

The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation

Francois Chapon, Romain Couillet, Walid Hachem, Xavier Mestre

▶ **To cite this version:**

Francois Chapon, Romain Couillet, Walid Hachem, Xavier Mestre. The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation. 2012. hal-00713811v2

HAL Id: hal-00713811
https://hal.science/hal-00713811v2

Preprint submitted on 22 Jan 2013

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

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

THE OUTLIERS AMONG THE SINGULAR VALUES OF LARGE RECTANGULAR RANDOM MATRICES WITH ADDITIVE FIXED RANK DEFORMATION

FRANÇOIS CHAPON, ROMAIN COUILLET,
WALID HACHEM AND XAVIER MESTRE

ABSTRACT. Consider the matrix $\Sigma_n = n^{-1/2} X_n D_n^{1/2} + P_n$ where the matrix $X_n \in \mathbb{C}^{N\times n}$ has Gaussian standard independent elements, D_n is a deterministic diagonal nonnegative matrix, and P_n is a deterministic matrix with fixed rank. Under some known conditions, the spectral measures of $\Sigma_n\Sigma_n^*$ and $n^{-1}X_nD_nX_n^*$ both converge towards a compactly supported probability measure μ as $N, n \to \infty$ with $N/n \to c > 0$. In this paper, it is proved that finitely many eigenvalues of $\Sigma_n\Sigma_n^*$ may stay away from the support of μ in the large dimensional regime. The existence and locations of these outliers in any connected component of $\mathbb{R} - \text{supp}(\mu)$ are studied. The fluctuations of the largest outliers of $\Sigma_n\Sigma_n^*$ are also analyzed. The results find applications in the fields of signal processing and radio communications.

1. INTRODUCTION

1.1. **The model and the literature.** Consider a sequence of $N \times n$ matrices Y_n, $n = 1, 2, \ldots$, of the form $Y_n = X_n D_n^{1/2}$ where X_n is a $N \times n$ random matrix whose coefficients X_{ij} are independent and identically distributed (iid) complex Gaussian random variables such that $\Re(X_{11})$ and $\Im(X_{11})$ are independent, each with mean zero and variance $1/2$, and where D_n is a deterministic nonnegative diagonal $n \times n$ matrix. Writing $D_n = \text{diag}(d_j^n)_{j=1,\ldots,n}$ and denoting by $\boldsymbol{\delta}$ the Dirac measure, it is assumed that the spectral measure $\nu_n = n^{-1}\sum_{j=1}^n \boldsymbol{\delta}_{d_j^n}$ of D_n converges weakly to a compactly supported probability measure ν when $n \to \infty$. It is also assumed that the maximum of the distances from the diagonal elements of D_n to the support $\text{supp}(\nu)$ of ν goes to zero as $n \to \infty$. Assume that $N/n \to c$ when $n \to \infty$, where c is a positive constant. Then it is known that with probability one, the spectral measure of the Gram matrix $n^{-1}Y_nY_n^*$ converges weakly to a compactly supported probability measure μ (see [26], [16], [35], [36]) and, with probability one, $n^{-1}Y_nY_n^*$ has no eigenvalues in any compact interval outside $\text{supp}(\mu)$ for large n [3].
Let r be a given positive integer and consider a sequence of deterministic $N \times n$ matrices P_n, $n = 1, 2, \ldots$, such that $\text{rank}(P_n) = r$ and $\sup_n \|P_n\| < \infty$ where $\|\cdot\|$ is the spectral norm. Consider the matrix $\Sigma_n = n^{-1/2}Y_n + P_n$. Since the additive deformation P_n has a fixed rank, the spectral measure of $\Sigma_n\Sigma_n^*$ still converges to μ

Date: January 2012.
2000 *Mathematics Subject Classification.* Primary 15A52, Secondary 15A18, 60F15.
Key words and phrases. Random Matrix Theory, Stieltjes Transform, Fixed rank deformation, Extreme eigenvalues, Gaussian fluctuations.
The work of the first three authors was supported by the French Ile-de-France region, DIM LSC fund, Digiteo project DESIR.

(see, *e.g.*, [2, Lemma 2.2]). However, a finite number of eigenvalues of $\Sigma_n\Sigma_n^*$ (often called "outliers" in similar contexts) may stay away of the support of μ. In this paper, minimal conditions ensuring the existence and the convergence of these outliers towards constant values outside $\mathrm{supp}(\mu)$ are provided, and these limit values are characterized. The fluctuations of the outliers lying at the right of $\mathrm{supp}(\mu)$ are also studied.

The behavior of the outliers in the spectrum of large random matrices has aroused an important research effort. In the statistics literature, one of the first contributions to deal with this subject was [23]. It raised the question of the behavior of the extreme eigenvalues of a sample covariance matrix when the population covariance matrix has all but finitely many of its eigenvalues equal to one (leading to a mutliplicative fixed rank deformation). This problem has been studied thoroughly in [5, 6, 32]. Other contributions (see [11]) study the outliers of a Wigner matrix subject to an additive fixed rank deformation. The asymptotic fluctuations of the outliers have been addressed in [5, 33, 32, 1, 11, 12, 7].

Recently, Benaych-Georges and Nadakuditi proposed in [8, 9] a generic method for characterizing the behavior of the outliers for a large palette of random matrix models. For our model, this method shows that the limiting locations as well as the fluctuations of the outliers are intimately related to the asymptotic behavior of certain bilinear forms involving the resolvents $(n^{-1}Y_nY_n^* - xI_N)^{-1}$ and $(n^{-1}Y_n^*Y_n - xI_n)^{-1}$ of the undeformed matrix for real values of x. When $D_n = I_n$, the asymptotic behavior of these bilinear forms can be simply identified (see [9]) thanks to the fact that the probability law of Y_n is invariant by left or right multiplication by deterministic unitary matrices. For general D_n, other tools need to be used. In this paper, these bilinear forms are studied with the help of an integration by parts formula for functionals of Gaussian vectors and the Poincaré-Nash inequality. These tools belong to the arsenal of random matrix theory, as shown in the recent monograph [31] and in the references therein. In order to be able to use them in our context, we make use of a regularizing function ensuring that the moments of the bilinear forms exist for certain $x \in \mathbb{R}_+ = [0, \infty)$.

The study of the spectrum outliers of large random matrices has a wide range of applications. These include communication theory [20], fault diagnosis in complex systems [14], financial portfolio management [34], or chemometrics [29]. The matrix model considered in this paper is widely used in the fields of multidimensional signal processing and radio communications. Using the invariance of the probability law of X_n by multiplication by a constant unitary matrix, D_n can be straightforwardly replaced with a nonnegative Hermitian matrix R_n. In the model $X_nR_n^{1/2} + P_n$ where $R_n^{1/2}$ is any square root of R_n, matrix P_n often represents n snapshots of a discrete time radio signal sent by r sources and received by an array of N antennas, while $X_nR_n^{1/2}$ is a temporally correlated and spatially independent "noise" (spatially correlated and temporally independent noises can be considered as well). In this framework, the results of this paper can be used for detecting the signal sources, estimating their powers, or determining their directions. These subjects are explored in the applicative paper [40].

The remainder of the article is organized as follows. The assumptions and the main results are provided in Section 2. The general approach as well as the basic mathematical tools needed for the proofs are provided in Secion 3. These proofs are given in Sections 4 and 5, which concern respectively the first order (convergence) and the second order (fluctuations) behavior of the outliers.

2. Problem description and main results

Given a sequence of integers $N = N(n)$, $n = 1, 2, \ldots$, we consider the sequence of $N \times n$ matrices $\Sigma_n = n^{-1/2} Y_n + P_n = n^{-1/2} X_n D_n^{1/2} + P_n$ with the following assumptions:

Assumption 1. *The ratio $c_n = N(n)/n$ converges to a positive constant c as $n \to \infty$.*

Assumption 2. *The matrix $X_n = [X_{ij}]_{i,j=1}^{N,n}$ is a $N \times n$ random matrix whose coefficients X_{ij} are iid complex random variables such that $\Re(X_{11})$ and $\Im(X_{11})$ are independent, each with probability distribution $\mathcal{N}(0, 1/2)$.*

Assumption 3. *The sequence of $n \times n$ deterministic diagonal nonnegative matrices $D_n = \operatorname{diag}(d_j^n)_{j=1}^n$ satisfies the following:*

1. *The probability measure $\nu_n = n^{-1} \sum_{j=1}^n \boldsymbol{\delta}_{d_j^n}$ converges weakly to a probability measure ν with compact support.*
2. *The distances $\boldsymbol{d}(d_j^n, \operatorname{supp}(\nu))$ from d_j^n to $\operatorname{supp}(\nu)$ satisfy*

$$\max_{j \in \{1,\ldots,n\}} \boldsymbol{d}\left(d_j^n, \operatorname{supp}(\nu)\right) \xrightarrow[n\to\infty]{} 0.$$

The asymptotic behavior of the spectral measure of $n^{-1} Y_n Y_n^*$ under these assumptions has been thoroughly studied in the literature. Before pursuing, we recall the main results which describe this behavior. These results are built around the Stieltjes Transform, defined, for a positive finite measure μ over the Borel sets of $\mathbb{R}$, as

$$m(z) = \int_{\mathbb{R}} \frac{1}{t - z} \mu(dt) \tag{1}$$

analytic on $\mathbb{C} - \operatorname{supp}(\mu)$. It is straightforward to check that $\Im m(z) \geq 0$ when $z \in \mathbb{C}_+ = \{z : \Im(z) > 0\}$, and $\sup_{y>0} |y m(\imath y)| < \infty$. Conversely, any analytic function $m(z)$ on $\mathbb{C}_+$ that has these two properties admits the integral representation (1) where μ is a positive finite measure. Furthermore, for any continuous real function φ with compact support in $\mathbb{R}$,

$$\int \varphi(t)\, \mu(dt) = \frac{1}{\pi} \lim_{y \downarrow 0} \int \varphi(x) \Im m(x + \imath y)\, dx \tag{2}$$

which implies that the measure μ is uniquely defined by its Stieltjes Transform. Finally, if $\Im(z m(z)) \geq 0$ when $z \in \mathbb{C}_+$, then $\mu((-\infty, 0)) = 0$ [25].

These facts can be generalized to Hermitian matrix-valued nonnegative finite measures [10, 15]. Let $m(z)$ be a $\mathbb{C}^{r \times r}$-valued analytic function on $z \in \mathbb{C}_+$. Letting $\Im X = (X - X^*)/(2\imath)$, assume that $\Im m(z) \geq 0$ and $\Im(z m(z)) \geq 0$ in the order of the Hermitian matrices for any $z \in \mathbb{C}_+$, and that $\sup_{y>0} \|y m(\imath y)\| < \infty$. Then $m(z)$ admits the representation (1) where μ is now a $r \times r$ matrix-valued nonnegative finite measure such that $\mu((-\infty, 0)) = 0$. One can also check that

$\mu([0,\infty)) = -\lim_{y\to\infty} \imath y\, m(-\imath y)$.

The first part of the following theorem has been shown in [26, 36], and the second part in [3]:

Theorem 2.1. *Under Assumptions 1, 2 and 3, the following hold true:*

(1) *For any $z \in \mathbb{C}_+$, the equation*

$$\mathbf{m} = \left(-z + \int \frac{t}{1 + c\mathbf{m}t} \nu(dt) \right)^{-1} \tag{3}$$

admits a unique solution $\mathbf{m} \in \mathbb{C}_+$. The function $\mathbf{m} = \mathbf{m}(z)$ so defined on $\mathbb{C}_+$ is the Stieltjes Transform of a probability measure μ whose support is a compact set of $\mathbb{R}_+$.
Let $(\lambda_i^n)_{i=1,\ldots,N}$ be the eigenvalues of $n^{-1}Y_nY_n^*$, and let $\theta_n = N^{-1}\sum_{i=1}^N \boldsymbol{\delta}_{\lambda_i^n}$ be the spectral measure of this matrix. Then for every bounded and continuous real function f,

$$\int f(t)\theta_n(dt) \xrightarrow[n\to\infty]{a.s.} \int f(t)\mu(dt). \tag{4}$$

(2) *For any interval $[x_1, x_2] \subset \mathbb{R} - \operatorname{supp}(\mu)$,*

$$\sharp\{i \,:\, \lambda_i^n \in [x_1, x_2]\} = 0 \text{ with probability } 1 \text{ for all large } n.$$

We now consider the additive deformation P_n:

Assumption 4. *The deterministic $N \times n$ matrices P_n have a fixed rank equal to r. Moreover, $\mathbf{p}_{\max} = \sup_n \|P_n\| < \infty$.*

In order for some of the eigenvalues of $\Sigma_n\Sigma_n^*$ to converge to values outside $\operatorname{supp}(\mu)$, an extra assumption involving in some sense the interaction between P_n and D_n is needed. Let $P_n = U_n^{\mathrm{GS}}(R_n^{\mathrm{GS}})^*$ be the Gram-Schmidt factorization of P_n where U_n^{GS} is an isometry $N \times r$ matrix and where $(R_n^{\mathrm{GS}})^*$ is an upper triangular matrix in row echelon form whose first nonzero coefficient of each row is positive. The factorization so defined is then unique. Define the $r \times r$ Hermitian nonnegative matrix-valued measure Λ_n^{GS} as

$$\Lambda_n^{\mathrm{GS}}(dt) = (R_n^{\mathrm{GS}})^* \begin{bmatrix} \boldsymbol{\delta}_{d_1^n}(dt) & & \\ & \ddots & \\ & & \boldsymbol{\delta}_{d_n^n}(dt) \end{bmatrix} R_n^{\mathrm{GS}}.$$

Assumption 3 shows that $\mathbf{d}_{\max} = \sup_n \|D_n\| < \infty$. Moreover, it is clear that the support of Λ_n^{GS} is included in $[0, \mathbf{d}_{\max}]$ and that $\Lambda_n^{\mathrm{GS}}([0, \mathbf{d}_{\max}]) \leq \mathbf{p}_{\max}^2 I_r$. Since the sequence $\Lambda_n^{\mathrm{GS}}([0, \mathbf{d}_{\max}])$ is bounded in norm, for every sequence of integers increasing to infinity, there exists a subsequence n_k and a nonnegative finite measure Λ_* such that $\int f(t)\Lambda_{n_k}^{\mathrm{GS}}(dt) \to \int f(t)\Lambda_*(dt)$ for every function $f \in \mathcal{C}([0, \mathbf{d}_{\max}])$, with $\mathcal{C}([0, \mathbf{d}_{\max}])$ being the set of continuous functions on $[0, \mathbf{d}_{\max}]$. This fact is a straightforward extension of its analogue for scalar measures.

