Central limit theorems for eigenvalues of deformations of Wigner matrices*

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

Abstract

In this paper, we explain the dependance of the fluctuations of the largest eigenvalues of a Deformed Wigner model with respect to the eigenvectors of the perturbation matrix. We exhibit quite general situations that will give rise to universality or non universality of the fluctuations.

1 Introduction

In a previous paper [C-D-F], we have studied the a.s. behaviour of extremal eigenvalues of finite rank deformation of Wigner matrices and 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. We exhibit a striking non-universality phenomenon at the fluctuations level. Indeed, we prove that the fluctuations of the largest eigenvalue vary with the particular distribution of the entries of the Wigner matrix. Let us recall these results. 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

$$\mathbf{M}_N = \frac{\mathbf{W}_N}{\sqrt{N}} + \mathbf{A}_N \tag{1.1}$$

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

(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 $\theta_1 > \cdots > \theta_J$ independent of N with some j_0 such that $\theta_{j_0} = 0$. We assume that the non-null eigenvalues θ_j of $\mathbf{A}_N$ are of fixed multiplicity k_j (with $\sum_{j \neq j_0} k_j = r$) i.e. $\mathbf{A}_N$ is similar to the diagonal matrix

$$\operatorname{diag}(\theta_1 I_{k_1}, \ldots, \theta_{j_0-1} I_{k_{j_0-1}}, 0_{N-r}, \theta_{j_0+1} I_{k_{j_0+1}}, \ldots, \theta_J I_{k_J}). \tag{1.2}$$

The Poincaré inequality assumption was fundamental in the approach of [C-D-F]. In the present paper, we only rely on the results of [C-D-F] without making use of the Poincaré inequality. Hence, we refer the reader to [C-D-F] and the references therein for details on such an inequality. Nevertheless, note that the Poincaré inequality assumption implies that μ has moments of any order (cf. Corollary 3.2 and Proposition 1.10 in [L]) and this last property will be used later on.
In the following, given an arbitrary Hermitian matrix M of size N, we will denote by $\lambda_1(M) \geq \cdots \geq$

*This work was partially supported by the Agence Nationale de la Recherche grant ANR-08-BLAN-0311-03.

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

‡UPMC Univ Paris 06 and CNRS, UMR 7599, Laboratoire de Probabilités et Modèles Aléatoires, Site Chevaleret, 16 rue Clisson, F-75013 Paris. E-mail: catherine.donati@upmc.fr

§Institut de Mathématiques de Bordeaux, Université Bordeaux 1, 351 Cours de la Libération, F-33405 Talence Cedex. E-mail: delphine.feral@math.u-bordeaux1.fr

$\lambda_N(M)$ its N ordered eigenvalues.
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 [B]). Thus, as in the classical Wigner model ($\mathbf{A}_N \equiv 0$), 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 μ_{sc} whose density is given by

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

Nevertheless, the asymptotic behavior of the extremal eigenvalues may be affected by the perturbation $\mathbf{A}_N$. When $\mathbf{A}_N \equiv 0$, it is well-known that 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 2σ (resp. -2σ) of the semicircle support (cf. [B]). This result fails when some of the θ_j's are sufficiently far from zero. Define

$$\rho_{\theta_j} = \theta_j + \frac{\sigma^2}{\theta_j}. \tag{1.4}$$

Observe that $\rho_{\theta_j} > 2\sigma$ (resp. $< -2\sigma$) when $\theta_j > \sigma$ (resp. $< -\sigma$) (and $\rho_{\theta_j} = \pm 2\sigma$ if $\theta_j = \pm\sigma$).
For definiteness, we set $k_1 + \cdots + k_{j-1} := 0$ if $j = 1$. In [C-D-F], we have established the following universal convergence result.

Theorem 1.1. *(**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) $\forall 1 \le j \le J_{+\sigma}, \forall 1 \le i \le k_j, \quad \lambda_{k_1+\cdots+k_{j-1}+i}(\mathbf{M}_N) \longrightarrow \rho_{\theta_j}$ *a.s.*

(2) $\lambda_{k_1+\cdots+k_{J_{+\sigma}}+1}(\mathbf{M}_N) \longrightarrow 2\sigma$ *a.s.*

(3) $\lambda_{k_1+\cdots+k_{J-J_{-\sigma}}}(\mathbf{M}_N) \longrightarrow -2\sigma$ *a.s.*

(4) $\forall j \ge J - J_{-\sigma} + 1, \forall 1 \le i \le k_j, \quad \lambda_{k_1+\cdots+k_{j-1}+i}(\mathbf{M}_N) \longrightarrow \rho_{\theta_j}$ *a.s.*

In the particular case of the rank one diagonal deformation $\mathbf{A}_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ such that $\theta > \sigma$, we investigated the fluctuations of the largest eigenvalue of $\mathbf{M}_N$ (with $\mathbf{W}_N$ satisfying **(i)**) around its limit ρ_θ. We obtained the following non-universality result.

Theorem 1.2. *(**CLT**) Let $\mathbf{A}_N = \mathrm{diag}(\theta, 0, \cdots, 0)$ and assume that $\theta > \sigma$. Define*

$$c_\theta = \frac{\theta^2}{\theta^2 - \sigma^2} \quad \text{and} \quad v_\theta = \frac{t}{4}\Big(\frac{m_4 - 3\sigma^4}{\theta^2}\Big) + \frac{t}{2}\frac{\sigma^4}{\theta^2 - \sigma^2} \tag{1.5}$$

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

$$c_\theta\sqrt{N}\Big(\lambda_1(\mathbf{M}_N) - \rho_\theta\Big) \xrightarrow{\mathcal{L}} \Big\{\mu * \mathcal{N}(0, v_\theta)\Big\}. \tag{1.6}$$

Remark 1.1. *The strong assumption on the distribution μ (Poincaré inequality) of the entries of $\mathbf{W}_N$ is a technical assumption we needed to prove the a.s. result, Theorem 1.1(we conjecture it is true under more general assumptions, cf. [C-D-F]) but the proof of the fluctuations of Theorem 1.2 only requires standard assumptions (existence of the fourth moment) once we know the a.s. convergence.*

On the other hand, in collaboration with S. Péché, the third author of the present article has stated in [Fe-Pe] the universality of the fluctuations of some Deformed Wigner models under a full deformation $\mathbf{A}_N$ defined by $(A_N)_{ij} = \theta/N$ for all $1 \le i, j \le N$ (see also [Fu-K]). Thus in the non-Gaussian setting, the fluctuations of the largest eigenvalue depend, not only on the spectrum of the deformation $\mathbf{A}_N$, but also on the particular definition of the matrix $\mathbf{A}_N$.

In this paper, we try to explain this dependance of the fluctuations of the largest eigenvalues of the Deformed Wigner model $\mathbf{M}_N$ with respect to the eigenvectors of the matrix $\mathbf{A}_N$. We investigate two quite general situations for which we exhibit a phenomenon of different nature.
First, when the eigenvectors associated to one of the largest eigenvalues of $\mathbf{A}_N$, say $\theta_j > \sigma$, are

not "spread" namely belong to a subspace generated by a fixed number K_j (independent of N) of canonical vectors of $\mathbb{C}^N$ and are independent of N, we establish that the limiting distribution in the fluctuations of $\lambda_{k_1+\cdots+k_{j-1}+i}(\mathbf{M}_N)$, $1 \leq i \leq k_j$, around ρ_{θ_j} is not universal and we give it explicitely in terms of these eigenvectors and of the distribution of the entries of the Wigner matrix.
Secondly, if K_j defined above depends on N, if there is no "leading" direction among the eigenvectors associated to θ_j, we establish the universality of the fluctuations of $\lambda_{k_1+\cdots+k_{j-1}+i}(\mathbf{M}_N)$, $1 \leq i \leq k_j$. We detail these results in the following section. Actually, we assume that the eigenvectors associated to the largest eigenvalues $\theta_1, \ldots, \theta_{J_{+\sigma}}$ of $\mathbf{A}_N$ belong to a subspace generated by $k(= k(N))$ canonical vectors of $\mathbb{C}^N$. In our approach, we need to isolate a $N-k \times N-k$ Deformed Wigner matrix M_{N-k} where the eigenvalues of the perturbation are all smaller than σ; we use several well known limiting results when $N-k$ tends to infinity involving M_{N-k}. Hence, our study does not include the full deformation case of [Fe-Pe] where $k = N$. Moreover for technical reasons we have to assume that $k \ll \sqrt{N}$ but we conjecture that our result still holds if $k \ll N$.

The same kind of questions has been previously studied for the spiked population models by [B-Ya2]. 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 $N \to \infty$) 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 Σ_p having all but finitely many eigenvalues equal to one. The analogue of Theorem 1.1 was established by J. Baik and J. Silverstein in [Bk-S1]: when some eigenvalues of Σ_p are far from one, the corresponding largest eigenvalues of S_N a.s. split from the limiting Marchenko-Pastur support. Fluctuations of the eigenvalues that jump have been recently found by Z. Bai and J. F. Yao in [B-Ya2]: the setting considered in [B-Ya2] is the multiplicative analogue of the particular case "k finite independent of N" in our Theorem 2.1; note that they exhibit universal fluctuations (we refer the reader to [B-Ya2] for the precise restrictions made on the definition of the covariance matrix Σ_p).
Note that the first steps of our approach as well as the approach of [B-Ya2] are in a spirit close to the one of [P] and [B-B-P].

The paper is organized as follows. In Section 2, we present the main results of this paper and give a summary of our approach. In Section 3, we introduce preliminary lemmas and fundamental results which will be of basic use later on. Section 4 is devoted to the proof of Theorem 2.1 and Theorem 2.2. Finally, we prove some technical results in an Appendix. All along the paper, the parameter t is such that $t = 4$ (resp. $t = 2$) in the real (resp. complex) setting and we let $m_4 := \int x^4 d\mu(x)$.

2 Main results

As in Theorem 1.1, we denote by $J_{+\sigma}$ (resp. $J_{-\sigma}$) the number of j's such that $\theta_j > \sigma$ (resp. $\theta_j < -\sigma$). Set $k_{+\sigma} := k_1 + \cdots + k_{J_{+\sigma}}$. We also denote by $(e_i;\, i = 1, \ldots, N)$ the canonical basis of $\mathbb{C}^N$.

