathrm{supp}(\mu_i) \cup \{0\}\}$, with the extension of $z \mapsto 1/F_{\mu_{\mathrm{MP}},\nu}(1/z)$ and $z \mapsto 1/\Omega_{\mu_{\mathrm{MP}},\nu}(1/z)$ respectively.

The above characterization of the support are explicitely given in [28, 33] for μ_1 and μ_3. Now, it can be deduced for μ_2 by the following arguments. In a W^*-probability space endowed with a faithful state, the support of the distribution of a random variable x corresponds to the spectrum of x. Thus, considering μ_2 as the distribution of $b^{1/2}ab^{1/2}$ where a and b are free bounded operators whose distributions are μ_{MP} and ν respectively, one can easily see that for $c < 1$, 0 belongs to the support of $\mu_{\mathrm{MP}} \boxtimes \nu$ if and only if 0 belongs to the support of ν. The latter equivalence and Lemma 6.1 in [7] readily yield (10).

When the support of ν has a finite number of connected components, we have the following description of the support of the μ_i's in terms of a finite union of closed disjoint intervals.

Theorem 12. *[37, 33] Assume that the support of ν is a finite union of disjoint (possibly degenerate) closed bounded intervals. For any $i = 1, 3$, there exists a nonnul integer number p and $u_1 < v_1 < u_2 < \ldots < u_p < v_p$ (depending on i) such that*

$$\mathcal{O}_i =]-\infty, u_1[\cup_{l=1}^{p-1}]v_l, u_{l+1}[\cup]v_p, +\infty[.$$

We have

$$supp(\nu) \subset \cup_{l=1}^{p} [u_l, v_l]$$

and for each $l \in \{1, \ldots, p\}$, $[u_l, v_l] \cap supp(\nu) \neq \emptyset$.
Moreover,

$$supp(\mu_i) = \cup_{l=1}^{p} [\phi_i(u_l^-), \phi_i(v_l^+)],$$

with

$$\phi_i(u_1^-) < \phi_i(v_1^+) < \phi_i(u_2^-) < \phi_i(v_2^+) < \cdots < \phi_i(u_p^-) < \phi_i(v_p^+),$$

where $\phi_i(u_l^-) = \lim_{u \uparrow u_l} \phi_i(u)$ and $\phi_i(v_l^+) = \lim_{u \downarrow v_l} \phi_i(u)$.

Finally, for each $l \in \{1, \ldots, p\}$,

$$\mu_i([\phi_i(u_l^-), \phi_i(v_l^+)]) = \nu([u_l, v_l]). \tag{12}$$

Using the characterization of the support (10), Remark 3.6 in [32] and the fact that from [16] the only possible mass of $\mu_2 = \mu_{\mathrm{MP}} \boxtimes \nu$ is at zero, one may check that the above result still holds for μ_2 allowing $u_1 = v_1 = 0$ or $\phi_2(u_1) = \phi_2(v_1) = 0$ in Theorem 12. Note that the latter cases occur only when ν has a Dirac mass at zero since from [16], $\mu_{\mathrm{MP}} \boxtimes \nu(\{0\}) = \max(\mu_{\mathrm{MP}}(\{0\}), \nu(\{0\}))$ and therefore, since $c \leq 1$, μ_2 has a Dirac mass at zero if and only if ν has a Dirac mass at zero. (12) can be seen as a consequence of the matricial exact separation phenomenon described in Section 4.1.1 b) below letting N go to infinity.

3.2.1 Behavior of the density at edges

P. Biane proved in [28] that $\mu_1 = \mu_{sc} \boxplus \nu$ has a continuous density. Choi and Silverstein [40] and Dozier and Silverstein [43] proved respectively that, away from zero, $\mu_2 = \mu_{\mathrm{MP}} \boxtimes \nu$ and

$\mu_3 = (\sqrt{\mu_{\rm MP}} \boxtimes_c \sqrt{\nu})^2$ possess a continuous density. Let us denote any of theses densities by p. Using the notations of Theorem 12, we have

$$\mathrm{supp}(\nu) \subset \cup_{l=1}^{p}[u_l; v_l] = \mathbb{R} \setminus \mathcal{O}_i$$

and for each $l \in \{1, \ldots, p\}$, $[u_l, v_l] \cap \mathrm{supp}(\nu) \neq \emptyset$. If $a = u_l$ or v_l are not in $\mathrm{supp}(\nu)$ that is if $\mathrm{supp}(\nu)$ does not stick to the frontier of $\mathbb{R} \setminus \mathcal{O}_i$ at these points, then the previous authors established that the density exhibits behavior closely resembling that of $\sqrt{|x-d|}$ for x near $d = \phi_i(a)$. We will say that such an edge $\phi_i(a)$ is regular. This is for instance obviously always the case dealing with a discrete measure ν.
Nevertheless for some measures ν with a density decreasing quite fast to zero at an edge of the support of ν, such an edge may coincide with some u_l or v_l, that is $\mathrm{supp}(\nu)$ may stick to the frontier of $\mathbb{R} \setminus \mathcal{O}_i$ at this point. Then, at the corresponding edge of the support of μ_i, the density p may exhibit different behaviour. This can be seen for instance in the following example investigated by Lee and Schnelli [53] :

$$d\nu(x) := Z^{-1}(1+x)^a(1-x)^b f(x) 1_{[-1,1]}(x)\mathrm{d}x$$

where $a < 1, b > 1$, f is a strictly positive $\mathcal{C}^1$-function and Z is a normalization constant. Indeed let σ_0 be such that

$$\int \frac{1}{(1-x)^2}\mathrm{d}\nu(x) = \frac{1}{\sigma_0^2}.$$

Let us consider

$$\mathbb{R} \setminus \mathcal{O}_1 = \mathrm{supp}(\nu) \cup \{u \in \mathbb{R} \setminus \mathrm{supp}(\nu), \int \frac{1}{(u-x)^2}\mathrm{d}\nu(x) \geq \frac{1}{\sigma^2}\}.$$

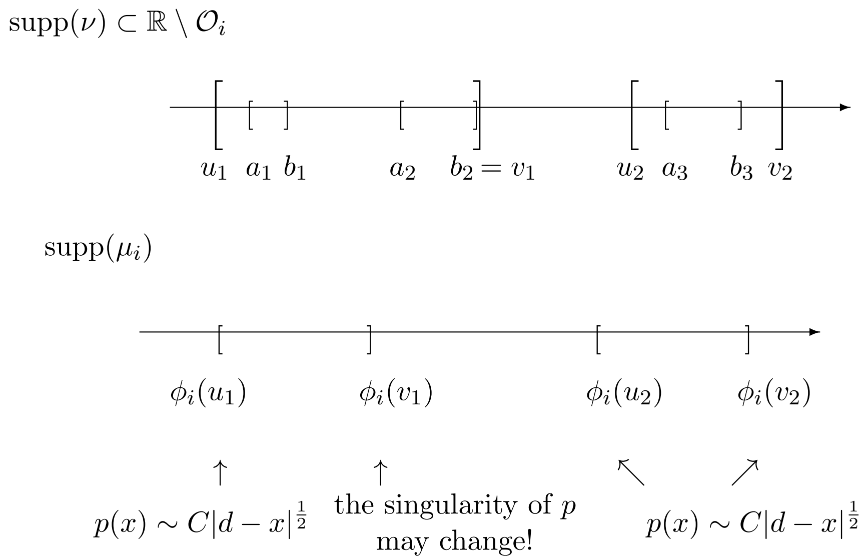
It can be easily seen that for all $\sigma > \sigma_0$, $\mathbb{R} \setminus \mathcal{O}_1 = [u_\sigma, v_\sigma]$ with $u_\sigma < -1 < 1 < v_\sigma$, so that

$$\mathrm{supp}\,(\mu_{sc} \boxplus \nu) = [\phi_1(u_\sigma), \phi_1(v_\sigma)]; \tag{13}$$

thus, $\phi_1(v_\sigma)$ is a regular edge and we have $p(x) \sim C(\phi_1(v_\sigma) - x)^{\frac{1}{2}}$.
Now, for all $\sigma \leq \sigma_0$, one can see that $\mathbb{R} \setminus \mathcal{O}_1 = [u_\sigma, 1]$, with $u_\sigma < -1$, so that $\mathrm{supp}\,(\mu_{sc} \boxplus \nu) = [\phi_1(u_\sigma), \phi_1(1)]$; it turns out that the density exhibits the following behaviour at the right edge

$$p(x) \sim C(\phi_1(1) - x)^b. \tag{14}$$

We illustrate by the following picture the difference of behaviour of the density p at edges of μ_i depending on wether the support of ν sticks to the frontier of $\mathbb{R} \setminus \mathcal{O}_i$ or not. We consider a measure ν whose support has three connected components $[a_i, b_i], i = 1, 2, 3$. Then, we know that $\mathbb{R} \setminus \mathcal{O}_i$ has at most three connected components and each of them contains at least a connected component of the support of ν. We draw one possible case where $[a_1, b_1]$ and $[a_2, b_2]$ are in the same connected component $[u_1, v_1]$ of $\mathbb{R} \setminus \mathcal{O}_i$ and $b_2 = v_1$ whereas $[a_3, b_3]$ is in an other connected component $[u_2, v_2]$ of $\mathbb{R} \setminus \mathcal{O}_i$.