Assumption 5. *Any two accumulation points Λ_1 and Λ_2 of the sequences Λ_n^{GS} satisfy $\Lambda_1(dx) = W\Lambda_2(dx)W^*$ where W is a $r \times r$ unitary matrix.*

This assumption on the interaction between P_n and D_n appears to be the least restrictive assumption ensuring the convergence of the outliers to fixed values outside $\operatorname{supp}(\mu)$ as $n \to \infty$. If we consider some other factorization $P_n = U_nR_n^*$ of P_n

where U_n is an isometry matrix with size $N \times r$, and if we associate to the R_n the sequence of $r \times r$ Hermitian nonnegative matrix-valued measures Λ_n defined as

$$\Lambda_n(dt) = R_n^* \begin{bmatrix} \boldsymbol{\delta}_{d_1^n}(dt) & & \\ & \ddots & \\ & & \boldsymbol{\delta}_{d_n^n}(dt) \end{bmatrix} R_n \tag{5}$$

then it is clear that $\Lambda_n(dt) = W_n \Lambda_n^{\mathrm{GS}}(dt) W_n^*$ for some $r \times r$ unitary matrix W_n. By the compactness of the unitary group, Assumption 5 is satisfied for Λ_n^{GS} if and only if it is satisfied for Λ_n. The main consequence of this assumption is that for any function $f \in \mathcal{C}([0, \mathbf{d}_{\max}])$, the eigenvalues of the matrix $\int f(t)\Lambda_n(dt)$ arranged in some given order will converge.

An example taken from the fields of signal processing and wireless communications might help to have a better understanding the applicability of Assumption 5. In these fields, the matrix P_n often represents a multidimensional radio signal received by an array of N antennas. Frequently this matrix can be factored as $P_n = n^{-1/2} U_n A_n S_n^*$ where U_n is a deterministic $N \times r$ isometry matrix, A_n is a deterministic $r \times r$ matrix such that $A_n A_n^*$ converges to a matrix M as $n \to \infty$ (one often assumes $A_n A_n^* = M$ for each n), and $S_n = [S_{ij}]_{i,j=1}^{n,r}$ is a $n \times r$ random matrix independent of X_n with iid elements satisfying $\mathbb{E}S_{1,1} = 0$ and $\mathbb{E}|S_{1,1}|^2 = 1$ (in the wireless communications terminology, $U_n A_n$ is the so called MIMO channel matrix and S_n is the so called signal matrix, see [38]). Taking $R_n = n^{-1/2} S_n A_n^*$ in (5) and applying the law of large numbers, one can see that for any $f \in \mathcal{C}([0, \mathbf{d}_{\max}])$, the integral $\int f(t)\Lambda_n(dt)$ converges to $\int f(t)\Lambda_*(dt)$ with $\Lambda_*(dt) = \nu(dt) \times M$. Clearly, the accumulation points of the measures obtained from any other sequence of factorizations of P_n are of the form $\nu(dt) \times WMW^*$ where W is an $r \times r$ unitary matrix.

It is shown in [37] that the limiting spectral measure μ has a continuous density on $\mathbb{R}^* = \mathbb{R} - \{0\}$ (see Prop. 3.1 below). Our first order result addresses the problem of the presence of isolated eigenvalues of $\Sigma_n \Sigma_n^*$ in any compact interval outside the support of this density. Of prime importance will be the $r \times r$ matrix functions

$$H_*(z) = \int \frac{\mathbf{m}(z)}{1 + c\mathbf{m}(z)t} \Lambda_*(dt)$$

where Λ_* is an accumulation point of a sequence Λ_n. Since $|1 + c\mathbf{m}(z)t| = |z(1 + c\mathbf{m}(z)t)|/|z| \geq |\Im(z(1 + c\mathbf{m}(z)t))|/|z| \geq \Im(z)/|z|$ on $\mathbb{C}_+$, the function $H_*(z)$ is analytic on $\mathbb{C}_+$. It is further easy to show that $\Im(H_*(z)) \geq 0$ and $\Im(zH_*(z)) \geq 0$ on $\mathbb{C}_+$, and $\sup_{y>0} \|yH_*(\imath y)\| < \infty$. Hence $H_*(z)$ is the Stieltjes Transform of a matrix-valued nonnegative finite measure carried by $[0, \infty)$. Note also that, under Assumption 5, the eigenvalues of $H_*(z)$ remain unchanged if Λ_* is replaced by another accumulation point.

The support of μ may consist in several connected components corresponding to as many "bulks" of eigenvalues. Our first theorem specifies the locations of the outliers between any two bulks and on the right of the last bulk. It also shows that there are no outliers on the left of the first bulk:

Theorem 2.2. *Let Assumptions 1, 2 and 3 hold true. Denote by* $(\hat{\lambda}_i^n)_{i=1,\ldots,N}$ *the eigenvalues of* $\Sigma_n \Sigma_n^*$. *Let* (a, b) *be any connected component of* $\mathrm{supp}(\mu)^c = \mathbb{R} - \mathrm{supp}(\mu)$. *Then the following facts hold true:*

(1) *Let (P_n) be a sequence satisfying Assumptions 4 and 5. Given an accumulation point Λ_* of a sequence Λ_n, let $H_*(z) = \int \mathbf{m}(z)(1+c\mathbf{m}(z)t)^{-1}\Lambda_*(dt)$. Then $H_*(z)$ can be analytically extended to (a,b) where its values are Hermitian matrices, and the extension is increasing in the order of Hermitian matrices on (a,b). The function $\mathcal{D}(x) = \det(H_*(x) + I_r)$ has at most r zeros on (a,b). Let $\rho_1, \ldots, \rho_k$, $k \leq r$ be these zeros counting multiplicities. Let $\mathcal{I}$ be any compact interval in (a,b) such that $\{\rho_1, \ldots, \rho_k\} \cap \partial\mathcal{I} = \emptyset$. Then*

$$\sharp\{i : \hat{\lambda}_i^n \in \mathcal{I}\} = \sharp\{i : \rho_i \in \mathcal{I}\} \text{ with probability } 1 \text{ for all large } n.$$

(2) *Let $A = \inf(\mathrm{supp}(\mu) - \{0\})$. Then for any positive $A' < A$ (assuming it exists) and for any sequence of matrices (P_n) satisfying Assumption 4,*

$$\sharp\{i : \hat{\lambda}_i^n \in (0, A']\} = 0 \text{ with probability } 1 \text{ for all large } n.$$

Given any sequence of positive real numbers $\rho_1 \leq \cdots \leq \rho_r$ lying in a connected component of $\mathrm{supp}(\mu)^c$ after the first bulk, it would be interesting to see whether there exists a sequence of matrices P_n that produces outliers converging to these ρ_k. The following theorem answers this question positively:

Theorem 2.3. *Let Assumptions 1, 2 and 3 hold true. Let $\rho_1 \leq \cdots \leq \rho_r$ be a sequence of positive real numbers lying in a connected component (a,b) of $\mathrm{supp}(\mu)^c$, and such that $a > A$. Then there exists a sequence of matrices P_n satisfying Assumptions 4 and 5 such that for any compact interval $\mathcal{I} \subset (a,b)$ with $\{\rho_1, \ldots, \rho_r\} \cap \partial\mathcal{I} = \emptyset$,*

$$\sharp\{i : \hat{\lambda}_i^n \in \mathcal{I}\} = \sharp\{i : \rho_i \in \mathcal{I}\} \text{ with probability } 1 \text{ for all large } n.$$

It would be interesting to complete the results of these theorems by specifying the indices of the outliers $\hat{\lambda}_i^n$ that appear between the bulks. This demanding analysis might be done by following the ideas of [11] or [39] relative to the so called separation of the eigenvalues of $\Sigma_n\Sigma_n^*$. Another approach dealing with the same kind of problem is developed in [4].

A case of practical importance at least in the domain of signal processing is described by the following assumption:

Assumption 6. *The accumulation points Λ_* are of the form $\nu(dt) \times W\mathbf{\Omega}W^*$ where*

$$\mathbf{\Omega} = \begin{bmatrix} \omega_1^2 I_{j_1} & & \\ & \ddots & \\ & & \omega_t^2 I_{j_t} \end{bmatrix} > 0, \quad \omega_1^2 > \cdots > \omega_t^2, \quad j_1 + \cdots + j_t = r$$

and where W is a unitary matrix.

Because of the specific structure of S_n in the factorization $P_n = n^{-1/2}U_nA_nS_n^*$, the MIMO wireless communication model described above satisfies this assumption, the ω_i^2 often referring to the powers of the radio sources transmitting their signals to an array of antennas.

Another case where Assumption 6 is satisfied is the case where P_n is a random matrix independent of X_n, where its probability distribution is invariant by right multiplication with a constant unitary matrix, and where the r non zero singular values of P_n converge almost surely towards constant values.

When this assumption is satisfied, we obtain the following corollary of Theorem 2.2 which exhibits some sort of phase transition analogous to the so-called BBP phase transition [5]:

Corollary 2.1. *Assume the setting of Theorem 2.2-(1), and let Assumption 6 hold true. Then the function* $g(x) = \mathbf{m}(x)\,(cx\mathbf{m}(x) - 1 + c)$ *is decreasing on* (a, b). *Depending on the value of* ω_ℓ^2, $\ell = 1, \ldots, t$, *the equation* $\omega_\ell^2 g(x) = 1$ *has either zero or one solution in* (a, b). *Denote by* $\rho_1, \ldots, \rho_k$, $k \leq r$ *these solutions counting multiplicities. Then the conclusion of Theorem 2.2-(1) hold true for these* ρ_i.

We now turn to the second order result, which will be stated in the simple and practical framework of Assumption 6. Actually, a stronger assumption is needed:

Assumption 7. *The following facts hold true:*

$$\sup_n \sqrt{n}|c_n - c| < \infty,$$

$$\limsup_n \sqrt{n}\left|\int \frac{1}{t-x}\nu_n(dt) \;-\; \int \frac{1}{t-x}\nu(dt)\right| < \infty \text{ for all } x \in \mathbb{R} - \mathrm{supp}(\nu).$$

Moreover, there exists a sequence of factorizations of P_n *such that the measures* Λ_n *associated with these factorizations by* (5) *converge to* $\nu(dt) \times \boldsymbol{\Omega}$ *and such that*

$$\limsup_n \sqrt{n}\left\|\int \frac{1}{t-x}\Lambda_n(dt) \;-\; \int \frac{1}{t-x}\nu(dt) \times \boldsymbol{\Omega}\right\| < \infty \text{ for all } x \in \mathbb{R} - \mathrm{supp}(\nu)$$

Note that one could have considered the above superior limits to be zero, which would simplify the statement of Theorem 2.4 below. However, in practice this is usually too strong a requirement, see *e.g.* the wireless communications model discussed after Assumption 5 for which the fluctuations of Λ_n are of order $n^{-1/2}$. On the opposite, slower fluctuations of Λ_n would result in a much more intricate result for Theorem 2.4, which we do not consider here.

Before stating the second order result, a refinement of the results of Theorem 2.1–(1) is needed:

Proposition 2.1 ([36, 22, 18])**.** *Assume that* D_n *is a* $n \times n$ *diagonal nonnegative matrix. Then, for any* n, *the equation*

$$m_n = \left[-z\left(1 + \frac{1}{n}\operatorname{Tr} D_n \widetilde{T}_n\right)\right]^{-1} \quad \text{where} \quad \widetilde{T}_n = \left[-z\,(I_n + c_n m_n D_n)\right]^{-1}$$

admits a unique solution $m_n \in \mathbb{C}_+$ *for any* $z \in \mathbb{C}_+$. *The function* $m_n = m_n(z)$ *so defined on* $\mathbb{C}_+$ *is the Stieltjes Transform of a probability measure* μ_n *whose support is a compact set of* $\mathbb{R}_+$. *Moreover, the* $n \times n$ *diagonal matrix-valued function* $\widetilde{T}_n(z) = [-z(I_n + c_n m_n(z) D_n)]^{-1}$ *is analytic on* $\mathbb{C}_+$ *and* $n^{-1}\operatorname{Tr}\widetilde{T}_n(z)$ *coincides with the Stieltjes Transform of* $c_n\mu_n + (1 - c_n)\boldsymbol{\delta}_0$.
Let Assumption 2 hold true, and assume that $\sup_n \|D_n\| < \infty$, *and* $0 < \liminf c_n \leq \limsup c_n < \infty$. *Then the resolvents* $Q_n(z) = (n^{-1}Y_nY_n^* - zI_N)^{-1}$ *and* $\widetilde{Q}_n(z) = (n^{-1}Y_n^*Y_n - zI_n)^{-1}$ *satisfy*

$$\frac{1}{N}\operatorname{Tr}\left(Q_n(z) - m_n(z)I_N\right) \xrightarrow[n\to\infty]{a.s.} 0 \quad \text{and} \quad \frac{1}{n}\operatorname{Tr}\left(\widetilde{Q}_n(z) - \widetilde{T}_n(z)\right) \xrightarrow[n\to\infty]{a.s.} 0 \qquad (6)$$

for any $z \in \mathbb{C}_+$. *When in addition Assumptions 1 and 3 hold true,* $m_n(z)$ *converges to* $\mathbf{m}(z)$ *provided in the statement of Theorem 2.1 uniformly on the compact subsets of* $\mathbb{C}_+$.

