The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and non-universality of the fluctuations

M. Capitaine*, C. Donati-Martin† and D. Féral‡

Abstract

In this paper, we investigate the asymptotic spectrum of complex or real Deformed Wigner matrices $(M_N)_N$ defined by $M_N = W_N/\sqrt{N} + A_N$ where W_N is a $N \times N$ Hermitian (resp. symmetric) Wigner matrix whose entries have a symmetric law satisfying a Poincaré inequality. The matrix A_N is Hermitian (resp. symmetric) and deterministic with all but finitely many eigenvalues equal to zero. We first show that, as soon as the first largest or last smallest eigenvalues of A_N are sufficiently far from zero, the corresponding eigenvalues of M_N almost surely exit the limiting semicircle compact support as the size N becomes large. The corresponding limits are universal in the sense that they only involve the variance of the entries of W_N. On the other hand, when A_N is diagonal with a sole simple non-null (fixed) eigenvalue large enough, we prove that the fluctuations of the largest eigenvalue are not universal and vary with the particular distribution of the entries of W_N.

1 Introduction

This paper lies in the lineage of recent works studying the influence of some perturbations on the asymptotic spectrum of classical random matrix models. Such questions come from Statistics (cf. [Jo]) and appeared in the framework of empirical covariance matrices, also called non-white Wishart matrices or spiked population models, considered by J. Baik, G. Ben Arous and S. Péché [Bk-B-P] and by J. Baik and J. Silverstein [Bk-S1]. The work [Bk-B-P] deals with random sample covariance matrices $(S_N)_N$ defined by

$$S_N = \frac{1}{N} Y_N^* Y_N \tag{1.1}$$

where Y_N is a $p \times N$ complex matrix whose sample column vectors are i.i.d, centered, Gaussian and of covariance matrix a deterministic Hermitian matrix Σ_p having all but finitely many eigenvalues equal to one. Besides, the size of the samples N and the size of the population $p = p_N$ are assumed of the same order (as $N \to \infty$). The authors of [Bk-B-P] first noticed that, as in the classical case (known as the Wishart model) where $\Sigma_p = I_p$ is the identity matrix, the global limiting behavior of the spectrum of S_N is not affected by the matrix Σ_p. Thus, the limiting spectral measure is the well-known Marchenko-Pastur law. On the other hand, they pointed out a phase transition phenomenon for the fluctuations of the largest eigenvalue according to the value of the largest eigenvalue(s) of Σ_p. The approach of [Bk-B-P] does not extend to the real Gaussian setting and the whole analog of their result is still an open question. Nevertheless, D. Paul was able to establish in [P] the Gaussian fluctuations of the largest eigenvalue of the real Gaussian matrix S_N when the largest eigenvalue of Σ_p is simple and sufficiently larger than one. More recently, J. Baik and J. Silverstein investigated in [Bk-S1] the almost sure limiting behavior of the extremal eigenvalues of complex or real non necessarily Gaussian

*Institut de Mathématiques de Toulouse, Equipe de Statistique et Probabilités, F-31062 Toulouse Cedex 09. E-mail: capitain@math.ups-tlse.fr

†Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 6, Site Chevaleret, 13 rue Clisson, F-75013 Paris. E-mail: donati@ccr.jussieu.fr

‡Institut de Mathématiques de Toulouse, Equipe de Statistique et Probabilités, F-31062 Toulouse Cedex 09. E-mail: dferal@math.ups-tlse.fr

matrices. Under assumptions on the first four moments of the entries of Y_N, they showed in particular that when exactly k eigenvalues of Σ_p are far from one, the k first eigenvalues of S_N are almost surely outside the limiting Marchenko-Pastur support. Fluctuations of the eigenvalues that jump are universal and have been recently found by Z. Bai and J. F. Yao in [B-Ya2] (we refer the reader to [B-Ya2] for the precise restrictions made on the definition of the covariance matrix Σ_p). Note that the problem of the fluctuations in the very general setting of [Bk-S1] is still open.

Our purpose here is to investigate the asymptotic behavior of the first extremal eigenvalues of some complex or real Deformed Wigner matrices. These models can be seen as the additive analogue of the spiked population models and are defined by a sequence $(M_N)_N$ given by

$$M_N = \frac{1}{\sqrt{N}} W_N + A_N := X_N + A_N \tag{1.2}$$

where W_N is a Wigner matrix such that the common distribution of its entries satisfied some technical conditions (given in (i) below) and A_N is a deterministic matrix of finite rank. We establish the analog of the main result of [Bk-S1] namely that, once A_N has exactly k (fixed) eigenvalues far enough from zero, the k first eigenvalues of M_N jump almost surely outside the limiting semicircle support. This result is universal (as the one of [Bk-S1]) since the corresponding almost sure limits only involve the variance of the entries of W_N. On the other hand, at the level of the fluctuations, we exhibit a striking phenomenon in the particular case where A_N is diagonal with a sole simple non-null eigenvalue large enough. Indeed, we find that in this case, the fluctuations of the largest eigenvalue of M_N are not universal and strongly depend on the particular law of the entries of W_N. More precisely, we prove that the limiting distribution of the (properly rescaled) largest eigenvalue of M_N is the convolution of the distribution of the entries of W_N with a Gaussian law. In particular, if the entries of W_N are not Gaussian, the fluctuations of the largest eigenvalue of M_N are not Gaussian.

In the following section, we first give the precise definition of the Deformed Wigner matrices (1.2) considered in this paper and we recall the known results on their asymptotic spectrum. Then, we present our results and sketch their proof. We also outline the organization of the paper.

2 Model and results

Throughout this paper, we consider complex or real Deformed Wigner matrices $(M_N)_N$ of the form (1.2) where the matrices W_N and A_N are defined as follows:

(i) W_N is a $N \times N$ Wigner Hermitian (resp. symmetric) matrix such that the N^2 random variables $(W_N)_{ii}$, $\sqrt{2}\Re e((W_N)_{ij})_{i<j}$, $\sqrt{2}\Im m((W_N)_{ij})_{i<j}$ (resp. the $\frac{N(N+1)}{2}$ random variables $\frac{1}{\sqrt{2}}(W_N)_{ii}$, $(W_N)_{ij}$, $i<j$) are independent identically distributed with a symmetric distribution μ of variance σ^2 and satisfying a Poincaré inequality (see Section 3).

(ii) A_N is a deterministic Hermitian (resp. symmetric) matrix of fixed finite rank r and built from a family of J fixed real numbers $\theta_1 > \cdots > \theta_J$ independent of N with some j_o such that $\theta_{j_o} = 0$. We assume that the non-null eigenvalues θ_j of A_N are of fixed multiplicity k_j (with $\sum_{j \neq j_o} k_j = r$) i.e. A_N is similar to the diagonal matrix

$$D_N = \operatorname{diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_{j_o-1}, \ldots, \theta_{j_o-1}}_{k_{j_o-1}}, \underbrace{0, \ldots\ldots, 0}_{N-r}, \underbrace{\theta_{j_o+1}, \ldots, \theta_{j_o+1}}_{k_{j_o+1}}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}). \tag{2.1}$$

Before going into details of the results, we want to point out that the condition made on μ (namely that μ satisfies a Poincaré inequality) is just a technical condition: we conjecture that our results still hold under weaker assumptions (see Remark 2.1 below). Nevertheless, a lot of measures satisfy a Poincaré inequality (we refer the reader to [B-G] for a characterization of such measures on $\mathbb{R}$, see also [A. et al]). For instance, consider $\mu(dx) = \exp(-|x|^\alpha)dx$ with $\alpha \geq 1$.
Futhermore, note that this condition implies that μ has moments of any order (cf. Corollary 3.2 and Proposition 1.10 in [L]).

Let us now introduce some notations. When the entries of W_N are further assumed to be Gaussian that is, in the complex (resp. real) setting when W_N is of the so-called GUE (resp. GOE), we will write W_N^G instead of W_N. Then $X_N^G := \frac{1}{\sqrt{N}} W_N^G$ will be said of the GU(O)E($N, \frac{\sigma^2}{N}$) and we will let $M_N^G = X_N^G + A_N$ be the corresponding Deformed GU(O)E model.
In the following, given an arbitrary Hermitian matrix B of order N, we will denote by $\lambda_1(B) \geq \cdots \geq \lambda_N(B)$ its N ordered eigenvalues and by $\mu_B = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i(B)}$ its empirical measure. For notational convenience, we will also define $\lambda_0(B) = +\infty$ and $\lambda_{N+1}(B) = -\infty$.

The Deformed Wigner model is built in such a way that the Wigner Theorem is still satisfied. Thus, as in the classical Wigner model ($A_N \equiv 0$), the spectral measure (μ_{M_N}) converges a.s. to the semicircle law μ_{sc} whose density is given by

$$\frac{d\mu_{sc}}{dx}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\,\mathbf{1}_{[-2\sigma,2\sigma]}(x). \tag{2.2}$$

This result follows from Lemma 2.2 of [B]. Note that it only relies on the two first moment assumptions on the entries of W_N and the fact that the A_N's are of finite rank.

On the other hand, the asymptotic behavior of the extremal eigenvalues may be affected by the perturbation A_N. Recently, S. Péché studied in [Pe] the Deformed GUE under a finite rank perturbation A_N defined by (ii). Following the method of [Bk-B-P], she highlighted the effects of the non-null eigenvalues of A_N at the level of the fluctuations of the largest eigenvalue of M_N^G. To explain this in more detail, let us recall that when $A_N \equiv 0$, it was established in [T-W] that as $N \to \infty$,

$$\sigma^{-1} N^{2/3}\Big(\lambda_1(X_N^G) - 2\sigma\Big) \xrightarrow{\mathcal{L}} F_2 \tag{2.3}$$

where F_2 is the well-known GUE Tracy-Widom distribution (see [T-W] for the precise definition). Dealing with the Deformed GUE M_N^G, it appears that this result is modified as soon as the first largest eigenvalue(s) of A_N are quite far from zero. In the particular case of a rank one perturbation A_N having a fixed non-null eigenvalue $\theta > 0$, [Pe] proved that the fluctuations of the largest eigenvalue of M_N^G are still given by (2.3) when θ is small enough and precisely when $\theta < \sigma$. The limiting law is changed when $\theta = \sigma$. As soon as $\theta > \sigma$, [Pe] established that the largest eigenvalue $\lambda_1(M_N^G)$ fluctuates around

$$\rho_\theta = \theta + \frac{\sigma^2}{\theta} \tag{2.4}$$

(which is $> 2\sigma$ since $\theta > \sigma$) as

$$\sqrt{N}\Big(\lambda_1(M_N^G) - \rho_\theta\Big) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\theta^2) \tag{2.5}$$

where

$$\sigma_\theta = \sigma\sqrt{1 - \frac{\sigma^2}{\theta^2}}. \tag{2.6}$$

Similar results are conjectured for the Deformed GOE but S. Péché emphasized that her approach fails in the real framework. Indeed, it is based on the explicit Fredholm determinantal representation for the distribution of the largest eigenvalue(s) that is specific to the complex setting. Nevertheless, M. Maïda [M] obtained a large deviation principle for the largest eigenvalue of the Deformed GOE M_N^G under a rank one deformation A_N; from this result she could deduce the almost sure limit with respect to the non-null eigenvalue of A_N. Thus, under a rank one perturbation A_N such that $D_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ where $\theta > 0$, [M] showed that

$$\lambda_1(M_N^G) \xrightarrow{a.s} \rho_\theta, \quad \text{if } \theta > \sigma \tag{2.7}$$

and

$$\lambda_1(M_N^G) \xrightarrow{a.s} 2\sigma, \quad \text{if } \theta \leq \sigma. \tag{2.8}$$

Note that the approach of [M] extends with minor modifications to the Deformed GUE. Following the investigations of [Bk-S1] in the context of general spiked population models, one can conjecture that such a phenomenon holds in a more general and non necessarily Gaussian setting. The first result of our paper, namely the following Theorem 2.1, is related to this question. Before being more explicit, let us recall that when $A_N \equiv 0$, the whole spectrum of the rescaled complex or real Wigner matrix $X_N = \frac{1}{\sqrt{N}} W_N$ belongs almost surely to the semicircle support $[-2\sigma, 2\sigma]$ as N goes to infinity and that (cf. [B-Yi] or Theorem 2.12 in [B])

$$\lambda_1(X_N) \xrightarrow{a.s} 2\sigma \quad \text{and} \quad \lambda_N(X_N) \xrightarrow{a.s} -2\sigma. \tag{2.9}$$

Note that this last result holds true in a more general setting than the one considered here (see [B-Yi] for details) and in particular only requires the finiteness of the fourth moment of the law μ. Moreover, one can readily extend the previous limits to the first extremal eigenvalues of X_N i.e.

$$\text{for any fixed } k \geq 1, \quad \lambda_k(X_N) \xrightarrow{a.s} 2\sigma \quad \text{and} \quad \lambda_{N-k}(X_N) \xrightarrow{a.s} -2\sigma. \tag{2.10}$$

Here, we prove that, under the assumptions (i)-(ii), (2.10) fails when some of the θ_j's are sufficiently far from zero: as soon as some of the first largest (resp. the last smallest) non-null eigenvalues θ_j's of A_N are taken strictly larger then σ (resp. strictly smaller than $-\sigma$), the same part of the spectrum of M_N almost surely exits the semicircle support $[-2\sigma, 2\sigma]$ as $N \to \infty$ and the new limits are the ρ_{θ_j}'s defined by

$$\rho_{\theta_j} = \theta_j + \frac{\sigma^2}{\theta_j}. \tag{2.11}$$

Observe that ρ_{θ_j} is $> 2\sigma$ (resp. $< -2\sigma$) when $\theta_j > \sigma$ (resp. $< -\sigma$) (and $\rho_{\theta_j} = \pm 2\sigma$ if $\theta_j = \pm\sigma$). Here is the precise formulation of our result. For definiteness, we set $k_1 + \cdots + k_{j-1} := 0$ if $j = 1$.

Theorem 2.1. *Let $J_{+\sigma}$ (resp. $J_{-\sigma}$) be the number of j's such that $\theta_j > \sigma$ (resp. $\theta_j < -\sigma$).*

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

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

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

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

Remark 2.1. *Let us notice that, following [Bk-S1], one can expect that this theorem holds true in a more general setting than the one considered here, namely only requires four first moment conditions on the law μ of the Wigner entries. As we will explain in the following, the assumption (i) that μ satisfies a Poincaré inequality is actually fundamental in our reasoning since we will need several variance estimates.*

This theorem will be proved in Section 4. The second part of this work is devoted to the study of the particular rank one diagonal deformation $A_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ such that $\theta > \sigma$. We investigate the fluctuations of the largest eigenvalue of any real or complex Deformed model M_N satisfying (i) around its limit ρ_θ (given by the previous theorem). We obtain the following result.

Theorem 2.2. *Let $A_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ and assume that $\theta > \sigma$. Define*

$$v_\theta = \frac{t}{4}\Big(\frac{m_4 - 3\sigma^4}{\theta^2}\Big) + \frac{t}{2}\frac{\sigma^4}{\theta^2 - \sigma^2} \tag{2.12}$$

where $t = 4$ (resp. $t = 2$) when W_N is real (resp. complex) and $m_4 := \int x^4 d\mu(x)$. Then

$$\sqrt{N}\Big(\lambda_1(M_N) - \rho_\theta\Big) \xrightarrow{\mathcal{L}} (1 - \frac{\sigma^2}{\theta^2})\Big\{\mu * \mathcal{N}(0, v_\theta)\Big\}. \tag{2.13}$$

Note that when $m_4 = 3\sigma^4$ as in the Gaussian case, the variance of the limiting distribution in (2.13) is equal to σ_θ^2 (resp. $2\sigma_\theta^2$) in the complex (resp. real) setting (with σ_θ given by (2.6)).

Remark 2.2. *Since μ is symmetric, it readily follows from Theorem 2.2 that when $A_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ and $\theta < -\sigma$, the smallest eigenvalue of M_N fluctuates as*

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

In particular, one derives the analog of (2.5) for the Deformed GOE that is

Theorem 2.3. *Let A_N be an arbitrary deterministic symmetric matrix of rank one having a non-null eigenvalue θ such that $\theta > \sigma$. Then the largest eigenvalue of the Deformed GOE fluctuates as*

$$\sqrt{N}\Big(\lambda_1(M_N^G) - \rho_\theta\Big) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\sigma_\theta^2). \tag{2.14}$$

Obviously, thanks to the orthogonal invariance of the GOE, this result is a direct consequence of Theorem 2.2.

