

# Central limit theorems for eigenvalues of deformations of Wigner matrices<sup>1</sup>

M. Capitaine<sup>a</sup>, C. Donati-Martin<sup>b</sup> and D. Féral<sup>c</sup>

<sup>a</sup>CNRS, Institut de Mathématiques de Toulouse, Equipe de Statistique et Probabilités, F-31062 Toulouse Cedex 09, France.

E-mail: [mireille.capitaine@math.univ-toulouse.fr](mailto:mireille.capitaine@math.univ-toulouse.fr)

<sup>b</sup>UPMC Univ. Paris 06 and CNRS, UMR 7599, Laboratoire de Probabilités et Modèles Aléatoires, Case 188, 4 place Jussieu, F-75252 Paris Cedex 05, France. E-mail: [catherine.donati@upmc.fr](mailto:catherine.donati@upmc.fr)

<sup>c</sup>Institut de Mathématiques de Bordeaux, Université Bordeaux 1, 351 Cours de la Libération, F-33405 Talence Cedex, France.

E-mail: [delphine.feral@math.u-bordeaux1.fr](mailto:delphine.feral@math.u-bordeaux1.fr)

Received 24 March 2009; revised 14 December 2010; accepted 27 December 2010

---

**Abstract.** In this paper, we study the fluctuations of the extreme eigenvalues of a spiked finite rank deformation of a Hermitian (resp. symmetric) Wigner matrix when these eigenvalues separate from the bulk. We exhibit quite general situations that will give rise to universality or non-universality of the fluctuations, according to the delocalization or localization of the eigenvectors of the perturbation. Dealing with the particular case of a spike with multiplicity one, we also establish a necessary and sufficient condition on the associated normalized eigenvector so that the fluctuations of the corresponding eigenvalue of the deformed model are universal.

**Résumé.** Dans ce papier, nous étudions les fluctuations des valeurs propres extrêmes d'une matrice de Wigner hermitienne (resp. symétrique) déformée par une perturbation de rang fini dont les valeurs propres non nulles sont fixées, dans le cas où ces valeurs propres extrêmes se détachent du reste du spectre. Nous décrivons des situations générales d'universalité ou de non-universalité des fluctuations correspondant au caractère localisé ou délocalisé des vecteurs propres de la perturbation. Lorsque l'une des valeurs propres de la perturbation est de multiplicité un, nous établissons de plus une condition nécessaire et suffisante sur le vecteur propre associé pour que les fluctuations de la valeur propre correspondante du modèle déformé soient universelles.

MSC: 60B20; 15A18; 60F05

Keywords: Random matrices; Deformed Wigner matrices; Extremal eigenvalues; Fluctuations; Localized eigenvectors; Universality

---

## 1. Introduction

Adding a finite rank perturbation to a GUE matrix, S. Péché [19] pointed out a sharp phase transition phenomenon: according to the largest eigenvalue of the perturbation, the largest eigenvalue of the perturbed matrix should either stick to the bulk and fluctuate according to the Tracy–Widom (or generalized Tracy–Widom) law or should be extracted away from the bulk and have then fluctuations of Gaussian nature. In the lineage of this work, in a previous paper [8], we have studied the limiting behavior of extremal eigenvalues of finite rank deformations of Wigner matrices. We established their almost sure convergence. The limiting values depend only on the spectrum of the deformation  $A_N$  and on the variance of the distribution of the entries of the Wigner matrix. On the contrary the fluctuations of these eigenvalues strongly depend on the eigenvectors of  $A_N$ . Indeed, in the particular case of a rank one diagonal deformation whose non-null eigenvalue is large enough, we established a central limit theorem for the largest eigenvalue

---

<sup>1</sup>This work was partially supported by the Agence Nationale de la Recherche grant ANR-08-BLAN-0311-03.

which deviates from the rest of the spectrum and proved that the fluctuations of the largest eigenvalue vary with the particular distribution of the entries of the Wigner matrix. Thus, this fluctuations result differs from that of the full rank one deformation case investigated in [10] and [12] since this latter case exhibited universal limiting distributions.

Let us recall these results in the complex setting, having in mind that similar results hold in the real symmetric case. In the following, given an arbitrary Hermitian matrix  $M$  of size  $N$ , we will denote by  $\lambda_1(M) \geq \dots \geq \lambda_N(M)$  its  $N$  ordered eigenvalues; we will denote the centered Gaussian distribution with variance  $v$  by  $\mathcal{N}(0, v)$ .

The random matrices under study are complex Hermitian matrices  $(\mathbf{M}_N)_N$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that

$$\mathbf{M}_N = \frac{\mathbf{W}_N}{\sqrt{N}} + \mathbf{A}_N. \quad (1.1)$$

$\mathbf{A}_N$  is a  $N \times N$  deterministic Hermitian matrix of fixed finite rank and whose spectrum does not depend on  $N$ . The matrix  $\mathbf{W}_N$  is a  $N \times N$  Wigner Hermitian matrix such that the  $N^2$  random variables  $(W_N)_{ii}$ ,  $\sqrt{2}\Re((W_N)_{ij})_{i < j}$ ,  $\sqrt{2}\Im((W_N)_{ij})_{i < j}$  are independent identically distributed with a centered distribution  $\mu$  of variance  $\sigma^2$ .

As the rank of the  $\mathbf{A}_N$ 's is assumed to be finite, the Wigner Theorem is still satisfied for the Deformed Wigner model  $(\mathbf{M}_N)_N$  (cf. Lemma 2.2 of [1]): the spectral measure  $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(\mathbf{M}_N)}$  of  $\mathbf{M}_N$  converges a.s. towards the semicircle law  $\mu_{sc}$  whose density is given by

$$\frac{d\mu_{sc}}{dx}(x) = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - x^2} \mathbf{1}_{[-2\sigma, 2\sigma]}(x). \quad (1.2)$$

When  $\mathbf{A}_N \equiv 0$ , it is well-known that once  $\mu$  has a finite fourth moment, the first largest (resp. last smallest) eigenvalues of the rescaled Wigner matrix  $\mathbf{W}_N/\sqrt{N}$  tend almost surely to the right (resp. left) endpoint  $2\sigma$  (resp.  $-2\sigma$ ) of the semicircle support (cf. [1]). The corresponding fluctuations, which have been first obtained by Tracy and Widom [23] in the Gaussian case and then extended by Soshnikov [21] for any symmetric probability measure  $\mu$  having sub-Gaussian moments, are governed by the so-called Tracy–Widom distributions. Note that the exponential decay condition (with symmetry assumption) has been replaced by a finite number of moments in [16,20]. Under the subexponential decay assumption, the symmetry assumption on  $\mu$  in [21] was replaced in [22] by the vanishing third moment condition and very recently, Erdős, Yau and Yin [9] proved the edge universality under the subexponential decay assumption alone.

Let us describe how the asymptotic behavior of the extremal eigenvalues of the perturbed Wigner matrix may be affected by the perturbation by considering the particular case of a rank one perturbation  $\mathbf{A}_N$  with non-null eigenvalue  $\theta$ . For a large class of probability measures  $\mu$ , it turns out that the largest eigenvalue  $\lambda_1(\mathbf{M}_N)$  still tends to the right-endpoint  $2\sigma$  if  $\theta \leq \sigma$  whereas  $\lambda_1(\mathbf{M}_N)$  jumps above the bulk to  $\rho_\theta = \theta + \frac{\sigma^2}{\theta}$  if  $\theta > \sigma$ . This was proved by Pécché in her pionnering work [19] when  $\mu$  is Gaussian, extended in [10] when  $\mu$  is symmetric and has sub-Gaussian moments but in the particular case of the full rank one deformation  $\mathbf{A}_N$  given by

$$(\mathbf{A}_N)_{ij} = \theta/N \quad \text{for all } 1 \leq i, j \leq N \quad (1.3)$$

and finally established in [8] for general  $\mathbf{A}_N$  when  $\mu$  is symmetric and satisfies a Poincaré inequality.

Moreover, considering the perturbation matrix defined by (1.3), Féral and Pécché [10] proved that the fluctuations of  $\lambda_1(\mathbf{M}_N)$  are the same as in the Gaussian setting investigated in [19] and in this sense are universal. Here is their result when  $\theta > \sigma$ :

**Proposition 1.1.** *If  $\mu$  is symmetric and has sub-Gaussian moments*

$$\sqrt{N}(\lambda_1(\mathbf{M}_N) - \rho_\theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\theta^2),$$

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

The proof of this result relies on the computations of moments of  $\mathbf{M}_N$  of high order (depending on  $N$ ) and the knowledge of the fluctuations in the Gaussian case, established by Pécché [19].

On the other hand, for the strongly localized perturbation matrix of rank 1 given by

$$\mathbf{A}_N = \text{diag}(\theta, 0, \dots, 0)$$

with  $\theta > \sigma$ , we proved in [8] that the fluctuations of  $\lambda_1(\mathbf{M}_N)$  vary with the particular distribution of the entries of the Wigner matrix so that this phenomenon can be seen as an example of a non-universal behavior:

**Proposition 1.2.** *Let  $\mu$  be symmetric and satisfy a Poincaré inequality. Define*

$$c_\theta = \frac{\theta^2}{\theta^2 - \sigma^2} \quad \text{and} \quad v_\theta = \frac{1}{2} \left( \frac{m_4 - 3\sigma^4}{\theta^2} \right) + \frac{\sigma^4}{\theta^2 - \sigma^2}, \quad (1.4)$$

where  $m_4 := \int x^4 d\mu(x)$ . Then

$$c_\theta \sqrt{N} (\lambda_1(\mathbf{M}_N) - \rho_\theta) \xrightarrow{\mathcal{L}} \{ \mu * \mathcal{N}(0, v_\theta) \}. \quad (1.5)$$

In the present paper, we consider perturbations  $\mathbf{A}_N$  of higher rank of Wigner matrices associated to some symmetric probability measure  $\mu$  satisfying a Poincaré inequality. The a.s. convergence of the extreme eigenvalues has already been described in [8] (see Theorem 3.1 below). Whenever the largest eigenvalues of  $\mathbf{M}_N$  are extracted away from the bulk, we describe their fluctuations which depend on the localization of the eigenvectors of  $\mathbf{A}_N$ , as already seen in the above rank 1 examples. We investigate two quite general situations for which we exhibit a phenomenon of different nature. To explain this, let us focus on the largest eigenvalue  $\theta_1$  of  $\mathbf{A}_N$ . We assume that  $\theta_1 > \sigma$  so that the largest eigenvalues of  $\mathbf{M}_N$  converges a.s towards  $\rho_{\theta_1} = \theta_1 + \frac{\sigma^2}{\theta_1} > 2\sigma$ .

First, when the eigenvectors associated to the largest eigenvalue  $\theta_1$  of  $\mathbf{A}_N$  are localized, we establish that the limiting distribution in the fluctuations of  $\lambda_i(\mathbf{M}_N)$ ,  $1 \leq i \leq k_1$ , around  $\rho_{\theta_1}$  is not universal and we give it explicitly in terms of these eigenvectors and of the distribution of the entries of the Wigner matrix, see Theorem 3.2.

Secondly, if the eigenvectors are sufficiently delocalized, we establish the universality of the fluctuations of  $\lambda_i(\mathbf{M}_N)$ ,  $1 \leq i \leq k_1$ , see Theorem 3.3.

Actually, in the rank one case, this study allows us to exhibit a necessary and sufficient condition on a normalized eigenvector of  $\mathbf{A}_N$  associated to the largest eigenvalue  $\theta_1$  for the universality of the fluctuations (see Theorem 3.4 below). Moreover if such an eigenvector of  $\mathbf{A}_N$  is not localized but does not satisfy the criteria of universality, the largest eigenvalue of  $\mathbf{M}_N$  may fluctuate according to a mixture of  $\mu$  and normal distributions generalizing (1.5). We will describe some of such intermediate situations.

We detail the definition of localization/delocalization and these results in the Section 3.

Our approach is close to the proof of (1.5), with more involved computations and are in the spirit of the works of [7] and [18]. It is valid in both real and complex settings. Actually, we assume that the eigenvectors associated to the largest eigenvalues of  $\mathbf{A}_N$  belong to a subspace generated by  $k$  ( $= k(N)$ ) canonical vectors of  $\mathbb{C}^N$  and the method requires that  $N - k \rightarrow \infty$  (and even  $\frac{k}{\sqrt{N}} \rightarrow 0$ ). In particular, this approach does not cover the case of Proposition 1.1 studied by [10] where  $k = N$ .

The Deformed Wigner matrix model may be seen as the additive analogue of the spiked population models. These are random sample covariance matrices  $(S_N)_N$  defined by  $S_N = \frac{1}{N} Y_N^* Y_N$  where  $Y_N$  is a  $p \times N$  complex (resp. real) matrix (with  $N$  and  $p = p_N$  of the same order as  $N \rightarrow \infty$ ) whose entries satisfy first four moments conditions; the sample column vectors are assumed to be i.i.d., centered and of covariance matrix a deterministic Hermitian (resp. symmetric) matrix  $\Sigma_p$  having all but finitely many eigenvalues equal to one. In their pioneering article on that topic [6], Baik–Ben Arous–Péché pointed out a phase transition phenomenon for the fluctuations of the largest eigenvalue of  $S_N$  according to the largest eigenvalue of  $\Sigma_p$ , in the complex Gaussian setting; their results were extended in [18] to the real case when the largest eigenvalue of  $\Sigma_p$  is simple and sufficiently larger than 1 and in [17] to singular Wishart matrices. In the non-Gaussian case, the fluctuations of the extreme eigenvalues have been recently studied by Bai–Yao [5] and Féral–Péché [11].

The paper is organized as follows. In Section 2, we present the matricial models under study and the notations that will be used throughout the paper. In Section 3, we present the main results of this paper. We give a summary of our approach in Section 4. Section 5 is devoted to the proofs of Theorems 3.2, 3.3 and 3.4. Finally, we recall some basic facts on matrices, a CLT for random sesquilinear forms and prove some technical results in an Appendix.

## 2. Model and notations

The random matrices under study are complex Hermitian (or real symmetric) matrices  $(\mathbf{M}_N)_N$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that

$$\mathbf{M}_N = \frac{\mathbf{W}_N}{\sqrt{N}} + \mathbf{A}_N, \quad (2.1)$$

where the matrices  $\mathbf{W}_N$  and  $\mathbf{A}_N$  are defined as follows:

(i)  $\mathbf{W}_N$  is a  $N \times N$  Wigner Hermitian (resp. symmetric) matrix such that the  $N^2$  random variables  $(W_N)_{ii}$ ,  $\sqrt{2}\Re((W_N)_{ij})_{i < j}$ ,  $\sqrt{2}\Im((W_N)_{ij})_{i < j}$  (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  $\mu$  of variance  $\sigma^2$  and satisfying a Poincaré inequality; the latter condition means that there exists a positive constant  $C$  such that for any  $C^1$  function  $f : \mathbb{R} \rightarrow \mathbb{C}$  such that  $f$  and  $f'$  are in  $L^2(\mu)$ ,

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

with  $\mathbf{V}(f) = \mathbb{E}(|f - \mathbb{E}(f)|^2)$ .

Note that when  $\mu$  is Gaussian,  $W_N$  is a  $\text{GU}(\text{O})\text{E}(N \times N, \sigma^2)$  matrix.

(ii)  $\mathbf{A}_N$  is a deterministic Hermitian (resp. symmetric) matrix of fixed finite rank  $r$  and built from a family of  $J$  fixed real numbers  $\theta_1 > \dots > \theta_J$  independent of  $N$  with some  $j_0$  such that  $\theta_{j_0} = 0$ . We assume that the non-null eigenvalues  $\theta_j$  of  $\mathbf{A}_N$  are of fixed multiplicity  $k_j$  (with  $\sum_{j \neq j_0} k_j = r$ ). Let  $J_{+\sigma}$  be the number of  $j$ 's such that  $\theta_j > \sigma$ . We denote by  $k_{+\sigma} := k_1 + \dots + k_{J_{+\sigma}}$ . We introduce  $k \geq k_{+\sigma}$  as the minimal number of canonical vectors among the canonical basis  $(e_i; i = 1, \dots, N)$  of  $\mathbb{C}^N$  needed to express all the eigenvectors associated to the largest eigenvalues  $\theta_1, \dots, \theta_{J_{+\sigma}}$  of  $\mathbf{A}_N$ . Without loss of generality (using the invariance of the distribution of the Wigner matrix  $\mathbf{W}_N$  by conjugation by a permutation matrix), we can assume that these  $k_{+\sigma}$  eigenvectors belong to  $\text{Vect}(e_1, \dots, e_k)$ .

All along the paper we assume that  $k \ll \sqrt{N}$ .

