

Deformed ensembles, polynomials in random matrices and free probability theory

Mireille Capitaine

▶ **To cite this version:**

Mireille Capitaine. Deformed ensembles, polynomials in random matrices and free probability theory. Probability [math.PR]. Université Paul Sabatier - Toulouse 3, 2017. tel-01978065

HAL Id: tel-01978065
https://hal.science/tel-01978065

Submitted on 11 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paul Sabatier Toulouse 3



Manuscrit
présenté pour l'obtention de

l'Habilitation à Diriger des Recherches

Mireille Capitaine



Deformed ensembles, polynomials in random matrices and free probability theory

---

Soutenue le 7 décembre 2017,

après avis des rapporteurs

Pr. James A. Mingo	Queen's University (Canada)
Pr. Roland Speicher	Universität des Saarlandes (Allemagne)
Pr. Dan Voiculescu	University of California at Berkeley (USA)

devant le jury formé de

Pr. Catherine Donati-Martin	Université de Versailles-St Quentin
Pr. Michel Ledoux	Université Paul Sabatier
Pr. Thierry Lévy	Université Pierre et Marie Curie
Pr. James A. Mingo	Queen's University (Canada)
Pr. Sandrine Péché	Université Paris Diderot
Pr. Roland Speicher	Universität des Saarlandes (Allemagne)
Pr. Dan Voiculescu	University of California at Berkeley (USA)

A mes parents,
à André, Aurélien et Marie,
avec toute ma reconnaissance et mon amour.

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude à Michel Ledoux. Il ne pourrait y avoir de mots à la hauteur de ma reconnaissance et mon admiration. Je mesure la chance que j'ai eu de le rencontrer: il m'a fait connaître le monde de la recherche, y a guidé mes premiers pas et n'a cessé de m'apporter écoute, soutien, encouragements et précieux conseils durant toutes ces années. Merci vraiment pour tout Michel!

Je voudrais remercier très chaleureusement Jamie Mingo, Roland Speicher et Dan Voiculescu qui ont donné de leur temps précieux pour examiner ce manuscrit et me font l'honneur de participer à ce jury. J'ai eu le plaisir de les rencontrer à plusieurs reprises lors de conférences sur les probabilités libres: je suis toujours très impressionnée par leur excellence mathématique mais aussi leur grande gentillesse.

Je suis très heureuse que Catherine Donati-Martin, Thierry Lévy et Sandrine Péché m'honorent de leur présence dans ce jury. Catherine, de notre longue et si précieuse collaboration est née une amitié très, très forte; je n'imaginais pas cette soutenance sans ta présence! C'est toujours avec un plaisir sincère, Thierry, que je te retrouve au gré des conférences et que j'assiste à tes exposés toujours remarquables. Travailler avec toi, Sandrine, a été très enrichissant et très agréable; j'espère que nous aurons bientôt l'occasion de collaborer sur d'autres projets.

J'ai appris beaucoup de mes collaborateurs, je remercie chacun d'entre eux; j'ai une pensée toute particulière et affectueuse pour Delphine, Muriel et Serban avec qui j'ai travaillé sur plusieurs projets, Charles avec qui j'ai partagé quelques années le bureau, et qui sont devenus des amis sincères.

Maxime, Yuriy, vous encadrer a été une expérience très enrichissante mathématiquement et humainement. Je vous souhaite beaucoup de plaisir dans votre métier de chercheur!

Au-delà des mathématiques, j'aimerais également remercier tous mes collègues et amis de l'IMT mais aussi des communautés des matrices aléatoires et probabilités libres.
Les jours précédents cette soutenance ont été endeuillés par le départ de maman. Je tiens à remercier très chaleureusement et sincèrement Agnès, Anthony, Béatrice, Catherine, Cécile, Charles, Delphine, Franck, François, Françoise, Gersende, Guillaume, Jean-Michel, Laure, Maxime, Michel, Muriel, Patrick, les deux Pierre, Reda, Sabine, Sébastien, Serban, Sylvie, Thierry C-D., Yuriy, pour leurs marques d'affection. Béatrice, Catherine, Cécile, Laure, Muriel, merci de tout coeur pour votre soutien permanent.

Franck, Laure, Michel, Patrick, Philippe, merci pour votre confiance et bienveillance: vous ne vous êtes jamais lassés de frapper à ma porte pour m'inciter à écrire cette habilitation! Eh bien, c'est fait ...

Acknowledgments

I would first like to express my deepest gratitude to Michel Ledoux. There could be no words to the height of my gratitude and my admiration. I measure the chance I had of meeting him: he made me discover the world of research, guided my first steps and kept on listening, supporting, encouraging and advising me during all these years. Thank you for everything Michel!

I would like to warmly thank Jamie Mingo, Roland Speicher and Dan Voiculescu, for agreeing to examine this manuscript and do me the honor to participate in this jury. I had the pleasure to meet them on several occasions during conferences on Free Probability Theory: I am always very impressed by their mathematical excellence but also their great kindness.

I am very pleased that Catherine Donati-Martin, Thierry Lévy and Sandrine Péché honored me with their presence in this jury. Catherine, from our long and so dear collaboration was born a very, very strong friendship; I did not imagine this defense without your presence! Thierry, it is always with sincere pleasure that I find you at conferences and that I attend your ever-remarkable presentations. Sandrine, working with you was very enriching and very enjoyable; I hope that we will soon have the opportunity to collaborate on other projects.

I have learned a lot from my collaborators, I thank each one of them; I have a special and affectionate thought for Delphine, Muriel and Serban with whom I worked on several projects, for Charles with whom I shared a few years an office, and who became all true friends.

Maxime, Yuriy, supervising your work was a very enriching experience mathematically and humanely. I wish you much pleasure in your job as a researcher!

Beyond mathematics, I would also like to thank all my colleagues and friends at IMT and from the random matrices and free probability theory communities.
A shadow was cast over the days before this defense by the death of my mother. I would like to warmly and sincerely thank Agnès, Anthony, Béatrice, Catherine, Cécile, Charles, Delphine, Franck, François, Françoise, Gersende, Guillaume, Jean-Michel, Laure, Maxime, Michel, Muriel, Patrick, les deux Pierre, Reda, Sabine, Sébastien, Serban, Sylvie, Thierry C-D., Yuriy, for their marks of affection.
Béatrice, Catherine, Cécile, Laure, Muriel, thank you with all my heart for your permanent affection.

Franck, Laure, Michel, Patrick, Philippe thank you for your trust and kindness: you were never tired of knocking at my door to make me write this habilitation! Well, it's done ...

Publications

[C1] (1995) Onsager-Machlup functional for some smooth norms on Wiener space, M. Capitaine, *Probab. Theory Relat. Fields, 102*, 189-201.

[C2] (1997) Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces, M. Capitaine, E.P. Hsu et M. Ledoux, *Elect. Comm. Prob.2*, paper 7.

[C3] (1998) Sur une inégalité de Sobolev logarithmique pour une diffusion unidimensionnelle, M. Capitaine, *Séminaire de Probabilités XXXII*, 6-13.

[C4] (2000) On the Onsager-Machlup functional for elliptic diffusion processes, M. Capitaine, *Séminaire de Probabilités XXXIV*, 313-328.

[C5] (2001) The Lévy area process for the free Brownian motion, M. Capitaine, C. Donati-Martin, *J. Funct. Anal. 179*, 153-169.

[C6] (2003) Large deviation bounds for matrix Brownian motion, P. Biane, M. Capitaine, A. Guionnet, *Inv. Math.* 152, 433-459.

[C7] (2004) Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to Beta random matrices, M. Capitaine, M. Casalis, *Indiana Univ. Math. J.*, Vol. 53, No 2, 397-431.

[C8] (2005) Free Wishart Processes, M. Capitaine, C. Donati-Martin, *J. Theor. Probab.*, Vol. 18, No. 2, 413-438.

[C9] (2006) Cumulants for random matrices as convolutions on the symmetric group, M. Capitaine, M. Casalis, *Probab. Theory Relat. Fields* 136, 19-36.

[C10] (2007) Strong asymptotic freeness for Wigner and Wishart matrices, M. Capitaine, C. Donati-Martin, *Indiana Univ. Math. J.* Vol. 56, No. 2, 767-803.

[C11] (2007) Cumulants for random matrices as convolutions on the symmetric group II, M. Capitaine, M. Casalis, *J. Theor. Probab.*, Vol. 20, No. 3, 505-533.

[C12] (2007) Geometric interpretation of the cumulants for random matrices previously defined as convolutions on the symmetric group, M. Capitaine, M. Casalis, *Séminaire de Probabilités XLI*, 93-119.

[C13] (2009) The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations, M. Capitaine, C. Donati-Martin, D. Féral, *Ann. Probab.*, 37 (1), 1-47.

[C14] (2011) Free Bessel laws, T. Banica, S. Belinschi, M. Capitaine, B. Collins, *Canad. J. Math* 63, 3-37, (ayant reçu le prix G. de B. Robinson 2012).

[C15] (2011) Free convolution with a semi-circular distribution and eigenvalues of spiked deformations of Wigner matrices, M. Capitaine, C. Donati-Martin, D. Féral, M. Février, *Electron. J. Probab.*, 16, 1750-1792.

[C16] (2012) Central limit theorems for eigenvalues of deformations of Wigner matrices, M. Capitaine, C. Donati-Martin, D. Féral, *Ann. Inst. Henri Poincaré Probab. Stat.* 48 (1), 107-133.

[C17] (2013) Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices, M. Capitaine, *J. Theor. Probab.* 26 (3) 595-648.

[C18] (2014) Exact separation phenomenon for the eigenvalues of large Information-Plus-Noise type matrices. Application to spiked models, M. Capitaine, *Indiana Univ. Math. J.* 63 (6), 1875-1910.

[C19] (2016) Fluctuations at the edges of the spectrum of the full rank deformed GUE, M. Capitaine, S. Péché, *Probab. Theory Relat. Fields*, Vol. 165, Issue 1, 117-161.

[C20] (2016) Outlier eigenvalues for deformed i.i.d random matrices, C. Bordenave, M. Capitaine, *Comm. Pure Appl. Math.* , Vol. 69, Issue 11, 2131-2194.

[C21] (2017) Outliers in the spectrum of large deformed unitarily invariant models, S.T. Belinschi, H. Bercovici, M. Capitaine, M. Février, to appear in *Ann. Probab.*.

[C22] (2017) Spectrum of deformed random matrices and free probability, M. Capitaine, C. Donati-Martin, to appear in *SMF volume Panoramas et Synthèses.*

[C23] (2017) Spectral properties of polynomials in independent Wigner and deterministic matrices, S. T. Belinschi, M. Capitaine, *J. Funct. Anal.*, http://dx.doi.org/10.1016/j.jfa.2017.07.010.

[C24] Limiting eigenvectors of outliers for Spiked Information-Plus-Noise type matrices, M. Capitaine, arxiv 1701.08069.

[C25] On the outlying eigenvalues of a polynomial in large independent random matrices, S. T. Belinschi, H. Bercovici, M. Capitaine, arXiv:1703.08102.

Summary of research

I obtained my PhD in Probability Theory, supervised by Michel Ledoux, from the University of Toulouse, Paul-Sabatier, in 1996. The main objectives of the PhD consisted in analyzing the role of the topology in various results concerning diffusion processes, such as the support theorem, the large deviations principle, or the determination of the Onsager-Machlup functional. The investigation led besides to some logarithmic Sobolev inequalities for elliptic diffusions. The general framework of this investigation was therefore at the interface of stochastic calculus, Malliavin calculus, stochastic differential geometry and semigroup theory. This work gave rise to 4 publications [C1-C4].

I then entered the world of free probability theory, created by Dan Voiculescu, and random matrices theory, naturally via stochastic calculus. Catherine Donati-Martin and I proved in [C5] that Lyons' deterministic approach to differential equations driven by rough paths can be applied successfully to the free Brownian motion. We constructed a Lévy area process for the free Brownian motion and in this way, a typical geometric rough path lying above the free Brownian path. In [C8], we defined free Wishart processes of parameter $\lambda > 0$ and proved a free additivity property and invertibility for $\lambda > 1$. For $\lambda \geq 1$, we showed that a free Wishart process is a solution of a SDE of square Bessel process type, driven by a free complex Brownian motion. In the case $\lambda > 1$, we established existence and uniqueness of a strong solution of a such a SDE.

In [C6], together with Philippe Biane and Alice Guionnet, we proved large deviation bounds for the convergence of Hermitian matrix valued Brownian motion towards free Brownian motion. As a consequence, we obtained upper and lower bounds on the microstates entropy introduced by Dan Voiculescu. The basic idea in the proof was to construct exponential martingales based on the Clark-Ocone formula. As an additional contribution, to obtain the lower bound, we established a characterization of free Brownian motion given by a free version of Lévy's theorem.

My interests then turned towards more combinatorial aspects of free probability theory as developed by Roland Speicher. The motivation behind the joint works [C7], [C9], [C11] and [C12] with Muriel Casalis, was to draw the dotted arrows of the following diagram

$$\begin{array}{ccc} \text{Freeness} & \longleftrightarrow & \text{Free cumulants} \\ N \to +\infty \uparrow & & \wedge \\ & & \vdots \\ N \times N \text{ independent matrices} & < \cdots > & \text{Matricial cumulants} \end{array}$$

We defined and studied cumulants of matrices and showed that free cumulants can be naturally seen as their limiting values when the dimension goes to infinity. We analysed the convolution relations on the geodesics on Cayley graphs involving moments and cumulants in free probability, at the matricial level dealing with random matrices whose distribution is invariant under the action of the unitary or orthogonal/symplectic group. These matricial convolution relations involve respectively the symmetric groups $(\mathcal{S}_n)_{n\in\mathbb{N}}$ and the Gelfand pairs $((S_{2n}, H_n))_{n\in\mathbb{N}}$, H_n denoting the hyperoctaedral group. It turns out that these combinatorial formulas for moments of random matrices allowed Teo Banica, Serban Belinschi, Benoit Collins and myself to provide matricial models for a family of real probability measures π_{st}, that we introduced in [C14] and called free Bessel laws. These are related to the free Poisson law π via the formula $\pi_{s1} = \pi^{\boxtimes s}$ and $\pi_{1t} = \pi^{\boxplus t}$. Our study includes definition and basic properties, analytic aspects (supports, atoms, densities), combinatorial aspects (functional transforms, moments, partitions) and a discussion of the relation with random matrices and quantum groups.

During the period 2009-17, my research activities aimed at developing the contribution of free probability theory to the analysis of the spectral properties of deformed ensembles and polynomials in random matrices. In particular, the main principle of subordination in free probability is emphasized as a main tool in the understanding of the localization of the outliers and the corresponding eigenvectors of many matricial models.

This investigation corresponds to my research activities in the last years illustrated by the contributions $[C10]$, $[C13]$ and $[C15 - C25]$. This report is devoted to these papers exclusively.

Introduction

There is currently a quite precise knowledge of the asymptotic spectral properties (i.e. when the dimension of the matrix tends to infinity) of a number of "classical" random matrix models (Wigner matrices, Wishart matrices, invariant ensembles...). This understanding covers both the so-called global regime (asymptotic behavior of the spectral measure) and the local regime (asymptotic behavior of the extreme eigenvalues and eigenvectors, spacings...).

Practical problems (in the theory of statistical learning, signal detection etc.) naturally lead to wonder about the spectrum reaction of a given random matrix after a deterministic perturbation. For example, in the signal theory, the deterministic perturbation is seen as the signal, the perturbed matrix is perceived as a "noise", and the question is to know whether the observation of the spectral properties of "signal plus noise" can give access to significant parameters on the signal. Theoretical results on these "deformed" random models may allow to establish statistical tests on these parameters. A typical illustration is the so-called BBP phenomenon (after Baik, Ben Arous, Péché) which put forward outliers (eigenvalues that move away from the rest of the spectrum) and their Gaussian fluctuations for spiked covariance matrices.

My interest in these issues has been concerned with deformed ensembles and polynomials in several random matrices with a particular emphasis towards a universal understanding (i.e. independent of the model) of the various output of deformation. This investigation has been supported at three levels: a common analysis of three classical Hermitian deformed models; the study of the spectral properties of non-commutative Hermitian polynomials in independent random matrices; the study of non-Hermitian deformed models. The crucial aspect of the contributions, with my collaborators, towards these objectives has been the analysis of the phenomenon of outliers via the (operator-valued) subordination functions from free probability theory.

This report focuses on localization and fluctuations of eigenvalues of complex (even if some similar results hold in the real setting) random deformed models and polynomials in asymptotically free random matrices with a particular interest in the outliers and the corresponding eigenvectors. After a very short reminder on the classical Hermitian random models (Chapter 1), Chapter 2 presents the pionner works on three finite rank deformations of these classical models. Chapter 3 tries to bring to light a general methodology, based on scalar free subordination properties, for studying the spectral properties of the corresponding full rank deformations. This universal understanding culminates in Chapter 4 dealing with non-commutative polynomials in random Hermitian matrices; this investigation is achieved by an even more general methodology based on a linearization procedure and operator-valued subordination properties. Chapter 5 investigates full rank additive perturbations of iid random matrices and is the first stage on the way of the analysis of outliers for non-Hermitian random models involving several non-Hermitian matrices. Appendix is a very short reminder on some facts on free probability theory used throughout this exposition.

This presentation of a huge and very active topic of the current research is by far not exhaustive, and mainly focuses on the recent developments in the area in connection with my own contributions. In particular, I did not try to achieve a complete bibliography of the relevant literature and only emphasize some sample results and references.

Chapter 1

Asymptotic spectrum of classical Hermitian random matrix models

Three fundamental classes of Hermitian matrices have been extensively studied in random matrix theory: Wigner matrices, sample covariance matrices and invariant ensembles. The study of spectral properties of such random matrices is twofold: one can consider the spectrum as a whole and study the global properties of the spectrum. This is the field where the first famous results of random matrix theory were obtained. One can also investigate local properties of the spectrum and study interactions between neighboring eigenvalues or the behavior of extreme eigenvalues. We focus on some few pionner works on the asymptotic behavior of the spectral measure and extreme eigenvalues which turn to be relevant for results of the next chapters. We refer to the monographies [4, 11, 52, 68, 102, 112] for a thorough introduction to random matrix theory.
Given an arbitrary Hermitian matrix B of size N, we denote by $\lambda_1(B) \geq \cdots \geq \lambda_N(B)$ its N ordered eigenvalues and by μ_B its empirical spectral measure

$$\mu_B = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(B)}.$$

1.1 Global behaviour of the spectrum of standard models

1.1.1 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:

$$S_N = \frac{1}{p} X_N X_N^* \tag{1.1}$$

where X_N is a random matrix with independent entries.
We shall assume here that $N \leq p(N)$, X_N is a $N \times p(N)$ matrix, $(X_N)_{ij}$ $i = 1, \ldots, N$, $j = 1, \ldots, p$ are i.i.d complex variables, with variance 1 and mean zero. If the variables are Gaussian, $S_N =: S_N^G$ is called a Wishart matrix, or from the invariant ensemble viewpoint (see below), a matrix from the Laguerre Unitary Ensemble (LUE).

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.

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 1 ([101] (see also[11]))**.** *If $c_N := \frac{N}{p(N)} \to c \in]0;1]$ when $N \to \infty$,*

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

where

$$\mu_{\mathrm{MP}}(dx) = \frac{1}{2\pi c x}\sqrt{(b-x)(x-a)}\,1_{[a,b]}(x)\mathrm{d}x, \tag{1.2}$$

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

1.1.2 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. Here, we will consider Hermitian Wigner matrices of the following form :

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

where H_N is an Hermitian matrix whose diagonal entries are iid real random variables and those above the diagonal are iid complex random variables, with variance σ^2. If the entries are independent Gaussian variables, $W_N =: W_N^G$ is a matrix from the Gaussian Unitary Ensemble (GUE).

Wigner proved that a precise description of the limiting spectrum of these matrices can be achieved.

Theorem 2. *[153, 154] (see also[11])*

$$\mu_{W_N} \xrightarrow{w} \mu_{sc} \quad a.s. \text{ when } \quad N \to +\infty$$

where

$$\frac{d\mu_{sc}}{dx}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\,1_{[-2\sigma,2\sigma]}(x) \tag{1.3}$$

is the so-called semi-circular distribution.

Two classical methods used to be investigated for the proof of these asymptotic global behaviours of the spectrum, after a truncation procedure of the entries: the moment method, boiling down to a combinatorial problem, based on the class of test functions $\{x \mapsto x^r, r \in \mathbb{N}\}$, and the Cauchy-Stieltjes method, boiling down to an analytical problem, based on the class of test functions $\{x \mapsto (z-x)^{-1}, z \in \mathbb{C}, \Im z > 0\}$. A review of these methods can be found in [4, 8, 11].

1.1.3 Invariant ensembles

A complex invariant ensemble is the distribution $\mathcal{L}(M_N)$ of an Hermitian $N \times N$ random matrix M_N such that

$$\mathcal{L}(M_N) = \mathcal{L}(U_N M_N U_N), \forall U_N \in \mathbb{U}(N).$$

M_N can be written $M_N = U_N D_N U_N^*$ almost surely, where U_N is distributed according to the Haar measure on the unitary group, D_N is a diagonal matrix, U_N and D_N are independent (see Proposition 6.1 in [48]). The archetypal ensembles of invariant ensembles for Wigner matrices and sample covariance matrices correspond to the GUE and LUE cases.
Note that, in particular, Quantum Field Theory uses invariant ensembles whose probability law is absolutely continuous with respect to the Lebesgue measure dH on the set of $N \times N$ Hermitian matrices with density $\frac{1}{C_N} \exp(-NTrV(H))$ for some function V. Note that, in the GUE case, $V(x) = \frac{x^2}{2\sigma^2}$. Then the empirical spectral measure of the corresponding model converges weakly towards the equilibrium measure in the external field V (see [112] Chap 11), which is the semi-circular distribution when V is quadratic. The proof is closely related to the theory of logarithmic potential and the mean field approximation in statistical mechanics (see [124]).

One can also investigate local properties of the spectrum and study interactions between neighboring eigenvalues or the behavior of extreme eigenvalues. We focus here on extremal eigenvalues since this is the main subject of study in the coming chapters.

1.2 Extremal eigenvalues

A priori, the convergence of the spectral measure does not prevent an asymptotically negligeable fraction of eigenvalues from going away from the limiting support (called *outliers* in the following). Actually, a further analysis of the moment method allows one to deduce the following behaviour of extreme eigenvalues of the above standard models.

Theorem 3. *[16] (see also [11]) Let W_N be a Wigner matrix as defined in Section 1.1.2. Assume that the entries of H_N has finite fourth moment, then almost surely,*

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

Similarly, it turns out that sample covariance matrices defined by (1.1) no exhibit outliers.

Theorem 4 ([71, 13, 156] (see also[11]))**.** *Assume that the entries of X_N has finite fourth moment, then almost surely,*

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

In multivariate analysis, certain statistics are defined in terms of the extreme eigenvalues of random matrices, which makes the limiting distribution of normalized extreme eigenvalues of special interest. In [142], Tracy and Widom derived the limiting distribution (called the Tracy-Widom law) of the largest eigenvalue of a GUE matrix.

Theorem 5. *Let W_N^G be a GUE matrix. Let $q : \mathbb{R} \to \mathbb{R}$ be the unique solution of the differential equation*

$$q''(x) = xq(x) + 2q(x)^3$$

such that $q(x) \sim_{x\to+\infty} Ai(x)$ where Ai is the Airy function, unique solution on $\mathbb{R}$ of the differential equation $f''(x) = xf(x)$ satisfying $f(x) \sim_{x\to+\infty} (4\pi\sqrt{x})^{1/2} \exp(-2/3x^{3/2})$. Then

$$\lim_{N\to+\infty} \mathbb{P}\left(N^{2/3} \left(\frac{\lambda_1(W_N^G)}{\sqrt{N}} - 2\sigma \right) \le s \right) = F_2(s),$$

where $F_2(s) = \exp\left(-\int_s^{+\infty} (x-s) q^2(x) dx \right)$.

This result has been extended to the LUE by Johansson [84] when $0 < c < +\infty$ and El Karoui [61] when $c = 0$ and $c = +\infty$. Borodin and Forrester [41] extended this result for the smallest eigenvalue of a Laguerre ensemble around the soft edge. (When the leftmost edge is the origin, it is common to refer to it as the hard edge. In contrast, any positive edge is also called a soft edge.) Moreover, dealing with general invariant ensembles, Deift and al [54, 53] proved that the largest eigenvalue still fluctuates around the supremum of the support of the limiting distribution according to the Tracy Widom distribution. The proofs are based on the fact that, in these frameworks, the process of eigenvalues is a determinantal process (see Section 4.2 [4] about general determinantal processes). Note that Ramírez, Rider and Virág [119] developed a stochastic operator approach as pioneered by Edelman and Sutton [60, 136] which hinges on tridiagonal representations.

For a wide class of Wigner or sample covariance matrices, the universality of the largest eigenvalue was conjectured. The first main step to prove this universality conjecture has been achieved by Soshnikov [132, 133] since he established that the largest eigenvalue of a Wigner matrix or a sample covariance matrix associated to a symmetric probability measure which admits sub-gaussian tails fluctuates according to the Tracy-Widom law. In these works, universality of the fluctuations are reduced to the universality of large traces. By a Lindeberg method, Tao and Vu [139] proved a variant of the universality results of Soshnikov for the largest eigenvalues, assuming moment conditions rather than symmetry conditions. In [96] and [55], with a Green function comparison method developped by Erdös, Yau and their co-authors [62, 65], a necessary and sufficient condition on off-diagonal entries of the Wigner matrix or on entries of the sample covariance matrix is established for the distribution of the largest eigenvalue to weakly converge to the Tracy-Widom distribution. We also refer to these papers for references on investigations on edge universality.

All these results stress the robustness of the Tracy-Widom distribution at soft edges of the spectrum of random matricial models.

Note that the local hard edge regime of sample covariance matrices involves the so-called Bessel kernel (see [112] chap. 7 and references therein).

1.3 Eigenvectors

It is well known that the matrix whose columns are the eigenvectors of a GUE matrix or a LUE matrix can be chosen to be distributed according to the Haar measure on the unitary group. In the non-Gaussian case, the exact distribution of the eigenvectors cannot be computed. However, the eigenvectors of general Wigner matrices or sample covariance matrices have been the object of a growing interest and in several papers, a delocalization and universality property were shown for the eigenvectors of these standard models (see among others [40, 45, 63, 64, 88, 141] and references therein). Heuristically, delocalization for a random matrix means that its normalized eigenvectors look like the vectors uniformly distributed over the unit sphere. Let us state for instance the following sample result.

Theorem 6. *(Isotropic delocalization, Theorem 2.16 from [40]). Let W_N be a $N \times N$ Wigner matrix as defined in Section 1.1.2 satisfying some technical assumptions. Let $v(1), \dots, v(N)$ denote the normalized eigenvectors of W_N. Then, for any $C_1 > 0$ and $0 < \epsilon < 1/2$, there exists $C_2 > 0$ such that*

$$\sup_{1 \leq i \leq N} |\langle v(i), u \rangle| \leq \frac{N^\epsilon}{\sqrt{N}},$$

for any fixed unit vector $u \in \mathbb{C}^N$, with probability at least $1 - C_2 N^{-C_1}$.

The following chapter investigates the impact of a finite rank deformation on the asymptotic spectral properties of the previous classical random matrices. The global behavior of the spectrum is unchanged but the situation is drastically different for extreme eigenvalues. Indeed, some eigenvalues may converge out of the bulk and then, the corresponding eigenvectors are localized; their fluctuations may not be universal (toning down the robustness of the Tracy-Widom distribution) and depend on the spectral properties of the deformation and the law of the entries of the non-perturbed matrix.

Chapter 2

Finite rank deformations of classical random matrix models

The study of deformations of random matrix models has been motivated by statistical investigations [86], wireless communications [50], imaging [70], physics [44] or financial applications [127]. In this chapter, we outline some results on finite rank deformations of the classical Hermitian ensembles.