It is worth noticing that, according to the Cramer-Lévy Theorem (cf. [F], Theorem 1 p. 525), the limiting distribution (2.13) is not Gaussian if μ is not Gaussian. Thus, (2.13) depends on the particular law μ of the entries of the Wigner matrix W_N which implies the non-universality of the fluctuations of the largest eigenvalue of rank one diagonal deformation of symmetric or Hermitian Wigner matrices (as conjectured in Remark 1.7 of [Fe-Pe]).

The latter also shows that in the non-Gaussian setting, the fluctuations of the largest eigenvalue depend, not only on the spectrum of the deformation A_N, but also on the particular definition of the matrix A_N. Indeed, in collaboration with S. Péché, the third author of the present article has recently stated in [Fe-Pe] the universality of the fluctuations of some Deformed Wigner models under a full deformation A_N defined by $(A_N)_{ij} = \frac{\theta}{N}$ for all $1 \leq i, j \leq N$ (see also [Fu-K]). Before giving some details on this work, we have to precise that [Fe-Pe] considered Deformed models such that the entries of the Wigner matrix W_N have sub-Gaussian moments. Nevertheless, thanks to the analysis made in [R], one can observe that the assumptions of [Fe-Pe] can be reduced and that it is for example sufficient to assume that the $W_{i,j}$'s have moment of order 19 (the precise condition of [R] is given by (2.15) below). Thus, the conclusions of [Fe-Pe] apply to the setting considered in our paper. The main result of [Fe-Pe] establishes the universality of the fluctuations of the largest eigenvalue of the complex Deformed model M_N associated to a full deformation A_N and for any value of the parameter θ. In particular, when $\theta > \sigma$, it is proved therein the universality of the Gaussian fluctuations (2.5). Notice that the approach of [Fe-Pe] is completely different from the one developed below in Section 5 to derive Theorem 2.2. It is mainly based on a combinatorial method inspired by the work [So] (which handles the non-Deformed Wigner model) and the known fluctuations for the Deformed GUE (given by [Pe]). The combinatorial arguments of [Fe-Pe] also work (with minor modifications) in the real framework and yields the universality of the fluctuations if $\theta < \sigma$. In the case where $\theta > \sigma$ which is of particular interest here, the analysis made in [Fe-Pe] reduced the universality problem in the real setting to the knowledge of the particular Deformed GOE model which was unknown up to now (note that this remark is also valid in the case where $\theta = \sigma$). Thus, thanks to our previous Theorem 2.3 and the analysis of [Fe-Pe] and [R], we are now in position to claim the following universality.

Theorem 2.4. *Let A_N be a full perturbation given by $(A_N)_{ij} = \frac{\theta}{N}$ for all (i, j). Assume that $\theta > \sigma$. Let W_N be an arbitrary real Wigner matrix with the underlying measure μ being symmetric with a variance σ^2 and such that there exists some $p > 18$ satisfying*

$$\mu([x, +\infty[) \leq \frac{1}{x^p}. \tag{2.15}$$

Then the largest eigenvalue of the Deformed model M_N has the Gaussian fluctuations (2.14).

Remark 2.3. *To be complete, let us notice that the previous result still holds when we allow the distribution ν of the diagonal entries of W_N being different from μ provided that ν is symmetric and satisfies (2.15).*

The fundamental tool of this paper is the Stieltjes transform. For $z \in \mathbb{C}\backslash\mathbb{R}$, we denote the resolvent of the matrix M_N by

$$G_N(z) = (zI_N - M_N)^{-1}$$

and the Stieltjes transform of the expectation of the empirical measure of the eigenvalues of M_N by

$$g_N(z) = \mathbb{E}(\mathrm{tr}_N(G_N(z)))$$

where tr_N is the normalized trace. We also denote by

$$g_\sigma(z) = \mathbb{E}((z-s)^{-1})$$

the Stieltjes transform[1] of a random variable s with semi-circular distribution μ_{sc}.

Theorem 2.1 is the analog of the main statement of [Bk-S1] established in the context of general spiked population models. The conclusion of [Bk-S1] requires numerous results obtained previously by J. Silverstein and co-authors in [Ch-S], [B-S1] and [B-S2] (a summary of all this literature can be found in [B] pp. 671–675). From very clever and tedious manipulations of some Stieltjes transforms and the use of the matricial representation (1.1), these works highligh a very close link between the spectra of the Wishart matrices and the covariance matrix (for quite general covariance matrix which include the spiked population model). Our approach mimics the one of [Bk-S1]. Thus, using the fact that the Deformed Wigner model is the additive analog of the spiked population model, several arguments can be quite easily adapted here (this point has been explained in Chapter 4 of the PhD Thesis [Fe]). Actually, the main point in the proof consists in establishing that for any $\varepsilon > 0$, almost surely,

$$\mathrm{Spect}(M_N) \subset K_\sigma^\varepsilon(\theta_1, \cdots, \theta_J) \tag{2.16}$$

for all N large, where we have defined

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

and

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

This point is the analog of the main result of [B-S1]. The analysis of [B-S1] is based on technical and numerous considerations of Stieltjes transforms strongly related to the Wishart context and that can not be directly transposed here. Thus, our approach to prove such an inclusion of the spectrum of M_N is very different from the one of [B-S1]. Indeed, we use the methods developed by U. Haagerup and S. Thorbjørnsen in [H-T], by H. Schultz [S] and by the two first authors of the present article [C-D]. 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

$$g_\sigma(z) - g_N(z) + \frac{1}{N} L_\sigma(z) = O(\frac{1}{N^2}), \tag{2.17}$$

where L_σ is the Stieltjes transform of a distribution Λ_σ with compact support in $K_\sigma(\theta_1, \cdots, \theta_J)$. Indeed such an estimation allows us through the inverse Stieltjes transform and some variance estimates to deduce that $\mathrm{tr}_N\, 1_{{}^cK^\varepsilon(\theta_1,\cdots,\theta_J)}(M_N) = O(1/N^{\frac{4}{3}})$ a.s.. Thus the number of eigenvalues of M_N in ${}^cK^\varepsilon(\theta_1, \cdots, \theta_J)$ is almost surely a $O(1/N^{\frac{1}{3}})$ and since for each N this number has to be an integer, we deduce that it is actually equal to zero as N goes to infinity.

Dealing with the particular diagonal perturbation $A_N = \mathrm{diag}(\theta, 0, \ldots, 0)$, we obtain the fluctuations of the largest eigenvalue $\lambda_1(M_N)$ by an approach close to the one of [P] and the ideas of [B-B-P]. The reasoning relies on the writing of $\lambda_1(M_N)$ in terms of the resolvent of a non-Deformed Wigner matrix.

The paper is organized as follows. In Section 3, we introduce preliminary lemmas which will be of basic use later on. Section 4 is devoted to the proof of Theorem 2.1. We first establish an equation (called "Master equation") satisfied by g_N up to some correction of order $\frac{1}{N^2}$ (see Section 4.1). Then

[1] Note that in some papers for which we make reference, the Stieltjes transform is defined with the opposite sign.

we explain how this master equation gives rise to an estimation of type (2.17) and thus to the inclusion (2.16) of the spectrum of M_N in $K^{\varepsilon}(\theta_1, \cdots, \theta_J)$ (see Sections 4.2 and 4.3). In Section 4.4, we use this inclusion to relate the asymptotic spectra of A_N and M_N and then deduce Theorem 2.1. The last section states Theorem 2.2.

Acknowledgments. The authors are very grateful to Jack Silverstein and Jinho Baik for providing them their proof of Theorem 5.2 (that is a fundamental argument in the proof of Theorem 2.2) which is presented in the Appendix of the present article.

3 Basic lemmas

We assume that the distribution μ of the entries of the Wigner matrix W_N satisfies a Poincaré inequality: there exists a positive constant C such that for any $\mathcal{C}^1$ function $f : \mathbb{R} \to \mathbb{C}$ such that f and f' are in $L^2(\mu)$,

$$\mathbf{V}(f) \leq C \int |f'|^2 d\mu,$$

with $\mathbf{V}(f) = \mathbb{E}(|f - \mathbb{E}(f)|^2)$.
For any matrix M, define $||M||_2 = (\mathrm{Tr}(M^*M))^{1/2}$ the Hilbert-Schmidt norm. Let $\Psi : (M_N(\mathbb{C})_{sa}) \to \mathbb{R}^{N^2}$ (resp. $\Psi : (M_N(\mathbb{C})_s) \to \mathbb{R}^{\frac{N(N+1)}{2}}$) be the canonical isomorphism which maps an Hermitian (resp. symmetric) matrix to the real parts and the imaginary parts of its entries (resp. to the entries) $(M)_{ij}, i \leq j$.

Lemma 3.1. *Let M_N be the complex (resp. real) Wigner Deformed matrix introduced in Section 2. For any $\mathcal{C}^1$ function $f : \mathbb{R}^{N^2}$(resp. $\mathbb{R}^{\frac{N(N+1)}{2}}$) $\to \mathbb{C}$ such that f and the gradient $\nabla(f)$ are both polynomially bounded,*

$$\mathbf{V}[f \circ \Psi(M_N)] \leq \frac{C}{N} \mathbb{E}\{\|\nabla \left[f \circ \Psi(M_N)\right]\|_2^2\}. \tag{3.1}$$

Proof: According to Lemma 3.2 in [C-D],

$$\mathbf{V}[f \circ \Psi(X_N)] \leq \frac{C}{N} \mathbb{E}\{\|\nabla \left[f \circ \Psi(X_N)\right]\|_2^2\}. \tag{3.2}$$

Note that even if the result in [C-D] is stated in the Hermitian case, the proof is valid and the result still holds in the symmetric case. Now (3.1) follows putting $g(x_{ij}; i \leq j) := f(x_{ij} + (A_N)_{ij}; i \leq j)$ in (3.2) and noticing that the $(A_N)_{ij}$ are uniformly bounded in i, j, N. □

This lemma will be useful to estimate many variances.
Now, we recall some useful properties of the resolvent (see [K-K-P], [C-D]). For any Hermitian matrix M we denote its spectrum by $\mathrm{Spect}(M)$.

Lemma 3.2. *For a $N \times N$ Hermitian or symmetric matrix M, for any $z \in \mathbb{C}\backslash\mathrm{Spect(M)}$, we denote by $G(z) := (zI_N - M)^{-1}$ the resolvent of M.*
Let $z \in \mathbb{C}\backslash\mathbb{R}$,

(i) $\|G(z)\| \leq |\Im m(z)|^{-1}$ *where* $\|.\|$ *denotes the operator norm.*

(ii) $|G(z)_{ij}| \leq |\Im m(z)|^{-1}$ *for all* $i, j = 1, \ldots N$.

(iii) For $p \geq 2$,

$$\frac{1}{N} \sum_{i,j=1}^{N} |G(z)_{ij}|^p \leq (|\Im m(z)|^{-1})^p.$$

(iv) The derivative with respect to M of the resolvent $G(z)$ satisfies:

$$G'_M(z).B = G(z)BG(z) \quad \textit{for any matrix } B.$$

(v) Let $z \in \mathbb{C}$ such that $|z| > \|M\|$; we have

$$\|G(z)\| \leq \frac{1}{|z| - \|M\|}.$$

Proof: We just mention that *(v)* comes readily noticing that the eigenvalues of the normal matrix $G(z)$ are the $\frac{1}{z-\lambda_i(M)}, i = 1, \ldots, N$. □

We will also need the following estimations on the Stieltjes transform g_σ of the semi-circular distribution μ_{sc}.

Lemma 3.3. *g_σ is analytic on $\mathbb{C}\backslash[-2\sigma, 2\sigma]$ and*

- $\forall z \in \{z \in \mathbb{C} : \Im m z \neq 0\}$,

$$\sigma^2 g_\sigma^2(z) - z g_\sigma(z) + 1 = 0. \tag{3.3}$$

$$|g_\sigma(z)| \leq |\Im m z|^{-1}. \tag{3.4}$$

$$|g_\sigma(z)^{-1}| \leq |z| + \sigma^2 |\Im m z|^{-1}. \tag{3.5}$$

$$|g'_\sigma(z)| = |\int \frac{1}{(z-t)^2} \, d\mu_\sigma(t)| \leq |\Im m(z)|^{-2}. \tag{3.6}$$

$$\text{For } a > 0, \theta \in \mathbb{R}, \qquad \left| \frac{1}{a g_\sigma(z) - z + \theta} \right| \leq |\Im m(z)|^{-1}. \tag{3.7}$$

- $\forall z \in \{z \in \mathbb{C} : |z| > 2\sigma\}$,

$$|g_\sigma(z)| \leq \frac{1}{|z| - 2\sigma}. \tag{3.8}$$

$$|g'_\sigma(z)| = |\int \frac{1}{(z-t)^2} \, d\mu_\sigma(t)| \leq \frac{1}{(|z| - 2\sigma)^2}. \tag{3.9}$$

$$|g_\sigma(z)|^{-1} \leq |z| + \frac{\sigma^2}{|z| - 2\sigma}. \tag{3.10}$$

Proof: For the equation (3.3), we refer the reader to Section 3.1 of [B]. (3.7) is a consequence of $\Im m(g_\sigma(z))\Im m(z) < 0$. Other inequalities derive from (3.3) and the definition of g_σ. □

4 Almost sure convergence of the first extremal eigenvalues

Sections 4.1, 4.2 and 4.3 below describe the different steps of the proof of the inclusion (2.16). We choose to develop the case of the complex Deformed Wigner model and just to point out some differences with the real model case (at the end of Section 4.3) since the approach would be basically the same. In these sections, we will often refer the reader to the paper [C-D] where the authors deal with several independent non Deformed Wigner matrices. The reader needs to fix $r = 1$, $m = 1$, $a_0 = 0$, $a_1 = \sigma$ and to change the notations $\lambda = z$, $G_N = g_N$, $G = g_\sigma$ in [C-D] in order to use the different proofs we refer to in the present framework. We shall denote by P_k any polynomial of degree k with positive coefficients and by C, K any constants; P_k, C, K can depend on the fixed eigenvalues of A_N and may vary from line to line. We also adopt the following convention to simplify the writing: we sometimes state in the proofs below that a quantity $\Delta_N(z)$, $z \in \mathbb{C}\backslash\mathbb{R}$ is $O(\frac{1}{N^p})$, $p = 1, 2$. This means precisely that

$$|\Delta_N(z)| \leq (|z| + K)^l \frac{P_k(|\Im m(z)|^{-1})}{N^p}$$

for some k and some l and we give the precise majoration in the statements of the theorems or propositions.
Section 4.4 explains how to deduce Theorem 2.1 from the inclusion (2.16).

4.1 The master equation

4.1.1 A first master inequality

In order to obtain a master equation for $g_N(z)$, we first consider the Gaussian case, i.e. $X_N = X_N^G$ is distributed as the GUE($N, \frac{\sigma^2}{N}$) distribution.[2]

Let us recall the integration by part formula for the Gaussian distribution.

Lemma 4.1. *Let Φ be a complex valued $\mathcal{C}^1$ function on $(M_N(\mathbb{C})_{sa})$ and $X_N \sim GUE(N, \frac{\sigma^2}{N})$. Then,*

$$\mathbb{E}[\phi'(X_N).H] = \frac{N}{\sigma^2}\mathbb{E}[\phi(X_N)\operatorname{Tr}(X_N H)] \tag{4.1}$$

for any Hermitian matrix H, or by linearity for $H = E_{jk}$, $1 \leq j,k \leq N$ where E_{jk}, $1 \leq j,k \leq N$ is the canonical basis of the complex space of $N \times N$ matrices.