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

$$\text{diag}(U_k^*, I_{N-k}) \mathbf{A}_N \text{diag}(U_k, I_{N-k}) = \text{diag}(\theta_j I_{k_j}, (\theta_l I_{k_l})_{l \leq J_{+\sigma}, l \neq j}, Z_{N-k_{+\sigma}}), \quad (2.2)$$

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

Define  $K_j = K_j(N)$  as the minimal number of canonical vectors among  $(e_1, \dots, e_k)$  needed to express all the orthonormal eigenvectors  $v_i^j$ ,  $1 \leq i \leq k_j$ , of  $\mathbf{A}_N$  associated to  $\theta_j$ . Without loss of generality, we can assume that the  $v_i^j$ ,  $1 \leq i \leq k_j$ , belong to  $\text{Vect}(e_1, \dots, e_{K_j})$ . Considering now the vectors  $v_i^j$  as vectors in  $\mathbb{C}^{K_j}$ , we define the  $K_j \times k_j$  matrix

$$U_{K_j \times k_j} := (v_1^j, \dots, v_{k_j}^j) \quad (2.3)$$

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

$$U_{K_j \times k_j}^* U_{K_j \times k_j} = I_{k_j}. \quad (2.4)$$

All along the paper, the parameter  $t$  is such that  $t = 4$  (resp.  $t = 2$ ) in the real (resp. complex) setting and we let  $m_4 := \int x^4 d\mu(x)$ .

Given an arbitrary Hermitian or symmetric matrix  $M$  of size  $N$ , we will denote by  $\lambda_1(M) \geq \dots \geq \lambda_N(M)$  its  $N$  ordered eigenvalues.

### 3. Main results

We first recall the a.s. convergence of the extreme eigenvalues. Define

$$\rho_{\theta_j} = \theta_j + \frac{\sigma^2}{\theta_j}. \quad (3.1)$$

Observe that  $\rho_{\theta_j} > 2\sigma$  (resp.  $< -2\sigma$ ) when  $\theta_j > \sigma$  (resp.  $< -\sigma$ ) (and  $\rho_{\theta_j} = \pm 2\sigma$  if  $\theta_j = \pm\sigma$ ).

For definiteness, we set  $k_1 + \dots + k_{j-1} := 0$  if  $j = 1$ . In [8], we have established the following universal convergence result.

**Theorem 3.1 (a.s. behaviour).** *Let  $J_{+\sigma}$  (resp.  $J_{-\sigma}$ ) be the number of  $j$ 's such that  $\theta_j > \sigma$  (resp.  $\theta_j < -\sigma$ ).*

- (1)  $\forall 1 \leq j \leq J_{+\sigma}, \forall 1 \leq i \leq k_j, \lambda_{k_1+\dots+k_{j-1}+i}(\mathbf{M}_N) \longrightarrow \rho_{\theta_j}$  a.s.
- (2)  $\lambda_{k_1+\dots+k_{J_{+\sigma}+1}}(\mathbf{M}_N) \longrightarrow 2\sigma$  a.s.
- (3)  $\lambda_{k_1+\dots+k_{J_{-\sigma}}}(\mathbf{M}_N) \longrightarrow -2\sigma$  a.s.
- (4)  $\forall j \geq J - J_{-\sigma} + 1, \forall 1 \leq i \leq k_j, \lambda_{k_1+\dots+k_{j-1}+i}(\mathbf{M}_N) \longrightarrow \rho_{\theta_j}$  a.s.

#### 3.1. Fluctuations around $\rho_{\theta_j}$

From Theorem 3.1, for all  $1 \leq i \leq k_j, \lambda_{k_1+\dots+k_{j-1}+i}(\mathbf{M}_N)$  converges to  $\rho_{\theta_j}$  a.s. We shall describe their fluctuations in the extreme two cases:

*Case (a): Localization of the eigenvectors associated to  $\theta_j$ :* The sequence  $K_j(N)$  is bounded,

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

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

*Case (b): Delocalization of the eigenvectors associated to  $\theta_j$ :*  $K_j = K_j(N) \rightarrow \infty$  when  $N \rightarrow \infty$  and  $U_k$  satisfies

$$\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}| \longrightarrow 0 \quad \text{as } N \rightarrow \infty. \quad (3.2)$$

The main results of our paper are the following two theorems. Let  $c_{\theta_j}$  be defined by

$$c_{\theta_j} = \frac{\theta_j^2}{\theta_j^2 - \sigma^2}. \quad (3.3)$$

In Case (a) (which includes the particular setting of Proposition 1.2), the fluctuations of the corresponding rescaled largest eigenvalues of  $\mathbf{M}_N$  are not universal.

**Theorem 3.2.** *In Case (a): the  $k_j$ -dimensional vector*

$$(c_{\theta_j} \sqrt{N} (\lambda_{k_1+\dots+k_{j-1}+i}(\mathbf{M}_N) - \rho_{\theta_j}); i = 1, \dots, k_j)$$

*converges in distribution to  $(\lambda_i(V_{k_j \times k_j}); i = 1, \dots, k_j)$  where  $\lambda_i(V_{k_j \times k_j})$  are the ordered eigenvalues of the matrix  $V_{k_j \times k_j}$  of size  $k_j$  defined in the following way. Let  $W_{\tilde{K}_j}$  be a Wigner matrix of size  $\tilde{K}_j$  with distribution given by  $\mu$  (cf.*

*(i) and  $H_{\tilde{K}_j}$  be a centered Hermitian Gaussian matrix of size  $\tilde{K}_j$  independent of  $W_{\tilde{K}_j}$  with independent entries  $H_{pl}$ ,  $p \leq l$ , with variance*

$$\begin{cases} v_{pp} = \mathbb{E}(H_{pp}^2) = \frac{1}{4} \left( \frac{m_4 - 3\sigma^4}{\theta_j^2} \right) + \frac{1}{2} \frac{\sigma^4}{\theta_j^2 - \sigma^2}, & p = 1, \dots, \tilde{K}_j, \\ v_{pl} = \mathbb{E}(|H_{pl}|^2) = \frac{\sigma^4}{\theta_j^2 - \sigma^2}, & 1 \leq p < l \leq \tilde{K}_j. \end{cases} \quad (3.4)$$

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

$$V_{k_j \times k_j} = \tilde{U}_{\tilde{K}_j \times k_j}^* (W_{\tilde{K}_j} + H_{\tilde{K}_j}) \tilde{U}_{\tilde{K}_j \times k_j}. \quad (3.5)$$

Case (b) exhibits universal fluctuations.

**Theorem 3.3.** *In Case (b): the  $k_j$ -dimensional vector*

$$(c_{\theta_j} \sqrt{N} (\lambda_{k_1 + \dots + k_{j-1} + i}(\mathbf{M}_N) - \rho_{\theta_j}); i = 1, \dots, k_j)$$

converges in distribution to  $(\lambda_i(V_{k_j \times k_j}); i = 1, \dots, k_j)$  where the matrix  $V_{k_j \times k_j}$  is distributed as the  $\text{GU}(\text{O})\text{E}(k_j \times k_j, \frac{\theta_j^2 \sigma^2}{\theta_j^2 - \sigma^2})$ .

**Remark 3.1.** *Note that since  $\mu$  is symmetric, analogue results can be deduced from Theorems 3.2 and 3.3 dealing with the lowest eigenvalues of  $\mathbf{M}_N$  and the  $\theta_j$  such that  $\theta_j < -\sigma$ .*

**Example.**

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

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

### 3.2. Further result for a spike $\theta_j > \sigma$ of multiplicity 1

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

**Theorem 3.4.** *If  $k_j = 1$ ,  $\theta_j > \sigma$ , then the fluctuations of  $\lambda_{k_1 + \dots + k_{j-1} + 1}(\mathbf{M}_N)$  are universal, namely*

$$\sqrt{N} (\lambda_{k_1 + \dots + k_{j-1} + 1}(\mathbf{M}_N) - \rho_{\theta_j}) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{t}{2} \sigma_{\theta_j}^2\right), \quad \text{where } \sigma_{\theta_j} = \sigma \sqrt{1 - \frac{\sigma^2}{\theta_j^2}}$$

if and only if

$$\max_{l \leq K_j} |(U_k)_{l1}| \rightarrow 0 \quad \text{when } N \rightarrow \infty. \quad (3.6)$$

Moreover, our approach allows us to describe the fluctuations of  $\lambda_{k_1 + \dots + k_{j-1} + 1}(\mathbf{M}_N)$  for some particular situations where the corresponding eigenvector of  $\mathbf{A}_N$  is not localized but does not satisfy the criteria of universality  $\max_{l \leq K_j} |(U_k)_{l1}| \rightarrow 0$  (that is somehow for intermediate situations between Cases (a) and (b)). Let  $m$  be a fixed integer number. Assume that for any  $l = 1, \dots, m$ ,  $(U_k)_{l1}$  is independent of  $N$ , whereas  $\max_{m < l \leq K_j} |(U_k)_{l1}| \rightarrow 0$  when  $N$  goes to infinity. We will prove at the end of Section 5 that  $c_{\theta_j} \sqrt{N} (\lambda_{k_1 + \dots + k_{j-1} + 1}(\mathbf{M}_N) - \rho_{\theta_j})$  converges in distribution towards the mixture of  $\mu$ -distributed or Gaussian random variables  $\sum_{i,l=1}^m a_{il} \xi_{il} + \mathcal{N}$  in the complex case,  $\sum_{1 \leq l \leq i \leq m} a_{il} \xi_{il} + \mathcal{N}$  in the real case, where  $\xi_{il}$ ,  $(i, l) \in \{1, \dots, m\}^2$ ,  $\mathcal{N}$  are independent random variables such that

- for any  $(i, l) \in \{1, \dots, m\}^2$ , the distribution of  $\xi_{il}$  is  $\mu$ ;

$$a_{il} = \begin{cases} \sqrt{2} \Im((U_k)_{l1} \overline{(U_k)_{i1}}) & \text{if } i < l, \\ \sqrt{t} \Re((U_k)_{l1} \overline{(U_k)_{i1}}) & \text{if } i > l, \\ \sqrt{\frac{t}{2}} |(U_k)_{l1}|^2 & \text{if } i = l; \end{cases}$$

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

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

#### 4. Sketch of the approach

Before we proceed to the proof of Theorems 3.2 and 3.3, let us give the sketch of our approach which are similar in both cases. To this aim, we define for any random variable  $\lambda$ ,

$$\xi_N(\lambda) = c_{\theta_j} \sqrt{N}(\lambda - \rho_{\theta_j}) \quad (4.1)$$

with  $c_{\theta_j}$  given by (3.3). We also set  $\hat{k}_{j-1} := k_1 + \dots + k_{j-1}$  with the convention that  $\hat{k}_0 = 0$ .

The reasoning made in the setting of Proposition 1.2 (for which  $k = k_{+\sigma} = 1$ ) relies (following ideas previously developed in [7] and [18]) on the writing of the rescaled eigenvalue  $\xi_N(\lambda_1(\mathbf{M}_N))$  in terms of the resolvent of an underlying non-Deformed Wigner matrix. The conclusion then essentially follows from a CLT on random sesquilinear forms established by J. Baik and J. Silverstein in the Appendix of [8] (which corresponds to the following Theorem A.2 in the scalar case). In the general case, to prove the convergence in distribution of the vector  $(\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N))); i = 1, \dots, k_j)$ , we will extend, as [5], the previous approach in the following sense. We will show that each of these rescaled eigenvalues is an eigenvalue of a  $k_j \times k_j$  random matrix which may be expressed in terms of the resolvent of a  $N - k \times N - k$  Deformed Wigner matrix whose eigenvalues do not jump asymptotically outside  $[-2\sigma; 2\sigma]$ ; then, the matrix  $V_{k_j \times k_j}$  will arise from a multidimensional CLT on random sesquilinear forms. Nevertheless, due to the multidimensional situation to be considered now, additional considerations are required. Let us give more details.

Consider an arbitrary random variable  $\lambda$  which converges in probability towards  $\rho_{\theta_j}$ . Then, applying factorizations of type (A.1), we prove that  $\lambda$  is an eigenvalue of  $\mathbf{M}_N$  iff  $\xi_N(\lambda)$  is (on some event having probability going to 1 as  $N \rightarrow \infty$ ) an eigenvalue of a  $k_j \times k_j$  matrix  $\check{X}_{k_j, N}(\lambda)$  of the form

$$\check{X}_{k_j, N}(\lambda) = V_{k_j, N} + R_{k_j, N}(\lambda), \quad (4.2)$$

where  $V_{k_j, N}$  converges in distribution towards  $V_{k_j \times k_j}$  and the remaining term  $R_{k_j, N}(\lambda)$  turns out to be negligible. Now, when  $k_j > 1$ , since the matrix  $\check{X}_{k_j, N}(\lambda)$  (in (4.2)) depends on  $\lambda$ , the previous reasoning with  $\lambda = \lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N)$  for any  $1 \leq i \leq k_j$  does not allow us to readily deduce that the  $k_j$  normalized eigenvalues  $\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N))$ ,  $1 \leq i \leq k_j$ , are eigenvalues of a same matrix of the form  $V_{k_j, N} + o_{\mathbb{P}}(1)$  and then that

$$(\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N))); 1 \leq i \leq k_j) = (\lambda_i(V_{k_j, N}); 1 \leq i \leq k_j) + o_{\mathbb{P}}(1). \quad (4.3)$$

Note that the authors do not develop this difficulty in [5], pp. 464–465. Hence, in the last step of the proof (Step 4 in Section 5), we detail the additional arguments which are needed to get (4.3) when  $k_j > 1$ .

Our approach will cover Cases (a) and (b) and we will handle both cases once this will be possible. In fact, the main difference appears in the proof of the convergence in distribution of the matrix  $V_{k_j, N}$  which gives rise to the “occurrence or non-occurrence” of the distribution  $\mu$  in the limiting fluctuations and then justifies the non-universality (resp. universality) in Case (a) (resp. Case (b)).

The proof is organized in four steps as follows. In Steps 1 and 2, we explain how to obtain (4.2): we exhibit the matrix  $\check{X}_{k_j, N}$  and bring its leading term  $V_{k_j, N}$  to light in Step 2. We establish the convergence in distribution of the matrix  $V_{k_j, N}$  in Step 3. Step 4 is devoted to the concluding arguments of the proof.

#### 5. Proofs of Theorems 3.2, 3.3 and 3.4

As far as possible, we handle both the proofs of Theorem 3.2 and Theorem 3.3. We will proceed in four steps. First, let us introduce a few notations.

For any matrix  $M \in \mathcal{M}_N(\mathbb{C})$ , we denote by  $\text{Tr}$  (resp.  $\text{tr}_N$ ) the classical (resp. normalized) trace. For a rectangular matrix,  $\|M\|$  is the operator norm of  $M$  and  $\|M\|_{\text{HS}} := (\text{Tr}(MM^*))^{1/2}$  the Hilbert–Schmidt norm.

For an Hermitian matrix, we denote by  $\text{Spect}(M)$  the spectrum of  $M$ . For  $z \in \mathbb{C} \setminus \text{Spect}(M)$ ,  $G_M(z) = (zI_N - M)^{-1}$  denotes the resolvent of  $M$  (we suppress the index  $M$  when there is no confusion). We have the following:

$$\text{For } x > \lambda_1(M); \quad \|G(x)\| \leq \frac{1}{x - \lambda_1(M)}. \quad (5.1)$$

For a  $m \times q$  matrix  $B$  (or  $\mathbf{B}$ ) and some integers  $1 \leq p \leq m$  and  $1 \leq l \leq q$ , we denote respectively by  $[B]_{p \times l}^{\nearrow}$ ,  $[B]_{p \times l}^{\searrow}$ ,  $[B]_{p \times l}^{\swarrow}$  and  $[B]_{p \times l}^{\nwarrow}$  the upper left, upper right, lower left and lower right corner of size  $p \times l$  of the matrix  $B$ . If  $p = l$ , we will often replace the indices  $p \times l$  by  $p$  for convenience. Moreover if  $p = m$ , we may replace  $\nearrow$  or  $\searrow$  by  $\rightarrow$  and  $\swarrow$  or  $\nwarrow$  by  $\leftarrow$ . Similarly if  $l = q$ , we may replace  $\nearrow$  or  $\nwarrow$  by  $\uparrow$  and  $\swarrow$  or  $\searrow$  by  $\downarrow$ .