4 Outliers of general spiked models

In Section 1.3, we presented the seminal works on the behavior of the largest eigenvalues for finite rank deformations of standard models. It turns out that the previous analysis in Section 3 allows to understand the appearence of outliers of general spikes models that is when A_N is a deformation with full rank and provides the good way to generalize the pioneering works. Actually, 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. In the finite rank case this criterion reduces to a critical threshold.
Note that in the iid case, we do not assume anymore that the entries are Gaussian; the results stated in this section are obtained under different technical assumptions on the entries on the non-deformed model that we do not precise here and we refer the reader to the corresponding papers.
In order to adopt universal notations for the three types of deformations, we set

$$\tilde{A}_N = \begin{cases} A_N & \text{for additive or multiplicative deformations} \\ A_N A_N^* & \text{for Information-plus-noise type deformation.} \end{cases}$$

Thus, for each type of deformation, we assume the following on the perturbation $\tilde{A}_N$:

- $\mu_{\tilde{A}_N}$ weakly converges towards a probability measure ν whose support is compact.
- The eigenvalues of $\tilde{A}_N$ are of two types :

– $N-r$ (r fixed) eigenvalues $\alpha_i(N)$ such that

$$\max_{i=1}^{N-r} \operatorname{dist}(\alpha_i(N), \operatorname{supp}(\nu)) \underset{N\to\infty}{\longrightarrow} 0$$

– a finite number J of fixed (independent of N) eigenvalues called spikes $\theta_1 > \ldots > \theta_J$ (> 0 for multiplicative deformations and information-plus-noise type models), $\forall i = 1, \ldots, J$, $\theta_i \notin \operatorname{supp}(\nu)$, each θ_j having a fixed multiplicity k_j, $\sum_j k_j = r$.

For technical reasons, for information-plus-noise type model, we assume moreover that A_N is a rectangular matrix of the type

$$A_N = \begin{pmatrix} a_1(N) & & & (0) \\ & & (0) & \\ & & \ddots & (0) \\ & (0) & & \\ & & a_N(N) & (0) \end{pmatrix} \tag{15}$$

4.1 Location of the outliers

Here is a naive intuition for general additive deformed models in order to make the reader understand the occurence and role of free subordination functions. We have the following subordination property

$$g_{\mu\boxplus\nu}(z) = g_\nu(\omega_{\mu,\nu}(z))$$

For an Hermitian deformed model such that

$$M_N = W_N + A_N; \quad \mu_{W_N} \to \mu; \mu_{A_N} \to \nu, \mu_{M_N} \to \mu \boxplus \nu,$$

the intuition is that (see (17) below)

$$g_{\mu_{M_N}}(z) \approx g_{\mu_{A_N}}(\omega_{\mu,\nu}(z)).$$

Assume that A_N has a spiked eigenvalue θ outside its limiting support. If $\rho \notin \operatorname{supp}(\mu \boxplus \nu)$ is a solution of $\omega_{\mu,\nu}(\rho) = \theta$, $g_{\mu_{M_N}}(\rho) \approx g_{\mu_{A_N}}(\omega_{\mu,\nu}(\rho))$ explodes!

Therefore the conjecture is that the spikes θ's of the perturbation A_N that may generate outliers in the spectrum of M_N belong to $\omega_{\mu,\nu}(\mathbb{R} \setminus \operatorname{supp}(\mu \boxplus \nu))$ and more precisely that for large N, the θ's such that the equation

$$\omega_{\mu,\nu}(\rho) = \theta_i$$

has solutions ρ outside $\operatorname{supp}(\mu \boxplus \nu)$ generate eigenvalues of M_N in a neighborhood of each of these ρ.

This intuition in fact corresponds to true results for both models: i.i.d and isotropic. Nevertheless, their proofs are different. In the following, we present the distinct approaches.

4.1.1 The i.i.d case

In this section, we will denote by $\phi, \varphi, \mathcal{O}$ any of $\phi_i, \varphi_i, \mathcal{O}_i$ for $i = 1, 2, 3$ introduced in Section 3.2 related to the investigated deformed model. We choose to present a unified result covering the three types of deformations. Nevertheless, results concerning the Information-Plus-Noise type model involve some technical additionnal asumptions. We brievely precise them in remarks following the unified result and refer the reader to the corresponding papers.

a) A deterministic equivalent

In the three deformed models, a deterministic measure plays a central role in the study of the spectrum of the deformed models. This measure is a very good approximation of the spectral measure μ_{M_N} in the sense that almost surely, for large N, each interval in the complement of the support of this deterministic measure contains no eigenvalue of M_N. This was first established by Bai and Silverstein [5] in the multiplicative case. We now express this deterministic measure ν_N in the three models :

i) Additive deformation of a Wigner matrix [37]:

$$\nu_N = \mu_{sc} \boxplus \mu_{A_N}.$$

ii) Multiplicative deformation of a sample covariance matrix ([5])

$$\nu_N = \mu_{\mathrm{MP}} \boxtimes \mu_{A_N}.$$

iii) Information plus noise model ([57] in the Gaussian case, [8])

$$\nu_N = (\sqrt{\mu_{\mathrm{MP}}} \boxtimes_c \sqrt{\mu_{A_N A_N^*}})^2.$$

We denote by K_N the support of ν_N and for $\epsilon > 0$, $(K_N)_\epsilon$ denotes an ϵ neighborhood of K_N. We have the following result :

Proposition 2. *$\forall \epsilon > 0$, $\forall [a, b] \subset (K_N)_\epsilon$,*

$$\mathbb{P}(\textit{ for large } N,\ M_N \textit{ has no eigenvalue in } [a, b]) = 1 \tag{16}$$

Remark 4. *For non-Gaussian Information plus noise model, the result is proved only for $a > 0$.*

The proof of (16) in [37, Theorem 5.1] relies on the estimation

$$\mathbb{E}\left(g_{\mu_{M_N}}(z)\right) - g_{\nu_N}(z) = \frac{1}{N} L_N(z) + O(\frac{1}{N^2}) \tag{17}$$

with an explicit formula for L_N. (17) is established using an integration by parts formula in the Gaussian case (and an approximate integration by parts formula in the general Wigner case). Note that the same method was used in [57] for Gaussian Information-Plus-Noise type matrices.

To prove the inclusion of the spectrum (16) for sample covariance matrices and information-plus-noise type matrices [5, 8], Bai and Silverstein provide a different approach although it also makes use of Stieltjes transform.

b) An exact separation phenomenon

A next step in the analysis of the spectrum of deformed iid models is an exact separation phenomenon between the spectrum of M_N and the spectrum of $\tilde{A}_N$, involving the subordination functions : to a gap in the spectrum of $\tilde{A}_N$, it corresponds, through the function ϕ defined in Section 3.2, a gap in the spectrum of M_N which splits the spectrum of M_N exactly as that of $\tilde{A}_N$. Let $[a, b]$ be a compact set such that for some $\delta > 0$, for all large N, $[a-\delta, b+\delta] \subset \mathbb{R} \setminus \mathrm{supp}\,(\nu_N)$. Then, almost surely, for large N, $[\varphi(a), \varphi(b)]$ is in the complement of the spectrum of $\tilde{A}_N$. Hence, with the convention that for any $N \times N$ matrix Z, $\lambda_0(Z) = +\infty$ and $\lambda_{N+1}(Z) = -\infty$, there is $i_N \in \{0, \dots, N\}$ such that

$$\lambda_{i_N+1}(A_N) < \varphi(a) \quad \text{and} \quad \lambda_{i_N}(A_N) > \varphi(b). \tag{18}$$

Moreover, $[a, b]$ splits the spectrum of M_N exactly as $[\varphi(a), \varphi(b)]$ splits the spectrum of $\tilde{A}_N$ as stated by the following.