We apply the above lemma to the function $\Phi(X_N) = (G_N(z))_{ij} = ((zI_N - X_N - A_N)^{-1})_{ij}$, $z \in \mathbb{C}\backslash\mathbb{R}$, $1 \leq i,j \leq N$. In order to simplify the notation, we write $(G_N(z))_{ij} = G_{ij}$. We obtain, for $H = E_{ij}$:

$$\begin{aligned}\mathbb{E}((GHG)_{ij}) &= \frac{N}{\sigma^2}\mathbb{E}(G_{ij}\operatorname{Tr}(X_N H)]\\ \mathbb{E}(G_{ii}G_{jj}) &= \frac{N}{\sigma^2}\mathbb{E}(G_{ij}(X_N)_{ji}]\end{aligned}$$

Now, we consider the normalized sum $\frac{1}{N^2}\sum_{ij}$ of the previous identities to obtain:

$$\mathbb{E}((\operatorname{tr}_N G)^2) = \frac{1}{\sigma^2}\mathbb{E}(\operatorname{tr}_N(GX_N)).$$

Then, since

$$GX_N = (z - X_N - A_N)^{-1}(X_N + A_N - zI_N - A_N + zI_N) = -I_N - GA_N + zG,$$

we obtain the following master equation:

$$\mathbb{E}((\operatorname{tr}_N G)^2) + \frac{1}{\sigma^2}(-z\mathbb{E}(\operatorname{tr}_N G) + 1 + \mathbb{E}(\operatorname{tr}_N GA_N)) = 0.$$

Now, it is well known (see [C-D], [H-T] and Lemma 3.1) that:

$$\operatorname{Var}(\operatorname{tr}_N(G)) \leq \frac{C|\Im mz|^{-4}}{N^2},$$

thus, we obtain:

Proposition 4.1. *The Stieltjes transform g_N satisfies the following inequality:*

$$|\sigma^2 g_N^2(z) - zg_N(z) + 1 + \frac{1}{N}\mathbb{E}(\operatorname{Tr}(G_N(z)A_N))| \leq C\frac{|\Im mz|^{-4}}{N^2} \tag{4.2}$$

Note that since A_N is of finite rank, $\mathbb{E}(\operatorname{Tr}(G_N(z)A_N)) \leq C$ where C is a constant independent of N (depending on the eigenvalues of A_N and z).

We now explain how to obtain the corresponding equation (4.2) in the Wigner case. Since the computations are the same as in [C-D][3] and [K-K-P][4], we just give some hints of the proof.

Step 1: The integration by part formula for the Gaussian distribution is replaced by the following tool:

[2]Throughout this section, we will drop the subscript G in the interest of clarity.
[3]This paper treats the case of several independent non Deformed Wigner matrices.
[4]The authors considered a non Deformed Wigner matrix in the symmetric real setting.

Lemma 4.2. *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,*

$$\mathbb{E}(\xi\phi(\xi)) = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!}\mathbb{E}(\phi^{(a)}(\xi)) + \epsilon \tag{4.3}$$

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

We apply this lemma with the function $\phi(\xi)$ given, as before, by $\phi(\xi) = G_{ij}$ and ξ is now one of the variable $\Re e((X_N)_{kl})$, $\Im m((X_N)_{kl})$. Note that, since the above random variables are symmetric, only the odd derivatives in (4.3) give a non null term. Moreover, as we are concerned by estimation of order $\frac{1}{N^2}$ of g_N, we only need to consider (4.3) up to the third derivative (see [C-D]). The computation of the first derivative will provide the same term as in the Gaussian case.

Step 2: Study of the third derivative.
We refer to [C-D] or [K-K-P] for a detailed study of the third derivative. Using some bounds on G_N, see Lemma 3.2, we can prove that the only term arising from the third derivative in the master equation, giving a contribution of order $\frac{1}{N}$, is:

$$\frac{1}{N}\mathbb{E}\left[\left(\frac{1}{N}\sum_{k=1}^{N} G_{kk}^2\right)^2\right].$$

In conclusion, the first master equation in the Wigner case reads as follows:

Theorem 4.1. *For $z \in \mathbb{C}\backslash\mathbb{R}$, $g_N(z)$ satisfies*

$$\left|\sigma^2 g_N(z)^2 - z g_N(z) + 1 + \frac{1}{N}\mathbb{E}[\mathrm{Tr}(G_N(z)A_N)] + \frac{1}{N}\frac{\kappa_4}{2}\mathbb{E}\left[\left(\frac{1}{N}\sum_{k=1}^{N}(G_N(z))_{kk}^2\right)^2\right]\right| \leq \frac{P_6(|\Im m(z)|^{-1})}{N^2} \tag{4.4}$$

where κ_4 is the fourth cumulant of the distribution μ.

4.1.2 Estimation of $|g_N - g_\sigma|$

Since

$$|\mathbb{E}[\mathrm{Tr}(G_N(z)A_N)]| \vee |\mathbb{E}\left[\left(\frac{1}{N}\sum_{k=1}^{N}(G_N(z))_{kk}^2\right)^2\right]| \leq P_4(|\Im m(z)|^{-1}),$$

Theorem 4.1 implies that for any $z \in \mathbb{C}\backslash\mathbb{R}$,

$$|\sigma^2 g_N(z)^2 - z g_N(z) + 1| \leq \frac{P_6(|\Im m(z)|^{-1})}{N} \tag{4.5}$$

To estimate $|g_N - g_\sigma|$ from the equation (3.3) satisfied by the Stieltjes transform g_σ on the one hand and from the equation (4.5) on the other hand, we follow the method initiated in [H-T] and [S]. We don't develop it here since it follows exactly the lines of Section 3.4 in [C-D] but we briefly recall the main arguments and results which will be useful later on. We define the open connected set

$$\mathcal{O}_N^{'} = \{z \in \mathbb{C}, \Im m(z) > 0, \frac{P_6(|\Im m z|^{-1})}{N}(\sigma^2|\Im m(z)|^{-1} + |z|) < \frac{1}{4|\Im m z|^{-1}}\}.$$

One can prove that for any z in $\mathcal{O}_N^{'}$,

- $g_N(z) \neq 0$ and

$$\frac{1}{|g_N(z)|} \leq 2(\sigma^2|\Im m(z)|^{-1} + |z|) \tag{4.6}$$

- $\Lambda_N(z) := \sigma^2 g_N(z) + \frac{1}{g_N(z)}$ is such that

$$|\Lambda_N(z) - z| \leq \frac{P_6(|\Im m z|^{-1})}{N} 2(\sigma^2 |\Im m z|^{-1} + |z|) \tag{4.7}$$

  and we have

$$\Im m(\Lambda_N(z)) \geq \frac{\Im m(z)}{2} > 0. \tag{4.8}$$

- Writing the equation (3.3) at the point $\Lambda_N(z)$, we easily get that

$$g_N(z) = g_\sigma(\Lambda_N(z)) \tag{4.9}$$

  on the non empty open subset $\mathcal{O}^{"}_N = \{z \in \mathcal{O}^{'}_N, \Im m z > \sqrt{2}\sigma\}$ and then on $\mathcal{O}^{'}_N$ by the principle of uniqueness of continuation.

This allows us to get an estimation of $|g_N(z) - g_\sigma(z)|$ on $\mathcal{O}^{'}_N$ and then to deduce:

Proposition 4.2. *For any $z \in \mathbb{C}$ such that $\Im m(z) > 0$,*

$$|g_N(z) - g_\sigma(z)| \leq (|z| + K) \frac{P_9(|\Im m z|^{-1})}{N}. \tag{4.10}$$

4.1.3 Study of the additional term $\mathbb{E}[\mathrm{Tr}(A_N G_N(z))]$

¿From now on and until the end of Section 4.1, we denote by $\gamma_1, \ldots, \gamma_r$ the non-null eigenvalues of A_N ($\gamma_i = \theta_j$ for some $j \neq j_0$) in order to simplify the writing. Let $U_N := U$ be a unitary matrix such that $A_N = U^* \Delta U$ where Δ is the diagonal matrix with entries $\Delta_{ii} = \gamma_i,\ i \leq r;\ \Delta_{ii} = 0,\ i > r$. We set

$$h_N(z) = \mathbb{E}[\mathrm{Tr}(A_N G_N(z))] \tag{4.11}$$

$$h_N(z) = \sum_{k=1}^{r} \gamma_k \sum_{i,j=1}^{N} U^*_{ik} U_{kj} \mathbb{E}[G_{ji}]$$

Our aim is to express $h_N(z)$ in terms of the Stieltjes transform $g_N(z)$ for N large, using the integration by part formula. Note that since we want an estimation of order $O(\frac{1}{N^2})$ in the master inequality (4.4), we only need an estimation of $h_N(z)$ of order $O(\frac{1}{N})$. As in the previous subsection, we first write the equation in the Gaussian case and then study the additional term (third derivative) in the Wigner case.

a) Gaussian case

Apply the formula (4.1) to $\Phi(X_N) = G_{jl}$ and $H = E_{il}$ to get

$$\mathbb{E}[G_{ji} G_{ll}] = \frac{N}{\sigma^2} \mathbb{E}[G_{jl} (X_N)_{li}]$$

and

$$\frac{1}{N} \sum_{l=1}^{N} \mathbb{E}[G_{ji} G_{ll}] = \frac{1}{\sigma^2} \mathbb{E}[(G X_N)_{ji}].$$

Expressing GX_N in terms of GA_N, we obtain:

$$I_{ji} := \sigma^2 \mathbb{E}[G_{ji}\, \mathrm{tr}_N(G)] + \delta_{ij} - z \mathbb{E}[G_{ji}] + \mathbb{E}[(GA_N)_{ji}] = 0. \tag{4.12}$$

Now, we consider the sum $\sum_{i,j} U^*_{ik} U_{kj} I_{ji},\ k = 1, \ldots r$ fixed and we denote $\alpha_k = \sum_{i,j} U^*_{ik} U_{kj} G_{ji} = (UGU^*)_{kk}$. Then, we have the following equality, using that U is unitary:

$$\sigma^2 \mathbb{E}[\alpha_k\, \mathrm{tr}_N(G)] + 1 - z \mathbb{E}[\alpha_k] + \sum_{i,j} U^*_{ik} U_{kj} \mathbb{E}[(GA_N)_{ji}] = 0.$$

Now,

$$\sum_{i,j} U^*_{ik} U_{kj} \mathbb{E}[(GA_N)_{ji}] = \mathbb{E}[(UGA_N U^*)_{kk}] = \mathbb{E}[(UGU^* \Delta UU^*)_{kk}]$$
$$= \gamma_k \mathbb{E}[(UGU^*)_{kk}] = \gamma_k \mathbb{E}[\alpha_k].$$

Therefore,

$$\sigma^2 \mathbb{E}[\alpha_k \operatorname{tr}_N(G)] + 1 + (\gamma_k - z)\mathbb{E}[\alpha_k] = 0.$$

Since α_k is bounded and $\operatorname{Var}(\operatorname{tr}_N(G)) = O(\frac{1}{N^2})$, we obtain

$$\mathbb{E}[\alpha_k](\sigma^2 g_N(z) + \gamma_k - z) + 1 = O(\frac{1}{N}). \tag{4.13}$$

Then using (4.10) we deduce that $\mathbb{E}[\alpha_k](\sigma^2 g_\sigma(z) + \gamma_k - z) + 1 = O(\frac{1}{N})$ and using (3.7)

$$h_N(z) = \sum_{k=1}^r \gamma_k \mathbb{E}[\alpha_k] = \sum_{k=1}^r \frac{\gamma_k}{z - \sigma^2 g_\sigma(z) - \gamma_k} + O(\frac{1}{N}). \tag{4.14}$$

b) The general Wigner case

We shall prove that (4.13) still holds. We now rely on Lemma 4.2 to obtain the analogue of (4.12)

$$J_{ij} := \sigma^2 \mathbb{E}[G_{ji} \operatorname{tr}_N(G)] + \delta_{ij} - z\mathbb{E}[G_{ji}] + \mathbb{E}[(GA_N)_{ji}] + \frac{\kappa_4}{6N}\left[\frac{1}{N}\sum_{l=1}^N \mathbb{E}[A_{i,j,l}]\right] = O(\frac{1}{N^2}). \tag{4.15}$$

The term $A_{i,j,l}$ is a fixed linear combination of the third derivative of $\Phi := G_{jl}$ with respect to $Re(X_N)_{il}$ (i.e. in the direction $e_{il} = E_{il} + E_{li}$) and $\Im m(X_N)_{il}$ (i.e. in the direction $f_{il} := \sqrt{-1}(E_{il} - E_{li})$). We don't need to write the exact form of this term since we just want to show that this term will give a contribution of order $O(\frac{1}{N})$ in the equation for $h_N(z)$. Let us write the derivative in the direction e_{il}:

$$\mathbb{E}[(Ge_{il}Ge_{il}Ge_{il}G)_{jl}]$$

which is the sum of eight terms of the form:

$$\mathbb{E}[G_{ji_1}G_{i_2i_3}G_{i_4i_5}G_{i_6l}] \tag{4.16}$$

where if $i_{2q+1} = i$ (resp. l), then $i_{2q+2} = l$ (resp. i), $q = 0, 1, 2$.

Lemma 4.3. *Let $1 \leq k \leq r$ fixed, then*

$$F(N) := |\sum_{i,j=1}^N U^*_{ik} U_{kj} \frac{1}{N}\sum_{l=1}^N \mathbb{E}[A_{i,j,l}]| \leq C|\Im mz|^{-4} \tag{4.17}$$

for a numerical constant C.

Proof: $F(N)$ is the sum of eight terms corresponding to (4.16). Let us write for example the term corresponding to $i_1 = i$, $i_3 = i$, $i_5 = i$:

$$\frac{1}{N}\sum_{i,j,l} U^*_{ik} U_{kj} E[G_{ji}G_{li}G_{li}G_{ll}]$$
$$= \mathbb{E}\left[\frac{1}{N}\sum_{i,l} U^*_{ik}(UG)_{ki}G_{li}G_{li}G_{ll}\right]$$
$$= \mathbb{E}\left[\frac{1}{N}\sum_{i} U^*_{ik}(UG)_{ki}(G^T G^D G^T)_{ii}\right]$$

where the superscript T denotes the transpose of the matrix and G^D is the diagonal matrix with entries G_{ii}. From the bounds $\|G_N(z)\| \leq |\Im mz|^{-1}$ and $\|U\| = 1$, we get the bound given in the

lemma.
We give the majoration for the term corresponding to $i_1 = l$, $i_3 = l$, $i_5 = l$:

$$\begin{aligned}
&\frac{1}{N}\sum_{i,j,l} U^*_{ik}U_{kj}\mathbb{E}[G_{jl}G^3_{il}] \\
&= \mathbb{E}\left[\frac{1}{N}\sum_{i,l} U^*_{ik}(UG)_{kl}G^3_{il}\right]
\end{aligned}$$

Its absolute value is bounded by $\mathbb{E}\left[\frac{1}{N}\sum_{i,l}|G_{il}|^3\right]|\Im m z|^{-1}$ and thanks to lemma 3.2 by $|\Im m z|^{-4}$. The other terms are treated in the same way. $\square$

As in the Gaussian case, we now consider the sum $\sum_{i,j} U^*_{ik}U_{kj}J_{ji}$. From Lemma 4.3 and the bound (using Cauchy-Schwarz inequality)

$$\sum_{i,j=1}^{N} |U^*_{ik}U_{kj}| \leq N,$$

we still get (4.13) and thus (4.14). More precisely, we proved

Proposition 4.3. *Let $h_N(z) = \mathbb{E}[\mathrm{Tr}(A_N G_N(z))]$, then*

$$|h_N(z) - \sum_{k=1}^{r}\frac{\gamma_k}{z - \sigma^2 g_\sigma(z) - \gamma_k}| \leq \frac{P_{11}(|\Im m(z)|^{-1})}{N}(K + |z|).$$

4.1.4 Convergence of $\mathbb{E}\left[\left(\frac{1}{N}\sum_{k=1}^{N} G^2_{kk}\right)^2\right]$

We now study the last term in the master inequality of Theorem 4.1. For the non Deformed Wigner matrices, it is shown in [K-K-P] that

$$R_N(z) := \mathbb{E}\left[\left(\frac{1}{N}\sum_{k=1}^{N} G^2_{kk}\right)^2\right] \underset{N \longrightarrow \infty}{\longrightarrow} g^4_\sigma(z).$$