The function $m_n(z) = \left(-z + \int t(1+c_n m_n(z)t)^{-1}\nu_n(dt)\right)^{-1}$ is a finite n approximation of $\mathbf{m}(z)$. Notice that since $N^{-1}\operatorname{Tr} Q_n(z)$ is the Stieltjes Transform of the spectral measure θ_n of $n^{-1}Y_nY_n^*$, Convergence (4) stems from (6).
We shall also need a finite n approximation of $H_*(z)$ defined as

$$H_n(z) = \int \frac{m_n(z)}{1+c_n m_n(z)t}\Lambda_n(dt).$$

With these definitions, we have the following preliminary proposition:

Proposition 2.2. *Let Assumptions 1, 3-7 hold true. Let g be the function defined in the statement of Corollary 2.1 and let $B_\mu = \sup(\operatorname{supp}(\mu))$. Assume that the equation $\omega_1^2 g(x) = 1$ has a solution in (B_μ, ∞), and denote $\rho_1 > \cdots > \rho_p$ the existing solutions (with respective multiplicities $j_1, \ldots, j_p$) of the equations $\omega_\ell^2 g(x) = 1$ in (B_μ, ∞). Then the following facts hold true:*

- $\boldsymbol{\Delta}(\rho_i) = 1 - c\int \left(\frac{\mathbf{m}(\rho_i)t}{1+c\mathbf{m}(\rho_i)t}\right)^2 \nu(dt)$ *is positive for every $i = 1, \ldots, p$.*
- *Denoting by $H_{1,n}(z), \ldots, H_{p,n}(z)$ the first p upper left diagonal blocks of $H_n(z)$, where $H_{i,n}(z) \in \mathbb{C}^{j_i \times j_i}$, $\limsup_n \|\sqrt{n}(H_{i,n}(\rho_i) + I_{j_i})\| < \infty$ for every $i = 1, \ldots, p$.*

We recall that a GUE matrix (*i.e.*, a matrix taken from the Gaussian Unitary Ensemble) is a random Hermitian matrix G such that $G_{ii} \sim \mathcal{N}(0,1)$, $\Re(G_{ij}) \sim \mathcal{N}(0,1/2)$ and $\Im(G_{ij}) \sim \mathcal{N}(0,1/2)$ for $i < j$, and such that all these random variables are independent. Our second order result is provided by the following theorem:

Theorem 2.4. *Let Assumptions 1-7 hold true. Keeping the notations of Proposition 2.2, let*

$$M_i^n = \sqrt{n}\left(\begin{bmatrix} \hat{\lambda}^n_{j_1+\cdots+j_{i-1}+1} \\ \vdots \\ \hat{\lambda}^n_{j_1+\cdots+j_i} \end{bmatrix} - \rho_i \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}\right)$$

where $j_0 = 0$ and where the eigenvalues $\hat{\lambda}_i^n$ of $\Sigma_n\Sigma_n^*$ are arranged in decreasing order. Let $G_1, \ldots, G_p$ be independent GUE matrices such that G_i is a $j_i \times j_i$ matrix. Then, for any bounded continuous $f\colon \mathbb{R}^{j_1+\ldots+j_p} \to \mathbb{R}$,

$$\mathbb{E}\left[f\left(M_1^n, \ldots, M_p^n\right)\right] - \mathbb{E}\left[f\left(\Xi_1^n, \ldots, \Xi_p^n\right)\right] \underset{n\to\infty}{\longrightarrow} 0$$

where $\Xi_i^n \in \mathbb{R}^{j_i}$ is the random vector of the decreasingly ordered eigenvalues of the matrix

$$\frac{1}{\omega_i^2 g(\rho_i)'}\left(\boldsymbol{\alpha}_i G_i + \sqrt{n}\left(H_{i,n}(\rho_i) + I_{j_i}\right)\right),$$

where

$$\boldsymbol{\alpha}_i^2 = \frac{\mathbf{m}^2(\rho_i)}{\boldsymbol{\Delta}(\rho_i)}\left[\int \frac{t^2 + 2\omega_i^2 t}{(1+c\mathbf{m}(\rho_i)t)^2}\nu(dt) + c\left(\int \frac{\omega_i^2 \mathbf{m}(\rho_i)t}{(1+c\mathbf{m}(\rho_i)t)^2}\nu(dt)\right)^2\right].$$

Some remarks can be useful at this stage. The first remark concerns Assumption 7, which is in some sense analogous to [7, Hypothesis 3.1]. This assumption is mainly needed to show that the $\sqrt{n}\|H_{i,n}(\rho_i) + I_{j_i}\|$ are bounded, guaranteeing the tightness of the vectors M_i^n. Assuming that Λ_n and Λ_n' both satisfy the third

item of Assumption 7, denoting respectively by $H_{i,n}(\rho_i)$ and $H'_{i,n}(\rho_i)$ the matrices associated to these measures as in the statement of Theorem 2.4, it is possible to show that $\sqrt{n}(\sigma(H_{i,n}(\rho_i) - \sigma(H'_{i,n}(\rho_i)) \to 0$ as $n \to \infty$. Thus the results of this theorem do not depend on the particular measure Λ_n satisfying Assumption 7. Finally, we note that Assumption 7 can be lightened at the expense of replacing the limit values ρ_i with certain finite n approximations of the outliers, as is done in the applicative paper [40].

The second remark pertains to the Gaussian assumption on the elements of X_n. We shall see below that the results of Theorems 2.2–2.4 are intimately related to the first and second order behaviors of bilinear forms of the type $u_n^* Q_n(x) v_n$, $\tilde{u}_n^* \widetilde{Q}_n(x) \tilde{v}_n$, and $n^{-1/2} u_n^* Y_n \widetilde{Q}_n(x) \tilde{v}_n$ where u_n, v_n, $\tilde{u}_n$ and $\tilde{v}_n$ are deterministic vectors of bounded norm and of appropriate dimensions, and where x is a real number lying outside the support of μ. In fact, it is possible to generalize Theorems 2.2 and 2.3 to the case where the elements of X_n are not necessarily Gaussian. This can be made possible by using the technique of [21] to analyze the first order behavior of these bilinear forms. On the other hand, the Gaussian assumption plays a central role in Theorem 2.4. Indeed, the proof of this theorem is based on the fact that these bilinear forms asymptotically fluctuate like Gaussian random variables when centered and scaled by $\sqrt{n}$. Take $u_n = e_{1,N}$ and $\tilde{v}_n = e_{1,n}$ where $e_{k,m}$ is the k^{th} canonical vector of $\mathbb{R}^m$. We show below (see Proposition 2.1 and Lemmas 4.3 and 4.6) that the elements $\tilde{q}_{ij}^n$ of the resolvent $\widetilde{Q}_n(x)$ are close for large n to the elements $\tilde{t}_{ij}^n$ of the deterministic matrix $\widetilde{T}_n(x)$. We therefore write informally

$$e_{1,N}^* Y_n \widetilde{Q}(x) e_{1,n} = \sum_{j=1}^n (d_j^n)^{1/2} \tilde{q}_{j1}^n X_{1j} \approx (d_1^n)^{1/2} \tilde{t}_{11}^n X_{11} + \sum_{j=2}^n (d_j^n)^{1/2} \tilde{q}_{j1}^n X_{1j}.$$

It can be shown furthermore that $\tilde{t}_{11}^n = \mathcal{O}(1)$ for large n and that the sum $\sum_{j=2}^n$ is tight. Hence, $e_{1,N}^* Y_n \widetilde{Q}(x) e_{1,n}$ is tight. However, when X_{11} is not Gaussian, we infer that $e_{1,N}^* Y_n \widetilde{Q}(x) e_{1,n}$ does not converge in general towards a Gaussian random variable. In this case, if we choose $P_n = \omega^2 e_{1,N} e_{1,n}^*$ (see Section 5), Theorem 2.4 no longer holds. Yet, we conjecture that an analogue of this theorem can be recovered when $e_{1,N}$ and $e_{1,n}$ are replaced with delocalized vectors, following the terminology of [12]. In a word, the elements of these vectors are "spread enough" so that the Gaussian fluctuations are recovered.

A word about the notations. In the remainder of the paper, we shall often drop the subscript or the superscript n when there is no ambiguity. A constant bound that may change from an inequality to another but which is independent of n will always be denoted K. Element (i, j) of matrix M is denoted M_{ij} or $[M]_{ij}$. Element i of vector x is denoted $[x]_i$. Convergences in the almost sure sense, in probability and in distribution will be respectively denoted $\xrightarrow{\text{a.s.}}$, $\xrightarrow{\mathcal{P}}$, and $\xrightarrow{\mathcal{D}}$.

3. Preliminaries and useful results

We start this section by providing the main ideas of the proofs of Theorems 2.2 and 2.3.

3.1. **Proof principles of the first order results.** The proof of Theorem 2.2-(1), to begin with, is based on the idea of [8, 9]. We start with a purely algebraic result. Let $P_n = U_n R_n^*$ be a factorization of P_n where U_n is a $N \times r$ isometry matrix. Assume that $x > 0$ is not an eigenvalue of $n^{-1}Y_nY_n^*$. Then x is an eigenvalue of $\Sigma_n\Sigma_n^*$ if and only if $\det \widehat{S}_n(x) = 0$ where $\widehat{S}_n(x)$ is the $2r \times 2r$ matrix

$$\widehat{S}_n(x) = \begin{bmatrix} \sqrt{x}U_n^*Q_n(x)U_n & I_r + n^{-1/2}U_n^*Y_n\widetilde{Q}_n(x)R_n \\ I_r + n^{-1/2}R_n^*\widetilde{Q}_n(x)Y_n^*U_n & \sqrt{x}R_n^*\widetilde{Q}_n(x)R_n \end{bmatrix}$$

(for details, see the derivations in [9] or in [20, Section 3]). The idea is now the following. Set x in $\mathrm{supp}(\mu)^c$. Using an integration by parts formula for functionals of Gaussian vectors and the Poincaré-Nash inequality [31], we show that when n is large,

$$U_n^*Q_n(x)U_n \approx m_n(x)I_r,\ R_n^*\widetilde{Q}_n(x)R_n \approx R_n^*\widetilde{T}_n(x)R_n,$$
$$\text{and } n^{-1/2}R_n^*\widetilde{Q}_n(x)Y_n^*U_n \approx 0$$

by controlling the moments of the elements of the left hand members. To be able to do these controls, we make use of a certain regularizing function which controls the escape of the eigenvalues of $n^{-1}Y_nY_n^*$ out of $\mathrm{supp}(\mu)$. Thanks to these results, $\widehat{S}_n(x)$ is close for large n to

$$S_n(x) = \begin{bmatrix} \sqrt{x}m_n(x)I_r & I_r \\ I_r & \sqrt{x}R_n^*\widetilde{T}_n(x)R_n \end{bmatrix}.$$

Hence, we expect the eigenvalues of $\Sigma_n\Sigma_n^*$ in the interval $\mathcal{I}$, when they exist, to be close for large n to the zeros in $\mathcal{I}$ of the function

$$\begin{aligned} \det S_n(x) &= \det\left(xm_n(x)R_n^*\widetilde{T}_n(x)R_n - I_r\right) \\ &= (-1)^r \det\left(m_n(x)R_n^*(I_n + c_nm_n(x)D_n)^{-1}R_n^* + I_r\right) \\ &= (-1)^r \det\left(H_n(x) + I_r\right) \end{aligned}$$

which are close to the zeros of $\mathcal{D}(x) = \det(H_*(x) + I_r)$. By Assumption 5, these zeros are independent of the choice of the accumulation point Λ_*.

To prove Theorems 2.2-(2) and 2.3, we make use of the results of [37] and [27, 28] relative to the properties of μ and to those of the restriction of $\mathbf{m}(z)$ to $\mathbb{R} - \mathrm{supp}(\mu)$. The main idea is to show that

- $\mathbf{m}(x)(1 + c\mathbf{m}(x)t)^{-1} > 0$ for all $x \in \mathrm{supp}(\mu)^c \cap (-\infty, A)$ (these x lie at the left of the first bulk) and for all $t \in \mathrm{supp}(\nu)$.
- For any component $(a, b) \subset \mathrm{supp}(\mu)^c$ such that $a > A$ (*i.e.*, lying between two bulks or at the right of the last bulk), there exists a Borel set $E \subset \mathbb{R}_+$ such that $\nu(E) > 0$ and

$$q(x) = \int_E \frac{\mathbf{m}(x)}{1 + c\mathbf{m}(x)t}\nu(dt) < 0$$

for all $x \in (a, b)$.

Thanks to the first result, for any x lying if possible between zero and the left edge of the first bulk, $\mathcal{D}(x) > 0$, hence $\Sigma_n\Sigma_n^*$ has asymptotically no outlier at the left of the first bulk.
Coming to Theorem 2.3, let E be a set associated to (a, b) by the result above. We

build a sequence of matrices P_n of rank r, and such that the associated Λ_n have an accumulation point of the form $\Lambda_*(dt) = 1_E(t)\,\mathbf{\Omega}\,\nu(dt)$ where we choose $\mathbf{\Omega} = \mathrm{diag}(-q(\rho_1)^{-1}, \ldots, -q(\rho_r)^{-1})$. Theorem 2.2-(1) shows that the function $H_*(x) = q(x)\mathbf{\Omega}$ associated with this Λ_* is increasing on (a, b). As a result, $H_*(x)+I_r$ becomes singular precisely at the points $\rho_1, \ldots, \rho_r$.