We introduce $k \geq k_{+\sigma}$ as the minimal number of canonical vectors of $\mathbb{C}^N$ needed to express all the eigenvectors associated to the largest eigenvalues $\theta_1, \ldots, \theta_{J_{+\sigma}}$ of $\mathbf{A}_N$. Without loss of generality, we can assume that these $k_{+\sigma}$ eigenvectors belong to $\mathrm{Vect}(e_1, \ldots, e_k)$. This follows from the invariance of the distribution of the Wigner matrix $\mathbf{W}_N$ by conjugation by a permutation matrix.

All along the paper we assume that $k \ll \sqrt{N}$. Let us now fix j such that $1 \leq j \leq J_{+\sigma}$. We shall study two cases:

Case a) the orthonormal eigenvectors v_i^j, $1 \leq i \leq k_j$, of $\mathbf{A}_N$ associated to θ_j depend on a finite number K_j (independent of N) of canonical vectors among $(e_1, \ldots, e_k)$ and their coordinates are independent of N ("The eigenvectors don't spread out"). Without loss of generality, we can assume that the v_i^j , $1 \leq i \leq k_j$, belong to $\mathrm{Vect}(e_1, \ldots, e_{K_j})$;

Case b) the orthonormal eigenvectors v_i^j, $1 \leq i \leq k_j$, belong to $\mathrm{Vect}(e_1, \ldots, e_{K_j})$ where $K_j =$

$K_j(N) \to \infty$ when $N \to \infty$ and the coordinates satisfy:

$$\forall i \leq k_j, \forall l \leq K_j, \; |(v_i^j, e_l)| \to 0 \text{ as } N \to \infty.$$

("There is no leading direction among the eigenvectors").

Therefore, we assume that there exists a unitary matrix U_k of size k such that

$$\operatorname{diag}(U_k^*, I_{N-k})\mathbf{A}_N \operatorname{diag}(U_k, I_{N-k}) = \operatorname{diag}(\theta_j I_{k_j}, (\theta_l I_{k_l})_{l \leq J_{+\sigma}, l \neq j}, Z_{N-k_{+\sigma}}) \tag{2.1}$$

where $Z_{N-k_{+\sigma}}$ is an Hermitian matrix with eigenvalues strictly smaller than $\theta_{J_{+\sigma}}$.
In the Case **b)**, U_k satisfies

$$\max_{p=1}^{k_j} \max_{i=1}^{K_j} |(U_k)_{ip}| \longrightarrow 0 \quad \text{as} \quad N \to \infty. \tag{2.2}$$

Considering now the vectors v_j^i as vectors in $\mathbb{C}^{K_j}$, we define the $K_j \times k_j$ matrix

$$U_{K_j \times k_j} := \left(v_1^j, \ldots, v_{k_j}^j\right) \tag{2.3}$$

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

$$U_{K_j \times k_j}^* U_{K_j \times k_j} = I_{k_j}. \tag{2.4}$$

Example:

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

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 does not depend of N and in Case b) if $p = p(N) \to +\infty$. For $j = 2$, we are in Case a).

From Theorem 1.1, for all $1 \leq i \leq k_j$, $\lambda_{k_1+\cdots+k_{j-1}+i}(\mathbf{M}_N)$ converges to ρ_{θ_j} a.s.. The main results of our paper are the following two theorems. Let c_{θ_j} be defined by

$$c_{\theta_j} = \frac{\theta_j^2}{\theta_j^2 - \sigma^2}. \tag{2.5}$$

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

Theorem 2.1. *In Case a): the k_j-dimensional vector*

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

converges in distribution to $(\lambda_i(V_{k_j \times k_j}); i = 1, \ldots k_j)$ where $\lambda_i(V_{k_j \times k_j})$ are the ordered eigenvalues of the matrix $V_{k_j \times k_j}$ of size k_j defined in the following way. Let W_{K_j} be a Wigner matrix of size K_j with distribution given by μ (cf **(i)***) and H_{K_j} be a centered Hermitian Gaussian matrix of size K_j independent of W_{K_j} with independent entries H_{pl}, $p \leq l$ with variance*

$$\begin{cases} v_{pp} = E(H_{pp}^2) = \dfrac{t}{4}\Big(\dfrac{m_4 - 3\sigma^4}{\theta_j^2}\Big) + \dfrac{t}{2}\dfrac{\sigma^4}{\theta_j^2 - \sigma^2}, \; p = 1, \ldots, K_j, \\ v_{pl} = \mathbb{E}(|H_{pl}|^2) = \dfrac{\sigma^4}{\theta_j^2 - \sigma^2}, \; 1 \leq p < l \leq K_j. \end{cases} \tag{2.6}$$

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

$$V_{k_j \times k_j} = U_{K_j \times k_j}^*(W_{K_j} + H_{K_j})U_{K_j \times k_j}. \tag{2.7}$$

Case b) exhibits universal fluctuations.

Theorem 2.2. *In Case b): the k_j-dimensional vector*

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

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

Remark 2.1. *The condition $1\ll k\ll\sqrt{N}$ is just a technical condition and we conjecture that our result still holds if $1\ll k\ll N$.*

Remark 2.2. *Note that since μ is symmetric, analogue results can be deduced from Theorem 2.1 and Theorem 2.2 dealing with the lowest eigenvalues of $\mathbf{M}_N$ and the θ_j such that $\theta_j<-\sigma$.*

Before we proceed to the proof of Theorems 2.1 and 2.2, let us give the sketch of our approach which are similar in both cases. To this aim, we define for any random variable λ,

$$\xi_N(\lambda)=c_{\theta_j}\sqrt{N}(\lambda-\rho_{\theta_j}) \tag{2.8}$$

with c_{θ_j} given by (2.5). We also set $\hat{k}_{j-1}:=k_1+\ldots+k_{j-1}$ with the convention that $\hat{k}_0=0$.
The reasoning made in the setting of Theorem 1.2 (for which $k=k_{+\sigma}=1$) relies (following ideas previously developed in [P] and [B-B-P]) 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 [C-D-F] (which corresponds to the following Theorem 3.2 in the scalar case). In the general case, to prove the convergence in distribution of the vector $\left(\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N));i=1,\ldots,k_j\right)$, we will extend, as [B-Ya2], 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 λ which converges in probability towards ρ_{θ_j}. Then, applying factorizations of type (3.1), we prove that λ is an eigenvalue of $\mathbf{M}_N$ iff $\xi_N(\lambda)$ is (on some event having probability going to 1 as $N\to\infty$) an eigenvalue of a $k_j\times k_j$ matrix $\check{X}_{k_j,N}(\lambda)$ of the form

$$\check{X}_{k_j,N}(\lambda)=V_{k_j,N}+R_{k_j,N}(\lambda) \tag{2.9}$$

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 (2.9)) depends on λ, the previous reasoning with $\lambda=\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N)$ for any $1\le i\le k_j$ does not allow us to readily deduce that the k_j normalized eigenvalues $\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N))$, $1\le i\le k_j$ are eigenvalues of a same matrix of the form $V_{k_j,N}+o_{\mathbb{P}}(1)$ and then that

$$(\xi_N(\lambda_{\hat{k}_{j-1}+i}(\mathbf{M}_N));\,1\le i\le k_j)=(\lambda_i(V_{k_j,N});\,1\le i\le k_j)+o_{\mathbb{P}}(1). \tag{2.10}$$

Note that the authors do not develop this difficulty in [B-Ya2] (pp. 464-465). Hence, in the last step of the proof (Step 4 in Section 4), we detail the additional arguments which are needed to get (2.10) 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 μ in the limiting fluctuations and then justifies the non-universality (resp. universality) in Case a) (resp. b)).

The proof is organized in four steps as follows. In Steps 1 and 2, we explain how to obtain (2.9): 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.

3 Basic tools

In this section, we fix some notations and recall some basic facts on matrices and some results on random sesquilinear forms needed for the proofs of Theorems 2.1 and 2.2.

3.1 Linear algebra

For any matrix $M \in \mathcal{M}_N(\mathbb{C})$, we denote by Tr (resp. tr_N) the classical (resp. normalized) trace. $||M||$ is the operator norm of M and $||M||_{HS} := (\mathrm{Tr}(MM^*))^{1/2}$ the Hilbert-Schmidt norm. Spect(M) denotes the spectrum of M.
For $z \in \mathbb{C}\backslash\mathrm{Spect}(M)$, we denote by $G_M(z) = (zI_N - M)^{-1}$ the resolvent of M (we suppress the index M when there is no confusion).

Lemma 3.1. *Let M be an Hermitian matrix and $x \in \mathbb{R}$ such that $x > \lambda_1(M)$; we have*

$$\|G(x)\| \leq \frac{1}{x - \lambda_1(M)}.$$

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

Lemma 3.2. *(cf. Theorem 4.3.7 of [H-J]) Let B and C be two $N \times N$ Hermitian matrices. For any pair of integers j, k such that $1 \leq j, k \leq N$ and $j + k \leq N + 1$, we have*

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

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

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

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

Lemma 3.3. *(cf. Theorem 11.3 page 330 in [B-S2]) Let $A \in \mathcal{M}_k(\mathbb{C})$ and D be a nonsingular matrix of order $N - k$. Let also B and ${}^tC$ be two matrices of size $k \times (N - k)$. Then*

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D) \det(A - BD^{-1}C). \tag{3.1}$$

3.2 Results on 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 3.1. *(Lemma 2.7 [B-S1]) Let $B = (b_{ij})$ be a $N \times N$ Hermitian matrix and Y_N be a vector of size N which contains i.i.d standardized entries with bounded fourth moment. Then there is a constant $K > 0$ such that*

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

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 3.2. *(cf. [B-Ya2] or Appendix by J. Baik and J. Silverstein in [C-D-F] 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 $x_i = (x_{li}) \in \mathbb{C}^K$ and $X(l) = (x_{l1}, \ldots, x_{lN})^T$ 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) $\omega = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N a_{ii}^2$,

(ii) $\theta = \lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}A^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i,j=1}^N |a_{ij}|^2$,

(iii) $\tau = \lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}AA^T = \lim_{N\to\infty} \frac{1}{N} \sum_{i,j=1}^N a_{ij}^2$.