Moreover, Proposition 3.2 in [C-D], in the more general setting of several independent Wigner matrices, gives an estimate of $|R_N(z) - g^4_\sigma(z)|$. The above convergence holds true in the Deformed case. We just give some hints of the proof of the estimate of $|R_N(z) - g^4_\sigma(z)|$ since the computations are almost the same as in the non Deformed case. Let us set

$$d_N(z) = \frac{1}{N}\sum_{k=1}^{N} G^2_{kk}.$$

We start from the resolvent identity

$$\begin{aligned}
zG_{kk} &= 1 + \sum_{l=1}^{N}(M_N)_{kl}G_{lk} \\
&= 1 + \sum_{l=1}^{N}(A_N)_{kl}G_{lk} + \sum_{l=1}^{N}(X_N)_{kl}G_{lk}
\end{aligned}$$

and

$$zd_N(z) = \frac{1}{N}\sum_{k=1}^{N} G_{kk} + \frac{1}{N}\sum_{k=1}^{N}(A_N G)_{kk}G_{kk} + \frac{1}{N}\sum_{k,l=1}^{N}(X_N)_{kl}G_{lk}G_{kk}.$$

For the last term, we apply an integration by part formula (Lemma 4.2) to obtain (see [K-K-P], [C-D])

$$\mathbb{E}\left[\frac{1}{N}\sum_{k,l=1}^{N}(X_N)_{kl}G_{lk}G_{kk}\right] = \sigma^2\mathbb{E}\left[(\frac{1}{N}\sum_{k=1}^{N} G_{kk})d_N(z)\right] + O(\frac{1}{N}).$$

It remains to see that the additional term due to A_N is of order $O(\frac{1}{N})$.

$$\frac{1}{N}\sum_{k=1}^{N}(A_N G)_{kk}G_{kk} = \frac{1}{N}\sum_{p=1}^{r}\gamma_p(UGG^D U^*)_{pp}$$

and

$$|\frac{1}{N}\sum_{k=1}^{N}(A_N G)_{kk}G_{kk}| \le (\sum_{p=1}^{r}|\gamma_p|)\frac{|\Im m z|^{-2}}{N}.$$

We thus obtain (again with the help of a variance estimate)

$$z\mathbb{E}[d_N(z)] = g_N(z) + \sigma^2 g_N(z)\mathbb{E}[d_N(z)] + O(\frac{1}{N}).$$

Then using (4.10) and since $d_N(z)$ is bounded we deduce that

$$z\mathbb{E}[d_N(z)] = g_\sigma(z) + \sigma^2 g_\sigma(z)\mathbb{E}[d_N(z)] + O(\frac{1}{N}).$$

Thus (using (3.7))

$$\mathbb{E}[d_N(z)] = \frac{g_\sigma(z)}{z-\sigma^2 g_\sigma(z)} + O(\frac{1}{N}) \underset{N\longrightarrow\infty}{\longrightarrow} \frac{g_\sigma(z)}{z-\sigma^2 g_\sigma(z)} = g_\sigma^2(z).$$

Now, using some variance estimate,

$$\mathbb{E}[d_N^2(z)] = (\mathbb{E}[d_N(z)])^2 + O(\frac{1}{N}) = g_\sigma^4(z) + O(\frac{1}{N}). \square$$

We can now give our final master inequality for $g_N(z)$ following our previous estimates:

Theorem 4.2. *For $z \in \mathbb{C}$ such that $\Im m(z) > 0$, $g_N(z)$ satisfies*

$$\left|\sigma^2 g_N^2(z) - z g_N(z) + 1 + \frac{1}{N}E_\sigma(z)\right| \le \frac{P_{14}(|\Im m z|^{-1})}{N^2}(|z|+K)$$

where $E_\sigma(z) = \sum_{k=1}^{r}\frac{\gamma_k}{z-\sigma^2 g_\sigma(z)-\gamma_k} + \frac{\kappa_4}{2}g_\sigma^4(z)$, κ_4 is the fourth cumulant of the distribution μ.

Note that $E_\sigma(z)$ can be written in terms of the distinct eigenvalues θ_j's of A_N as

$$E_\sigma(z) = \sum_{j=1, j\neq j_0}^{J} k_j \frac{\theta_j}{z-\sigma^2 g_\sigma(z)-\theta_j} + \frac{\kappa_4}{2}g_\sigma^4(z).$$

Let us set

$$L_\sigma(z) = g_\sigma(z)^{-1}\mathbb{E}((z-s)^{-2})E_\sigma(z) \tag{4.18}$$

where s is a centered semicircular random variable with variance σ^2.

4.2 Estimation of $|g_\sigma(z) - g_N(z) + \frac{1}{N}L_\sigma(z)|$

The method is roughly the same as the one described in Section 3.6 in [C-D]. Nevertheless we choose to develop it here for the reader convenience. We have for any z in $\mathcal{O}_n^{'}$, by using (4.9),

$$
\begin{aligned}
&|g_\sigma(z) - g_N(z) + \tfrac{1}{N} L_\sigma(z)| \\
&\quad = \quad |g_\sigma(z) - g_\sigma(\Lambda_N(z)) + \frac{1}{N} L_\sigma(z)| \\
&\quad = \quad |\mathbb{E}\left[(z-s)^{-1}(\Lambda_N(z)-s)^{-1}(\Lambda_N(z)-z) + \frac{1}{N} g_\sigma(z)^{-1}(z-s)^{-2}E_\sigma(z)\right]| \\
&\quad \leq \quad |\mathbb{E}\left[(z-s)^{-1}(\Lambda_N(z)-s)^{-1}(\Lambda_N(z)-z+\frac{1}{N} g_\sigma(z)^{-1}E_\sigma(z))\right]| \\
&\qquad\quad +\mathbb{E}\left[|(z-s)^{-1}\{(z-s)^{-1}-(\Lambda_N(z)-s)^{-1}\}|\right]\frac{1}{N}|g_\sigma(z)^{-1}E_\sigma(z)| \\
&\quad \leq \quad 2|\Im m(z)|^{-2}|\Lambda_N(z)-z+\frac{1}{N}E_\sigma(z)g_\sigma(z)^{-1}| \\
&\qquad\quad +\frac{P_8(|\Im m(z)|^{-1})}{N}|\Lambda_N(z)-z|(|z|+K)
\end{aligned}
$$

where we made use of the estimates (3.5), (4.8), $\forall z \in \mathbb{C}\backslash\mathbb{R}, x \in \mathbb{R}$,

$$
\left|\frac{1}{z-x}\right| \leq |\Im m(z)|^{-1},
$$

$$
|E_\sigma(z)| \leq P_4(|\Im m(z)|^{-1}) \quad \text{(using (3.7))}. \tag{4.19}
$$

Let us write

$$
\begin{aligned}
&|\Lambda_n(z) - z + \tfrac{1}{N}E_\sigma(z)g_\sigma(z)^{-1}| \\
&\quad = \tfrac{1}{g_N(z)}\left\{\sigma^2 g_N^2(z) - z g_N(z) + 1 + \tfrac{1}{N}E_\sigma(z)\right\} + \tfrac{1}{g_N(z)g_\sigma(z)}\tfrac{1}{N}\left\{g_N(z)-g_\sigma(z)\right\}E_\sigma(z).
\end{aligned}
$$

We get from Theorem 4.2, (4.6), (4.10), (4.19), (3.5)

$$
|\Lambda_N(z) - z + \frac{1}{N}E_\sigma(z)g_\sigma(z)^{-1}| \leq (|z|+K)^3 \frac{P_{15}(|\Im m z|^{-1})}{N^2}.
$$

Finally, using also (4.7), we get for any z in $\mathcal{O}_n^{'}$,

$$
|g_\sigma(z) - g_N(z) + \frac{1}{N}L_\sigma(z)| \leq (|z|+K)^3 \frac{P_{17}(|\Im m(z)|^{-1})}{N^2}.
$$

Now, for $z \notin \mathcal{O}_n^{'}$, such that $\Im m(z) > 0$

$$
\begin{aligned}
1 \quad &\leq \quad 4\frac{P_6(|(\Im m(z))^{-1}|)}{N}(|z| + \sigma^2|\Im m(z)|^{-1})|\Im m(z)|^{-1} \\
&\leq \quad (|z|+K)\frac{P_8(|\Im m(z)^{-1}|)}{N}.
\end{aligned}
$$

We get

$$
\begin{aligned}
|g_\sigma(z) - g_N(z) + \frac{1}{N}L_\sigma(z)| \quad &\leq \quad |g_\sigma(z) - g_N(z)| + \frac{1}{N}|L_\sigma(z)| \\
&\leq \quad (|z|+K)\frac{P_8(|\Im m(z)|^{-1})}{N} \times \\
&\qquad \left[(|z|+K)\frac{P_9(|\Im m(z)|^{-1})}{N} + \frac{1}{N}P_7(|\Im m(z)|^{-1})(|z|+K)\right] \\
&\leq \quad (|z|+K)^2\frac{P_{17}(|\Im m(z)|^{-1})}{N^2}.
\end{aligned}
$$

Thus, for any z such that $\Im m(z) > 0$,

$$
|g_\sigma(z) - g_N(z) + \frac{1}{N}L_\sigma(z)| \leq (|z|+K)^3 \frac{P_{17}(|\Im m(z)|^{-1})}{N^2}. \tag{4.20}
$$

Let us denote for a while $g_N = g_N^{A_N}$ and $L_\sigma = L_\sigma^{A_N}$. Note that we get exactly the same estimation (4.20) dealing with $-A_N$ instead of A_N. Hence since $g_\sigma(z) = -g_\sigma(-z)$, $g_N^{-A_N}(z) = -g_N^{A_N}(-z)$ (using the symmetry assumption on μ) and $L_\sigma^{-A_N}(z) = L_\sigma^{A_N}(-z)$, it readily follows that (4.20) is also valid for any z such that $\Im m z < 0$. In conclusion,

Proposition 4.4. *For any $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$|g_\sigma(z) - g_N(z) + \frac{1}{N} L_\sigma(z)| \leq (|z| + K)^3 \frac{P_{17}(|\Im m(z)|^{-1})}{N^2}. \tag{4.21}$$

4.3 Inclusion of the spectrum of M_N

The following step now consists in deducing Proposition 4.6 from Proposition 4.4 (from which we will easily deduce the appropriate inclusion of the spectrum of M_N). Since this transition is based on the inverse Stieltjes transform, we start with establishing the fundamental Proposition 4.5 below concerning the nature of L_σ. Note that one can rewrite L_σ as

$$L_\sigma(z) = g_\sigma(z)^{-1} \times g'_\sigma(z) \left(\sum_{j=1}^{J} k_j \frac{\theta_j}{\frac{1}{g_\sigma(z)} - \theta_j} + \frac{\kappa_4}{2} g_\sigma^4(z) \right). \tag{4.22}$$

We recall that $J_{+\sigma}$ (resp. $J_{-\sigma}$) denotes the number of j's such that $\theta_j > \sigma$ (resp. $\theta_j < -\sigma$). As in the introduction, we define

$$\rho_{\theta_j} = \theta_j + \frac{\sigma^2}{\theta_j} \tag{4.23}$$

which is $> 2\sigma$ (resp. $< -2\sigma$) when $\theta_j > \sigma$ (resp. $< -\sigma$).

Proposition 4.5. *L_σ is the Stieltjes transform of a distribution Λ_σ with compact support*

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

As in [S], we will use the following characterization:

Theorem 4.3. *[T]*

- *Let Λ be a distribution on $\mathbb{R}$ with compact support. Define the Stieltjes transform of Λ, $l : \mathbb{C}\backslash\mathbb{R} \to \mathbb{C}$ by*

  $$l(z) = \Lambda\left(\frac{1}{z-x}\right).$$

  Then l is analytic in $\mathbb{C}\backslash\mathbb{R}$ and has an analytic continuation to $\mathbb{C}\backslash supp(\Lambda)$. Moreover

  (c_1) $l(z) \to 0$ as $|z| \to \infty$,

  (c_2) there exists a constant $C > 0$, an $n \in \mathbb{N}$ and a compact set $K \subset \mathbb{R}$ containing $supp(\Lambda)$ such that for any $z \in \mathbb{C}\backslash\mathbb{R}$,

  $$|l(z)| \leq C max\{dist(z, K)^{-n}, 1\},$$

  (c_3) for any $\phi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ with compact support

  $$\Lambda(\phi) = -\frac{1}{\pi} \lim_{y \to 0^+} \Im m \int_{\mathbb{R}} \phi(x) l(x + iy) dx.$$

- *Conversely, if K is a compact subset of $\mathbb{R}$ and if $l : \mathbb{C}\backslash K \to \mathbb{C}$ is an analytic function satisfying (c_1) and (c_2) above, then l is the Stieltjes transform of a compactly supported distribution Λ on $\mathbb{R}$. Moreover, $supp(\Lambda)$ is exactly the set of singular points of l in K.*

In our proof of Proposition 4.5, we will refer to the following lemma which gives several properties on g_σ.

Lemma 4.4. *g_σ is analytic and invertible on $\mathbb{C}\backslash[-2\sigma, 2\sigma]$ and its inverse z_σ satisfied*

$$z_\sigma(g) = \frac{1}{g} + \sigma^2 g, \quad \forall g \in g_\sigma(\mathbb{C}\backslash[-2\sigma, 2\sigma]).$$

(a) *The complement of the support of μ_σ is characterized as follows*

$$x \in \mathbb{R}\backslash[-2\sigma, 2\sigma] \Longleftrightarrow \exists g \in \mathbb{R}^* \text{ such that } |\frac{1}{g}| > \sigma \text{ and } x = z_\sigma(g).$$

(b) *Given $x \in \mathbb{R}\backslash[-2\sigma, 2\sigma]$ and $\theta \in \mathbb{R}$ such that $|\theta| > \sigma$, one has*

$$\frac{1}{g_\sigma(x)} = \theta \Longleftrightarrow x = \theta + \frac{\sigma^2}{\theta} := \rho_\theta.$$

This lemma can be easily proved using for example the explicit expression of g_σ (derived from (3.3)) namely for all $x \in \mathbb{R}\backslash[-2\sigma, 2\sigma]$,

$$g_\sigma(x) = \frac{x}{2\sigma^2}\left(1 - \sqrt{1 - \frac{4\sigma^2}{x^2}}\right).$$

Proof of Proposition 4.5: Using (4.22), one readily sees that the singular points of L_σ is the set $[-2\sigma; 2\sigma] \bigcup \left\{x \in \mathbb{R}\backslash[-2\sigma, 2\sigma] \text{ and } \frac{1}{g_\sigma(x)} \in \mathrm{Spect}(A_N)\right\}$. Hence (using point (b) of Lemma 4.4) the set of singular points of L_σ is exactly $K_\sigma(\theta_1, \cdots, \theta_J)$.

Now, we are going to show that L_σ satisfies (c_1) and (c_2) of Theorem 4.3. We have obviously that

$$|z - \sigma^2 g_\sigma(z) - \theta_j| \geq |\ |z - \theta_j| - |\sigma^2 g_\sigma(z)|\ |.$$

Now, let $\alpha > 0$ such that $\alpha > 2\sigma$ and for any $j = 1, \ldots, J$, $\alpha - |\theta_j| > \frac{\sigma^2}{\alpha - 2\sigma}$. For any $z \in \mathbb{C}$ such that $|z| > \alpha$,

$$|z - \theta_j| \geq |z| - |\theta_j| > \frac{\sigma^2}{\alpha - 2\sigma}$$

and according to (3.8)

$$|\sigma^2 g_\sigma(z)| \leq \frac{\sigma^2}{|z| - 2\sigma} \leq \frac{\sigma^2}{\alpha - 2\sigma}.$$

Thus we get that for $z \in \mathbb{C}$ such that $|z| > \alpha$,

$$|z - \sigma^2 g_\sigma(z) - \theta_j| \geq |z| - |\theta_j| - \frac{\sigma^2}{\alpha - 2\sigma}.$$

Using also (3.8), (3.9), (3.10), we get readily that for $|z| > \alpha$,

$$|L_\sigma(z)| \leq \left(|z| + \frac{\sigma^2}{|z| - 2\sigma}\right) \frac{1}{(|z| - 2\sigma)^2} \left\{\sum_{j=1}^J k_j \frac{|\theta_j|}{|z| - |\theta_j| \ - \frac{\sigma^2}{\alpha - 2\sigma}} + \frac{|\kappa_4|}{2(|z| - 2\sigma)^4}\right\}.$$

Then, it is clear than $|L_\sigma(z)| \to 0$ when $|z| \to +\infty$ and (c_1) is satisfied.