For simplicity in the writing we will define the  $k \times k$ , resp.  $N - k \times N - k$ , resp.  $k \times N - k$  matrix  $W_k$ , resp.  $W_{N-k}$ , resp.  $Y$ , by setting

$$\mathbf{W}_N = \begin{pmatrix} W_k & Y \\ Y^* & W_{N-k} \end{pmatrix}. \quad (5.2)$$

Given  $\mathbf{B} \in \mathcal{M}_N(\mathbb{C})$ , we will denote by  $\tilde{\mathbf{B}}$  the  $N \times N$  matrix given by

$$\tilde{\mathbf{B}} := \text{diag}(U_k^*, I_{N-k}) \mathbf{B} \text{diag}(U_k, I_{N-k}) = \begin{pmatrix} \tilde{B}_k & \tilde{B}_{k \times N-k} \\ \tilde{B}_{N-k \times k} & \tilde{B}_{N-k} \end{pmatrix}.$$

One obviously has that  $\tilde{B}_{N-k} = B_{N-k}$ .

In this way, we define the matrices  $\tilde{\mathbf{M}}_N$ ,  $\tilde{\mathbf{W}}_N$  and  $\tilde{\mathbf{A}}_N$ . In particular, we notice from (2.2) that

$$\tilde{\mathbf{A}}_N = \text{diag}(\theta_j I_{k_j}, (\theta_l I_{k_l})_{l \leq J+\sigma, l \neq j}, Z_{N-k+\sigma}) = \begin{pmatrix} \tilde{A}_k & \tilde{A}_{k \times N-k} \\ \tilde{A}_{N-k \times k} & A_{N-k} \end{pmatrix}. \quad (5.3)$$

Note also that since  $A_{N-k}$  is a submatrix of  $Z_{N-k+\sigma}$ , all its eigenvalues are strictly smaller than  $\sigma$ .

Let  $0 < \delta < (\rho_{\theta_j} - 2\sigma)/2$ . For any random variable  $\lambda$ , define the events

$$\begin{aligned} \Omega_N^{(1)}(\lambda) &= \left\{ \lambda_1 \left( \frac{\mathbf{W}_N}{\sqrt{N}} + \text{diag}(U_k, I_{N-k}) \text{diag}(0_{k+\sigma}, Z_{N-k+\sigma}) \text{diag}(U_k^*, I_{N-k}) \right) < 2\sigma + \delta; \lambda > \rho_{\theta_j} - \delta \right\}, \\ \Omega_N^{(2)} &= \left\{ \lambda_1 \left( \frac{W_{N-k}}{\sqrt{N}} + A_{N-k} \right) \leq 2\sigma + \delta \right\} \end{aligned}$$

and

$$\Omega_N(\lambda) = \Omega_N^{(1)}(\lambda) \cap \Omega_N^{(2)}. \quad (5.4)$$

On  $\Omega_N(\lambda)$ , neither  $\lambda$  nor  $\rho_{\theta_j}$  are eigenvalues of  $M_{N-k} := \frac{W_{N-k}}{\sqrt{N}} + A_{N-k}$ , thus the resolvent  $\hat{G}(x)$  of  $M_{N-k}$  is well defined at  $x = \lambda$  and  $x = \rho_{\theta_j}$ .

Note that from Theorem 3.1, for any random sequence  $\Lambda_N$  converging towards  $\rho_{\theta_j}$  in probability,

$$\lim_{N \rightarrow \infty} \mathbb{P}(\Omega_N(\Lambda_N)) = 1.$$

Let us now introduce on  $\Omega_N(\lambda)$  some auxiliary matrices that will be of basic use to the proofs.

$$B_{k,N} = W_k + \frac{1}{\sqrt{N}} \left( Y \hat{G}(\rho_{\theta_j}) Y^* - (N-k) \frac{\sigma^2}{\theta_j} I_k \right), \quad (5.5)$$

$$V_{k+\sigma,N} := [U_k^* B_{k,N} U_k]_{k+\sigma}^{\nwarrow}, \quad (5.6)$$



$$\tau_N(\lambda) = \frac{1}{N}(\lambda - \rho_{\theta_j})Y\hat{G}(\lambda)\hat{G}(\rho_{\theta_j})^2Y^*, \quad (5.7)$$

$$\phi_N = -\frac{1}{N}(Y\hat{G}(\rho_{\theta_j})^2Y^* - \sigma^2 \text{Tr} \hat{G}(\rho_{\theta_j})^2 I_k), \quad (5.8)$$

$$\psi_N = -\sigma^2 \frac{N-k}{N} \left( \text{tr}_{N-k} \hat{G}(\rho_{\theta_j})^2 - \frac{1}{\theta_j^2 - \sigma^2} \right) I_k, \quad (5.9)$$

$$c_{\theta_j} D_{k,N}(\lambda) = \tau_N(\lambda) + \phi_N + \psi_N, \quad (5.10)$$

$$T_N(\lambda) = \left[ U_k^* \left( W_k + \frac{1}{\sqrt{N}} Y \hat{G}(\lambda) Y^* \right) U_k \right]_{k+\sigma \times (k-k+\sigma)}^{\nearrow}, \quad (5.11)$$

$$\Delta_N(\lambda) = \left[ U_k^* Y \hat{G}(\lambda) \tilde{A}_{N-k \times k} \right]_{k+\sigma \times (k-k+\sigma)}^{\nearrow}, \quad (5.12)$$

$$\Gamma_{k+\sigma \times k-k+\sigma}(\lambda) = T_N(\lambda) + \Delta_N(\lambda), \quad (5.13)$$

$$\mathbf{Q}_{k,N}(\lambda) := \tilde{M}_k + \tilde{M}_{k \times N-k} \hat{G}(\lambda) \tilde{M}_{N-k \times k}, \quad (5.14)$$

$$\Sigma_{k-k+\sigma}(\lambda) = \left( \left[ \mathbf{Q}_{k,N}(\lambda) \right]_{k-k+\sigma}^{\searrow} - \lambda I_{k-k+\sigma} \right)^{-1}. \quad (5.15)$$

Note that we will justify that  $\Sigma_{k-k+\sigma}(\lambda)$  is well defined in the course of the proof of Proposition 5.1 below. Finally, set

$$\begin{aligned} \mathbf{X}_{k+\sigma,N}(\lambda) &= \left[ U_k^* B_{k,N} U_k \right]_{k+\sigma}^{\searrow} + \sqrt{N} \text{diag}(0_{k_j}, (\theta_l - \theta_j) I_{k_l}, l = 1, \dots, J+\sigma, l \neq j) \\ &\quad + \xi_N(\lambda) \left[ U_k^* D_{k,N}(\lambda) U_k \right]_{k+\sigma}^{\searrow} \\ &\quad + \left( \frac{\sigma^2}{\theta_j^2 - \sigma^2} \frac{\xi_N(\lambda)}{c_{\theta_j}} \frac{k}{N} - \frac{k}{\sqrt{N}} \frac{\sigma^2}{\theta_j} \right) I_{k+\sigma} \\ &\quad - \frac{1}{\sqrt{N}} \Gamma_{k+\sigma \times k-k+\sigma}(\lambda) \Sigma_{k-k+\sigma}(\lambda) \Gamma_{k+\sigma \times k-k+\sigma}(\lambda)^*. \end{aligned} \quad (5.16)$$

*Step 1:* We show that an eigenvalue of  $\mathbf{M}_N$  is an eigenvalue of a matrix of size  $k+\sigma$ . More, precisely, we have:

**Proposition 5.1.** *For any random variable  $\lambda$  and any  $k+\sigma \times k+\sigma$  random matrix  $\Delta_{k+\sigma}$ , on  $\Omega_N(\lambda)$ ,  $\lambda$  is an eigenvalue of  $\tilde{\mathbf{M}}_N + \text{diag}(\Delta_{k+\sigma}, 0)$  iff  $\xi_N(\lambda)$  is an eigenvalue of  $\mathbf{X}_{k+\sigma,N}(\lambda) + \sqrt{N} \Delta_{k+\sigma}$  where  $\mathbf{X}_{k+\sigma,N}(\lambda)$  is defined by (5.16). Moreover, the  $k-k+\sigma \times k-k+\sigma$  matrix  $\Sigma_{k-k+\sigma}(\lambda)$  defined by (5.15) is such that*

$$\| \Sigma_{k-k+\sigma}(\lambda) \| \leq 1/(\rho_{\theta_j} - 2\sigma - 2\delta).$$

**Proof.** Let  $\lambda$  be a random variable. On  $\Omega_N(\lambda)$ ,

$$\begin{aligned} \det(\mathbf{M}_N - \lambda I_N) &= \det(\tilde{\mathbf{M}}_N - \lambda I_N) \\ &= \det \begin{pmatrix} \tilde{M}_k - \lambda I_k & \tilde{M}_{k \times N-k} \\ M_{N-k \times k} & M_{N-k} - \lambda I_{N-k} \end{pmatrix} \\ &= \det(M_{N-k} - \lambda I_{N-k}) \det(\tilde{M}_k - \lambda I_k + \tilde{M}_{k \times N-k} \hat{G}(\lambda) \tilde{M}_{N-k \times k}). \end{aligned}$$

The last equality in the above equation follows from (A.1). Since on  $\Omega_N(\lambda)$ ,  $\lambda$  is not an eigenvalue of  $M_{N-k}$ , we can deduce that  $\lambda$  is an eigenvalue of  $\tilde{\mathbf{M}}_N$  if and only if it is an eigenvalue of

$$\mathbf{Q}_{k,N}(\lambda) = \tilde{M}_k + \tilde{M}_{k \times N-k} \hat{G}(\lambda) \tilde{M}_{N-k \times k}.$$

Now, note that we have also from (A.1) that

$$\begin{aligned} & \det\left(\left[\frac{\tilde{\mathbf{W}}_N}{\sqrt{N}}\right]_{N-k+\sigma}^{\searrow} + Z_{N-k+\sigma} - \lambda I_{N-k+\sigma}\right) \\ &= \det\left(\frac{W_{N-k}}{\sqrt{N}} + [Z_{N-k+\sigma}]_{N-k}^{\searrow} - \lambda I_{N-k}\right) \times \det\left([\mathbf{Q}_{k,N}(\lambda)]_{k-k+\sigma}^{\searrow} - \lambda I_{k-k+\sigma}\right). \end{aligned}$$

The matrix  $[\frac{\tilde{\mathbf{W}}_N}{\sqrt{N}}]_{N-k+\sigma}^{\searrow} + Z_{N-k+\sigma}$  is a submatrix of  $\frac{\tilde{\mathbf{W}}_N}{\sqrt{N}} + \text{diag}(0_{k+\sigma}, Z_{N-k+\sigma})$  whose eigenvalues are (on  $\Omega_N(\lambda)$ ) smaller than  $2\sigma + \delta$ . So, since on  $\Omega_N(\lambda)$ ,  $\lambda$  is greater than  $\rho_{\theta_j} - \delta > 2\sigma + \delta$ , we can conclude that  $\lambda$  cannot be an eigenvalue of  $[\frac{\tilde{\mathbf{W}}_N}{\sqrt{N}}]_{N-k+\sigma}^{\searrow} + Z_{N-k+\sigma}$ , and then neither of  $[\mathbf{Q}_{k,N}(\lambda)]_{k-k+\sigma}^{\searrow}$ . Thus, we can define

$$\Sigma_{k-k+\sigma}(\lambda) = \left([\mathbf{Q}_{k,N}(\lambda)]_{k-k+\sigma}^{\searrow} - \lambda I_{k-k+\sigma}\right)^{-1}.$$

Moreover on  $\Omega_N(\lambda)$ , one can see using (A.1) that if  $\lambda_0$  is an eigenvalue of  $[\mathbf{Q}_{k,N}(\lambda)]_{k-k+\sigma}^{\searrow} - \lambda I_{k-k+\sigma}$  then  $\lambda$  is an eigenvalue of

$$\left[\frac{\tilde{\mathbf{W}}_N}{\sqrt{N}}\right]_{N-k+\sigma}^{\searrow} + Z_{N-k+\sigma} - \text{diag}(\lambda_0 I_{k-k+\sigma}, 0_{N-k}).$$

Hence,

$$\lambda \leq \lambda_1 \left( \left[\frac{\tilde{\mathbf{W}}_N}{\sqrt{N}}\right]_{N-k+\sigma}^{\searrow} + Z_{N-k+\sigma} \right) + |\lambda_0|$$

and then

$$|\lambda_0| \geq \rho_{\theta_j} - \delta - 2\sigma - \delta,$$

so that finally

$$\|\Sigma_{k-k+\sigma}(\lambda)\| \leq \frac{1}{\rho_{\theta_j} - 2\sigma - 2\delta}. \quad (5.17)$$

Using once more (A.1), we get that on  $\Omega_N(\lambda)$ ,  $\lambda$  is an eigenvalue of  $\mathbf{Q}_{k,N}(\lambda)$  if and only if it is an eigenvalue of  $[\mathbf{Q}_{k,N}(\lambda)]_{k+\sigma}^{\nearrow} - [\mathbf{Q}_{k,N}(\lambda)]_{k+\sigma \times k-k+\sigma}^{\nearrow} \Sigma_{k-k+\sigma}(\lambda) [\mathbf{Q}_{k,N}(\lambda)]_{k-k+\sigma \times k+\sigma}^{\swarrow}$  or equivalently if and only if  $\xi_N(\lambda)$  is an eigenvalue of

$$c_{\theta_j} \sqrt{N} \left( [\mathbf{Q}_{k,N}(\lambda)]_{k+\sigma}^{\nearrow} - \rho_{\theta_j} I_{k+\sigma} - [\mathbf{Q}_{k,N}(\lambda)]_{k+\sigma \times k-k+\sigma}^{\nearrow} \Sigma_{k-k+\sigma}(\lambda) [\mathbf{Q}_{k,N}(\lambda)]_{k-k+\sigma \times k+\sigma}^{\swarrow} \right).$$

Now using

$$\hat{G}(\lambda) - \hat{G}(\rho_{\theta_j}) = -(\lambda - \rho_{\theta_j}) \hat{G}(\rho_{\theta_j}) \hat{G}(\lambda),$$

one can replace  $\hat{G}(\lambda)$  by  $\hat{G}(\rho_{\theta_j}) + [-(\lambda - \rho_{\theta_j}) \hat{G}(\rho_{\theta_j}) (\hat{G}(\rho_{\theta_j}) - (\lambda - \rho_{\theta_j}) \hat{G}(\rho_{\theta_j}) \hat{G}(\lambda))]$  and get the following writing

$$\frac{1}{\sqrt{N}} Y \hat{G}(\lambda) Y^* = \frac{1}{\sqrt{N}} Y \hat{G}(\rho_{\theta_j}) Y^* + \xi_N(\lambda) D_{k,N}(\lambda) - \xi_N(\lambda) \frac{N-k}{N} \frac{\sigma^2}{c_{\theta_j}(\theta_j^2 - \sigma^2)} I_k, \quad (5.18)$$

where

$$\begin{aligned} c_{\theta_j} D_{k,N}(\lambda) &= \frac{1}{N} (\lambda - \rho_{\theta_j}) Y \hat{G}(\lambda) \hat{G}(\rho_{\theta_j})^2 Y^* - \frac{1}{N} (Y \hat{G}(\rho_{\theta_j})^2 Y^* - \sigma^2 \text{Tr} \hat{G}(\rho_{\theta_j})^2 I_k) \\ &\quad - \sigma^2 \frac{N-k}{N} \left( \text{tr}_{N-k} \hat{G}(\rho_{\theta_j})^2 - \frac{1}{\theta_j^2 - \sigma^2} \right) I_k. \end{aligned}$$

Then

$$\begin{aligned}
c_{\theta_j} \sqrt{N} ([\mathbf{Q}_{k,N}(\lambda)]_{k+\sigma}^{\nearrow} - \rho_{\theta_j} I_{k+\sigma}) &= c_{\theta_j} \left\{ \left[ U_k^* \left( W_k + \frac{1}{\sqrt{N}} \left( Y \hat{G}(\rho_{\theta_j}) Y^* - (N-k) \frac{\sigma^2}{\theta_j} I_k \right) \right) U_k \right]_{k+\sigma}^{\nearrow} \right. \\
&\quad + \sqrt{N} \operatorname{diag}(0_{k_j}, (\theta_l - \theta_j) I_{k_l}, l = 1, \dots, J_{+\sigma}, l \neq j) \\
&\quad + \xi_N(\lambda) [U_k^* D_{k,N}(\lambda) U_k]_{k+\sigma}^{\nearrow} - \frac{k}{\sqrt{N}} \frac{\sigma^2}{\theta_j} I_{k+\sigma} + \frac{\sigma^2}{\theta_j^2 - \sigma^2} \frac{\xi_N(\lambda)}{c_{\theta_j}} \frac{k}{N} I_{k+\sigma} \left. \right\} \\
&\quad - \frac{\sigma^2}{\theta_j^2 - \sigma^2} \xi_N(\lambda) I_{k+\sigma}.
\end{aligned}$$

The proposition (adding an extra matrix  $\Delta_{k+\sigma}$  for future computations) readily follows.  $\square$

Throughout Steps 2 and 3,  $\Lambda_N$  denotes any random sequence converging in probability towards  $\rho_{\theta_j}$ . The aim of these two steps is to study the limiting behavior of the matrix  $\mathbf{X}_{k+\sigma,N}(\Lambda_N)$  (defined by (5.16)) as  $N$  goes to infinity.