Theorem 13. *With i_N satisfying* (18)*, one has*

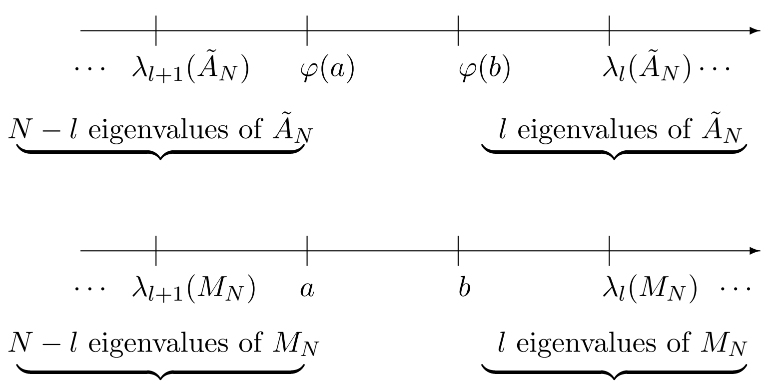
$$\mathbb{P}[\lambda_{i_N+1}(M_N) < a \;\; \text{and} \;\; \lambda_{i_N}(M_N) > b, \text{ for all large } N] = 1. \tag{19}$$

Remark 5. *For non-Gaussian Information plus noise model, if $c < 1$, the result is proved only for b such that $\varphi(b) > 0$.*

The following picture illustrates this exact separation phenomenon.

$$[a, b] \subset \; \mathbb{R} \setminus \mathrm{supp}(\nu_N) \longleftrightarrow [\varphi(a), \varphi(b)]$$

$$\text{gap in } \; \mathrm{Spect}(M_N) \longleftrightarrow \text{gap in } \; \mathrm{Spect}(\tilde{A}_N)$$



Again, this was first observed by Bai and Silverstein [6], in the case of sample covariance matrices. We refer to [37] for deformed Wigner matrices and to Loubaton-Vallet [58] (Gaussian case), Capitaine [33] for information plus noise type models. This exact separation phenomenon leads asymptotically to the relation (12) between the cumulative distribution function of the μ_i's and the cumulative distribution function of ν.

c) Convergence of eigenvalues

The following result gives the precise statement of the intuition given at the beginning of Section 4.1 and is a consequence of the inclusion of the spectrum and the exact separation.

Theorem 14. *[10, 66, 37, 33] Assume that the LSD ν of A_N has a finite number of connected components. For each spiked eigenvalue θ_j, we denote by $n_{j-1}+1, \ldots, n_{j-1}+k_j$ the descending ranks of θ_j among the eigenvalues of $\tilde{A}_N$. With the notations of Section 3.2,*

1) *If $\theta_j \in \mathcal{O}$, the k_j eigenvalues $(\lambda_{n_{j-1}+i}(M_N),\ 1 \leq i \leq k_j)$ converge almost surely outside the support of μ towards $\rho_{\theta_j} = \phi(\theta_j)$.*

2) *If $\theta_j \in \mathbb{R}\backslash\mathcal{O}$ then we let $[s_{l_j}, t_{l_j}]$ (with $1 \leq l_j \leq m$) be the connected component of $\mathbb{R}\backslash\mathcal{O}$ which contains θ_j.*

   a) *If θ_j is on the right (resp. on the left) of any connected component of $\operatorname{supp}(\nu)$ which is included in $[s_{l_j}, t_{l_j}]$ then the k_j eigenvalues $(\lambda_{n_{j-1}+i}(M_N),\ 1 \leq i \leq k_j)$ converge almost surely to $\phi(t_{l_j}^+)$ (resp. $\phi(s_{l_j}^-)$) which is a boundary point of the support of μ.*

   b) *If θ_j is between two connected components of $\operatorname{supp}(\nu)$ which are included in $[s_{l_j}, t_{l_j}]$ then the k_j eigenvalues $(\lambda_{n_{j-1}+i}(M_N),\ 1 \leq i \leq k_j)$ converge almost surely to the α_j-th quantile of μ (that is to q_{α_j} defined by $\alpha_j = \mu(]-\infty, q_{\alpha_j}])$) where α_j is such that $\alpha_j = 1 - \lim_N \frac{n_{j-1}}{N} = \nu(]-\infty, \theta_j])$.*

Remark 6. *For Information-Plus-Noise type models, in 2)a) of Theorem 14, if θ_j is on the left of any connected component of $\operatorname{supp}(\nu)$ which are included in the first connected component $[s_1, t_1]$ of $\mathbb{R}\backslash\mathcal{O}$, then the convergence of the corresponding eigenvalues of M_N towards $\phi(s_1^-)$ is established only for $\phi(s_1^-) = 0$.*

4.1.2 The isotropic case

Here we consider an additive spiked deformation of an isotropic matrix

$$M_N = A_N + U_N^* B_N U_N.$$

U_N is a unitary matrix whose distribution is the normalized Haar measure on the unitary group $\mathrm{U}(N)$. B_N is a deterministic Hermitian matrix of size $N \times N$ such that μ_{B_N} converges weakly to μ compactly supported as $N \to \infty$ and such that the eigenvalues of B_N converge uniformly to $\operatorname{supp}(\mu)$ as $N \to \infty$. A_N is a deterministic Hermitian $N \times N$ perturbation as defined at the beginning of this Section.

Note that if A_N has no outlier, that is if $\{\theta_1, \cdots, \theta_J\} = \emptyset$, then the general study of Collins and Male allows to deduce that neither does M_N (see Corollary 3.1 in [41]) meaning that for all large N, all the eigenvalues of X_N are inside a small neighborhood of the support of $\mu \boxplus \nu$.

Assume now that A_N has outliers. Then Belinschi, Bercovici, Capitaine and Février established in [19] the following result.

Theorem 15. *Set* $K = \mathrm{supp}(\mu \boxplus \nu)$,

$$K' = K \cup \left[\bigcup_{i=1}^{J} \omega_2^{-1}(\{\theta_i\}) \right],$$

whith ω_2 *defined as in Section 2.2.1. The following results hold almost surely for large* N*:*
Given $\varepsilon > 0$ *and denoting by* K'_ϵ *an* ϵ *neighborhood of* K*, we have*

$$\mathrm{spect}(M_N) \subset K'_\varepsilon.$$

Let ρ *be a fixed number in* $K' \setminus K$ *and* θ_i *be such that* $\omega_2(\rho) = \theta_i$*. For any* $\varepsilon > 0$ *such that* $(\rho - 2\varepsilon, \rho + 2\varepsilon) \cap K' = \{\rho\}$*, we have*

$$\mathrm{card}(\{\mathrm{spect}(M_N) \cap (\rho - \varepsilon, \rho + \varepsilon)\}) = k_i.$$

Here we explain the sketch of the proof. Fix $\alpha \in \mathrm{supp}(\nu)$. Due to the left and right invariance of the Haar measure on $\mathrm{U}(N)$ we may assume without loss of generality that both A_N and B_N are diagonal matrices. More precisely, we let A_N be the diagonal matrix

$$A_N = \mathrm{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1 \text{times}}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J \text{times}}, \alpha_1^{(N)}, \ldots, \alpha_{N-r}^{(N)}),$$

and write $A_N = A'_N + A''_N$, where

$$A'_N = \mathrm{Diag}(\underbrace{\alpha, \ldots, \alpha}_{r}, \alpha_1^{(N)}, \ldots, \alpha_{N-r}^{(N)}),$$

and

$$A''_N = \mathrm{Diag}(\underbrace{\theta_1 - \alpha, \ldots, \theta_1 - \alpha}_{k_1 \text{times}}, \ldots, \underbrace{\theta_J - \alpha, \ldots, \theta_J - \alpha}_{k_J \text{times}}, \underbrace{0, \ldots, 0}_{N-r}).$$