Now we follow the approach of [S](Lemma 5.5) to prove (c_2). Denote by $\mathcal{E}$ the convex envelope of $K_\sigma(\theta_1, \cdots, \theta_J)$ and define the interval

$$K := \{x \in \mathbb{R}; dist(x, \mathcal{E}) \leq 1\} = \left[\min\{x \in K_\sigma(\theta_1, \cdots, \theta_J)\} - 1; \max\{x \in K_\sigma(\theta_1, \cdots, \theta_J)\} + 1\right]$$

and

$$D = \{z \in \mathbb{C}; 0 < dist(z, K) \leq 1\}.$$

- Let $z \in D \cap \mathbb{C}\backslash\mathbb{R}$ with $\Re e(z) \in K$. We have $dist(z, K) = |\Im m z| \leq 1$. Using the upper bounds (3.4), (3.5), (3.6) and (3.7), we easily deduce that there exists some constant C_0 such that for any $z \in D \cap \mathbb{C}\backslash\mathbb{R}$ with $\Re e(z) \in K$.

$$|L_\sigma(z)| \leq C_0 |\Im m z|^{-7} = C_0 dist(z, K)^{-7} = C_0 \max(dist(z, K)^{-7}; 1).$$

- Let $z \in D \cap \mathbb{C}\backslash\mathbb{R}$ with $\Re e(z) \notin K$. Then $dist(z, K_\sigma(\theta_1, \cdots, \theta_J)) \geq 1$. Since L_σ is bounded on compact subsets of $\mathbb{C}\backslash K_\sigma(\theta_1, \cdots, \theta_J)$, we easily deduce that there exists some constant C_1 such that for any $z \in D$ with $\Re e(z) \notin K$,

$$|L_\sigma(z)| \leq C_1 \leq C_1 dist(z, K)^{-7} = C_1 \max(dist(z, K)^{-7}; 1).$$

- Since $|L_\sigma(z)| \to 0$ when $|z| \to +\infty$, L_σ is bounded on $\mathbb{C}\backslash D$. Thus, there exists some constant C_2 such that for any $z \in \mathbb{C}\backslash D$,

$$|L_\sigma(z)| \leq C_2 = C_2 \max(dist(z, K)^{-7}; 1).$$

Hence (c_2) is satisfied with $C = \max(C_0, C_1, C_2)$ and $n = 7$ and Proposition 4.5 follows from Theorem 4.3. □

We are now in position to deduce the following proposition from the estimate (4.21).

Proposition 4.6. *For any smooth function φ with compact support*

$$\mathbb{E}[\mathrm{tr}_N(\varphi(M_N))] = \int \varphi \, d\mu_{sc} + \frac{1}{N}\Lambda_\sigma(\varphi) + O(\frac{1}{N^2}). \tag{4.24}$$

Consequently, for φ smooth, constant outside a compact set and such that $supp(\varphi) \cap K_\sigma(\theta_1, \cdots, \theta_J) = \emptyset$,

$$\mathrm{tr}_N(\varphi(M_N)) = O(\frac{1}{N^{\frac{4}{3}}}) \quad a.s. \tag{4.25}$$

Proof: Using the inverse Stieltjes tranform, we get respectively that, for any φ in $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ with compact support,

$$\mathbb{E}[\mathrm{tr}_N(\varphi(M_N))] - \int \varphi d\mu_{sc} - \frac{\Lambda_\sigma(\varphi)}{N} = -\frac{1}{\pi} \lim_{y \to 0^+} \Im m \int_{\mathbb{R}} \varphi(x) r_N(x + iy) dx \tag{4.26}$$

where $r_N = g_\sigma(z) - g_N(z) + \frac{1}{N} L_\sigma(z)$ satisfies, according to Proposition 4.4, for any $z \in \mathbb{C}\backslash\mathbb{R}$,

$$|r_N(z)| \leq \frac{1}{N^2}(|z| + K)^\alpha P_k(|\Im m(z)^{-1}|) \tag{4.27}$$

where $\alpha = 3$ and $k = 17$.
We refer the reader to the Appendix of [C-D] where it is proved using the ideas of [H-T] that

$$\limsup_{y \to 0^+} |\int_{\mathbb{R}} \varphi(x) h(x + iy) dx| \leq C < +\infty \tag{4.28}$$

when h is an analytic function on $\mathbb{C}\backslash\mathbb{R}$ which satisfies

$$|h(z)| \leq (|z| + K)^\alpha P_k(|\Im m(z)^{-1}|). \tag{4.29}$$

Dealing with $h(z) = N^2 r_N(z)$, we deduce that

$$\limsup_{y \to 0^+} |\int_{\mathbb{R}} \varphi(x) r_N(x + iy) dx| \leq \frac{C}{N^2} \tag{4.30}$$

and then (4.24).
Following the proof of Lemma 5.6 in [S], one can show that $\Lambda_\sigma(1) = 0$. Then, the rest of the proof of

(4.25) sticks to the proof of Lemma 6.3 in [H-T] (using Lemma 3.1). $\square$

Following [H-T](Theorem 6.4), we set $K = K_\sigma(\theta_1, \cdots, \theta_J)+(-\frac{\varepsilon}{2}, \frac{\varepsilon}{2})$, $F = \{t \in \mathbb{R}; d(t, K_\sigma(\theta_1, \cdots, \theta_J)) \geq \varepsilon\}$ and take $\varphi \in \mathbb{C}^\infty(\mathbb{R}, \mathbb{R})$ such that $0 \leq \varphi \leq 1$, $\varphi(t) = 0$ for $t \in K$ and $\varphi(t) = 1$ for $t \in F$. Then according to (4.25), $\mathrm{tr}_N(\varphi(M_N)) = O(\frac{1}{N^{\frac{4}{3}}})$ a.s. Since $\varphi \geq 1_F$, it follows that $\mathrm{tr}_N(1_F(M_N)) = O(\frac{1}{N^{\frac{4}{3}}})$ a.s. and thus the number of eigenvalues of M_N in F is almost surely a $O(\frac{1}{N^{\frac{1}{3}}})$ as N goes to infinity. Since for each N this number has to be an integer we deduce that the number of eigenvalues of M_N in F is zero almost surely as N goes to infinity. The fundamental inclusion (2.16) follows, namely, for any $\varepsilon > 0$, almost surely

$$\mathrm{Spect(M_N)} \subset \mathrm{K}_\sigma(\theta_1, \cdots, \theta_\mathrm{J}) + (-\varepsilon, \varepsilon)$$

when N goes to infinity.

Such a method can be carried out in the case of Wigner real symmetric matrices; then the approximate Master equation is the following

$$\sigma^2 g_N(z)^2 - z g_N(z) + 1 + \frac{1}{N}\frac{\kappa_4}{2}\mathbb{E}\Big[\Big(\frac{1}{N}\sum_{k=1}^N G_{kk}(z)^2\Big)^2\Big] + \frac{\sigma^2}{N}\mathbb{E}\left(\mathrm{tr}_N\, G_N(z)^2\right) + \mathbb{E}\left(\mathrm{tr}_N\left[A_N G_N(z)\right]\right) = O(\frac{1}{N^2}).$$

Note that the additionnal term $\frac{\sigma^2}{N}\mathbb{E}\left(\mathrm{tr}_N\, G_N(z)^2\right)$ already appears in the non-Deformed GOE case in [S]. One can establish in a similar way the analog of (4.10) and then, following the proof of Corollary 3.3 in [S], deduce that

$$\mathbb{E}\left(\mathrm{tr}_N\, G_N(z)^2\right) = \mathbb{E}\left((z-s)^{-2}\right) + O(\frac{1}{N}),$$

where s is a centered semi-circular variable with variance σ^2. Hence by similar arguments as in the complex case, one get the master equation

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

where

$$E_\sigma(z) = \sum_{j=1, j\neq j_0}^J k_j \frac{\theta_j}{z - \sigma^2 g_\sigma(z) - \theta_j} + \frac{\kappa_4}{2} g_\sigma^4(z) + \mathbb{E}\left((z-s)^{-2}\right).$$

It can be proved that $L_\sigma(z) := g_\sigma(z)^{-1}\mathbb{E}((z-s)^{-2})E_\sigma(z)$ is the Stieltjes transform of a distribution Λ_σ with compact support $K_\sigma(\theta_1, \cdots, \theta_J)$ too. The last arguments hold likewise in the real symmetric case.

Hence we have established

Theorem 4.4. *Let $J_{+\sigma}$ (resp. $J_{-\sigma}$) be the number of j's such that $\theta_j > \sigma$ (resp. $\theta_j < -\sigma$). Then for any $\varepsilon > 0$, almost surely, there is no eigenvalue of M_N in*

$$\begin{aligned} & (-\infty, \rho_{\theta_J} - \epsilon) \cup (\rho_{\theta_J} + \epsilon, \rho_{\theta_{J-1}} - \epsilon) \cup \cdots \cup (\rho_{\theta_{J-J_{-\sigma}+1}} + \epsilon, -2\sigma - \epsilon) \\ \cup \quad & (2\sigma + \epsilon, \rho_{\theta_{J_{+\sigma}}} - \epsilon) \cup \cdots \cup (\rho_{\theta_2} + \epsilon, \rho_{\theta_1} - \epsilon) \cup (\rho_{\theta_1} + \epsilon, +\infty) \end{aligned} \tag{4.31}$$

when N is large enough.

Remark 4.1. *As soon as $\epsilon > 0$ is small enough that is when*

$$2\epsilon < \min\left(\rho_{\theta_{j-1}} - \rho_{\theta_j}, J - J_{-\sigma} + 2 \leq j \leq J \text{ or } 2 \leq j \leq J_{+\sigma}; -2\sigma - \rho_{\theta_{J-J_{-\sigma}+1}}; \rho_{\theta_{J_{+\sigma}}} - 2\sigma\right)$$

the union (4.31) is made of non-empty disjoint intervals.

4.4 Almost sure convergence of the first extremal eigenvalues

As announced in the introduction, Theorem 2.1 is the analog of the main statement of [Bk-S1] established for general spiked population models (1.1). The previous Theorem 4.4 is the main step of the proof since now, we can quite easily adapt the arguments needed for the conclusion of [Bk-S1] viewing the Deformed Wigner model (1.2) as the additive analog of the spiked population model (1.1).

Let us consider one of the positive eigenvalue θ_j of the A_N's. We recall that this implies that $\lambda_{k_1+\cdots+k_{j-1}+i}(A_N) = \theta_j$ for all $1 \leq i \leq k_j$. We want to show that if $\theta_j > \sigma$ (i.e. with our notations, if $j \in \{1, \cdots, J_{+\sigma}\}$), the corresponding eigenvalues of M_N almost surely jump above the right endpoint 2σ of the semicircle support as

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

whereas the rest of the asymptotic spectrum of M_N lies below 2σ with

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

Analog results hold for the negative eigenvalues θ_j (see points (c) and (d) of Theorem 2.1). To describe the phenomenon, one can say that, when N is large enough, the (first extremal) eigenvalues of M_N can be viewed as a "smoothed" deformation of the (first extremal) eigenvalues of A_N. So, our main purpose now is to establish the link between the spectra of the matrices M_N and A_N. According to the analysis made in the previous section (Proposition 4.5), we yet know that the θ_j's are related to the ρ_{θ_j}'s through the Stieltjes transform g_σ. More precisely, one has

$$\text{for all } j \text{ such that } |\theta_j| > \sigma, \ \frac{1}{g_\sigma(\rho_{\theta_j})} = \theta_j.$$

Actually, one can refine this analysis and state the following important Lemma 4.5 on g_σ. As before, we denote (recall Lemma 4.4) by z_σ its inverse which is given by

$$z_\sigma(g) = \frac{1}{g} + \sigma^2 g.$$

Using Lemma 4.4, one readily sees that the set ${}^cK_\sigma(\theta_1, \cdots, \theta_J)$ can be characterized as follows

$$x \in {}^cK_\sigma(\theta_1, \cdots, \theta_J) \Longleftrightarrow \exists g \in \mathcal{G}_\sigma \text{ such that } x = z_\sigma(g) \tag{4.32}$$

where

$$\mathcal{G}_\sigma := \Big\{ g \in \mathbb{R}^* : |\frac{1}{g}| > \sigma \text{ and } \frac{1}{g} \notin \text{Spect}(A_N) \Big\}.$$

Obviously, when $x \in {}^cK_\sigma(\theta_1, \cdots, \theta_J)$, one has $g = g_\sigma(x)$.

Lemma 4.5. *Let $[a, b]$ be a compact set contained in ${}^cK_\sigma(\theta_1, \cdots, \theta_J)$. Then,*

(i) $[\frac{1}{g_\sigma(a)}, \frac{1}{g_\sigma(b)}] \subset (\text{Spect}(A_N))^c$.

(ii) For all $0 < \hat{\sigma} < \sigma$, the interval $[z_{\hat{\sigma}}(g_\sigma(a)); z_{\hat{\sigma}}(g_\sigma(b))]$ is contained in ${}^cK_{\hat{\sigma}}(\theta_1, \cdots, \theta_J)$ and $z_{\hat{\sigma}}(g_\sigma(b)) - z_{\hat{\sigma}}(g_\sigma(a)) \geq b - a$.

Proof: The function $\frac{1}{g_\sigma}$ being increasing, (i) readily follows from (4.32).
Noticing that $\mathcal{G}_\sigma \subset \mathcal{G}_{\hat{\sigma}}$ for all $\hat{\sigma} < \sigma$ implies (recall also that g_σ decreases on $[a, b]$) that $[g_\sigma(b); g_\sigma(a)] \subset \mathcal{G}_{\hat{\sigma}}$. Relation (4.32) combined with the fact that the function $z_{\hat{\sigma}}$ is decreasing on $[g_\sigma(b); g_\sigma(a)]$ leads to

$$[z_{\hat{\sigma}}(g_\sigma(a)); z_{\hat{\sigma}}(g_\sigma(b))] \subset {}^cK_{\hat{\sigma}}(\theta_1, \cdots, \theta_J)$$

and the first part of (ii) is stated. Now, we have

$$\begin{aligned} l_\sigma(\hat{\sigma}) &:= z_{\hat{\sigma}}(g_\sigma(b)) - z_{\hat{\sigma}}(g_\sigma(a)) \\ &= \frac{1}{g_\sigma(b)} - \frac{1}{g_\sigma(a)} + \hat{\sigma}^2(g_\sigma(b) - g_\sigma(a)). \end{aligned}$$

Since g_σ decreases on $[a,b]$, we have $g_\sigma(b) - g_\sigma(a) \leq 0$ and thus l_σ is decreasing on $\mathbb{R}^+$. Then the last point of (ii) follows since $l_\sigma(\sigma) = b - a$. $\Box$

Thanks to this lemma and the previous Theorem 4.4, one can state the asymptotic relation between the spectrum of A_N and the one of M_N.
Let $[a,b]$ be an interval contained in ${}^cK_\sigma(\theta_1, \cdots, \theta_J)$. By Theorem 4.4, $[a,b]$ is outside the spectrum of M_N. Moreover, from Lemma 4.5 (i), it corresponds an interval $[a',b']$ outside the spectrum of A_N i.e. there is an integer $i_N \in \{0, \ldots, N\}$ such that

$$\lambda_{i_N+1}(A_N) < \frac{1}{g_\sigma(a)} := a' \quad \text{and} \quad \lambda_{i_N}(A_N) > \frac{1}{g_\sigma(b)} := b'. \tag{4.33}$$

a and a' (resp. b and b') are linked as follows

$$a = \rho_{a'} := a' + \frac{\sigma^2}{a'} \quad (\text{resp. } b = \rho_{b'}).$$

Our aim now is to prove that $[a,b]$ splits the eigenvalues of M_N exactly as $[a',b']$ splits the spectrum of A_N. In [B-S2], one talks about the exact separation phenomenon.