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

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

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

$$\begin{aligned} B_1(l,l') &= \omega(\mathbb{E}[\bar{x}_{l1}y_{l1}\bar{x}_{l'1}y_{l'1}] - \rho(l)\rho(l')) \\ B_2(l,l') &= (\theta-\omega)\mathbb{E}[\bar{x}_{l1}y_{l'1}]\mathbb{E}[\bar{x}_{l'1}y_{l1}] \\ B_3(l,l') &= (\tau-\omega)\mathbb{E}[\bar{x}_{l1}\bar{x}_{l'1}]\mathbb{E}[y_{l1}y_{l'1}]. \end{aligned} \tag{3.2}$$

4 Proofs of Theorem 2.1 and Theorem 2.2

As far as possible, we handle both the proofs of Theorem 2.1 and Theorem 2.2. We will proceed in four steps. First, let us introduce a few notations.

For a $m \times q$ matrix B (or $\mathbf{B}$) and some integers $1 \le p \le m$ and $1 \le l \le q$, we denote respectively by $[B]^{\nwarrow}_{p\times l}$, $[B]^{\nearrow}_{p\times l}$, $[B]^{\swarrow}_{p\times l}$ and $[B]^{\searrow}_{p\times l}$ 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 $\to$ 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

$$\mathbf{W}_N = \begin{pmatrix} W_k & Y \\ Y^* & W_{N-k} \end{pmatrix}. \tag{4.1}$$

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

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

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 (2.1) that

$$\tilde{\mathbf{A}}_N = \mathrm{diag}(\theta_j I_{k_j}, (\theta_l I_{k_l})_{l\le J_{+\sigma}, l\ne j}, Z_{N-k_{+\sigma}}) = \begin{pmatrix} \tilde{A}_k & \tilde{A}_{k\times N-k} \\ \tilde{A}_{N-k\times k} & A_{N-k} \end{pmatrix}. \tag{4.2}$$

Note also that since A_{N-k} is a submatrix of $Z_{N-k_{+\sigma}}$, all its eigenvalues are strictly smaller than σ.
Let $0 < \delta < (\rho_{\theta_j} - 2\sigma)/2$. For any random variable λ, define the events

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

and

$$\Omega_N(\lambda) = \Omega_N^{(1)}(\lambda) \bigcap \Omega_N^{(2)}. \tag{4.3}$$

On $\Omega_N(\lambda)$, neither λ nor ρ_{θ_j} are eigenvalues of $M_{N-k} := \frac{W_{N-k}}{\sqrt{N}} + A_{N-k}$ and the resolvents $\widehat{G}(\lambda) := (\lambda I_{N-k} - M_{N-k})^{-1}$ and $\widehat{G}(\rho_{\theta_j}) := (\rho_{\theta_j} I_{N-k} - M_{N-k})^{-1}$ are well defined. Note that from Theorem 1.1, for any random sequence Λ_N converging towards ρ_{θ_j} in probability, $\lim_{N\longrightarrow\infty} \mathbb{P}(\Omega_N(\Lambda_N)) = 1$.

STEP 1: Let λ be a random variable. On $\Omega_N(\lambda)$,

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

The last equality in the above equation follows from (3.1). Since on $\Omega_N(\lambda)$, λ is not an eigenvalue of M_{N-k}, we can deduce that λ is an eigenvalue of $\tilde{M}_N$ if and only if it is an eigenvalue of

$$\mathbf{Q}_{k,N}(\lambda) := \tilde{M}_k + \tilde{M}_{k\times N-k}\hat{G}(\lambda)\tilde{M}_{N-k\times k}. \tag{4.4}$$

Now, note that we have also from (3.1) that

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

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

$$\Sigma_{k-k_{+\sigma}}(\lambda) := \left([\mathbf{Q}_{k,N}(\lambda)]^{\searrow}_{k-k_{+\sigma}} - \lambda I_{k-k_{+\sigma}}\right)^{-1}. \tag{4.5}$$

Moreover on $\Omega_N(\lambda)$, one can see using (3.1) that if λ_0 is an eigenvalue of $[\mathbf{Q}_{k,N}(\lambda)]^{\searrow}_{k-k_{+\sigma}} - \lambda I_{k-k_{+\sigma}}$ then λ is an eigenvalue of

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

Hence,

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

and then

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

so that finally

$$\|\Sigma_{k-k_{+\sigma}}(\lambda)\| \le \frac{1}{\rho_{\theta_j} - 2\sigma - 2\delta}. \tag{4.6}$$

Using oncemore (3.1), we get that on $\Omega_N(\lambda)$, λ is an eigenvalue of $\mathbf{Q}_{k,N}(\lambda)$ if and only if it is an eigenvalue of $[\mathbf{Q}_{k,N}(\lambda)]^{\nwarrow}_{k_{+\sigma}} - [\mathbf{Q}_{k,N}(\lambda)]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\,\Sigma_{k-k_{+\sigma}}(\lambda)\,[\mathbf{Q}_{k,N}(\lambda)]^{\swarrow}_{k-k_{+\sigma}\times k_{+\sigma}}$ or equivalently if and only if $\xi_N(\lambda)$ is an eigenvalue of

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

Now using

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

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

$$\frac{1}{\sqrt{N}} Y\widehat{G}(\lambda)Y^* = \frac{1}{\sqrt{N}} Y\widehat{G}(\rho_{\theta_j})Y^* + \xi_N(\lambda) D_{k,N}(\lambda) - \xi_N(\lambda)\frac{N-k}{N}\frac{\sigma^2}{c_{\theta_j}(\theta_j^2-\sigma^2)} I_k \tag{4.7}$$

where

$$c_{\theta_j} D_{k,N}(\lambda) = \frac{1}{N}(\lambda-\rho_{\theta_j}) Y\widehat{G}(\lambda)\widehat{G}(\rho_{\theta_j})^2 Y^* - \frac{1}{N}\Big(Y\widehat{G}(\rho_{\theta_j})^2 Y^* - \sigma^2 \operatorname{Tr}\widehat{G}(\rho_{\theta_j})^2 I_k\Big) \\ -\sigma^2 \frac{N-k}{N}\Big(\operatorname{tr}_{N-k}\widehat{G}(\rho_{\theta_j})^2 - \frac{1}{\theta_j^2-\sigma^2}\Big) I_k.$$

Then

$$c_{\theta_j}\sqrt{N}\Big([\mathbf{Q}_{k,N}(\lambda)]^{\nwarrow}_{k_{+\sigma}} - \rho_{\theta_j} I_{k_{+\sigma}}\Big) = c_{\theta_j}\Bigg\{\Big[U_k^*\Big(W_k + \frac{1}{\sqrt{N}}\Big(Y\widehat{G}(\rho_{\theta_j})Y^* - (N-k)\frac{\sigma^2}{\theta_j} I_k\Big)\Big)U_k\Big]^{\nwarrow}_{k_{+\sigma}} \\ +\sqrt{N}\operatorname{diag}\big(0_{k_j}, (\theta_l-\theta_j)I_{k_l}, \, l=1,\ldots,J_{+\sigma}, l\neq j\big) \\ +\xi_N(\lambda)\,[U_k^* D_{k,N}(\lambda) U_k]^{\nwarrow}_{k_{+\sigma}} - \frac{k}{\sqrt{N}}\frac{\sigma^2}{\theta_j} I_{k_{+\sigma}} + \frac{\sigma^2}{\theta_j^2-\sigma^2}\frac{\xi_N(\lambda)}{c_{\theta_j}}\frac{k}{N} I_{k_{+\sigma}}\Bigg\} \\ -\frac{\sigma^2}{\theta_j^2-\sigma^2}\xi_N(\lambda).$$

The following proposition (adding an extra matrix $\Delta_{k_{+\sigma}}$ for future computations) readily follows:

Proposition 4.1. *For any random variable λ and any $k_{+\sigma}\times k_{+\sigma}$ random matrix $\Delta_{k_{+\sigma}}$, on $\Omega_N(\lambda)$, λ is an eigenvalue of $\tilde{\mathbf{M}}_N + \operatorname{diag}(\Delta_{k_{+\sigma}}, 0)$ iff $\xi_N(\lambda)$ is an eigenvalue of $\mathbf{X}_{k_{+\sigma},N}(\lambda) + \sqrt{N}\Delta_{k_{+\sigma}}$ where*

$$\mathbf{X}_{k_{+\sigma},N}(\lambda) := [U_k^* B_{k,N} U_k]^{\nwarrow}_{k_{+\sigma}} + \sqrt{N}\operatorname{diag}\big(0_{k_j}, (\theta_l-\theta_j)I_{k_l}, \, i=1,\ldots,J_{+\sigma}, l\neq j\big) + \xi_N(\lambda)\,[U_k^* D_{k,N}(\lambda)U_k]^{\nwarrow}_{k_{+\sigma}} \\ + \Big(\frac{\sigma^2}{\theta_j^2-\sigma^2}\frac{\xi_N(\lambda)}{c_{\theta_j}}\frac{k}{N} - \frac{k}{\sqrt{N}}\frac{\sigma^2}{\theta_j}\Big) I_{k_{+\sigma}} - \frac{1}{\sqrt{N}}\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\lambda)\Sigma_{k-k_{+\sigma}}(\lambda)\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\lambda)^* \qquad (4.8)$$

where

$$B_{k,N} = W_k + \frac{1}{\sqrt{N}}\Big(Y\widehat{G}(\rho_{\theta_j})Y^* - (N-k)\frac{\sigma^2}{\theta_j} I_k\Big), \qquad (4.9)$$

$$c_{\theta_j} D_{k,N}(\lambda) = \tau_N(\lambda) + \phi_N + \psi_N \quad \textit{with} \qquad (4.10)$$

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

$$\phi_N = -\frac{1}{N}\Big(Y\widehat{G}(\rho_{\theta_j})^2 Y^* - \sigma^2 \operatorname{Tr}\widehat{G}(\rho_{\theta_j})^2 I_k\Big)$$

$$\psi_N = -\sigma^2\frac{N-k}{N}\Big(\operatorname{tr}_{N-k}\widehat{G}(\rho_{\theta_j})^2 - \frac{1}{\theta_j^2-\sigma^2}\Big) I_k,$$

and

$$\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\lambda) = T_N(\lambda) + \Delta_{k_{+\sigma}}(\lambda) \quad \textit{with} \qquad (4.11)$$