3.2. **Sketch of the proof of the second order result.** The fluctuations of the outliers will be deduced from the fluctuations of the elements of the matrices $\widehat{S}_n(\rho_i)$ introduced above. The proof of Theorem 2.4 can be divided into two main steps. The first step (Lemma 5.4) consists in establishing a Central Limit Theorem on the $3p$–uple of random matrices

$$\sqrt{n}\left(\frac{U_{i,n}^* Y_n \widetilde{Q}_n(\rho_i) R_{i,n}}{\sqrt{n}}, U_{i,n}^*(Q_n(\rho_i) - m_n(\rho_i) I_N) U_{i,n}, \right.$$
$$\left. R_{i,n}^*(\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n}\right)_{i=1}^{p},$$

where $P_n = U_n R_n^*$ is a sequence of factorization such that Λ_n satisfies the third item of Assumption 7. We also write $U_n = [U_{1,n} \cdots U_{t,n}]$ and $R_n = [R_{1,n} \cdots R_{t,n}]$ where $U_{i,n} \in \mathbb{C}^{N \times j_i}$ and $R_{i,n} \in \mathbb{C}^{n \times j_i}$.

This CLT is proven by using the Gaussian tools introduced in Section 3.4, namely the integration by parts formula and the Poincaré-Nash inequality, and by relying on the invariance properties of the probability distribution of $n^{-1/2}X_n$. The fluctuations of the zeros of $\det \widehat{S}_n(x)$ outside $\mathrm{supp}(\mu)$ are then deduced from this CLT by adapting the approach of [7] (Lemma 5.5).

We now come to the basic mathematical tools needed for our proofs:

3.3. **Differentiation formulas.** Let $\partial/\partial z = (\partial/\partial x - \imath\partial/\partial y)/2$ and $\partial/\partial \bar{z} = (\partial/\partial x + \imath\partial/\partial y)/2$ for $z = x + \imath y$. Given a Hermitian matrix X with a spectral decomposition $X = \sum_\ell \lambda_\ell v_\ell v_\ell^*$, let $\mathrm{adj}(X) = \sum_k (\prod_{\ell \neq k} \lambda_\ell) v_k v_k^*$ be the classical adjoint of X, *i.e.*, the transpose of its cofactor matrix. Let ψ be a continuously differentiable real-valued function on $\mathbb{R}$. Then

$$\frac{\partial \det \psi\left(n^{-1}YY^*\right)}{\partial \bar{Y}_{ij}} = \frac{1}{n}\left[\mathrm{adj}\left(\psi\left(n^{-1}YY^*\right)\right)\, \psi'\left(n^{-1}YY^*\right)\, y_j\right]_i$$

where y_j is column j of Y, see [19, Lemma 3.9] for a proof.
We shall also need the expressions of the following derivatives of the elements of the resolvents Q and $\widetilde{Q}$ (see [18]):

$$\frac{\partial Q_{pq}}{\partial \bar{Y}_{ij}} = -\frac{1}{n}[QY]_{pj} Q_{iq}, \quad \frac{\partial \widetilde{Q}_{pq}}{\partial \bar{Y}_{ij}} = -\frac{1}{n}\widetilde{Q}_{pj}[Y\widetilde{Q}]_{iq}.$$

3.4. **Gaussian tools.** Our analysis fundamentally relies on two mathematical tools which are often used in the analysis of large random matrices with Gaussian elements. The first is the so called integration by parts (IP) formula for functionals of Gaussian vectors introduced in random matrix theory in [24, 30]. Let $\Gamma : \mathbb{R}^{2Nn} \to \mathbb{C}$ be a continuously differentiable function polynomially bounded together with its partial derivatives. Then

$$\mathbb{E}\left(Y_{ij}\Gamma(Y)\right) = d_j \mathbb{E}\left[\frac{\partial \Gamma(Y)}{\partial \bar{Y}_{ij}}\right]$$

for any $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, n\}$. The second tool is the Poincaré-Nash inequality (see for instance [13]). In our situation, it states that the variance $\mathbb{V}\mathrm{ar}(\Gamma(Y))$ satisfies

$$\mathbb{V}\mathrm{ar}(\Gamma(Y)) \leq \sum_{i=1}^{N} \sum_{j=1}^{n} d_j \mathbb{E}\left[\left| \frac{\partial \Gamma(Y)}{\partial Y_{ij}} \right|^2 + \left| \frac{\partial \Gamma(Y)}{\partial \bar{Y}_{ij}} \right|^2 \right] .$$

We now recall the results of Silverstein and Choi [37] which will be needed to prove Theorems 2.2-(2) and 2.3. Close results can be found in [27] and in [28].

3.5. **Analysis of the support of μ.**

Proposition 3.1 ([37], Th.1.1)**.** *For all $x \in \mathbb{R}^*$, $\lim_{z \in \mathbb{C}_+ \to x} \mathbf{m}(z)$ exists. The limit that we denote $\mathbf{m}(x)$ is continuous on $\mathbb{R}^*$. Moreover, μ has a continuous density f on $\mathbb{R}^*$ given by $f(x) = \pi^{-1} \Im \mathbf{m}(x)$.*

In [37], the support of μ is also identified. Since $\mathbf{m}(z)$ is the unique solution in $\mathbb{C}_+$ of (3) for $z \in \mathbb{C}_+$, it has a unique inverse on $\mathbb{C}_+$ given by

$$z(\mathbf{m}) = -\frac{1}{\mathbf{m}} + \int \frac{t}{1 + c\mathbf{m}t} \nu(dt)$$

The characterization of the support of μ is based on the following idea. On any open interval of $\mathrm{supp}(\mu)^c$, $\mathbf{m}(x) = \int (t-x)^{-1} \mu(dt)$ is a real, continuous and increasing function. Consequently, it has a real, continuous and increasing inverse. In [37], it is shown that the converse is also true. More precisely, let $B = \{\mathbf{m} : \mathbf{m} \neq 0, -(c\mathbf{m})^{-1} \in \mathrm{supp}(\nu)^c\}$, and let

$$\begin{array}{rccc} x & : & B & \longrightarrow & \mathbb{R} \\ & & \mathbf{m} & \longmapsto & x(\mathbf{m}) = -\dfrac{1}{\mathbf{m}} + \displaystyle\int \frac{t}{1 + c\mathbf{m}t} \, \nu(dt). \end{array} \tag{7}$$

Then the following proposition holds:

Proposition 3.2 ([37], Th. 4.1 and 4.2)**.** *For any $x_0 \in \mathrm{supp}(\mu)^c$, let $\mathbf{m}_0 = \mathbf{m}(x_0)$. Then $\mathbf{m}_0 \in B$, $x_0 = x(\mathbf{m}_0)$, and $x'(\mathbf{m}_0) > 0$. Conversely, let $\mathbf{m}_0 \in B$ such that $x'(\mathbf{m}_0) > 0$. Then $x_0 = x(\mathbf{m}_0) \in \mathrm{supp}(\mu)^c$, and $\mathbf{m}(x_0) = \mathbf{m}_0$.*

The following proposition will also be useful:

Proposition 3.3 ([37], Th. 4.4)**.** *Let $[\mathbf{m}_1, \mathbf{m}_2]$ and $[\mathbf{m}_3, \mathbf{m}_4]$ be two disjoint intervals of B satisfying $\forall \mathbf{m} \in (\mathbf{m}_1, \mathbf{m}_2) \cup (\mathbf{m}_3, \mathbf{m}_4)$, $x'(\mathbf{m}) > 0$. Then $[x_1, x_2]$ and $[x_3, x_4]$ are disjoint where $x_i = x(\mathbf{m}_i)$.*

The following result is also proven in [37]:

Proposition 3.4. *Assume that $\nu(\{0\}) = 0$. Then $\mu(\{0\}) = \max(0, 1 - c^{-1})$.*

We shall assume hereafter that $\nu(\{0\}) = 0$ without loss of generality (otherwise, it would be enough to change the value of c). The two following lemmas will also be needed:

Lemma 3.1. *Let $\mathcal{I}$ be a compact interval of $\mathrm{supp}(\mu)^c$, and let $D_{\mathcal{I}}$ be the closed disk having $\mathcal{I}$ as one of its diameters. Then there exists a constant K which depends on $\mathcal{I}$ only such that*

$$\forall\, t \in \mathrm{supp}(\nu),\ \forall\, z \in D_{\mathcal{I}},\ |1 + c\mathbf{m}(z)t| \geq K, \quad \textit{and}$$
$$\forall\, n \textit{ large enough},\ \forall\, t \in \mathrm{supp}(\nu_n),\ \forall\, z \in D_{\mathcal{I}},\ |1 + c_n m_n(z)t| \geq K.$$

From the second inequality, we deduce that if $\{0\} \notin \mathcal{I}$, the matrix function $\widetilde{T}_n(z)$ is analytic in a neighborhood of $\mathcal{I}$ for n large enough, and

$$\limsup_n \sup_{z \in D_{\mathcal{I}}} \|\widetilde{T}_n(z)\| < \infty. \tag{8}$$

Proof. When $z \in \mathbb{C}_+$, $\Im \mathbf{m}(z) > 0$ and $\Im(-(c\mathbf{m}(z))^{-1}) > 0$, and we have the opposite inequalities when $\Im z < 0$. Applying Proposition 3.2 for $z \in \mathcal{I}$, we deduce that $|\mathbf{m}(z)|$ and $f(z) = \boldsymbol{d}(-(c\mathbf{m}(z))^{-1}, \mathrm{supp}(\nu))$ are positive on $D_{\mathcal{I}}$. Since these functions are continuous on this compact set, $\min |\mathbf{m}(z)| = K_1 > 0$ and $\min f(z) = K_2 > 0$ on $D_{\mathcal{I}}$. Consequently, for any $z \in D_{\mathcal{I}}$ and any $t \in \mathrm{supp}(\nu)$, $|1 + c\mathbf{m}(z)t| = |c\mathbf{m}(z)(-(c\mathbf{m}(z))^{-1} - t)| \geq |c\mathbf{m}(z)| f(z) \geq cK_1K_2 > 0$. We now prove the second inequality. Denote by $\boldsymbol{d}_{\mathrm{H}}(A, B)$ the Hausdorff distance between two sets A and B. Let $f_n(z) = \boldsymbol{d}(-(c_n m_n(z))^{-1}, \mathrm{supp}(\nu_n))$. We have

$$\begin{aligned} f_n(z) &\leq \boldsymbol{d}\Big(\frac{-1}{c_n m_n(z)}, \frac{-1}{c\mathbf{m}(z)}\Big) + \boldsymbol{d}\Big(\frac{-1}{c\mathbf{m}(z)}, \mathrm{supp}(\nu_n)\Big) \\ &\leq \boldsymbol{d}\Big(\frac{-1}{c_n m_n(z)}, \frac{-1}{c\mathbf{m}(z)}\Big) + f(z) + \boldsymbol{d}_{\mathrm{H}}(\mathrm{supp}(\nu_n), \mathrm{supp}(\nu)), \end{aligned}$$

and $f(z) \leq \boldsymbol{d}(-(c_n m_n(z))^{-1}, -(c\mathbf{m}(z))^{-1}) + f_n(z) + \boldsymbol{d}_{\mathrm{H}}(\mathrm{supp}(\nu_n), \mathrm{supp}(\nu))$ similarly. Since $m_n(z)$ converges uniformly to $\mathbf{m}(z)$ and $\inf |\mathbf{m}(z)| > 0$ on $D_{\mathcal{I}}$, we get that $\boldsymbol{d}(-(c_n m_n(z))^{-1}, -(c\mathbf{m}(z))^{-1}) \to 0$ uniformly on this disk. By Assumption 3, $\boldsymbol{d}_{\mathrm{H}}(\mathrm{supp}(\nu_n), \mathrm{supp}(\nu)) \to 0$. Hence $f_n(z)$ converges uniformly to $f(z)$ on $D_{\mathcal{I}}$ which proves the second inequality. □

Lemma 3.2. *Assume the setting of Lemma 3.1, and assume that $\{0\} \notin \mathcal{I}$. Then for any sequence of vectors $\tilde{u}_n \in \mathbb{C}^n$ such that $\sup_n \|\tilde{u}_n\| < \infty$, the quadratic forms $\tilde{u}_n^* \widetilde{T}_n(z) \tilde{u}_n$ are the Stieltjes Transforms of positive measures γ_n such that $\sup_n \gamma_n(\mathbb{R}) < \infty$ and $\gamma_n(\mathcal{I}) = 0$ for n large enough.*

Indeed, one can easily check the conditions that enable $\tilde{u}_n^* \widetilde{T}_n(z) \tilde{u}_n$ to be a Stieltjes Transform of a positive finite measure. The last result is obtained by analyticity in a neighborhood of $\mathcal{I}$. In fact, it can be checked that $\mathrm{supp}(\gamma_n) \subset \mathrm{supp}(\mu_n) \cup \{0\}$.

3.6. **A control over the support of θ_n.** In this paragraph, we adapt to our case an idea developed in [11] to deal with Wigner matrices whose elements distribution satisfies a Poincaré-Nash inequality.