We have $A''_N = P_N^* \Theta P_N$, where P_N is the $r \times N$ matrix representing the usual projection $\mathbb{C}^N \to \mathbb{C}^r$ onto the first r coordinates, and

$$\Theta = \mathrm{Diag}(\underbrace{\theta_1 - \alpha, \ldots, \theta_1 - \alpha}_{k_1 \text{times}}, \ldots, \underbrace{\theta_J - \alpha, \ldots, \theta_J - \alpha}_{k_J \text{times}}).$$

The matrices A'_N and B_N have no spikes, and therefore [41, Corollary 3.1] applies to the matrix $M'_N = A'_N + U_N^* B_N U_N$. Note that the limiting spectral measure is still $\mu \boxplus \nu$. The first key idea is due to Benaych-Georges and Nadakuditi [25] and consists in reducing the problem of locating outliers of the deformations to a convergence problem of a fixed size $r \times r$ random matrix, by using the Sylvester's determinant identity: for rectangular matrices X and Y such that XY and YX are square, we have

$$\det(I + XY) = \det(I + YX). \tag{20}$$

Given z outside the support of $\mu \boxplus \nu$, we have

$$\det(zI_N - (A_N + U_N^* B_N U_N)) = \det(zI_N - M'_N) \det(I_N - \left(zI_N - M'_N\right)^{-1} P_N^* \Theta P_N).$$

so that using Sylvester's identity,

$$\det(zI_N - (A_N + U_N^* B_N U_N)) = \det(zI_N - M'_N) \det(I_r - P_N \left(zI_N - M'_N\right)^{-1} P_N^* \Theta).$$

We conclude that the eigenvalues of $A_N + U_N^* B_N U_N$ outside $\mu \boxplus \nu$ are precisely the zeros of the function $\det(F_N(z))$, where

$$F_N(z) = I_r - P_N \left(zI_N - M'_N\right)^{-1} P_N^* \Theta. \tag{21}$$

The key idea is now to establish an approximate matricial subordination result. Biane [29] proved the stronger result that for any $\mathbf{a}$ and $\mathbf{b}$ free selfadjoint random variables in a tracial W*-probability space, there exists an analytic self-map $\omega\colon \mathbb{C}^+ \to \mathbb{C}^+$ of the upper half-plane so that

$$\mathbb{E}_{\mathbb{C}[\mathbf{a}]}\left[(z - (\mathbf{a} + \mathbf{b}))^{-1}\right] = (\omega(z) - \mathbf{a})^{-1}, \quad z \in \mathbb{C}^+. \tag{22}$$

Here $\mathbb{E}_{\mathbb{C}[\mathbf{a}]}$ denotes the conditional expectation onto the von Neumann algebra generated by $\mathbf{a}$. It can be proved that an approximate version does hold in the sense that the compression

$$P_N \left[\mathbb{E}\left[(z - (A'_N + U_N^* B_N U_N))^{-1}\right]^{-1} + A'_N\right] P_N$$

is close to $\omega_2(z) I_r$, as N goes to infinity. Thus, it turns out that almost surely the sequence $\{F_N\}_N$ converges uniformly on compact subsets of $\mathbb{C} \setminus \operatorname{supp}(\mu \boxplus \nu)$ to the analytic function F defined by

$$F(z) = \operatorname{Diag}\left(\underbrace{1 - \frac{\theta_1 - \alpha}{\omega_2(z) - \alpha}}_{k_1 \text{times}}, \ldots, \underbrace{1 - \frac{\theta_J - \alpha}{\omega_2(z) - \alpha}}_{k_J \text{times}}\right),$$

where ω_2 is the subordination function from (7). The set of points z such that $F(z)$ is not invertible is precisely $\bigcup_{i=1}^{J} \omega_2^{-1}(\{\theta_i\})$.

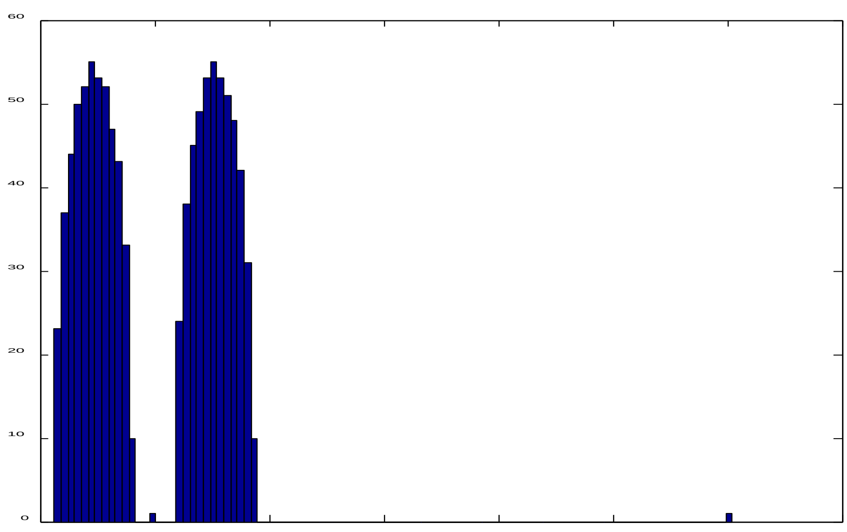
It follows from this result that a remarkable new phenomenon arises: a single spike of A_N can generate asymptotically a finite or even a countably infinite set of outliers of M_N. This arises from the fact that the restriction to the real line of some subordination functions may be many-to-one, that is, with the above notation, the set $\omega_i^{-1}(\{\theta\})$ may have cardinality strictly greater than 1, unlike the subordination function related to free convolution with a semicircular distribution that was used in section 4.1.1. The following numerical simulation, due to Charles Bordenave, illustrates the appearance of two outliers arising from a single spike. We take $N = 1000$ and $M_N = A_N + U_N B_N U_N^*$, where $B_N = \mathrm{diag}(\underbrace{-1, \ldots, -1}_{\frac{N}{2}}, \underbrace{1, \ldots, 1}_{\frac{N}{2}})$, and

$$A_N = \begin{bmatrix} \frac{W_{N-1}^G}{2} & 0_{(N-1)\times 1} \\ 0_{1\times(N-1)} & 10 \end{bmatrix},$$

with W_{N-1}^G being sampled from a standard 999×999 G.U.E.. This is not a spiked deformed GUE model and now, the spike $\theta = 10$ is associated to the matrix approximating the semicircular distribution. We have the subordination identities

$$g_{\mu_{sc}\boxplus\nu}(z) = g_\nu(\omega_2(z)) = g_{\mu_{sc}}(\omega_1(z))$$

where ω_2 is injective on $\mathbb{R} \setminus \mathrm{supp}\,(\mu_{sc} \boxplus \nu)$ but ω_1 may be many to one. Actually $\omega_1(\rho) = 10$ has 2 solutions ρ_1 and ρ_2.



Actually, one may consider additive models where both A_N and B_N have spiked outliers ([19]). Let us consider

$$M_N = U_N B_N U_N^* + A_N$$

where U_N is a Haar unitary matrix, A_N and B_N are deterministic diagonal matrices such that $\mu_{B_N} \underset{N\to+\infty}{\longrightarrow} \mu$ and $\mu_{A_N} \underset{N\to+\infty}{\longrightarrow} \nu$, both μ and ν being compactly supported. We also assume that there exist a spiked $\theta \notin \text{supp}(\nu)$ which is an eigenvalue of A_N with multiplicity k and a spiked $\alpha \notin \text{supp}(\mu)$ which is an eigenvalue of B_N with multiplicity l, whereas the other eigenvalues are uniformly close to the limiting supports. We have the subordination properties

$$g_{\mu \boxplus \nu}(z) = g_\nu(\omega_2(z)) = g_\mu(\omega_1(z)).$$

If there exists $\rho \in \mathbb{R} \setminus \text{supp}(\mu \boxplus \nu)$ such that

$$\begin{cases} \omega_1(\rho) = \alpha \\ \omega_2(\rho) = \theta \end{cases}$$

then for all large N, there are $k + l$ outliers of M_N in a neighborhood of ρ.

Such results are established for multiplicative perturbations of unitarily invariant matricial models, based on similar ideas, with the subordination function replaced by its multiplicative counterpart.