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 is a Wigner matrix (see Section 1.1.1),
$$M_N = W_N + A_N.$$

ii) Multiplicative perturbation: A_N is a non negative Hermitian $N \times N$ matrix and S_N is 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 and X_N is a random matrix with i.i.d. entries defined as in Section 1.1.2, σ is some positive real number,
$$M_N = \left(\sigma \frac{X_N}{\sqrt{p}} + A_N\right)\left(\sigma \frac{X_N}{\sqrt{p}} + A_N\right)^*.$$

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 on the unitary group and B deterministic, and in iii) X_N by a random matrix whose distribution is biunitarily invariant.

In Sections 2.2 and 2.3, we list some pioneer results on extreme eigenvalues of such models dealing with a perturbation A_N with finite rank in i) or iii) or such that $I_N - A_N$ is of finite rank in ii). In Section 2.4, we precise our first contributions [C13] and [C16] in the study of non-Gaussian models, addressing the question of spectral universality. The next chapter will put forward tools from free probability theory to provide a unified understanding, for all these models, of the appearence and localization of outliers as well as of the asymptotic behaviour of the corresponding eigenvectors.

2.1 Asymptotic spectral measure

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),* $\mu_{M_N} \xrightarrow{w} \mu_{sc}$ *(defined by* (1.3)*),*
- *In case ii) and iii),* $\mu_{M_N} \xrightarrow{w} \mu_{\mathrm{MP}}$ *(defined by* (1.2)*).*

This follows from the rank inequalities (see [11, Appendix A.6]).

2.2 Convergence of extreme eigenvalues: the BBP phase transition

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 of Gaussian type or unitarily invariant models. For each model, a phase transition (called BBP phase transition) is pointed out involving different thresholds and different limiting values of outlying eigenvalues or eigenvectors projection. In the following chapter, we will provide a unified understanding of all these results.

2.2.1 Multiplicative deformations

A complete study 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 LUE matrix as defined in Section 1.1.1 and the perturbation A_N[1] is given by

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

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. Baik, Ben Arous and Péché proved the following result.

Theorem 7. *(BBP phase transition)[17, 18]*
Let $\omega_c = 1 + \sqrt{c}$,

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

- *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* $N \to +\infty$

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

The same phenomenon of phase transition was established by Benaych-Georges and Nadakuditi [29] in the case of a deformation of an unitarily invariant model of the form :

$$M_N = (I_N + P_N)^{1/2} U_N B_N U_N^* (I_N + P_N)^{1/2},$$

- B_N is a deterministic $N \times N$ Hermitian matrix such that:
  - (B1) B_N is non negative definite,
  - (B2) when N goes to infinity, $\mu_{B_N} := \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i(B_N)}$ weakly converges to some probability measure μ with compact support $[a; b]$,
  - (B3) the smallest and largest eigenvalue of B_N converge to a and b respectively.
- U_N is a random $N \times N$ unitary matrix distributed according to the Haar measure.
- P_N is a deterministic Hermitian matrix having r non-zero eigenvalues $\gamma_1 \geq \cdots \geq \gamma_s > 0 > \gamma_{s+1} \geq \cdots \geq \gamma_r > -1$, r, $\gamma_i, i = 1, \ldots r$, fixed independent of N .

Note that μ_{M_N} weakly converges to μ. Benaych-Georges and Nadakuditi established the following result.

Theorem 8 ([29]). *Define* $T_\mu \colon \mathbb{C} \setminus \mathrm{supp}(\mu) \to \mathbb{C}$, $\quad T_\mu(z) = \int_{\mathbb{R}} \frac{t d\mu(t)}{z-t}$. *Then, we have for each* $1 \leq i \leq s$, *almost surely,*

$$\lambda_i(M_N) \to_{N\to+\infty} \begin{cases} T_\mu^{-1}(1/\gamma_i) & \text{if } \gamma_i > 1/\lim_{z\downarrow b} T_\mu(z), \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $i > s$, *a.s.,* $\lambda_i(M_N) \to_{N\to+\infty} b$.
Similarly, for the smallest eigenvalues, we have for each $0 \leq j < r - s$, *a.s,*

$$\lambda_{N-j}(M_N) \to_{N\to+\infty} \begin{cases} T_\mu^{-1}(1/\gamma_{r-j}) & \text{if } \gamma_{r-j} < 1/\lim_{z\uparrow a} T_\mu(z), \\ a & \text{otherwise,} \end{cases}$$

while for each fixed $j \geq r - s$, *a.s.* $\lambda_{N-j}(M_N) \to_{N\to+\infty} a$.

2.2.2 Additive deformations

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

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

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 [115]. (Such models were first investigated by Füredi and Komlós [69].)

Theorem 9.
- *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 [29] for

$$M_N = U_N B_N U_N^* + A_N, \tag{2.1}$$

where B_N and U_N satisfy the same conditions as in Section 2.2.1 except (B1) and A_N is a deterministic Hermitian matrix having r non-zero eigenvalues $\gamma_1 \geq \cdots \geq \gamma_s > 0 > \gamma_{s+1} \geq \cdots \geq \gamma_r$, where r and $\gamma_i, i = 1, \ldots r$ are fixed (independent of N). The spectral measure μ_{M_N} weakly converges towards the limiting spectral measure μ of B_N but we have the following result.

Theorem 10 ([29])**.** *Define $g_\mu \colon \mathbb{C} \setminus \mathrm{supp}(\nu) \to \mathbb{C}, \quad g_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(t)}{z-t}$. Then, we have for each $1 \leq i \leq s$, almost surely,*

$$\lambda_i(M_N) \to_{N\to+\infty} \begin{cases} g_\mu^{-1}(1/\gamma_i) & \text{if } \gamma_i > 1/\lim_{z\downarrow b} g_\mu(z), \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $i > s$, a.s., $\lambda_i(M_N) \to_{N\to+\infty} b$.
Similarly, for the smallest eigenvalues, we have for each $0 \leq j < r - s$, a.s.,

$$\lambda_{N-j}(M_N) \to_{N\to+\infty} \begin{cases} g_\mu^{-1}(1/\gamma_{r-j}) & \text{if } \gamma_{r-j} < 1/\lim_{z\uparrow a} g_\mu(z), \\ a & \text{otherwise,} \end{cases}$$

while for each fixed $j \geq r - s$, a.s $\lambda_{N-j}(M_N) \to_{N\to+\infty} a$.

2.2.3 Information plus noise type matrices

A phase transition phenomenon of the same kind was established by Loubaton and Vallet in [98] 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 fixed non zero eigenvalues $\theta_1 \geq \cdots \geq \theta_r$.

Theorem 11. *([98]) 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,*

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

This result was extended as follows in [30] to the case $M_N = (V_N + A_N)(V_N + A_N)^*$ where V_N is a $N \times p$ 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 b (resp a) of the support of μ.

Theorem 12 ([30])**.** *For each $1 \leq i \leq r$, almost surely,*

$$\lambda_i(M_N) \quad \underset{\substack{N\to+\infty \\ N/p\to c\in]0;1]}}{\longrightarrow} \quad \begin{cases} D_\mu^{-1}\left(\frac{1}{\theta_i^2}\right) & \text{if } \theta_i^2 > 1/D_\mu(b^+), \\ b & \text{otherwise.} \end{cases}$$

where

$$D_\mu(z) = \left[\int \frac{z}{z^2-t^2} d\mu(t)\right] \times \left[c\int \frac{z}{z^2-t^2} d\mu(t) + \frac{1-c}{z}\right], \quad \text{for } z > b$$

and D_μ^{-1} denotes its functional inverse on $[b, +\infty[$.

2.2.4 Eigenvectors associated to outliers

In the spiked deformed models, when some eigenvalues separate from the bulk, one may wonder how the corresponding eigenvectors behave. There are some pionneering results concerning finite rank perturbations: [114] in the real Gaussian sample covariance matrix setting, and [29, 30] dealing with finite rank additive or multiplicative perturbations of unitarily invariant matrices. Here is a result from [29] for eigenvector projection onto those of the perturbation corresponding to the largest outlier of a finite rank additive perturbation of a unitarily invariant model.

Theorem 13. *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 Theorem 9 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 $\lambda_1(M_N)$,

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

Thus, an eigenvector associated to an outlier of the deformed model asymptotically lies on a cone around the one corresponding to the spike of the perturbation that generates this outlier.

2.3 Fluctuations of extremal eigenvalues

The BBP phase transition pointed out in [17] for the spiked sample covariance matrix defined in Section 2.2.1, describes three distinct regimes. Assume that $N/p = c$.

- *The subcritical regime*

  If the spiked eigenvalues π_i all remain at some positive distance below $1+\sqrt{c}$ then the largest sample eigenvalue behave exactly as in the non-deformed case, developing Tracy-Widom fluctuations with order $N^{2/3}$ around the right endpoint of the Marchenko-Pastur support.

- *The critical regime.*

  If $\pi_k = \cdots = \pi_1 = 1+\sqrt{c}$, the fluctuations order around the right endpoint of the Marchenko-Pastur support is still $N^{2/3}$ but the limiting distribution is changed to a deformation of the Tracy-Widom law.

- *The supercritical regime.*

  If $\pi_k = \cdots = \pi_1$ exceed $1 + \sqrt{c}$, the largest eigenvalue will separate from the rest and fluctuates according to the distribution of the largest eigenvalue of a $k \times k$ GUE matrix around its outlying limit. Moreover the fluctuation order is $\sqrt{N}$.

Péché [115] obtained the following similar striking phase transition phenomenon for the fluctuations of the largest eigenvalue of a finite rank deformation of a GUE.

Theorem 14. *(with the notations of Section 2.2.2)*

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

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_{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 GUE 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}.

Note that the additive deformation of a unitarily invariant model investigated in [29] enters in the framework of [28] so that a similar supercritical regime can be deduced for the matrix model (2.1) under some assumptions. (See also [30] for Information-Plus-Noise type models in the isotropic case).

The proofs of [17, 115] are based on the explicit 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. Note that Bloemental and Virag [42, 43] used an approach based on the continuum operator limit at the general beta soft edge, developed by Ramirez, Rider and Virág [119]. Both of these approaches rely heavily on the Gaussian nature of the non-deformed model.

In [29] and [30], Benaych-Georges and Nadakuditi exhibited a clever, now famous, key idea consisting 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.

In order to study the phase transition for deformed non-Gaussian or non-isotropic matrix ensembles, and in particular address the question of spectral universality, we investigated in the papers [C13] and [C16] different approaches presented in the following Section 2.4. Note that the subsequent paper [98] is in the lineage of our work.

2.4 Our first contributions in [C13] and [C16] for finite rank perturbations of Wigner matrices

Let H_N be a $N \times N$ Hermitian matrix such that the random variables $(H_N)_{ii}$, $\sqrt{2}\Re[(H_N)_{ij}]$, $\sqrt{2}\Im[(H_N)_{ij}]$ $i < j$, are independent identically distributed with a symmetric distribution τ of variance σ^2 and satisfying a Poincaré inequality. Let $M_N = W_N + A_N$ where $W_N = \frac{1}{\sqrt{N}}H_N$. Let A_N be a deterministic Hermitian matrix of fixed finite rank r whose distinct eigenvalues are J fixed real numbers $\theta_1 > \cdots > \theta_J$ independent of N with some j_0 such that $\theta_{j_0} = 0$. We assume that the non-null eigenvalues θ_j of A_N are of fixed multiplicity k_j (with $\sum_{j \neq j_0} k_j = r$). Define

$$\rho_{\theta_j} = \theta_j + \frac{\sigma^2}{\theta_j}. \tag{2.2}$$

Observe that $\rho_{\theta_j} > 2\sigma$ (resp. $< -2\sigma$) when $\theta_j > \sigma$ (resp. $< -\sigma$) (and $\rho_{\theta_j} = \pm 2\sigma$ if $\theta_j = \pm\sigma$).
For definiteness, we set $k_1 + \cdots + k_{j-1} := 0$ if $j = 1$.

2.4.1 Almost sure convergence of outliers

In [C13], we first extend Theorem 9 as follows.

Theorem 15. *Let* $J_{+\sigma}$ *(resp.* $J_{-\sigma}$*) be the number of* j*'s such that* $\theta_j > \sigma$ *(resp.* $\theta_j < -\sigma$*).*

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

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

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

(4) $\forall j \geq J - J_{-\sigma} + 1, \forall 1 \leq i \leq k_j, \quad \lambda_{k_1+\cdots+k_{j-1}+i}(M_N) \longrightarrow \rho_{\theta_j} \quad a.s.$

The main point in the proof consists in establishing that for any $\varepsilon > 0$, almost surely,

$$\mathrm{Spect}(M_N) \subset K_\sigma^\varepsilon(\theta_1, \cdots, \theta_J) \tag{2.3}$$

for all N large, where

$$K_\sigma^\varepsilon(\theta_1, \cdots, \theta_J) = K_\sigma(\theta_1, \cdots, \theta_J) + (-\varepsilon, \varepsilon)$$

and

$$K_\sigma(\theta_1, \cdots, \theta_J) := \left\{\rho_{\theta_J}; \cdots; \rho_{\theta_{J-J_{-\sigma}+1}}\right\} \cup [-2\sigma, 2\sigma] \cup \left\{\rho_{\theta_{J_{+\sigma}}}; \cdots; \rho_{\theta_1}\right\}.$$

To prove such an inclusion of the spectrum of M_N, we use the methods developed by U. Haagerup and S. Thorbjørnsen in [77], by H. Schultz [125]. Define $g_{\mu_{sc}}(z) = \int_{\mathbb{R}} \frac{d\mu_{sc}(t)}{z-t}$ where μ_{sc} is the semicircular distribution defined by (1.3) and $g_N(z) = \int_{\mathbb{R}} \frac{d\mu_{M_N}(t)}{z-t}$. The key point of this approach is to obtain a precise estimation at any point $z \in \mathbb{C}\backslash\mathbb{R}$ of the following type

$$g_\sigma(z) - g_N(z) + \frac{1}{N} L_\sigma(z) = O(\frac{1}{N^2}), \tag{2.4}$$

where L_σ is the Stieltjes transform of a distribution Λ_σ with compact support in $K_\sigma(\theta_1, \cdots, \theta_J)$ and $O(\frac{1}{N^2})$ can be bounded by $\frac{P(\Im z)}{N^2}$ where P is some polynomial P with nonnegative coefficients. Indeed such an estimation allows us through the inverse Stieltjes transform and some variance estimates to deduce that a.s., $\mathrm{tr}_N(1_{^cK_\sigma^\varepsilon(\theta_1,\cdots,\theta_J)}(M_N)) = O(N^{-4/3})$, denoting the normalized trace by tr_N. Thus the number of eigenvalues of M_N in ${}^cK_\sigma^\varepsilon(\theta_1, \cdots, \theta_J)$ is almost surely a $O(N^{-1/3})$ and since for each N this number has to be an integer, we deduce that it is actually equal to zero for large N. The main technical tool to establish (2.4) is the following approximate integration by part lemma (see [87]).

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

$$\mathbb{E}(\xi\phi(\xi)) = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!} \mathbb{E}(\phi^{(a)}(\xi)) + \epsilon \tag{2.5}$$

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

Actually, given an interval $[a, b] \subset \mathbb{R} \setminus K_\sigma(\theta_1, \cdots, \theta_J)$, saying by (2.3) that $[a, b]$ does not contain eigenvalues of M_N can be improved: it corresponds to $[a, b]$ some interval $I_{a,b}$ lying outside the spectrum of A_N and such that the number of eigenvalues of M_N in one side of $[a, b]$ is equal to the one of A_N in the corresponding side of $I_{a,b}$. Following [10], we say that there is exact separation of eigenvalues of the matrices A_N and M_N. Now, we can adapt the arguments needed for the conclusion of [18] viewing the deformed Wigner model as the additive analogue of the spiked population model, to deduce Theorem 15. This whole approach will be generalized and developed in Section 3.3 dealing with full rank deformations.

2.4.2 Nonuniversality of the fluctuations of outliers, dependence of the fluctuations with respect to the eigenvectors of the perturbation

In [C13], we show that 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. Let us 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 $M_N^{(i)} = W_N + A_N^{(i)}$ $(i = 1, 2)$ around $\rho_\theta := \theta + \frac{\sigma^2}{\theta}$ are given as follows :

Proposition 2. *1. Delocalized case [67] : The largest eigenvalue $\lambda_1(M_N^{(1)})$ have Gaussian fluctuations,*

$$\sqrt{N}(\lambda_1(M_N^{(1)}) - \rho_\theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(1 - \sigma^2/\theta^2)). \tag{2.6}$$

2. Localized case [C13]: The largest eigenvalue $\lambda_1(M_N^{(2)})$ fluctuates as

$$\sqrt{N}(1 - \frac{\sigma^2}{\theta^2})(\lambda_1(M_N^{(2)}) - \rho_\theta) \xrightarrow{\mathcal{D}} \tau \star \mathcal{N}(0, v_\theta). \tag{2.7}$$

where τ 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 τ.

The proof of (2.6) by Féral and Péché is combinatorial and is based on the computation of large moments of $M_N^{(1)}$. The proof of (2.7) is close to the one of [114] and the ideas of [36]; the key idea is to associate the fluctuations of the outliers with quadratic forms involving the resolvent of the unperturbed Wigner matrix and to use a central limit theorem proved in an Appendix by J. Baik and J. Silverstein in [C13].

In [C16], we exhibit quite general situations with respect to the eigenvectors of the perturbation matrix that will give rise to universality or non universality of the fluctuations as follows.

Let M_N be a deformed Wigner matrix as defined at the beginning of Section 2.4. We denote by $k_{+\sigma} := k_1 + \cdots + k_{J_{+\sigma}}$. We introduce $k \geq k_{+\sigma}$ as the minimal number of canonical vectors among the canonical basis $(e_i;\, i = 1, \ldots, N)$ of $\mathbb{C}^N$ needed to express all the eigenvectors associated to the largest eigenvalues $\theta_1, \ldots, \theta_{J_{+\sigma}}$ of A_N. Without loss of generality (using the invariance of the distribution of the Wigner matrix W_N by conjugation by a permutation matrix), we can assume that these $k_{+\sigma}$ eigenvectors belong to $\mathrm{Vect}(e_1, \ldots, e_k)$. We assume that $k \ll \sqrt{N}$.

Let us now fix j such that $1 \leq j \leq J_{+\sigma}$ and let U_k be a unitary matrix of size k such that

$$\mathrm{diag}(U_k^*, I_{N-k}) A_N \mathrm{diag}(U_k, I_{N-k}) = \mathrm{diag}(\theta_j I_{k_j}, (\theta_l I_{k_l})_{l \leq J_{+\sigma}, l \neq j}, Z_{N-k_{+\sigma}}) \tag{2.8}$$

where $Z_{N-k_{+\sigma}}$ is an Hermitian matrix with eigenvalues strictly smaller than $\theta_{J_{+\sigma}}$.

Define $K_j = K_j(N)$ as the minimal number of canonical vectors among $(e_1, \ldots, e_k)$ needed to express all the orthonormal eigenvectors v_i^j, $1 \leq i \leq k_j$, of A_N associated to θ_j. Without loss of generality, we can

assume that the v_i^j , $1 \leq i \leq k_j$, belong to $\mathrm{Vect}(e_1, \ldots, e_{K_j})$. Considering now the vectors v_j^i as vectors in $\mathbb{C}^{K_j}$, we define the $K_j \times k_j$ matrix

$$U_{K_j \times k_j} := \Big(v_1^j, \ldots, v_{k_j}^j \Big) \tag{2.9}$$

namely $U_{K_j \times k_j}$ is the upper left corner of U_k of size $K_j \times k_j$. It satisfies

$$U_{K_j \times k_j}^* U_{K_j \times k_j} = I_{k_j}. \tag{2.10}$$

From Theorem 15, for all $1 \leq i \leq k_j$, $\lambda_{k_1 + \cdots + k_{j-1} + i}(M_N)$ converges to ρ_{θ_j} a.s.. We shall describe their fluctuations in the extreme two cases:

Case a) localization of the eigenvectors associated to θ_j: The sequence $K_j(N)$ is bounded,

$$\sup_N K_j(N) = \tilde{K_j}$$

and the the upper left corner $U_{\tilde{K_j} \times k_j}$ of U_k of size $\tilde{K_j} \times k_j$ converges towards some matrix $\tilde{U}_{\tilde{K_j} \times k_j}$ when N goes to infinity;

Case b) delocalization of the eigenvectors associated to θ_j: $K_j = K_j(N) \to \infty$ when $N \to \infty$ and U_k satisfies

$$\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}| \longrightarrow 0 \quad \text{as} \quad N \to \infty. \tag{2.11}$$

The main results are the following two theorems. Set $m_4 := \int x^4 d\tau(x)$ and define c_{θ_j} by

$$c_{\theta_j} = \frac{\theta_j^2}{\theta_j^2 - \sigma^2}. \tag{2.12}$$

In Case a) the fluctuations of the corresponding rescaled largest eigenvalues of M_N are not universal.

Theorem 16. *In Case a): the k_j-dimensional vector*

$$\Big(c_{\theta_j} \sqrt{N} (\lambda_{k_1 + \ldots + k_{j-1} + i}(M_N) - \rho_{\theta_j});\, i = 1, \ldots, k_j \Big)$$

converges in distribution to $(\lambda_i(V_{k_j \times k_j}); i = 1, \ldots k_j)$ where $\lambda_i(V_{k_j \times k_j})$ are the ordered eigenvalues of the matrix $V_{k_j \times k_j}$ of size k_j defined in the following way. Let $W_{\tilde{K}_j}$ be an Hermitian matrix of size $\tilde{K}_j$ such that the entries $\sqrt{2}\Re(W_{\tilde{K}_j})_{uv}, \sqrt{2}\Im(W_{\tilde{K}_j})_{uv}, u < v, (W_{\tilde{K}_j})_{uu}$ are independent random variables with distribution τ. Let $\mathcal{G}_{\tilde{K}_j}$ be a centered Hermitian Gaussian matrix of size $\tilde{K}_j$ independent of $W_{\tilde{K}_j}$ with independent entries $\mathcal{G}_{pl}$, $p \leq l$ with variance

$$\begin{cases} v_{pp} = E(\mathcal{G}_{pp}^2) = \dfrac{1}{2}\Big(\dfrac{m_4 - 3\sigma^4}{\theta_j^2} \Big) + \dfrac{\sigma^4}{\theta_j^2 - \sigma^2},\ p = 1, \ldots, \tilde{K}_j, \\ v_{pl} = \mathbb{E}(|\mathcal{G}_{pl}|^2) = \dfrac{\sigma^4}{\theta_j^2 - \sigma^2},\ 1 \leq p < l \leq \tilde{K}_j. \end{cases} \tag{2.13}$$

Then, $V_{k_j \times k_j}$ is the $k_j \times k_j$ matrix defined by

$$V_{k_j \times k_j} = \tilde{U}_{\tilde{K}_j \times k_j}^* (W_{\tilde{K}_j} + \mathcal{G}_{\tilde{K}_j}) \tilde{U}_{\tilde{K}_j \times k_j}. \tag{2.14}$$

Case b) exhibits universal fluctuations.

Theorem 17. *In Case b): the k_j-dimensional vector*

$$\Big(c_{\theta_j}\sqrt{N}(\lambda_{k_1+\ldots+k_{j-1}+i}(M_N)-\rho_{\theta_j});\, i=1,\ldots,k_j\Big)$$

converges in distribution to $(\lambda_i(V_{k_j\times k_j});\, i=1,\ldots k_j)$ where the matrix $V_{k_j\times k_j}$ is distributed as a $k_j\times k_j$ GUE matrix whose entries have a variance equal to $\frac{\theta_j^2\sigma^2}{\theta_j^2-\sigma^2}$.

The proof is a (quite technical) generalization of the proof of (2.7); using determinant identities, we show that each of these rescaled eigenvalues is an eigenvalue of a $k_j\times k_j$ random matrix which may be expressed in terms of the resolvent of a $(N-k)\times(N-k)$ deformed Wigner matrix whose eigenvalues do not jump asymptotically outside $[-2\sigma;2\sigma]$; then, the matrix $V_{k_j\times k_j}$ will arise from a multidimensional CLT from [14] on random sesquilinear forms.

Remark 1. *Note that since τ is symmetric, analogue results can be deduced from Theorem 16 and Theorem 17 dealing with the lowest eigenvalues of M_N and the θ_j such that $\theta_j<-\sigma$.*

Example:

$$A_N=\mathrm{diag}(A_p(\theta_1),\theta_2 I_{k_2},0_{N-p-k_2})$$

where $A_p(\theta_1)$ is a matrix of size p defined by $A_p(\theta_1)_{ij}=\theta_1/p$, with $\theta_1,\theta_2>\sigma$, $p\ll\sqrt{N}$. Then $k=p+k_2$, $k_1=1$, $K_1=p$, $K_2=k_2$. For $j=1$, we are in Case a) if p is bounded and in Case b) if $p=p(N)\to+\infty$. For $j=2$, we are in Case a).

Dealing with a spike $\theta_j>\sigma$ with multiplicity 1, it turns out that case b) is actually the unique situation where universality holds since we establish the following.