Proposition 3.5. *For any sequence of $n \times n$ deterministic diagonal nonnegative matrices $\widetilde{U}_n$ such that $\sup_n \|\widetilde{U}_n\| < \infty$,*

$$\left| \frac{1}{n} \operatorname{Tr} \mathbb{E} Q_n(z) - m_n(z) \right| \leq \frac{P(|z|) R(|\Im(z)|^{-1})}{n^2}, \text{ and}$$

$$\left| \frac{1}{n} \operatorname{Tr} \widetilde{U}_n \mathbb{E} \widetilde{Q}_n(z) - \frac{1}{n} \operatorname{Tr} \widetilde{U}_n \widetilde{T}_n(z) \right| \leq \frac{P(|z|) R(|\Im(z)|^{-1})}{n^2}$$

for $z \in \mathbb{C}_+$, where P and R are polynomials with nonnegative coefficients independent of n.

This proposition is obtained from a simple extension of the results of [18, Th. 3 and Prop.5] from $z \in (-\infty, 0)$ to $z \in \mathbb{C}_+$.
The following important result, due to Haagerup and Thorbjørnsen, is established in the proof of [17, Th.6.2]:

Lemma 3.3. *Assume that $h(z)$ is an analytic function on $\mathbb{C}_+$ that satisfies $|h(z)| \leq P(|z|)R(|\Im(z)|^{-1})$ where P and R are polynomials with nonnegative coefficients. Then for any function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$, the set of smooth real-valued functions with compact support in $\mathbb{R}$,*

$$\limsup_{y\downarrow 0} \left| \int_{\mathbb{R}} \varphi(x) h(x+\imath y) dx \right| < \infty.$$

Since $N^{-1}\operatorname{Tr} Q_n(z)$ is the Stieltjes Transform of the spectral measure θ_n, the inversion formula (2) shows that

$$\int \varphi(t)\,\theta_n(dt) = \frac{1}{\pi} \lim_{y\downarrow 0} \Im \int \varphi(x) \frac{1}{N} \operatorname{Tr} Q_n(x+\imath y)\, dx$$

for any function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$. Using then Proposition 3.5 and Lemma 3.3, we obtain the following result:

Proposition 3.6. *For any function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$,*

$$\left| \mathbb{E} \int \varphi(t)\,\theta_n(dt) \;-\; \int \varphi(t)\,\mu_n(dt) \right| \leq \frac{K}{n^2}.$$

4. Proofs of first order results

In all this section, $\mathcal{I}$ is a compact interval of a component (a,b) of $\operatorname{supp}(\mu)^c$, and z is a complex number such that $\Re(z) \in \mathcal{I}$ and $\Im(z)$ is arbitrary. Moreover, $u_n, v_n \in \mathbb{C}^N$ and $\tilde{u}_n, \tilde{v}_n \in \mathbb{C}^n$ are sequences of deterministic vectors such that $\sup_n \max(\|u_n\|, \|v_n\|, \|\tilde{u}_n\|, \|\tilde{v}_n\|) < \infty$, and $\widetilde{U}_n$ is a sequence of $n \times n$ diagonal deterministic matrix such that $\sup_n \|\widetilde{U}_n\| < \infty$.
We now introduce the regularization function alluded to in the introduction. Choose $\varepsilon > 0$ small enough so that $\mathcal{I} \cap \mathcal{S}_\varepsilon = \emptyset$ where $\mathcal{S}_\varepsilon = \{x \in \mathbb{R}, \boldsymbol{d}(x, \operatorname{supp}(\mu) \cup \{0\}) \leq \varepsilon\}$. Fix $0 < \varepsilon' < \varepsilon$, let $\psi : \mathbb{R} \to [0,1]$ be a continuously differentiable function such that

$$\psi(x) = \begin{cases} 1 & \text{if } x \in \mathcal{S}_{\varepsilon'} \\ 0 & \text{if } x \in \mathbb{R} - \mathcal{S}_\varepsilon \end{cases}$$

and let $\phi_n = \det \psi(n^{-1} Y_n Y_n^*)$. In all the subsequent derivations, quantities such as $u_n^* Q_n(z) u_n$ or $\tilde{u}_n^* \widetilde{Q}_n(z) \tilde{u}_n$ for $\Re(z) \in \mathcal{I}$ will be multiplied by ϕ_n in order to control their magnitudes when z is close to the real axis. By performing this regularization as is done in [19], we shall be able to define and control the moments of random variables such as $\phi_n\, u_n^* Q_n(z) u_n$ or $\phi_n\, \tilde{u}_n^* \widetilde{Q}_n(z) \tilde{u}_n$ with the help of the Gaussian tools introduced in Section 3.4.

We start with a series of lemmas. The first of these lemmas relies on Proposition 3.6 and on the Poincaré-Nash inequality. Its detailed proof is a minor modification of the proof of [19, Lemma 3] and is therefore omitted:

Lemma 4.1. *Given $0 < \varepsilon' < \varepsilon$, let φ be a smooth nonnegative function equal to zero on $\mathcal{S}_{\varepsilon'}$ and to one on $\mathbb{R} - \mathcal{S}_\varepsilon$. Then for any $\ell \in \mathbb{N}$, there exists a constant K_ℓ for which*

$$\mathbb{E}\left[\left(\operatorname{Tr} \varphi(n^{-1} Y_n Y_n^*) \right)^\ell \right] \leq \frac{K_\ell}{n^\ell}.$$

Remark 1. *Notice that this lemma proves Theorem 2.1-*(2)*. The proof provided in* [3] *is in fact more general, being not restricted to the Gaussian case.*

Lemma 4.2. *For any $\ell \in \mathbb{N}$, the following holds true:*

$$\mathbb{E}\left[\left(\sum_{i,j=1}^{N,n} d_j \left|\frac{\partial \phi_n}{\partial \bar{Y}_{ij}}\right|^2\right)^\ell\right] \leq \frac{K_\ell}{n^{2\ell}}.$$

Proof. Letting $n^{-1/2}Y = W\,\mathrm{diag}(\sqrt{\lambda}_1, \cdots, \sqrt{\lambda}_N)V^*$ be a singular value decomposition of $n^{-1/2}Y$, we have

$$\mathrm{adj}\left(\psi\Big(\frac{YY^*}{n}\Big)\right)\psi'\Big(\frac{YY^*}{n}\Big)\frac{Y}{\sqrt{n}} = W\Xi V^* \text{ where } \Xi = \mathrm{diag}\Big(\sqrt{\lambda}_k\psi'(\lambda_k)\prod_{\ell\neq k}\psi(\lambda_\ell)\Big)_{k=1}^N$$

and we observe that $\mathrm{Tr}\,\Xi^2 \leq KZ_n$ where $Z_n = \sharp\{k \,:\, \lambda_k \in \mathcal{S}_\varepsilon - \mathcal{S}_{\varepsilon'}\}$. Using the first identity in Section 3.3 and recalling that $|\,\mathrm{Tr}(AB)| \leq \|A\|\,\mathrm{Tr}\,B$ when A is a square matrix and B is a Hermitian nonnegative matrix, we have

$$\mathbb{E}\left[\left(\sum_{i,j=1}^{N,n} d_j \left|\frac{\partial \phi_n}{\partial \bar{Y}_{ij}}\right|^2\right)^\ell\right] = \frac{1}{n^\ell}\mathbb{E}\left[\left(\mathrm{Tr}\left(\mathrm{adj}(\psi)\psi'\frac{YDY^*}{n}\,\mathrm{adj}(\psi)\psi'\right)\right)^\ell\right] \leq \frac{K}{n^\ell}\mathbb{E}Z_n^\ell$$

and the result follows from Lemma 4.1 with a proper choice of φ. □

Lemma 4.3. *The following inequalities hold true:*

$$\mathbb{E}\,|\phi_n\, u_n^* Q_n(z) v_n - \mathbb{E}[\phi_n\, u_n^* Q_n(z) v_n]|^4 \leq \frac{K}{n^2},$$
$$\mathbb{E}\left|\phi_n\, \tilde{u}_n^* \widetilde{Q}_n(z) \tilde{v}_n - \mathbb{E}[\phi_n\, \tilde{u}_n^* \widetilde{Q}_n(z) \tilde{v}_n]\right|^4 \leq \frac{K}{n^2},$$
$$\mathbb{V}\mathrm{ar}\,(\phi_n\,\mathrm{Tr}\,Q_n(z)) \leq K.$$

Proof. We shall only prove the first inequality. By the polarization identity, this inequality is shown whenever we show that $\mathbb{E}\,|\phi\, u^*Qu - \mathbb{E}[\phi\, u^*Qu]|^4 \leq K/n^2$. Let us start by showing that $\mathbb{V}\mathrm{ar}(\phi\, u^*Qu) \leq K/n$. By the Poincaré-Nash inequality, we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}\,(\phi u^*Qu) &\leq 2\sum_{i,j=1}^{N,n} d_j\mathbb{E}\left|\frac{\partial \phi u^*Qu}{\partial \bar{Y}_{ij}}\right|^2 \\ &\leq 4\sum_{i,j=1}^{N,n} d_j\mathbb{E}\left|\phi\frac{\partial u^*Qu}{\partial \bar{Y}_{ij}}\right|^2 + 4\sum_{i,j=1}^{N,n} d_j\mathbb{E}\left|u^*Qu\frac{\partial \phi}{\partial \bar{Y}_{ij}}\right|^2.\end{aligned}$$

Using the expression of $\partial Q_{pq}/\partial \bar{Y}_{ij}$ in Section 3.3, we have

$$\partial u^*Qu/\partial \bar{Y}_{ij} = -n^{-1}u^*Qy_j\,[Qu]_i,$$

hence

$$\sum_{i,j=1}^{N,n} d_j\mathbb{E}\left|\phi\frac{\partial u^*Qu}{\partial \bar{Y}_{ij}}\right|^2 = \frac{1}{n}\mathbb{E}\left[\phi^2\, u^*Q\frac{YDY^*}{n}Qu\, u^*Q^2u\right] \leq \frac{K}{n}$$

since the argument of the expectation is bounded for $\Re(z) \in \mathcal{I}$. From the first identity in Section 3.3, $\sum_{i,j} d_j\mathbb{E}\left|u^*Qu\partial\phi/\partial \bar{Y}_{ij}\right|^2 \leq K\sum_{i,j} d_j\mathbb{E}\left|\partial\phi/\partial \bar{Y}_{ij}\right|^2$ which

is bounded by K/n^2 by Lemma 4.2. It results that $\mathbb{V}\mathrm{ar}(\phi\, u^*Qu) \leq K/n$. Now, writing $\overset{\circ}{X} = X - \mathbb{E}X$,

$$\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\Big|^4 = \big(\mathbb{V}\mathrm{ar}(\phi\, u^*Qu)\big)^2 + \mathbb{V}\mathrm{ar}\Big(\Big(\overbrace{\phi\, u^*Qu}^{\circ}\Big)^2\Big) \leq K/n^2 + \mathbb{V}\mathrm{ar}\Big(\Big(\overbrace{\phi\, u^*Qu}^{\circ}\Big)^2\Big).$$

By the Poincaré-Nash inequality,

$$\mathbb{V}\mathrm{ar}\Big(\Big(\overbrace{\phi\, u^*Qu}^{\circ}\Big)^2\Big) \leq 2\sum_{i,j=1}^{N,n} d_j\mathbb{E}\Big|\partial\Big(\overbrace{\phi\, u^*Qu}^{\circ}\Big)^2 / \,\partial\bar{Y}_{ij}\Big|^2$$

$$\leq 16\sum_{i,j=1}^{N,n} d_j\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\ \phi\,\frac{\partial u^*Qu}{\partial\bar{Y}_{ij}}\Big|^2 + 16\sum_{i,j=1}^{N,n} d_j\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\ u^*Qu\,\frac{\partial\phi}{\partial\bar{Y}_{ij}}\Big|^2 := V_1 + V_2.$$

By developing the derivative in V_1 similarly to above, $V_1 \leq Kn^{-1}\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\Big|^2 \leq Kn^{-2}$. By the Cauchy-Schwarz inequality and Lemma 4.2,

$$V_2 \leq K\sum_{i,j=1}^{N,n} d_j\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\,\frac{\partial\phi}{\partial\bar{Y}_{ij}}\Big|^2 \leq \frac{K}{n^2}\Big(\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\Big|^4\Big)^{1/2}.$$

Writing $a_n = n^2\mathbb{E}\Big|\overbrace{\phi\, u^*Qu}^{\circ}\Big|^4$, we have shown that $\sqrt{a_n} \leq K/\sqrt{a_n} + K/n$. Assume that a_n is not bounded. Then there exists a sequence n_k of integers such that $a_{n_k} \to \infty$, which raises a contradiction. The first inequality in the statement of this lemma is shown. The other two inequalities can be shown similarly. □

Lemma 4.4. *The following holds true:*

$$1 - \mathbb{E}\phi_n \leq \frac{K_\ell}{n^\ell} \quad \textit{for any } \ell \in \mathbb{N}.$$

Proof. For $0 < \varepsilon_1 < \varepsilon'$ where ε' is defined in the construction of ψ, let φ be a smooth nonnegative function equal to zero on $\mathcal{S}_{\varepsilon_1}$ and to one on $\mathbb{R} - \mathcal{S}_{\varepsilon'}$. Then $1 - \phi_n \leq \big(\mathrm{Tr}(\varphi(n^{-1}YY^*))\big)^\ell$ for any $\ell \in \mathbb{N}$, and the result stems from Lemma 4.1. □

Lemma 4.5. *The following inequalities hold true (recall that $\Re(z) \notin \mathrm{supp}(\mu)$):*

$$\left|\mathbb{E}\left[\phi_n\,\mathrm{Tr}\,Q_n(z)\right] - Nm_n(z)\right| \leq \frac{K}{n}, \quad \textit{and} \quad \left|\mathrm{Tr}\,\widetilde{U}_n\left(\mathbb{E}[\phi_n\widetilde{Q}_n(z)] - \widetilde{T}_n(z)\right)\right| \leq \frac{K}{n}.$$