Theorem 4.5. *With i_N satisfying (4.33), one has*

$$\mathbb{P}[\lambda_{i_N+1}(M_N) < a \text{ and } \lambda_{i_N}(M_N) > b, \text{ for all large } N] = 1. \tag{4.34}$$

Remark 4.2. *This result is the analog of the main statement of [B-S2] (cf. Theorem 1.2 of [B-S2]) established in the spiked population setting (and in fact for quite general sample covariance matrices).*

Intuitively, the statement of Theorem 4.5 seems rather natural when σ is close to zero. Indeed, when N goes to ∞, since the spectrum of $\frac{1}{\sqrt{N}}W_N$ is concentrated in $[-2\sigma, 2\sigma]$ (recall (2.9)), the spectrum of M_N would be close to the one of A_N as soon as σ will be close to zero (in other words, the spectrum of M_N is, viewed as a deformation of the one of A_N, continuous in σ in a neighborhood of zero). Actually, this can be justified regardless of the size of σ thanks to the following classical result (due to Weyl).

Lemma 4.6. *(cf. Theorem 4.3.7 of [H-J]) Let B and C be two $N \times N$ Hermitian matrices. For any pair of integers j, k such that $1 \leq j, k \leq N$ and $j + k \leq N + 1$, we have*

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

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

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

Note that this lemma is the additive analogue of Lemma 1.1 of [B-S2] needed for the investigation of the spiked population model.

Proof of Theorem 4.5: With our choice of $[a,b]$ and the very definition of the spectrum of the A_N's, one can consider $\epsilon' > 0$ small enough such that, for all large N,

$$\lambda_{i_N+1}(A_N) < \frac{1}{g_\sigma(a)} - \epsilon' \quad \text{and} \quad \lambda_{i_N}(A_N) > \frac{1}{g_\sigma(b)} + \epsilon'.$$

Given $L > 0$ and $k \geq 0$ (their size will be determined later), we introduce the matrix $W_N^{k,L} = \frac{1}{\sqrt{1+\frac{k}{L}}} W_N$ and $M_N^{k,L} = A_N + \frac{1}{\sqrt{N}} W_N^{k,L}$. We also define

$$\sigma_{k,L} = \frac{1}{\sqrt{1 + \frac{k}{L}}}\sigma, \quad a_{k,L} = z_{\sigma_{k,L}}(g_\sigma(a)) \text{ and } b_{k,L} = z_{\sigma_{k,L}}(g_\sigma(b))$$

where we recall that $z_{\sigma_{k,L}}(g) = \frac{1}{g} + \sigma_{k,L}^2 g$. Note that for all $L > 0$, one has: $M_N^{0,L} = M_N$, $a_{0,L} = a$ and $b_{0,L} = b$.
We first choose the size of L as follows. We take L_0 large enough such that for all $L \geq L_0$,

$$\max\left(\frac{1}{L}\sigma^2(|g_\sigma(a)| + |g_\sigma(b)|); \frac{3\sigma}{L}\right) < \frac{b-a}{4} \tag{4.35}$$

¿From the very definition of the $a_{k,L}$'s and $b_{k,L}$'s, one can easily see that $b_{k,L} - a_{k,L} \geq b - a$ (using the last point of (ii) in Lemma 4.5) and that this choice of L_0 ensures that, for all $L \geq L_0$ and for all $k \geq 0$,

$$|a_{k+1,L} - a_{k,L}| < \frac{b-a}{4} \quad \text{and} \quad |b_{k+1,L} - b_{k,L}| < \frac{b-a}{4} \tag{4.36}$$

Now, we fix L such that $L \geq L_0$ and we write $a_k = a_{k,L}$, $b_k = b_{k,L}$ and $\sigma_k = \sigma_{k,L}$.

Lemma 4.6 first gives that

$$\lambda_{i_N+1}(M_N^{k,L}) \leq a_k - \epsilon' - \sigma_k^2 g_\sigma(a) + \frac{1}{\sqrt{1+\frac{k}{L}}}\lambda_1(\frac{1}{\sqrt{N}}W_N) \quad \text{for} \quad i_N < N$$

$$\text{and} \quad \lambda_{i_N}(M_N^{k,L}) \geq b_k + \epsilon' - \sigma_k^2 g_\sigma(b) + \frac{1}{\sqrt{1+\frac{k}{L}}}\lambda_N(\frac{1}{\sqrt{N}}W_N) \quad \text{for} \quad i_N > 0.$$

Furthermore, according to (2.9), the two first extremal eigenvalues of $\frac{1}{\sqrt{N}}W_N$ are such that, almost surely, at least for N large enough

$$0 < \max(-\lambda_N(\frac{1}{\sqrt{N}}W_N), \lambda_1(\frac{1}{\sqrt{N}}W_N)) < 3\sigma.$$

Thus, for all k, almost surely, at least for N large enough (N does not depend on k),

$$0 < \frac{1}{\sqrt{1+\frac{k}{L}}} \times \max(-\lambda_N(\frac{1}{\sqrt{N}}W_N), \lambda_1(\frac{1}{\sqrt{N}}W_N)) < 3\sigma_k.$$

As $\sigma_k \to 0$ when $k \to +\infty$, there is K large enough such that for all $k \geq K$,

$$\max(|3\sigma_k - \sigma_k^2 g_\sigma(a)|, |3\sigma_k + \sigma_k^2 g_\sigma(b)|) < \epsilon'$$

and then a.s for N large enough

$$\lambda_{i_N+1}(M_N^{k,L}) < a_k \quad \text{if} \ \ i_N < N, \tag{4.37}$$

$$\lambda_{i_N}(M_N^{k,L}) > b_k \quad \text{if} \ \ i_N > 0 \tag{4.38}$$

Since (4.37) respectively (4.38) are obviously satisfied too for $i_N = N$ resp. $i_N = 0$, we have established that for any $i_N \in \{0, \ldots, N\}$, for all $k \geq K$,

$$\mathbb{P}\big[\lambda_{i_N+1}(M_N^{k,L}) < a_k \text{ and } \lambda_{i_N}(M_N^{k,L}) > b_k \text{ for all large } N\big] = 1.$$

In particular,

$$\mathbb{P}\big[\lambda_{i_N+1}(M_N^{K,L}) < a_K \text{ and } \lambda_{i_N}(M_N^{K,L}) > b_K \text{ for all large } N\big] = 1. \tag{4.39}$$

Now, we shall show that with probability 1: for N large, $[a_K, b_K]$ and $[a, b]$ split the eigenvalues of, respectively, $M_N^{K,L}$ and M_N having equal amount of eigenvalues to the left sides of the intervals. To this aim, we will proceed by induction on k and show that, for all $k \geq 0$, $[a_k, b_k]$ and $[a, b]$ split the eigenvalues of $M_N^{k,L}$ and M_N (recall that $M_N = M_N^{0,L}$) in exactly the same way. To begin, let us consider for all $k \geq 0$, the set

$$\mathrm{E}_k = \{\text{no eigenvalues of } M_N^{k,L} \text{ in } [a_k, b_k], \text{ for all large } N\}.$$

By Lemma 4.5 (ii) and Theorem 4.4, we know that $\mathbb{P}(\mathrm{E}_k) = 1$ for all k. In particular, from the fact that $\mathbb{P}(\mathrm{E}_0) = 1$, one has for all $\omega \in \mathrm{E}_0$, for all large N:

$$\exists j_N(\omega) \in \{0, \cdots, N\} \text{ such that } \lambda_{j_N(\omega)+1}(M_N) < a \text{ and } \lambda_{j_N(\omega)}(M_N) > b. \tag{4.40}$$

Extending the random variable j_N by setting for instance $j_N := -1$ on ${}^c\mathrm{E}_0$, we want to show that for all k,

$$\mathbb{P}[\lambda_{j_N+1}(M_N^{k,L}) < a_k \text{ and } \lambda_{j_N}(M_N^k) > b_k, \text{ for all large } N] = 1. \tag{4.41}$$

This can be done by induction calling, once more time, on Lemma 4.6. By (4.40), this is true for $k = 0$. Now, let us assume that (4.41) holds true. We shall show that this still holds replacing k by $k+1$. One has

$$M_N^{k+1,L} = M_N^{k,L} + \Big(\frac{1}{\sqrt{1+\frac{k+1}{L}}} - \frac{1}{\sqrt{1+\frac{k}{L}}}\Big)\frac{1}{\sqrt{N}}W_N$$

so, by Lemma 4.6,

$$\lambda_{j_N+1}(M_N^{k+1,L}) \leq \lambda_{j_N+1}(M_N^{k,L}) + \frac{1}{L}\Big(-\lambda_N\big(\frac{1}{\sqrt{N}}W_N\big)\Big).$$

But, for N large enough, $0 < -\lambda_N(\frac{1}{\sqrt{N}}W_N) \leq 3\sigma$ a.s, so by the choice (4.35) on L,

$$\lambda_{j_N+1}(M_N^{k+1,L}) < a_k + \frac{b-a}{4} := \hat{a}_k.$$

Similarly, one can show that

$$\text{a.s.} \quad \lambda_{j_N}(M_N^{k+1,L}) > b_k - \frac{b-a}{4} := \hat{b}_k.$$

Now, by (4.36), one readily observes $\hat{a}_k - a_{k+1} = a_k - a_{k+1} + \frac{b-a}{4} > 0$ and similarly that $\hat{b}_k - b_{k+1} < 0$ which implies that

$$[\hat{a}_k, \hat{b}_k] \subset [a_{k+1}, b_{k+1}].$$

So, as $\mathbb{P}(\mathrm{E}_{k+1}) = 1$, we deduce that with probability 1:

$$\lambda_{j_N+1}(M_N^{k+1,L}) < a_{k+1} \text{ and } \lambda_{j_N}(M_N^{k+1,L}) > b_{k+1}, \quad \text{for all } N \text{ large enough.}$$

As a consequence, (4.41) holds for all $k \geq 0$ and in particular for $k = K$. Comparing this with (4.39), we deduce that $j_N = i_N$ a.s. and

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

This ends the proof of Theorem 4.5. □

Now, we are in position to prove the main Theorem 2.1.

Proof of Theorem 2.1: Our reasoning is close to the last Section 4 of [Bk-S1]. It is enough to establish parts (a) and (b) since the assertions (c) and (d) can then be deduced by taking -M_N instead of M_N.
The proof of (a) is mainly based on successive applications of Theorem 4.5. Fix an integer $1 \leq j \leq J_{+\sigma}$, and let us consider for $\epsilon > 0$, the interval $[a, b] = [\rho_{\theta_j} + \epsilon, \rho_{\theta_{j-1}} - \epsilon]$ which is included in the union (4.31) (at least for ϵ small enough). We define $K_j = k_1 + \cdots + k_j$ (with $\theta_0 := +\infty$ and the convention $\lambda_0(M_N) = \lambda_0(A_N) = +\infty$). Since $1/g_\sigma(\rho_{\theta_k}) = \theta_k$ for $k = j-1$ and j and since the function $1/g_\sigma$ is continuous and increasing on $[a, b]$, the compact interval $[a, b]$ satisfies (4.33) with $i_N = K_{j-1}$ (with the convention that $i_N = 0$ if $j = 1$). Hence, by Theorem 4.5, one has

$$\mathbb{P}[\lambda_{K_{j-1}}(M_N) \geq \rho_{\theta_{j-1}} - \epsilon \text{ and } \lambda_{K_{j-1}+1}(M_N) \leq \rho_{\theta_j} + \epsilon, \text{ for } N \text{ large}] = 1.$$

Similar arguments imply that, for all $j \in \{1, \ldots, J_{+\sigma} - 1\}$,

$$\mathbb{P}[\lambda_{K_j}(M_N) \geq \rho_{\theta_j} - \epsilon \text{ and } \lambda_{K_j+1}(M_N) \leq \rho_{\theta_{j+1}} + \epsilon, \text{ for } N \text{ large}] = 1.$$

As a result, we deduce that for all $1 \leq j \leq J_{+\sigma} - 1$,

$$\mathbb{P}[\rho_{\theta_j} - \epsilon \leq \lambda_{K_j}(M_N) \leq \cdots \leq \lambda_{K_{j-1}+1}(M_N) \leq \rho_{\theta_j} + \epsilon, \text{ for } N \text{ large}] = 1. \tag{4.42}$$

So, letting ϵ go to zero, we obtain (a) for each integer j of $\{1, \cdots, J_{+\sigma} - 1\}$.
Let us now quickly consider the case where $j = J_{+\sigma}$. Note first that, from the preceding discussion, we still have (for ϵ small enough)

$$\mathbb{P}[\lambda_{K_{J_{+\sigma}-1}+1}(M_N) \leq \rho_{\theta_{J_{+\sigma}}} + \epsilon, \text{ for } N \text{ large}] = 1.$$

Then, using the fact that $1/g_\sigma$ increases continuously on $]2\sigma, +\infty[$ with $1/g_\sigma(]2\sigma, +\infty[) =]\sigma, +\infty[$, one can show that once $\epsilon > 0$ is small enough, the compact set $[a, b] = [2\sigma + \epsilon, \rho_{\theta_{J_{+\sigma}}} - \epsilon]$ satisfies the assumptions of Theorem 4.5 with $i_N = K_{J_{+\sigma}}$. This leads to

$$\mathbb{P}[\lambda_{K_{J_{+\sigma}}}(M_N) \geq \rho_{\theta_{J_{+\sigma}}} - \epsilon \text{ and } \lambda_{K_{J_{+\sigma}}+1}(M_N) \leq 2\sigma + \epsilon, \text{ for } N \text{ large}] = 1.$$

Thus, letting $\epsilon \to 0$, we deduce that (4.42) holds for $j = J_{+\sigma}$ and the assertion (a) is established. For point (b), it remains to prove that

$$\liminf_N \lambda_{K_{J_{+\sigma}}+1}(M_N) \geq 2\sigma \quad \text{a.s.}$$

Such an inequality follows from the fact that the spectral measure of M_N converges almost surely towards the semicircle law μ_{sc} which is compactly supported in $[-2\sigma, 2\sigma]$. This completes the proof of Theorem 2.1. □

5 Fluctuations

The (complex or real) Wigner matricial models under consideration are the same as previously (i.e defined by (i) in Section 2) but now we assume that the perturbation A_N is diagonal with unique eigenvalue $\theta > \sigma$: $A_N = \mathrm{diag}(\theta, 0, \cdots, 0)$. According to the previous section, the a.s convergence of $\lambda_1(M_N)$ towards ρ_θ is universal in the sense that it does not depend on μ. In this section, we are going to show that the fluctuations of $\lambda_1(M_N)$ around this universal limit are not universal any more. Indeed, we are going to prove that $\sqrt{N}(1 - \frac{\sigma^2}{\theta^2})^{-1}(\lambda_1(M_N) - \rho_\theta)$ converges in distribution towards the convolution of μ and a Gaussian distribution. Hence, the limiting distribution clearly varies with μ and in particular cannot be Gaussian unless μ is Gaussian.

5.1 Basic tools

We start with the following results which will be of basic use later on. Note that in the following, a complex random variable x will be said *standardized* if $\mathbb{E}(x) = 0$ and $\mathbb{E}(|x|^2) = 1$.

Theorem 5.1. *(Lemma 2.7 [B-S1]) Let $B = (b_{ij})$ be a $N \times N$ Hermitian matrix and Y_N be a vector of size N which contains i.i.d standardized entries with bounded fourth moment. Then there is a constant $K > 0$ such that*

$$\mathbb{E}|Y_N^* B Y_N - \mathrm{Tr} B|^2 \leq K \mathrm{Tr}(BB^*).$$

Theorem 5.2. *(cf. [B-Ya2] or Appendix by J. Baik and J. Silverstein) Let $B = (b_{ij})$ be a $N \times N$ random Hermitian matrix and $Y_N = (y_1, \ldots, y_N)$ be an independent vector of size N which contains i.i.d standardized entries with bounded fourth moment and such that $\mathbb{E}(y_1^2) = 0$ if y_1 is complex. Assume that*

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

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

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

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

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

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

Proof: This result is in fact a particular case of a more general result of [B-Ya2] (Theorems 7.1 and 7.2) which follows from the method of moments. We give an alternative elegant proof by J. Baik and J. Silverstein in the Appendix of the present paper.