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

$$\Delta_{k_{+\sigma}}(\lambda) = \Big[U_k^* Y\hat{G}(\lambda)\tilde{A}_{N-k\times k}\Big]^{\nearrow}_{k_{+\sigma}\times(k-k_{+\sigma})}.$$

Moreover, the $k-k_{+\sigma}\times k-k_{+\sigma}$ matrix $\Sigma_{k-k_{+\sigma}}(\lambda)$ defined by (4.5) is such that

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

Let us make some comments on our approach in order to explain why we proceed in two steps namely we apply twice a factorization of type (3.1) to deal with a $k\times k$ matrix and then with a $k_{+\sigma}\times k_{+\sigma}$ matrix. This approach makes the accommodating resolvent of the Deformed Wigner matrix $\frac{W_{N-k}}{N} + A_{N-k}$ arise. A single application of a factorization of type (3.1) to go from a $N\times N$ to $k_{+\sigma}\times k_{+\sigma}$ matrix would require to deal with the matrix $[\tilde{\mathbf{W}}_N/\sqrt{N}]^{\searrow}_{N-k_{+\sigma}} + Z_{N-k_{+\sigma}}$ whose limiting spectral behaviour is a priori unknown. Moreover, the independance of the matrix $[\tilde{\mathbf{W}}_N/\sqrt{N}]^{\nearrow}_{k\times N-k}$ and $\frac{W_{N-k}}{N}+A_{N-k}$ arising from our two steps approach will be of fundamental use in the following, whereas their analogues in a single step approach are $[\tilde{\mathbf{W}}_N/\sqrt{N}]^{\nearrow}_{k_{+\sigma}\times N-k_{+\sigma}}$ and $[\tilde{\mathbf{W}}_N/\sqrt{N}]^{\searrow}_{N-k_{+\sigma}}+Z_{N-k_{+\sigma}}$ which are not independant.

Throughout Steps 2 and 3, Λ_N denotes any random sequence converging in probability towards ρ_{θ_j}. The aim of these two steps is to study the limiting behavior of the matrix $\mathbf{X}_{k_{+\sigma},N}(\Lambda_N)$ 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 4.2. *Assume that $k\ll\sqrt{N}$. For any random sequence Λ_N converging in probability towards ρ_{θ_j}, on $\Omega_N(\Lambda_N)$,*

$$\mathbf{X}_{k_{+\sigma},N}(\Lambda_N)=V_{k_{+\sigma},N}+\sqrt{N}\mathrm{diag}\left(0_{k_j},(\theta_l-\theta_j)I_{k_l},\,l=1,\ldots,J_{+\sigma},l\neq j\right)+(1+|\xi_N(\Lambda_N)|)^2 o_{\mathbb{P}}(1), \tag{4.12}$$

with $V_{k_{+\sigma},N}$ given by

$$V_{k_{+\sigma},N}:=[U_k^*B_{k,N}U_k]^{\nwarrow}_{k_{+\sigma}}. \tag{4.13}$$

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

$$[U_k^*D_{k,N}(\Lambda_N)U_k]^{\nwarrow}_{k_{+\sigma}}=o_{\mathbb{P}}(1). \tag{4.14}$$

Proof of Lemma 4.1: We refer to Proposition 4.1 for the definition of $D_{k,N}(\Lambda_N)$, τ_N, ϕ_N and ψ_N. Let $K=\mathrm{diag}(I_{k_{+\sigma}},0_{k-k_{+\sigma}})$. On $\Omega_N(\Lambda_N)$,

$$\begin{aligned}\|\,[U_k^*\tau_NU_k]^{\nwarrow}_{k_{+\sigma}}\,\|_{HS} &= \frac{1}{N}|\Lambda_N-\rho_{\theta_j}|\{\mathrm{Tr}(KU_k^*Y\widehat{G}(\Lambda_N)\widehat{G}(\rho_{\theta_j})^2Y^*U_kKU_k^*Y\widehat{G}(\rho_{\theta_j})^2\widehat{G}(\Lambda_N)Y^*U_kK)\}^{\frac12}\\ &\leq \frac{1}{N}|\Lambda_N-\rho_{\theta_j}|\|\widehat{G}(\rho_{\theta_j})\|^2\|\widehat{G}(\Lambda_N)\|\|Y^*U_kKU_k^*Y\|^{\frac12}\{\mathrm{Tr}(KU_k^*YY^*U_kK)\}^{\frac12}\\ &\leq \frac{1}{N}|\Lambda_N-\rho_{\theta_j}|\|\widehat{G}(\rho_{\theta_j})\|^2\|\widehat{G}(\Lambda_N)\|\mathrm{Tr}(KU_k^*YY^*U_kK)\\ &\leq \frac{1}{N}|\Lambda_N-\rho_{\theta_j}|\frac{1}{(\rho_{\theta_j}-2\sigma-2\delta)^3}\mathrm{Tr}(KU_k^*YY^*U_kK).\end{aligned}$$

We have,

$$\begin{aligned}\frac{1}{N}\mathrm{Tr}(KU_k^*YY^*U_kK) &= \frac{1}{N}\sum_{p=1}^{N-k}\sum_{i=1}^{k_{+\sigma}}\sum_{l,q=1}^{k}\overline{(U_k)_{l,i}}W_{l,k+p}\overline{W_{q,k+p}}\,(U_k)_{q,i}\\ &= \sum_{i=1}^{k_{+\sigma}}\frac{1}{N}\sum_{p=1}^{N-k}\sum_{l,q=1,l\neq q}^{k}\overline{(U_k)_{l,i}}W_{l,k+p}\overline{W_{q,k+p}}\,(U_k)_{q,i}\\ &\quad+\sum_{i=1}^{k_{+\sigma}}\frac{1}{N}\sum_{p=1}^{N-k}\sum_{l=1}^{k}|(U_k)_{l,i}|^2|W_{l,k+p}|^2.\end{aligned}$$

Since $\{\sum_{l,q=1,l\neq q}^k\overline{(U_k)_{l,i}}W_{l,k+p}\overline{W_{q,k+p}}\,(U_k)_{q,i};1\leq p\leq N-k\}$ are i.i.d random variables with mean zero and such that the second moments are bounded in N, we can deduce by the law of large numbers that $\frac{1}{N}\sum_{p=1}^{N-k}\sum_{l,q=1,l\neq q}^k\overline{(U_k)_{l,i}}W_{l,k+p}\overline{W_{q,k+p}}\,(U_k)_{q,i}$ converges in L^2 to zero and thus in probability. Similarly, since $\{\sum_{l=1}^k|(U_k)_{l,i}|^2|W_{l,k+p}|^2;1\leq p\leq N-k\}$ are i.i.d random variables with mean σ^2 and bounded (in N) second moments, by the law of large numbers $\frac{1}{N}\sum_{p=1}^{N-k}\sum_{l=1}^k|(U_k)_{l,i}|^2|W_{l,k+p}|^2$ converges in L^2 towards σ^2 and thus in probability. It follows that when $N\to+\infty$,

$$\frac{1}{N}\mathrm{Tr}(KU_k^*YY^*U_kK)\xrightarrow{\mathbb{P}}k_{+\sigma}\sigma^2 \tag{4.15}$$

Hence $[U_k^* \tau_N U_k]^{\nwarrow}_{k_{+\sigma}} = o_{\mathbb{P}}(1)$.

It follows from Lemma 5.1 in the Appendix that

$$[U_k^* \psi_N U_k]^{\nwarrow}_{k_{+\sigma}} := -\sigma^2 \frac{N-k}{N} \Big[\operatorname{tr}_{N-k} \widehat{G}(\rho_{\theta_j})^2 - \frac{1}{\theta_j^2 - \sigma^2} \Big] I_{k_{+\sigma}} = o_{\mathbb{P}}(1).$$

Now, we have

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

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

Lemma 4.2. *Let $\mathcal{U}$ denote the $N-k \times k_{+\sigma}$ matrix $[Y^* U_k]^{\leftarrow}_{k_{+\sigma}}$. Then, the rows $(\mathcal{U}_{i\cdot}; i \leq N-k)$ are centered i.i.d vectors in $\mathbb{C}^{k_{+\sigma}}$, with a distribution depending on N. Moreover, we have for all $1 \leq p, q \leq k_{+\sigma}$:*

$$\mathbb{E}(\mathcal{U}_{1p} \bar{\mathcal{U}}_{1q}) = \delta_{p,q} \sigma^2 \quad \textit{with} \quad \mathbb{E}(\mathcal{U}_{1p} \mathcal{U}_{1q}) = 0 \ \textit{ in the complex case,}$$

$$\mathbb{E}[|\mathcal{U}_{ip}|^2 |\mathcal{U}_{iq}|^2] = (1 + \frac{t}{2} \delta_{p,q}) \sigma^4 + [\mathbb{E}(|W_{12}|^4) - (1 + \frac{t}{2}) \sigma^4] \sum_{l=1}^{k} |(U_k)_{l,p}|^2 |(U_k)_{l,q}|^2. \quad (4.16)$$

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 3.1 and using Lemma 3.1,

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

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

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

where we denote by G the matrix $\widehat{G}(\rho_{\theta_j})$ for simplicity. ¿From Lemma 4.2, 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,

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

and

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

2) $i=j=k=l$. In this case, using (4.16), 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

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

Therefore,

$$\mathbb{E}\left(|\mathcal{U}(p)^*\widehat{G}(\rho_{\theta_j})^2\mathcal{U}(q)|^2\mathbf{1}_{\Omega_N^{(2)}}\right) \leq (C+\sigma^4)\ \mathbb{E}\operatorname{Tr}(\widehat{G}(\rho_{\theta_j})^4\mathbf{1}_{\Omega_N^{(2)}}). \tag{4.17}$$