Theorem 18. *If $k_j=1$, $\theta_j>\sigma$, then the fluctuations of $\lambda_{k_1+\cdots+k_{j-1}+1}(M_N)$ are universal, namely*

$$\sqrt{N}(\lambda_{k_1+\cdots+k_{j-1}+1}(M_N)-\rho_{\theta_j})\xrightarrow{\mathcal{L}}\mathcal{N}(0,\sigma_{\theta_j}^2)\ \textit{where}\ \sigma_{\theta_j}=\sigma\sqrt{1-\frac{\sigma^2}{\theta_j^2}},$$

if and only if

$$\max_{l\le K_j}|(U_k)_{l1}|\to 0\ \textit{when}\ N\to\infty. \tag{2.15}$$

Moreover, our approach allows us to describe the fluctuations of $\lambda_{k_1+\cdots+k_{j-1}+1}(M_N)$ for some particular situations where the corresponding eigenvector of A_N is not localized but does not satisfy the criteria of universality $\max_{l\le K_j}|(U_k)_{l1}|\to 0$ (that is somehow for intermediate situations between Case a) and Case b)). Let m be a fixed integer number. Assume that for any $l=1,\ldots,m$, $(U_k)_{l1}$ is independent of N, whereas $\max_{m<l\le K_j}|(U_k)_{l1}|\to 0$ when N goes to infinity. We can prove that $c_{\theta_j}\sqrt{N}(\lambda_{k_1+\cdots+k_{j-1}+1}(M_N)-\rho_{\theta_j})$ converges in distribution towards the mixture of τ-distributed or gaussian random variables $\sum_{i,l=1}^m a_{il}\xi_{il}+\mathcal{N}$ where $\xi_{il},(i,l)\in\{1,\ldots,m\}^2$, $\mathcal{N}$ are independent random variables such that

- for any $(i,l)\in\{1,\ldots,m\}^2$, the distribution of ξ_{il} is τ;
- $a_{il}=\begin{cases}\sqrt{2}\Im(\overline{(U_k)_{l1}}(U_k)_{i1}) \text{ if } i<l\\ \sqrt{2}\Re(\overline{(U_k)_{l1}}(U_k)_{i1}) \text{ if } i>l\\ |(U_k)_{l1}|^2 \text{ if } i=l;\end{cases}$

- $\mathcal{N}$ is a centered gaussian variable with variance

$$\frac{1}{2}\frac{\left[m_4-3\sigma^4\right]\sum_{l=1}^m |(U_k)_{l1}|^4}{\theta_j^2}+\frac{\sigma^4}{\theta_j^2-\sigma^2}+\left[1-\left(\sum_{l=1}^m |(U_k)_{l1}|^2\right)^2\right]\sigma^2.$$

2.5 Some subsequent works on outliers of finite rank deformed models

- In [117], Pizzo, Renfrew and Soshnikov extend Theorems 15, 16, 17 to more general Wigner matrices. In the paper [121], Renfrew and Soshnikov extend quite significantly Theorem 17 as follows. They assume that the entries of the Wigner matrix are centered independent random variables, the variance of off-diagonal entries are all equal to σ^2, the off-diagonal (resp. diagonal) entries have uniformly bounded fourth (resp. second) moments and satisfy a Lindeberg type condition for the fourth (resp. second) moments. We use the notations introduced at the beginning of Section 2.4 and in (2.12), (2.2). Assume that $|\theta_j|>\sigma$. Let $u_N^1,\ldots,u_N^{k_j}$ be an orthonormal set of eigenvectors of the perturbation associated with θ_j. Define the $N\times N$ matrix M_3 by

$$(M_3)_{ij}=\mu_{3,ij}(1-\delta_{ij})$$

where

$$\mu_{3,ij}=\mathbb{E}\left[|W_{ij}|^2W_{ij}\right].$$

Then Renfree and Soshnikov prove that, in the case where the l^∞ norm of every orthonormal eigenvector of the perturbation corresponding to θ_j goes to zero as $N\to+\infty$, the difference between the k_j dimensional vector

$$\left\{c_{\theta_j}\sqrt{N}\left(\lambda_{k_1+\cdots+k_{j-1}+i}-\rho_j\right), i=1,\ldots,k_j\right\}$$

and the vector formed by the ordered eigenvalues of a $k_j\times k_j$ GUE matrix (with the variance of the matrix entries given by $\frac{\theta_j^2\sigma^2}{\theta_j^2-\sigma^2}$) plus a deterministic matrix with entries given by $\frac{1}{\theta_j^2N}(u_N^l)^*M_3u_N^p$ converges to zero in probability. Note that this additional matrix $\frac{1}{\theta_j^2N}\left((u_N^l)^*M_3u_N^p\right)_{1\leq l,p\leq k_j}$ appeared neither in [115], [C16] or [117] nor in [67] since in [115] the entries of the Wigner matrix are Gaussian and in [C16] and [67] the distribution entries of the Wigner matrix is symmetric so that $M_3\equiv 0$ and, in [117], $k=o(\sqrt{N})$ so that

$$\frac{1}{\theta_j^2N}\left((u_N^l)^*M_3u_N^p\right)_{l,p}\to_{N\to+\infty}0.$$

The first step in proving the above result is to associate the fluctuations of the outliers with quadratic forms of the resolvent R_N of the unperturbed Wigner matrix. Then, the proof comes down to study the asymptotic of

$$\sqrt{N}\left(\langle u_N^l,R_N(\rho_j)u_N^p\rangle-\frac{1}{\theta_j}\delta_{lp}\right)_{1\leq l,p\leq k_j}.$$

- Dealing with finite rank deformation of Wigner matrices, Knowles and Yin in [89], allow the eigenvalues θ_i of the deformation to depend on N under the condition that

$$||\theta_i|-1|\geq(\log N)^{C\log\log N}N^{-1/3},$$

meaning that $\rho_{\theta_i} - 2$ is much greater than $N^{-2/3}$ (setting $\sigma = 1$). If $r = 1$, the outlier associated with θ_1 fluctuates around ρ_{θ_1} on the scale $N^{-1/2}(|\theta_1| - 1)^{1/2}$ and the asymptotic distribution is explicitely computed by the authors. In [89], the authors establish actually such a result for any r but if

$$\forall j, |\theta_j| > 1, \min_{j \neq i} |\theta_i - \theta_j| \geq (\log N)^{C \log \log N} N^{-1/2}((|\theta_i| - 1)^{-1/2}$$

which means that the outlier associated with θ_i does not overlap with the other outliers since their separation is greater than the scale on which they fluctuate. [90] extends the results of [89] in two directions: the authors allow overlapping outliers and describe the joint distribution of all outliers. Their proof relies on the isotropic local semicircle law

The next chapter will put forward tools from free probability theory to provide a unified understanding, for all the previous matrix models and dealing with full rank deformations, of several phenomena such as the exact separation phenomenon, the appearence, localization and fluctuations of outliers as well as of the asymptotic behaviour of the corresponding eigenvectors. The main principle of subordination in free probability is emphasized as a main tool in this understanding.

Chapter 3

Full rank perturbation of classical random matrix models

This chapter first tries to provide a unified understanding based on free subordination properties of the spectral properties (limiting support, behavior of the density at edges, exact separation phenomenon) of the three deformed models i), ii) and iii) introduced at the beginning of Chapter 2, dealing now with *full* rank deformations. This chapter actually brings to light a general methodology to localize outliers and to describe the asymptotic behaviour of the corresponding eigenvectors for all the preceding models, both in the *independent case* as well as for the corresponding deformed *isotropic* models, namely 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.
It was the purpose of [$C22$] to put forward an unified understanding based on subordination in free probability. This investigation relies notably on [$C15$], [$C17$], [$C18$], [$C19$], [$C21$] and [$C24$].

3.1 Convergence of spectral measures

This section is devoted to the study of the limiting spectral distribution (LSD) of the deformed matrix M_N for the different models i) to iii). We now assume that for some probability ν, the spectrum of the deformation A_N satisfies

$$\begin{aligned} &\mu_{A_N} \underset{N\to+\infty}{\longrightarrow} \nu \qquad \text{weakly, for cases i) and ii)} \\ &\mu_{A_N A_N^*} \underset{N\to+\infty}{\longrightarrow} \nu \quad \text{weakly, for case iii).} \end{aligned}$$

The following theorem is a review of the pionner results concerning the limiting spectral distribution for the three deformed models in the independent case. The spectral distribution is characterized via an equation satisfied by its Stieltjes transform. (For any probability measure τ on $\mathbb{R}$, we denote by g_τ the Stieltjes transform of τ defined for any $z \in \mathbb{C}^+ := \{z \in \mathbb{C}, \Im z > 0\}$ by $g_\tau(z) = \int \frac{1}{z-t} d\tau(t)$.)

Theorem 19. *Convergence of the spectral measure μ_{M_N}*

i) Deformed Wigner matrices ([111], [4] Theorem 5.4.5)

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

with

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_1}(z) = \int \frac{1}{z - \sigma^2 g_{\mu_1}(z) - t} \mathrm{d}\nu(t). \tag{3.1}$$

ii) Sample covariance matrices ([101, 130])

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

with

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_2}(z) = \int \frac{1}{z - t(1 - c + cz g_{\mu_2}(z))} \mathrm{d}\nu(t). \tag{3.2}$$

iii) Information-Plus-Noise type matrices ([56, 155])

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

with

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_3}(z) = \int \frac{1}{(1 - c\sigma^2 g_{\mu_3}(z))z - \frac{t}{1 - c\sigma^2 g_{\mu_3}(z)} - \sigma^2(1 - c)} \mathrm{d}\nu(t). \tag{3.3}$$

Remark 2. *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 the non-deformed model) and only depend on A_N through the limiting spectral measure ν.*

Remark 3. *If $\nu = \delta_0$ in* (3.1)*, we recover the equation satisfied by the Stieltjes transform of the semi-circular distribution. The same is true for the Marchenko-Pastur distribution in* (3.2) *with $\nu = \delta_1$ and in* (3.3) *with $\nu = \delta_0$.*

Remark 4. *Such functional equations for Stieltjes transforms have been obtained for deformations of isotropic models by [113] and [143].*

3.2 Study of LSD of deformed ensembles through free probability theory

In this section, we take an other look of the LSD described in Section 3.1, in the light of free probability theory introduced in the Appendix. 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 3.3. 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 be discussed in Section 3.5.

3.2.1 Free probabilistic interpretation of LSD

As noticed in Remark 2, the limiting spectral distributions of the deformed models investigated in Section 3.1 are **universal** in the sense that they do not depend on the distribution of the entries of the non-deformed model. Therefore, choosing Gaussian entries and applying Theorem 43 and Theorem 44, 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 19.

- 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 (3.1), (3.2) and (3.3) satisfied by the limiting Stieltjes transforms correspond to free subordination properties introduced in Section 5.5 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. We use here the notations introduced in Section 5.5.

- Deformed Wigner matrices

$$\forall z \in \mathbb{C}^+, \quad g_{\mu_1}(z) = g_\nu(\omega_{\mu_{sc},\nu}(z)).$$

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

- Sample covariance matrices

$$\forall z \in \mathbb{C}^+, \psi_{\mu_1}\left(\frac{1}{z}\right) = \psi_\nu(F_{\mu_{\mathrm{MP}},\nu}\left(\frac{1}{z}\right)) \tag{3.4}$$

with

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

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

- Information-Plus-Noise type matrices

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

$$\forall z \in \mathbb{C}^+, H^{(c)}_{\sqrt{\mu_3}}\left(\frac{1}{z}\right) = H^{(c)}_{\sqrt{\nu}}\left(\Omega_{\mu_{\mathrm{MP}},\nu}\left(\frac{1}{z}\right)\right), \tag{3.5}$$

with

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

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

3.2.2 Limiting supports of LSD

For each deformed model involving independent entries, several authors studied the limiting support [33, 47, 57, 97, 98], [C18]. 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 $\mu_1 = \mu_{sc} \boxplus \nu$

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

$$\mathcal{O}_1 := \{u \in \mathbb{R} \setminus \operatorname{supp}(\nu), \phi_1'(u) > 0\},$$
$$u \in \mathbb{R} \setminus \operatorname{supp}(\nu), \quad \phi_1(u\ ) = u + \sigma^2 g_\nu(u),$$
$$x \in \mathbb{R} \setminus \operatorname{supp}(\mu_1), \quad \varphi_1(x\ ) = x - \sigma^2 g_{\mu_1}(x).$$

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

$$\mathbb{R} \setminus \{\operatorname{supp}(\mu_2)\} \underset{\phi_2}{\overset{\varphi_2}{\rightleftarrows}} \mathcal{O}_2 \subset \mathbb{R} \setminus \{\operatorname{supp}(\nu)\}, \tag{3.6}$$

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

$$u \in \mathbb{R} \setminus \operatorname{supp}(\nu), \quad \phi_2(u) = u + cu \int \frac{t}{u-t} \mathrm{d}\nu(t),$$
$$x \in \mathbb{R} \setminus \operatorname{supp}(\mu_2), \quad \varphi_2(x) = \begin{cases} \frac{x}{(1-c)+cxg_{\mu_2}(x)} & \text{if } \ c < 1 \\ \frac{1}{g_{\mu_2}(x)} & \text{if } \ c = 1. \end{cases}$$

Note that φ_2 is well defined on $\mathbb{R} \setminus \operatorname{supp}(\mu_2)$ since its denominator never vanishes according to Lemma 6.1 in [11].

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

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

$$\mathcal{O}_3 = \left\{u \in \mathbb{R} \setminus \operatorname{supp}(\nu), \phi_3^{'}(u) > 0, g_\nu(u) > -\frac{1}{\sigma^2 c}\right\},$$
$$u \in \mathbb{R} \setminus \operatorname{supp}(\nu), \quad \phi_3(u) = u(1 + c\sigma^2 g_\nu(u))^2 + \sigma^2(1-c)(1 + c\sigma^2 g_\nu(u)), \tag{3.7}$$
$$x \in \mathbb{R} \setminus \operatorname{supp}(\mu_3), \quad \varphi_3(x) = x(1 - c\sigma^2 g_{\mu_3}(x))^2 - (1-c)\sigma^2(1 - c\sigma^2 g_{\mu_3}(x)).$$

Note that φ_1 corresponds to the extension of $\omega_{\mu_{sc},\nu}$ on $\mathbb{R} \setminus \operatorname{supp}(\mu_1)$ and for $i = 2, 3$, φ_i coincides on $\mathbb{R} \setminus \{\operatorname{supp}(\mu_i) \cup \{0\}\}$, 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 [33] and [C18] 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 [11] readily yield (3.6).

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 20. *[C15] [C18] 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 $l \in \{1, \ldots, p\}$,

$$\mu_i([\phi_i(u_l^-), \phi_i(v_l^+)]) = \nu([u_l, v_l]). \tag{3.8}$$

Using the characterization of the support (3.6), Remark 3.6 in [C17] and the fact that from [20] 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 20. Note that the latter cases occur only when ν has a Dirac mass at zero since from [20], $\mu_{\mathrm{MP}} \boxtimes \nu(\{0\}) = \max(\mu_{\mathrm{MP}}(\{0\}), \nu(\{0\}))$ 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. (3.8) can be seen as a consequence of the matricial exact separation phenomenon described in Theorem 3.13 below letting N go to infinity.

3.2.3 Behavior of the density at edges

Biane proved in [33] that $\mu_1 = \mu_{sc} \boxplus \nu$ has a continuous density. Choi and Silverstein [47] and Dozier and Silverstein [57] proved respectively that, away from zero, $\mu_2 = \mu_{\mathrm{MP}} \boxtimes \nu$ and $\mu_3 = (\sqrt{\mu_{\mathrm{MP}}} \boxplus_c \sqrt{\nu})^2$ possess a continuous density. Let us denote any of theses densities by f.
Using the notations of Theorem 20, 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 f may exhibit different behaviour. This can be seen for instance in the following example investigated by Lee and Schnelli [92]:

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

where $a < 1, b > 1$, h 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

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

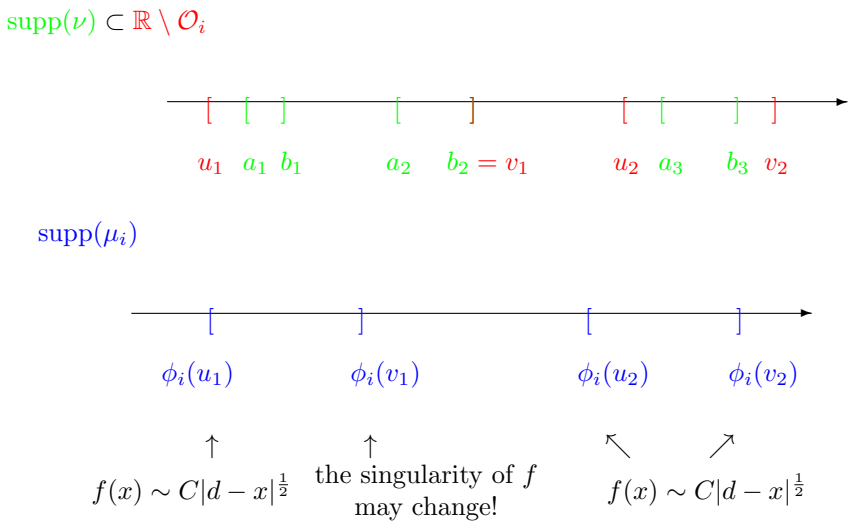
It can be easily seen that for all $\sigma > \sigma_0$, $\mathbb{R} \setminus \mathcal{O}_1 = [u_\sigma, v_\sigma]$ with $u_\sigma < -1 < 1 < v_\sigma$, so that

$$\mathrm{supp}\,(\mu_{sc} \boxplus \nu) = [\phi_1(u_\sigma), \phi_1(v_\sigma)]; \tag{3.9}$$

thus, $\phi_1(v_\sigma)$ is a regular edge and we have $f(x) \sim C(\phi_1(v_\sigma) - x)^{\frac{1}{2}}$.
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

$$f(x) \sim C(\phi_1(1) - x)^b. \tag{3.10}$$

We illustrate by the following picture the difference of behaviour of the density f 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 $\mathbb{R} \setminus \mathcal{O}_i$.



3.3 Outliers of general spiked models

In Section 2.2, 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 3.2 allows us to understand the appearence of outliers of general spikes models that is dealing with full rank deformations and provides the correct way to generalize the pioneering works. Actually, the relevant criterion for a spiked eigenvalue of the deformation to generate an outlier in the spectrum of the deformed model is to belong to the set $\mathcal{O}_i$ related to the subordination functions and introduced in Section 3.2.2. In the finite rank case this criterion reduces to a critical threshold.

In order to adopt universal notations for the three types of deformations, we set

$$\tilde{A}_N = \begin{cases} A_N & \text{for additive or multiplicative deformations} \\ A_N A_N^* & \text{for Information-plus-noise type deformation.} \end{cases} \tag{3.11}$$

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, $\sum_j k_j = r$.

3.3.1 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 free subordination property (see Section 5.5)

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

For an Hermitian deformed model such that

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

the intuition is that

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

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

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

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: independent and isotropic case. Nevertheless, their proofs are different. In the following, we present the distinct approaches.

The independent case

In this section, we will denote by $\phi, \varphi, \mathcal{O}, \mu$, any of $\phi_i, \varphi_i, \mathcal{O}_i, \mu_i$, for $i = 1, 2, 3$ introduced in Section 3.2.2 related to the investigated deformed model. We choose to present a unified result covering the three types of deformations.

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 [9] in the multiplicative case. We now express this deterministic measure ν_N in the three models :

i) Additive deformation of a Wigner matrix [C15]:

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

ii) Multiplicative deformation of a sample covariance matrix ([9])

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

iii) Information plus noise model ([97] in the Gaussian case, [12], [C24])

$$\nu_N = (\sqrt{\mu_{\mathrm{MP}}} \boxplus_{c_N} \sqrt{\mu_{A_N A_N^*}})^2 \text{ (with } \mathrm{c_N = N/p}).$$

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 which holds for the three models (we refer to the original papers for the precise technical assumptions on the models):

Proposition 3. *$\forall \epsilon > 0$, $\forall [a, b] \subset (K_N)_\epsilon$,*

$$\mathbb{P}(\textit{ for large } N, \textit{ } M_N \textit{ has no eigenvalue in } [a, b]) = 1 \tag{3.12}$$

For deformed Wigner matrices, (3.12) is established in in [C15, Theorem 5.1] under the assumption that the real part and the imaginary part of the off-diagonal entries of the Wigner matrix are iid with variance one half and the common distribution is symmetric and satisfies a Poincaré inequality; its proof uses the strategy developed by [77] and [78] already described in Section 2.4.1. Note that [C23] that will be presented in the following Chapter (dealing with much more general polynomial models) extends to more general Wigner matrices. The same method was used in [97] for Gaussian Information-Plus-Noise type matrices.

To prove the inclusion of the spectrum (3.12) for sample covariance matrices and information-plus-noise type matrices [9, 12], Bai and Silverstein provide a different approach strongly related to the models although it also makes use of Stieltjes transform. Note that, in an Appendix in [C24], we present alternative versions of the earlier result of [12] for non-Gaussian Information-Plus-Noise type matrices where we remove some technical assumptions that were difficult to handle but assume that the real part and the imaginary part of the entries of the random matrix are independent with variance one half . Note also that in the context of [12], (3.12) is proved only for $a > 0$ whereas it includes the case $a = 0$ in the framework of [C24].

b) An exact separation phenomenon

A next step in the analysis of the spectrum of deformed models dealing with independent entries 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 3.2.2, 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, $[a-\delta, b+\delta] \subset \mathbb{R} \setminus \text{supp}\ (\nu_N)$. 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 $i_N \in \{0, \ldots, N\}$ such that

$$\lambda_{i_N+1}(\tilde{A}_N) < \varphi(a) \quad \text{and} \quad \lambda_{i_N}(\tilde{A}_N) > \varphi(b). \tag{3.13}$$

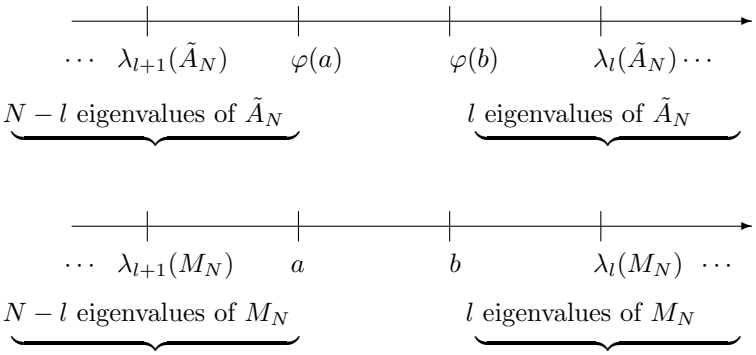
Moreover, $[a, b]$ splits the spectrum of M_N exactly as $[\varphi(a), \varphi(b)]$ splits the spectrum of $\tilde{A}_N$ as stated by the following.

Theorem 21. *With i_N satisfying* (3.13)*, one has*

$$\mathbb{P}[\lambda_{i_N+1}(M_N) < a \ \textit{and}\ \lambda_{i_N}(M_N) > b, \textit{for all large } N] = 1. \tag{3.14}$$

The following picture illustrates this exact separation phenomenon.

$$[a, b] \subset \mathbb{R} \setminus \text{supp}(\nu_N) \longleftrightarrow [\varphi(a), \varphi(b)]$$
$$\text{gap in Spect}(M_N) \longleftrightarrow \text{gap in Spect}(\tilde{A}_N)$$



This was first observed by Bai and Silverstein [10], in the case of sample covariance matrices. We refer to [C15] for deformed Wigner matrices and to Loubaton-Vallet [98] (Gaussian case), [C18], [C24] for Information-plus-noise type models when the real part and the imaginary part of the entries of the random matrix are independent with variance one half. This exact separation phenomenon leads asymptotically to the relation (3.8) 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 3.3.1 and is a consequence of the inclusion of the spectrum and the exact separation.

Theorem 22. *[15, 120] [C15], [C18] 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 3.2.2,*

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 $\mathrm{supp}(\nu)$ 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 $\mathrm{supp}(\nu)$ 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 $\alpha_j = 1 - \lim_N \frac{n_{j-1}}{N} = \nu(]-\infty, \theta_j])$.*

The isotropic case

Here we consider an additive spiked deformation of an isotropic matrix

$$M_N = A_N + U_N^* B_N U_N. \tag{3.15}$$

U_N is a unitary matrix whose distribution is the normalized Haar measure on the unitary group $\mathbb{U}(N)$. 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 $\mathrm{supp}(\mu)$ as $N \to \infty$. A_N is a deterministic Hermitian $N \times N$ perturbation as defined at the beginning of Section 3.3.

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 a.s. neither does M_N (see Corollary 3.1 in [48]) meaning that for all large N, all the eigenvalues of M_N are inside a small neighborhood of the support of $\mu \boxplus \nu$.

Assume now that A_N has outliers. Then we established in [C21] the following result.

Theorem 23. *Set $K = \mathrm{supp}(\mu \boxplus \nu)$,*

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

with ω_2 defined as in Section 5.5. 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 $\mathbb{U}(N)$ 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 outliers, and therefore [48, 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 [29] 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

$$\det(I + XY) = \det(I + YX). \tag{3.16}$$

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

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

so that using Sylvester's identity, we obtain that

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

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

$$F_N(z) = I_r - P_N (zI_N - M'_N)^{-1} P_N^* \Theta. \tag{3.17}$$

The key idea is now to establish an approximate matricial subordination result. Biane [34] 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

$$\mathbb{E}_{\mathbb{C}[\mathbf{a}]} \left[(z - (\mathbf{a} + \mathbf{b}))^{-1} \right] = (\omega(z) - \mathbf{a})^{-1}, \quad z \in \mathbb{C}^+. \tag{3.18}$$

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, where ω_2 is the subordination function from (5.15). For this task we use the following lemma.

Lemma 2. *Assume that $\varepsilon > 0$, and $T \in M_N(\mathbb{C})$ satisfies the inequality*

$$|k^*(TY - YT)h| \leq \varepsilon \|Y\|$$

for every rank one matrix $Y \in M_N(\mathbb{C})$ and all unit vectors $h, k \in \mathbb{C}^N$. Then for any w in the numerical range $W(T) = \{h^*Th : \|h\| = 1\}$, we have $\|T - wI_N\| \leq 2\varepsilon$.

Thanks to this lemma, we establish that there exists $(\omega^{(N)}(z))_N$ in $\mathbb{C}$ such that

$$\lim_{N\to\infty} \|\mathbb{E}\left[(zI_N - (A'_N + U_N^* B_N U_N))^{-1}\right]^{-1} + A'_N - \omega^{(N)}(w) I_N\| = 0,$$

and thus

$$\mathbb{E}\left[(zI_N - (A'_N + U_N^* B_N U_N))^{-1}\right] \approx \left(A_N - \omega^{(N)}(z) I_N\right)^{-1}.$$

Then we use asymptotic freeness and subordination property to prove that $\omega^{(N)}(z)$ converges towards $\omega_2(z)$.
Thus, it turns out that almost surely the sequence $\{F_N\}_N$ converges uniformly on compact subsets of $\mathbb{C} \setminus \text{supp}(\mu \boxplus \nu)$ to the analytic function F defined by

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

The set of points z such that $F(z)$ is not invertible is precisely $\bigcup_{i=1}^J \omega_2^{-1}(\{\theta_i\})$. Theorem 23 follows by Hurwitz's Theorem.

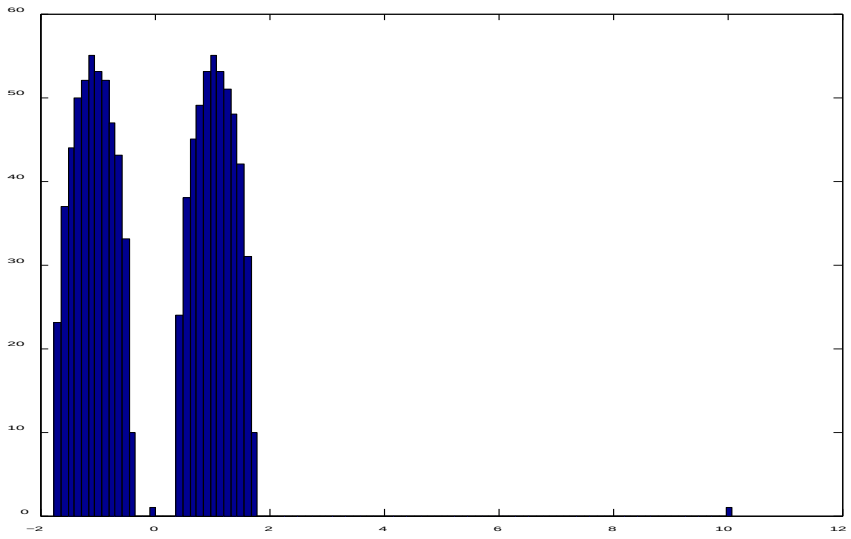
It follows from this result that a remarkable new phenomenon arises: a single spike of A_N can generate asymptotically several outliers of M_N. This arises from the fact that the restriction to the real line of some subordination functions may be many-to-one, that is, with the above notation, the set $\omega_i^{-1}(\{\theta\})$ may have cardinality strictly greater than 1, unlike the subordination function related to free convolution with a semicircular distribution that was used in Section 3.3.1. The following numerical simulation 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 $B_N = \text{diag}(\underbrace{-1, \ldots, -1}_{\frac{N}{2}}, \underbrace{1, \ldots, 1}_{\frac{N}{2}})$, 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 GUE.. 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 $\mathbb{R} \setminus \text{supp}\,(\mu_{sc} \boxplus \nu)$ but ω_1 may be many to one. The equation $\omega_1(\rho) = 10$ has 2 solutions ρ_1 and ρ_2.



Actually, in [C21] we also consider additive models where both A_N and B_N have spiked outliers ([C21]). Namely, let us consider the matricial model (3.15) where we 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 in [C21] for multiplicative perturbations of unitarily invariant matricial models, based on similar ideas, with the subordination function replaced by its multiplicative counterpart.

3.3.2 Eigenvectors

For a general perturbation, dealing with sample covariance matrices, S. Péché and O. Ledoit [116] 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 in our studies [C17], [C21], [C24], we 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 [C17], [C21] holds for spiked additive deformations in the independent case as well as in the isotropic case. Let A_N be a deterministic deformation as defined as the beginning of Section 3.3; dealing with either a deformed Wigner matrix or a deformed isotropic matrix as defined by (3.15), we have the following

Theorem 24. *Set* $K = \mathrm{supp}(\mu \boxplus \nu)$,

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

and let ω_2 *be the subordination function satisfying* (5.15). *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,*

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

Similar results are established for spiked multiplicative deformations in the independent case as well as in the isotropic case [C17], [C21] and for information-plus-noise type models in [C24] in the independent case. See the following tabular in Section 3.3.3. 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 3.2.2.