Theorem 5.3. *(Theorem 1.1 in [B-Ya1]) Let* f *be an analytic function on an open set of the complex plane including* $[-2\sigma; 2\sigma]$*. If the entries* $((W_N)_{ij})_{1\leq i\leq j\leq N}$ *of a general Wigner matrix* W_N *satisfy the conditions*

- $\mathbb{E}(|(W_N)_{ij}|^4) = M$ *for* $i \neq j$,
- *For any* $\eta > 0$, $\lim_{N\to+\infty} \frac{1}{\eta^4 n^2}\sum_{i,j}\mathbb{E}\left[|(W_N)_{ij}|^4 1_{\{|(W_N)_{ij}|\geq \eta\sqrt{n}\}}\right] = 0$,

then $N\{\mathrm{tr}_N(f(\frac{1}{\sqrt{N}}W_N)) - \int f d\mu_{sc}\}$ *converges in distribution towards a Gaussian variable.*

In our setting, μ satisfies a Poincaré inequality and thus, as already noticed in Section 2, μ satisfies $\int |x|^q d\mu(x) < +\infty$ for any q in $\mathbb{N}$. Hence, the general Wigner matrices we consider obviously satisfy the conditions of Theorem 5.3. Nevertheless, in the following study of fluctuations, we do not use the Poincaré inequality; thus one can expect that Theorem 2.2 is still valid under assumptions on the only first moments of μ provided one can prove the a.s convergence of $\lambda_1(M_N)$ towards ρ_θ under these weaker assumptions.

5.2 Proof of Theorem 2.2

The approach is the same for the complex and real settings and is close to the one of [P] and the ideas of [B-B-P]. Let $\widehat{M}_{N-1}$ be the $N-1\times N-1$ matrix obtained from M_N removing the first row and the first column. Thus, $\frac{\sqrt{N}}{\sqrt{N-1}}\widehat{M}_{N-1}$ is a non-Deformed Wigner matrix associated with the measure μ. We will denote by $\lambda_1(\widehat{M}_{N-1})$ the largest eigenvalue and by $\lambda_{N-1}(\widehat{M}_{N-1})$ the lowest eigenvalue of the matrix $\widehat{M}_{N-1}$. Let $0 < \delta < \frac{\rho_\theta - 2\sigma}{4}$. Let us define the event

$$\Omega_N = \left\{\lambda_1(\widehat{M}_{N-1}) \leq 2\sigma + \delta; \lambda_{N-1}(\widehat{M}_{N-1}) \geq -2\sigma - \delta; \lambda_1(M_N) \geq \rho_\theta - \delta\right\}.$$

According to [B-Yi] and Theorem 2.1 of the previous section, $\lim_{N\to+\infty}\mathbb{P}(\Omega_N) = 1$. Thus, it is sufficient to restrict ourselves to the event Ω_N in order to study the convergence in distribution of $\sqrt{N}(\lambda_1(M_N) - \rho_\theta)$.

Let $V = {}^t(v_1, \ldots, v_N)$ be an eigenvector corresponding to $\lambda_1(M_N)$.

$$M_N V = \lambda_1(M_N)V \Longleftrightarrow \left\{\begin{array}{l} \theta v_1 + \frac{(W_N)_{11}}{\sqrt{N}}v_1 + \sum_{j=2}^N (M_N)_{1j}v_j = \lambda_1(M_N)v_1 \\ (M_N)_{i1}v_1 + \sum_{j=2}^N (\widehat{M}_{N-1})_{ij}v_j = \lambda_1(M_N)v_i, \ \ \forall i = 2,\ldots,N. \end{array}\right.$$

Define the following vectors in $\mathbb{C}^{N-1}$:

$$\widehat{V} = {}^t(v_2,\ldots,v_N)$$

and

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

On Ω_N, $\lambda_1(M_N)$ is not an eigenvalue of $\widehat{M}_{N-1}$ and one can write the eigen-equations using the resolvent $\widehat{G}(\lambda_1(M_N)) := (\lambda_1(M_N)I_{N-1} - \widehat{M}_{N-1})^{-1}$ as follows:

$$\widehat{V} = v_1\widehat{G}(\lambda_1(M_N))\check{M}_{\cdot 1}. \tag{5.1}$$

$$\lambda_1(M_N)v_1 = \theta v_1 + \frac{(W_N)_{11}}{\sqrt{N}}v_1 + v_1\check{M}_{\cdot 1}^*\widehat{G}(\lambda_1(M_N))\check{M}_{\cdot 1}. \tag{5.2}$$

Since v_1 is obviously non equal to zero, one gets from (5.2)

$$\lambda_1(M_N) = \theta + \frac{(W_N)_{11}}{\sqrt{N}} + \check{M}_{\cdot 1}^* \widehat{G}(\lambda_1(M_N))\check{M}_{\cdot 1}. \tag{5.3}$$

On Ω_N, ρ_θ is not an eigenvalue of $\widehat{M}_{N-1}$ and the resolvent $\widehat{G}(\rho_\theta) := (\rho_\theta I_{N-1} - \widehat{M}_{N-1})^{-1}$ is well defined too. Thus, (5.3) is equivalent to

$$\lambda_1(M_N) - \rho_\theta = \frac{(W_N)_{11}}{\sqrt{N}} + \check{M}_{\cdot 1}^* \widehat{G}(\rho_\theta)\check{M}_{\cdot 1} - \frac{\sigma^2}{\theta} + \check{M}_{\cdot 1}^* \left[\widehat{G}(\lambda_1(M_N)) - \widehat{G}(\rho_\theta)\right] \check{M}_{\cdot 1}. \tag{5.4}$$

Using $\widehat{G}(\lambda_1(M_N)) - \widehat{G}(\rho_\theta) = -(\lambda_1(M_N) - \rho_\theta)\widehat{G}(\rho_\theta)\widehat{G}(\lambda_1(M_N))$ and $g_\sigma(\rho_\theta) = \frac{1}{\theta}$, one gets

$$\begin{aligned}
&\lambda_1(M_N) - \rho_\theta \\
&\quad = \frac{(W_N)_{11}}{\sqrt{N}} + \check{M}_{\cdot 1}^* \widehat{G}(\rho_\theta)\check{M}_{\cdot 1} - \sigma^2 g_\sigma(\rho_\theta) \\
&\qquad + \check{M}_{\cdot 1}^* \left[-(\lambda_1(M_N) - \rho_\theta)\widehat{G}(\rho_\theta)\left(\widehat{G}(\rho_\theta) - (\lambda_1(M_N) - \rho_\theta)\widehat{G}(\rho_\theta)\widehat{G}(\lambda_1(M_N))\right)\right] \check{M}_{\cdot 1}.
\end{aligned}$$

Finally, defining $f_\theta(z) := \dfrac{1}{\rho_\theta - z} 1_{|z| \leq 2\sigma + \delta}$, we can easily deduce from the previous equality the following identity on Ω_N:

$$\{1 + c_N + \delta_1(N) + \delta_2(N)\} \sqrt{N}(\lambda_1(M_N) - \rho_\theta) = (W_N)_{11} + \sqrt{\frac{N}{N-1}} d_N + \sqrt{\frac{N}{N-1}} \delta_3(N)$$

where

$$\begin{aligned}
c_N &= \sigma^2 tr_{N-1}\left[f_\theta^2(\widehat{M}_{N-1})\right] \\
d_N &= \sqrt{N-1}\left\{\check{M}_{\cdot 1}^*\left(\widehat{G}(\rho_\theta) 1_{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta}\right)\check{M}_{\cdot 1} - \sigma^2 tr_{N-1}\left(\widehat{G}(\rho_\theta) 1_{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta}\right)\right\} \\
\delta_1(N) &= -(\lambda_1(M_N) - \rho_\theta)\check{M}_{\cdot 1}^* \left[\widehat{G}(\rho_\theta)\right]^2 \widehat{G}(\lambda_1(M_N))\check{M}_{\cdot 1} 1_{\Omega_N} \\
\delta_2(N) &= \check{M}_{\cdot 1}^* \left[\widehat{G}(\rho_\theta) 1_{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta}\right]^2 \check{M}_{\cdot 1} - \sigma^2 tr_{N-1}\left[\widehat{G}(\rho_\theta) 1_{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta}\right]^2 \\
\delta_3(N) &= \sigma^2 \sqrt{N-1}\left\{tr_{N-1}\left(f_\theta(\widehat{M}_{N-1})\right) - \int f_\theta d\mu_{sc}\right\}.
\end{aligned}$$

First

$$\begin{aligned}
|\delta_1(N)| &\leq |\lambda_1(M_N) - \rho_\theta| \|\check{M}_{\cdot 1}\|^2 \|\widehat{G}(\rho_\theta)\|^2 \|\widehat{G}(\lambda_1(M_N))\| 1_{\Omega_N} \\
&\leq \frac{1}{(\rho_\theta - 2\sigma - 2\delta)(\rho_\theta - 2\sigma - \delta)^2}\left\{\frac{1}{N}\sum_{j=2}^N |(W_N)_{j1}|^2\right\} |\lambda_1(M_N) - \rho_\theta|,
\end{aligned}$$

(using point (v) of Lemma 3.2). By the law of large numbers $\frac{1}{N}\sum_{j=2}^N |(W_N)_{j1}|^2$ converges a.s. towards σ^2 and according to Theorem 2.1, $|\lambda_1(M_N) - \rho_\theta|$ converges a.s. to zero. Hence $\delta_1(N)$ converges obviously in probability towards zero.

Now, since f_θ is analytic on an open set including $[-2\sigma; 2\sigma]$, we deduce from Theorem 5.3 the convergence in probability of $\delta_3(N)$ towards zero and of c_N towards $\sigma^2 \int f_\theta^2 d\mu_{sc} = \dfrac{\sigma^2}{\theta^2 - \sigma^2}$.

According to Theorem 5.1 and using Lemma 3.2(v),

$$\begin{aligned}
\mathbb{E}(|\delta_2(N)|^2) &\leq \frac{K}{N-1}\mathbb{E}\left(\mathrm{tr}_N\left[\widehat{G}(\rho_\theta) 1_{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta}\right]^4\right) \\
&\leq \frac{K}{N-1}\mathbb{E}\left(\|\widehat{G}(\rho_\theta)\|^4 1_{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta}\right) \\
&\leq \frac{K}{N-1}\frac{1}{(\rho_\theta - 2\sigma - \delta)^4}.
\end{aligned}$$

The convergence in probability of $\delta_2(N)$ towards zero readily follows by Tchebychev inequality.

Let us check that $\widehat{G}(\rho_\theta)1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}$ satisfies the conditions of Theorem 5.2.

(i) $\|\widehat{G}(\rho_\theta)1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}\| \leq \frac{1}{\rho_\theta - 2\sigma-\delta}$ by Lemma 3.2 (v).

(ii) According to Theorem 5.3, $\mathrm{tr}_{N-1}\, f_\theta^2(\widehat{M}_{N-1})$ converges in probability towards $\int f_\theta^2 d\mu_{sc}$. Since on the event $\{\|\widehat{M}_{N-1}\| \leq 2\sigma+\delta\}$, with limiting probability 1, $\mathrm{tr}_{N-1}[\widehat{G}(\rho_\theta)1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}]^2$ coincide with $\mathrm{tr}_{N-1}\, f_\theta^2(\widehat{M}_{N-1})$, it also converges in probability towards $\int f_\theta^2 d\mu_{sc}$.

(iii) It is proved in Proposition 3.1 in [C-D] that for any $z \in \mathbb{C}$ such that $\Im m z > 0$, $\frac{1}{N-1}\sum_{i=1}^{N-1}([\widehat{G}(z)]_{ii})^2$ converges in probability towards $g_\sigma^2(z)$. The same result holds for $\frac{1}{N-1}\sum_{i=1}^{N-1}([\widehat{G}(z)]_{ii})^2 1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}$. For any $\epsilon > 0$ and any $\alpha > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\left|\tfrac{1}{N-1}\textstyle\sum_{i=1}^{N-1}([\widehat{G}(\rho_\theta)]_{ii})^2 1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta} - g_\sigma^2(\rho_\theta)\right| > \epsilon\right) \\
&\quad\leq \mathbb{P}\left(|\frac{1}{N-1}\sum_{i=1}^{N-1}\left\{([\widehat{G}(\rho_\theta)]_{ii})^2 - ([\widehat{G}(\rho_\theta+i\alpha)]_{ii})^2\right\}1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}| > \frac{\epsilon}{3}\right) \\
&\qquad + \mathbb{P}\left(|\frac{1}{N-1}\sum_{i=1}^{N-1}([\widehat{G}(\rho_\theta+i\alpha)]_{ii})^2 1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta} - g_\sigma^2(\rho_\theta+i\alpha)| > \frac{\epsilon}{3}\right) \\
&\qquad + \mathbb{P}\left(|g_\sigma^2(\rho_\theta) - g_\sigma^2(\rho_\theta+i\alpha)| > \frac{\epsilon}{3}\right).
\end{aligned}$$

Since

$$\begin{aligned}
&\left\{([\widehat{G}(\rho_\theta)]_{ii})^2 - ([\widehat{G}(\rho_\theta+i\alpha)]_{ii})^2\right\}1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta} \\
&\qquad = [\widehat{G}(\rho_\theta) - \widehat{G}(\rho_\theta+i\alpha)]_{ii}[\widehat{G}(\rho_\theta)+\widehat{G}(\rho_\theta+i\alpha)]_{ii}1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta} \\
&\qquad = i\alpha[\widehat{G}(\rho_\theta)\widehat{G}(\rho_\theta+i\alpha)]_{ii}[\widehat{G}(\rho_\theta)+\widehat{G}(\rho_\theta+i\alpha)]_{ii}1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta},
\end{aligned}$$

we get by using Lemma 3.2 (v)

$$|([\widehat{G}(\rho_\theta)]_{ii})^2 - ([\widehat{G}(\rho_\theta+i\alpha)]_{ii})^2|1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta} \leq \frac{2\alpha}{(\rho_\theta-2\sigma-\delta)^3}.$$

Similarly, we get that

$$|g_\sigma^2(\rho_\theta) - g_\sigma^2(\rho_\theta+i\alpha)| \leq \frac{2\alpha}{(\rho_\theta-2\sigma)^3}.$$

Thus, choosing α such that $\frac{2\alpha}{(\rho_\theta-2\sigma-\delta)^3} < \frac{\epsilon}{3}$, we readily deduce the convergence in probability of

$$\frac{1}{N-1}\sum_{i=1}^{N-1}([\widehat{G}(\rho_\theta)]_{ii})^2 1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}$$

towards $g^2(\rho_\theta)$.

Since $\widehat{G}(\rho_\theta)1_{\|\widehat{M}_{N-1}\|\leq 2\sigma+\delta}$ and $\check{M}_{\cdot 1}$ are independent, we can deduce from Theorem 5.2 that d_N converges in distribution towards a Gaussian law with mean zero and variance

$$v_\theta := \sigma^4\left\{(\mathbb{E}(|\frac{(W_N)_{12}}{\sigma}|^4) - 1 - t/2)\frac{1}{\theta^2}) + \frac{t}{2}\frac{1}{\theta^2-\sigma^2}\right\}$$

where $t = 4$ in the real setting and $t = 2$ in the complex one. One readily verifies that v_θ satisfies (2.12).