*Step 2:* We first focus on the negligible terms in  $\mathbf{X}_{k+\sigma,N}(\Lambda_N)$  and establish the following.

**Proposition 5.2.** *Assume that  $k \ll \sqrt{N}$ . For any random sequence  $\Lambda_N$  converging in probability towards  $\rho_{\theta_j}$ , on  $\Omega_N(\Lambda_N)$ ,*

$$\mathbf{X}_{k+\sigma,N}(\Lambda_N) = V_{k+\sigma,N} + \sqrt{N} \operatorname{diag}(0_{k_j}, (\theta_l - \theta_j) I_{k_l}, l = 1, \dots, J_{+\sigma}, l \neq j) + (1 + |\xi_N(\Lambda_N)|)^2 \mathfrak{o}_{\mathbb{P}}(1) \quad (5.19)$$

with  $V_{k+\sigma,N}$  defined by (5.6).

The proof of this proposition is quite long and is divided in several lemmas. Although our final result in the case  $k$  infinite holds only for  $k \ll \sqrt{N}$ , we will give some estimates for  $k \ll N$  once this is possible.

**Lemma 5.1.** *Let  $k \ll N$ . Then, on  $\Omega_N(\Lambda_N)$ ,*

$$[U_k^* D_{k,N}(\Lambda_N) U_k]_{k+\sigma}^{\nearrow} = \mathfrak{o}_{\mathbb{P}}(1). \quad (5.20)$$

**Proof.**  $D_{k,N}(\Lambda_N)$ ,  $\tau_N$ ,  $\phi_N$  and  $\psi_N$  are respectively defined by (5.10), (5.7), (5.8) and (5.9).

Since  $Y$  is a submatrix of  $\mathbf{W}_N$ ,  $\limsup_N \frac{\|Y\|}{\sqrt{N}} \leq 2$  and  $\frac{\|Y\|}{\sqrt{N}} = \mathfrak{O}_{\mathbb{P}}(1)$ . Thus,

$$\|\tau_N\| = \mathfrak{O}_{\mathbb{P}}(\Lambda_N - \rho_{\theta_j}) = \mathfrak{o}_{\mathbb{P}}(1), \quad (5.21)$$

where we used that  $\|\hat{G}(\lambda)\| \leq \frac{1}{(\rho_{\theta_j} - 2\sigma - 2\delta)}$  for  $\lambda = \rho_{\theta_j}$  or  $\lambda = \Lambda_N$ . Therefore,

$$\|[U_k^* \tau_N U_k]_{k+\sigma}^{\nearrow}\| = \mathfrak{o}_{\mathbb{P}}(1).$$

It follows from Lemma A.3 in the Appendix that

$$[U_k^* \psi_N U_k]_{k+\sigma}^{\nearrow} := -\sigma^2 \frac{N-k}{N} \left[ \operatorname{tr}_{N-k} \hat{G}(\rho_{\theta_j})^2 - \frac{1}{\theta_j^2 - \sigma^2} \right] I_{k+\sigma} = \mathfrak{O}_{\mathbb{P}}(1).$$

Now, we have

$$\begin{aligned}
\mathbb{E}(\|[U_k^* \phi_N U_k]_{k+\sigma}^{\nearrow} \mathbf{1}_{\Omega_N(\Lambda_N)}\|_{\text{HS}}^2) &\leq \mathbb{E}(\|[U_k^* \phi_N U_k]_{k+\sigma}^{\nearrow} \mathbf{1}_{\Omega_N^{(2)}}\|_{\text{HS}}^2) \\
&\leq \sum_{p,q=1}^{k+\sigma} \frac{1}{N^2} \mathbb{E}(|\mathcal{U}(p)^* \hat{G}(\rho_{\theta_j})^2 \mathcal{U}(q) - \sigma^2 \operatorname{Tr} \hat{G}(\rho_{\theta_j})^2 \delta_{p,q}|^2 \mathbf{1}_{\Omega_N^{(2)}}),
\end{aligned}$$

where for any  $p = 1, \dots, k_{+\sigma}$ , we let  $\mathcal{U}(p) = {}^t[(Y^*U_k)_{1,p}, \dots, (Y^*U_k)_{N-k,p}]$ . We first state some properties of the vectors  $\mathcal{U}(p)$ .

**Lemma 5.2.** *Let  $\mathcal{U}$  denote the  $N - k \times k_{+\sigma}$  matrix  $[Y^*U_k]_{k_{+\sigma}}^{\leftarrow}$ . Then, the rows  $(\mathcal{U}_i; i \leq N - k)$  are centered i.i.d. vectors in  $\mathbb{C}^{k_{+\sigma}}$ , with a distribution depending on  $N$ . Moreover, we have for all  $1 \leq p, q \leq k_{+\sigma}$ :*

$$\begin{aligned} \mathbb{E}(\mathcal{U}_{1p}\bar{\mathcal{U}}_{1q}) &= \delta_{p,q}\sigma^2 \quad \text{with } \mathbb{E}(\mathcal{U}_{1p}\mathcal{U}_{1q}) = 0 \text{ in the complex case,} \\ \mathbb{E}[|\mathcal{U}_{ip}|^2|\mathcal{U}_{iq}|^2] &= \left(1 + \frac{t}{2}\delta_{p,q}\right)\sigma^4 + \left[\mathbb{E}(|W_{12}|^4) - \left(1 + \frac{t}{2}\right)\sigma^4\right] \sum_{l=1}^k |(U_k)_{l,p}|^2 |(U_k)_{l,q}|^2. \end{aligned} \tag{5.22}$$

Since  $\sum_{l=1}^k |(U_k)_{l,p}|^4 \leq 1$ , the fourth moment of  $\mathcal{U}_{1p}$  is uniformly bounded.

We skip the proof of this lemma which follows from straightforward computations using the independence of the entries of  $Y$  and the fact that  $U_k$  is unitary.

Then, according to Theorem A.1 and using (5.1),

$$\begin{aligned} \frac{1}{N^2} \mathbb{E}(|\mathcal{U}(p)^* \hat{G}(\rho_{\theta_j})^2 \mathcal{U}(p) - \sigma^2 \text{Tr} \hat{G}(\rho_{\theta_j})^2|^2 \mathbf{1}_{\Omega_N(\Lambda_N)}) &\leq \frac{K}{N} \mathbb{E}(\text{tr}_N \hat{G}(\rho_{\theta_j})^4 \mathbf{1}_{\Omega_N^{(2)}}) \\ &\leq \frac{K}{N} \mathbb{E}(\|\hat{G}(\rho_{\theta_j})\|^4 \mathbf{1}_{\Omega_N^{(2)}}) \\ &\leq \frac{K}{N} \frac{1}{(\rho_{\theta_j} - 2\sigma - \delta)^4}. \end{aligned}$$

Besides for  $p \neq q$ , using the independence between  $(\mathcal{U}(p), \mathcal{U}(q))$  and  $\hat{G}(\rho_{\theta_j})$ , we have:

$$\begin{aligned} \mathbb{E}(|\mathcal{U}(p)^* \hat{G}(\rho_{\theta_j})^2 \mathcal{U}(q)|^2 \mathbf{1}_{\Omega_N^{(2)}}) &= \sum_{i,j,l,m}^{N-k} \mathbb{E}[\bar{\mathcal{U}}_{ip}(G^2)_{ij} \mathcal{U}_{jq} \mathcal{U}_{lp} \overline{(G^2)_{lm}} \bar{\mathcal{U}}_{mq} \mathbf{1}_{\Omega_N^{(2)}}] \\ &= \sum_{i,j,l,m}^{N-k} \mathbb{E}[\bar{\mathcal{U}}_{ip} \mathcal{U}_{jq} \mathcal{U}_{lp} \bar{\mathcal{U}}_{mq}] \mathbb{E}[(G^2)_{ij} \overline{(G^2)_{lm}} \mathbf{1}_{\Omega_N^{(2)}}], \end{aligned}$$

where we denote by  $G$  the matrix  $\hat{G}(\rho_{\theta_j})$  for simplicity. From Lemma 5.2, for  $p \neq q$ , the only terms giving a non-null expectation in the above equation are those for which:

(1)  $i = l, j = m$  and  $i \neq j$ . In this case,

$$\mathbb{E}[\bar{\mathcal{U}}_{ip} \mathcal{U}_{jq} \mathcal{U}_{lp} \bar{\mathcal{U}}_{jq}] = \mathbb{E}[\bar{\mathcal{U}}_{ip} \mathcal{U}_{ip}] \mathbb{E}[\mathcal{U}_{jq} \bar{\mathcal{U}}_{jq}] = \sigma^4$$

and

$$\sum_{i,j,i \neq j}^{N-k} \mathbb{E}[(G^2)_{ij} \overline{(G^2)_{ij}} \mathbf{1}_{\Omega_N^{(2)}}] \leq \mathbb{E} \text{Tr}(G^4 \mathbf{1}_{\Omega_N^{(2)}}).$$

(2)  $i = j = k = l$ . In this case, using (5.22), there is a constant  $C > 0$  such that

$$\mathbb{E}[\bar{\mathcal{U}}_{ip} \mathcal{U}_{iq} \mathcal{U}_{ip} \bar{\mathcal{U}}_{iq}] = \mathbb{E}[|\mathcal{U}_{ip}|^2 |\mathcal{U}_{iq}|^2] \leq C.$$

Moreover

$$\sum_{i=1}^{N-k} \mathbb{E}[G_{ii}^2 \bar{G}_{ii}^2 \mathbf{1}_{\Omega_N^{(2)}}] \leq \mathbb{E} \text{Tr}(G^4 \mathbf{1}_{\Omega_N^{(2)}}).$$

Therefore,

$$\mathbb{E}(|\mathcal{U}(p)^* \hat{G}(\rho_{\theta_j})^2 \mathcal{U}(q)|^2 \mathbf{1}_{\Omega_N^{(2)}}) \leq (C + \sigma^4) \mathbb{E} \operatorname{Tr}(\hat{G}(\rho_{\theta_j})^4 \mathbf{1}_{\Omega_N^{(2)}}). \quad (5.23)$$

Hence,

$$\begin{aligned} \frac{1}{N^2} \mathbb{E}(|\mathcal{U}(p)^* \hat{G}(\rho_{\theta_j})^2 \mathcal{U}(q)|^2 \mathbf{1}_{\Omega_N^{(2)}}) &\leq \frac{C + \sigma^4}{N} \mathbb{E}(\|\hat{G}(\rho_{\theta_j})\|^4 \mathbf{1}_{\Omega_N^{(2)}}) \\ &\leq \frac{C + \sigma^4}{N} \frac{1}{(\rho_{\theta_j} - 2\sigma - \delta)^4}. \end{aligned}$$

Thus

$$\mathbb{E}(\| [U_k^* \phi_N U_k]_{k+\sigma}^\frown \|^2 \mathbf{1}_{\Omega_N(\Lambda_N)}) \leq (C + \sigma^4) \frac{k_{+\sigma}^2}{N} \frac{1}{(\rho_{\theta_j} - 2\sigma - \delta)^4}.$$

The convergence in probability of  $[U_k^* \phi_N U_k]_{k+\sigma}^\frown$  towards zero readily follows by Tchebychev inequality. Lemma 5.1 is established.  $\square$

For simplicity, we now write

$$\Sigma(\Lambda_N) = \Sigma_{k-k+\sigma}(\Lambda_N).$$

Let us define

$$\begin{aligned} R_{k,N}(\Lambda_N) &:= -\frac{k}{\sqrt{N}} \frac{\sigma^2}{\theta_j} I_{k+\sigma} + \frac{\sigma^2}{\theta_j^2 - \sigma^2} \frac{\xi_N(\Lambda_N)}{c_{\theta_j}} \frac{k}{N} I_{k+\sigma} \\ &\quad - \frac{1}{\sqrt{N}} \Gamma_{k+\sigma \times k-k+\sigma}(\Lambda_N) \Sigma(\Lambda_N) \Gamma_{k+\sigma \times k-k+\sigma}(\Lambda_N)^*. \end{aligned} \quad (5.24)$$

To get Proposition 5.2, it remains to prove that if  $k \ll \sqrt{N}$ ,

$$R_{k,N}(\Lambda_N) = (1 + |\xi_N(\Lambda_N)|)^2 \mathbf{o}_{\mathbb{P}}(1). \quad (5.25)$$

Once  $k \ll \sqrt{N}$ , we readily have that

$$-\frac{k}{\sqrt{N}} \frac{\sigma^2}{\theta_j} I_{k+\sigma} + \frac{\sigma^2}{\theta_j^2 - \sigma^2} \frac{\xi_N(\Lambda_N)}{c_{\theta_j}} \frac{k}{N} I_{k+\sigma} = (1 + |\xi_N(\Lambda_N)|)^2 \mathbf{o}_{\mathbb{P}}(1).$$

Hence, (5.25) will follow if we prove

**Lemma 5.3.** *Assume that  $k \ll \sqrt{N}$ . Let  $\Gamma_{k+\sigma \times k-k+\sigma}(\lambda)$  and  $\Sigma(\lambda)$  be defined as (5.13) and (5.15). On  $\Omega_N(\Lambda_N)$ ,*

$$\frac{1}{\sqrt{N}} \Gamma_{k+\sigma \times k-k+\sigma}(\Lambda_N) \Sigma(\Lambda_N) \Gamma_{k+\sigma \times k-k+\sigma}(\Lambda_N)^* = (1 + |\xi_N(\Lambda_N)|)^2 \mathbf{o}_{\mathbb{P}}(1). \quad (5.26)$$

**Proof.** For the proof, we use the following decomposition ( $T_N(\lambda)$  and  $\Delta_N(\lambda)$  being defined by (5.11) and (5.12)):

$$\begin{aligned} &\Gamma_{k+\sigma \times k-k+\sigma}(\Lambda_N) \Sigma(\Lambda_N) \Gamma_{k+\sigma \times k-k+\sigma}(\Lambda_N)^* \\ &= T_N \Sigma T_N^* + T_N \Sigma \Delta_N(\Lambda_N)^* + \Delta_N(\Lambda_N) \Sigma \Delta_N(\Lambda_N)^* + \Delta_N(\Lambda_N) \Sigma T_N^*, \end{aligned} \quad (5.27)$$

where (using (5.18))

$$\begin{aligned} T_N &:= T_N(\Lambda_N) \\ &= \left[ U_k^* \left( W_k + \frac{1}{\sqrt{N}} Y \hat{G}(\Lambda_N) Y^* \right) U_k \right]_{k+\sigma \times (k-k+\sigma)}^{\nearrow} \\ &= [U_k^* B_{k,N} U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} + \xi_N(\Lambda_N) [U_k^* D_{k,N}(\Lambda_N) U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \end{aligned}$$

and we replaced  $\Sigma(\Lambda_N)$  by  $\Sigma$ . We will prove the following lemma on  $T_N$ .

**Lemma 5.4.** *If  $k \ll N$ ,*

$$\| [U_k^* D_{k,N}(\Lambda_N) U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \| = \mathfrak{o}_{\mathbb{P}}(1). \quad (5.28)$$

*If  $k \ll \sqrt{N}$ ,*

$$\| [U_k^* B_{k,N}(\Lambda_N) U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \|_{\text{HS}} = \mathfrak{o}_{\mathbb{P}}(N^{1/4}) \quad (5.29)$$

*and therefore, for  $k \ll \sqrt{N}$ ,*

$$\| T_N \| = \mathfrak{o}_{\mathbb{P}}(N^{1/4}) (1 + |\xi_N(\Lambda_N)|).$$

**Proof.** To prove (5.28), we use the decomposition

$$c_{\theta_j} [U_k^* D_{k,N}(\Lambda_N) U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} = [U_k^* \tau_N U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} + [U_k^* \phi_N U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow}.$$

As in the proof of Lemma 5.1, we have

$$\mathbb{E}(\| [U_k^* \phi_N U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \|_{\text{HS}}^2 \mathbf{1}_{\Omega_N(\Lambda_N)}) \leq (C + \sigma^4) \frac{k k + \sigma}{N} \frac{1}{(\rho_{\theta_j} - 2\sigma - \delta)^4},$$

so that, for  $k \ll N$  and using Tchebychev inequality, we can deduce that

$$\| [U_k^* \phi_N U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \|_{\text{HS}} \mathbf{1}_{\Omega_N(\Lambda_N)} = \mathfrak{o}_{\mathbb{P}}(1).$$

From (5.21),

$$\| [U_k^* \tau_N U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \| = \mathfrak{o}_{\mathbb{P}}(1)$$

and therefore

$$\| [U_k^* D_{k,N}(\Lambda_N) U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \| = \mathfrak{o}_{\mathbb{P}}(1).$$

Thus, (5.28) is established.