Hence,

$$\frac{1}{N^2}\mathbb{E}\left(|\mathcal{U}(p)^*\widehat{G}(\rho_{\theta_j})^2\mathcal{U}(q)|^2\mathbf{1}_{\Omega_N^{(2)}}\right) \leq \frac{C+\sigma^4}{N}\mathbb{E}\left(\|\widehat{G}(\rho_{\theta_j})\|^4\mathbf{1}_{\Omega_N^{(2)}}\right) \leq \frac{C+\sigma^4}{N}\,\frac{1}{(\rho_{\theta_j}-2\sigma-\delta)^4}.$$

Thus

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

The convergence in probability of $[U_k^*\phi_N U_k]_{k_{+\sigma}}^{\nwarrow}$ towards zero readily follows by Tchebychev inequality. Lemma 4.1 is established. □

For simplicity, we now write

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

Let us define

$$R_{k,N}(\Lambda_N) := -\frac{k}{\sqrt{N}}\frac{\sigma^2}{\theta_j}I_{k_{+\sigma}} + \frac{\sigma^2}{\theta_j^2-\sigma^2}\frac{\xi_N(\Lambda_N)}{c_{\theta_j}}\frac{k}{N}I_{k_{+\sigma}} - \frac{1}{\sqrt{N}}\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)\Sigma(\Lambda_N)\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)^*. \tag{4.18}$$

To get Proposition 4.2, it remains to prove that if $k \ll \sqrt{N}$,

$$R_{k,N}(\Lambda_N) = (1+|\xi_N(\Lambda_N)|)^2 o_{\mathbb{P}}(1). \tag{4.19}$$

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

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

Hence, (4.19) will follow if we prove

Lemma 4.3. *Assume that $k \ll \sqrt{N}$. On $\Omega_N(\Lambda_N)$,*

$$\frac{1}{\sqrt{N}}\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)\Sigma(\Lambda_N)\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)^* = (1+|\xi_N(\Lambda_N)|)^2 o_{\mathbb{P}}(1). \tag{4.20}$$

For the proof, we use the following decomposition (recall the notations of Proposition 4.1):

$$\begin{aligned}&\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)\Sigma(\Lambda_N)\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)^*\\ &\quad= T_N\Sigma T_N^* + T_N\Sigma\Delta_{k_{+\sigma}}(\Lambda_N)^* + \Delta_{k_{+\sigma}}(\Lambda_N)\Sigma\Delta_{k_{+\sigma}}(\Lambda_N)^* + \Delta_{k_{+\sigma}}(\Lambda_N)\Sigma T_N^*.\end{aligned} \tag{4.21}$$

where (using (4.7))

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

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

Lemma 4.4. *If $k \ll \sqrt{N}$,*

$$\| \, [U_k^* D_{k,N}(\Lambda_N) U_k]^{\nearrow}_{k_{+\sigma} \times k-k_{+\sigma}} \, \|_{HS} = o_{\mathbb{P}}(N^{\frac{1}{4}}). \tag{4.22}$$

$$\| \, [U_k^* B_{k,N}(\Lambda_N) U_k]^{\nearrow}_{k_{+\sigma} \times k-k_{+\sigma}} \, \|_{HS} = o_{\mathbb{P}}(N^{\frac{1}{4}}). \tag{4.23}$$

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

$$\|T_N\| = o_{\mathbb{P}}(N^{\frac{1}{4}})(1 + |\xi_N(\Lambda_N)|).$$

Proof of Lemma 4.4: To prove (4.22), we use the decomposition

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

As in the proof of Lemma 4.1, we have

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

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

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

Let $K = \mathrm{diag}(I_{k_{+\sigma}}, 0)$ and $L = \mathrm{diag}(0, I_{k-k_{+\sigma}})$. On $\Omega_N(\Lambda_N)$,

$$\begin{aligned}
\| \, [U_k^* \tau_N U_k]^{\nearrow}_{k_{+\sigma} \times k-k_{+\sigma}} \, \|_{HS} \quad &= \quad \frac{1}{N} |\Lambda_N - \rho_{\theta_j}| \{ \mathrm{Tr}(K U_k^* Y \widehat{G}(\Lambda_N) \widehat{G}(\rho_{\theta_j})^2 Y^* U_k L U_k^* Y \widehat{G}(\rho_{\theta_j})^2 \widehat{G}(\Lambda_N) Y^* U_k K) \}^{\frac{1}{2}} \\
&\leq \quad \frac{1}{N} |\Lambda_N - \rho_{\theta_j}| \|\widehat{G}(\rho_{\theta_j})\|^2 \|\widehat{G}(\Lambda_N)\| \|Y^* U_k L U_k^* Y\|^{\frac{1}{2}} \{ \mathrm{Tr}(K U_k^* Y Y^* U_k K) \}^{\frac{1}{2}} \\
&\leq \quad \frac{1}{N} |\Lambda_N - \rho_{\theta_j}| \|\widehat{G}(\rho_{\theta_j})\|^2 \|\widehat{G}(\Lambda_N)\| \{ \mathrm{Tr}(L U_k^* Y Y^* U_k L) \}^{\frac{1}{2}} \{ \mathrm{Tr}(K U_k^* Y Y^* U_k K) \}^{\frac{1}{2}} \\
&\leq \quad \frac{1}{N} |\Lambda_N - \rho_{\theta_j}| \frac{1}{(\rho_{\theta_j} - 2\sigma - 2\delta)^3} \{ \mathrm{Tr}(L U_k^* Y Y^* U_k L) \}^{\frac{1}{2}} \{ \mathrm{Tr}(K U_k^* Y Y^* U_k K) \}^{\frac{1}{2}}.
\end{aligned}$$

According to (4.15), $\frac{1}{\sqrt{N}} \{ \mathrm{Tr}(K U_k^* Y Y^* U_k K) \}^{\frac{1}{2}}$ converges in probability towards $\sqrt{k_{+\sigma}} \sigma$.

$$\begin{aligned}
\frac{1}{N} \mathrm{Tr}(L U_k^* Y Y^* U_k L) \quad &= \quad \frac{1}{N} \sum_{i=k_{+\sigma}+1}^{k} \sum_{p=1}^{N-k} \sum_{l,q=1, l \neq q}^{k} \overline{(U_k)_{l,i}} W_{l,k+p} \, \overline{W_{q,k+p}} \, (U_k)_{q,i} \\
&\quad + \frac{1}{N} \sum_{p=1}^{N-k} \sum_{i=k_{+\sigma}+1}^{k} \sum_{l=1}^{k} |(U_k)_{l,i}|^2 |W_{l,k+p}|^2.
\end{aligned}$$

$\{ \sum_{i=k_{+\sigma}+1}^{k} \sum_{l,q=1,l\neq q}^{k} \overline{(U_k)_{l,i}} W_{l,k+p} \, \overline{W_{q,k+p}} \, (U_k)_{q,i}, 1 \leq p \leq N-k \}$ are i.i.d variables with mean zero such that

$$\mathbb{E}\left(\left| \sum_{i=k_{+\sigma}+1}^{k} \sum_{l,q=1,l\neq q}^{k} \overline{(U_k)_{l,i}} W_{l,k+p} \, \overline{W_{q,k+p}} \, (U_k)_{q,i} \right|^2 \right) \leq C(k - k_{+\sigma})^2.$$

for some constant C. Hence

$$\mathbb{E}\left(\left| \frac{1}{N} \sum_{p=1}^{N-k} \sum_{i=k_{+\sigma}+1}^{k} \sum_{l,q=1,l\neq q}^{k} \overline{(U_k)_{l,i}} W_{l,k+p} \, \overline{W_{q,k+p}} \, (U_k)_{q,i} \right|^2 \right) \leq C \frac{(k - k_{+\sigma})^2}{N},$$

so that the first term in the previous sum converges in L^2 towards 0 and thus in probability. Moreover, since $\sum_{i=k_{+\sigma}+1}^{k} \frac{1}{k-k_{+\sigma}} \sum_{l=1}^{k} |(U_k)_{l,i}|^2 |W_{l,k+p}|^2$ are i.i.d random variables with mean σ^2 and bounded second moments (in N), there exists some constant C such that

$$\mathbb{E}\left(\left| \frac{1}{N} \sum_{p=1}^{N-k} \sum_{i=k_{+\sigma}+1}^{k} \frac{1}{k - k_{+\sigma}} \sum_{l=1}^{k} |(U_k)_{l,i}|^2 |W_{l,k+p}|^2 - \sigma^2 \right|^2 \right) \leq \frac{C}{N},$$

so that $\frac{1}{N}\sum_{p=1}^{N-k}\sum_{i=k_{+\sigma}+1}^{k}\frac{1}{k-k_{+\sigma}}\sum_{l=1}^{k}|(U_k)_{l,i}|^2|W_{l,k+p}|^2$ converges in L^2 towards σ^2 and thus in probability. It follows that $\frac{1}{N}\operatorname{Tr}(LU_k^*YY^*U_kL)=O_{\mathbb{P}}(k)$ and then that under the assumption $k\ll\sqrt{N}$,

$$\|\,[U_k^*\tau_N U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\,\|_{HS}=o_{\mathbb{P}}(\sqrt{k})=o_{\mathbb{P}}(N^{\frac14}).$$

Note that we also have for $k\ll N$,

$$\|\,[U_k^*\tau_N U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\,\|_{HS}=|\xi_N(\lambda)|o_{\mathbb{P}}(1),$$

and therefore

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

Thus, (4.22) is established.
For (4.23), recall that $[U_k^*B_{k,N}U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}=[U_k^*W_kU_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}+\frac{1}{\sqrt{N}}[U_k^*Y\widehat{G}(\rho_{\theta_j})Y^*U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}$. Since

$$\mathbb{E}\left(\|[U_k^*W_kU_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\|^2_{HS}\right)\le k_{+\sigma}k\sigma^2$$

one has that

$$\mathbb{P}\left(\|[U_k^*W_kU_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\|_{HS}>\epsilon N^{\frac14}\right)\le\frac{k_{+\sigma}k\sigma^2}{\epsilon^2\sqrt{N}}.$$