Theorem 25. *Let* $n_{j-1}+1, \ldots, n_{j-1}+k_j$ *the descending ranks of* θ_j *among the eigenvalues of* $\tilde{A}_N$ *defined by* (3.11) *and* $\xi(j)$ *a unit eigenvector associated to one of the eigenvalues* $(\lambda_{n_{j-1}+q}(M_N),\ 1 \leq q \leq k_j)$. *Then when* N *goes to infinity,*

- *For any* $\theta_l \neq \theta_j$,
$$\|P_{\mathrm{Ker}\ (\theta_l I_N - \tilde{A}_N)}(\xi(j))\| \to 0 \quad \textit{almost surely}$$
- $$\|P_{\mathrm{Ker}\ (\theta_j I_N - \tilde{A}_N)}(\xi(j))\|^2 \to \tau_j \quad \textit{almost surely}$$

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

Here are the common basic ideas [C17] of the proof of these results.
First note that if $u_1, \ldots, u_N$ and $w_1, \ldots, w_N$ are respectively a basis of eigenvectors associated with $\lambda_1(\tilde{A}_N), \ldots, \lambda_N(\tilde{A}_N)$ and with $\lambda_1(M_N), \ldots, \lambda_N(M_N)$, we have

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

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 $Tr\left[h(M_N) f(\tilde{A}_N)\right]$ for a fit choice of h and f. Then, a concentration of measure phenomenon reduces the problem to the study of $\mathbb{E}(Tr\left[h(M_N) f(\tilde{A}_N)\right])$.

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

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

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 $\{1,\dots,N\}$.

3.3.3 Unified understanding

In conclusion, solving the problem of outliers consists in solving an equation involving the relevant 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 1.1) or Y_N is unitarily invariant (resp. biunitarily invariant for the information-plus-noise type model). We use notations introduced in Section 5.5.

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

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 in the isotropic case has been proved only for finite rank perturbation A_N.

3.4 Fluctuations at edges of spiked deformed models

In this section, we present some results on fluctuations of outliers and eigenvalues at soft edges of the limiting support of spiked deformed models, with particular stress on understanding the phenomena through free probability theory. Note that, while global fluctuations of eigenvalues of large random matrices have been described in a series of papers by Mingo et al. [49, 103, 104] in terms of the so-called second order freeness, second order freeness does not seem to be the apropriate tool for the analysis of fluctuations of outliers.

3.4.1 Deterministic full rank deformations of Gaussian models

The seminal works concerning full rank additive deformations of a GUE matrix made strong assumptions on the rate of convergence of μ_{A_N} to ν. In [126], 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 3.2.3) is $N^{-2/3}$. [126] considers the case where μ_{A_N} concentrate quite fast to the measure ν. In particular, there are no spike. More precisely [126] makes a technical assumption on the uniform convergence of the Stieltjes transform of μ_{A_N} to g_ν:

$$\sup_{z \in K} |g_{\mu_{A_N}}(z) - g_\nu(z)| \leq N^{-2/3-\epsilon}, \tag{3.19}$$

where K is some compact subset of the complex plane at a positive distance of the support of ν. Then, [126] proves that the eigenvalues converging to u_0 have universal asymptotic behavior, characterized by the Tracy-Widom distribution. In [2] and [1], [37, 5] 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 [C19], we establish the following results.

Let d be a regular right edge of $\mathrm{supp}(\mu_{sc} \boxplus \nu)$. 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 $\mathrm{supp}(\mu_{sc} \boxplus \mu_{A_N})$ in $]d - \eta, d + \eta[$ and the asymptotic distribution of eigenvalues in the vicinity of d_N is universal as the following Theorem 26 states.

Theorem 26. *([C19]) 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*

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

converges in distribution as $N \to \infty$ to the so-called Tracy-Widom GUE distribution for the k largest eigenvalues (see [142]).

To define the limiting correlation function at an outlier, we consider for $k = 1, 2, \ldots$, the distribution $G_k(\cdot)$ given by

$$G_k(x) = \frac{1}{Z_k} \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} \prod_{1 \le i < j \le k} |\xi_i - \xi_j|^2 \cdot \prod_{i=1}^{k} e^{-\frac{1}{2}\xi_i^2} d\xi_1 \cdots d\xi_k. \tag{3.20}$$

G_k is the distribution of the largest eigenvalue of a $k \times k$ GUE. 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 [C15], we prove that the spectrum of M_N exhibits k_i eigenvalues in a neighborhood of

$$\rho_{\theta_i} = \theta_i + \sigma^2 \int \frac{\mathrm{d}\nu(x)}{\theta_i - x}. \tag{3.21}$$

Dealing with mobile edges related to $\mu_{sc} \boxplus \mu_{A_N}$, [C19] obtains the following universal result.

Theorem 27. *Let θ_i be such that $\int \frac{\mathrm{d}\nu(x)}{(\theta_i - x)^2} < 1/\sigma^2$ 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 $]\rho_{\theta_i} - \epsilon, \rho_{\theta_i} + \epsilon[$. Let us denote by λ_{max} the largest of the k_i outliers around ρ_{θ_i}. There exists $c > 0$ depending on θ_i and ν only such that*

$$\lim_{N \to \infty} \mathbb{P}\left(\sqrt{N} c \left(\lambda_{max} - \frac{L_i(N) + D_i(N)}{2} \right) \le x \right) = G_{k_i}(x).$$

Actually, 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.

In [1, 2], dealing with A_N such that μ_{A_N} is a finite combination of Dirac delta masses, the authors also consider the case where u_0 is a point where two connected components of supp$(\mu_{sc} \boxplus \nu)$ merge so that, denoting by p the density of $\mu_{sc} \boxplus \nu$, $p(u) > 0, \forall u \in (u_0 - \epsilon; u_0 + \epsilon) \setminus \{u_0\}$ for some $\epsilon > 0$ and $p(u_0) = 0$ (u_0 is a so-called cusp point). In this case, the limiting eigenvalue statistics are described by the so-called Pearcey kernel defined by

$$K_P(x, y) := \frac{1}{2i\pi} \int_{\Gamma_0} dt \int_{-i\infty}^{i\infty} ds e^{-t^4 + xt + s^4 - sy} \frac{1}{s - t}. \tag{3.22}$$

The contour Γ_0 is formed by two curves lying respectively to the right and left of 0: one goes from $\infty e^{i\frac{\pi}{4}}$ to $\infty e^{-i\frac{\pi}{4}}$ and the other from $-\infty e^{i\frac{\pi}{4}}$ to $-\infty e^{-i\frac{\pi}{4}}$. In [C19], we also investigates the fluctuations in a neighborhood of an isolated point of vanishing density. Let $u_0 \in \mathbb{R}$ be such that $p(u_0) = 0$ and that there exists $\epsilon > 0$ such that, $\forall u \in]u_0 - \epsilon; u_0 + \epsilon[\setminus\{u_0\}$, $p(u) > 0$. Assume that for any θ_i such that $\int \frac{d\nu(s)}{(\theta_i - s)^2} = 1/\sigma^2$, we have $\theta_i + \sigma^2 g_\nu(\theta_i) \neq u_0$. Set $t_0 = \Psi_{\sigma,\nu}^{-1}(u_0)$ where

$$\Psi_{\sigma,\nu}(t) = t + \sigma^2 \int_{\mathbb{R}} \frac{(t - x) d\nu(x)}{(t - x)^2 + v_{\sigma,\nu}(t)^2}$$

and the function $v_{\sigma,\nu} : \mathbb{R} \to \mathbb{R}^+$ is defined by

$$v_{\sigma,\nu}(u) = \inf \left\{ v \ge 0, \int_{\mathbb{R}} \frac{d\nu(x)}{(u - x)^2 + v^2} \le \frac{1}{\sigma^2} \right\}.$$

The function $\Psi_{\sigma,\nu}$ has been introduced by Biane in [33]: it is an homeomorphism from $\mathbb{R} \setminus \mathcal{O}_1$ (where $\mathcal{O}_1$ is defined in Section 3.2.2) onto the support of $\mu_{sc} \boxplus \nu$. Assuming that the equation $\int \frac{d\mu_{A_N}(x)}{(t-x)^2} - 1/\sigma^2 = 0$

admits a unique solution $t \in \mathbb{C}$ in a neighborhood of t_0, we prove that for η small enough, for all large N, there exists u_N in $]u_0 - \eta; u_0 + \eta[$ such that $p_N(u_N) = 0$ and $\forall u \in]u_0 - \eta; u_0 + \eta[\backslash\{u_N\}$, $p_N(u) > 0$, where p_N denotes the density of $\mu_{sc} \boxplus \mu_{A_N}$.
Last we derive the asymptotic behavior of eigenvalues at the vicinity of u_N.

Theorem 28. *[C19] Let k be a fixed integer and $f : \mathbb{R}^k \to R$ be a symmetric bounded function with compact support. There exists $\kappa > 0$ such that*

$$\mathbb{E} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le N} f\left(\kappa N^{\frac{3}{4}}(\lambda_{i_1} - u_N), \kappa N^{\frac{3}{4}}(\lambda_{i_2} - u_N), \ldots, \kappa N^{\frac{3}{4}}(\lambda_{i_k} - u_N)\right)$$
$$\underset{N \to \infty}{\to} \int_{\mathbb{R}^k} \frac{1}{k!} f(x_1, \ldots, x_k) \det(K_P(x_i, x_j))_{i,j=1}^k \prod_{i=1}^k dx_i.$$

The basic tool of [C19] is a saddle point analysis of the correlation functions of the deformed GUE, the contours involving the image of $\mathbb{R}$ by the continuous extension of the subordination function ω_2 defined by (5.15) with $\mu = \mu_{sc}$ and $\nu = \mu_{A_N}$.

[79, 80, 81] investigate the behavior of the eigenvalues of Gaussian Sample covariance matrices in the case of a full rank perturbation A_N of the identity, in the vicinity of the soft edges, the hard edge when existing, and the cusp points in the support of $\mu_{\mathrm{MP}} \boxtimes \nu$. They obtain the analog of Theorems 26 and 28 dealing with mobile edges namely the edges of the deterministic equivalent $\mu_{\mathrm{MP}} \boxtimes \mu_{A_N}$. As for deformed GUE model, 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.

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 [19, 94, 95] or make use of anisotropic local laws [91]. We do not detail the results since they do not fall under the scope of free probability theory.

3.4.2 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 [85] and Lee and Schnelli [93] that we revisited through free convolutions.

- Johansson [85] considered
$$M_N = W_N^G + A_N$$
where W_N^G is a GUE matrix as defined in Section 1.1 and
$$A_N = N^{-1/6}\mathrm{diag}(\mathrm{y}_1, \ldots, \mathrm{y}_\mathrm{N})$$
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

$$\sigma^{-1} N^{2/3} \left(\lambda_{\max}(M_N) - \tilde{d}_N \right) \xrightarrow{\mathcal{D}} X + Y \tag{3.23}$$

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

$$\tilde{d}_N = \tilde{t}_N + \sigma^2 \int \frac{1}{\tilde{t}_N - x/N^{1/6}} \mathrm{d}\tau(x), \tag{3.24}$$

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

$$\int \frac{1}{(\tilde{t}_N - x/N^{1/6})^2} \mathrm{d}\tau(x) = \frac{1}{\sigma^2}. \tag{3.25}$$

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

$$d_N = t_N + \sigma^2 \frac{1}{N} \sum_{i=1}^{N} \frac{1}{t_N - y_i/N^{1/6}}, \tag{3.26}$$

where t_N in the vicinity of σ satisfies

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{(t_N - y_i/N^{1/6})^2} = \frac{1}{\sigma^2}. \tag{3.27}$$

It is easy to see that (3.24), (3.25), (3.26) and (3.27) 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

$$\sigma^{-1} N^{2/3} \left\{ d_N - \tilde{d}_N \right\} \xrightarrow{\mathcal{D}} N(0, v^2/\sigma^2). \tag{3.28}$$

Now (3.23) readily follows since by Theorem 26, 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 [93] 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

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

with $a < 1, b > 1$ and $h > 0$ is a $\mathcal{C}^1$-function. Assume that W_N is a GUE. Let σ_0 be defined by $\int \frac{1}{(1-x)^2} \mathrm{d}\nu(x) = \frac{1}{\sigma_0^2}$. According to Section 3.2.2, we have $\mathrm{supp}\,(\mu_{sc} \boxplus \nu) = [d_\sigma^-, d_\sigma^+]$. Morevover, by (3.9), for all $\sigma > \sigma_0$, d_σ^+ is a regular edge, the density p of $\mu_{sc} \boxplus \nu$ satisfies $p(x) \sim C(d_\sigma^+ - x)^{\frac{1}{2}}$. Therefore denoting by $d_\sigma^+(N)$ the (random) upper right edge of $\mathrm{supp}\,(\mu_{sc} \boxplus \mu_{A_N})$, Theorem 26 yields the

convergence in distribution of $\alpha^{-1}N^{2/3}(\lambda_1(M_N) - d_\sigma^+(N))$ towards the Tracy-Widom distribution. 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 (3.10), $p(x) \sim C(d_\sigma^+ - x)^b$. 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 $c = c(\nu, \sigma)$).

The universal understanding, provided by free probability theory, of the asymptotic spectrum of models involving several matrices, culminates in the following chapter dealing with non-commutative polynomials in random independent Hermitian matrices; this investigation is achieved by a general methodology based on a linearization procedure and operator-valued subordination properties.

Chapter 4

Polynomials in independent random matrices

We refer to the Appendix where we recall some definitions and properties on freeness with amalgamation introduced by Voiculescu in [148]. Below, we briefly remind the linearization procedure introduced by Anderson in [3, Proposition 3] and the algorithm elaborated by Belinschi, Mai and Speicher in [23] to compute the distribution of a selfadjoint polynomial in free non commutative random variables. We refer to the book [105] by Mingo and Speicher and references therein. Then, we present our contributions in [C23] and [C25] dealing with strong asymptotic freeness, deterministic equivalent measures and outlying eigenvalues of Hermitian polynomial in independent random matrices.

4.1 Linearization

A powerful tool to deal with non commutative polynomials in random matrices or in operators is the so-called "linearization trick" that goes back to Haagerup and Thorbjørnsen [77, 78] in the context of operator algebras and random matrices (see [105]). We use the procedure introduced in [3, Proposition 3], which has several advantages, to be described below.

Given a polynomial $P \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$, we call *linearization* of P any $L_P \in M_m(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \ldots, X_k\rangle$ such that

$$L_P := \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix} \in M_m(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \ldots, X_k\rangle$$

where

1. $m \in \mathbb{N}$,
2. $Q \in M_{m-1}(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \ldots, X_k\rangle$ is invertible,
3. u is a row vector and v is a column vector, both of size $m-1$ with entries in $\mathbb{C}\langle X_1, \ldots, X_k\rangle$,
4. the polynomial entries in Q, u and v all have degree ≤ 1,
5. $$P = -uQ^{-1}v,$$

It is shown in [3] that, given a polynomial $P \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$, there exist $m \in \mathbb{N}$ and a linearization $L_P \in M_m(\mathbb{C}) \otimes \mathbb{C}\langle X_1, \ldots, X_k\rangle$. The algebra of polynomials in non-commuting indeterminates $X_1, \ldots, X_k$ becomes a $*$-algebra by anti-linear extension of $(X_{i_1} X_{i_2} \cdots X_{i_m})^* = X_{i_m} \cdots X_{i_2} X_{i_1}$. It turns out that if P is self-adjoint, L_P can be chosen to be self-adjoint.

Example: $P(X_1, X_2) = X_1X_2 + X_2X_1 + X_1^2$.

Note that $L_P = \begin{pmatrix} 0 & X_1 & X_2 + \frac{1}{2}X_1 \\ X_1 & 0 & -1 \\ X_2 + \frac{1}{2}X_1 & -1 & 0 \end{pmatrix}$ is a selfadjoint linearization of P.

$$L_P = \gamma \otimes 1 + \alpha_1 \otimes X_1 + \alpha_2 \otimes X_2,$$

where

$$\alpha_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \ \alpha_1 = \begin{pmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}, \ \gamma = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

The following results are fundamental for our purpose.

Lemma 3. *Let $P = P^* \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$ and let $L_P \in M_m(\mathbb{C}\langle X_1, \ldots, X_k\rangle)$ be a linearization of P with the properties outlined above. Let e_{11} be the $m \times m$ matrix whose single nonull entry equals one and occurs in the row 1 and column 1. Let $y = (y_1, \ldots, y_k)$ be a k-tuple of self-adjoint operators in a $\mathcal{C}^*$-algebra $\mathcal{A}$. Then, for any $z \in \mathbb{C}$, $ze_{11} \otimes 1_{\mathcal{A}} - L_P(y)$ is invertible if and only if $z1_{\mathcal{A}} - P(y)$ is invertible and we have*

$$(ze_{11} \otimes 1_{\mathcal{A}} - L_P(y))^{-1} = \begin{pmatrix} (z1_{\mathcal{A}} - P(y))^{-1} & \star \\ \star & \star \end{pmatrix}.$$

Lemma 4. *[C23] Let $P = P^* \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$ and let $L_P \in M_m(\mathbb{C}\langle X_1, \ldots, X_k\rangle)$ be a linearization of P with the properties outlined above. Let $y_n = (y_1^{(n)}, \ldots, y_k^{(n)})$ be a k-tuple of self-adjoint operators in a $\mathcal{C}^*$-algebra $\mathcal{A}$ such that $\sup_n \max_{i=1}^k \|y_n^{(i)}\| = C < +\infty$. Let $z_0 \in \mathbb{C}$ be such that, for all large n, the distance from z_0 to the spectrum of $P(y_n)$ is greater than δ. Then, there exists a constant $\epsilon > 0$, depending only on δ, L_P and C such that the distance from 0 to $sp(z_0e_{11} \otimes 1_{\mathcal{A}} - L_P(y_n))$ is at least ϵ.*

4.2 An algorithm to compute the distribution of a selfadjoint polynomial in free non commutative random variables

Belinschi, Mai and Speicher [23] (see also [6]) elaborated a general algorithm to deal with the distribution of selfadjoint polynomials in free variables. The main idea will be to relate such a polynomial with an operator-valued linear polynomial, and then use operator-valued convolution to deal with the latter. The distribution of a self-adjoint polynomial $P(x_1, \ldots, x_n)$ in free variables $x_1, \ldots, x_n$ in a $\mathcal{C}^*$-probability space $(\mathcal{A}, \phi)$ can be deduced from the operator-valued distribution with respect to $E = id_m \otimes \phi$ of a corresponding linearization $L_P := b_0 \otimes 1 + b_1 \otimes x_1 + \ldots + b_n \otimes x_n \in M_m(\mathbb{C}) \otimes \mathcal{A}$. Indeed, according to Lemma 3,

$$\phi((z1_{\mathcal{A}} - P(x_1, \ldots, x_n))^{-1}) = \lim_{\epsilon \downarrow 0} [id_m \otimes \phi((\Lambda_\epsilon(z) - L_P(x_1, \ldots, x_n))^{-1})]_{11}$$

for all $z \in \mathbb{C}^+$ where

$$\Lambda_\epsilon(z) = \begin{pmatrix} z & (0) & (0) \\ 0 & i\epsilon & (0) \\ & & \ddots (0) \\ (0) & & \\ (0) & (0) & i\epsilon \end{pmatrix} \tag{4.1}$$

and one can then obtain the desired distribution of $P(x_1, \ldots, x_n)$ by applying the Stieltjes inversion formula. The freeness of $x_1, \ldots, x_n$ implies the freeness over $M_m(\mathbb{C})$ of $b_1 \otimes x_1, \ldots, b_n \otimes x_n$. We have finally reduced the determination of the distribution of $P(x_1, \ldots, x_n)$ to a problem involving operator-valued additive free convolution. However, in spite of the existence of operator-valued Cauchy and R-transforms [148, 150] there are rarely any non-trivial operator-valued examples where an explicit solution can be written down. It is the subordination formulation of those convolutions which comes to the rescue as follows.

For any $\xi \in M_m(\mathbb{C}) \otimes \mathcal{A}$ such that $\xi = \xi^*$, define the $M_m(\mathbb{C})$-valued Cauchy transform

$$G_\xi(\kappa) = id_m \otimes \phi((\kappa \otimes 1_{\mathcal{A}} - \xi)^{-1}),$$

for any $\kappa \in \mathbb{H}^+(M_m(\mathbb{C})) := \{b \in M_m(\mathbb{C}), \Im b > 0\}$. Note that, for $j = 1, \ldots, n$, the $M_m(\mathbb{C})$-valued Cauchy transform $G_{b_j \otimes x_j}$ is completely determined by the scalar-valued Cauchy transform $g_{\mu_{x_j}} = \int \frac{1}{z-t} d\mu_{x_j}$ via

$$G_{b_j \otimes x_j}(\kappa) = \lim_{\epsilon \downarrow 0} -\frac{1}{\pi} \int_{\mathbb{R}} (\kappa - t b_j)^{-1} \Im m(g_{\mu_{x_j}}(t + i\epsilon)) dt.$$

Now, by Theorem 45,

$$G_{b_1 \otimes x_1 + b_2 \otimes x_2}(\kappa) = G_{b_1 \otimes x_1}(\omega_m(\kappa))$$

where $\omega_m(\kappa)$ is the unique fixed point of

$$f_\kappa : \mathbb{H}^+(M_m(\mathbb{C})) \to \mathbb{H}^+(M_m(\mathbb{C})), \quad f_\kappa(w) = h_{y_2}(h_{y_1}(w) + \kappa) + \kappa \tag{4.2}$$

$$\text{where } \; y_i = b_i \otimes x_i, \; h_{y_i}(\kappa) = E\left[(\kappa - y_i)^{-1}\right]^{-1} - b$$

$$\text{and } \; \omega_m(\kappa) = \lim_{k \to +\infty} f_\kappa^{\circ k}(w), \quad \text{for any } \; w \in \mathbb{H}^+(M_m(\mathbb{C})).$$

Then one can compute $\omega_m(\kappa) = \lim_{k \to +\infty} f_\kappa^{\circ k}(w)$, for some $w \in \mathbb{H}^+(M_m(\mathbb{C}))$ and deduce $G_{b_1 \otimes x_1 + b_2 \otimes x_2}(\kappa)$. One can get $G_{b_1 \otimes x_1 + \cdots b_n \otimes x_n}(\kappa)$ by iterating this procedure. The Cauchy transform of L_P is then given by

$$G_{L_P}(\kappa) = G_{L_P - b_0 \otimes 1}(\kappa - b_0).$$

Thus, the algorithm of Belinschi, Mai and Speicher [23] to compute the distribution of a self-adjoint polynomial $P(x_1, \ldots, x_n)$ in free variables $x_1, \ldots, x_n$ in a $\mathcal{C}^*$-probability space $(\mathcal{A}, \phi)$ can be summarized as follows.

- STEP 1: Find a self-adjoint linearization $L_P := b_0 \otimes 1 + b_1 \otimes x_1 + \ldots + b_n \otimes x_n \in M_m(\mathcal{A})$.

Fix $\kappa \in \mathbb{H}^+(M_N(\mathbb{C}))$.

- STEP 2: Compute $G_{b_j \otimes x_j}(\kappa) = \lim_{\epsilon \downarrow 0} -\frac{1}{\pi} \int_{\mathbb{R}} (\kappa - t b_j)^{-1} \Im m(g_{\mu_{x_j}}(t + i\epsilon)) dt$.
- STEP 3: Compute $\omega_m(\kappa) = \lim_{k \to +\infty} f_\kappa^{\circ k}(w)$, for some $w \in \mathbb{H}^+(M_m(\mathbb{C}))$, where f_κ is defined by (4.2). Then compute $G_{b_1 \otimes x_1 + b_2 \otimes x_2}(\kappa) = G_{b_1 \otimes x_1}(\omega_m(\kappa))$. Iterate and get $G_{b_1 \otimes x_1 + \cdots b_n \otimes x_n}(\kappa)$.
- STEP 4: Compute $G_{L_P}(\kappa) = G_{L_P - b_0 \otimes 1}(\kappa - b_0)$.
- STEP 5: For all $z \in \mathbb{C}^+$, compute $g_P(z) = \lim_{\epsilon \downarrow 0} [G_{L_P}(\Lambda_\epsilon(z))]_{11}$ where $\Lambda_\epsilon(z)$ is defined by (4.1).
- STEP 6: Apply the Stieltjes inversion formula.

4.3 Deterministic equivalent measures

In the engineering literature there exists a version of the notion of a deterministic equivalent measure (going back to Girko [73], see also [82]). This deterministic equivalent measure is obtained by replacing the Cauchy transform g_N of the considered $N \times N$ random matrix model (for which no analytic solution exists) by a function $\tilde{g}_N$ which is defined as the solution of a specified system of equations. In [134], Speicher and Vargas give a more conceptual approach and show that the only meaningful way to get a closed system of equations when dealing with random matrices is to replace the random matrices by free variables. In Chapter 3, we introduced such deterministic equivalent measures for three classical deformed models. Actually, the results, previously stated in Proposition 3 in Chapter 3, from [9] for sample covariance matrices and [12] for Information-plus-noise type models, and [C15] for additive deformations of a Wigner matrix, can be extended to selfadjoint polynomial in Wigner matrices together with deterministic matrices. Indeed, in [C23], we prove that a sequence of deterministic measures plays a central role in the study of the spectrum of such general models. These measures are computed with the tools of free probability and are in some sense obtained by taking partially the limit when the dimension goes to infinity, only for the Wigner matrices, as conceptualized by [134]. We establish that almost surely, for large dimension, each interval lying at some distance from the supports of these deterministic measures contains no eigenvalue of the Hermitian polynomial model.

Here are the matricial models we deal with. Let t and r be fixed nonzero integer numbers independent from N.

• $(A_N^{(1)}, \ldots, A_N^{(t)})$ is a $t-$tuple of $N \times N$ deterministic matrices such that for any $u = 1, \ldots, t$,

$$\sup_N \|A_N^{(u)}\| < \infty, \tag{4.3}$$

where $\|\cdot\|$ denotes the spectral norm.

• We consider r independent $N \times N$ random Hermitian matrices $X_N^{(v)} = [X_{ij}^{(v)}]_{i,j=1}^N$, $v = 1, \ldots, r$, where, for each v, $[X_{ij}^{(v)}]_{i\geq 1, j\geq 1}$ is an infinite array of random variables such that $X_{ii}^{(v)}$, $\sqrt{2}Re(X_{ij}^{(v)}), i < j$, $\sqrt{2}Im(X_{ij}^{(v)}), i < j$, are independent, centered with variance 1 and satisfy:

1. There exists a K_v and a random variable $Z^{(v)}$ with finite fourth moment for which there exists $x_0 > 0$ and an integer number $N_0 > 0$ such that, for any $x > x_0$ and any integer number $N > N_0$, we have
$$\frac{1}{N^2} \sum_{1\leq i,j\leq N} \mathbb{P}\left(|X_{ij}^{(v)}| > x\right) \leq K_v \mathbb{P}\left(|Z^{(v)}| > x\right). \tag{4.4}$$

2. $$\sup_{(i<j)\in\mathbb{N}^2} \mathbb{E}(|X_{ij}^{(v)}|^3) < +\infty.$$

Remark 5. *Note that assumption such as* (4.4) *appears in [51]. It obviously holds if the random variables* $X_{ii}^{(v)}$, $\sqrt{2}Re(X_{ij}^{(v)}), i < j$, $\sqrt{2}Im(X_{ij}^{(v)}), i < j$, *are identically distributed with finite fourth moment.*

Here is the result roughly stating that there is no eigenvalue of some polynomial model in such matrices outside the support of a deterministic measure.