For (5.29), recall that  $[U_k^* B_{k,N} U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} = [U_k^* W_k U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} + \frac{1}{\sqrt{N}} [U_k^* Y \hat{G}(\rho_{\theta_j}) Y^* U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow}$ . Since  $\|W_k\| = \mathfrak{O}_{\mathbb{P}}(\sqrt{k})$ , we have  $\| [U_k^* W_k U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \|^2 = \mathfrak{O}_{\mathbb{P}}(\sqrt{k})$ . Hence, as  $k \ll \sqrt{N}$ , we can deduce that  $\| [U_k^* W_k U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \|_{\text{HS}} = \mathfrak{O}_{\mathbb{P}}(N^{1/4})$ .

Now, let us prove the same estimate for the remaining term. Using the same proof as in (5.23), one can get that for  $p \neq q$ , for some constant  $C > 0$ ,

$$\mathbb{E}(|\mathcal{U}(p)^* \hat{G}(\rho_{\theta_j}) \mathcal{U}(q)|^2 \mathbf{1}_{\Omega_N^{(2)}}) \leq C \mathbb{E} \text{Tr}(\hat{G}(\rho_{\theta_j})^2 \mathbf{1}_{\Omega_N^{(2)}})$$

and then that for some constant  $C > 0$ ,

$$\mathbb{E} \left[ \left\| \frac{1}{\sqrt{N}} [U_k^* Y \hat{G}(\rho_{\theta_j}) Y^* U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \right\|_{\text{HS}}^2 \mathbf{1}_{\Omega_N^{(2)}} \right] \leq C k k_{+\sigma} \frac{1}{(\rho_{\theta_j} - 2\sigma - \delta)^2}.$$

Then using that

$$\begin{aligned} & \mathbb{P} \left( \left\| \frac{1}{\sqrt{N}} [U_k^* Y \hat{G}(\rho_{\theta_j}) Y^* U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \right\|_{\text{HS}} \mathbf{1}_{\Omega_N^{(2)}} > \varepsilon N^{1/4} \right) \\ & \leq \frac{1}{\varepsilon^2 \sqrt{N}} \mathbb{E} \left[ \left\| \frac{1}{\sqrt{N}} [U_k^* Y \hat{G}(\rho_{\theta_j}) Y^* U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \right\|_{\text{HS}}^2 \right] \end{aligned}$$

we deduce since  $k \ll \sqrt{N}$  that

$$\left\| \frac{1}{\sqrt{N}} [U_k^* Y \hat{G}(\rho_{\theta_j}) Y^* U_k]_{k+\sigma \times k-k+\sigma}^{\nearrow} \right\|_{\text{HS}} \mathbf{1}_{\Omega_N^{(2)}} = o_{\mathbb{P}}(N^{1/4}).$$

Thus (5.29) and Lemma 5.4 are proved.  $\square$

Using that

$$\|\Sigma\| \leq \frac{1}{\rho_{\theta_j} - 2\sigma - 2\delta}, \quad (5.30)$$

one can readily notice that Lemma 5.4 leads to

$$\frac{1}{\sqrt{N}} T_N \Sigma T_N^* = (1 + |\xi_N(\Lambda_N)|)^2 o_{\mathbb{P}}(1). \quad (5.31)$$

We now consider the remaining terms in the r.h.s. of (5.27). We first show the following result where we recall that  $\Delta_N(\rho_{\theta_j}) = [U_k^* Y \hat{G}(\rho_{\theta_j}) \tilde{A}_{N-k \times k}]_{k+\sigma \times k-k+\sigma}^{\nearrow}$ .

**Lemma 5.5.**  $\frac{1}{\sqrt{N}} T_N \Sigma \Delta_N(\rho_{\theta_j})^*$ ,  $\frac{1}{\sqrt{N}} \Delta_N(\rho_{\theta_j}) \Sigma \Delta_N(\rho_{\theta_j})^*$  and  $\frac{1}{\sqrt{N}} \Delta_N(\rho_{\theta_j}) \Sigma T_N^*$  are all equal to some  $(1 + |\xi_N(\Lambda_N)|) o_{\mathbb{P}}(1)$ .

**Proof.** We will show that, on  $\Omega_N(\Lambda_N)$ , for any  $u > 0$ ,

$$\Delta_N(\rho_{\theta_j}) = o_{\mathbb{P}}(N^u). \quad (5.32)$$

One can readily see that this leads to the announced result combining Lemma 5.4, (5.30) and (5.32).

First, using the fact that  $U_k^* Y$  is independent of  $\mathbf{1}_{\Omega_N^{(2)}} \hat{G}(\rho_{\theta_j})$  and that for any  $p$ , the random vector  $\mathcal{U}(p) = {}^t[(Y^* U_k)_{1,p}, \dots, (Y^* U_k)_{N-k,p}]$  has independent centered entries with variance  $\sigma^2$ , one has that

$$\begin{aligned} \mathbb{E}(\mathbf{1}_{\Omega_N(\Lambda_N)} \text{Tr} \Delta_N(\rho_{\theta_j}) \Delta_N(\rho_{\theta_j})^*) & \leq \mathbb{E}(\mathbf{1}_{\Omega_N^{(2)}} \text{Tr} \Delta_N(\rho_{\theta_j}) \Delta_N(\rho_{\theta_j})^*) \\ & = k_{+\sigma} \sigma^2 \mathbb{E} \{ \mathbf{1}_{\Omega_N^{(2)}} \text{Tr} [\hat{G}^2(\rho_{\theta_j}) \tilde{A}_{N-k \times k-k+\sigma} \tilde{A}_{N-k \times k-k+\sigma}^*] \} \\ & \leq k_{+\sigma} \sigma^2 \mathbb{E} \{ \mathbf{1}_{\Omega_N^{(2)}} \|\hat{G}(\rho_{\theta_j})\|^2 \text{Tr} \tilde{A}_{N-k \times k-k+\sigma} \tilde{A}_{N-k \times k-k+\sigma}^* \} \\ & \leq \frac{k_{+\sigma} \sigma^2}{(\rho_{\theta_j} - 2\sigma - \delta)^2} \text{Tr} \mathbf{A}_N^2 \\ & = \frac{k_{+\sigma} \sigma^2}{(\rho_{\theta_j} - 2\sigma - \delta)^2} \sum_{l=1}^J k_l \theta_l^2. \end{aligned}$$

Therefore,  $\mathbb{P}(\mathbf{1}_{\Omega_N(\Lambda_N)} \|\Delta_N(\rho_{\theta_j})\|_{\text{HS}} \geq \varepsilon N^u) \leq \varepsilon^{-2} N^{-2u} \mathbb{E}(\mathbf{1}_{\Omega_N(\Lambda_N)} \|\Delta_N(\rho_{\theta_j})\|_{\text{HS}}^2)$  goes to zero as  $N$  tends to infinity. Hence (5.32) holds true on  $\Omega_N(\Lambda_N)$  and the proof of Lemma 5.5 is complete.  $\square$

Let us now prove that

**Lemma 5.6.**  $\Delta_N(\Lambda_N) = \Delta_N(\rho_{\theta_j}) + \mathcal{O}_{\mathbb{P}}(|\xi_N(\Lambda_N)|)$ .

**Proof.** We have

$$\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j}) = -(\Lambda_N - \rho_{\theta_j}) [U_k^* Y \hat{G}(\rho_{\theta_j}) \hat{G}(\Lambda_N) \tilde{A}_{N-k \times k}]_{k_{+\sigma} \times k_{-k_{+\sigma}}}^{\nearrow}.$$

Let us define  $\nabla_{k_{+\sigma}} = [U_k^* Y \hat{G}(\rho_{\theta_j}) \hat{G}(\Lambda_N) \tilde{A}_{N-k \times k}]_{k_{+\sigma} \times k_{-k_{+\sigma}}}^{\nearrow}$ . Then for some constant  $C > 0$  depending on the matrix  $\tilde{A}_{N-k \times k}$ ,

$$\begin{aligned} \text{Tr}(\nabla_{k_{+\sigma}} \nabla_{k_{+\sigma}}^*) &\leq C \|\hat{G}(\rho_{\theta_j})\|^2 \|\hat{G}(\Lambda_N)\|^2 \text{Tr}(\mathcal{U}^* \mathcal{U}) \\ &\leq \frac{C}{(\rho_{\theta_j} - 2\sigma - 2\delta)^4} \text{Tr}(\mathcal{U}^* \mathcal{U}), \end{aligned}$$

where we denote as before  $\mathcal{U} = [Y^* U_k]_{k_{+\sigma}}^{\leftarrow}$ . Thus letting  $C' := C c_{\theta_j}^{-2}$ ,

$$\|\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j})\|_{\text{HS}}^2 \leq C' (\xi_N(\Lambda_N))^2 \frac{1}{(\rho_{\theta_j} - 2\sigma - 2\delta)^4} \frac{1}{N} \text{Tr}(\mathcal{U}^* \mathcal{U}).$$

From Lemma 5.2, it follows that

$$\frac{1}{N} \text{Tr}(\mathcal{U}^* \mathcal{U}) \xrightarrow{\mathbb{P}} k_{+\sigma} \sigma^2$$

implying that  $\|\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j})\|_{\text{HS}} = \mathcal{O}_{\mathbb{P}}(|\xi_N(\Lambda_N)|)$ .  $\square$

We are now in position to conclude the proof of Lemma 5.3. Indeed, writing

$$\Delta_N(\Lambda_N) \Sigma T_N^* = (\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j})) \Sigma T_N^* + \Delta_N(\rho_{\theta_j}) \Sigma T_N^*$$

and

$$\begin{aligned} \Delta_N(\Lambda_N) \Sigma \Delta_N(\Lambda_N) &= \Delta_N(\rho_{\theta_j}) \Sigma \Delta_N(\rho_{\theta_j})^* \\ &\quad + (\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j})) \Sigma \Delta_N(\rho_{\theta_j})^* \\ &\quad + (\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j})) \Sigma (\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j}))^* \\ &\quad + \Delta_N(\rho_{\theta_j}) \Sigma (\Delta_N(\Lambda_N) - \Delta_N(\rho_{\theta_j}))^*, \end{aligned}$$

we deduce from Lemmas 5.4, 5.6 and (5.30), (5.32) that  $\frac{1}{\sqrt{N}} \Delta_N(\Lambda_N) \Sigma T_N^*$  and  $\frac{1}{\sqrt{N}} \Delta_N(\Lambda_N) \Sigma \Delta_N(\Lambda_N)^*$  are both equal to some  $(1 + |\xi_N(\Lambda_N)|) \mathcal{O}_{\mathbb{P}}(1)$ . Using also (5.31), we can deduce that

$$\frac{1}{\sqrt{N}} \Gamma_{k_{+\sigma} \times k_{-k_{+\sigma}}}(\Lambda_N) \Sigma \Gamma_{k_{+\sigma} \times k_{-k_{+\sigma}}}(\Lambda_N)^* = (1 + |\xi_N(\Lambda_N)|)^2 \mathcal{O}_{\mathbb{P}}(1) \quad (5.33)$$

which gives (5.26) and completes the proof of Lemma 5.3.  $\square$

Combining all the preceding, we have established Proposition 5.2. We now prove that provided it converges in distribution, with a probability going to one as  $N$  goes to infinity,  $\xi_N(\Lambda_N)$  is actually an eigenvalue of a matrix of size  $k_j$ .



**Lemma 5.7.** For all  $u > 0$ ,

$$\frac{\|V_{k+\sigma, N}\|_{\text{HS}}}{N^u} = o_{\mathbb{P}}(1).$$

**Proof.** Straightforward computations lead to the existence of some constant  $C$  such that

$$\mathbb{E}(\|[U_k^* W_k U_k]_{k+\sigma}\|_{\text{HS}}) \leq C.$$

The convergence of  $\|[U_k^* W_k U_k]_{k+\sigma}\|/N^u$  in probability towards zero readily follows by Tchebychev inequality. Following the proof in Lemma 5.1 of the convergence in probability of  $[U_k^* \Phi_N U_k]_{k+\sigma}$  towards zero, one can get that

$$\mathbb{E}\left(\left\|[U_k^* \frac{1}{\sqrt{N}} \mathbf{1}_{\Omega_N^{(2)}}(Y \hat{G}(\rho_{\theta_j}) Y^* - \sigma^2 \text{Tr} \hat{G}(\rho_{\theta_j}) I_k) U_k]_{k+\sigma}\right\|_{\mathbf{1}_{\Omega_N(\Lambda_N)}}^2\right) \leq \frac{(C + \sigma^4) k_{+\sigma}^2}{(\rho_{\theta_j} - 2\sigma - \delta)^2},$$

and the convergence in probability towards zero of the term inside the above expectation follows by Tchebychev inequality. Since moreover according to Lemma A.3,

$$\frac{1}{\sqrt{N}} \mathbf{1}_{\Omega_N^{(2)}} \left( \text{Tr} \hat{G}(\rho_{\theta_j}) - (N - k) \frac{1}{\theta_j} \right) = o_{\mathbb{P}}(1),$$

we can deduce that

$$N^{-u} \left\|[U_k^* \frac{1}{\sqrt{N}} \mathbf{1}_{\Omega_N^{(2)}} \left( Y \hat{G}(\rho_{\theta_j}) Y^* - (N - k) \frac{\sigma^2}{\theta} I_k \right) U_k]_{k+\sigma}\right\|_{\mathbf{1}_{\Omega_N^{(2)}}} = o_{\mathbb{P}}(1).$$

The proof of Lemma 5.7 is complete.  $\square$

**Proposition 5.3.** Let  $\Delta_{k_j}$  be an arbitrary  $k_j \times k_j$  random matrix. If  $\xi_N(\Lambda_N)$  converges in distribution, then, with a probability going to one as  $N$  goes to infinity, it is an eigenvalue of  $\mathbf{X}_{k+\sigma, N}(\Lambda_N) + \text{diag}(\Delta_{k_j}, 0)$  iff  $\xi_N(\Lambda_N)$  is an eigenvalue of a matrix  $\check{X}_{k_j, N}(\Lambda_N) + \Delta_{k_j}$  of size  $k_j$ , satisfying

$$\check{X}_{k_j, N}(\Lambda_N) = V_{k_j, N} + o_{\mathbb{P}}(1), \tag{5.34}$$

where  $V_{k_j, N}$  is the  $k_j \times k_j$  element in the block decomposition of  $V_{k+\sigma, N}$  defined by (5.6); namely

$$V_{k_j, N} = U_{K_j \times k_j}^* [B_{k, N}]_{K_j}^{\wedge} U_{K_j \times k_j}$$

with  $U_{K_j \times k_j}$  and  $B_{k, N}$  defined respectively by (2.3) and (5.5).