4.2 Eigenvectors

For a general perturbation, dealing with sample covariance matrices, S. Péché and O. Ledoit [65] introduced a tool to study the average behaviour of the eigenvectors but it seems that this did not allow them to focus on the eigenvectors associated with the eigenvalues that separate from the bulk. It turns out that further studies [32, 19, 34] point out that the angle between the eigenvectors of the outliers of the deformed model and the eigenvectors associated to the corresponding original spikes is determined by free subordination functions.

The following theorem [32, 19] holds for spiked additive deformations in the i.i.d case as well as in the isotropic case. Let A_N be a deterministic deformation as defined as the beginning of Section 4; dealing with either a deformed Wigner matrix or a deformed isotropic matrix as defined at the beginning of Section 4.1.2, we have the following

Theorem 16. *Set $K = \text{supp}(\mu \boxplus \nu)$,*

$$K' = K \cup \left[\bigcup_{i=1}^{p} \omega_2^{-1}(\{\theta_i\}) \right],$$

and let ω_2 be the subordination function satisfying (7). Let ρ be a fixed number in $K' \setminus K$ and θ_i be such that $\omega_2(\rho) = \theta_j$. Let $\varepsilon > 0$ be such that $(\rho - 2\varepsilon, \rho + 2\varepsilon) \cap K' = \{\rho\}$. Let ξ be a unit eigenvector associated to an eigenvalue of M_N in $(\rho - \epsilon, \rho + \epsilon)$. Then when N goes to infinity,

$$\|P_{\text{Ker }(\theta_l I_N - A_N)}(\xi)\|^2 \to \frac{\delta_{jl}}{\omega_2'(\rho)} \quad \textit{almost surely}$$

Similar results are established for spiked multiplicative deformations in the i.i.d case as well as in the isotropic case [32, 19] and for information-plus-noise type models in [34] in the i.i.d case. See the following tabular. Note that in the i.i.d case everything is explicit and can be rewritten as follows using the ϕ_i's defined in Section 3.2.

Theorem 17. *Let $n_{j-1}+1,\dots,n_{j-1}+k_j$ the descending ranks of θ_j among the eigenvalues of $\tilde{A}_N$ and $\xi(j)$ a unit eigenvector associated to one of the eigenvalues $(\lambda_{n_{j-1}+q}(M_N),\ 1\le q\le k_j)$. Then when N goes to infinity,*

- *For any $\theta_l \neq \theta_j$,*
$$\|P_{\mathrm{Ker}\ (\theta_l I_N-\tilde{A}_N)}(\xi(j))\| \to 0 \quad \textit{almost surely}$$

- $$\|P_{\mathrm{Ker}\ (\theta_j I_N-\tilde{A}_N)}(\xi(j))\|^2 \to \alpha_j \quad \textit{almost surely}$$

$$\textit{where } \alpha_j = \begin{cases} \phi_1'(\theta_j) = 1-\sigma^2\int \frac{1}{(\theta_j-x)^2}\mathrm{d}\nu(x) \ \textit{for deformed Wigner matrices} \\[2ex] \frac{\theta_j\phi_2'(\theta_j)}{\phi_2(\theta_j)} = \frac{1-c\int\frac{x^2}{(\theta_j-x)^2}\mathrm{d}\nu(x)}{1+c\int\frac{x}{(\theta_j-x)}\mathrm{d}\nu(x)} \ \textit{for sample covariance matrices} \\[2ex] \frac{\phi_3'(\theta_j)}{1+\sigma^2 c g_\nu(\theta_j)} \ \textit{for information-plus-noise type matrices} \end{cases}$$

Here are the common basic ideas of the proof of these results [32].
First note that if $u_1,\dots,u_N$ and $w_1,\dots,w_N$ are respectively a basis of eigenvectors associated with $\lambda_1(\tilde{A}_N),\dots,\lambda_N(\tilde{A}_N)$ and with $\lambda_1(M_N),\dots,\lambda_N(M_N)$, we have

$$\mathrm{Tr}\left[h(M_N)f(\tilde{A}_N)\right] = \sum_{k,l} h(\lambda_k(M_N))f(\lambda_l(\tilde{A}_N))|\langle u_l, w_k\rangle|^2.$$

Thus, since the θ_l's separate from the rest of the spectrum of $\tilde{A}_N$ and the outliers of M_N separate from the rest of the spectrum of M_N, one can deduce the asymptotic norm of the projection onto an eigenspace associated to a spike θ_i, of an eigenvector associated to an outlier of M_N from the study of the asymptotic behaviour of $\mathrm{Tr}\left[h(M_N)f(\tilde{A}_N)\right]$ for a fit choice of h and f. Then, a concentration of measure phenomenon reduces the problem to the study of $\mathbb{E}(\mathrm{Tr}\left[h(M_N)f(\tilde{A}_N)\right])$.
The third key point is to approximate the function h by its convolution by the Poisson Kernel in order to exhibit the resolvent of the deformed model

$$\mathbb{E}\left[\mathrm{Tr}\left[h(M_N)f(\tilde{A}_N)\right]\right] = -\lim_{y\to 0^+}\frac{1}{\pi}\Im\int \mathbb{E}\left(\mathrm{Tr}\left[G_N(t+iy)f(\tilde{A}_N)\right]\right)h(t)\mathrm{d}t$$

where $G_N(z)=(zI_N-M_N)^{-1}$. Finally, writing $\tilde{A}_N = U^*DU$, with D diagonal and U unitary, defining $\tilde{G}_N := UG_NU^*$, the result follows from sharp estimations of $\mathbb{E}(\{\tilde{G}_N\}_{kk}(z))$, for any k in $\{1,\dots,N\}$.

4.3 Unified understanding

In conclusion, solving the problem of outliers consists in solving an equation involving the free subordination function and the spikes of the perturbation. Moreover, the norm of the orthogonal projection of an eigenvector associated to an outlier of the deformed model onto the eigenspace of the corresponding spike of the perturbation is asymptotically determined by the free subordination function. This is summarized in the following tabular.
In the tabular, Y_N denotes a Hermitian random matrix of iid type ($Y_N = W_N$, S_N or $\sigma \frac{X_N}{\sqrt{p}}$ according to the deformations, see section 1.3) or Y_N is unitarily invariant (resp. biunitarily invariant for the information plus noise model).

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

Note that up to now, the formula in the lower right corner of the previous tabular, 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 finite rank perturbation A_N.

5 Fluctuations at edges of spiked deformed models

In this section, we present results on fluctuations of outliers and eigenvalues at soft edges of the limiting support, with particular stress on understanding the phenomena through free probability theory.

5.1 The Gaussian case

5.1.1 Additive deformation

Let us consider the deformed G.U.E.. It is known from Johansson [48] (see also [31]) that the joint eigenvalue density induced by the latter model can be explicitely computed. Furthermore it induces a so-called "determinantal random point field".
When A_N is of finite rank, Péché [64] obtained a striking phase transition phenomenon for the fluctuations of the largest eigenvalue of the deformed G.U.E..

Theorem 18. *1. If $\theta_1 < \sigma$, $\sigma^{-1}N^{2/3}(\lambda_1(M_N) - 2\sigma)$ converges in distribution to the G.U.E. Tracy Widom distribution F_2.*

2. If $\theta_1 = \sigma$, $\sigma^{-1}N^{2/3}(\lambda_1(M_N) - 2\sigma)$ converges in distribution to a "generalized" Tracy Widom distribution F_{3,k_1}.

3. If $\theta_1 > \sigma$, $N^{1/2}(\lambda_1(M_N) - 2\sigma)$ converges in distribution to the largest eigenvalue of a G.U.E. matrix of size k_1 and parameter $\sigma_{\theta_1} = \sigma\sqrt{1-(\sigma/\theta_1)^2}$. In particular, if $k_1 = 1$, $N^{1/2}(\lambda_1(M_N) - 2\sigma)$ converges in distribution to a centered normal distribution with variance σ_{θ_1}.

The proof is based on the expression of the distribution of the largest eigenvalue in terms of Fredholm determinant and then as a contour integral. The asymptotic properties rely on a saddle point analysis.