Let $0 < \epsilon < 1$. Since $\delta_1(N) + \delta_2(N)$ converges in probability towards zero, the probability of the event $\widetilde{\Omega}_N = \Omega_N \cap \{|\delta_1(N) + \delta_2(N)| \leq \epsilon\}$ tends to 1. Now, since $c_N \geq 0$ we have the following identity on $\widetilde{\Omega}_N$:

$$\sqrt{N}(\lambda_1(M_N) - \rho_\theta) = \frac{1}{u_N}\left\{(W_N)_{11} + \sqrt{\frac{N}{N-1}}d_N + \sqrt{\frac{N}{N-1}}\delta_3(N)\right\}$$

with $u_N := 1 + c_N + \delta_1(N) + \delta_2(N)$ converging in distribution towards $(1 - \frac{\sigma^2}{\theta^2})^{-1}$. Moreover, since $(W_N)_{11}$ and d_N are independent, $(W_N)_{11} + \sqrt{\frac{N}{N-1}}d_N + \sqrt{\frac{N}{N-1}}\delta_3(N)$ converges in distribution towards the convolution of μ and a Gaussian distribution $\mathcal{N}(0, v_\theta)$. Finally, we can conclude that $\sqrt{N}(\lambda_1(M_N) - \rho_\theta)$ converges in distribution towards $(1 - \frac{\sigma^2}{\theta^2})\,\{\mu * \mathcal{N}(0, v_\theta)\}$. $\square$

Appendix by J. Baik and J. Silverstein

This Appendix presents the proof by J. Baik and J. Silverstein of the CLT (given by Theorem 5.2) needed in the previous section for the proof of Theorem 2.2. Their proof is based on a writing of the expression

$$(1/\sqrt{N})(Y_N^* B Y_N - \mathrm{Tr} B) \tag{A.1}$$

as a sum of martingale difference, and uses the following CLT.

Theorem A.1. *(Theorem 35.12 of [Bi]) For each N, suppose $Z_{N1}, Z_{N2}, \ldots, Z_{Nr_N}$ is a real martingale difference sequence with respect to the increasing σ-field $\{\mathcal{F}_{N,j}\}$ having second moments. If as $N \to \infty$,*

$$\sum_{j=1}^{r_N} \mathbb{E}(Z_{Nj}^2 | \mathcal{F}_{N,j-1}) \xrightarrow{P} v^2 \tag{A.2}$$

where v^2 is a positive constant, and for each $\epsilon > 0$,

$$\sum_{j=1}^{r_N} \mathbb{E}(Z_{Nj}^2\, 1_{|Z_{Nj}| \geq \epsilon}) \to 0 \tag{A.3}$$

then

$$\sum_{j=1}^{r_N} Z_{Nj} \xrightarrow{\mathcal{L}} \mathcal{N}(0, v^2).$$

Proof of Theorem 5.2: First, one can write $(A.1)$ as a sum of martingale differences:

$$(1/\sqrt{N})(Y_N^* B Y_N - \mathrm{Tr} B) = (1/\sqrt{N}) \sum_{i=1}^{N} \left((|y_i|^2 - 1) b_{ii} + \bar{y}_i \sum_{j<i} y_j b_{ij} + \bar{y}_i \sum_{j>i} y_j b_{ij} \right)$$

$$= (1/\sqrt{N}) \sum_{i=1}^{N} \left((|y_i|^2 - 1) b_{ii} + \bar{y}_i \sum_{j<i} y_j b_{ij} + y_i \sum_{j<i} \bar{y}_j b_{ji} \right) = \sum_{i=1}^{N} Z_i$$

where

$$Z_i = Z_{Ni} = (1/\sqrt{N})((|y_i|^2 - 1) b_{ii} + \bar{y}_i \sum_{j<i} y_j b_{ij} + y_i \sum_{j<i} \bar{y}_j \bar{b}_{ij}).$$

Let $\mathcal{F}_{N,i}$ (resp. $\mathcal{F}_{N,0}$) be the σ-field generated by $y_1, \ldots, y_i$ and B (resp. by B). Let also $\mathbb{E}_i(\cdot)$ denote conditional expectation with respect to $\mathcal{F}_{N,i}$.
It is clear that Z_i is measurable with respect to $\mathcal{F}_{N,i}$ and satisfies $\mathbb{E}_{i-1}(Z_i) = 0$.

We will show the conditions of Theorem A.1 are met.

To verify the Lindeberg condition (A.3), we need to show this property is closed under addition. This will follow from the following fact. For random variables X_1, X_2, and positive ϵ

$$\mathbb{E}(|X_1+X_2|^2\, 1_{(|X_1+X_2|\geq\epsilon)}) \quad\leq\quad 4\big(\mathbb{E}(|X_1|^2\, 1_{(|X_1|\geq\epsilon/2)})+\mathbb{E}(|X_2|^2\, 1_{(|X_2|\geq\epsilon/2)})\big). \tag{A.4}$$

Indeed, we have

$$\begin{aligned}\mathbb{E}(|X_1|^2\, 1_{(|X_1+X_2|\geq\epsilon)}) &\leq \mathbb{E}(|X_1|^2\, 1_{(|X_1|\geq\epsilon/2)})+\mathbb{E}(|X_1|^2\, 1_{(|X_1|<\epsilon/2,|X_2|\geq\epsilon/2)})\\ &\leq \mathbb{E}(|X_1|^2\, 1_{(|X_1|\geq\epsilon/2)})+(\epsilon^2/4)\mathbb{P}(|X_2|\geq\epsilon/2)\\ &\leq \mathbb{E}(|X_1|^2\, 1_{(|X_1|\geq\epsilon/2)})+\mathbb{E}(|X_2|^2\, 1_{(|X_2|\geq\epsilon/2)}).\end{aligned}$$

The same bound starting with X_2 leads (A.4).

Write $Z_i = X_1^i + X_2^i$, with $X_1^i = (1/\sqrt{N})(|y_i|^2-1)b_{ii}$. Then for $\epsilon>0$,

$$\sum_{i=1}^N \mathbb{E}(|X_1^i|^2\, 1_{(|X_1^i|\geq\epsilon)}) \leq a^2\mathbb{E}((|y_1|^2-1)^2\, 1_{(||x_1|^2-1|\geq\sqrt{N}\epsilon/a)}) \to 0 \tag{A.5}$$

as $N\to\infty$, by dominated convergence theorem.

We have

$$\begin{aligned}\mathbb{E}|\sum_{j<i} y_j b_{ij}|^4 &= \mathbb{E}(|y_1|^4\sum_{j<i}|b_{ij}|^4)+2\mathbb{E}(\sum_{*}|b_{ij_1}|^2|b_{ij_2}|^2)+\mathbb{E}(|y_1^2|^2\sum_{*} b_{ij_1}^2\bar{b}_{ij_2}^2)\\ &\leq \mathbb{E}|y_1|^4\mathbb{E}[\max_j (B^2)_{jj}(B^2)_{ii}]+(2+\mathbb{E}|y_1^2|^2)\mathbb{E}[(B^2)_{ii}^2]\\ &\leq a^4\left[\mathbb{E}|y_1|^4+2+\mathbb{E}|y_1^2|^2\right]\end{aligned}$$

where the sum $\sum_*$ is over $\{j_1<i,\, j_2<i,\, j_1\neq j_2\}$. Therefore $\mathbb{E}|X_2^i|^4 = o(N^{-1})$ so that for any $\epsilon>0$,

$$\sum_{i=1}^N \mathbb{E}(|X_2^i|^2\, 1_{(|X_2^i|\geq\epsilon)}) \leq (1/\epsilon^2)\sum_{i=1}^N \mathbb{E}|X_2^i|^4 \to 0 \quad \text{as } N\to\infty. \tag{A.6}$$

Thus, by (A.5), (A.6) and (A.4), $\{Z_i\}$ satisfies (A.3).

Now, we shall verify condition (A.2). We have

$$\begin{aligned}\sum_{i=1}^N \mathbb{E}_{i-1}Z_i^2 &= (1/N)\sum_{i=1}^N \Big\{(\mathbb{E}|y_1|^4-1)b_{ii}^2+\mathbb{E}\bar{y}_1^2(\sum_{j<i} y_j b_{ij})^2+\mathbb{E}y_1^2(\sum_{j<i}\bar{y}_j\bar{b}_{ij})^2\\ &\quad +2\mathbb{E}(|y_1|^2\bar{y}_1)b_{ii}\sum_{j<i} y_j b_{ij}+2\mathbb{E}(|y_1|^2 y_1)b_{ii}\sum_{j<i}\bar{y}_j\bar{b}_{ij}+2(\sum_{j<i} y_j b_{ij})(\sum_{j<i}\bar{y}_j\bar{b}_{ij})\Big\}.\end{aligned} \tag{A.7}$$

Let B_L denote the strictly lower triangular part of B. We have

$$\mathbb{E}[(1/N)\sum_{i=1}^N b_{ii}\sum_{j<i} y_j b_{ij}] = 0$$

and using Cauchy-Schwarz

$$\begin{aligned}&\mathbb{E}|(1/N)\sum_{i=1}^N b_{ii}\sum_{j<i} y_j b_{ij}|^2\\ =\ &\mathbb{E}|(1/N)\sum_{j=1}^{N-1} y_j\sum_{i>j} b_{ii}b_{ij}|^2 = (1/N^2)\mathbb{E}(\sum_{j=1}^{N-1}\sum_{i>j} b_{ii}b_{ij}\sum_{\underline{i}>j} b_{\underline{i}\,\underline{i}}\bar{b}_{\underline{i}j}) = (1/N^2)\mathbb{E}(\sum_{i\underline{i}} b_{ii}b_{\underline{i}\,\underline{i}}(B_L B_L^*)_{i\underline{i}})\\ \leq\ &\mathbb{E}\big[(\max_i b_{ii})^2(1/n)(\sum_{i\underline{i}}|(B_L B_L^*)_{i\underline{i}}|^2)^{1/2}\big] = \mathbb{E}\big[(\max_i b_{ii})^2(1/N)\mathrm{Tr}((B_L B_L^*)^2)^{1/2}\big]\\ \leq\ &\mathbb{E}\big[(\max_i b_{ii})^2(1/\sqrt{N})\|B_L\|^2\big].\end{aligned}$$

We apply the following bound (due to R. Mathias, see [Mt]): $\|B_L\| \leq \gamma_N \|B\|$, where $\gamma_N = O(\ln N)$ and the bound $||B|| \leq a$ to conclude that

$$1/N \sum_{i=1}^{N} b_{ii} \sum_{j<i} y_j b_{ij} \xrightarrow{P} 0$$

as $N \to \infty$.
Then (recall that $\mathbb{E}y_1^2 = 0$ when y_1 is complex), (A.7) can be written as

$$\begin{aligned} \sum_{i=1}^{N} \mathbb{E}_{i-1} Z_i^2 &= (1/N) \sum_{i=1}^{N} [(\mathbb{E}|y_1|^4 - 1) b_{ii}^2 + t(\sum_{j<i} y_j b_{ij})(\sum_{j<i} \bar{y}_j \bar{b}_{ij})] + o_P(1) \\ &= (1/N) \sum_{i=1}^{N} (\mathbb{E}|y_1|^4 - 1) b_{ii}^2 + t(1/N) Y_N^* B_L^* B_L Y_N + o_P(1) \end{aligned} \tag{A.8}$$

where $t = 4$ when y_1 real, and is 2 when y_1 is complex.
Besides, from Lemma 2.7 in [B-S1] (recalled in Theorem 5.1) we have

$$\mathbb{E}|(1/N)(Y_N^* B_L^* B_L Y_N - \mathrm{Tr} B_L^* B_L)|^2 \leq (1/N^2) \mathbb{E}(\mathrm{Tr}(B_L^* B_L)^2) \leq K \mathbb{E}\|B\|^4 \frac{\ln^4 N}{N} \to 0$$

as $N \to \infty$. So, as

$$\mathrm{Tr} B_L^* B_L = \sum_{j<i} |b_{ij}|^2 = (1/2)(\mathrm{Tr} B^2 - \sum_i b_{ii}^2)$$

(A.8) implies that condition (A.2) holds with

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

Thus, by Theorem A.1, we deduce that $(1/\sqrt{N})(Y_N^* B Y_N - \mathrm{Tr} B)$ converges in distribution to a Gaussian variable with mean zero and variance v^2. $\square$

References

[A. et al] Ané C., Blachère S., Chafaï D., Fougères P., Gentil I., Malrieu F., Roberto C., Scheffer G., *Sur les inégalités de Sobolev logarithmiques.* (French) [Logarithmic Sobolev inequalities] Panoramas et Synthèses [Panoramas and Syntheses] **10**, S.M.F Paris, 2000.

[A] Arnold L., *On Wigner's semi-circle law for the eigenvalues of random matrices*, J. Math. Anal. Appl. **20**, 262–268 (1967).

[B] Bai Z.D., *Methodologies in spectral analysis of large-dimensional random matrices, a review*, Statist. Sinica **9**, 611–677 (1999).

[B-S1] Bai Z.D. and Silverstein J.W., *No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices*, Ann. Probab. **26**, 316–345 (1998).

[B-S2] Bai Z.D. and Silverstein J.W., *Exact separation of eigenvalues of large dimensional sample covariance matrices*, Ann. Probab. **27**, 1536–1555 (1999).

[B-Yi] Bai Z.D. and Yin Y.Q., *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*, Ann. Probab. **16**, 1729–1741 (1998).

[B-Ya1] Bai Z.D. and Yao J.F., *On the convergence of the spectral empirical process of Wigner matrices*, Bernoulli **11** (6), 1059–1092 (2005).

[B-Ya2] Bai Z.D. and Yao J.F., *Central limit theorems for eigenvalues in a spiked population model*, available at http://perso.univ-rennes1.fr/jian-feng.yao/papers.html, to appear in Ann. Inst. H. Poincaré (2007).

[Bk-B-P] Baik J., Ben Arous G. and Péché S., *Phase transition of the largest eigenvalue for non-null complex sample covariance matrices*, Ann. Probab. **33**, 1643–1697 (2005).

[Bk-S1] Baik J. and Silverstein J.W., *Eigenvalues of large sample covariance matrices of spiked population models*, J. of Multi. Anal. **97**, 1382–1408 (2006).

[Bi] Billingsley P., *Probability and Measure*, 3rd edition Wiley, New York, 1995.

[B-B-P] Biroli G., Bouchaud J.P. and Potters M., *On the top eigenvalue of heavy-tailed random matrices*, ArXiv cond-mat/0609070 (2006).

[B-G] Bobkov S. and Gotze F., *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities*, J. Funct. Anal. **163**, 1–28 (1999).

[C-D] Capitaine M. and Donati-Martin C., *Strong asymptotic freeness of Wigner and Wishart matrices*, Indiana Univ. Math. J. **56**, 767–804 (2007).

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

[F] Feller W., An introduction to probability theory and its applications, Vol. II. Second edition, John Wiley & Sons Inc., New York (1971).

[Fe] Féral D., *Grandes déviations et fluctuations des valeurs propres maximales de matrices aléatoires*, PhD Thesis of University of Toulouse, Available at http://www.lsp.ups-tlse.fr/Fp/Feral (2006).

[Fe-Pe] Féral D. and Péché S., *The largest eigenvalue of rank one deformation of large Wigner matrices*, Arxiv PR/0605624, to appear in Comm. Math. Phys. (2007).

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

[H-T] Haagerup U. and Thorbjørnsen S., *A new application of random matrices: $Ext(C^*_{red}(F_2))$ is not a group*, Ann. Math. **162**, 711–775 (2005).

[H-J] Horn R.A. and Johnson C.R., Matrix Analysis, Cambridge University Press (1991).

[Jo] Johnstone I., *On the distribution of the largest principal component*, Ann. Stat. **29**, 295–327 (2001).

[K-K-P] Khorunzhy A., Khoruzhenko B., Pastur L., *Asymptotic properties of large random matrices with independent entries*, J. Math. Phys. **37**, 5033–5060 (1996).

[L] Ledoux L., The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs, Volume 89, A.M.S, 2001.

[M] Maïda M., *Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles*, ArXiv PR/0609738 (2006).

[Mt] Mathias R., *The Hadamard operator norm of a circulant and applications*, SIAM J. Matrix Anal. Appl. **14**, 1152–1167 (1993).

[P] Paul D., *Asymptotics of the leading sample eigenvalues for a spiked covariance model*, Available at http://www-stat.stanford.edu/ debashis/ (2004).

[Pe] Péché S., *The largest eigenvalues of small rank perturbations of Hermitian random matrices*, Prob. Th. Rel. Fields **134**, 127–174 (2006).

[R] Ruzmaikina A., *Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries*, Comm. Math. Phys. **261** (2), 277–296 (2006).

[S] Schultz H., *Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases.*, Prob. Th. Rel. Fields **131**, 261–309 (2005).

[So] Soshnikov A., *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. **207**, 697–733 (1999).

[T] Tillmann H.G., *Randverteilungen analytischer Funcktionen und Distributionen*, Math. Zeitschr. Bd. **59**, 61–83 (1953).

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