**Proof.** Since  $\xi_N(\Lambda_N)$  converges in distribution, we can write the matrix  $\mathbf{X}_{k+\sigma, N}(\Lambda_N)$  given by (5.19) as

$$\mathbf{X}_{k+\sigma, N}(\Lambda_N) = \sqrt{N} \text{diag}(0_{k_j}, ((\theta_l - \theta_j) I_{k_l})_{l \neq j}) + \check{R}_{k+\sigma, N}(\Lambda_N),$$

where  $\check{R}_{k+\sigma, N}(\Lambda_N) := V_{k+\sigma, N} + o_{\mathbb{P}}(1)$ . Let us decompose  $\mathbf{X}_{k+\sigma, N}(\Lambda_N)$  in blocks as

$$\mathbf{X}_{k+\sigma, N}(\Lambda_N) = \begin{pmatrix} X_{k_j, N} & X_{k_j \times k+\sigma-k_j, N} \\ X_{k+\sigma-k_j \times k_j, N} & X_{k+\sigma-k_j, N} \end{pmatrix}.$$

We first show that  $\xi_N(\Lambda_N)$  is not an eigenvalue of  $X_{k+\sigma-k_j, N}$ . Let  $\alpha = \inf_{l \neq j} |\theta_l - \theta_j| > 0$ . Since,

$$X_{k+\sigma-k_j, N} = \sqrt{N} \text{diag}(((\theta_l - \theta_j) I_{k_l})_{l \neq j}) + \check{R}_{k+\sigma-k_j, N},$$

if  $\mu$  is an eigenvalue of  $X_{k+\sigma-k_j}$ , then

$$|\mu|/\sqrt{N} \geq \alpha - \|\check{R}_{k+\sigma-k_j,N}\|/\sqrt{N}.$$

Now, using Lemma 5.7,

$$\|\check{R}_{k+\sigma-k_j,N}\|/\sqrt{N} = o_{\mathbb{P}}(1).$$

Hence  $\xi_N(\Lambda_N)$  cannot be an eigenvalue of  $X_{k+\sigma-k_j,N}$ . Therefore, we can define

$$\begin{aligned} \check{X}_{k_j,N} &= X_{k_j,N} - X_{k_j \times k_{+\sigma-k_j},N} (X_{k_{+\sigma-k_j},N} - \xi_N(\Lambda_N) I_{k_{+\sigma-k_j}})^{-1} X_{k_{+\sigma-k_j} \times k_j,N} \\ &= V_{k_j,N} - \check{R}_{k_j \times k_{+\sigma-k_j},N} (X_{k_{+\sigma-k_j},N} - \xi_N(\Lambda_N) I_{k_{+\sigma-k_j}})^{-1} \check{R}_{k_{+\sigma-k_j} \times k_j,N} + o_{\mathbb{P}}(1). \end{aligned}$$

To get (5.34), it remains to show that

$$\|\check{R}_{k_j \times k_{+\sigma-k_j},N} (X_{k_{+\sigma-k_j},N} - \xi_N(\Lambda_N) I_{k_{+\sigma-k_j}})^{-1} \check{R}_{k_{+\sigma-k_j} \times k_j,N}\| = o_{\mathbb{P}}(1).$$

This follows from the previous computations showing that (for some constant  $C > 0$ )

$$\|(X_{k_{+\sigma-k_j},N} - \xi_N(\Lambda_N) I_{k_{+\sigma-k_j}})^{-1}\| \leq (C + o_{\mathbb{P}}(1))/\sqrt{N},$$

combined with the definition of  $\check{R}_{k_{+\sigma},N}(\Lambda_N)$  and Lemma 5.7. The statement of the proposition then follows from (A.1).  $\square$

*Step 3:* We now examine the convergence of the  $k_j \times k_j$  matrix  $V_{k_j,N} = U_{K_j \times k_j}^* [B_{k,N}]_{K_j}^{\searrow} U_{K_j \times k_j}$

**Proposition 5.4.** *The  $k_j \times k_j$  matrix  $V_{k_j,N} = U_{K_j \times k_j}^* [B_{k,N}]_{K_j}^{\searrow} U_{K_j \times k_j}$  converges in distribution to a  $\text{GU}(\text{O})\text{E}(k_j \times k_j, \frac{\theta_j^2 \sigma^2}{\theta_j^2 - \sigma^2})$  if and only if  $\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}|$  converges to zero when  $N$  goes to infinity.*

**Proof.** Assume that  $\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}|$  converges to zero when  $N$  goes to infinity. We decompose the proof of the convergence of  $U_{K_j \times k_j}^* [B_{k,N}]_{K_j}^{\searrow} U_{K_j \times k_j}$  in distribution to a  $\text{GU}(\text{O})\text{E}(k_j \times k_j, \frac{\theta_j^2 \sigma^2}{\theta_j^2 - \sigma^2})$  into the two following lemmas.

**Lemma 5.8.** *If  $\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}|$  converges to zero when  $N$  goes to infinity then the  $k_j \times k_j$  matrix  $U_{K_j \times k_j}^* [W_k]_{K_j}^{\searrow} U_{K_j \times k_j}$  converges in distribution to a  $\text{GU}(\text{O})\text{E}(k_j \times k_j, \sigma^2)$ .*

**Proof.** First we consider the complex case. Let  $\alpha_{pq} \in \mathbb{C}$ ,  $1 \leq p < q \leq k_j$ , and  $\alpha_{pp} \in \mathbb{R}$ ,  $1 \leq p \leq k_j$ , and define

$$L_N(\alpha) := \sum_{1 \leq p < q \leq k_j} (\alpha_{pq} (U_k^* W_k U_k)_{pq} + \overline{\alpha_{pq} (U_k^* W_k U_k)_{pq}}) + \sum_{1 \leq p \leq k_j} 2\alpha_{pp} (U_k^* W_k U_k)_{pp}.$$

We have

$$L_N(\alpha) = \sum_{i=1}^{K_j} D_i (W_N)_{ii} + \sum_{1 \leq i < l \leq K_j} R_{il} (\sqrt{2} \Re e((W_N)_{il})) + \sum_{1 \leq i < l \leq K_j} I_{il} (\sqrt{2} \Im m((W_N)_{il})),$$

where

$$\begin{aligned} D_i &= 2\Re\left(\sum_{1 \leq p < q \leq k_j} \alpha_{pq}(U_k)_{iq} \overline{(U_k)_{ip}}\right), \\ R_{il} &= \sqrt{2}\Re\left(\sum_{1 \leq p < q \leq k_j} \alpha_{pq}((U_k)_{lq} \overline{(U_k)_{ip}} + (U_k)_{iq} \overline{(U_k)_{lp}})\right), \\ I_{il} &= \sqrt{2}\Im\left(\sum_{1 \leq p < q \leq k_j} \overline{\alpha_{pq}}(\overline{(U_k)_{lq}}(U_k)_{ip} - \overline{(U_k)_{iq}}(U_k)_{lp})\right). \end{aligned}$$

Hence  $L_N(\alpha) = \sum_{m=1}^{K_j^2} \beta_{m,N} \phi_m$  where  $\phi_m$  are i.i.d. random variables with distribution  $\mu$  and  $\beta_{m,N}$  are real constants (depending on the  $\alpha_{pq}$ ) which satisfy  $\max_{m=1}^{K_j^2} |\beta_{m,N}| \rightarrow 0$  when  $N$  goes to infinity. Therefore the cumulants of  $L_N(\alpha)$  are given by  $C_n^{(N)} = \sum_{m=1}^{K_j^2} \beta_{m,N}^n C_n(\mu)$  for any  $n \in \mathbb{N}^*$  where  $C_n(\mu)$  denotes the  $n$ th cumulant of  $\mu$  (all are finite since  $\mu$  has moments of any order). In particular  $C_1^{(N)} = 0$ . We are going to prove that the variance of  $L_N(\alpha)$  is actually constant, given by

$$\frac{C_2^{(N)}}{\sigma^2} = \sum_{m=1}^{K_j^2} \beta_{m,N}^2 = 2 \sum_{1 \leq p < q \leq k_j} |\alpha_{pq}|^2 + 4 \sum_{1 \leq p \leq k_j} |\alpha_{pp}|^2. \quad (5.35)$$

One may rewrite  $L_N(\alpha)$  as

$$L_N(\alpha) = \text{Tr}(H U_k^* W_k U_k),$$

where  $H$  is the  $k \times k$  Hermitian matrix defined by

$$H_{pq} = \alpha_{qp} \quad \text{if } p > q \quad \text{and} \quad H_{pp} = 2\alpha_{pp}.$$

Hence using

$$\mathbb{E}[(W_k)_{ji} (W_k)_{qp}] = \delta_{jp} \delta_{iq} \sigma^2$$

it is easy to see that

$$\begin{aligned} \mathbb{E}[L_N(\alpha)^2] &= \sigma^2 \text{Tr}[(U_k H U_k^*)^2] \\ &= \sigma^2 \text{Tr} H^2. \end{aligned}$$

Then (5.35) readily follows. In the following, we let  $const = \sum_{m=1}^{K_j^2} \beta_{m,N}^2$ .

Since  $|C_n^{(N)}| \leq const \max_{m=1}^{K_j^2} |\beta_{m,N}|^{n-2} |C_n(\mu)|$ ,  $C_n^{(N)}$  converges to zero for each  $n \geq 3$ . Thus we can deduce from Janson's theorem [15] that  $L_N(\alpha)$  converges to a centered Gaussian distribution with variance  $\sigma^2 \times (2 \sum_{1 \leq p < q \leq k_j} |\alpha_{pq}|^2 + 4 \sum_{1 \leq p \leq k_j} |\alpha_{pp}|^2)$  and the proof of Lemma 5.8 is complete in the complex case.

Dealing with symmetric matrices, one needs to consider the random variable

$$L_N(\alpha) := \sum_{1 \leq p < q \leq k_j} \alpha_{pq} (U_k^t W_k U_k)_{pq} + \sum_{1 \leq p \leq k_j} \alpha_{pp} (U_k^t W_k U_k)_{pp}$$

for any real numbers  $\alpha_{pq}$ ,  $p \leq q$ . One can similarly prove that  $L_N(\alpha)$  converges to a centered Gaussian distribution with variance  $\sigma^2 (2 \sum_{1 \leq p < q \leq k_j} \alpha_{pq}^2 + 2 \sum_{1 \leq p \leq k_j} \alpha_{pp}^2)$ .  $\square$

**Remark 5.1.** Note that Lemma 5.8 is true under the assumption of the existence of a fourth moment. This can be shown by using a Taylor development of the Fourier transform of  $L_N(\alpha)$ .

**Lemma 5.9.** If  $\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}|$  converges to zero when  $N$  goes to infinity then the  $k_j \times k_j$  matrix  $\frac{1}{\sqrt{N}} U_{K_j \times k_j}^* [(Y \hat{G}(\rho_{\theta_j}) Y^* - (N-k) \frac{\sigma^2}{\theta_j^2} I_k)]_{K_j}^{\wedge} U_{K_j \times k_j}$  converges towards a  $\text{GU}(\text{O})\text{E}(k_j \times k_j, \frac{\sigma^4}{\theta_j^2 - \sigma^2})$ .

**Proof.** We shall apply a slightly modified version of Theorem A.2 (see Theorem 7.1 in [5]) but requiring the finiteness of sixth (instead of fourth) moments. Let  $K = k_j(k_j + 1)/2$ . The set  $\{1, \dots, K\}$  is indexed by  $l = (p, q)$  with  $1 \leq p \leq q \leq k_j$ , taking the lexicographic order. We define a sequence of i.i.d. centered vectors  $(x_i, y_i)_{i \leq N-k}$  in  $\mathbb{C}^K \times \mathbb{C}^K$  by  $x_{li} = \mathcal{U}_{ip}$  and  $y_{li} = \mathcal{U}_{iq}$  for  $l = (p, q)$  where  $\mathcal{U}$  is defined in Lemma 5.2. The matrix  $A$  of size  $N - k$  is the matrix  $\hat{G}(\rho_{\theta_j})$  and is independent of  $\mathcal{U}$ . Note that we are not exactly in the context of Theorem 7.1 of [5] since the i.i.d. vectors  $(x_i, y_i)_i$  depend on  $N$  (and should be rather denoted by  $(x_{i,N}, y_{i,N})_i$ ) but their distribution satisfies:

1.  $\rho(l) = \mathbb{E}[\bar{x}_{l1} y_{l1}] = \delta_{p,q} \sigma^2$  for  $l = (p, q)$  is independent of  $N$ .
2.  $\mathbb{E}[\bar{x}_{l1} y_{l'1}] = \delta_{p,q'} \sigma^2$  if  $l = (p, q), l' = (p', q')$  (see  $B_2$  in (A.2)).
3. Complex case:  $\mathbb{E}[\bar{x}_{l1} \bar{x}_{l'1}] = \mathbb{E}[y_{l1} y_{l'1}] = 0$  if  $l = (p, q), l' = (p', q')$  (see  $B_3$  in (A.2)).  
Real case:  $\mathbb{E}[\bar{x}_{l1} \bar{x}_{l'1}] = \sigma^2 \delta_{p,p'}$  and  $\mathbb{E}[y_{l1} y_{l'1}] = \sigma^2 \delta_{q,q'}$  if  $l = (p, q), l' = (p', q')$ .
4. (see  $B_1$  in (A.2))

$$\left\{ \begin{array}{l} \mathbb{E}[\bar{x}_{l1} y_{l1} \bar{x}_{l'1} y_{l'1}] = \sigma^4 (\delta_{p,q} \delta_{p',q'} + \delta_{p,q'} \delta_{p',q}) \\ \quad + [\mathbb{E}(|W_{12}|^4) - 2\sigma^4] \sum_{i=1}^{K_j} (U_k)_{i,q} \overline{(U_k)_{i,p}} (U_k)_{i,q'} \overline{(U_k)_{i,p'}} \quad \text{in the complex case,} \\ \mathbb{E}[\bar{x}_{l1} y_{l1} \bar{x}_{l'1} y_{l'1}] = \sigma^4 (\delta_{p,q} \delta_{p',q'} + \delta_{p,q'} \delta_{p',q} + \delta_{p,p'} \delta_{q,q'}) \\ \quad + [\mathbb{E}(|W_{12}|^4) - 3\sigma^4] \sum_{i=1}^{K_j} (U_k)_{i,q} \overline{(U_k)_{i,p}} (U_k)_{i,q'} \overline{(U_k)_{i,p'}} \quad \text{in the real case.} \end{array} \right.$$

Under the assumption that  $\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{i,p}|$  converges to zero when  $k$  goes to infinity, the last term in the r.h.s. of the two above equations tends to 0.

It can be seen that the proof of Theorem 7.1 still holds in this case once we verify that for  $\varepsilon > 0$  and for  $z = x$  or  $y$ , for any  $l$ ,

$$\mathbb{E}[|z_{l1}|^4 \mathbf{1}_{(|z_{l1}| \geq \varepsilon N^{1/4})}] \longrightarrow 0 \quad \text{as } N \rightarrow \infty. \quad (5.36)$$

We postpone the proof of (5.36) to the end of the proof. Assuming that (5.36) holds true, we obtain the CLT Theorem 7.1 ([5]): the Hermitian matrix  $Z_N = (Z_N(p, q))$  of size  $k_j$  defined by

$$Z_N(p, q) = \frac{1}{\sqrt{N-k}} \left[ \sum_{i,i'=1, \dots, N-k} \bar{\mathcal{U}}_{ip} \hat{G}(\rho_{\theta_j})_{ii'} \mathcal{U}_{i'q} - \delta_{p,q} \sigma^2 \text{Tr}(\hat{G}(\rho_{\theta_j})) \right]$$

converges to an Hermitian Gaussian matrix  $G$ . The Laplace transform of  $G$  (considered as a vector of  $\mathbb{C}^K$ , that is of  $\{G_{pq}, 1 \leq p \leq q \leq k_j\}$ ) is given for any  $c \in \mathbb{C}^K$  by

$$\mathbb{E}[\exp(c^T G)] = \exp \left[ \frac{1}{2} c^T B c \right],$$

where the  $K \times K$  matrix  $B = (B(l, l'))$  is given by:  $B = \lim_N B_1(N) + B_2 + B_3$  with

$$B_1(N)(l, l') = \omega (\mathbb{E}[\bar{x}_{l1} y_{l1} \bar{x}_{l'1} y_{l'1}] - \rho(l) \rho(l')),$$

$$B_2(l, l') = (\theta - \omega) \mathbb{E}[\bar{x}_{l1} y_{l'1}] \mathbb{E}[\bar{x}_{l'1} y_{l1}],$$

$$B_3(l, l') = (\tau - \omega) \mathbb{E}[\bar{x}_{l1} \bar{x}_{l'1}] \mathbb{E}[y_{l1} y_{l'1}]$$

and the coefficients  $\omega, \theta, \tau$  are defined in Theorem A.2. Here  $A = \hat{G}(\rho_{\theta_j})$  so that  $\omega = 1/\theta_j^2$  and  $\theta = 1/(\theta_j^2 - \sigma^2)$  (see the Appendix).

From Lemma 5.2,

$$B_2(l, l') = (\theta - \omega)\sigma^4 \delta_{p,q'} \delta_{p',q} = (\theta - \omega)\sigma^4 \mathbf{1}_{p=q=p'=q'}.$$

Moreover in the complex case,  $B_3 \equiv 0$  and in the real case,

$$B_3(l, l') = (\theta - \omega)\sigma^4 \delta_{l,l'}.$$

From 4., in the real case,

$$\lim_{N \rightarrow \infty} B_1(N)(l, l') = \delta_{l,l'} \omega \sigma^4 (1 + \delta_{p,q}),$$

and in the complex case,

$$\lim_{N \rightarrow \infty} B_1(N)(l, l') = \delta_{l,l'} \omega \sigma^4 \delta_{p,q}.$$

It follows that  $B$  is a diagonal matrix given by:

$$\begin{cases} B(l, l) = (1 + \delta_{p,q})\theta\sigma^4 = (1 + \delta_{p,q})\frac{\sigma^4}{\theta_j^2 - \sigma^2} & \text{in the real case,} \\ B(l, l) = \delta_{p,q}\theta\sigma^4 = \delta_{p,q}\frac{\sigma^4}{\theta_j^2 - \sigma^2} & \text{in the complex case.} \end{cases}$$

In the real case, the matrix  $B$  is exactly the covariance of the limiting Gaussian distribution  $G$ . It follows that  $G$  is the distribution of the  $\text{GOE}(k_{+\sigma} \times k_{+\sigma}, \sigma^4/(\theta_j^2 - \sigma^2))$ .