Theorem 29. *([C23]) Let $(\mathcal{A},\tau)$ be a C^*-probability space equipped with a faithful tracial state and $x=(x_1,\ldots,x_r)$ be a semi-circular system in $(\mathcal{A},\tau)$. Let $a_N=(a_N^{(1)},\ldots,a_N^{(t)})$ be a t-tuple of noncommutative random variables which is free from x in $(\mathcal{A},\tau)$ and such that the distribution of (a_N,a_N^*) in $(\mathcal{A},\tau)$ coincides with the distribution of $(A_N^{(1)},\ldots,A_N^{(t)},(A_N^{(1)})^*,\ldots,(A_N^{(t)})^*)$ in $(M_N(\mathbb{C}),\frac{1}{N}Tr)$. Let P be a self-adjoint polynomial in $2t+r$ noncommutative indeterminates $X_1,\ldots,X_r,X_{r+1},\ldots,X_{r+t},X_{r+1}^*,\ldots,X_{r+t}^*$, where for any $i=1,\ldots,r$, $X_i=X_i^*$. Let $[b,c]$ be a real interval such that there exists $\delta>0$ such that, for any large N, $[b-\delta,c+\delta]$ lies outside the support of the distribution of the noncommutative random variable $P\left(x_1,\ldots,x_r,a_N^{(1)},\ldots,a_N^{(t)},(a_N^{(1)})^*,\ldots,(a_N^{(t)})^*\right)$ in $(\mathcal{A},\tau)$. Then, almost surely, for all large N, there is no eigenvalue of the $N\times N$ Hermitian matrix $P\left(\frac{X_N^{(1)}}{\sqrt{N}},\ldots,\frac{X_N^{(r)}}{\sqrt{N}},A_N^{(1)},\ldots,A_N^{(t)},(A_N^{(1)})^*,\ldots,(A_N^{(t)})^*\right)$ in $[b,c]$.*

Remark 6. *When $r=t=1$, $A_N^{(1)}=(A_N^{(1)})^*$ and $P(X_1,X_2,X_2^*)=X_1+\frac{X_2+X_2^*}{2}$, the distribution of $P(x_1,a_N^{(1)},(a_N^{(1)})^*)$ is the free convolution $\mu_{sc}\boxplus\mu_{A_N^{(1)}}$ where $\mu_{A_N^{(1)}}=\frac{1}{N}\sum_{i=1}^N\lambda_i(A_N^{(1)})$, denoting by $\lambda_i(A_N^{(1)})$, $i=1,\ldots,N$, the eigenvalues of $A_N^{(1)}$.*

Note that it is sufficient to prove Theorem 29 for Hermitian matrices $A_N^{(1)},\ldots,A_N^{(t)}$ by considering their Hermitian and anti-Hermitian parts. Moreover, using a truncation and Gaussian convolution procedure, it is sufficient to prove Theorem 29 when we assume that the $X_{ij}^{(v)}$'s satisfy:

(H) the variables $\sqrt{2}\Re eX_{ij}^{(v)}$, $\sqrt{2}\Im mX_{ij}^{(v)}$, $(i<j)\in\mathbb{N}^2$, $X_{ii}^{(v)}$, $i\in\mathbb{N}$, $v=1,\ldots,r$, are independent, centered with variance 1 and satisfy a Poincaré inequality with common constant C_{PI}.

Note that this implies that for any $p\in\mathbb{N}$,

$$\max_{v=1,\ldots,r}\sup_{(i,j)\in\mathbb{N}^2}\mathbb{E}\left(|X_{ij}^{(v)}|^p\right)<+\infty. \tag{4.5}$$

We adopt the strategy from [77] and [78] based on a linearization trick and sharp estimates on matricial Stieltjes transforms. Hence, the proof of Theorem 29 is based on the following key lemma.

Lemma 5. *([C23]) Let $(\mathcal{A},\tau)$ be a C^*-probability space equipped with a faithful tracial state and $x=(x_1,\ldots,x_r)$ be a semi-circular system in $(\mathcal{A},\tau)$. Let $a_N=(a_N^{(1)},\ldots,a_N^{(t)})$ be a t-tuple of noncommutative self-adjoint random variables which is free from x in $(\mathcal{A},\tau)$, such that the distribution of a_N coincides with the distribution of $(A_N^{(1)},\ldots,A_N^{(t)})$ in $(M_N(\mathbb{C}),\frac{1}{N}Tr_N)$. Then, for all $m\in\mathbb{N}$, all self-adjoint matrices $\gamma,\alpha_1,\ldots,\alpha_r,\beta_1,\ldots,\beta_t$ of size $m\times m$ and all $\epsilon>0$, almost surely, for all large N, we have*

$$sp(\gamma\otimes I_N+\sum_{v=1}^r\alpha_v\otimes\frac{X_N^{(v)}}{\sqrt{N}}+\sum_{u=1}^t\beta_u\otimes A_N^{(u)})$$

$$\subset sp(\gamma\otimes 1_{\mathcal{A}}+\sum_{v=1}^r\alpha_v\otimes x_v+\sum_{u=1}^t\beta_u\otimes a_N^{(u)})+]-\epsilon,\epsilon[. \tag{4.6}$$

Here, $sp(T)$ denotes the spectrum of the operator T, I_N the identity matrix and $1_{\mathcal{A}}$ denotes the unit of $\mathcal{A}$.

Indeed, given a noncommutative polynomial P, choosing in Lemma 5 the γ, $(\alpha_v)_{v=1,\ldots,r}$, $(\beta_u)_{u=1,\ldots,t}$ corresponding to a self-adjoint linearization of P as defined in [3] and using Lemma 4 allow to deduce

Theorem 29.
The proof of (4.6) requires sharp estimates of $g_N(z) - \tilde{g}_N(z)$ where for $z \in \mathbb{C} \setminus \mathbb{R}$,

$$g_N(z) = \mathbb{E}\frac{1}{m}Tr_m \otimes \frac{1}{N}Tr_N[(zI_m \otimes I_N - \gamma \otimes I_N - \sum_{v=1}^{r} \alpha_v \otimes \frac{X_N^{(v)}}{\sqrt{N}}(\omega) - \sum_{u=1}^{t} \beta_u \otimes A_N^{(u)})^{-1}]$$

and

$$\tilde{g}_N(z) = \frac{1}{m}Tr_m \otimes \tau[(zI_m \otimes 1_{\mathcal{A}} - \gamma \otimes 1_{\mathcal{A}} - \sum_{v=1}^{r} \alpha_v \otimes x_v - \sum_{u=1}^{t} \beta_u \otimes a_N^{(u)})^{-1}].$$

We establish that there exists a polynomial Q with nonnegative coefficients such that, for $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\left| g_N(z) - \tilde{g}_N(z) + \tilde{E}_N(z) \right| \leq \frac{Q(|\Im mz|^{-1})}{N\sqrt{N}}, \tag{4.7}$$

where $\tilde{E}_N$ is the Stieltjes transform of a compactly supported Schwartz distribution ∇_N on $\mathbb{R}$ whose support is included in the spectrum of $\gamma \otimes 1_{\mathcal{A}} + \sum_{v=1}^{r} \alpha_v \otimes x_v + \sum_{u=1}^{t} \beta_u \otimes a_N^{(u)}$ and such that $\nabla_N(1) = 0$. But this required sharp estimate makes necessary a fit use of free operator-valued subordination maps. In particular, we need an explicit development of the Stieltjes transform up to the order $\frac{1}{N\sqrt{N}}$ but the stability under perturbation argument used in [100] does not provide this development from the approximate matricial subordination equation (obtained by the approximate integration by part formula in Lemma 1). Therefore we use a strategy based on an invertibility property of matricial subordination maps related to $M_m(\mathbb{C})$-valued semicircular variable.

4.4 Strong asymptotic freeness

Let k be an integer greater than zero. Denote by $\mathcal{P}$ the set of polynomials in $2k$ noncommutative indeterminates. A sequence of families of variables $(a_N)_{N\geq 1} = (a_1(N), \ldots, a_k(N))_{n\geq 1}$ in C^*-probability spaces $(\mathcal{A}_N, \tau_n)$ converge, when N goes to infinity, respectively in distribution if the map $P \in \mathcal{P} \mapsto \tau_n(P(a_N, a_N^*))$ converges pointwise and strongly in distribution if moreover the map $P \in \mathcal{P} \mapsto \|P(a_N, a_N^*)\|$ converges pointwise.

Assume that the states τ and $(\tau_N)_{N\in\mathbb{N}}$ are faithful. Then, (see Proposition 2.1 [48]) a sequence of k-tuples a_N of selfadjoint noncommutative random variables converges strongly in distribution to a k-tuple a of selfadjoint noncommutative random variables if and only if for any self-adjoint polynomial $P \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$, $\mu_{P(a_N)}$ converges in weak-* topology to $\mu_{P(a)}$ and the support of $\mu_{P(a_N)}$ converges in the Hausdorff topology to the support of $\mu_{P(a)}$, that is: for any $\epsilon > 0$, there exists N_0 such that for any $N \geq N_0$,

$$\mathrm{supp}(\mu_{\mathrm{h_N}}) \subset \mathrm{supp}(\mu_{\mathrm{h}}) + (-\epsilon, +\epsilon).$$

(For any noncommutative random variable y, μ_y denotes its distribution and the symbol supp means the support of the measure.) In particular, the strong convergence in distribution of a single self-adjoint variable is its convergence in distribution together with the Hausdorff convergence of its spectrum.

In [77], Haagerup and Thorbjørsen proved the strong asymptotic freeness of independent GUE matrices $X_N^{(i)}$, $i = 1, \ldots, r$, namely: almost surely, for any polynomial P in r noncommutative indeterminates,

$$\left\| P\left(\frac{X_N^{(1)}}{\sqrt{N}}, \ldots, \frac{X_N^{(r)}}{\sqrt{N}}\right) \right\| \to_{N\to+\infty} \|P(x_1, \ldots, x_r)\|, \tag{4.8}$$

where $(x_1, \ldots, x_r)$ is a semi-circular system. Note that this result led to the proof of the very important result in the theory of operator algebras that $Ext(\mathcal{C}^*_{red}(F_2))$ is not a group. (4.8) was proved for Gaussian random matrices with real or symplectic entries by Schultz [78], for Wigner matrices with symmetric distribution of the entries satisfying a Poincaré inequality by Capitaine and Donati-Martin [C10] and for Wigner matrices under i.i.d assumptions and fourth moment hypotheses by Anderson [3].

Voiculescu [146, 149] proved the asymptotic freeness of independent GUE matrices with an extra family of deterministic matrices with limiting distribution. Male [100] established the strong asymptotic freeness of independent GUE matrices with an extra independent family of matrices with strong limiting distribution. Note that in [48], Collins and Male established the strong asymptotic freeness of Haar and deterministic matrices with strong limiting distribution.
Sticking to the proofs in [77] and [78], we can deduced from Lemma 5 the following strong asymptotic freeness of independent non-Gaussian Wigner matrices and an extra family of deterministic matrices with strong limiting distribution.

Theorem 30. *([C23]) Let $(\mathcal{A}, \tau)$ be a C^*-probability space equipped with a faithful tracial state. Let $x = (x_1, \ldots, x_r)$ be a semi-circular system and $a = (a_1, \ldots, a_t)$ be a t-tuple of noncommutative random variables which is free from x in $(\mathcal{A}, \tau)$. Let $(A_N^{(1)}, \ldots, A_N^{(t)}, (A_N^{(1)})^*, \ldots, (A_N^{(t)})^*)$ and $X_N^{(v)} = [X_{ij}^{(v)}]_{i,j=1}^N$, $v = 1, \ldots, r$, be matrices as defined in Section 4.3. Assume moreover that $(A_N^{(1)}, \ldots, A_N^{(t)}, (A_N^{(1)})^*, \ldots, (A_N^{(t)})^*)$ converges strongly towards $a = (a_1, \ldots, a_t, a_1^*, \ldots, a_t^*)$ in $(\mathcal{A}, \tau)$, that is for any polynomial P in 2t noncommutative indeterminates,*

$$\frac{1}{N} TrP\left(A_N^{(1)}, \ldots, A_N^{(t)}, (A_N^{(1)})^*, \ldots, (A_N^{(t)})^*\right) \to_{N\to+\infty} \tau\left(P(a_1, \ldots, a_t, a_1^*, \ldots, a_t^*\right)$$

and

$$\left\| P\left(A_N^{(1)}, \ldots, A_N^{(t)}, (A_N^{(1)})^*, \ldots, (A_N^{(t)})^*\right)\right\| \to_{N\to+\infty} \|P(a_1, \ldots, a_t, a_1^*, \ldots, a_t^*)\|_{\mathcal{A}}\,.$$

Then, almost surely, for any polynomial P in $r + 2t$ noncommutative variables,

$$\lim_{N\to+\infty} \frac{1}{N} TrP\left(\frac{X_N^{(1)}}{\sqrt{N}}, \ldots, \frac{X_N^{(r)}}{\sqrt{N}}, A_N^{(1)}, \ldots, A_N^{(t)}, (A_N^{(1)})^*, \ldots, (A_N^{(t)})^*\right)$$
$$= \tau\left(P\left(x_1, \ldots, x_r, a_1, \ldots, a_t, a_1^*, \ldots, a_t^*\right)\right) \tag{4.9}$$

and

$$\lim_{N\to+\infty} \left\| P\left(\frac{X_N^{(1)}}{\sqrt{N}}, \ldots, \frac{X_N^{(r)}}{\sqrt{N}}, A_N^{(1)}, \ldots, A_N^{(t)}, (A_N^{(1)})^*, \ldots, (A_N^{(t)})^*\right)\right\|$$
$$= \|P\left(x_1, \ldots, x_r, a_1, \ldots, a_t, a_1^*, \ldots, a_t^*\right)\|_{\mathcal{A}}\,. \tag{4.10}$$

4.5 Outlying eigenvalues

Let μ be a compactly supported probability measure on $\mathbb{R}$, p be a positive integer, and $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_p$ in $\mathbb{R} \setminus \mathrm{supp}(\mu)$ be a sequence of fixed real numbers. We consider, for all $N \in \mathbb{N}, N \geq 1$, a random selfadjoint matrix $A_N \in M_N(\mathbb{C})$ which satisfies the following conditions:

(A1) the sequence of empirical spectral measures $\{\mu_{A_N}\}_{N=1}^{\infty}$ converges weakly to μ,

(A2) $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_p$ are p eigenvalues of A_N and

(A3) the other eigenvalues of A_N which may be random , say $\gamma_j(A_N), j = 1\ldots, N-p$, satisfy almost surely that for every $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that

$$\gamma_j(A_N) \in \mathrm{supp}(\mu) + (-\varepsilon, \varepsilon), \quad j = 1, \ldots, N-p, N \geq N(\varepsilon).$$

In other words, only the p eigenvalues $\theta_1, \ldots, \theta_p$ prevent $\{A_N\}_{N=1}^{\infty}$ from converging strongly in distribution to μ.

In [C25], we investigate two polynomial matricial models, both involving A_N.

- Our first model involves a sequence $\{B_N\}_{N=1}^{\infty}$ of random Hermitian matrices such that

  (B1) almost surely, B_N converges strongly in distribution to a compactly supported probability measure ν on $\mathbb{R}$,

  (B2) for each N, the distribution of B_N is invariant under conjugation by any $N \times N$ unitary matrix.

  We consider the matricial model

  $$M_N = P(B_N, A_N) \tag{4.11}$$

  for any selfadjoint polynomial P in $\mathbb{C}\langle X_1, X_2\rangle$.

- Our second model deals with a $N \times N$ random Hermitian Wigner matrix $X_N = [X_{ij}]_{i,j=1}^N$, where $[X_{ij}]_{i\geq 1, j\geq 1}$ is an infinite array of random variables satifying

  (X1) X_{ii}, $\sqrt{2}Re(X_{ij}), i < j$, $\sqrt{2}Im(X_{ij}), i < j$, are independent, centered with variance 1,

  (X2) there exists a K and a random variable Z with finite fourth moment for which there exists $x_0 > 0$ and an integer number $n_0 > 0$ such that, for any $x > x_0$ and any integer number $n > n_0$, we have

  $$\frac{1}{n^2} \sum_{1 \leq i,j \leq n} \mathbb{P}\left(|X_{ij}| > x\right) \leq K\mathbb{P}\left(|Z| > x\right).$$

  (X3)

  $$\sup_{(i<j)\in\mathbb{N}^2} \mathbb{E}(|X_{ij}|^3) < +\infty.$$

  We consider the matricial model

  $$M_N = P(\frac{X_N}{\sqrt{N}}, A_N) \tag{4.12}$$

  for any selfadjoint polynomial P in $\mathbb{C}\langle X_1, X_2\rangle$.

According to results of Voiculescu [146] (see also [151]), there exist selfadjoint elements a, b in a II_1-factor $(\mathcal{A}, \tau)$ such that, almost surely, the sequence $\{(A_N, B_N)\}_{N=1}^{\infty}$ converges in distribution to (a, b). More specifically, a and b are freely independent and $\mu = \mu_a, \nu = \mu_b$. In particular, if P is a selfadjoint polynomial in $\mathbb{C}\langle X_1, X_2\rangle$,

$$\lim_{N\to\infty} \mu_{P(B_N, A_N)} = \mu_{P(b,a)}$$

almost surely in the weak* topology. When $p = 0$, [48] shows that, almost surely, the sequence $\{P(B_N, A_N)\}_{N=1}^{\infty}$ converges strongly in distribution to $P(b, a)$.
According to (2.10) in [C23] and [4, Theorem 5.4.5], if P is a selfadjoint polynomial in $\mathbb{C}\langle X_1, X_2\rangle$,

$$\lim_{N\to\infty} \mu_{P(\frac{X_N}{\sqrt{N}}, A_N)} = \mu_{P(x,a)}$$

almost surely in the weak* topology, where a and x are freely independent selfadjoint non commutative random variables, $\mu = \mu_a$ and x is the standard semicircular variable (i.e $d\mu_x = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x)$). As in the first model, when $p = 0$, Theorem 30 and Proposition 2.1 [48] show that, almost surely, the sequence $\{P(\frac{X_N}{\sqrt{N}}, A_N)\}_{N=1}^{\infty}$ converges strongly in distribution to $P(x, a)$.

Let Y_N be either B_N or $\frac{X_N}{\sqrt{N}}$ and y be either b or x. As precised below, when $p = 0$, when N is large, the spectrum of $P(Y_N, A_N)$ has no outlier, that is, almost surely for N large enough, the spectrum is included in a small neighborhood of the support of $\mu_{P(y,a)}$.
Our main result applies to $p > 0$. The set of outliers is calculated from the spikes $\theta_1, \ldots, \theta_p$ using linearization and Voiculescu's matrix subordination function [150] as follows.
Let $L_P \in M_m(\mathbb{C}\langle X_1, X_2\rangle)$, $L_P = \gamma \otimes 1 + \alpha \otimes X_1 + \beta \otimes X_2$, be a selfadjoint linearization of P as introduced in Section 4.1. By Theorem 45, there exists an analytic function $\omega\colon \mathbb{H}^+(M_m(\mathbb{C})) \to \mathbb{H}^+(M_m(\mathbb{C}))$ such that

$$\begin{aligned}(\mathrm{id}_m \otimes \tau)\left[(\omega(w) \otimes 1 - \beta \otimes a)^{-1}\right] \\ = \quad (\mathrm{id}_m \otimes \tau)\left[(w \otimes 1 - (\beta \otimes a + \alpha \otimes b))^{-1}\right]\end{aligned}$$

for every $w \in \mathbb{H}^+(M_m(\mathbb{C}))$.

Theorem 31. *[C25] Define the function $\omega_o(z) = (\omega(ze_{11} - \gamma) + iI_m)^{-1}$. ω_o is analytic on $\mathbb{C}^+$ and has an analytic extension to a neighbourhood of $\mathbb{R} \setminus sp(P(b, a))$. Define*

$$H_j(z) = \det[(\theta_j\beta + i)\omega_o(z) - I_m]$$

and denote by $m_{\theta_j}(t)$ the order of t as a zero of $H_j(z)$ at $z = t$, $1 \le j \le p$.
Suppose that $t \in \mathbb{R} \setminus sp(P(b, a))$ and set $m(t) = \sum_{j=1}^{p} m_{\theta_j}(t)$. Then there exists $\delta_0 > 0$ such that for any $\delta \in (0, \delta_0)$, almost surely for large N, the random matrix $P(B_N, A_N)$ has $m(t)$ eigenvalues in the interval $(t - \delta, t + \delta)$, counting multiplicity.

Remark 7. *Note that if we know in addition that $z \mapsto \omega(ze_{1,1} - \gamma)$ is analytic at $z = t$, then the function $H_j(z)$ can be replaced with $z \mapsto \det[\theta_j\beta - \omega(te_{11} - \gamma)]$. Then $m(t)$ is equal to the multiplicity of t as a zero of*

$$z \mapsto \prod_{j=1}^{p} \det[\theta_j\beta - \omega(ze_{11} - \gamma)].$$

This is always the case when b is a semicircular variable and then, when the unitarily invariant matrix B_N is a GUE matrix. When B_N is replaced by a Wigner matrix X_N satisfying the hypotheses (X1)–(X3), we obtain the following

Theorem 32. *([C25]) Let a, x be free selfadjoint elements in a C^*-probability space $(\mathcal{A}, \tau)$ with distribution μ and the standard semicircular distribution μ_{sc} respectively. Let $\omega\colon \mathbb{H}^+(M_m(\mathbb{C})) \to \mathbb{H}^+(M_m(\mathbb{C}))$ be the $M_m(\mathbb{C})$-valued map such that for all $\kappa \in \mathbb{H}^+(M_m(\mathbb{C}))$,*

$$\omega(\kappa) = \kappa - \alpha(\mathrm{Id}_{M_m(\mathbb{C})} \otimes \tau)\left[(\kappa \otimes 1_{\mathcal{A}} - \alpha \otimes x - \beta \otimes a)^{-1}\right]\alpha.$$

Let $t \in \mathbb{R} \setminus sp(P(x, a))$. Denote by $m_{\theta_i}(t)$ the order of t as a zero of $\det(\omega(ze_{11} - \gamma) - \theta_i\beta)$ at $z = t$, $1 \le i \le p$. Then, for ϵ small enough, almost surely for large N, there are exactly $\sum_{i=1}^{p} m_{\theta_i}(t)$ eigenvalues of $P(\frac{X_N}{\sqrt{N}}, A_N)$ in an ϵ-neighborhood of t.

After a linearization procedure of P of the form $L \in M_m(\mathbb{C}\langle X_1, X_2\rangle)$ described in Section 4.1, the firts step of both proofs of Theorem 31 and Theorem 32 consists in reducing the problem, in the spirit of [31], to the convergence of an $M_{mp}(\mathbb{C})$-valued map F_N, involving a random matrix-valued generalized resolvent. Establishing these convergences is the main part of the proofs. Such a convergence is proved directly for the first model $P(U_N D_N U_N^*, A_N)$ by extending the arguments of [C21] and making use of the fundamental operator-valued subordination properties described in Section 5.7. For the second model $P(\frac{X_N}{\sqrt{N}}, A_N)$, the convergence of F_N is then obtained by a comparison argument to the Gaussian case.

Remark 8. *Suppose that $\mu = \delta_0$. Then the function ω is computed as follows:*

$$\omega(w) = \left\{\mathrm{id}_m \otimes \tau[(w \otimes 1 - \alpha \otimes b)^{-1}]\right\}^{-1}, \quad w \in M_n(\mathbb{C}).$$

As an illustration, consider the random matrix

$$M = A_N \frac{X_N}{\sqrt{N}} + \frac{X_N}{\sqrt{N}} A_N + \frac{X_N^2}{N},$$

where X_N is a standard standard GUE matrix of size N and

$$A_N = \mathrm{Diag}(\theta, 0, \ldots, 0), \quad \theta \in \mathbb{R} \setminus \{0\}.$$

In this case, A_N has rank one, and thus $\mu = \delta_0$. It follows that the limit spectral measure Π of M is the same as the limit spectral measure of X_N^2/N. Thus, Π is the Marchenko-Pastur distribution with parameter 1:

$$d\Pi(x) = \frac{\sqrt{(4-x)\,x}}{2\pi x} 1_{(0,4)}(x) dx.$$

The polynomial P is $P(X_1, X_2) = X_1X_2 + X_2X_1 + X_2^2$, $\mu = \delta_0$ and ν is the standard semi-circular distribution. An economical linearization of P is provided by $L = \gamma \otimes 1 + \alpha \otimes X_1 + \beta \otimes X_2$, where

$$\gamma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Denote by

$$G_\Pi(z) = \int_0^4 \frac{1}{z-t}\, d\Pi(t) = \frac{z + \sqrt{z^2 - 4z}}{2z}, \quad z \in \mathbb{C} \setminus [0,4],$$

the Cauchy transform of the measure ρ. This function satisfies the quadratic equation $zG_\Pi(z)^2 - zG_\Pi(z) + 1 = 0$. Suppose now that $x \notin [0,4]$. Since $\mu = \delta_0$, we have $a = 0$. Denoting by $E = \mathrm{id}_3 \otimes \tau : \mathrm{M}_3(\mathcal{A}) \to \mathrm{M}_3(\mathbb{C})$ the usual expectation and using Remark 8, we have

$$\omega(xe_{11} - \gamma) = E((xe_{11} - \gamma - \alpha \otimes b)^{-1})^{-1}, \quad x \in \mathbb{R} \setminus [0,4].$$

The inverse of $xe_{11} - \gamma - \alpha \otimes b$ is then calculated explicitly and application of the expected value to its entries yields

$$\omega(xe_{11} - \gamma) = \begin{bmatrix} \frac{1}{G_\Pi(x)} & 0 & 0 \\ 0 & \frac{1}{xG_\Pi(x)} - 1 & \frac{1}{2xG_\Pi(x)} + \frac{1}{2} \\ 0 & \frac{1}{2xG_\Pi(x)} + \frac{1}{2} & \frac{1}{4xG_\Pi(x)} - \frac{1}{4} \end{bmatrix}.$$

The equation $\det[\beta\theta - \omega(xe_{11} - \gamma)] = 0$ is easily seen to reduce to

$$\theta^2 G_\Pi(x)^2 - (1 - G_\Pi(x)) = 0. \tag{4.13}$$

This equation has two solutions, namely

$$\frac{2\theta^4}{-(3\theta^2+1) \pm \sqrt{4\theta^2+1}(\theta^2+1)},$$

one of which is negative. The positive solution belongs to $[4, +\infty)$ precisely when $|\theta| > \sqrt{2}$. Thus, the matrix M_N exhibits one (negative) outlier when $0 < |\theta| \leq \sqrt{2}$ and two outliers (one negative and one > 4) when $|\theta| > \sqrt{2}$. This is illustrated by the simulation presented in Figure 4.1.

Note that Shlyakhtenko [129] considered a framework which makes it possible to understand this kind of result as a manifestation of infinitesimal freeness. In fact, the results of [129] also very interestingly allow one to detect the presence of spikes from the behaviour of the bulk of the eigenvalues of the polynomial model, even when there is no outlying eigenvalues.

In the unitarily invariant case, following the strategy of [C17], we also show that the eigenvectors associated to these outlying eigenvalues have projections of computable size onto the eigenspaces of A_N. The results are stated precisely as follows.