Hence, as $k\ll\sqrt{N}$, we can deduce that $\|[U_k^*W_kU_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\|_{HS}=o_{\mathbb{P}}(N^{\frac14})$.
Now, let us prove the same estimate for the remaining term. Using the same proof as in 4.17, one can get that for $p\neq q$, for some constant $C>0$,

$$\mathbb{E}\left(|\mathcal{U}(p)^*\widehat{G}(\rho_{\theta_j})\mathcal{U}(q)|^2\mathbf{1}_{\Omega_N^{(2)}}\right)\le C\mathbb{E}\operatorname{Tr}(\widehat{G}(\rho_{\theta_j})^2\mathbf{1}_{\Omega_N^{(2)}})$$

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

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

Then using that

$$\mathbb{P}\left(\|\frac{1}{\sqrt{N}}[U_k^*Y\widehat{G}(\rho_{\theta_j})Y^*U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\|_{HS}\mathbf{1}_{\Omega_N^{(2)}}>\epsilon N^{\frac14}\right)\le\frac{1}{\epsilon^2\sqrt{N}}\mathbb{E}[\|\frac{1}{\sqrt{N}}[U_k^*Y\widehat{G}(\rho_{\theta_j})Y^*U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\|^2_{HS}]$$

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

$$\|\frac{1}{\sqrt{N}}[U_k^*Y\widehat{G}(\rho_{\theta_j})Y^*U_k]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}\|_{HS}\mathbf{1}_{\Omega_N^{(2)}}=o_{\mathbb{P}}(N^{\frac14}).$$

Thus (4.23) and Lemma 4.4 are proved. □

Using that

$$\|\Sigma\|\le\frac{1}{\rho_{\theta_j}-2\sigma-2\delta},\tag{4.24}$$

one can readily notice that Lemma 4.4 leads to

$$\frac{1}{\sqrt{N}}T_N\Sigma T_N^*=(1+|\xi_N(\Lambda_N)|)^2o_{\mathbb{P}}(1).\tag{4.25}$$

We now consider the remaining terms in the r.h.s of (4.21). We first show the following result where we recall that $\Delta_{k_{+\sigma}}(\rho_{\theta_j})=[U_k^*Y\hat{G}(\rho_{\theta_j})\tilde{A}_{N-k\times k}]^{\nearrow}_{k_{+\sigma}\times k-k_{+\sigma}}$.

Lemma 4.5. *$\frac{1}{\sqrt{N}}T_N\Sigma\Delta_{k_{+\sigma}}(\rho_{\theta_j})^*$, $\frac{1}{\sqrt{N}}\Delta_{k_{+\sigma}}(\rho_{\theta_j})\Sigma\Delta_{k_{+\sigma}}(\rho_{\theta_j})^*$ and $\frac{1}{\sqrt{N}}\Delta_{k_{+\sigma}}(\rho_{\theta_j})\Sigma T_N^*$ are all equal to some $(1+|\xi_N(\Lambda_N)|)o_{\mathbb{P}}(1)$.*

Proof of Lemma 4.5 : We will show that, on $\Omega_N(\Lambda_N)$, for any $u>0$,

$$\Delta_{k_{+\sigma}}(\rho_{\theta_j}) = o_{\mathbb{P}}(N^u). \tag{4.26}$$

One can readily see that this leads to the announced result combining Lemma 4.4, (4.24) and (4.26). 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 $\mathcal{U}(p) =^t [(Y^*U_k)_{1,p},\ldots,(Y^*U_k)_{N-k,p}]$ has independent centered entries with variance σ^2, one has that

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

Therefore, $\mathbb{P}(\mathbf{1}_{\Omega_N(\Lambda_N)}\|\Delta_{k_{+\sigma}}(\rho_{\theta_j})\|_{HS}\ge \epsilon N^u) \le \epsilon^{-2}N^{-2u}\mathbb{E}(\mathbf{1}_{\Omega_N(\Lambda_N)}\|\Delta_{k_{+\sigma}}(\rho_{\theta_j})\|^2_{HS})$ goes to zero as N tends to infinity. Hence (4.26) holds true on $\Omega_N(\Lambda_N)$ and the proof of Lemma 4.5 is complete. □

Let us now prove that

Lemma 4.6. $\Delta_{k_{+\sigma}}(\Lambda_N) = \Delta_{k_{+\sigma}}(\rho_{\theta_j}) + O_{\mathbb{P}}(|\xi_N(\Lambda_N)|)$.

Proof of Lemma 4.6: We have

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

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

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

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

$$\|\Delta_{k_{+\sigma}}(\Lambda_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\|^2_{HS} \le C'(\xi_N(\Lambda_N))^2\frac{1}{(\rho_{\theta_j}-2\sigma-2\delta)^4}\frac{1}{N}\operatorname{Tr}(\mathcal{U}^*\mathcal{U}).$$

It follows from (4.15) that

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

implying that $\|\Delta_{k_{+\sigma}}(\Lambda_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\|_{HS} = O_{\mathbb{P}}(|\xi_N(\Lambda_N)|)$. □

We are now in position to conclude the proof of Lemma 4.3. Indeed, writing

$$\Delta_{k_{+\sigma}}(\Lambda_N)\Sigma T_N^* = \left(\Delta_{k_{+\sigma}}(\Delta_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\right)\Sigma T_N^* + \Delta_{k_{+\sigma}}(\rho_{\theta_j})\Sigma T_N^*$$

and

$$\begin{aligned}
\Delta_{k_{+\sigma}}(\Lambda_N)\Sigma\Delta_{k_{+\sigma}}(\Lambda_N) &= \Delta_{k_{+\sigma}}(\rho_{\theta_j})\Sigma\Delta_{k_{+\sigma}}(\rho_{\theta_j})^*\\
&\quad+\left(\Delta_{k_{+\sigma}}(\Lambda_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\right)\Sigma\Delta_{k_{+\sigma}}(\rho_{\theta_j})^*\\
&\quad+\left(\Delta_{k_{+\sigma}}(\Lambda_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\right)\Sigma\left(\Delta_{k_{+\sigma}}(\Lambda_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\right)^*\\
&\quad+\Delta_{k_{+\sigma}}(\rho_{\theta_j})\Sigma\left(\Delta_{k_{+\sigma}}(\Lambda_N)-\Delta_{k_{+\sigma}}(\rho_{\theta_j})\right)^*,
\end{aligned}$$

we deduce from Lemmas 4.4, 4.6 and (4.24), (4.26) that $\frac{1}{\sqrt{N}}\Delta_{k_{+\sigma}}(\Lambda_N)\Sigma T_N^*$ and $\frac{1}{\sqrt{N}}\Delta_{k_{+\sigma}}(\Lambda_N)\Sigma\Delta_{k_{+\sigma}}(\Lambda_N)^*$ are both equal to some $(1+|\xi_N(\Lambda_N)|)o_{\mathbb{P}}(1)$. Using also (4.25), we can deduce that

$$\frac{1}{\sqrt{N}}\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)\Sigma\Gamma_{k_{+\sigma}\times k-k_{+\sigma}}(\Lambda_N)^* = (1+|\xi_N(\Lambda_N)|)^2 o_{\mathbb{P}}(1) \tag{4.27}$$

which gives (4.20) and completes the proof of Lemma 4.3. □

Combining all the preceding, we have established Proposition 4.2. 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 4.7. *For all $u>0$,*

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

Proof: Straightforward computations lead to the existence of some constant C such that

$$\mathbb{E}\left(\|\,[U_k^* W_k U_k]_{k_{+\sigma}}\,\|_{HS}\right) \leq C.$$

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 4.1 of the convergence in probability of $[U_k^*\Phi_N U_k]_{k_{+\sigma}}$ towards zero, one can get that

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

and the convergence in probability towards zero of the term inside the above expectation follows by Tchebychev inequality. Since moreover according to Lemma 5.1,

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

we can deduce that

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

The proof of Lemma 4.7 is complete. □

Proposition 4.3. *Let Δ_{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 $\mathbf{X}_{k_{+\sigma},N}(\Lambda_N)+\operatorname{diag}(\Delta_{k_j},0)$ 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*

$$\check{X}_{k_j,N}(\Lambda_N) = V_{k_j,N} + o_{\mathbb{P}}(1) \tag{4.28}$$

where $V_{k_j,N}$ is the $k_j\times k_j$ element in the block decomposition of $V_{k_{+\sigma},N}$ defined by (4.13); namely

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

with $U_{K_j\times k_j}$ and $B_{k,N}$ defined respectively by (2.3) and (4.9).

Proof of Proposition 4.3: Since $\xi_N(\Lambda_N)$ converges in distribution, we can write the matrix $\mathbf{X}_{k_{+\sigma},N}(\Lambda_N)$ given by (4.12) as

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

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,

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

if μ 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 4.7,

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

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 (4.28), 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 4.7. The statement of the proposition then follows from (3.1). □

STEP 3: We now examine the convergence of the $k_j\times k_j$ matrix $V_{k_j,N} = U^*_{K_j\times k_j}[B_{k,N}]^{\nwarrow}_{K_j}U_{K_j\times k_j}$ in the two cases: K_j independent of N and $K_j\to\infty$.

a) K_j and the matrix $U_{K_j\times k_j}$ are independent of N

Proposition 4.4. *The Hermitian (resp. symmetric) matrix $[B_{k,N}]^{\nwarrow}_{K_j}$ converges in distribution towards the law of $W_{K_j}+H_{K_j}$ where W_{K_j} is a Wigner matrix of size K_j with distribution given by μ (cf* **(i)***) and H_{K_j} is a centered Gaussian Hermitian (resp. symmetric) matrix of size K_j independent of W_{K_j}, with independent entries H_{pl}, $p\leq l$ with variance*

$$\begin{cases} v_{pp} = E(H_{pp}^2) = \dfrac{t}{4}\Big(\dfrac{m_4-3\sigma^4}{\theta_j^2}\Big) + \dfrac{t}{2}\dfrac{\sigma^4}{\theta_j^2-\sigma^2},\ p=1,\ldots,K_j,\\ v_{pl} = \mathbb{E}(|H_{pl}|^2) = \dfrac{\sigma^4}{\theta_j^2-\sigma^2},\ 1\leq p<l\leq K_j.\end{cases} \tag{4.29}$$

The proof follows from Theorem 3.2 and is omitted. We shall detail the proof of a similar result in the infinite case (cf. below the proof of Lemma 4.9).

b) $K_j(=K_j(N))\to\infty$ and $U_{K_j\times k_j}(=U_{K_j\times k_j}(N))$ satisfies (2.2)

Proposition 4.5. *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 $U^*_{K_j\times k_j}[B_{k,N}]^{\nwarrow}_{K_j}U_{K_j\times k_j}$ converges in distribution to a GU(O)E($k_j\times k_j$, $\frac{\theta_j^2\sigma^2}{\theta_j^2-\sigma^2}$).*

We decompose the proof into the two following lemmas.