The seminal works concerning full rank deformations of a G.U.E. matrix made strong assumptions on the rate of convergence of μ_{A_N} to ν. In [69], the author investigates the local edge regime which deals with the behavior of the eigenvalues near any regular extremity point u_0 of a connected component of $\mathrm{supp}(\mu_{sc} \boxplus \nu)$. The typical size of the fluctuations of the eigenvalues at regular edges (see Section 3.2.1) is $N^{-2/3}$. [69] considers the case where μ_{A_N} concentrate quite fast to the measure ν. In particular, there are no spike. More precisely [69] makes a technical assumption on the uniform convergence of the Stieltjes transform of μ_{A_N} to g_ν:

$$\sup_{z\in K} |g_{\mu_{A_N}}(z) - g_\nu(z)| \leq N^{-2/3-\epsilon}, \tag{23}$$

where K is some compact subset of the complex plane at a positive distance of the support of ν. Then, [69] proves that the joint distribution of the eigenvalues converging to u_0 have universal asymptotic behavior, characterized by the Tracy-Widom distribution. In [2] and [1], [30, 4] the authors consider the case where A_N has two distinct eigenvalues $\pm a$ of equal multiplicity. They proved the Tracy-Widom fluctuations at edges (which are all regular since ν is discrete).

It turns out that the above strong assumptions made on the rate of convergence of μ_{A_N} to ν can be removed by studying the asymptotic distribution of eigenvalues in the vicinity of mobile edges namely the edges of the deterministic equivalent $\mu_{sc} \boxplus \mu_{A_N}$ of the empirical eigenvalue distribution of the deformed GUE. In [38], the authors establish the following results.

Let d be a regular right edge of $\text{supp}(\mu_{sc} \boxplus \nu)$. Assume moreover that for any θ_j such that $\int \frac{\mathrm{d}\nu(s)}{(\theta_j - s)^2} = 1/\sigma^2$, we have $d \neq \phi_1(\theta_j) = \theta_j - \sigma^2 g_\nu(\theta_j)$. It turns out that for η small enough, for all large N, there exists a unique right edge d_N of $\text{supp}(\mu_{sc} \boxplus \mu_{A_N})$ in $]d-\eta, d+\eta[$. and the asymptotic distribution of eigenvalues in the vicinity of d_N is universal as the following Theorem 19 states.

Theorem 19. *Let k be a given fixed integer. Let $\lambda_{max} \geq \lambda_{max-1} \geq \cdots \lambda_{max-k+1}$ denote the k largest of those eigenvalues of M_N converging to d. There exists $\alpha > 0$ depending on d_N only such that the vector*

$$\frac{N^{2/3}}{\alpha} \left(\lambda_{max} - d_N, \lambda_{max-1} - d_N, \ldots, \lambda_{max-k+1} - d_N\right)$$

converges in distribution as $N \to \infty$ to the so-called Tracy-Widom G.U.E. distribution for the k largest eigenvalues (see [72]).

We now turn to the behavior of outliers. Let θ_i be a spiked eigenvalue with multiplicity k_i, such that $\int \frac{1}{(\theta_i - x)^2} \mathrm{d}\nu(x) < 1/\sigma^2$. Recall that in [37], the authors prove that the spectrum of M_N exhibits k_i eigenvalues in a neighborhood of

$$\rho_{\theta_i} = \theta_i + \sigma^2 \int \frac{\mathrm{d}\nu(x)}{\theta_i - x}. \tag{24}$$

Oncemore, dealing with mobile edges related to $\mu_{sc} \boxplus \mu_{A_N}$, [38] obtains the following universal result.

Theorem 20. *Let θ_i be such that $\int \frac{\mathrm{d}\nu(x)}{(\theta_i - x)^2} < 1/\sigma^2$ and $\rho_{\theta_i} = \phi_1(\theta_i)$. Then, for $\epsilon > 0$ small enough, for all large N, $supp(\mu_{sc} \boxplus \mu_{A_N})$ has a unique connected component $[L_i(N), D_i(N)]$ inside $]\rho_{\theta_i} - \epsilon, \rho_{\theta_i} + \epsilon[$. Moreover, the k_i outliers of M_N close to ρ_{θ_i} fluctuate at rate $\sqrt{N}$ around $\frac{L_i(N)+D_i(N)}{2}$ as the eigenvalues of a $k_i \times k_i$ GUE.*

The basic tool is a saddle point analysis of the correlation functions of the deformed G.U.E., the contours involving the image of $\mathbb{R}$ by the continuous extension of the subordination function ω_2 defined by (7) with $\mu = \mu_{sc}$ and $\nu = \mu_{A_N}$.

5.1.2 Sample covariance matrices

As in the above section, the distribution of the eigenvalues is explicit, with a determinantal structure. The analysis of the fluctuations relies on an expression of the distribution of the extremal eigenvalues in terms of a Fredholm determinant and then an asymptotic analysis based on a saddle point method, or a steepest descent method.
The first result was obtained by Baik, Ben Arous and Péché [12] who described the fluctuations of the largest eigenvalues at the right edge and revealed the phase transition phenomenon, in the case of a finite rank perturbation A_N of the identity. We refer to [12] for the precise statement, their result being an analogue of Theorem 18.
The full rank case was investigated by Hachem, Hardy and Najim [47] for extremal eigenvalues sticking to the bulk (the support of $\mu_{\mathrm{MP}} \boxtimes \nu$). As in Theorem 19, around regular soft edges, the associated extremal eigenvalues, properly rescaled, in the vicinity of mobile edges namely the edges of the deterministic equivalent $\mu_{\mathrm{MP}} \boxtimes \mu_{A_N}$, converge in law to the Tracy-Widom distribution at the scale $N^{2/3}$.

5.1.3 Random perturbations

If one let the perturbation matrix A_N be random then the mobile edges of the equivalent measure become random and may lead to different rates of convergence and different asymptotic distributions. We present two examples established respectively by Johansson [49] and Lee and Schnelli [54] that we revisited through free convolutions.

- Johansson [49] considered
$$M_N = W_N^G + A_N$$
where W_N^G is a G.U.E. matrix as defined in Section 1.1.1 and
$$A_N = N^{-1/6}\mathrm{diag}(\mathrm{y_1}, \ldots, \mathrm{y_N})$$
where the y_i's are iid real random variables with distribution τ and independent from W_N^G. Let us assume that τ is compactly supported and set $v^2 = \int x^2 \mathrm{d}\tau(x)$. Note that almost surely μ_{A_N} converges weakly to δ_0 and μ_{M_N} converges weakly to μ_{sc}. Denote by $\tilde{d}_N$ the *deterministic* upper right edge of $\mu_{sc} \boxplus \frac{\tau}{N^{1/6}}$ where $\frac{\tau}{N^{1/6}}$ denotes the pushforward of τ by the map $x \mapsto \frac{x}{N^{1/6}}$. Johansson established that
$$\sigma^{-1} N^{2/3}\left(\lambda_{\max}(M_N) - \tilde{d}_N\right) \xrightarrow{\mathcal{D}} X + Y \tag{25}$$
where X and Y are independent random variables, X has the Tracy-Widom distribution and Y has distribution $N(0, \frac{v^2}{\sigma^2})$. Note that the upper right edge $\tilde{d}_N$ of $\mu_{sc} \boxplus \frac{\tau}{N^{1/6}}$ is defined by
$$\tilde{d}_N = \tilde{t}_N + \sigma^2 \int \frac{1}{\tilde{t}_N - x/N^{1/6}} \mathrm{d}\tau(x), \tag{26}$$

where $\tilde{t}_N$ in the vicinity of σ satisfies

$$\int \frac{1}{(\tilde{t}_N - x/N^{1/6})^2} \mathrm{d}\tau(x) = \frac{1}{\sigma^2}. \tag{27}$$

Consider now the *random* upper right edge d_N of $\mu_{sc} \boxplus \mu_{A_N}$. It is defined by

$$d_N = t_N + \sigma^2 \frac{1}{N} \sum_{i=1}^{N} \frac{1}{t_N - y_i/N^{1/6}}, \tag{28}$$

where t_N in the vicinity of σ satisfies

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{(t_N - y_i/N^{1/6})^2} = \frac{1}{\sigma^2}. \tag{29}$$

It is easy to see that (26), (27), (28) and (29) yield

$$d_N - \tilde{d}_N = \sigma^2 Z_N + O((t_N - \tilde{t}_N)^2) \;\text{ and }\; t_N - \tilde{t}_N = O(Z_N),$$

where

$$N^{2/3} Z_N = N^{2/3} \left\{ g_{\mu_{A_N}}(\tilde{t}_N) - g_{\frac{\tau}{N^{1/6}}}(\tilde{t}_N) \right\}$$

converges weakly to a centered Gaussian distribution with variance v^2/σ^4. Thus, it comes readily that

$$\sigma^{-1} N^{2/3} \left\{ d_N - \tilde{d}_N \right\} \xrightarrow{\mathcal{D}} N(0, v^2/\sigma^2). \tag{30}$$

Now (25) readily follows since by Theorem 19, given A_N, $\sigma^{-1} N^{2/3} (\lambda_1(M_N) - d_N)$ converges weakly to the Tracy-Widom distribution.