Proof. Let ε be defined in the construction of ψ. Choose a small $\varepsilon_1 > \varepsilon$ in such a way that $\mathcal{S}_{\varepsilon_1} \cap \mathcal{I} = \emptyset$. Let ζ be a $\mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$ nonnegative function equal to one on $\mathcal{S}_\varepsilon$ and to zero on $\mathbb{R} - \mathcal{S}_{\varepsilon_1}$, so that

$$\phi\frac{1}{n}\,\mathrm{Tr}\,Q = \phi\int\frac{\zeta(t)}{t-z}\theta_n(dt)\ .$$

Using this equality, and recalling that $\phi \in [0,1]$, we have

$$\left|\mathbb{E}\phi\frac{1}{n}\,\mathrm{Tr}\,Q - \mathbb{E}\int\frac{\zeta(t)}{t-z}\theta_n(dt)\right| \leq \mathbb{E}\left[(1-\phi)\left|\int\frac{\zeta(t)}{t-z}\theta_n(dt)\right|\right] \leq \frac{1-\mathbb{E}\phi}{\boldsymbol{d}(z,\mathcal{S}_{\varepsilon_1})} \leq \frac{K_\ell}{n^\ell}$$

for any $\ell \in \mathbb{N}$. Moreover, we have

$$\left|\mathbb{E}\int \frac{\zeta(t)}{t-z}\theta_n(dt) - m_n(z)\right| = \left|\mathbb{E}\int \frac{\zeta(t)}{t-z}\theta_n(dt) - \int \frac{\zeta(t)}{t-z}\mu_n(dt)\right| \leq \frac{K}{n^2}$$

by Proposition 3.3, and the first inequality is proved.
By performing a spectral factorization of $n^{-1}Y^*Y$, one can check that $n^{-1}\operatorname{Tr}\widetilde{U}\widetilde{Q}(z)$ is the Stieltjes Transform of a positive measure τ_n such that $\sup_n \tau_n(\mathbb{R}) < \infty$ and $\operatorname{supp}(\tau_n) \subset \operatorname{supp}(\theta_n) \cup \{0\}$. By Lemma 3.2, $n^{-1}\operatorname{Tr}\widetilde{U}\widetilde{T}(z)$ is the Stieltjes Transform of a positive measure γ_n such that $\sup_n \gamma_n(\mathbb{R}) < \infty$ and $\gamma_n(\mathcal{I}) = 0$ for all large n. With the help of the second inequality of Proposition 3.5, we have a result similar to that of Proposition 3.6, namely that $|\mathbb{E}\int \varphi d\tau_n - \int \varphi d\gamma_n| \leq K/n^2$ for any function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R},\mathbb{R})$. We can then prove the second inequality similarly to the first one. □

Lemma 4.6. *The following inequalities hold true:*

$$|\mathbb{E}[\phi_n\, u_n^* Q_n(z) v_n] - u_n^* v_n\, m_n(z)| \leq K/n,$$
$$\left|\mathbb{E}[\phi_n\, \tilde{u}_n^* \widetilde{Q}_n(z)\tilde{v}_n] - \tilde{u}_n^* \widetilde{T}_n(z)\tilde{v}_n\right| \leq K/n.$$

In [21], it is proven in a more general setting that $\left|\mathbb{E}u_n^* Q_n(z)u_n - \|u_n\|^2 m_n(z)\right| \leq P(|z|)R(|\Im(z)|^{-1})/\sqrt{n}$ for any $z \in \mathbb{C}_+$. Observing that $u_n^* Q_n(z) u_n$ and $\|u_n\|^2 m_n(z)$ are Stieltjes Transforms of positive measures, and mimicking the proof of the previous lemma, we can establish this lemma with the rate $\mathcal{O}(n^{-1/2})$, which is in fact enough for our purposes. However, in order to give a flavor of the derivations that will be carried out in the next section, we consider here another proof that uses the IP formula and the Poincaré-Nash inequality. To that end, we introduce new notations:

$$\beta(z) = \phi_n \frac{1}{n}\operatorname{Tr} Q_n(z),\ \alpha(z) = \mathbb{E}\beta(z),\ \hat{\beta}(z) = \beta(z) - \phi_n\alpha(z), \quad \text{and}$$
$$\tilde{\alpha}(z) = \frac{1}{n}\operatorname{Tr} D_n[-z(I_n + \alpha(z)D_n)]^{-1}.$$

Proof. We start with the first inequality. By the IP formula, we have

$$\mathbb{E}[Q_{pi}Y_{ij}\bar{Y}_{\ell j}\phi] = -\frac{d_j}{n}\mathbb{E}[[Qy_j]_p Q_{ii}\bar{Y}_{\ell j}\phi] + \delta(\ell - i)d_j\mathbb{E}[Q_{pi}\phi] + \frac{d_j}{n}\mathbb{E}[Q_{pi}\bar{Y}_{\ell j}[\operatorname{adj}(\psi)\psi' y_j]_i].$$

Taking the sum over i, we obtain

$$\mathbb{E}[[Qy_j]_p\bar{Y}_{\ell j}\phi] = -d_j\mathbb{E}[[Qy_j]_p\bar{Y}_{\ell j}\beta] + d_j\mathbb{E}[Q_{p\ell}\phi] + \frac{d_j}{n}\mathbb{E}[\bar{Y}_{\ell j}[Q\operatorname{adj}(\psi)\psi' y_j]_p].$$

Writing $\beta = \hat{\beta} + \phi\alpha$, we get

$$\mathbb{E}[[Qy_j]_p\bar{Y}_{\ell j}\phi] = \frac{d_j}{1+\alpha d_j}\mathbb{E}[Q_{p\ell}\phi] - \frac{d_j}{1+\alpha d_j}\mathbb{E}[[Qy_j]_p\bar{Y}_{\ell j}\hat{\beta}]$$
$$+ \frac{d_j}{n(1+\alpha d_j)}\mathbb{E}[\bar{Y}_{\ell j}[Q\operatorname{adj}(\psi)\psi' y_j]_p].$$

Taking the sum over j, we obtain

$$\mathbb{E}\left[\left[Q\frac{YY^*}{n}\right]_{p\ell}\phi\right] = -z\tilde{\alpha}\mathbb{E}[Q_{p\ell}\phi] - \mathbb{E}\left[\hat{\beta}\left[Q\frac{YD(I+\alpha D)^{-1}Y^*}{n}\right]_{p\ell}\right] + \frac{1}{n}\mathbb{E}\left[\left[Q\,\mathrm{adj}(\psi)\psi'\frac{YD(I+\alpha D)^{-1}Y^*}{n}\right]_{p\ell}\right].$$

We now use the identity $zQ = n^{-1}QYY^* - I$, which results in

$$z\mathbb{E}[Q_{p\ell}\phi] = \mathbb{E}\left[\left[Q\frac{YY^*}{n}\right]_{p\ell}\phi\right] - \delta(p-\ell)\mathbb{E}[\phi],$$

$$\mathbb{E}[Q_{p\ell}\phi] = \frac{\delta(p-\ell)}{-z(1+\tilde{\alpha})}\mathbb{E}[\phi] + \frac{\text{2nd and 3rd terms of next to last equation}}{z(1+\tilde{\alpha})}.$$

Multiplying each side by $[u^*]_p[v]_\ell$ and taking the sum over p and ℓ, we finally obtain

$$\mathbb{E}[u^*Qv\phi] = \mathbb{E}[\phi]\frac{u^*v}{-z(1+\tilde{\alpha})} - [-z(1+\tilde{\alpha})]^{-1}\mathbb{E}\left[\hat{\beta}u^*Q\frac{YD(I+\alpha D)^{-1}Y^*}{n}v\right] + \frac{1}{n}[-z(1+\tilde{\alpha})]^{-1}\mathbb{E}\left[u^*Q\,\mathrm{adj}(\psi)\psi'\frac{YD(I+\alpha D)^{-1}Y^*}{n}v\right]. \quad (9)$$

Let us evaluate the three terms at the right hand side of this equality. From Lemma 4.5, we have $\alpha = c_n m_n + \mathcal{O}(n^{-2})$. Using in addition the bound (8), we obtain $\tilde{\alpha} = n^{-1}\operatorname{Tr}(D(-z(I + c_n m_n D + (\alpha - c_n m_n)D)^{-1}) = n^{-1}\operatorname{Tr} D\widetilde{T} + \mathcal{O}(n^{-2})$. Since $m_n(z) = (-z(1+n^{-1}\operatorname{Tr} D\widetilde{T}(z)))^{-1}$, we obtain that $(-z(1+\tilde{\alpha}))^{-1} = m_n(z) + \mathcal{O}(n^{-2})$. Using in addition Lemma 4.4, we obtain that the first right hand side term of (9) is $u^*v\,m_n(z) + \mathcal{O}(n^{-2})$. Due to the presence of ϕ in the expression of $\hat{\beta}$, the second term is bounded by $K\mathbb{E}|\hat{\beta}|$. Moreover, $\hat{\beta} = n^{-1}\phi\operatorname{Tr} Q - n^{-1}\mathbb{E}[\phi\operatorname{Tr} Q] + (1-\phi)n^{-1}\mathbb{E}[\phi\operatorname{Tr} Q]$. By Lemmas 4.4 and 4.3, $\mathbb{E}|\hat{\beta}| = \mathcal{O}(n^{-1})$. The third term can be shown to be bounded by $Kn^{-1}\mathbb{E}\operatorname{Tr}\varphi(n^{-1}YY^*) = \mathcal{O}(n^{-2})$ where φ is as in the statement of Lemma 4.1. This proves the first inequality in the statement of the lemma.

The second result in the statement of the lemma is proven similarly. The proof requires the second inequality of Lemma 4.5. □

The proof of the following lemma can be done along the same lines and will be omitted:

Lemma 4.7. *The following inequalities hold true:*

$$\left|\mathbb{E}\phi_n u_n^* Y_n \widetilde{Q}_n(z)\tilde{v}_n\right| \le K/\sqrt{n}$$

$$\mathbb{E}\left|\phi_n u_n^* Y_n \widetilde{Q}_n(z)\tilde{v}_n\right|^4 \le K.$$

We now prove Theorem 2.2-(1).

Proof of Theorem 2.2-(1). Our first task is to establish the properties of $H_*(z)$ given in the statement of Theorem 2.2-(1). By Lemma 3.1 and the fact that $\operatorname{supp}(\Lambda_*) \subset \operatorname{supp}(\nu)$, the function $H_*(z)$ can be analytically extended to (a,b). The comments preceding Theorem 2.2 show that any $H_*(z)$ is the Stieltjes Transform of a matrix-valued nonnegative measure Γ. Since $H_*(z)$ is analytic on (a,b), it

is increasing on this interval in the order of Hermitian matrices, and the properties of this function given in the statement of Theorem 2.2-(1) are established.

We now prove the convergence stated in Theorem 2.2-(1), formalizing the argument introduced in Section 3.1. In the remainder, we restrict ourselves to the probability one set where $n^{-1}YY^*$ has no eigenvalues for large n in a large enough closed interval in (a,b). Given a compact interval $\mathcal{I} \subset (a,b)$, let $D_{\mathcal{I}}^{\circ}$ be the open disk with diameter $\mathcal{I}$. Observe that the functions $\widehat{S}_n(x)$ and $S_n(x)$ introduced in Section 3.1 can be extended analytically to a neighborhood of $D_{\mathcal{I}}$, and the determinants of these functions do not cancel outside the real axis. Let $\widehat{L}_n = \sharp\{i \,:\, \hat{\lambda}_i^n \in D_{\mathcal{I}}^{\circ}\}$, $L_n = \sharp\{$ zeros of $\det S_n(z)$ in $D_{\mathcal{I}}^{\circ}\}$, and $L = \sharp\{i \,:\, \rho_i \in D_{\mathcal{I}}^{\circ}\}$. We need to prove that $\widehat{L}_n = L$ with probability one for n large. By the argument of [8, 9], $\widehat{L}_n$ is equal to the number of zeros of $\det \widehat{S}_n(z)$ in $\mathcal{I}^{\circ}$. By the well known argument principle for holomorphic functions,

$$\widehat{L}_n = \frac{1}{2\imath\pi} \oint_{\partial D_{\mathcal{I}}^{\circ}} \frac{(\det \widehat{S}_n(z))'}{\det \widehat{S}_n(z)}\, dz,$$

$$L_n = \frac{1}{2\imath\pi} \oint_{\partial D_{\mathcal{I}}^{\circ}} \frac{(\det S_n(z))'}{\det S_n(z)}\, dz = \frac{1}{2\imath\pi} \oint_{\partial D_{\mathcal{I}}^{\circ}} \frac{(\det(H_n(z)+I_r))'}{\det(H_n(z)+I_r)}\, dz \quad \text{and}$$

$$L = \frac{1}{2\imath\pi} \oint_{\partial D_{\mathcal{I}}^{\circ}} \frac{(\det(H_*(z)+I_r))'}{\det(H_*(z)+I_r)}\, dz$$

where $\partial D_{\mathcal{I}}^{\circ}$ is seen as a positively oriented contour.
For any $1 \le k, \ell \le r$, let $h_{n,k,\ell}(z) = [U_n^*(Q_n(z) - m_n(z)I_N)U_n]_{k,\ell}$. Let V be a small neighborhood of $D_{\mathcal{I}}$, the closure of $D_{\mathcal{I}}^{\circ}$. Let z_m be a sequence of complex numbers in V having an accumulation point in V. By Lemmas 4.1, 4.3 and 4.6 and the Borel Cantelli lemma, $h_{n,k,\ell}(z_m) \overset{\text{as}}{\to} 0$ as $n \to \infty$ for every m. Moreover, for n large, the $h_{n,k,\ell}$ are uniformly bounded on any compact subset of V. By the normal family theorem, every n-sequence of $h_{n,k,\ell}$ contains a further subsequence which converges uniformly on the compact subsets of V to a holomorphic function h_*. Since $h_*(z_m) = 0$ for every m, we obtain that almost surely, $h_{n,k,\ell}$ converges uniformly to zero on the compact subsets of V, and the same can be said about $\|U_n^*(Q_n(z) - m_n(z)I_N)U_n\|$. Using in addition Lemmas 3.1 and 4.7 we obtain the same result for $\|R_n^*(\widetilde{Q}_n(z) - \widetilde{T}_n(z))R_n\|$ and $n^{-1/2}\|U_n^* Y_n \widetilde{Q}_n(z) R_n\|$.