Lemma 4.8. *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 $U^*_{K_j\times k_j}[W_k]^{\nwarrow}_{K_j}U_{K_j\times k_j}$ converges in distribution to a GU(O)E($k_j\times k_j,\sigma^2$).*

Proof of Lemma 4.8: 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

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

We have

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

where

$$D_i = 2\Re e\Big(\sum_{1\le p\le q\le k_j} \alpha_{pq}(U_k)_{iq}\overline{(U_k)}_{ip}\Big),$$

$$R_{il} = \sqrt{2}\Re e\Big(\sum_{1\le p\le q\le k_j} \alpha_{pq}((U_k)_{lq}\overline{(U_k)}_{ip} + (U_k)_{iq}\overline{(U_k)}_{lp})\Big),$$

$$I_{il} = \sqrt{2}\Im m\Big(\sum_{1\le p\le q\le k_j} \overline{\alpha_{pq}}(\overline{(U_k)}_{lq}(U_k)_{ip} - \overline{(U_k)}_{iq}(U_k)_{lp})\Big).$$

Hence $L_N(\alpha) = \sum_{m=1}^{K_j^2} \beta_{m,N}\phi_m$ where ϕ_m are i.i.d random variables with distribution μ and $\beta_{m,N}$ are real constants (depending on the α_{pq}) which satisfy $\max_{m=1}^{K_j^2}|\beta_{m,N}| \to 0$ 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 μ (all are finite since μ 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

$$\frac{C_2^{(N)}}{\sigma^2} = \sum_{i=1}^{K_j} D_i^2 + \sum_{1\le i<j\le K_j} R_{ij}^2 + \sum_{1\le i<j\le K_j} I_{ij}^2 = 2\sum_{1\le p<q\le k_j}|\alpha_{pq}|^2 + 4\sum_{1\le p\le k_j}|\alpha_{pp}|^2. \tag{4.30}$$

Let us rewrite the l.h.s as

$$\sum_{i=1}^{K_j} D_i^2 + \sum_{1\le i<l\le K_j} R_{il}^2 + \sum_{1\le i<l\le K_j} I_{il}^2 = \sum_{\substack{1\le p\le q\le k_j\\ 1\le p'\le q'\le k_j}} \Pi_{p,q,p',q'}$$

where

$$\begin{aligned}
\frac{\Pi_{p,q,p',q'}}{2} &= 2\sum_{i=1}^{K_j}\Re e\left(\alpha_{pq}(U_k)_{iq}\overline{(U_k)}_{ip}\right)\Re e\left(\alpha_{p'q'}(U_k)_{iq'}\overline{(U_k)}_{ip'}\right)\\
&+\sum_{1\le i<l\le K_j}\Re e\left(\alpha_{pq}((U_k)_{lq}\overline{(U_k)}_{ip}+(U_k)_{iq}\overline{(U_k)}_{lp})\right)\Re e\left(\alpha_{p'q'}((U_k)_{lq'}\overline{(U_k)}_{ip'}+(U_k)_{iq'}\overline{(U_k)}_{lp'})\right)\\
&+\sum_{1\le i<l\le K_j}\Im m\left(\alpha_{pq}((U_k)_{lq}\overline{(U_k)_{ip}}-(U_k)_{iq}\overline{(U_k)_{lp}})\right)\Im m\left(\alpha_{p'q'}((U_k)_{lq'}\overline{(U_k)_{ip'}}-(U_k)_{iq'}\overline{(U_k)_{lp'}})\right).
\end{aligned}$$

So that

$$\begin{aligned}
\Pi_{p,q,p',q'} &= \sum_{1\le i,l\le K_j}\Big\{\Re e\left(\alpha_{pq}((U_k)_{lq}\overline{(U_k)}_{ip}+(U_k)_{iq}\overline{(U_k)}_{lp})\right)\Re e\left(\alpha_{p'q'}((U_k)_{lq'}\overline{(U_k)}_{ip'}+(U_k)_{iq'}\overline{(U_k)}_{lp'})\right)\\
&\quad+\Im m\left(\alpha_{pq}((U_k)_{lq}\overline{(U_k)_{ip}}-(U_k)_{iq}\overline{(U_k)_{lp}})\right)\Im m\left(\alpha_{p'q'}((U_k)_{lq'}\overline{(U_k)_{ip'}}-(U_k)_{iq'}\overline{(U_k)_{lp'}})\right)\Big\}\\
&= 2\Pi^{(1)}_{p,q,p',q'} + 2\Pi^{(2)}_{p,q,p',q'}
\end{aligned}$$

where

$$\begin{aligned}
\Pi^{(1)}_{p,q,p',q'} &= \sum_{1\le i,l\le K_j}\Big\{\Re e\left(\alpha_{pq}(U_k)_{lq}\overline{(U_k)}_{ip}\right)\Re e\left(\alpha_{p'q'}(U_k)_{lq'}\overline{(U_k)}_{ip'}\right)\\
&\quad+\Im m\left(\alpha_{pq}(U_k)_{lq}\overline{(U_k)_{ip}}\right)\Im m\left(\alpha_{p'q'}(U_k)_{lq'}\overline{(U_k)_{ip'}}\right)\Big\}\\
&= \Re e\Big\{\alpha_{pq}\overline{\alpha_{p'q'}}\sum_{1\le i,l\le K_j}(U_k)_{ip'}\overline{(U_k)}_{ip}(U_k)_{lq}\overline{(U_k)}_{lq'}\Big\}\\
&= |\alpha_{pq}|^2\delta_{(p,q),(p',q')}
\end{aligned}$$

and

$$\begin{aligned}
\Pi^{(2)}_{p,q,p',q'} &= \sum_{1\le i,l\le K_j} \Big\{\Re e\left(\alpha_{pq}(U_k)_{lq}\overline{(U_k)}_{ip}\right)\Re e\left(\alpha_{p'q'}(U_k)_{iq'}\overline{(U_k)}_{lp'}\right)\\
&\quad -\Im m\left(\alpha_{pq}(U_k)_{lq}\overline{(U_k)}_{ip}\right)\Im m\left(\alpha_{p'q'}(U_k)_{iq'}\overline{(U_k)}_{lp'}\right)\Big\}\\
&= \Re e\Big\{\alpha_{pq}\alpha_{p'q'}\sum_{1\le i,l\le K_j}(U_k)_{iq'}\overline{(U_k)}_{ip}(U_k)_{lq}\overline{(U_k)}_{lp'}\Big\}\\
&= |\alpha_{pp}|^2\delta_{(p,q),(p',q')}\delta_{p,q}
\end{aligned}$$

Then (4.30) readily follows. In the following, we let $const = \sum_{m=1}^{{K_j}^2}\beta^2_{m,N}$.
Since $|C_n^{(N)}| \le const\max_{m=1}^{K_j^2}|\beta_{m,N}|^{n-2}|C_n(\mu)|$, $C_n^{(N)}$ converges to zero for each $n\ge 3$. Thus we can deduce from Janson's theorem [J] that $L_N(\alpha)$ converges to a centered gaussian distribution with variance $\sigma^2(2\sum_{1\le p<q\le k_j}|\alpha_{pq}|^2 + 4\sum_{1\le p\le k_j}|\alpha_{pp}|^2)$ and the proof of Lemma 4.8 is complete in the complex case.

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

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

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

Remark 4.1. *Note that Lemma 4.8 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 4.9. *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 $\frac{1}{\sqrt{N}}U^*_{K_j\times k_j}\left[\left(Y\widehat{G}(\rho_{\theta_j})Y^* - (N-k)\frac{\sigma^2}{\theta_j}I_k\right)\right]^{\nwarrow}_{K_j}U_{K_j\times k_j}$ converges towards a GU(O)E($k_j\times k_j, \frac{\sigma^4}{\theta_j^2-\sigma^2}$).*

Proof of Lemma 4.9: We shall apply a slightly modified version of Theorem 3.2 (see Theorem 7.1 in [B-Ya2]) 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 $1\le p\le q\le k_j$, taking the lexicographic order. We define a sequence of i.i.d centered vectors $(x_i,y_i)_{i\le N-k}$ 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 4.2. The matrix A of size $N-k$ is the matrix $\widehat{G}(\rho_{\theta_j})$ and is independent of $\mathcal{U}$. Note that we are not exactly in the context of Theorem 7.1 of [B-Ya2] 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 (3.2)).
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 (3.2)).
   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 (3.2))

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

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 $\epsilon > 0$ and for $z = x$ or y, for any l,

$$\mathbb{E}[|z_{l1}|^4 \mathbf{1}_{(|z_{l1}|\geq \epsilon N^{1/4})}] \longrightarrow 0 \quad \text{as } N \to \infty. \tag{4.31}$$

We postpone the proof of (4.31) to the end of the proof. Assuming that (4.31) holds true, we obtain the CLT theorem 7.1 ([B-Ya2]): the Hermitian matrix $Z_N = (Z_N(p,q))$ of size k_j defined by

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

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

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

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

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

and the coefficients ω, θ, τ are defined in Theorem 3.2. Here $A = \widehat{G}(\rho_{\theta_j})$ so that $\omega = 1/\theta_j^2$ and $\theta = 1/(\theta_j^2 - \sigma^2)$ (see the Appendix).
¿From Lemma 4.2,

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

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,

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

and in the complex case,

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

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 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 $\frac{1}{2}\lim_{N\to\infty} \tilde{B}_{12}$ (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 $\tilde{B}_{12}(3) = 0$,