- Another example is provided by [54] who considered the following deformed model

$$W_N + \mathrm{diag}(v_1, \ldots, v_N)$$

where W_N is a Wigner matrix and v_i are i.i.d random variables independent with W_N, with distribution

$$d\nu(x) = Z^{-1}(1+x)^a (1-x)^b f(x) 1_{[-1,1]}(x) dx$$

with $a < 1, b > 1$ and $f > 0$ is a $\mathcal{C}^1$-function. Assume that W_N is a G.U.E.. Let σ_0 be defined by $\int \frac{1}{(1-x)^2} \mathrm{d}\nu(x) = \frac{1}{\sigma_0^2}$. According to Section 3.2.1, we have $\mathrm{supp}\,(\mu_{sc} \boxplus \nu) = [d_\sigma^-, d_\sigma^+]$. Morevover, by (13), for all $\sigma > \sigma_0$, d_σ^+ is a regular edge, $p(x) \sim C(d_\sigma^+ - x)^{\frac{1}{2}}$. Therefore denoting by $d_\sigma^+(N)$ the (random) upper right edge of $\mathrm{supp}\,(\mu_{sc} \boxplus \mu_{A_N})$, Theorem 19 yields

$$\alpha^{-1} N^{2/3} (\lambda_1(M_N) - d_\sigma^+(N)) \xrightarrow{\mathcal{D}} TW.$$

A similar study as in Johansson's example shows that the random edge $d_\sigma^+(N)$ fluctuates as

$$\sqrt{N}(d_\sigma^+(N) - d_\sigma^+) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2\left(1 - \sigma^2\left(g_{\mu_{sc}\boxplus\nu}(d_\sigma^+)\right)^2\right)\right).$$

Thus, we can deduce that

$$\sqrt{N}(\lambda_1(M_N) - d_\sigma^+) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2\left(1 - \sigma^2\left(g_{\mu_{sc}\boxplus\nu}(d_\sigma^+)\right)^2\right)\right).$$

Note that for all $\sigma < \sigma_0$, according to (14), $p(x) \sim C(d_\sigma^+ - x)^b$. Lee and Schnelli also investigate the fluctuations at the non-regular edge d_σ^+ and establish that

$$N^{\frac{1}{b+1}}(\lambda_1(M_N) - d_\sigma^+) \xrightarrow{\mathcal{D}} G_{b+1}(s)$$

as N goes to infinity, where $G_{b+1}(s) = (1 - \exp((\frac{s}{c})^{b+1}))\mathbf{1}_{[0;+\infty[}(s)$ (Weibull distribution with parameters $b+1$ and $c = c(\nu, \sigma)$).

5.2 The general i.i.d case

We finish by some remarks concerning non Gaussian frameworks. We do not detail the results since they do not fall under the scope of free probability theory.

- Universality at soft edges.
  Several recent works proved the universality of the Tracy-Widom fluctuations at soft edges for quite general deformed Wigner matrices or sample covariance matrices without outliers. The methods pursue a Green function comparison strategy [15, 55, 56] or make use of anisotropic local laws [52].

- Non universal fluctuations of outliers
  A new phenomenon arises for the fluctuations of the outliers : the limiting distribution can depend on the distribution of the entries (non universality), according to the localization/delocalization of the eigenvectors of A_N. Note that in the Gaussian case in the previous subsection, the eigenvectors of the perturbation are irrelevant for the fluctuations, due to the unitary invariance in Gaussian models. We illustrate this dependence on the eigenvectors on A_N in a very simple situation, in the additive case. Consider two finite rank perturbations of rank 1, with one non null eigenvalue $\theta > \sigma$. The first one $A_N^{(1)}$ is a matrix with all entries equal to θ/N (delocalized eigenvector associated to θ). The second one $A_N^{(2)}$ is a diagonal matrix (localized eigenvector). The fluctuations of the largest eigenvalue λ_1 of the matrix $M_N^{(i)} = X_N + A_N^{(i)}$ ($i = 1, 2$) around $\rho_\theta := \theta + \frac{\sigma^2}{\theta}$ are given as follows :

  Proposition 3. *Fluctuations of outliers*

  1. *Delocalized case [44] : The largest eigenvalue $\lambda_1(M_N^{(1)})$ have Gaussian fluctuations :*

$$\sqrt{N}(\lambda_1(M_N^{(1)}) - \rho_\theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(1 - \sigma^2/\theta^2)) \tag{31}$$

2. Localized case [35] : The largest eigenvalue $\lambda_1(M_N^{(2)})$ fluctuates as

$$\sqrt{N}(1 - \frac{\sigma^2}{\theta^2})(\lambda_1(M_N^{(2)}) - \rho_\theta) \xrightarrow{\mathcal{D}} \mu \star \mathcal{N}(0, v_\theta). \tag{32}$$

where μ is the distribution of the entries of the Wigner matrix, the variance v_θ of the Gaussian distribution depends on θ and the second and fourth moments of μ.

The proof of (31) is combinatorial and is based on the computation of large moments of $\mathrm{Tr}(M_N)$. The proof of (32) relies on a determinant identity, analogous to (20), boiling down to the behavior of a fixed rank determinant and a CLT for quadratic forms.

We refer to [10] (sample covariance case), [36, 67, 68, 51] for fluctuations of ouliers for more general perturbations, in the case of iid models, and [24] for unitarily invariant models.

References

[1] M. Adler, M. Cafasso and P. van Moerbeke. From the Pearcey to the Airy process. *Electron. J. Probab.* 16 , no. 36, (2011), 1048-1064.

[2] M. Adler and P. van Moerbeke. PDEs for the Gaussian ensemble with external source and the Pearcey distribution. *Comm. Pure Appl. Math.* 60, no. 9, (2007),1261–1292.

[3] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, Cambridge University Press, Cambridge, 2010.

[4] A.I. Aptekarev, P.M. Bleher, A.B.J Kuijlaars. Large n limit of Gaussian random matrices with external source, part II. *Commun. Math. Phys.* 25, (2005), 367–389.

[5] Z. Bai and J. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* 26 , no. 1 (1998), 316–345.

[6] Z. Bai and J. Silverstein. Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.* 27 , no. 3 (1999), 1536–1555.

[7] Z. Bai and J. Silverstein. *Spectral analysis of large dimensional random matrices.* Second edition. Springer Series in Statistics. Springer, New York, 2010.

[8] Z. Bai and J. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices. *Random Matrices Theory Appl.* 1 (2012), no. 1, 1150004, 44 pp.

[9] 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.

[10] Z. Bai and J. Yao. On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.* 106 (2012), 167–177.

[11] 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.

[12] J. Baik, G. Ben Arous and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* 33 no. 5, (2005), 1643–1697.

[13] J. Baik and J. W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.*, 97(6), (2006), 1382–1408.

[14] Z. Bao, L. Erdös, K. Schnelli. Local law of addition of random matrices on optimal scale arXiv:1509.07080v2

[15] 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.

[16] S. T. Belinschi. The atoms of the free multiplicative convolution of two probability distributions. *Integral Equations and Operator Theory* , 46 (4), (2003) 377– 386.

[17] S. Belinschi, F. Benaych-Georges, and A. Guionnet. Regularization by free additive convolution, square and rectangular cases. *Complex Anal. Oper. Theory*, 3(3), (2009), 611–660.

[18] S. T. Belinschi and H. Bercovici, A new approach to subordination results in free probability, *J. Anal. Math.* 101, (2007), 357–365.

[19] S. T. Belinschi, H. Bercovici, M. Capitaine, M. Février : Outliers in the spectrum of large deformed unitarily invariant models. arXiv:1412.4916

[20] S. Belinschi, T. Mai, R. Speicher. Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. *J. Reine angew. Math.*, to appear.