Since $\det X$ is a polynomial in the elements of matrix X, $\det \widehat{S}_n(z) - \det S_n(z)$ converges almost surely to zero on $\partial D_{\mathcal{I}}^{\circ}$, and this convergence is uniform. By analyticity, the same can be said about the derivatives of these quantities. Moreover, $\det S_n(z)$ converges to $(-1)^r \det(H_*(z) + I_r)$ (which is the same for all accumulation points Λ_*) uniformly on $\partial D_{\mathcal{I}}^{\circ}$. Similarly, $(\det S_n(z))'$ converges to $(-1)^r(\det(H_*(z)+I_r))'$ uniformly on $\partial D_{\mathcal{I}}^{\circ}$. Furthermore, by construction of the interval $\mathcal{I}$, we have $\inf_{z\in\partial D_{\mathcal{I}}^{\circ}} |\det(H_*(z)+I_r)| > 0$ which implies that $\liminf_n \inf_{z\in\partial D_{\mathcal{I}}^{\circ}} |\det S_n(z)| > 0$. It follows that $\widehat{L}_n = L_n$ and $L_n = L$ for n large enough. This concludes the proof of Theorem 2.2-(1).

Proofs of Theorems 2.2-(2) and 2.3. We start with the following lemma:

Lemma 4.8. *Let $A = \inf(\mathrm{supp}(\mu) - \{0\})$ and let (a,b) be a component of $\mathrm{supp}(\mu)^c$. Then the following facts hold true:*

(i) *If $b \le A$, then $\mathbf{m}(x)(1+c\mathbf{m}(x)t)^{-1} > 0$ for all $x \in (a,b)$ and all $t \in \mathrm{supp}(\nu)$.*

(ii) *Alternatively, if $a > A$, then there exists a Borel set $E \subset \mathbb{R}_+$ such that $\nu(\partial E) = 0$ and*

$$q(x) = \int_E \frac{\mathbf{m}(x)}{1 + c\mathbf{m}(x)t}\nu(dt) < 0$$

for all $x \in (a, b)$.

Proof. The proof is based on the results of Section 3.5. To have an illustration of some of the proof arguments, the reader may refer to Figures 1 and 2 which provide typical plots of $x(\mathbf{m})$ for $c < 1$ and $c > 1$ respectively.
We start by fixing a point x_0 in (a, b), we write $\mathbf{m}_0 = \mathbf{m}(x_0)$ and we choose in the remainder of the proof $E = [0, -(c\mathbf{m}_0)^{-1}]$ with the convention $E = \emptyset$ when $\mathbf{m}_0 > 0$. We already assumed that $\nu(\{0\}) = 0$ in Section 3.5. Since $-(c\mathbf{m}_0)^{-1} \in \text{supp}(\nu)^c$ by Proposition 3.2, $\nu(\{-(c\mathbf{m}_0)^{-1}\}) = 0$. Hence $\nu(\partial E) = 0$. To prove the lemma, we shall show that

$$\nu(E) > 0 \Leftrightarrow a > A. \tag{10}$$

To see why (10) proves the lemma, consider first $a > A$. Then $\mathbf{m}_0 < 0$ since $\nu(E) > 0$. Moreover $1 + c\mathbf{m}_0 t \geq 0$ for any $t \in E$. It results that $q(x_0) < 0$. Consider now another point $x_1 \in (a, b)$, and let $\mathbf{m}_1 = \mathbf{m}(x_1)$ and $E_1 = [0, -(c\mathbf{m}_1)^{-1}]$. By the same argument as for x_0, we get that $\int_{E_1} \mathbf{m}(x)(1 + c\mathbf{m}(x)t)^{-1}\nu(dt) < 0$. But Proposition 3.2 shows that the closed interval between $\mathbf{m}_0$ and $\mathbf{m}_1$ belongs to the set B. It results that $\nu(E \triangle E_1) = 0$ where $E \triangle E_1$ is the symmetric difference between E and E_1. Hence $q(x_1) < 0$ and (ii) is true.
Assume now that $b \leq A$. If $\mathbf{m}_0 > 0$, then for all $x_1 \in (a, b)$, $\mathbf{m}_1 = \mathbf{m}(x_1) > 0$ since the segment between $\mathbf{m}_0$ and $\mathbf{m}_1$ belongs to B, and (i) is true. Assume that $\mathbf{m}_0 < 0$. Then since $\nu(E) = 0$, for any $t \in \text{supp}(\nu)$, $t > -(c\mathbf{m}_0)^{-1}$, hence $\mathbf{m}_0(1 + c\mathbf{m}_0 t)^{-1} > 0$. If we take another point $x_1 \in (a, b)$, then the associated set E_1 will also satisfy $\nu(E_1) = 0$ since $\nu(E \triangle E_1) = 0$. Hence we also have $\mathbf{m}_1(1 + c\mathbf{m}_1 t)^{-1} > 0$ for any $t \in \text{supp}(\nu)$, which proves (i).
We now prove (10) in the $\Leftarrow$ direction, showing that $x_0 > A \Rightarrow \nu(E) > 0$. Since $\mathbf{m}(z)$ is the Stieltjes Transform of a probability measure supported by $\mathbb{R}_+$, the function $\mathbf{m}(x)$ decreases to zero as $x \to -\infty$. Furthermore, $(0, \infty)$ belongs to B. Hence, by Proposition 3.2, $x_0 > A \Rightarrow \mathbf{m}_0 < 0$. Assume that $\nu(E) = 0$. Then $(-\infty, \mathbf{m}_0] \subset B$. Since $t > -(c\mathbf{m}_0)^{-1}$ in the integral in (7), $x(\mathbf{m}) \to 0$ as $\mathbf{m} \to -\infty$ by the dominated convergence theorem. By Propositions 3.1, 3.2 and 3.3, $x(\mathbf{m})$ should be increasing from 0 to x_0 on $(-\infty, \mathbf{m}_0]$. This contradicts $x_0 > A$.
We now prove (10) in the $\Rightarrow$ direction. To that end, we consider in turn the cases $c < 1$, $c > 1$ and $c = 1$.
Assume $c < 1$. Since $\mathbf{m}(z)$ is the Stieltjes Transform of a probability measure supported by $\mathbb{R}_+$, the function $\mathbf{m}(x)$ decreases to zero as $x \to -\infty$. Hence $x(\mathbf{m}) \to -\infty$ as $\mathbf{m} \to 0^+$. From (7) we notice that $\mathbf{m}x(\mathbf{m}) \to (1 - c)/c > 0$ as $\mathbf{m} \to \infty$, hence $x(\mathbf{m})$ reaches a positive maximum on $(0, \infty)$. By Propositions 3.2 and 3.3, this maximum is A, and we have $x < A \Rightarrow \mathbf{m}(x) > 0$. Therefore, $x_0 < A \Rightarrow \nu(E) = 0$.
Consider now the case $c > 1$. We shall also show that $x_0 < A \Rightarrow \nu(E) = 0$. By Proposition 3.4, the measure μ has a Dirac at zero with weight $1 - c^{-1}$. Hence, either $x_0 < 0$, or $A > 0$ and $0 < x_0 < A$. Since $\mathbf{m}(z)$ is the Stieltjes Transform of a probability measure supported by $\mathbb{R}_+$, it holds that $x < 0 \Rightarrow \mathbf{m}(x) > 0$. Hence, $\nu(E) = 0$ when $x_0 < 0$. We now consider the second case. Since $(0, x_0] \subset \text{supp}(\mu)^c$, the image of this interval by $\mathbf{m}$ belongs to B. By Proposition 3.4, $\lim_{x \to 0^+} \mathbf{m}(x) =$

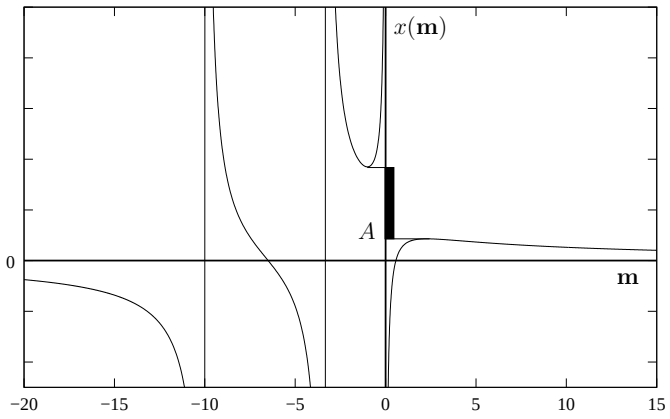


FIGURE 1. Plot of $x(\mathbf{m})$ for $c = 0.1$ and $\nu = 0.5(\boldsymbol{\delta}_1 + \boldsymbol{\delta}_3)$. The thick segment represents $\mathrm{supp}(\mu)$.

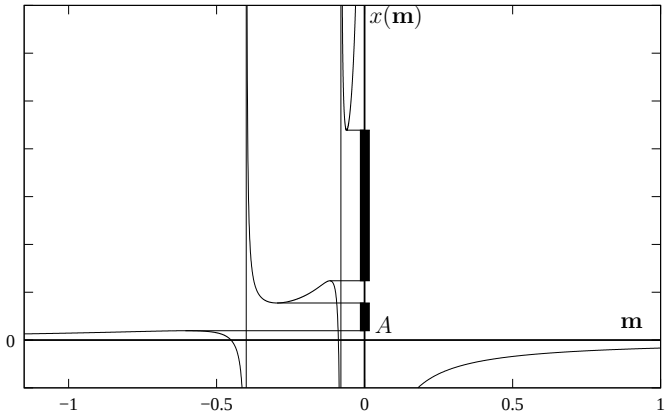


FIGURE 2. Plot of $x(\mathbf{m})$ for $c = 5$ and $\nu = 0.5(\boldsymbol{\delta}_{1/2} + \boldsymbol{\delta}_{5/2})$. The thick segments represent $\mathrm{supp}(\mu)$.

$-\infty$. Hence this image is $(-\infty, \mathbf{m}_0]$. This implies that $\nu(E) = 0$.
We finally consider the case $c = 1$. We show here that $A = 0$, which will result in $x_0 < A \Rightarrow \mathbf{m}_0 > 0 \Rightarrow \nu(E) = 0$ as above. Assume $A > 0$ and let $x_1 \in (0, A)$. By Proposition 3.4, $\mu(\{0\}) = 0$ hence $\mathbf{m}(x_1) = \int (t - x_1)^{-1} \mu(dt) > 0$. But from (7), we easily show that $x(\mathbf{m})$ increases from $-\infty$ to 0 as $\mathbf{m}$ increases from 0 to ∞, which raises a contradiction. This concludes the proof of Lemma 4.8. □

This lemma shows that for any $x < \inf(\mathrm{supp}(\mu) - \{0\})$, $H_*(x) \geq 0$, hence $\mathcal{D}(x) > 0$ for those x. This proves Theorem 2.2-(2).
Turning to Theorem 2.3, let E be a Borel set associated to (a, b) by Lemma 4.8-(ii) and let $q(x)$ be the function defined in the statement of that lemma. The argument preceding Theorem 2.2 shows that the extension of $q(x)$ to $\mathbb{C}_+$ is the Stieltjes Transform of a positive measure. It results that $q(x)$ is negative and increasing on (a, b). Let $\boldsymbol{\Omega} = \mathrm{diag}(\omega_1^2, \ldots, \omega_r^2)$ where $\omega_k^2 = -1/q(\rho_k)$. Then it is clear that

the function $\mathcal{D}(x) = \det(q(x)\mathbf{\Omega} + I_r)$ has r roots in (a, b) which coincide with the ρ_k. Theorem 2.3 will be proven if we find a sequence of matrices P_n for which $H_*(z) = q(z)\mathbf{\Omega}$, *i.e.*, $\Lambda_*(dt) = 1_E(t)\,\mathbf{\Omega}\,\nu(dt)$.
Rearrange the elements of D_n in such a way that all the d_j^n which belong to E are in the top left corner of this matrix. Let $M_n = [M_{ij}^n]$ be a random $\lfloor n\nu(E)\rfloor \times r$ matrix with iid elements such that $\sqrt{n}M_{11}^n$ has mean zero and variance one. Let Z_n be the $n \times r$ matrix obtained by adding $n - \lfloor n\nu(E)\rfloor$ rows of zeros below M_n. Then the law of large numbers shows in conjunction with a normal family theorem argument that there is a set of probability one over which $zm_n(z)Z_n^*\widetilde{T}_n(z)Z_n$ converges to $q(z)I_r$ uniformly on the compact subsets of (a, b). Consequently, there exists a sequence of deterministic matrices B_n such that $zm_n(z)B_n^*\widetilde{T}_n(z)B_n \to q(z)I_r$ uniformly on these compact subsets. Matrix $P_n = A_nB_n^*$ with $A_n = \begin{bmatrix} \mathbf{\Omega}^{1/2} \\ 0_{(N-r)\times r} \end{bmatrix}$ satisfies the required property. Theorem 2.3 is proven.