In the complex case, from the form of  $B$ , we can conclude that the coordinates of  $G$  are independent ( $B$  diagonal),  $G_{pp}$  has variance  $\sigma^4/(\theta_j^2 - \sigma^2)$  and for  $p \neq q$ ,  $\Re e(G_{pq})$  and  $\Im m(G_{pq})$  are independent with the same variance (since  $B(l, l) = 0$  for  $p \neq q$ ). It remains to compute the variance of  $\Re e(G_{pq})$ . Since the Laplace transform of  $\Re e(Z_N(p, q))$  and  $\Im m(Z_N(p, q))$  can be expressed as a Laplace transform of  $Z_N(p, q)$  and  $\overline{Z_N(p, q)}$ , we shall apply Theorem 7.1 to  $(Z_N(p, q), \overline{Z_N(p, q)})$  that is to the vectors  $x_i = (\mathcal{U}_{ip}, \mathcal{U}_{iq})$  and  $y_i = (\mathcal{U}_{iq}, \mathcal{U}_{ip})$  in  $\mathbb{C}^2$ . We denote by  $\tilde{B}$  the associated ‘‘covariance’’ matrix of size 2. The variance of  $\Re e(G_{pq})$  is given by  $\frac{1}{2} \lim_{N \rightarrow \infty} \tilde{B}_{12}$  (since  $\tilde{B}_{11} = \tilde{B}_{22} = 0$  from the previous computations) with

$$\tilde{B}_{12} = \tilde{B}_{12}(1) + \tilde{B}_{12}(2) + \tilde{B}_{12}(3),$$

where here  $\tilde{B}_{12}(3) = 0$ ,

$$\tilde{B}_{12}(1) = \omega \mathbb{E}[|\mathcal{U}_{1p}|^2 |\mathcal{U}_{1q}|^2] \rightarrow \omega \sigma^4 \quad \text{and} \quad \tilde{B}_{12}(2) = (\theta - \omega) \mathbb{E}[|\mathcal{U}_{1p}|^2] \mathbb{E}[|\mathcal{U}_{1q}|^2] = (\theta - \omega) \sigma^4.$$

Therefore,  $\text{var}(\Re e(G_{pq})) = \theta \sigma^4 / 2 = \sigma^4 / (2(\theta_j^2 - \sigma^2))$ .

We thus obtain Lemma 5.9 by using that  $\text{Tr}(\hat{G}(\rho_{\theta_j})) = (N - k) \text{tr}_{N-k}(\hat{G}(\rho_{\theta_j}))$  and  $\text{tr}_{N-k}(\hat{G}(\rho_{\theta_j})) \rightarrow 1/\theta_j$ .

It remains to prove (5.36). The variable  $\alpha_N := |z_{11}|^4 \mathbf{1}_{(|z_{11}| \geq \varepsilon N^{1/4})}$  tends to 0 in probability. It is thus enough to prove uniform integrability of the sequence  $\alpha_N$ , a sufficient condition is given by  $\sup_N \mathbb{E}[\alpha_N^{6/4}] < \infty$ . It is easy to see that for any  $1 \leq p \leq k_j$ ,  $\sup_N \mathbb{E}[|\mathcal{U}_{1p}|^6] < \infty$  since the Wigner matrix  $\mathbf{W}_N$  has finite sixth moment and  $U_k$  is unitary. This proves (5.36) and finishes the proof of Lemma 5.9.  $\square$

Assume now that the matrix  $V_{k_j, N} = U_{K_j \times k_j}^* [B_{k, N}]_{K_j}^{\wedge} U_{K_j \times k_j}$  converges in distribution towards a  $\text{GU}(\text{O})\text{E}(k_j \times k_j, \frac{\theta_j^2 \sigma^2}{\theta_j^2 - \sigma^2})$  whereas  $\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}|$  does not converge to zero when  $N$  goes to infinity. There exists  $p_0 \in \{1, \dots, k_j\}$  such that  $\max_{i=1}^{K_j} |(U_k)_{ip}|$  does not converge to zero. Let  $i_N$  be such that  $\max_{i=1}^{K_j} |(U_k)_{ip_0}| = |(U_k)_{i_N p_0}|$ . Now we have

$$(V_{k_j, N})_{p_0 p_0} = |(U_k)_{i_N p_0}|^2 W_{i_N i_N} + X_N,$$

where  $X_N$  is a random variable which is independent with  $|(U_k)_{i_N p_0}|^2 W_{i_N i_N}$ . One can find a subsequence such that  $|(U_k)_{i_{\phi(N)} p_0}|^2 W_{i_{\phi(N)} i_{\phi(N)}}$  converges in distribution towards  $c\xi$  where  $c > 0$  and  $\xi$  is  $\mu$ -distributed. This leads to a contradiction using Cramer–Lévy’s Theorem since  $(V_{k_j, \phi(N)})_{p_0 p_0}$  converges towards a Gaussian variable. The proof of Proposition 5.4 is complete.  $\square$

In the Case (a), condition of Proposition 5.4 are obviously not satisfied and we have the following asymptotic result.

**Proposition 5.5.** *In Case (a), the Hermitian (resp. symmetric) matrix  $V_{k_j, N}$  converges in distribution towards the law of  $V_{k_j \times k_j}$  of size  $k_j$  defined in the following way. Let  $W_{\tilde{K}_j}$  be a Wigner matrix of size  $\tilde{K}_j$  with distribution given by  $\mu$  (cf. (i)) and  $H_{\tilde{K}_j}$  be a centered Hermitian (resp. symmetric) Gaussian matrix of size  $\tilde{K}_j$  independent of  $W_{\tilde{K}_j}$  with independent entries  $H_{pl}$ ,  $p \leq l$ , with variance*

$$\begin{cases} v_{pp} = \mathbb{E}(H_{pp}^2) = \frac{t}{4} \left( \frac{m_4 - 3\sigma^4}{\theta_j^2} \right) + \frac{t}{2} \frac{\sigma^4}{\theta_j^2 - \sigma^2}, & p = 1, \dots, \tilde{K}_j, \\ v_{pl} = \mathbb{E}(|H_{pl}|^2) = \frac{\sigma^4}{\theta_j^2 - \sigma^2}, & 1 \leq p < l \leq \tilde{K}_j. \end{cases} \quad (5.37)$$

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

$$V_{k_j \times k_j} = \tilde{U}_{\tilde{K}_j \times k_j}^* (W_{\tilde{K}_j} + H_{\tilde{K}_j}) \tilde{U}_{\tilde{K}_j \times k_j}. \quad (5.38)$$

The proof follows from Theorem A.2 and is omitted since we have detailed the similar proof of Lemma 5.9.

*Step 4:* We are now in position to prove that

$$\left( \xi_N(\lambda_{\hat{k}_{j-1}+1}(\mathbf{M}_N)), \dots, \xi_N(\lambda_{\hat{k}_{j-1}+k_j}(\mathbf{M}_N)) \right) \xrightarrow{\mathcal{L}} (\lambda_1(V_{k_j \times k_j}), \dots, \lambda_{k_j}(V_{k_j \times k_j})). \quad (5.39)$$

To prove (5.39), our strategy will be indirect: we start from the matrix  $V_{k_j, N}$  and its eigenvalues  $(\lambda_i(V_{k_j, N}); 1 \leq i \leq k_j)$  and we will reverse the previous reasoning to raise to the normalized eigenvalues  $\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N))$ ,  $1 \leq i \leq k_j$ . This approach works in both Cases (a) and (b) as we now explain.

First, for any  $1 \leq i \leq k_j$ , we define  $\Lambda_N^{(i)}$  such that

$$\xi_N(\Lambda_N^{(i)}) = \lambda_i(V_{k_j, N}),$$

that is  $\Lambda_N^{(i)} = \rho_{\theta_j} + \lambda_i(V_{k_j, N})/c_{\theta_j} \sqrt{N}$ .

Since  $V_{k_j, N}$  converges in distribution towards  $V_{k_j \times k_j}$ ,  $\lambda_i(V_{k_j, N})$  also converges in distribution towards  $\lambda_i(V_{k_j \times k_j})$ . Hence  $\xi_N(\Lambda_N^{(i)})$  converges in distribution and  $\Lambda_N^{(i)}$  converges in probability towards  $\rho_{\theta_j}$ . Let  $\check{X}_{k_j}^{(i)} \equiv \check{X}_{k_j, N}(\Lambda_N^{(i)}) = V_{k_j, N} + o_{\mathbb{P}}(1)$  as defined in Proposition 5.3. This fit choice of  $\Lambda_N^{(i)}$  gives that

$$\lambda_i(\check{X}_{k_j}^{(i)}) = \xi_N(\Lambda_N^{(i)}) + \varepsilon_i \quad \text{with } \varepsilon_i = o_{\mathbb{P}}(1).$$

Hence,  $\xi_N(\Lambda_N^{(i)})$  is an eigenvalue of  $\check{X}_{k_j}^{(i)} - \varepsilon_i I_{k_j}$ .

According to Propositions 5.1 and 5.3, on an event  $\check{\Omega}_N$  whose probability goes to one as  $N$  goes to infinity, there exists some  $l_i$  such that

$$\Lambda_N^{(i)} = \lambda_{l_i} \left( \mathbf{M}_N - \frac{\varepsilon_i}{\sqrt{N}} \text{diag}(I_{k_j}, 0_{N-k_j}) \right).$$

The following lines hold on  $\check{\Omega}_N$ . By using Weyl’s inequalities (Lemma A.1), one has for all  $i \in \{1, \dots, k_j\}$  that

$$\left| \lambda_{l_i} \left( \mathbf{M}_N - \frac{\varepsilon_i}{\sqrt{N}} \text{diag}(I_{k_j}, 0_{N-k_j}) \right) - \lambda_{l_i}(\mathbf{M}_N) \right| \leq \frac{|\varepsilon_i|}{\sqrt{N}}.$$

We then deduce that

$$(\xi_N(\lambda_{l_1}(\mathbf{M}_N)), \dots, \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))) = (\lambda_1(V_{k_j, N}), \dots, \lambda_{k_j}(V_{k_j, N})) + o_{\mathbb{P}}(1) \quad (5.40)$$

and thus

$$(\xi_N(\lambda_{l_1}(\mathbf{M}_N)), \dots, \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))) \xrightarrow{\mathcal{L}} (\lambda_1(V_{k_j \times k_j}), \dots, \lambda_{k_j}(V_{k_j \times k_j})). \quad (5.41)$$

Now, to get (5.39), it is sufficient to prove that

$$\mathbb{P}(l_i = \hat{k}_{j-1} + i; i = 1, \dots, k_j) \rightarrow 1 \quad \text{as } N \rightarrow \infty. \quad (5.42)$$

Indeed, one can notice that on the event  $\{l_i = \hat{k}_{j-1} + i; i = 1, \dots, k_j\}$  the following equality holds true

$$(\xi_N(\lambda_{\hat{k}_{j-1}+1}(\mathbf{M}_N)), \dots, \xi_N(\lambda_{\hat{k}_{j-1}+k_j}(\mathbf{M}_N))) = (\xi_N(\lambda_{l_1}(\mathbf{M}_N)), \dots, \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))). \quad (5.43)$$

Hence, if (5.42) is satisfied then (5.43) combined with (5.41) imply (5.39).

We turn now to the proof of (5.42). The key point is to notice that the  $k_j$  eigenvalues of  $V_{k_j \times k_j}$  have a joint density. This fact is well-known if  $V_{k_j \times k_j}$  is a matrix from the GU(O)E and so when  $K_j$  is infinite (Case (b)). When  $K_j$  is bounded (Case (a)), we call on the following arguments. One can decompose the matrix  $\tilde{U}_{\tilde{K}_j \times k_j}^* H_{\tilde{K}_j} \tilde{U}_{\tilde{K}_j \times k_j}$  appearing in the definition (5.38) of  $V_{k_j \times k_j}$  in the following way

$$\tilde{U}_{\tilde{K}_j \times k_j}^* H_{\tilde{K}_j} \tilde{U}_{\tilde{K}_j \times k_j} = Q_{k_j} + \check{H}_{k_j}$$

with  $\check{H}_{k_j}$  distributed as GU(O)E (using the fact that  $\tilde{U}_{\tilde{K}_j \times k_j}^* \tilde{U}_{\tilde{K}_j \times k_j} = I_{k_j}$ ) and  $Q_{k_j}$  independent from  $\check{H}_{k_j}$ . Hence, the law of  $V_{k_j \times k_j}$  is that of the sum of two random independent matrices: the first one being the matrix  $\check{H}_{k_j}$  distributed as GU(O)E associated to a Gaussian measure with some variance  $\tau$  and the second one being a matrix  $Z_{k_j}$  of the form  $\tilde{U}_{\tilde{K}_j \times k_j}^* W_{\tilde{K}_j} \tilde{U}_{\tilde{K}_j \times k_j} + Q_{k_j}$ . Using the density of the GU(O)E matrix  $\check{H}_{k_j}$  with respect to the Lebesgue measure  $dM$  on Hermitian (resp. symmetric) matrices, decomposing  $dM$  on  $\mathbb{U}_N \times (\mathbb{R}^N)_{\leq}$  (denoting by  $\mathbb{U}_N$  the unitary (resp. orthogonal) group), one can easily see that the distribution of the eigenvalues of  $\check{H}_{k_j} + Z_{k_j}$  is absolutely continuous with respect to the Lebesgue measure  $d\lambda$  on  $\mathbb{R}^n$  with a density given by:

$$f(\lambda_1, \dots, \lambda_N) = \exp\left(-\frac{N}{\tau t} \sum_{i=1}^N \lambda_i^2\right) \prod_{i < j} (\lambda_i - \lambda_j)^{4/t} \mathbb{E}\left(\exp\left\{-\frac{N}{\tau t} \text{Tr} Z_{k_j}^2\right\} I((\lambda_1, \dots, \lambda_N), Z_{k_j})\right) d\lambda,$$

where  $I((\lambda_1, \dots, \lambda_N), Z_{k_j}) = \int \exp(\frac{2}{\tau t} N \text{Tr}(U \text{diag}(\lambda_1, \dots, \lambda_N) U^* Z_{k_j})) m(dU)$  denoting by  $m$  the Haar measure on the unitary (resp. orthogonal) group.

Thus, we deduce that the  $k_j$  eigenvalues of  $V_{k_j \times k_j}$  are distinct (with probability one). Using Portmanteau's Lemma with (5.41) then implies that the event

$$\check{\Omega}'_N := \{\xi_N(\lambda_{l_1}(\mathbf{M}_N)) > \dots > \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))\} \cap \check{\Omega}_N$$

is such that  $\lim_N \mathbb{P}(\check{\Omega}'_N) = 1$ . By Theorem 3.1, we notice that the event

$$\tilde{\Omega}'_N := \{\lambda_{\hat{k}_{j-1}}(\mathbf{M}_N) > \rho_{\theta_j} + \delta > \lambda_{l_1}(\mathbf{M}_N)\} \cap \check{\Omega}'_N \cap \{\lambda_{l_{k_j}}(\mathbf{M}_N) > \rho_{\theta_j} - \delta > \lambda_{\hat{k}_{j-1}+k_j+1}(\mathbf{M}_N)\}$$

also satisfies  $\lim_N \mathbb{P}(\tilde{\Omega}'_N) = 1$ , for  $\delta$  small enough. This leads to (5.42) since  $\tilde{\Omega}'_N \subset \{l_i = i + \hat{k}_{j-1}, i = 1, \dots, k_j\}$ .