Proposition 4. *([C25]) Assume in addition that the spikes of A_N satisfy $\theta_1 > \cdots > \theta_p$, that is, each eigenvalue θ_j has multiplicity one. Assume that $\det H_{i_0}(t) = 0$. Denote by E_{A_N} the spectral measure of A_N (thus, if S is a Borel set in $\mathbb{C}$, then $E_{A_N}(S)$ is the orthogonal projection onto the linear span of all eigenvectors of A_N corresponding to eigenvalues in S).*
Then, for ε small enough, almost surely

$$\lim_{N\to\infty} \left\| E_{A_N}(\{\theta_i\}) \left[E_{P(U_N B_N U_N^*, A_N)}((t-\varepsilon, t+\varepsilon)) - \delta_{i,i_0} \mathcal{C}_i(t) I_N \right] E_{A_N}(\{\theta_i\}) \right\| = 0, \tag{4.14}$$

where $\mathcal{C}_i(t) = \lim_{z\to t}(z-t)\left[(\omega(ze_{11} - \gamma) - \theta_i\beta)^{-1}\right]_{1,1}$ is the residue of the analytic function $z \mapsto \left[(\omega(ze_{11} - \gamma) - \theta_i\beta)^{-1}\right]_{1,1}$ at $z = t$.

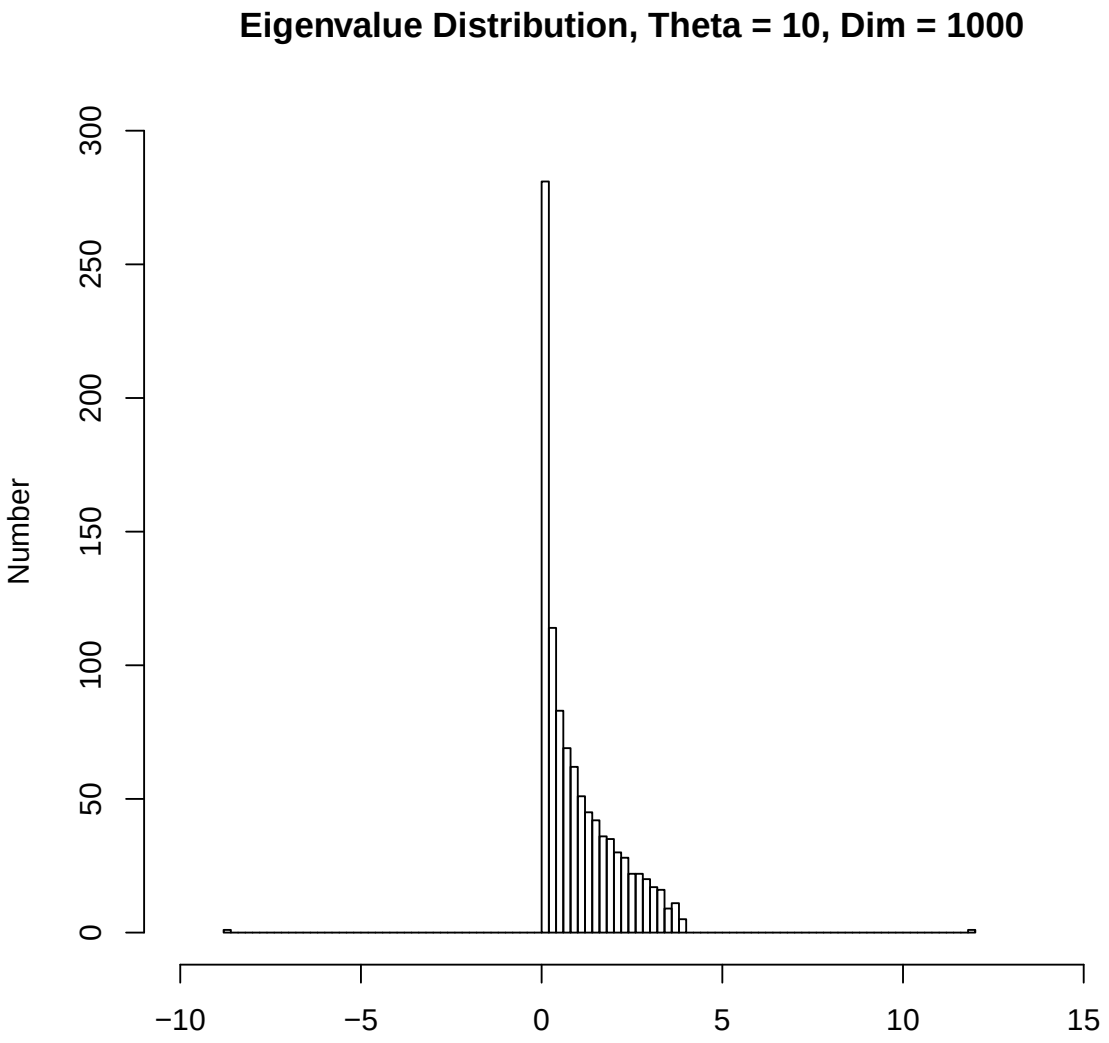


Figure 4.1: One sample from the model described in Remark 8 corresponding to $\theta = 10$, with matrix size $N = 1000$.

The following chapter is only the first stage on the way of the analysis of outliers phenomena for general non-Hermitian random models involving several non-Hermitian matrices and the described work does not use free probability theory. Many things remain to understand and generalize; nevertheless, even at this stage, free probability theory may bring some hint to understand some phenomena (see the end of Section 5.3.3).

Chapter 5

Large deformed non-Hermitian matrices

This chapter presents the joint work [C20] with Charles Bordenave investigating full rank additive perturbations of iid random matrices. Under mild assumptions, as N grows, the empirical distribution of the eigenvalues of full rank additive perturbations of iid random matrices converges weakly to a limit probability measure β on the complex plane. This work is devoted to the study of the outlier eigenvalues, i.e. eigenvalues in the complement of the support of β. Even in the simplest cases, a variety of interesting phenomena can occur. We give a sufficient condition to guarantee that outliers are stable (in the sense that outliers of the deformed model and the perturbation coincide asymptotically as pointed out by Tao in the finite rank case) and provide examples where their fluctuations vary with the particular distribution of the entries of the random matrix or the Jordan decomposition of the perturbation. We also exhibit concrete examples where the outlier eigenvalues converge in distribution to the zeros of a Gaussian analytic function.

5.1 A concrete motivation: numerical instability of eigenvalues

The instability of the eigenvalues of badly conditioned matrices has dramatic consequences in the numerical computation of eigenvalues. As an example, take $N \geq 1$ be an integer and consider the standard nilpotent matrix

$$A_N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \end{pmatrix}. \tag{5.1}$$

Its eigenvalues are obviously all zero. Let U_N be a Haar-distributed orthogonal matrix and consider the unitarily equivalent matrix $B_N = U_N A_N U_N^*$. If we ask a computer to compute the eigenvalues of B_N, we obtain a surprising answer. Figure 5.1 is a plot of these numerically computed eigenvalues of B_N.
In the spirit of von Neumann and Goldstine [152], Spielman and Teng [135] or Edelman and Rao [59], a possible way to try to explain this phenomenon is to approximate numerical rounding errors by randomness and study the spectrum of the matrix

$$A_N + \sigma Y_N,$$

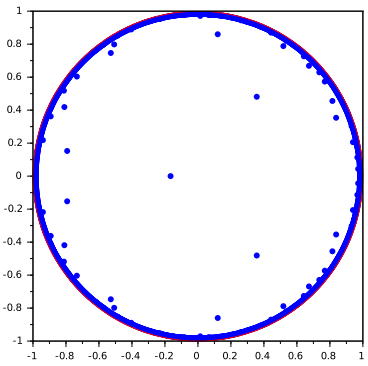


Figure 5.1: The blue dots are the numerically computed eigenvalues of $B_N = U_N A_N U_N^*$, the nilpotent matrix (5.1) conjugated by a Haar-distributed orthogonal matrix U_N $N = 2000$.

where Y_N is a random matrix normalized to have an operator norm of order 1 and σ is small positive parameter. As we shall see, in Section 5.3.4, in the limit $N \to \infty$ and then $\sigma \to 0$, one obtain a reasonable explanation of the picture of Figure 5.1.

5.2 Finite rank perturbations of i.i.d random matrices

Ginibre (1965) introduced the basic non-Hermitian ensemble of random matrix theory. A so-called Ginibre matrix is a $N \times N$ matrix comprised of independent complex Gaussian entries. More generally, an *iid random matrix* is a $N \times N$ random matrix $X_N = (X_{ij})_{1 \leq i,j \leq N}$ whose entries are independent identically distributed complex entries with mean zero and variance 1. The following theorem is the culmination of the work of many authors [7, 11, 72, 74, 102, 110, 138, 140].

Theorem 33. *Let X_N be an iid random matrix. Then the empirical spectral distribution of $\frac{X_N}{\sqrt{N}}$ converges almost surely to the circular measure μ_c where $d\mu_c = \frac{1}{\pi} \mathbf{1}_{|z| \leq 1} dz$.*

One can prove that when the fourth moment is finite, there are no significant outliers to the circular law.

Theorem 34. *(see Theorem 1.4 in [137]) Let X_N be an iid random matrix whose entries have finite fourth moment: $\mathbb{E}(X_{11}|^4) < +\infty$. Then the spectral radius $\rho(\frac{X_N}{\sqrt{N}}) = \sup_{1 \leq j \leq N} \left| \lambda_j \left(\frac{X_N}{\sqrt{N}} \right) \right|$ converges to 1 almost surely as N goes to infinity.*

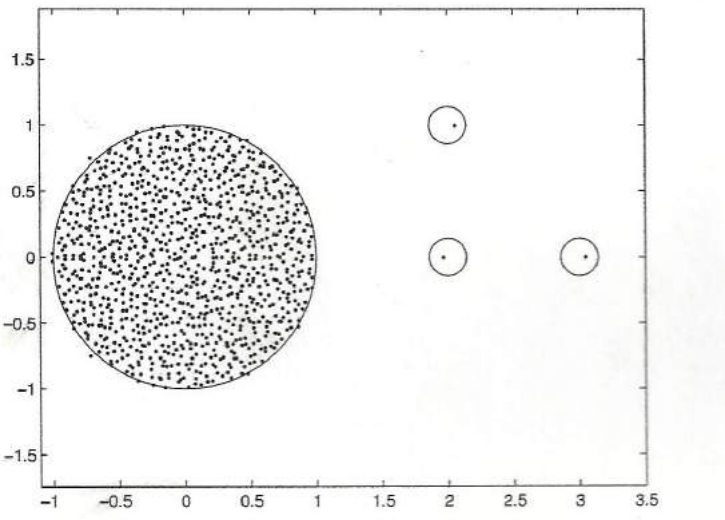
We are interested in the impact of a perturbation on the spectrum. Dealing with low rank perturbations, the global behaviour of the spectrum does not change.

Theorem 35. *[140](Corollary 1.17) Let X_N be an iid random matrix and for each N let A_N be a deterministic matrix with rank $o(N)$ obeying the Frobenius norm bound*

$$\|A_N\|_F = (\mathrm{Tr} A_N A_N^*)^{1/2} = O(N^{1/2}).$$

Then the empirical spectral distribution of $\frac{X_N}{\sqrt{N}} + A_N$ converges almost surely to the circular measure μ_c

However, a low rank perturbation A_N can now create outliers as we can see in the following simulation from Tao [137] describing eigenvalues of a matrix $\frac{X_N}{\sqrt{N}} + A_N$, $N = 1000$, $\mathcal{L}(X_{ij}) = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, with $A_N = \mathrm{diag}(2+i, 3, 2, 0, \ldots, 0)$.



Actually, when A_N has bounded rank and bounded operator norm and the entries of the iid matrix have finite fourth moment, Tao proved that outliers are stable in the sense that outliers of M_N and A_N coincide asymptotically.

Theorem 36. *[137] Let X_N be an iid random matrix whose entries have finite fourth moment: $\mathbb{E}(X_{11}|^4) < +\infty$. Let A_N be a deterministic matrix with rank $O(1)$ and operator norm $O(1)$. Let $\epsilon > 0$, and suppose that for all sufficiently large N, there are :*

- *no eigenvalues of A_N in $\{z \in \mathbb{C} : 1 + \epsilon < |z| < 1 + 3\epsilon\}$,*
- *$j = O(1)$ eigenvalues $\lambda_1(A_N), \ldots, \lambda_j(A_N)$ in $\{z \in \mathbb{C} : |z| \geq 1 + 3\epsilon\}$.*

Then, a.s , for sufficiently large N, there are precisely j eigenvalues of $\frac{X_N}{\sqrt{N}} + A_N$ in $\{z \in \mathbb{C} : |z| \geq 1 + 2\epsilon\}$ and after labeling these eigenvalues properly, as N goes to infinity, for each $1 \leq i \leq j$,

$$\lambda_i(\frac{X_N}{\sqrt{N}} + A_N) = \lambda_i(A_N) + o(1).$$

Now, our issue is to investigate this problem dealing with full rank perturbations.

5.3 Full rank perturbations

We will consider the deformed model:

$$M_N = A_N + \sigma Y_N, \tag{5.2}$$

where $\sigma > 0$, Y_N is a $N \times N$ random matrix and A_N is a $N \times N$ deterministic matrix. The matrix M_N can be thought as a random perturbation of the matrix A_N. We set

$$Y_N = \frac{X_N}{\sqrt{N}}, \tag{5.3}$$

and we shall consider the following set of statistical assumptions on the matrices $X_N = (X_{ij})_{1 \leq i,j \leq N}$:

(X1) $(X_{ij})_{i,j \geq 1}$ are independent and identically distributed complex random variables with $\mathbb{E}X_{ij} = 0$, $\mathbb{E}|X_{ij}|^2 = 1$.

(X2) $\mathbb{E}|X_{ij}|^4 < \infty$.

(X3) There exists $c > 0$ such that for all $k \geq 1$ integer $\mathbb{E}|X_{ij}|^k \leq (ck)^c$.

Our first assumptions on the matrices A_N are as follows:

(A1) There exists $M > 0$ such that for all N, $\|A_N\| \leq M$.

(A2) For all $z \in \mathbb{C}$, $\mu_{(A_N - zI_N)(A_N - zI_N)^*}$ converges weakly to a probability measure ν_z.

Example: If A_N converges in $\star$-moments to an operator a in a $\mathcal{C}^*$-non commutative probability space $(\mathcal{A}, \tau)$, i.e. for all $\varepsilon_l \in \{1, *\}$,

$$\frac{1}{N}\mathrm{Tr}\left(A_N^{\varepsilon_1} \dots A_N^{\varepsilon_k}\right) \overset{N \to +\infty}{\longrightarrow} \tau\left(a^{\varepsilon_1} \dots a^{\varepsilon_k}\right),$$

then ν_z is the distribution of $(a - z)(a - z)^*$.

5.3.1 The limiting empirical spectral distribution

In general, the limiting spectral distribution β of $M_N = \sigma \frac{X_N}{\sqrt{N}} + A_N$ is no more the circular law.
For a probability measure τ on $\mathbb{C}$ such that $\int \log(1 + |\lambda|) d\tau(\lambda) < \infty$, we denote by h_τ its logarithmic potential defined for $z \in \mathbb{C}$, by

$$h_\tau(z) = -\int_{\mathbb{C}} \log|\lambda - z| d\tau(\lambda).$$

There are various possible characterizations of the limit measure β, the usual relies on its logarithmic potential. It is expressed in terms of Cauchy-Stieltjes transform of the limit measures of shifted singular values of M_N as follows.
For any $z \in \mathbb{C}$, denote by

$$M_N^z = \sigma Y_N + A_N - zI_N.$$

According to Dozier and Silverstein [56], almost surely the empirical spectral measure $\mu_{M_N^z M_N^{z\,*}}$ of $M_N^z M_N^{z\,*}$ converges weakly towards a nonrandom distribution μ_z which is characterized in terms of its Stieltjes transform which satisfies the following equation: for any $w \in \mathbb{C}^+$,

$$g_{\mu_z}(w) = \int \frac{1}{(1 - \sigma^2 g_{\mu_z}(w))w - \frac{t}{1 - \sigma^2 g_{\mu_z}(w)}} d\nu_z(t). \tag{5.4}$$

According to [140, 131], see also [38], almost surely the empirical spectral measure of μ_{M_N} converges weakly to a probability measure β on $\mathbb{C}$ which is characterized by its logarithmic potential

$$h_\beta(z) = -\frac{1}{2}\int \log(t) d\mu_z(t).$$

For more references, we refer to the surveys [139, 38].

Remark 9. *If A_N converges in $\star$-moments to an operator a, set*

$$b = \sigma c + a,$$

where c is a circular element free of a. Then, μ_z is the distribution of $(b-z)(b-z)^*$, β is the Brown's measure of b.

Concerning Brown's measure, we refer to [83, 35, 131, 39]. Explicit computation of β are rare, see Biane and Lehner [35].

Observe that z is an eigenvalue of $B \in M_N(\mathbb{C})$ if and only if 0 is an eigenvalue of $(B-zI)(B-zI)^*$. Recall that

$$\forall z \in \mathbb{C}, \quad \mu_{(M_N - zI_N)(M_N - zI_N)^*} \xrightarrow{w} \mu_z, \quad \mu_{M_N} \xrightarrow{w} \beta$$

We will assume that a similar property holds for β and μ_z:

$$(A3) \qquad \mathrm{supp}(\beta) = \{\mathrm{z} \in \mathbb{C} : 0 \in \mathrm{supp}(\mu_\mathrm{z})\}.$$

This is true for instance in most cases when A_N is normal. Under assumption (A3) the support of β takes a particularly simple expression.

Proposition 5. *([C20]) If (A3) holds then*

$$\mathrm{supp}(\beta) = \left\{ \mathrm{z} \in \mathbb{C} : 0 \in \mathrm{supp}(\nu_\mathrm{z}) \text{ or } \int \lambda^{-1} \mathrm{d}\nu_\mathrm{z}(\lambda) \geq \sigma^{-2} \right\}.$$

We are interested by describing the individual eigenvalues of M_N outside $\{z \in \mathbb{C}, d(z, \mathrm{supp}(\beta)) \leq \epsilon\}$ for some $\epsilon > 0$.

5.3.2 Examples, simulations

We start with three examples and simulations that illustrate different phenomena we will study: lack of outliers, stable or unstable outliers ("stable"meaning that outliers of the deformed model and the perturbation coincide asymptotically) . Let us consider the three following perturbations.

(1)

$$C_N = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \ddots \\ 1 & 0 & \cdots & \cdots \end{pmatrix}. \tag{5.5}$$

(2)

$$B_N = \begin{pmatrix} C_{N-5} & 0 \\ 0 & D_5 \end{pmatrix}, \quad D_5 = \mathrm{diag}(0; 0; \mathrm{i}/3; 1; 2). \tag{5.6}$$

(3)

$$\mathcal{N}_N = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \ddots \\ 0 & 0 & \cdots & \cdots \end{pmatrix}. \tag{5.7}$$

C_N is a normal matrix with eigenvalues $e^{\frac{2i\pi l}{N}}$, $1 \leq l \leq N$. When $A_N = C_N$, the limiting empirical spectral distribution β of $M_N = \sigma\frac{X_N}{\sqrt{N}} + A_N$, is the Brown's measure of $\sigma c + a$ where a is a Haar unitary and c is a circular element free with a and

$$\begin{aligned}\text{supp}(\beta) &= \left\{z \in \mathbb{C} : \frac{1}{2\pi}\int_0^{2\pi} \frac{dx}{|e^{ix} - z|^2} \geq \sigma^{-2}\right\} \\ &= \left\{z \in \mathbb{C} : \sqrt{(1-\sigma^2)_+} \leq |z| \leq \sqrt{1+\sigma^2}\right\}.\end{aligned}$$

Now, the difference between the circulant matrix C_N and any of the matrices $\mathcal{N}_N$, B_N, has rank $o(N)$. Therefore, for $A_N \in \{C_N, \mathcal{N}_N, B_N\}$, for any $z \in \mathbb{C}$, $\nu_z = \lim_{N\to} \mu_{(A_N - zI_N)(A_N - zI_N)^*}$ is the same. Thus the μ_z's and the limiting empirical spectral distribution β of $M_N = \sigma\frac{X_N}{\sqrt{N}} + A_N$ (uniquely determined by the ν_z's) are the same.

Here are simulations of the eigenvalues of $M_N = \sigma\frac{X_N}{\sqrt{N}} + A_N$ dealing with $N = 500$, $\sigma^2 = 1/2$, X_{ij} real Gaussian and $A_N \in \{C_N, \mathcal{N}_N, B_N\}$. Then $\text{supp}(\beta) = \{z \in \mathbb{C} : 1/\sqrt{2} \leq |z| \leq \sqrt{3/2}\}$.

(1) When $A_N = C_N$, we see that there is no outlier.

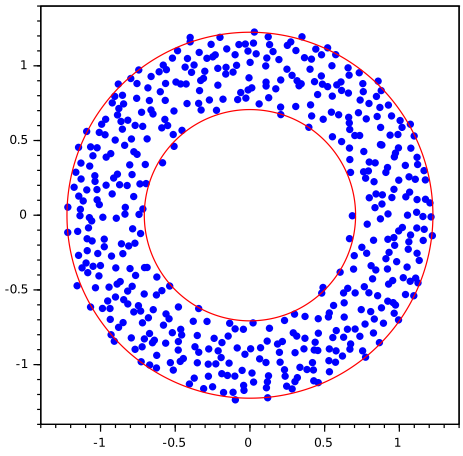


Figure 5.2: no outlier

(2) Now, when $A_N = B_N$, there are outliers in a neigborhood of 0, 2 and $i/3$ so that the outliers of the deformed model seem stable in the sense that they stay close to the outliers of A_N (see Figure 5.3).

(3) Finally, when $A_N = \mathcal{N}_N$, there are unstable outliers in the bounded component of the complement of $\text{supp}(\beta)$ in the sense that there are not in a neighborhood of those of A_N (see Figure 5.4).

The first aim is to understand these phenomena.

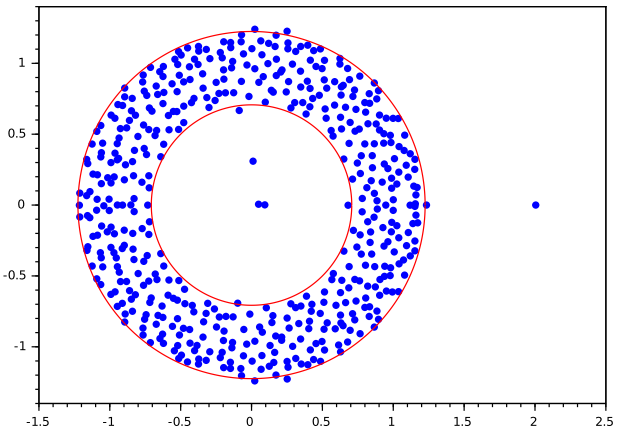


Figure 5.3: stable outliers

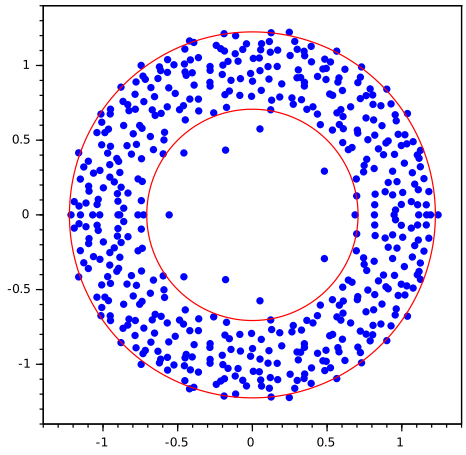


Figure 5.4: unstable outliers

5.3.3 Stable outliers

We first give a sufficient condition to guarantee that outliers of the deformed model are stable. For this purpose, we introduce the notion of well-conditioned matrix which is related to the phenomenon of lack of

outlier and of well-conditioned decomposition of A_N which lead to the statement of a sufficient condition for the stability of the outliers. We will denote by $s_1(B) \geq \cdots \geq s_N(B)$ the singular values of any $N \times N$ matrix B. For any set K and any $\epsilon > 0$, $B(K, \epsilon)$ stands for the set $\{z \in \mathbb{C}, d(z, K) \leq \epsilon\}$.

No outlier

Definition 1. *Let $\Gamma \subset \mathbb{C} \setminus \text{supp}(\beta)$ be a compact set. A_N is well-conditioned in Γ if for any $z \in \Gamma$, there exists $\eta > 0$ such that for all N large enough, $s_N(A_N - zI_N) > \eta$.*

Theorem 37. *([C20]) Assume that A_N is well-conditioned in Γ, Then, a.s. for all N large enough, M_N has no eigenvalue in Γ.*

Corollary 1. *If for any $z \in \mathbb{C} \setminus \text{supp}(\beta)$, there exists $\eta > 0$ such that for all N large enough, $s_N(A_N - zI_N) > \eta$, then, for any $\varepsilon > 0$, a.s. for all N large enough, all eigenvalues of M_N are in $B(\text{supp}(\beta), \varepsilon)$.*

The proof relies on previous results on Information-plus-noise type models [57, 12], [C18].

The circulant matrix C_N defined by (5.5) is well-conditioned in $\mathbb{C} \setminus \text{supp}(\beta)$ where

$$\text{supp}(\beta) = \left\{ \text{z} \in \mathbb{C} : \sqrt{(1-\sigma^2)_+} \leq |\text{z}| \leq \sqrt{1+\sigma^2} \right\}.$$

Indeed, for any z outside the support of β, the singular values of $C_N - zI_N$ are equal to $|e^{\frac{2i\pi l}{N}} - z| \geq |1 - |z|| > 0, \text{for } \ l = 1, \ldots, N$. This explains simulations for (1).
Contrariwise, $A_N = B_N$ defined by (5.6) provides an example of badly-conditioned perturbation in neighborhoods of 0, $i/3$ and 2 since $s_N(A_N) = s_N(A_N - i/3I_N) = s_N(A_N - 2I) = 0$.
An example of badly-conditioned perturbation in the inner disk is the nilpotent matrix $\mathcal{N}_N$. Indeed, for $|z| > 1$, $s_N(\mathcal{N}_N - zI_N) \geq |z| - 1$; therefore $\mathcal{N}_N$ is well conditioned out of the maximum circle but for $|z| < 1$, $s_N(\mathcal{N}_N - zI_N) = o(1)$.

Let us introduce now the notion of *well-conditioned decomposition* of A_N which will allow us to exhibit a sufficient condition for stability of outliers.

Definition 2. *Let $\Gamma \subset \mathbb{C} \backslash \text{supp}(\beta)$ be a compact set. A_N admits a well-conditioned decomposition if : $A_N = A'_N + A''_N$ where*

- *There exists $M > 0$ such that for all N, $\|A'_N\| + \|A''_N\| \leq M$.*
- *For any $z \in \Gamma$, there exists $\eta > 0$ such that for all N large enough, $s_N(A'_N - zI_N) > \eta$ (i.e A'_N is well-conditioned in Γ) and A''_N has rank $r = O(1)$.*

Stable outliers

Theorem 38. *([C20]) Let $\Gamma \subset \mathbb{C} \backslash \text{supp}(\beta)$ be a compact set with continuous boundary. Assume that A_N admits a well-conditioned decomposition: $A_N = A'_N + A''_N$. If for some $\varepsilon > 0$ and all N large enough,*

$$\min_{z \in \partial\Gamma} \left| \frac{\det(A_N - z)}{\det(A'_N - z)} \right| \geq \varepsilon, \tag{5.8}$$

then a.s. for all N large enough, the number of eigenvalues of A_N and M_N in Γ is equal.