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

Thus, $var(\Re e(G_{pq})) = \theta\sigma^4/2 = \sigma^4/(2(\theta_j^2 - \sigma^2))$. We thus obtain Lemma 4.9 by using that $\mathrm{Tr}(\widehat{G}(\rho_{\theta_j})) = (N-k)\,\mathrm{tr}_{N-k}(\widehat{G}(\rho_{\theta_j}))$ and $\mathrm{tr}_{N-k}(\widehat{G}(\rho_{\theta_j})) \to 1/\theta_j$.
It remains to prove (4.31). The variable $\alpha_N := |z_{l1}|^4 \mathbf{1}_{(|z_{l1}| \geq \epsilon N^{1/4})}$ tends to 0 in probability. It is thus enough to prove uniform integrability of the sequence α_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 (4.31) and finishes the proof of Lemma 4.9. □

STEP 4: We are now in position to prove that

$$\Big(\xi_N(\lambda_{\hat{k}_{j-1}+1}(\mathbf{M}_N)), \ldots, \xi_N(\lambda_{\hat{k}_{j-1}+k_j}(\mathbf{M}_N))\Big) \xrightarrow{\mathcal{L}} \big(\lambda_1(V_{k_j\times k_j}), \ldots, \lambda_{k_j}(V_{k_j\times k_j})\big). \tag{4.32}$$

To prove (4.32), 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 ρ_{θ_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 4.3. This fit choice of $\Lambda_N^{(i)}$ gives that

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

Hence, $\xi_N(\Lambda_N^{(i)})$ is an eigenvalue of $\tilde{X}_{k_j}^{(i)} - \epsilon_i I_{k_j}$.
According to Propositions 4.1 and 4.3, on an event $\check{\Omega}_N$ whose probability goes to one as N goes to infinity, there exists some l_i such that

$$\Lambda_N^{(i)} = \lambda_{l_i}\Big(\mathbf{M}_N - \frac{\epsilon_i}{\sqrt{N}}\mathrm{diag}(I_{k_j}, 0_{N-k_j})\Big).$$

The following lines hold on $\check{\Omega}_N$. By using Weyl's inequalities (Lemma 3.2), one has for all $i \in \{1, \ldots, k_j\}$ that

$$\Big|\lambda_{l_i}\Big(\mathbf{M}_N - \frac{\epsilon_i}{\sqrt{N}}\mathrm{diag}(I_{k_j}, 0_{N-k_j})\Big) - \lambda_{l_i}(\mathbf{M}_N)\Big| \leq \frac{|\epsilon_i|}{\sqrt{N}}.$$

We then deduce that

$$\Big(\xi_N(\lambda_{l_1}(\mathbf{M}_N)), \ldots, \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))\Big) = \Big(\lambda_1(V_{k_j,N}), \ldots, \lambda_{k_j}(V_{k_j,N})\Big) + o_{\mathbb{P}}(1) \tag{4.33}$$

and thus

$$\Big(\xi_N(\lambda_{l_1}(\mathbf{M}_N)), \ldots, \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))\Big) \xrightarrow{\mathcal{L}} \Big(\lambda_1(V_{k_j\times k_j}), \ldots, \lambda_{k_j}(V_{k_j\times k_j})\Big). \tag{4.34}$$

Now, to get (4.32), it is sufficient to prove that

$$\mathbb{P}\Big(l_i = \hat{k}_{j-1} + i;\, i = 1, \ldots, k_j\Big) \to 1, \quad \text{as } N \to \infty. \tag{4.35}$$

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

$$\Big(\xi_N(\lambda_{\hat{k}_{j-1}+1}(\mathbf{M}_N)), \ldots, \xi_N(\lambda_{\hat{k}_{j-1}+k_j}(\mathbf{M}_N))\Big) = \Big(\xi_N(\lambda_{l_1}(\mathbf{M}_N)), \ldots, \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N))\Big). \tag{4.36}$$

Hence, if (4.35) is satisfied then (4.36) combined with (4.34) imply (4.32).
We turn now to the proof of (4.35). 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 finite (Case a)) and independent of N, we call on the following arguments. One can decompose the matrix $U^*_{K_j \times k_j} H_{K_j} U_{K_j \times k_j}$ appearing in the definition (2.7) of $V_{k_{+\sigma}}$ in the following way

$$U^*_{K_j \times k_j} H_{K_j} U_{K_j \times k_j} = Q_{k_j} + \check{H}_{k_j}$$

with $\check{H}_{k_j}$ distributed as GU(O)E (using the fact that $U^*_{K_j \times k_j} U_{K_j \times k_j} = I_{k_j}$) 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 τ and the second one being a matrix Z_{k_j} of the form $U^*_{K_j \times k_j} W_{K_j} U_{K_j \times k_j} + Q_{k_j}$. 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:

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

where $I((\lambda_1, \ldots, \lambda_N), Z_{k_j}) = \int \exp\left(\frac{2}{\tau t} N \operatorname{Tr}(U \operatorname{diag}(\lambda_1, \ldots, \lambda_N) U^* Z_{k_j})\right) m(dU)$ 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 (4.34) then implies that the event

$$\check{\Omega}'_N := \left\{ \xi_N(\lambda_{l_1}(\mathbf{M}_N)) > \cdots > \xi_N(\lambda_{l_{k_j}}(\mathbf{M}_N)) \right\} \bigcap \check{\Omega}_N$$

is such that $\lim_N \mathbb{P}(\check{\Omega}'_N) = 1$. By Theorem 1.1, we notice that the event

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

also satisfies $\lim_N \mathbb{P}(\tilde{\Omega}'_N) = 1$, for δ small enough. This leads to (4.35) since $\tilde{\Omega}'_N \subset \{l_i = i + \hat{k}_{j-1}, i = 1, \ldots, k_j\}$.
The proof of Theorems 2.1 and 2.2 is complete.

5 Appendix

We recall the CLT for the empirical distribution of a Wigner matrix.

Theorem 5.1. *(Theorem 1.1 in [B-Ya1]) Let f be an analytic function on an open set of the complex plane including $[-2\sigma, 2\sigma]$. If the entries $((W_N)_{il})_{1 \leq i \leq l \leq N}$ of a general Wigner matrix $\mathbf{W}_N$ of variance σ^2 satisfy the conditions*

(i) for $i \neq l$, $\mathbb{E}(|(W_N)_{il}|^4) = const$,

(ii) for any $\eta > 0$, $\lim_{N \to +\infty} \frac{1}{\eta^4 n^2} \sum_{i,l} \mathbb{E}\left[|(W_N)_{il}|^4 1\!\!1_{\{|(W_N)_{il}| \geq \eta\sqrt{N}\}} \right] = 0$,

then $N\Big(\operatorname{tr}_N(f(\frac{1}{\sqrt{N}} \mathbf{W}_N)) - \int f d\mu_{sc} \Big)$ converges in distribution towards a Gaussian variable, where μ_{sc} is the semicircle distribution of variance σ^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 $N - k \to \infty$.

Lemma 5.1. *Each of the following convergence holds in probability as $N \to \infty$:*

i) $\sqrt{N} \left(\operatorname{tr}_{N-k} \hat{G}(\rho_{\theta_j}) - 1/\theta_j \right) \longrightarrow 0$, *ii)* $\operatorname{tr}_{N-k} \hat{G}^2(\rho_{\theta_j}) \longrightarrow \int \frac{1}{(\rho_{\theta_j}-x)^2} d\mu_{sc}(x) = 1/(\theta_j^2 - \sigma^2)$,

iii) $\frac{1}{N-k} \sum_{i=1}^{N-k} (\hat{G}(\rho_{\theta_j})_{ii})^2 \longrightarrow \left(\int \frac{d\mu_{sc}(x)}{\rho_{\theta_j}-x}\right)^2 = 1/\theta_j^2$.

Proof of Lemma 5.1: We denote by G the resolvent of the non-Deformed Wigner matrix $W_{N-k}/\sqrt{N}$.
i) By Theorem 5.1, one knows that $\sqrt{N}\left(\operatorname{tr}_{N-k} G(\rho_{\theta_j}) - \int \frac{d\mu_{sc}(x)}{\rho_{\theta_j}-x}\right)$ converges in probability towards 0. Now, we have $\int \frac{d\mu_{sc}(x)}{\rho_{\theta_j}-x} = \frac{1}{\theta_j}$ (see [H-P] p. 94). It is thus enough to show that

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

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

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

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 Lemma 3.1) so that we deduce that

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

Using Theorem 5.1 and the fact that $\mathbb{P}(\tilde{\Omega}_N) \to 1$, we obtain the announced result.

ii) It is sufficient to show that $\operatorname{tr}_{N-k} \hat{G}^2(\rho_{\theta_j}) - \operatorname{tr}_{N-k} G^2(\rho_{\theta_j}) \to 0$ in probability since, by Theorem 5.1, one knows that $\operatorname{tr}_{N-k} G^2(\rho_\theta)$ converges in probability towards $\int \frac{1}{(\rho_{\theta_j}-x)^2} d\mu_{sc}(x)$.
Using the fact that $\operatorname{Tr}(BC) = \operatorname{Tr}(CB)$, it is not hard to see that

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

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 [C-D-F]. Indeed, it was shown in Section 5.2 of [C-D-F] 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 [C-D-F] 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 [C-D-F]. □

References

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

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

[B-S2] Bai Z.D. and Silverstein J.W., Spectral Analysis of Large Dimensional Random Matrices, Science Press, Beijing, 2006.

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

[B-Ya2] Bai Z.D. and Yao J.F., *Central limit theorems for eigenvalues in a spiked population model*, Ann. Inst. H. Poincaré **44** (3), 447-474 (2008).

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

[B-B-P] Biroli G., Bouchaud J.P. and Potters M., *On the top eigenvalue of heavy-tailed random matrices*, Europhys. Lett. EPL **78** (1), Art 10001, 5 pp (2007).

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

[Fe-Pe] Féral D. and Péché S., *The largest eigenvalue of rank one deformation of large Wigner matrices*, Comm. Math. Phys. **272**, 185–228 (2007).

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

[H-P] Hiai F. and Petz D., The semicircle law, free random variables and entropy, Mathematical Surveys and Monographs, Volume 77, A.M.S, 2000.

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

[J] Janson S., *Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs*, Ann. Probab. **16**, 305–312 (1988).

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

[P] Paul D., *Asymptotics of sample eigenstructure for a large dimensional spiked covariance model*, Statist. Sinica **17**, 1617–1641 (2007).