The proof of Theorems 3.2 and 3.3 is complete.  $\square$

According to Theorem 3.3, in order to establish Theorem 3.4, we only need to prove that the condition (3.6) is actually necessary for universality of the fluctuations. Hence assume that  $\sqrt{N}(\lambda_{k_1+\dots+k_{j-1}+1}(\mathbf{M}_N) - \rho_{\theta_j}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \frac{t}{2}\sigma_{\theta_j}^2)$ . Propositions 5.1 and 5.3 lead to

$$c_{\theta_j} \sqrt{N}(\lambda_{k_1+\dots+k_{j-1}+1}(\mathbf{M}_N) - \rho_{\theta_j}) = V_{1,N} + o_{\mathbb{P}}(1),$$

where

$$V_{1,N} = U_{K_j \times 1}^* [B_{k,N}]_{K_j}^{\searrow} U_{K_j \times 1}.$$

It follows that  $V_{1,N}$  converges towards the Gaussian distribution  $\mathcal{N}(0, \frac{t}{2} \frac{\theta_j^2 \sigma^2}{\theta_j^2 - \sigma^2})$  and then according to Proposition 5.4,  $\max_{i=1}^{K_j} |(U_k)_{i1}|$  converges to zero when  $N$  goes to infinity.  $\square$

Let  $\theta_j$  such that  $\theta_j > \sigma$  and  $k_j = 1$ . Let us prove now the description given in Section 3.2 of the fluctuations of  $\lambda_{k_1+\dots+k_{j-1}+1}(\mathbf{M}_N)$  for some intermediate situations between Cases (a) and (b). Let  $m$  be a fixed integer number. Assume that for any  $l = 1, \dots, m$   $(U_k)_{l1}$  is independent of  $N$ , whereas  $\max_{m < l \leq K_j} |(U_k)_{l1}| \rightarrow 0$  when  $N$  goes to infinity. Following the proofs of Lemmas 5.8 and 5.9, one can check that  $V_{1,N}$  converges in distribution towards  $\sum_{i,l=1}^m a_{il} \xi_{il} + \mathcal{N}$  in the complex case,  $\sum_{1 \leq l \leq i \leq m} a_{il} \xi_{il} + \mathcal{N}$  in the real case, where  $\xi_{il}$ ,  $(i, l) \in \{1, \dots, m\}^2$ ,  $\mathcal{N}$  are independent random variables such that

- for any  $(i, l) \in \{1, \dots, m\}^2$ , the distribution of  $\xi_{il}$  is  $\mu$ ;

$$a_{il} = \begin{cases} \sqrt{2} \Im(\overline{(U_k)_{l1}} (U_k)_{i1}) & \text{if } i < l, \\ \sqrt{t} \Re(\overline{(U_k)_{l1}} (U_k)_{i1}) & \text{if } i > l, \\ \sqrt{\frac{t}{2}} |(U_k)_{l1}|^2 & \text{if } i = l; \end{cases}$$

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

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

Now, following the lines of Step 4 (using the results of Steps 1 and 2), we can conclude that

$$c_{\theta_j} \sqrt{N}(\lambda_{k_1+\dots+k_{j-1}+1}(\mathbf{M}_N) - \rho_{\theta_j})$$

converges in distribution towards the mixture of  $\mu$ -distributed or Gaussian random variables  $\sum_{i,l=1}^m a_{il} \xi_{il} + \mathcal{N}$  in the complex case,  $\sum_{1 \leq l \leq i \leq m} a_{il} \xi_{il} + \mathcal{N}$  in the real case.  $\square$

## Appendix

In this section, we recall some basic facts on matrices and some results on random sesquilinear forms needed for the proofs of Theorems 3.2 and 3.3.

### A.1. Linear algebra

For Hermitian matrices, denoting by  $\lambda_i$  the decreasing ordered eigenvalues, we have the Weyl's inequalities:

**Lemma A.1 (cf. Theorem 4.3.7 of [14]).** *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

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

In the computation of determinants, we shall use the following formula.

**Lemma A.2 (cf. Theorem A.3 in [3]).** *Let  $A \in \mathcal{M}_k(\mathbb{C})$  and  $D$  be a non-singular matrix of order  $N - k$ . Let also  $B$  and  ${}^t C$  be two matrices of size  $k \times (N - k)$ . Then*

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D) \det(A - BD^{-1}C). \quad (\text{A.1})$$

### A.2. CLT for random sesquilinear forms

In the following, a complex random variable  $x$  will be said *standardized* if  $\mathbb{E}(x) = 0$  and  $\mathbb{E}(|x|^2) = 1$ .

**Theorem A.1 (Lemma 2.7 of [2]).** *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*

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

This theorem is still valid if the i.i.d. standardized coordinates  $Y(i)$  of  $Y_N$  have a distribution depending on  $N$  such that  $\sup_N \mathbb{E}(|Y(i)|^4) < \infty$ .

**Theorem A.2 (cf. [5] or Appendix by J. Baik and J. Silverstein in [8] in the scalar case).** *Let  $A = (a_{ij})$  be a  $N \times N$  Hermitian matrix and  $\{(x_i, y_i), i \leq N\}$  a sequence of i.i.d. centered vectors in  $\mathbb{C}^K \times \mathbb{C}^K$  with finite fourth moment. We write  $x_i = (x_{i1}) \in \mathbb{C}^K$  and  $X(l) = (x_{l1}, \dots, x_{lN})^T$  for  $1 \leq l \leq K$  and a similar definition for the vectors  $\{Y(l), 1 \leq l \leq K\}$ . Set  $\rho(l) = \mathbb{E}[\bar{x}_{l1} y_{l1}]$ . Assume that the following limits exist:*

- (i)  $\omega = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N a_{ii}^2$ ,
- (ii)  $\theta = \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr} A^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i,j=1}^N |a_{ij}|^2$ ,
- (iii)  $\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr} A A^T = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i,j=1}^N a_{ij}^2$ .

Then the  $K$ -dimensional random vector  $\frac{1}{\sqrt{N}}(X(l)^* A Y(l) - \rho(l) \text{Tr } A)$  converges in distribution to a Gaussian complex-valued vector  $G$  with mean zero. The Laplace transform of  $G$  is given by

$$\forall c \in \mathbb{C}^K, \quad \mathbb{E}[\exp(c^T G)] = \exp\left(\frac{1}{2} c^T B c\right),$$

where the  $K \times K$  matrix  $B = (B(l, l'))$  is given by  $B = B_1 + B_2 + B_3$  with:

$$\begin{aligned} B_1(l, l') &= \omega (\mathbb{E}[\bar{x}_{l1} y_{l1} \bar{x}_{l'1} y_{l'1}] - \rho(l) \rho(l')), \\ B_2(l, l') &= (\theta - \omega) \mathbb{E}[\bar{x}_{l1} y_{l'1}] \mathbb{E}[\bar{x}_{l'1} y_{l1}], \\ B_3(l, l') &= (\tau - \omega) \mathbb{E}[\bar{x}_{l1} \bar{x}_{l'1}] \mathbb{E}[y_{l1} y_{l'1}]. \end{aligned} \quad (\text{A.2})$$

### A.3. CLT for the empirical distribution of a Wigner matrix and applications

**Theorem A.3 (Theorem 1.1 in [4]).** *Let  $f$  be an analytic function on an open set of the complex plane including  $[-2\sigma, 2\sigma]$ . If the entries  $((W_N)_{il})_{1 \leq i \leq l \leq N}$  of a general Wigner matrix  $\mathbf{W}_N$  of variance  $\sigma^2$  satisfy the conditions*

- (i) for  $i \neq l$ ,  $\mathbb{E}(|(W_N)_{il}|^4) = \text{const}$ ,
- (ii) for any  $\eta > 0$ ,  $\lim_{N \rightarrow +\infty} \frac{1}{\eta^4 n^2} \sum_{i,l} \mathbb{E}[|(W_N)_{il}|^4 \mathbf{1}_{\{|(W_N)_{il}| \geq \eta \sqrt{N}\}}] = 0$ ,

then  $N(\text{tr}_N(f(\frac{1}{\sqrt{N}}\mathbf{W}_N)) - \int f d\mu_{\text{sc}})$  converges in distribution towards a Gaussian variable, where  $\mu_{\text{sc}}$  is the semi-circle distribution of variance  $\sigma^2$ .

We now prove some convergence results of the resolvent  $\hat{G}$  used in the previous proofs.

Let  $1 \leq j \leq J_{+\sigma}$  and  $k$  such that  $\frac{k}{\sqrt{N}} \rightarrow 0$ .

**Lemma A.3.** *Each of the following convergence holds in probability as  $N \rightarrow \infty$ :*

- (i)  $\sqrt{N}(\text{tr}_{N-k} \hat{G}(\rho_{\theta_j}) - 1/\theta_j) \rightarrow 0$ ,
- (ii)  $\text{tr}_{N-k} \hat{G}^2(\rho_{\theta_j}) \rightarrow \int \frac{1}{(\rho_{\theta_j}-x)^2} d\mu_{\text{sc}}(x) = 1/(\theta_j^2 - \sigma^2)$ ,
- (iii)  $\frac{1}{N-k} \sum_{i=1}^{N-k} (\hat{G}(\rho_{\theta_j})_{ii})^2 \rightarrow (\int \frac{d\mu_{\text{sc}}(x)}{\rho_{\theta_j}-x})^2 = 1/\theta_j^2$ .

**Proof.** We denote by  $G$  the resolvent of the non-deformed Wigner matrix  $W_{N-k}/\sqrt{N}$ .

(i) By Theorem A.3, one knows that  $\sqrt{N}(\text{tr}_{N-k} G(\rho_{\theta_j}) - \int \frac{d\mu_{\text{sc}}(x)}{\rho_{\theta_j}-x})$  converges in probability towards 0. Now, we have  $\int \frac{d\mu_{\text{sc}}(x)}{\rho_{\theta_j}-x} = \frac{1}{\theta_j}$  (see [13], p. 94). It is thus enough to show that

$$\text{tr}_{N-k} \hat{G}(\rho_{\theta_j}) - \text{tr}_{N-k} G(\rho_{\theta_j}) = o_{\mathbb{P}}(1/\sqrt{N}).$$

Let then  $U_{N-k} := U$  (resp.  $D_{N-k}$ ) be a unitary (resp. diagonal) matrix such that  $A_{N-k} = U^* D_{N-k} U$ . Then, one has

$$\begin{aligned} |\text{tr}_{N-k} \hat{G}(\rho_{\theta_j}) - \text{tr}_{N-k} G(\rho_{\theta_j})| &= |\text{tr}_{N-k} (\hat{G}(\rho_{\theta_j}) A_{N-k} G(\rho_{\theta_j}))| \\ &= |\text{tr}_{N-k} (D_{N-k} U^* G(\rho_{\theta_j}) \hat{G}(\rho_{\theta_j}) U)| \\ &:= |\text{tr}_{N-k} (D_{N-k} \Lambda(\rho_{\theta_j}))| \leq (r/(N-k)) \|D_{N-k}\| \|\Lambda(\rho_{\theta_j})\|, \end{aligned}$$

where  $r$  is the finite rank of the perturbed matrix  $A_{N-k}$ .

One has  $\|D_{N-k}\| \leq \|A_N\| := c$  (with  $c = \max(\theta_1, |\theta_J|)$  independent from  $N$ ). Moreover on the event  $\tilde{\Omega}_N := \Omega_N^{(2)} \cap \{\|W_{N-k}/\sqrt{N}\| < 2\sigma + \delta\}$ ,  $\|\Lambda(\rho_{\theta_j})\| \leq (\rho_{\theta_j} - 2\sigma - \delta)^{-2}$  (use (5.1)) so that we deduce that

$$|\text{tr}_{N-k} (\hat{G}(\rho_{\theta_j})) - \text{tr}_{N-k} (G(\rho_{\theta_j}))| \mathbf{1}_{\tilde{\Omega}_N} \leq \frac{rc}{N-k} (\rho_{\theta_j} - 2\sigma - \delta)^{-2} \rightarrow 0.$$

Using Theorem A.3 and the fact that  $\mathbb{P}(\tilde{\Omega}_N) \rightarrow 1$ , we obtain the announced result.

(ii) It is sufficient to show that  $\text{tr}_{N-k} \hat{G}^2(\rho_{\theta_j}) - \text{tr}_{N-k} G^2(\rho_{\theta_j}) \rightarrow 0$  in probability since, by Theorem A.3, one knows that  $\text{tr}_{N-k} G^2(\rho_{\theta_j})$  converges in probability towards  $\int \frac{1}{(\rho_{\theta_j}-x)^2} d\mu_{\text{sc}}(x)$ .

Using the fact that  $\text{Tr}(BC) = \text{Tr}(CB)$ , it is not hard to see that

$$\begin{aligned} \text{tr}_{N-k} \hat{G}^2(\rho_{\theta_j}) - \text{tr}_{N-k} G^2(\rho_{\theta_j}) &= \text{tr}_{N-k} ((\hat{G}(\rho_{\theta_j}) + G(\rho_{\theta_j}))(\hat{G}(\rho_{\theta_j}) - G(\rho_{\theta_j}))) \\ &= \text{tr}_{N-k} (\hat{G}(\rho_{\theta_j}) A_{N-k} G(\rho_{\theta_j}) (G(\rho_{\theta_j}) + \hat{G}(\rho_{\theta_j}))) \\ &= \text{tr}_{N-k} (D_{N-k} U G(\rho_{\theta_j}) (G(\rho_{\theta_j}) + \hat{G}(\rho_{\theta_j})) \hat{G}(\rho_{\theta_j}) U^*) \\ &:= \text{tr}_{N-k} (D_{N-k} \Lambda'(\rho_{\theta_j})), \end{aligned}$$

where the matrices  $D_{N-k}$  and  $U$  have been defined in (i). We then conclude in a similar way as before since on the event  $\tilde{\Omega}_N$ ,  $\|\Lambda'(\rho_{\theta_j})\| \leq 2(\rho_{\theta_j} - 2\sigma - \delta)^{-3}$ .

For point (iii), we refer the reader [8]. Indeed, it was shown in Section 5.2 of [8] that the announced convergence holds in the case  $k = 1$  and for  $G$  instead of  $\hat{G}$ . It is easy to adapt the arguments of [8] which mainly follow from the fact that, for any  $z \in \mathbb{C}$  such that  $\Im m(z) > 0$ ,  $\frac{1}{N-k} \sum_{i=1}^{N-k} (\hat{G}(z)_{ii})^2$  converges towards  $g_{\sigma}^2(z)$ . But this latter convergence was proved in Section 4.1.4 of [8].  $\square$

## Acknowledgments

We would like to thank the anonymous referees for their pertinent comments which led to an overall improvement of the paper.

## References

- [1] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica* **9** (1999) 611–677. [MR1711663](#)
- [2] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** (1998) 316–345. [MR1617051](#)
- [3] Z. D. Bai and J. W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*, 2nd edition. *Springer Ser. Statist.* Springer, New York, 2010.
- [4] Z. D. Bai and J. F. Yao. On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli* **11** (2005) 1059–1092. [MR2189081](#)
- [5] Z. D. Bai and J. F. Yao. Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. H. Poincaré Probab. Statist.* **44** (2008) 447–474. [MR2451053](#)
- [6] J. Baik, G. Ben Arous and S. Péché. Phase transition of the largest eigenvalue for non-null complex sample covariance matrices. *Ann. Probab.* **33** (2005) 1643–1697. [MR2165575](#)
- [7] G. Biroli, J. P. Bouchaud and M. Potters. On the top eigenvalue of heavy-tailed random matrices. *Europhys. Lett.* **78** (2007) Art 10001. [MR2371333](#)
- [8] M. Capitaine, C. Donati-Martin and D. Féral. The largest eigenvalue of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37** (2009) 1–47. [MR2489158](#)
- [9] L. Erdős, H.-T. Yau and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. Preprint, 2010. Available at [arXiv:1007.4652](#).
- [10] D. Féral and S. Péché. The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.* **272** (2007) 185–228. [MR2291807](#)
- [11] D. Féral and S. Péché. The largest eigenvalues of sample covariance matrices for a spiked population: Diagonal case. *J. Math. Phys.* **50** (2009) 073302. [MR2548630](#)
- [12] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica* **1** (1981) 233–241. [MR0637828](#)
- [13] F. Hiai and D. Petz. *The Semicircle Law, Free Random Variables and Entropy. Mathematical Surveys and Monographs* **77**. Amer. Math. Soc., Providence, RI, 2000. [MR1746976](#)
- [14] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge Univ. Press, New York, 1991. [MR1084815](#)
- [15] S. Janson. Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs. *Ann. Probab.* **16** (1988) 305–312. [MR0920273](#)
- [16] O. Khorunzhiy. High moments of large Wigner random matrices and asymptotic properties of the spectral norm. Preprint, 2009. Available at [arXiv:0907.3743v5](#).
- [17] A. Onatski. The Tracy–Widom limit for the largest eigenvalues of singular complex Wishart matrices. *Ann. Appl. Probab.* **18** (2008) 470–490. [MR2398763](#)
- [18] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** (2007) 1617–1641. [MR2399865](#)
- [19] S. Péché. The largest eigenvalues of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* **134** (2006) 127–174. [MR2221787](#)
- [20] A. Ruzmaikina. Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries. *Comm. Math. Phys.* **261** (2006) 277–296. [MR2191882](#)
- [21] A. Soshnikov. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207** (1999) 697–733. [MR1727234](#)
- [22] T. Tao and V. Vu. Random matrices: Universality of the local eigenvalue statistics up to the edge. *Comm. Math. Phys.* **298** (2010) 549–572. [MR2669449](#)
- [23] C. A. Tracy and H. Widom. Level spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159** (1994) 151–174. [MR1257246](#)