Let us give a concrete application of Theorem 38 with a specific decomposition of $A_N = A'_N + A''_N$. Assume that for all N, there exists a subset $J \subset \{1, \cdots, N\}$ of cardinal at most r such that for any $\varepsilon > 0$, for all N large enough and $k \in J$, $\lambda_k(A_N) \notin B(\mathrm{supp}(\beta), \varepsilon)$. We consider a triangular decomposition of A_N:

$$A_N = P \left(\begin{array}{c|c} T'' & * \\ \hline 0 & T' \end{array} \right) P^{-1},$$

where P is an invertible matrix, T' is an upper triangular matrix of size $N - |J|$ with the eigenvalues $\lambda_k(A_N)$, $k \notin J$, on the diagonal, and T'' is an upper triangular matrix of size $|J|$ with diagonal entries $\lambda_k(A_N)$, $k \in J$. Fix some $a \in \mathrm{supp}(\beta)$, we decompose A_N as $A_N = A'_N + A''_N$ with

$$A'_N = P \left(\begin{array}{c|c} aI_J & * \\ \hline 0 & T' \end{array} \right) P^{-1} \quad \text{and} \quad A''_N = P \left(\begin{array}{c|c} T'' - aI_J & 0 \\ \hline 0 & 0 \end{array} \right) P^{-1}. \tag{5.9}$$

We take $a \in \mathrm{supp}(\beta)$ to avoid addition of outliers of A'_N. Note that

$$\frac{\det(A_N - z)}{\det(A'_N - z)} = \prod_{k \in J} \frac{\lambda_k(A_N) - z}{a - z}. \tag{5.10}$$

The next statement will be an easy consequence of Theorem 38. It generalizes Tao [137, Theorem 1.7] where $A'_N = 0$. When A'_N is a Wigner random matrix, it is a special case of O'Rourke and Renfrew [108, Theorem 2.4].

Corollary 2. *Assume that assumptions (X1-X2) hold and that A'_N, A''_N given by (5.9) provide a well-conditioned decomposition of A_N in $\Gamma = \mathbb{C} \setminus \mathrm{supp}(\beta)$. Fix $\varepsilon > 0$. Assume that for all N large enough, $\forall k \in J$, $\lambda_k(A_N) \notin B(\mathrm{supp}(\beta), 3\varepsilon)$. Then, a.s. for all N large enough, there are exactly $|J|$ eigenvalues of M_N in $\mathbb{C} \setminus B(\mathrm{supp}(\beta), 2\varepsilon)$. Moreover, if we index them by $\lambda_k(M_N)$, $k \in J$, after labeling properly, a.s.*

$$\max_{k \in J} |\lambda_k(M_N) - \lambda_k(A_N)| \to 0.$$

This result allows us to understand the simulation involving B_N in (2). Indeed, one can write

$$\begin{aligned} A_N &= \begin{pmatrix} C_{N-5} & 0 \\ 0 & I_5 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -I_5 + D_5 \end{pmatrix} \\ &= A'_N + A''_N \end{aligned}$$

with

$$\left| \frac{\det(A_N - zI_N)}{\det(A'_N - zI_N)} \right| = \left| \frac{z^2(i/3 - z)(2 - z)}{(1 - z)^4} \right|.$$

Therefore, for any $\delta > 0$ small enough, for any z_0 in $\{0; i/3; 2\}$, there exists $\epsilon > 0$ such that

$$\min_{z \in \partial B(z_0; \delta)} \left| \frac{\det(A_N - zI_N)}{\det(A'_N - zI_N)} \right| \geq \epsilon.$$

Thus, the stability of the outliers follows.

Comparison with finite rank perturbations of the single ring model [31]

In [31], Benaych-Georges and Rochet consider matrices of the type $M_N = U_N D_N V_N + A_N$ with diagonal $D_N > 0$, μ_{D_N} converges towards a compactly supported probability measure ν on $\mathbb{R}^+$, U_N, V_N are independent Haar unitary matrices, A_N has a finite rank and a bounded operator norm.
The limiting empirical eigenvalues distribution of such a model is described by the so-called single ring theorem, see [75, 76, 123]. It is deterministic and its support is $\{z \in \mathbb{C}; a \leq |z| \leq b\}$ where $a = (\int x^{-2} d\nu(x))^{-1/2}$ and $b = (\int x^2 d\nu(x))^{1/2}$. Benaych-Georges and Rochet prove that if A_N has some eigenvalues out of the maximal circle of the single ring, then M_N has outliers in the neighborhood of these eigenvalues of A_N. Nevertheless, when $a > 0$, the eigenvalues of A_N which may be in the inner disk of the complement of the limiting support do not generate outliers in the spectrum of M_N.

Now, in the framework of [C20] dealing with full rank perturbations of iid matrices, there can be outlier eigenvalues in bounded components of the complement of $\mathrm{supp}(\beta)$.
Actually, the nature of the bounded connected component of the complement of the support of the limiting empirical eigenvalues distributions considering above is different: the first (in [31]) comes from the limiting support of the non-deformed model whereas the second one (in the framework of [C20]) is created by the deformation. Subordination-like properties of the Stieltjes transform $g_\mu(z) = \int d\mu(\lambda)/(z-\lambda)$ of limiting spectral measures may help to understand these phenomena as explained below.
In the case of [31], since the limiting empirical eigenvalues distribution μ is radial, we have $g_\mu(z) = \frac{1}{z}$ if $|z| > b$ and $g_\mu(z) = 0$ if $|z| < a$ so that roughly speaking

$$g_\mu(z) = g_{\delta_0}(\omega(z)) \quad \text{where} \quad \begin{cases} \omega(z) = z \text{ if } |z| > b \\ \text{“}\omega(z) = \infty\text{” if } |z| < a. \end{cases}$$

In our case, dealing for instance with diagonal perturbations of a Ginibre matrix, the limiting empirical eigenvalues distribution β is the Brown measure of $c + a$ where c is a circular element which is free with a whose Brown measure is α (see Śniady [131]). We have the following subordination property.

$$\forall z \in \mathbb{C} \setminus \mathrm{supp}(\beta),\ \mathrm{g_{c+a}(z)} = \mathrm{g_a}(\omega(\mathrm{z})) \ \text{ where } \ \omega(\mathrm{z}) = \mathrm{z}.$$

It can be deduced from [39, Proposition 4.3], see also [35, 150].
In both cases, the intuition is that

$$g_{\mu_{M_N}}(z) \approx g_{\mu_{A_N}}(\omega(z))$$

and that therefore they will be eigenvalues ρ of M_N that separate from the bulk whenever some of the equations $\omega(\rho) = \theta$ admits a solution ρ outside the limiting support, when θ describes the spectrum of A_N. Therefore, we understand on the one hand that in the framework of [31], there is no solution inside the inner disk of such an equation since there $\omega \equiv \infty$ and on the other hand that the outliers of the deformed model stay in the neighborhood of the eigenvalues of the perturbation which are located where ω is the identity function.

In the deformed Hermitian models, the outliers of the deformed model are not located in a neighborhood of the spikes of the deformation. It contrasts with Corollary 2. It is rather non-intuitive that additive perturbation by a Hermitian random matrix has more effect on outlier eigenvalues than additive perturbation by a non-Hermitian random matrix.

5.3.4 Unstable outliers

(3) provides a typical situation where (5.8) does not hold. Let us consider the nilpotent matrix defined in (5.7),

$$\mathcal{N}_\mathcal{N} = C_N - e_N e_1^*.$$

Set $A'_N = C_N$ and $A''_N = -e_N e_1^*$. For any $|z| \leq 1-\varepsilon$, we have

$$\left|\frac{\det(\mathcal{N}_N - zI_N)}{\det(C_N - zI_N)}\right| = \frac{|z|^n}{|1-z^N|} \leq \frac{(1-\varepsilon)^N}{1-(1-\varepsilon)^N} = o(1/\sqrt{N}).$$

We see numerically that the conclusion of Theorem 38 does not seem to hold. Actually, the simulations can be explained by the following results.

Theorem 39 (C20)**.** *Let $\Gamma \subset \mathbb{C}\backslash\mathrm{supp}(\beta)$ be a compact set with continuous boundary. Assume that $A_N = A'_N + A''_N$ is a well conditioned decomposition, A'_N is diagonal and $A''_N = v_N u_N^*$ where $u_N, v_N \in \mathbb{C}^N$, $\|u_N\|_\infty, \|v_N\|_\infty = O(1/\sqrt{N})$. Assume moreover that*

$$\max_{z\in\Gamma}\left|\frac{\det(A_N - z)}{\det(A'_N - z)}\right| = o\left(\frac{1}{\sqrt{N}}\right).$$

Define $R'_N(z) = (zI_N - A'_N)^{-1}$, $\varphi_N(z,w) = \frac{1}{N}\mathrm{TrR'_N(z)(R'_N(w))^*}$, $\psi_N(z,w) = \frac{1}{N}\mathrm{TrR'_N(z)R'_N(w)}$. Let $(g_N(z))_{z\in\Gamma}$ be a centered Gaussian process with covariance given by, for $z, w \in \Gamma$,

$$\begin{aligned}
\mathbb{E}g_N(z)\bar{g}_N(w) &= \frac{u_N^* R'_N(z)(R'_N(w))^* u_N v_N^*(R'_N(w))^* R'_N(z)v_N}{1-\sigma^2\varphi_N(z,w)} \\
\mathbb{E}g_N(z)g_N(w) &= \frac{u_N^* R'_N(z)R'_N(w)\bar{u}_N v_N^T(R'_N(w))^T R'_N(z)v_N \mathbb{E}X_{11}^2}{1-\mathbb{E}X_{11}^2\sigma^2\psi_N(z,w)}
\end{aligned}$$

The Lévy-Prohorov distance between the point process of eigenvalues of M_N in $\mathring{\Gamma}$ and the point process of zeros of g_N in $\mathring{\Gamma}$ goes to 0 as N goes to ∞.

The intensity of zeros of g_N can be computed explicitly thanks the Edelman-Kostlan's formula, see Theorem 3.1 in [58].

Note that if A'_N is a normal matrix and X_{ij} are standard complex Gaussian variables, then, by unitary invariance, we can always assume that A'_N is diagonal. We may rewrite the nilpotent matrix $\mathcal{N}_N$ in the orthonormal basis of eigenvectors of A'_N:

$$C_N = U_N D'_N U_N^* \quad \text{and} \quad \mathcal{N}_N = U_N(D'_N + D''_N)U_N^*$$

where

$$D'_N = \mathrm{diag}(e^{\frac{2i\pi}{N}}, \cdots, e^{\frac{2iN\pi}{N}}) \quad \text{and} \quad D''_N = -f_N f_1^\top.$$

with $f_l = (e^{\frac{2i\pi lk}{N}}/\sqrt{N})_{1\leq k\leq N}$. Hence the next corollary deals with the phenomenon illustrated by Figure 5.4.

Corollary 3. *$0 < \sigma < 1$, $A_N = \mathcal{N}_N$ and assume moreover that X_{ij} is gaussian. The point process of eigenvalues of M_N in $\mathring{B}(0,\sqrt{1-\sigma^2})$ converges weakly to the point process of the zeros of the centered Gaussian process $g(z)$ on $\mathring{B}(0,\sqrt{1-\sigma^2})$ such that for $z, w \in \Gamma$,*

$$\begin{aligned}
\mathbb{E}(g(z)g(w)) &= 0, \\
\mathbb{E}g(z)\overline{g(w)} &= \frac{\varphi(z,w)^2}{1-\sigma^2\varphi(z,w)} \text{ where } \varphi(z,w) = \frac{1}{1-z\bar{w}}.
\end{aligned}$$

We may notice that, as $\sigma \to 0$, the kernel $K(z,w)$ appearing in Corollary 3 does not vanish, it converges pointwise to the kernel $K_0(z,w) = \varphi(z,w)^2$ on the unit complex disc. The kernel K_0 is the kernel of the Gaussian analytic function

$$g(z) = \sum_{k=0}^{\infty} z^k \gamma_k \sqrt{k+1},$$

where γ_k are iid complex Gaussian variables with $\mathbb{E}\gamma_k^2 = 0$, $\mathbb{E}|\gamma_k|^2 = 1$. This Gaussian analytic function may thus be related to the numerical phenomenon illustrated by the plot of Figure 5.1.

5.3.5 Basic ideas of the proofs

We present here the ideas of the proofs when A_N'' has rank 1 and A_N' is diagonal. Let $\Gamma \subset \mathbb{C} \setminus \mathrm{supp}(\beta)$ be a compact set. Let $A_N = A_N' + A_N''$ be a well-conditioned decomposition such that $A_N'' = v_N u_N^*$, $u_N, v_N \in \mathbb{C}^N$, and A_N' is diagonal.
Define $M_N' := \sigma \frac{X_N}{\sqrt{N}} + A_N'$ so that $M_N = M_N' + v_N u_N^*$. For any $\lambda \in \Gamma$, define

$$R_N^{(M')}(\lambda) := (\lambda I_N - M_N')^{-1}, \quad R_N^{(A')}(\lambda) := (\lambda I_N - A_N')^{-1}.$$

Using the identity

$$det(B + PQ) = det(B) det(I_r + QB^{-1}P), \quad P, Q^T \in M_{N,r}(\mathbb{C}), B \in GL_N(\mathbb{C})$$

for

$$\begin{cases} B = \lambda I_N - M_N' \\ P = v_N \\ Q = u_N^* \\ B = \lambda I_N - A_N' \end{cases}$$

we obtain that

$$\begin{cases} \frac{\det(\lambda I_N - M_N)}{\det(\lambda I_N - M_N')} = 1 - u_N^* R_N^{(M')}(\lambda) v_N \\ \\ \frac{\det(\lambda I_N - A_N)}{\det(\lambda I_N - A_N')} = 1 - u_N^* R_N^{(A')}(\lambda) v_N. \end{cases}$$

Therefore the eigenvalues of M_N and A_N, in Γ, are respectively the zeroes of $1 - u_N^* R_N^{(M')} v_N$ and $1 - u_N^* R_N^{(A')} v_N$.

Now, since $M_N' := \sigma \frac{X_N}{\sqrt{N}} + A_N'$, we have $R_N^{(M')} = R_N^{(A')} + R_N^{(A')} \sigma \frac{X_N}{\sqrt{N}} R_N^{(M')}$ and thus,

$$1 - u_N^* R_N^{(M')} v_N = 1 - u_N^* R_N^{(A')} v_N - \sum_{k \geq 1} u_N^* \left(R_N^{(A')} \sigma \frac{X_N}{\sqrt{N}} \right)^k R_N^{(A')} v_N.$$

It turns out that when N goes to infinity $\sum_{k\geq 1} u_N^* \left(R_N^{(A')} \sigma \frac{X_N}{\sqrt{N}} \right)^k R_N^{(A')} v_N$ converges towards zero and $\left(\sqrt{N} u_N^* \left(R_N^{(A')}(z_i) \sigma \frac{X_N}{\sqrt{N}} \right)^k R_N^{(A')}(z_i) v_N \right)_{1 \leq i \leq m}$ is "asymptotically close to" a gaussian vector. Hence we have

$$\overbrace{1-u_N^* R_N^{(M')} v_N}^{\text{whose zeroes=eigenvalues of } M_N \downarrow} = \overbrace{1-u_N^* R_N^{(A')} v_N}^{\text{whose zeroes=eigenvalues of } A_N \downarrow} - \overbrace{\sum_{k\geq 1} u_N^* \left(R_N^{(A')} \sigma \frac{X_N}{\sqrt{N}} \right)^k R_N^{(A')} v_N}^{0 \uparrow N\to+\infty}$$

We are naturally brought to consider the two following situations.

(i) If $\min_{z\in\partial\Gamma} \left| \frac{\det(\lambda I_N - A_N)}{\det(\lambda I_N - A'_N)} \right| = \min_{z\in\partial\Gamma} \left| 1 - u_N^* R_N^{(A')}(\lambda) v_N \right| \geq \varepsilon$, the term $1 - u_N^* R_N^{(A')}(\lambda) v_N$ plays the central role in the behaviour of the outliers of M_N and gives rise to stable outliers by Rouché's theorem.

(ii) If $\max_{\lambda\in\Gamma} \left| \frac{\det(\lambda I_N - A_N)}{\det(\lambda I_N - A'_N)} \right| = \max_{\lambda\in\Gamma} \left| 1 - u_N^* R_N^{(A')}(\lambda) v_N \right| = o(\frac{1}{\sqrt{N}})$, let us rewrite

$$\overbrace{\sqrt{N}\left(1 - u_N^* R_N^{(M')}(\lambda) v_N\right)}^{\text{whose zeroes=eigenvalues of } M_N} = \overbrace{\sqrt{N}\left(1 - u_N^* R_N^{(A')}(\lambda) v_N\right)}^{0 \uparrow N\to+\infty} - \sum_{k\geq 1} \underbrace{\sqrt{N} u_N^* \left(R_N^{(A')}(\lambda) \sigma \frac{X_N}{\sqrt{N}} \right)^k R_N^{(A')}(\lambda) v_N}_{\text{asymptotically close to a gaussian variable with independence for } k \neq k'}$$

Therefore, $\sum_{k\geq 1} \sqrt{N} u_N^* \left(R_N^{(A')} \sigma \frac{X_N}{\sqrt{N}} \right)^k R_N^{(A')} v_N$ plays the central role in the behaviour of the outliers of M_N and gives rise to unstable outliers through limit theorems for random analytic functions and their zeros from [128].

Remark 10. *In the unbounded component of $\mathbb{C} \setminus \mathrm{supp}(\beta)$, (ii) cannot hold.*

Nevertheless, as in Rajan and Abbott [118] and Tao [137], outliers may appear in the unbounded component of $\mathbb{C}\backslash\mathrm{supp}(\beta)$ when $\|A''_N\|$ is of order $\sqrt{N}$. It was observed in [118] and [137], for $A'_N = 0$ and a particular random choice of A''_N.

Theorem 40. *Assume that assumptions (X1-X3) and assumptions (A2-A3) hold. Assume that A''_N has rank 1, A'_N is diagonal and for some $\varepsilon > 0$ and all N large enough, all eigenvalues of A'_N lie in $\mathbb{C}\backslash B(\Gamma, \varepsilon)$. Let $\Gamma \subset \mathbb{C}\backslash\mathrm{supp}(\beta)$ be a compact set with continuous boundary. We set $R'_N(z) = (zI_N - A'_N)^{-1}$, and assume further that $\|A'_N\| = O(1)$ and $A''_N = \sqrt{N} v_N u_N^*$ where $\|u_N\|_\infty, \|v_N\|_\infty$ and $u_N^* R'_N(z) v_N$ are of order $O(1/\sqrt{N})$. Consider the centered Gaussian process $(g_N(z))_{z\in\Gamma}$ of Theorem 39. Then, the Lévy-Prohorov distance between the point process of eigenvalues of M_N in $\mathring{\Gamma}$ and the point process of zeros of $1 - \sqrt{N} u_N^* R'_N(z) v_N + \sigma g_N(z)$ in $\mathring{\Gamma}$ goes to 0 as N goes to ∞.*

This result is illustrated with Figure 5.5. As an application, we have for example the following corollary which is related to [137, Theorem 1.11].

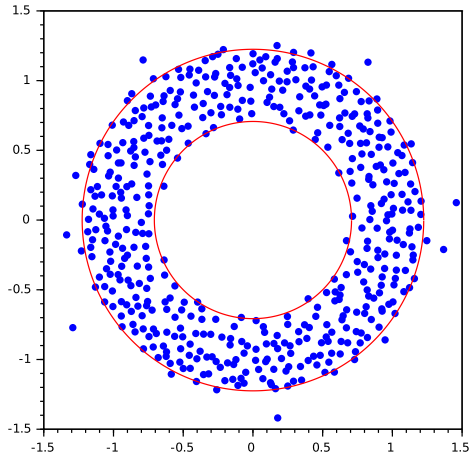


Figure 5.5: Eigenvalues of M_N where $A_N = \text{diag}\left(e^{\frac{2i\pi}{N}}, \cdots, e^{\frac{2iN\pi}{N}}\right) + \sqrt{N} f_N f_1^\top$, with $f_l = (e^{\frac{2i\pi lk}{N}}/\sqrt{N})_{1\leq k\leq N}$., $N = 500$, $\sigma^2 = 1/2$ and X_N has complex Gaussian entries. For $|z| > 1$, we have $f_N^T R'_N(z) f_1 = 1/(z^N - 1) = o(1/\sqrt{N})$. The outlier eigenvalues in the unbounded component of the complement of $\text{supp}(\beta) = \{z \in \mathbb{C} : 1/\sqrt{2} \leq |z| \leq \sqrt{3/2}\}$ converge in distribution to the zeros of $1 + \sigma g$ where g is the Gaussian analytic function with kernel $H(z, w) = \varphi(z, w)^2/(1 + \sigma^2\varphi(z, w))$ and φ given by $\varphi(z, w) = \frac{1}{1-z\bar{w}}$.

Corollary 4. *Assume that assumptions (X1-X3) hold, that $\sigma = 1$, $A'_N = 0$ and $A_N = A''_N = \theta_N v_N u_N^T$ with $\sqrt{N}/\theta_N \to \kappa \in \mathbb{C}$, $u_N, v_N \in \mathbb{R}^N$, $\|u_N\| = \|v_N\| = 1$, $\theta_N u_N^T v_N \to \lambda \in \mathbb{C}$ and $\|u_N\|_\infty, \|v_N\|_\infty$ are of order $O(1/\sqrt{N})$. Fix $\varepsilon > 0$ and set $\Gamma = \mathbb{C}\backslash B(0, 1 + \varepsilon)$. We consider $g(z) = \sum_{k\geq 0} \gamma_k z^{-k}$ with γ_k independent complex Gaussian variables with variance given by $\mathbb{E}|\gamma_k|^2 = 1$ and $\mathbb{E}\gamma_k^2 = (\mathbb{E}X_{11}^2)^{k+1}$. Then, as N goes to ∞, the point process of eigenvalues of M_N in Γ converges vaguely to the point process of zeros of $\kappa z(z - \lambda) + g(z)$ in Γ.*

5.3.6 Fluctuations of Stable Outliers

Now, we provide examples where the fluctuations of stable outliers vary with the particular distribution of the entries of X_N or the Jordan decomposition of A_N. We have studied the fluctuations of the convergence of outliers eigenvalues in the simplest case for the decomposition of A_N on its outlier eigenspace.

First, more precisely, we will suppose that A_N has the following decomposition, for some integer $r \geq 1$ and complex number θ_N,

$$A_N = \left(\begin{array}{c|c} \theta_N I_r & 0 \\ \hline 0 & \hat{A}_{N-r} \end{array}\right). \tag{5.11}$$

Theorem 41. *Suppose that assumptions (X1-X2) and assumptions (A1-A3) hold with A_N given by (5.11). Set $m_2 = \mathbb{E}(X_{11}^2)$. We suppose further that θ_N converges toward $\theta \in \mathbb{C} \setminus \text{supp}(\beta)$ when N*

goes to infinity and that for some $\eta > 0$ *and all large* N, $\hat{A}_{N-r} - \theta I_{N-r}$ *has no singular value in* $[0, \eta]$. *Finally, assume that* $\frac{m_2}{N-r} \mathrm{Tr}\left\{(\theta I_{N-r} - \hat{A}_{N-r})^{-1}(\theta I_{N-r} - \hat{A}_{N-r}^{\top})^{-1}\right\}$ *converges to* $\psi \in \mathbb{C}$. *We set* $\varphi = \int \lambda^{-1} d\nu_\theta(\lambda)$.

Then, for any $0 < \delta < \eta$, *almost surely for all large* N *there are exactly* r *eigenvalues* λ_i, $i = 1, \ldots, r$ *of* M_N *in* $B(\theta, \delta)$. *Moreover, the point process of* $\left(\sqrt{N}(\lambda_1 - \theta_N), \ldots, \sqrt{N}(\lambda_r - \theta_N)\right)$ *converges in distribution towards the point process of the eigenvalues of a* $r \times r$ *matrix* V *defined as*

$$V = \sigma\,(X_r + G)\,, \tag{5.12}$$

where X_r *is independent of* G, *a* $r \times r$ *Ginibre matrix whose entries are independent copies of a centered complex Gaussian variable* Z *whose covariance is characterized by,*

$$\mathbb{E}|Z|^2 = \frac{\sigma^2 \varphi}{1 - \sigma^2 \varphi} \quad \text{and} \quad \mathbb{E}Z^2 = \frac{\sigma^2 m_2 \psi}{1 - \sigma^2 \psi}.$$

Theorem 41 shows that the fluctuation of stable outliers are not universal (they may depend on the law of entries). We previously pointed out a similar phenomenon for deformed Wigner matrices in Section 2.4.2. Figure 5.6 illustrates Theorem 41.

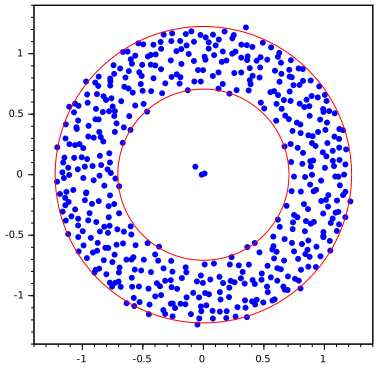


Figure 5.6: Eigenvalues of M_N for $A_N = \begin{pmatrix} C_{N-5} & 0 \\ 0 & D_5 \end{pmatrix}$, $D_5 = \mathrm{diag}(0; 0, \mathrm{i}/3; 1; 2)$ with $\sigma^2 = 1/2$, X_N has complex Gaussian entries, $N = 500$ and $r = 3$.

It was recently discovered by Benaych-Georges and Rochet [31] in the related study of the outliers in the single ring theorem that the fluctuations can be larger than $1/\sqrt{N}$ when the Jordan decomposition of an eigenvalue is not a diagonal matrix. Inspired by their work, we have also studied the fluctuations when for some integer $r \geq 2$ and complex number θ_N,

$$A_N = \left(\begin{array}{c|c} \hat{A}_r & 0 \\ \hline 0 & \hat{A}_{N-r} \end{array}\right), \text{ with} \qquad \hat{A}_r = P_N J_N P_N^{-1}, \quad P_N \in \mathrm{GL}_r(\mathbb{C}), \tag{5.13}$$

and $J_N \in M_r(\mathbb{C})$ is the Jordan matrix

$$J_N = \begin{pmatrix} \theta_N & 1 & & \\ & \theta_N & 1 & \\ & & \ddots & \ddots \\ & & & \end{pmatrix}.$$

Theorem 42. *Suppose that assumptions (X1-X2) and assumptions (A1-A3) hold with A_N given by (5.13). We suppose further that θ_N converges toward $\theta \in \mathbb{C} \setminus \mathrm{supp}(\beta)$ when N goes to infinity and that for some $\eta > 0$ and all large N, $\hat{A}_{N-r} - \theta I_{N-r}$ has no singular value in $[0, \eta]$. Assume finally that $\|P_N - P\| \to 0$ for some $P \in \mathrm{GL}_r(\mathbb{C})$, and that either $\mathbb{E}X_{11}^2 = 0$ or that $\frac{1}{N-r}\mathrm{Tr}\left\{(\theta I_{N-r} - \hat{A}_{N-r})^{-1}(\theta I_{N-r} - \hat{A}_{N-r}^{\top})^{-1}\right\}$ converges to $\psi \in \mathbb{C}$ (in the first case, we set $\psi = 0$).*

Then, for any $0 < \delta < \eta$, almost surely for all large N there are exactly r eigenvalues λ_i, $i = 1, \ldots, r$ of M_N in $B(\theta, \delta)$. Moreover, the point process of $\left(N^{1/(2r)}(\lambda_1 - \theta_N), \ldots, N^{1/(2r)}(\lambda_r - \theta_N)\right)$ converges in distribution towards the point process of the roots of the random polynomial

$$z^r - e_r^* P^{-1} V P e_1,$$

where V is defined by (5.12).

When $\hat{A}_{N-r} = 0$ and X_N is a complex Ginibre matrix, the above result is contained in [31, Theorem 2.6]. This result shows the strong correlation of the outlier eigenvalues in the setting of Theorem 42: properly rescaled they are asymptotically the r-th roots of the same random complex number. Figure 5.7 illustrates Theorem 42.

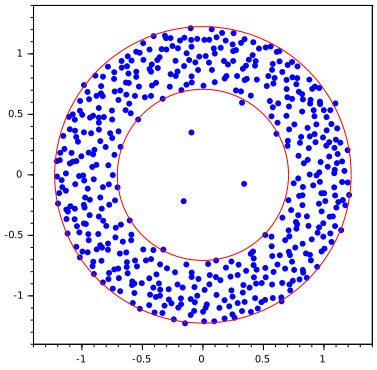


Figure 5.7: Eigenvalues of M_N for $A_N = F_N = \begin{pmatrix} C_{N-3} & & (0) & \\ & & 0 & 1 & 0 \\ (0) & & 0 & 0 & 1 \\ & & 0 & 0 & 0 \end{pmatrix}$, $N = 500$, $\sigma^2 = 1/2$ we have $N^{1/(2r)} = N^{1/6} \simeq 0.35$.