[21] S. Belinschi, P. Sniady, R. Speicher. Eigenvalues of non-hermitian random matrices and Brown measure of non-normal operators: hermitian reduction and linearization method. arXiv:1506.02017

[22] S. Belinschi, R. Speicher, J. Treilhard. Operator-valued free multiplicative convolution: analytic subordination theory and applications to random matrix theory. Int. Math. Res. Not. IMRN 2015, no. 14, 5933–5958.

[23] F. Benaych-Georges. Rectangular random matrices. Related convolution, *Probab. Theory Related Fields* Vol. 144, no. 3 (2009) 471–515.

[24] F. Benaych-Georges, A. Guionnet and M. Maida. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. *Electron. J. Probab.* 16 no. 60, (2011), 1621–1662.

[25] F. Benaych-Georges and R.R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* 227 no. 1, (2011), 494–521.

[26] F. Benaych-Georges and R.R. Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *J. Multivariate Anal.* 111 (2012), 120–135.

[27] H. Bercovici and D. Voiculescu, Free convolution of measures with unbounded support, *Indiana Univ. Math. J.*42 (1993), 733–773.

[28] P. Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.*, 46(3) (1997) 705–718.

[29] P. Biane, Processes with free increments, *Math. Z.* 227 (1998), 143–174.

[30] P.M. Bleher, A.B.J. Kuijlaars. Large n limit of Gaussian random matrices with external source, part I. *Commun. Math. Phys.* 252, (2004), 43–76.

[31] E. Brézin, S. Hikami Correlations of nearby levels induced by a random potential. *Nuclear Phys. B* 479, no. 3, 697-706, 1996.

[32] M. Capitaine. Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices *Journal of Theoretical Probability*, 26 (3) (2013), 595–648.

[33] M. Capitaine. Exact separation phenomenon for the eigenvalues of large Information-Plus-Noise type matrices. Application to spiked models. *Indiana Univ. Math. J.*, 63 (6), (2014), 1875–1910.

[34] M. Capitaine. Limiting eigenvectors of outliers for Spiked Information-Plus-Noise type matrices. Preprint (2015).

[35] M. Capitaine, C. Donati-Martin, D. Féral : The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. *Annals of Proba* 37 (2009), no. 1, 1–47.

[36] M. Capitaine, C. Donati-Martin, D. Féral : Central limit theorems for eigenvalues of deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* 48 no. 1, (2012), 107–133.

[37] M. Capitaine, C. Donati-Martin, D. Féral and M. Février. Free convolution with a semicircular distribution and eigenvalues of spiked deformations of Wigner matrices. *Electron. J. Probab.* 16 no. 64, (2011), 1750–1792.

[38] M. Capitaine, S. Péché. Fluctuations at the edges of the spectrum of the full rank deformed GUE. *Probab. Theory Relat. Fields*, 165(1), (2016), 117–161.

[39] G. P. Chistyakov and F. Götze. Limit theorems in free probability theory. I. *Ann. Probab.*, 36(1), (2008), 54-90.

[40] S. Choi and J.W. Silverstein. Analysis of the limiting spectral distribution of large dimensional random matrices. *J. Multivariate. Anal.*, 54, (1995), 295–309.

[41] B. Collins and C. Male, The strong asymptotic freeness of Haar and deterministic matrices, *Ann. Sci. Éc. Norm. Supér. (4)* 47 no. 1, (2014), 147–163.

[42] R.B. Dozier and J.W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices. *J. Multivariate. Anal.*, vol. 98, no. 4, (2007), 678–694.

[43] R.B. Dozier and J.W. Silverstein. Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices. *J. Multivariate. Anal.*, vol. 98, no. 6, (2007), 1099–1122.

[44] D. Féral and S. Péché. The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.* 272 no. 1, (2007), 185–228.

[45] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3), (1981) 233–241.

[46] S. Geman. A limit theorem for the norm of random matrices. *Ann. Probab.*, 8 (2), (1980) 252–261.

[47] W. Hachem, A. Hardy and J. Najim. Large Complex Correlated Wishart Matrices: Fluctuations and Asymptotic Independence at the Edges, to appear in *Annals of Probability* (2015).

[48] K. Johansson. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* 215, no. 3, (2001), 683-705.

[49] K. Johansson. From Gumbel to Tracy-Widom. *Probability Theory and Related Fields.* 138, (1), (2007), 75-112.

[50] I. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* 29 no. 2, (2001), 295–327.

[51] A. Knowles and J. Yin. The outliers of a deformed Wigner matrix. The Annals of Probability Vol. 42, No. 5, (2014), 1980–2031.

[52] A. Knowles and J. Yin. Anisotropic local laws for random matrices. (2014), arXiv:1410.3516,

[53] J. O. Lee and K. Schnelli. Local deformed semicircle law and complete delocalization for Wigner matrices with random potential. J. Math. Phys. 54, 103504 (2013).

[54] J. O. Lee and K. Schnelli. Extremal eigenvalues and eigenvectors of deformed Wigner matrices, *Probability Theory and Related Fields* pp 1-77 First online: 06 January 2015

[55] J. O. Lee and K. Schnelli. Edge Universality for Deformed Wigner Matrices. (2014), arXiv:1407.8015.

[56] J. O. Lee and K. Schnelli. Tracy-Widom Distribution for the Largest eigenvalue of real sample covariance matrices with general population. (2014), arXiv:1409.4979.

[57] P. Vallet, P. Loubaton and X. Mestre. Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case. *IEEE Trans. Inform. Theory* 58 no. 2, (2012), 1043–1068.

[58] P. Loubaton and P. Vallet. Almost sure localization of the eigenvalues in a Gaussian information-plus-noise model. Application to the spiked models . *Electron. J. Probab.* 16 no. 70, (2011), 1934–1959.

[59] Hans Maassen. Addition of freely independent random variables. *J. Funct. Anal.*, 106(2), (1992), 409-438.

[60] V. A. Marchenko and L. A. Pastur, Distribution of eigenvalues for some sets of random matrices, Math. USSR-Sb. 1 (1967), 457-483.

[61] L. A. Pastur. On the spectrum of random matrices. *Theoretical and Mathematical Physics*, Volume 10, Issue 1, (1972), 67-74.

[62] L. Pastur and V. Vasilchuk. On the Law of Addition of Random Matrices. *Commun. Math. Phys. 214*, (2000), 249–286 .

[63] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model *Statist. Sinica*, 17 (4), (2007), 1617–1642.

[64] S. Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* 134, no. 1, (2006), 127-173

[65] S. Péché and O. Ledoit. Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Relat. Fields* 151 (2011), 233–264.

[66] N. R. Rao and J. W. Silverstein. Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples. *IEEE Journal of Selected Topics in Signal Processing*, 4(3), (2010) 468–480.

[67] A. Pizzo, D. Renfrew, A. Soshnikov. On finite rank deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* 49, no. 1, (2013), 64-94.

[68] D. Renfrew, A. Soshnikov. On finite rank deformations of Wigner matrices II. Delocalized perturbations *Random Matrices Theory Appl.* 2 (2013), no. 1, 1250015, 36 pp.

[69] T. Shcherbina. On universality of local edge regime for the deformed Gaussian unitary ensemble. *Journal of Statistical Physics* 143(3), (2011), 455–481.

[70] D. Shlyakhtenko. Free probability of type B and asymptotics of finite-rank perturbations of random matrices. arXiv:1509.08841 (2015).

[71] J. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis* 55(2) (1995), 331–339.

[72] C. Tracy, H. Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys. 159* (1994), 151–174.

[73] V. Vasilchuk. On the law of multiplication of random matrices. *Mathematical Physics, Analysis and Geometry* Volume 4, Issue 1, (2001), 1–36

[74] D. Voiculescu. Addition of certain noncommuting random variables, *J. Funct. Anal.* **66** (1986), 323–346.

[75] D. Voiculescu, Multiplication of certain noncommuting random variables. *J. Operator Theory* 18(1987), 223–235.

[76] D. Voiculescu. Limit laws for random matrices and free products, *Invent. Math.* **104** (1991), 201–220.

[77] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory. I, *Comm. Math. Phys.* 155 (1993), 71–92.

[78] D. Voiculescu. The coalgebra of the free difference quotient and free probability, *Internat. Math. Res. Notices* 2000, 79–106.

[79] D. Voiculescu, K. J. Dykema, and A. Nica. *Free random variables*, American Mathematical Society, Providence, RI, 1992.

[80] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)* 62, (1955), 548–564..

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

[82] 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.