Note that some authors investigated other non-Hermitian models. Let us mention some works among others. In [66], Feldheim, Paquette and Zeitouni have recently studied the model (5.2) when σ decays polynomially of N and A_N is a block diagonal matrix with blocks of size $\log N$. Rochet [122] consider

the sum of a Hermitian random matrix and a finite rank matrix which is not necessarily Hermitian and O' Rourke and Renfrew [108] and Benaych-Georges, Cébron and Rochet [27] investigate finite rank perturbation of elliptic matrices.

Appendix : Free probability theory

We refer to [105], [107] and [151] for an introduction to free probability theory, created by Voiculescu. We only recall here the definitions and properties required to explain how free probability acts for our issues in random matrix 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 $(a_i)_{i=1,\dots,q}$ is a family of non-commutative random variables in $(\mathcal{A}, \phi)$, the distribution $\mu_{(a_i)_{i=1,\dots,q}}$ of $(a_i)_{i=1,\dots,q}$ is the linear functional on the algebra $\mathbb{C}\langle X_i | i = 1, \dots, q\rangle$ of polynomials in the non-commutating variables $(X_i)_{i=1,\dots,q}$ given by

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

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 $(\mathcal{A}_i)_{i=1,\dots q}$ in $(\mathcal{A}, \phi)$ is freely independent if for every $p \geq 1$, for every $(a_1, \dots, a_p)$ such that, for every k in $\{1, \dots, p\}$, $\phi(a_k) = 0$ and a_k is in $\mathcal{A}_{i(k)}$ for some $i(k)$ in $\{1, \dots, q\}$ with $i(k) \neq i(k+1)$, then $\phi(a_1, \dots, a_p) = 0$. 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 $(a_i^n)_{i=1,\dots,q}$ be a family of noncommutative random variables in a noncommutative probability space $(\mathcal{A}_n, \phi_n)$. The sequence of joint distributions $\mu_{(a_i^n)_{i=1,\dots,q}}$ converges as n tends to $+\infty$, if there exists a distribution μ such that $\mu_{(a_i^n)_{i=1,\dots,q}}(P)$ converges to $\mu(P)$ as n tends to $+\infty$ for every P in $\mathbb{C}\langle X_i | i = 1, \dots, q\rangle$. μ is called the limit distribution of $(a_i^n)_{i=1,\dots,q}$. If $(a_i)_{i=1,\dots,q}$ is a family of noncommutative random variables with distribution μ, we also say that $(a_i^n)_{i=1,\dots,q}$ converges towards $(a_i)_{i=1,\dots,q}$. A family of noncommutative random variables $(a_i^n)_{i=1,\dots,q}$ is said to be asymptotically free as n tends to ∞ if it has a limit distribution μ and if $(X_1, \dots, X_q)$ are free in $(\mathbb{C}\langle X_i | i = 1, \dots, q\rangle, \mu)$.

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 [32, 99, 144, 145] for the definitions and main properties of free convolutions. In the following sections, we briefly recall the analytic approach developed in [144, 145] 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 [26]. Then, we present the fundamental analytic subordination properties [21, 34, 147, 150] of these three convolutions. We also recall asymptotic freeness of some random matricial models [146]. Finally, we present the more general context of operator-valued free probability [148], [150],[23].

5.4 Free convolution

Free additive convolution

The Stieltjes transform g_μ 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 [144] 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.

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

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

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

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 [145, 32] 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.

Rectangular free convolution

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

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

$$\text{and} \quad T^{(c)}(z) = (cz+1)(z+1).$$

The transform $C_\tau^{(c)}$ [26] 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,

$$C^{(c)}_{\mu \boxplus_c \nu}(z) = C^{(c)}_\mu(z) + C^{(c)}_\nu(z).$$

5.5 Analytic subordinations

The analytic subordination phenomenon for free convolutions was first noted by Voiculescu in [147] for free additive convolution of compactly supported probability measures. Biane [34] 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 [22]. Note that such a subordination property allows to give a new definition of free additive convolution [46]. Finally, Belinschi, Benaych-Georges, and Guionnet [21] established such a phenomenon for the rectangular free convolution.

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. $$\omega_1(z) + \omega_2(z) - z = J_\mu(\omega_1(z)) = J_\nu(\omega_2(z)) = J_{\mu \boxplus \nu}(z), \quad z \in \mathbb{C}^+. \tag{5.14}$$

3. In particular (see [22]), 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 (5.14), we have for any $z \in \mathbb{C}^+$,

$$g_{\mu \boxplus \nu}(z) = g_\mu(\omega_1(z)) = g_\nu(\omega_2(z)). \tag{5.15}$$

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. $$\frac{F_1(z)F_2(z)}{z} = \eta_\mu(F_1(z)) = \eta_\nu(\omega_2(z)) = \eta_{\mu \boxtimes \nu}(z), \quad z \in \mathbb{C} \setminus [0, +\infty). \tag{5.16}$$

3. In particular (see [22]), 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 (5.16) yields

$$\psi_{\mu \boxtimes \nu}(z) = \psi_\mu(F_1(z)) = \psi_\nu(F_2(z)). \tag{5.17}$$

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 $\mathbb{C} \setminus \mathbb{R}^+$; 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 functions ω_i, F_i and Ω_i in the three previous subsections are called subordination functions.

5.6 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 $\mathcal{A}_N$ be the algebra of $N \times N$ random matrices $(\Omega, \mathcal{F}, P) \to M_N(\mathbb{C})$. Define

$$\Phi_N : \left\{ \begin{array}{l} \mathcal{A}_N \to \mathbb{C} \\ A \to \frac{1}{N} \operatorname{Tr} A \end{array} \right.$$

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 [146], Voiculescu shows that independent Gaussian Wigner matrices converge in distribution as their size goes to infinity to free semi-circular variables. The following result deals with matrices in generic position.

Theorem 43. *[Corollary 5.4.11 [4]] 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 the diagonal matrices $D_N(1)$ 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, when 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, Benaych-Georges [26] 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 44. *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 $\frac{1}{N} \sum_{s \text{ sing. val. of } A_N} \delta_s \to \mu$, $\frac{1}{N} \sum_{s \text{ sing. val. of } B_N} \delta_s \to \nu$ 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.

5.7 Operator-valued free probability theory

There exists an extension, operator-valued free probability theory, which still shares the basic properties of free probability but is much more powerful because of its wider domain of applicability. The concept of freeness with amalgamation and some of the relevant analytic transforms were introduced by Voiculescu in [148].

Definition 3. *Let $\mathcal{M}$ be an algebra and $\mathcal{B} \subset \mathcal{M}$ be a unital subalgebra. A linear map $E : \mathcal{M} \to \mathcal{B}$ is a conditional expectation if $E(b) = b$ for all $b \in \mathcal{B}$ and $E(b_1 a b_2) = b_1 E(a) b_2$ for all $a \in \mathcal{M}$ and b_1, b_2 in $\mathcal{B}$. Then $(\mathcal{M}, E)$ is called a $\mathcal{B}$-valued probability space. If in addition $\mathcal{M}$ is a C^*-algebra, $\mathcal{B}$ is a C^*-algebra of $\mathcal{M}$ and E is completely positive, then we have a $\mathcal{B}$-valued C^*-probability space.*

Example: Let $(\mathcal{A}, \phi)$ be a non commutative probability space. Define

$$M_2(\mathcal{A}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, a, b, c, d \in \mathcal{A} \right\}, \quad E := id_2 \otimes \phi \quad \text{that is}$$

$$E\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = \begin{pmatrix} \phi(a) & \phi(b) \\ \phi(c) & \phi(d) \end{pmatrix}.$$

$(M_2(\mathcal{A}), E)$ is an $M_2(\mathbb{C})$-valued probability space. $(\mathbb{C} \approx \mathbb{C}1_{\mathcal{A}})$

As in scalar-valued free probability, one defines [148] *freeness with amalgamation* over $\mathcal{B}$ via an algebraic relation similar to freeness, but involving E and noncommutative polynomials with coefficients in $\mathcal{B}$.

Definition 4. *Consider an operator-valued probability space $(\mathcal{M}, E : \mathcal{M} \to \mathcal{B})$.*
The $\mathcal{B}$-valued distribution of $a \in \mathcal{M}$ is given by all $\mathcal{B}$-valued moments $E(ab_1 a b_2 \cdots b_{n-1} a) \in \mathcal{B}$, $n \in \mathbb{N}, b_0, \ldots, b_{n-1} \in \mathcal{B}$.
Let $(A_i)_{i \in I}$ be a family of subalgebras with $\mathcal{B} \subset A_i$ for all $i \in I$. The subalgebras $(A_i)_{i \in I}$ are free with respect to E or free with amalgamation over B if $E(a_1 \cdots a_n) = 0$ whenever $a_j \in A_{i_j}$, $i_j \in I$, $E(a_j) = 0$, for all j and $i_1 \neq i_2 \neq \cdots \neq i_n$.

Random variables in $\mathcal{M}$ or subsets of $\mathcal{M}$ are free with amalgamation over $\mathcal{B}$ if the algebras generated by $\mathcal{B}$ and the variables or the algebras generated by $\mathcal{B}$ and the subsets, respectively, are so. Note that the subalgebra generated by $\mathcal{B}$ and some variable a is not just the linear span of monomials of the form ba^n, but, because elements from $\mathcal{B}$ and our variable a do not commute in general, we must also consider general monomials of the form $b_0 a b_1 a b_2 \cdots b_{n-1} a b_n$.

The previous results of free subordination property in the scalar case are approached from an abstract coalgebra point of view by Voiculescu in [150] and this approach extends the results to the $\mathcal{B}$-valued case. In [23], Belinschi, Mai and Speicher develop an analytic theory. In order to describe operator-valued subordination property, we need some notation. If $\mathcal{A}$ is a unital C^*-algebra and $b \in \mathcal{A}$, we denote by $\Re b = (b + b^*)/2$ and $\Im b = (b - b^*)/2i$ the real and imaginary parts of b, so $b = \Re b + i \Im b$. For a selfadjoint operator $b \in \mathcal{A}$, we write $b \geq 0$ if the spectrum of b is contained in $[0, +\infty)$ and $b > 0$ if the spectrum of b is contained in $(0, +\infty)$. The operator upper half-plane of $\mathcal{A}$ is the set $\mathbb{H}^+(\mathcal{A}) = \{b \in \mathcal{A} \colon \Im b > 0\}$.

Theorem 45. *[150],[23] Let $(\mathcal{M}, E : \mathcal{M} \to \mathcal{B})$ be an operator-valued C^* -probability space. Let $x_1, x_2 \in \mathcal{M}$ be selfadjoint variables which are free with amalgamation over $\mathcal{B}$.*
There exist an analytic map $\omega \colon \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B})$ such that

$$E\left[(\omega(b) - x_1)^{-1}\right] = E\left[(b - (x_1 + x_2))^{-1}\right] \quad \text{for all } b \in \mathbb{H}^+(\mathcal{B}).$$

Moreover, if $b \in \mathbb{H}^+(\mathcal{B})$, then $\omega(b)$ is the unique fixed point of the map

$$f_b : \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B}), \quad f_\beta(w) = h_{x_2}(h_{x_1}(w) + b) + b$$

$$\textit{where} \quad h_{x_i}(b) = E\left[(b - x_i)^{-1}\right]^{-1} - b$$

$$\textit{and} \quad \omega(b) = \lim_{k \to +\infty} f_b^{\circ k}(w), \quad \textit{for any} \quad w \in \mathbb{H}^+(\mathcal{B}).$$

Maps of the form $b \mapsto E\left[(b - X)^{-1}\right]$ for some selfadjoint X and (conditional) expectation E are also known as *operator-valued Cauchy-Stieltjes transforms.*

In our applications, the algebra $\mathcal{B}$ is $M_n(\mathbb{C})$ for some $n \in \mathbb{N}$. The following result from [106] explains why this case is relevant in our work using linearizations of polynomials.

Proposition 6. *Let $(\mathcal{A}, \phi)$ be a C^*-probability space, let m be a positive integer, and let $x_1, \ldots, x_n \in \mathcal{A}$ be freely independent. Then the map $id_m \otimes \phi \colon M_m(\mathcal{A}) \to M_m(\mathbb{C})$ is a unit preserving conditional expectation, and $\alpha_1 \otimes x_1, \ldots, \alpha_n \otimes x_n$ are free over $M_m(\mathbb{C})$ for any $\alpha_i \in M_m(\mathbb{C})$.*

Bibliography

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

[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, 1261–1292, 2007.

[3] G. W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.*, 41(3B):2103–2181, 2013.

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

[5] 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, 367–389, 2005.

[6] O. Arizmendi, I. Nechita, and C. Vargas. On the asymptotic distribution of block-modified random matrices. *J. Math. Phys.*, 57(1):015216, 2016.

[7] Z. D. Bai. Circular law. *Ann. Probab.* 25, 494–529, 1997.

[8] Z. D. Bai. Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica* 9, 611-677, 1999.

[9] 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, 316–345, 1998.

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

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

[12] 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, no. 1, 1150004, 44 pp, 2012.

[13] 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), 166–168, 1988.

[14] Z.D. Bai and J.F. Yao, Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. H. Poincaré* 44 (3), 447–474, 2008.

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

[16] 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), 1729–1741, 1988.

[17] 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, 1643–1697, 2005.

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

[19] Z. Bao, G. Pan, W. Zhou. Universality for the Largest Eigenvalue of Sample Covariance Matrices with General Population. *Ann. Statist.* 43.1, 382–421, 2015.

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

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

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

[23] 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.*, ISSN (Online) 1435-5345, ISSN (Print) 0075-4102, DOI: 10.1515/crelle-2014-0138, 2015.

[24] S. T. Belinschi and D. Shlyakhtenko. Free probability of type B: analytic interpretation and applications. *American Journal of Mathematics*, 134, No.1, 193–234, 2012.

[25] 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

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

[27] F. Benaych-Georges, G. Cébron and J. Rochet. Fluctuation of matrix entries and application to outliers of elliptic matrices, ArXiv:1602.02929.

[28] 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, 1621–1662, 2011.

[29] 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, 494–521, 2011.

[30] 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, 120–135, 2012.

[31] F. Benaych-Georges and J. Rochet. Outliers in the single ring theorem. *Probab. Theory Related Fields* Vol. 165, no. 1, 313–363, 2016.

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

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

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

[35] P. Biane and F. Lehner. Computation of some examples of Brown's spectral measure in free probability. *Colloq. Math.*, 90(2):181–211, 2001.

[36] G. Biroli, J.P. Bouchaud and M. Potters. On the top eigenvalue of heavy-tailed random matrices, *Europhys. Lett.* 78, 10001, 2007.

[37] 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.

[38] C. Bordenave and D. Chafaï. Around the circular law. *Probab. Surv.*, 9:1–89, 2012.

[39] C. Bordenave, P. Caputo and D. Chafaï. Spectrum of markov generators on sparse random graphs. *Comm. Pure Appl. Math.* 67(4):621–669, 2014.

[40] A. Bloemendal, L. Erdös, A. Knowles, H.-T. Yau, and J. Yin. Isotropic local laws for sample covariance and generalized wigner matrices. *Electron. J. Probab*, 19(33):1–53, 2014.

[41] A. Borodin and P.J. Forrester. Increasing subsequences and the hard-to-soft edge transition in matrix ensembles. *J. Phys.* A, 36 (12): 2963-2981, Random matrix theory, 2003.

[42] A. Bloemendal and Virag. Limits of spiked random matrices I. *Probab. Theory and Related Fields* 156, 3-4 , 795–825, 2013.

[43] A. Bloemendal and Virag. Limits of spiked random matrices II. *Ann. of Probab.*, Vol. 44, No. 4, 2726–2769, 2016.

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

[45] C. Cacciapuoti, A. Maltsevb, and B. Schlein. Local Marchenko-Pastur law at the hard edge of sample covariance matrices. *J. Math. Phys.* 54, 043302, 2013.

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

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

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

[49] B. Collins, J. Mingo, P. Sniady and R. Speicher. Second Order Freeness and Fluctuations of Random Matrices III. Higher order freeness and free cumulants. *Documenta Math.* 12: 1–70, 2007.

[50] R. Couillet and M. Debbah. *Random Matrix Methods for Wireless Communications.* Cambridge University Press, 2011.

[51] R. Couillet, J. W. Silverstein, Z. Bai and M. Debbah. Eigen-inference for energy estimation of multiple sources, *IEEE Trans. Inform. Theory*, 57, 4 (2011):2420–2439.

[52] P. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*. volume 3 of Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.

[53] P. Deift and D. Gioev. Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices. *Comm. Pure Appl. Math.* 60, no. 6, 867–910, 2007.

[54] P. Deift, T. Kriecherbauer, K. McLaughlin, S. Venakides and X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* 52, no. 12, 1335–1425, 1999.

[55] X. Ding and F. Yang A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices. Arxiv 1607.06873, 2017.

[56] 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, 678–694, 2007.

[57] 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, 1099–1122, 2007.

[58] A. Edelman and E. Kostlan. How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc.* (N.S.), 32(1):1–37, 1995.

[59] A. Edelman and N. R. Rao. Random matrix theory. *Acta Numer.*, 14:233–297, 2005.

[60] A. Edelman and B. D. Sutton. From random matrices to stochastic operators. *J. Stat. Phys.* 127, Issue 6: 1121-1165, 2007.

[61] N. El Karoui. On the largest eigenvalue of Wishart matrices with identity covariance when n, p and p=n tend to infinity. ArXiv math.ST/0309355, 2003.

[62] L. Erdös. Universality of Wigner random matrices: a Survey of Recent Results. *Russian Mathematical Surveys*, Volume 66, Number 3, 2011.

[63] L. Erdös, B. Schlein, B. and H.T Yau. Local semicicle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* 287(2), 641–655, 2009.

[64] L. Erdös, B. Schlein and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.*, 37(3):815–852, 2009.

[65] L. Erdös, H. Yau. *Dynamical approach to random matrix theory* (book in preparation).

[66] O. Feldheim, E. Paquette and O. Zeitouni. Regularization of non-normal matrices by Gaussian noise, *Internat. Math. Res. Notices* (18): 8724–8751, 2015.

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

[68] P.J Forrester. *Log-Gases and Random Matrices*. London Mathematical Society Monographs (LMS-34).

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

[70] J. Garnier and K. Solna. Applications of random matrix theory for sensor array imaging with measurement noise. *Random Matrices MSRI Publication*, Volume 65, 2014.

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

[72] J. Ginibre, Statistical Ensembles of Complex, Quaternion, and Real Matrices, *J. Math. Phys.* 6, 440–449, 1965.

[73] V. L. Girko. *Theory of stochastic canonical equations.* Vol. I, volume 535 of Mathematics and its Applications. Kluwer Academic Publishers, Dordrecht, 2001.

[74] F. Götze, A.N. Tikhomirov. The Circular Law for Random Matrices. *Ann. Probab.* 38, no. 4, 1444–1491, 2010.

[75] A. Guionnet, M. Krishnapur, and O. Zeitouni. The single ring theorem. *Ann. of Math.* (2), 174(2):1189–1217, 2011.

[76] A. Guionnet and O. Zeitouni. Support convergence in the single ring theorem. *Probab. Theory Related Fields*, 154(3-4):661–675, 2012.

[77] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $Ext(C^*_{\mathrm{red}}(\mathbb{F}_2))$ is not a group. *Ann. of Math.* (2), 162(2):711–775, 2005.

[78] U. Haagerup, H. Schultz, and S. Thorbjornsen. A random matrix approach to the lack of projections in $C^*_{\mathrm{red}}(\mathbb{F}_2)$. *Adv. Math.*, 204(1):183, 2006.

[79] W. Hachem, A. Hardy and J. Najim. Large Complex Correlated Wishart Matrices: Fluctuations and Asymptotic Independence at the Edges. *Ann. Probab.* Volume 44, Number 3, 2264-2348, 2016.

[80] W. Hachem, A. Hardy and J. Najim. Large Complex Correlated Wishart Matrices: The Pearcey Kernel and Expansion at the Hard Edge. *Electron. J. Probab.* Volume 21, Number (2016), paper no. 1, 36 pp.

[81] W. Hachem, A. Hardy and J. Najim. A Survey on the Eigenvalues Local Behavior of Large Complex Correlated Wishart Matrices. *ESAIM: Proceedings and surveys*, Vol. 51, p. 150–174, 2015.

[82] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. *Ann. Appl. Probab.*, 17(3):875930, 2007.

[83] U. Haagerup and F. Larsen. Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras. *J. Funct. Anal.*, 176(2):331–367, 2000.

[84] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.* 209, 437–476, 2000.

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

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

[87] A. M. Khorunzhy, B.A. Khoruzhenko and L.A. Pastur. Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.* 37 no. 10, 50335060, 1996.

[88] A. Knowles and J. Yin. Eigenvector distribution of Wigner matrices. *Probab. Theory and Related Fields.* 155(3-4), 543–582, 2013.

[89] A. Knowles and J. Yin, J.. The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.*, Volume 66, Issue 11, 16631749, 2013.

[90] A. Knowles and J. Yin. The outliers of a deformed Wigner matrix. *Ann. of Probab.* Vol. 42, No. 5, 1980–2031, 2014.

[91] A. Knowles and J. Yin. Anisotropic local laws for random matrices. *Probab. Theory and Related Fields.* First Online: DOI: 10.1007/s00440-016-0730-4, 2016.

[92] 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.

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

[94] J. O. Lee and K. Schnelli. Edge Universality for Deformed Wigner Matrices. *Rev. Math. Phys.* 27, 1550018, 2015. https://doi.org/10.1142/S0129055X1550018X.

[95] J. O. Lee and K. Schnelli. Tracy-Widom Distribution for the Largest eigenvalue of real sample covariance matrices with general population. *Ann. Appl. Probab.* Volume 26, Number 6, 3786-3839, 2016.

[96] J. O. Lee and J. Yin. A necessary and sufficient condition for edge universality of Wigner matrices. *Duke Math. J.* Volume 163, Number 1, 117–173, 2014.

[97] 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, 1043–1068, 2012.

[98] 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, 1934–1959, 2011.

[99] H. Maassen. Addition of freely independent random variables. *J. Funct. Anal.*, 106(2), 409–438, 1992.

[100] C. Male. The norm of polynomials in large random and deterministic matrices. With an appendix by Dimitri Shlyakhtenko. *Probab. Theory Related Fields* 154, no. 3-4 (2012), 477–532, 2012.

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

[102] M.L. Mehta. *Random Matrices and the Statistical Theory of Energy Levels*, Academic Press, New York, NY, 1967.

[103] J. Mingo, P. Sniady and R. Speicher. Second order freeness and fluctuations of random matrices: II. Unitary random matrices. *Adv. in Math.* 209: 212240, 2007.

[104] J. Mingo and R. Speicher. Second Order Freeness and Fluctuations of Random Matrices: I. Gaussian and Wishart matrices and Cyclic Fock spaces. *J. Funct. Anal.*, 235, pp. 226270, 2006.

[105] J.A. Mingo and R. Speicher. *Free probability and Random matrices.* Fields Institute Monographs Originally published by American Mathematical Society ISSN: 1069-5273. Springer.

[106] A. Nica, D. Shlyakhtenko, and R. Speicher, *Operator-Valued Distributions. I. Characterizations of Freeness.* Internat. Math. Res. Notices, **29**, 1509–1538, 2002.

[107] A. Nica and R. Speicher. *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, 2006.

[108] S. O'Rourke and D. Renfrew. Low rank perturbations of large elliptic random matrices. *Electron. J. Probab.* 19(43), 2013.

[109] O' Rourke, S., Vu, V., and Wang, K. Eigenvectors of random matrices: a survey. *J. Combin. Theory* Ser. A 144, 361–442, 2016.

[110] G. Pan and W. Zhou, Circular law, Extreme singular values and potential theory. *J. Multivariate Anal.* 101, no. 3, 645–656, 2010.

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

[112] L. Pastur and and M. Shcherbina. *Eigenvalue Distribution of Large Random Matrices.* Mathematical Surveys and Monographs, Volume 171, American Mathematical Society.

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

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

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

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

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

[118] K. Rajan and L. F. Abbott. Eigenvalue spectra of random matrices for neural networks. *Phys. Rev. Lett.*, 97:188104, Nov 2006.

[119] J. A. Ramírez, B. Rider and B. Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.* 24 919–944, 2011.

[120] 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), 468–480, 2010.

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

[122] J. Rochet. Complex outliers of Hermitian random matrices. *J. Theor. Probab.* doi:10.1007/s10959-016-0686-4, 2016.

[123] M. Rudelson and R. Vershynin. Invertibility of random matrices: Unitary and orthogonal perturbations. *J. Amer. Math. Soc.*, 27(2):293–338, 2014.

[124] E.B. Saff and V. Totik. *Logarithmic Potentials with External Fields*. Series: Grundlehren der mathematischen Wissenschaften, Vol. 316. Springer, 1997.

[125] H. Schultz. Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases., *Prob. Th. Rel. Fields* **131**, 261–309, 2005.

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

[127] K. Shen J. Yao and W. K. Li. On a spiked model for large volatility matrix estimation from noisy high-frequency data. arxiv 1702.03417, 2017.

[128] T. Shirai. Limit theorems for random analytic functions and their zeros. In *Functions in number theory and their probabilistic aspects*, RIMS Kôkyûroku Bessatsu, B34, pages 335–359. Res. Inst. Math. Sci. (RIMS), Kyoto, 2012.

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

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

[131] P. Sniady. Random regularization of Brown spectral measure. *J. Funct. Anal.*, 193(2):291–313, 2002.

[132] A. Soshnikov. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* 207, 697–733.

[133] A. Soshnikov, A. A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *J. Stat. Phys.* 108, no. 5-6, 1033–1056, 2002.

[134] R. Speicher and C. Vargas. Free deterministic equivalents, rectangular random matrix models, and operator-valued free probability theory. *Random Matrices: Theory and Applications*, 1(02):1150008, 2012.

[135] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms. In Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), pages 597–606. Higher Ed. Press, Beijing, 2002.

[136] B.D. Sutton. The stochastic Operator approach to Random matrix theory, PhD thesis (2005), Massachusetts Institute of Technology.

[137] T. Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. *Probab. Theory Related Fields*, 155(1-2):231–263, 2013.

[138] T. Tao and V. Vu. Random matrices: the circular law. *Commun. Contemp. Math.* 10, 261–307, 2008.

[139] T. Tao and V. Vu. From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices. *Bull. Amer. Math. Soc.* (N.S.), 46(3):377–396, 2009.

[140] T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, 38(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur.

[141] T. Tao and V. Vu. Random matrices: universal properties of eigenvectors. *Random Matrices Theory Appl.* 1(1), 1150,001, 27, 2012.

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

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

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

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

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

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

[148] D. Voiculescu. Operations on certain non-commutative operator-valued random variables. Recent advances in operator algebras (Orléans, 1992). *Astérisque* (1995), no. 232, 243–275, 1995.

[149] D. Voiculescu., A strengthened asymptotic freeness result for random matrices with applications to free entropy, *Internat. Math. Res. Notices*, no. 1, 41–63, 1998.

[150] D. Voiculescu.,The coalgebra of the free difference quotient and free probability, Internat. Math. Res. Notices, no. 2, 79–106, 2000.

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

[152] J. von Neumann and H. H. Goldstine. Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, 53:1021–1099, 1947.

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

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

[155] J.-s. Xie. The convergence on spectrum of sample covariance matrices for information-plus-noise type data. *Appl. Math. J. Chinese Univ.* Ser. B, 27(2):181–191, 2012.

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