Proof of Corollary 2.1. Observe from Proposition 2.1 that $\int(1+c\mathbf{m}(z)t)^{-1}\nu(dt) = -z\lim(n^{-1}\operatorname{Tr}\widetilde{T}_n(z)) = -cz\mathbf{m}(z) + 1 - c$. Consequently, in this particular case, $H_*(x)$ is unitarily equivalent to $-\mathbf{m}(x)\,(cx\mathbf{m}(x) - 1 + c)\,\mathbf{\Omega}$ on (a, b).

5. Proof of the second order result

We start by briefly showing Proposition 2.2.

Proof of Proposition 2.2. For any $i = 1, \ldots, p$, it is clear that $\mathbf{m}(\rho_i)^2 > 0$ and $\mathbf{m}'(\rho_i) > 0$. An immediate calculus then gives $\mathbf{m}'(\rho_i)\mathbf{\Delta}(\rho_i) = \mathbf{m}^2(\rho_i)$ which shows that $\mathbf{\Delta}(\rho_i) > 0$.
To prove the second fact, we shall establish more generally that $\limsup \sqrt{n}\|H_n(\rho_i) + g(\rho_i)\mathbf{\Omega}\| < \infty$. Invoking Equation (3) and its analogue $m_n(z) = (-z + \int t(1 + c_nm_n(z)t)^{-1}\nu_n(dt))^{-1}$, taking the difference and doing some straightforward derivations, we get that $(m_n(\rho_i) - \mathbf{m}(\rho_i))(\mathbf{\Delta}(\rho_i) + \varepsilon_1) = \varepsilon_2$ where $\varepsilon_1 \to 0$ and where $|\varepsilon_2| \leq K/\sqrt{n}$ thanks to the first two items of Assumption 7. Hence $|m_n(\rho_i) - \mathbf{m}(\rho_i)| \leq K/\sqrt{n}$. Now we have

$$H_n(\rho_i) + g(\rho_i)\mathbf{\Omega} = \int \left(\frac{m_n(\rho_i)}{1 + c_nm_n(\rho_i)t} - \frac{\mathbf{m}(\rho_i)}{1 + c\mathbf{m}(\rho_i)t} \right) \Lambda_n(dt) + \int \frac{\mathbf{m}(\rho_i)}{1 + c\mathbf{m}(\rho_i)t}\Lambda_n(dt) - \int \frac{\mathbf{m}(\rho_i)}{1 + c\mathbf{m}(\rho_i)t}\nu(dt) \times \mathbf{\Omega}.$$

which shows thanks to Assumption 7 that $\limsup \sqrt{n}\|H_n(\rho_i) + g(\rho_i)\mathbf{\Omega}\| < \infty$.

Proof of Theorem 2.4. In all the remainder of this section, we shall work on a sequence of factorizations $P_n = U_nR_n^*$ such that Λ_n satisfies the third item of Assumption 7. We also write $U_n = [U_{1,n} \cdots U_{t,n}]$ and $R_n = [R_{1,n} \cdots R_{t,n}]$ where $U_{i,n} \in \mathbb{C}^{N\times j_i}$ and $R_{i,n} \in \mathbb{C}^{n\times j_i}$.

We now enter the core of the proof Theorem 2.4. The following preliminary lemmas are proven in the appendix:

Lemma 5.1. *Let s be a fixed integer, and let $Z_N = [Z_{ij}]$ be a $N \times s$ complex matrix with iid elements with independent $\mathcal{N}(0, 1/2)$ real and imaginary parts. Let $\Upsilon_N = [\Upsilon_{ij}]$ be a deterministic Hermitian $N \times N$ matrix such that $\operatorname{Tr}\Upsilon_N = 0$, and let $F_N = [F_{ij}]$ be a complex deterministic $N \times s$ matrix. Assume that $F_N^*F_N \to \boldsymbol{\varsigma}^2I_s$,*

that $\limsup_N \|\Upsilon_N\| < \infty$*, and that* $N^{-1}\operatorname{Tr}\Upsilon_N^2 \to \boldsymbol{\sigma}^2$ *as* $N \to \infty$*. Let* M *be a* $s \times s$ *complex matrix with iid elements with independent* $\mathcal{N}(0, 1/2)$ *real and imaginary parts, and let* G *be a* $s \times s$ *GUE matrix independent of* M*. Then*

$$\left(N^{-1/2} Z_N^* \Upsilon_N Z_N, Z_N^* F_N\right) \xrightarrow[N\to\infty]{\mathcal{D}} (\boldsymbol{\sigma} G, \varsigma M)\,.$$

Lemma 5.2. *For* $x \in \operatorname{supp}(\mu)^c$*,*

$$\mathbb{E}\left[\phi_n \tilde{u}_n^* \widetilde{Q}_n(x)(n^{-1}Y_n^*Y_n)\widetilde{Q}_n(x)\tilde{u}_n\right] = c_n \frac{x^2 m_n(x)^2 \tilde{u}_n^* D_n \widetilde{T}_n^2(x)\tilde{u}_n}{1 - c_n x^2 m_n(x)^2 \frac{1}{n}\operatorname{Tr} D_n^2 \widetilde{T}_n^2(x)} + \mathcal{O}(n^{-1})$$

and

$$\mathbb{V}\mathrm{ar}\left(\phi_n \tilde{u}_n^* \widetilde{Q}_n(x)(n^{-1}Y_n^*Y_n)\widetilde{Q}_n(x)\tilde{u}_n\right) \le \frac{K}{n}.$$

Lemma 5.3. *For* $i = 1, \dots, p$*, let* A_i *be a deterministic Hermitian* $j_i \times j_i$ *matrix independent of* n*, where* p *and the* j_i *are as in the statement of Theorem 2.4. For* $i = 1, \dots, p$*, let* $M_{i,n}$ *a* $n \times j_i$ *matrix such that* $\sup_n \|M_{i,n}\| < \infty$*. Then for any* $t \in \mathbb{R}$*,*

$$\mathbb{E}\Big[\exp\Big(\imath\sqrt{n}t \sum_{i=1}^p \operatorname{Tr} A_i M_{i,n}^*(\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i))M_{i,n}\Big)\Big] = \exp(-t^2\tilde{\sigma}_n^2/2) + \mathcal{O}(n^{-1/2})$$

where

$$\tilde{\sigma}_n^2 = \sum_{i,k=1}^p c_n \rho_i \rho_k m_n(\rho_i) m_n(\rho_k) \times \frac{\operatorname{Tr} A_i M_{i,n}^* \widetilde{T}_n(\rho_i) D_n \widetilde{T}_n(\rho_k) M_{k,n} A_k M_{k,n}^* \widetilde{T}_n(\rho_k) D_n \widetilde{T}_n(\rho_i) M_{i,n}}{1 - c_n \rho_i \rho_k m_n(\rho_i) m_n(\rho_k) \frac{1}{n}\operatorname{Tr} D_n \widetilde{T}_n(\rho_i) D_n \widetilde{T}_n(\rho_k)}.$$

Replacing the $M_{i,n}$ with the blocks $R_{i,n}$ of R_n in the statement of Lemma 5.3 and observing that

$$R_n^* \widetilde{T}_n(\rho_i) D_n \widetilde{T}_n(\rho_k) R_n = \int \frac{t}{\rho_i \rho_k (1 + c_n m_n(\rho_i) t)(1 + c_n m_n(\rho_k) t)} \Lambda_n(dt),$$

we obtain from the third item of Assumption 7 that $\tilde{\sigma}_n^2 \to \sum_{i=1}^p \tilde{\boldsymbol{\sigma}}_i^2 \operatorname{Tr} A_i^2$ where

$$\tilde{\boldsymbol{\sigma}}_i^2 = \frac{c\omega_i^4}{\rho_i^2 \boldsymbol{\Delta}(\rho_i)} \Big(\int \frac{\mathbf{m}(\rho_i) t}{(1 + c\mathbf{m}(\rho_i)t)^2} \nu(dt)\Big)^2.$$

Invoking the Cramer-Wold device, this means that the p-uple of random matrices

$$\sqrt{n}\left(R_{i,n}^*(\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i))R_{i,n}\right)_{i=1}^p$$

converges in distribution towards $(\tilde{\boldsymbol{\sigma}}_i \widetilde{G}_i)_{i=1}^p$ where $\widetilde{G}_1, \dots, \widetilde{G}_p$ are independent GUE matrices with $\widetilde{G}_i \in \mathbb{C}^{j_i \times j_i}$.

Lemmas 5.1–5.3 lead to the following result which plays a central role in the proof of Theorem 2.4:

Lemma 5.4. *Consider the 3p-uple of random matrices*

$$L_n = \sqrt{n}\times$$
$$\left(\frac{U_{i,n}^* Y_n \widetilde{Q}_n(\rho_i) R_{i,n}}{\sqrt{n}}, U_{i,n}^*(Q_n(\rho_i) - m_n(\rho_i) I_N) U_{i,n},\ R_{i,n}^*(\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n}\right)_{i=1}^p .$$

Define the following quantities

$$\boldsymbol{\varsigma}_i^2 = \frac{\omega_i^2}{\boldsymbol{\Delta}(\rho_i)} \int \frac{\mathbf{m}^2(\rho_i) t}{(1 + c\mathbf{m}(\rho_i)t)^2} \nu(dt)$$
$$\boldsymbol{\sigma}_i^2 = \frac{1}{\boldsymbol{\Delta}(\rho_i)} \int \frac{\mathbf{m}^4(\rho_i) t^2}{(1 + c\mathbf{m}(\rho_i)t)^2} \nu(dt)$$
$$\tilde{\boldsymbol{\sigma}}_i^2 = \frac{c\omega_i^4}{\rho_i^2 \boldsymbol{\Delta}(\rho_i)} \left(\int \frac{\mathbf{m}(\rho_i) t}{(1 + c\mathbf{m}(\rho_i)t)^2} \nu(dt) \right)^2 .$$

Let $M_1, \ldots, M_p$ be random matrices such that $M_i \in \mathbb{C}^{j_i \times j_i}$ and has independent elements with independent $\mathcal{N}(0, 1/2)$ real and imaginary parts. Let $G_1, \widetilde{G}_1, \ldots, G_p, \widetilde{G}_p$ be GUE matrices such that $G_i, \widetilde{G}_i \in \mathbb{C}^{j_i \times j_i}$. Assume in addition that M_1, G_1, $\widetilde{G}_1$, …, M_p, G_p, $\widetilde{G}_p$ are independent. Then

$$L_n \xrightarrow[n\to\infty]{\mathcal{D}} \left(\boldsymbol{\varsigma}_i M_i, \boldsymbol{\sigma}_i G_i, \tilde{\boldsymbol{\sigma}}_i \widetilde{G}_i\right)_{i=1}^p .$$

Proof. Let $\alpha_n(\rho) = N^{-1} \operatorname{Tr} Q_n(\rho)$. By Lemmas 4.3 and 4.5, $\sqrt{n}(\alpha_n(\rho_i) - m_n(\rho_i)) \xrightarrow{\mathcal{P}} 0$. Therefore, we can replace the $m_n(\rho_i)$ in the expression of L_n by $\alpha_n(\rho_i)$, as we shall do in the rest of the proof.

Write $s = j_1 + \cdots + j_p$ and let Z_n be a $N \times s$ complex matrix with iid elements with independent $\mathcal{N}(0, 1/2)$ real and imaginary parts. Assume that Z_n and X_n are independent. Write $Z_n = \begin{bmatrix} Z_{1,n} \cdots Z_{p,n} \end{bmatrix}$ where the block $Z_{i,n}$ is $N \times j_i$. Let $n^{-1/2} X_n = W_n \Delta_n \widetilde{W}_n^*$ be a singular value decomposition of $n^{-1/2} X_n$. By assumption 2, the square matrices W_n and $\widetilde{W}_n$ are Haar distributed over their respective unitary groups, and moreover, W_n, Δ_n and $\widetilde{W}_n$ are independent. Let

$$\overline{L}_n =$$
$$\sqrt{n}\left(\frac{U_n^* Y_n \widetilde{Q}_n(\rho_i) R_{i,n}}{\sqrt{n}}, U_n^*(Q_n(\rho_i) - \alpha_n(\rho_i) I_N) U_n,\ R_{i,n}^*(\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n}\right)_{i=1}^p .$$

We have

$$\overline{L}_n \overset{\mathcal{L}}{=} \Big(\sqrt{N}(Z_n^* Z_n)^{-1/2} Z_n^* F_{i,n}, N^{1/2}(Z_n^* Z_n)^{-1/2} Z_n^* \Upsilon_{i,n} Z_n (Z_n^* Z_n)^{-1/2},$$
$$\sqrt{n} R_{i,n}^*(\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n} \Big)_{i=1}^p$$

where $F_{i,n} = c_n^{-1/2} \Delta_n \widetilde{W}_n^* D_n^{1/2} \widetilde{Q}_n(\rho_i) R_{i,n}$ and $\Upsilon_{i,n} = c_n^{-1/2}((\Delta_n \widetilde{W}_n^* D_n \widetilde{W}_n \Delta_n - \rho_i)^{-1} - \alpha_n(\rho_i) I_N)$. We shall now show that the term $\sqrt{N}(Z_n^* Z_n)^{-1/2} Z_n^* F_{i,n}$ can be replaced with $Z_n^* F_{i,n}$. By the law of large numbers, we have $N^{-1} Z_n^* Z_n \xrightarrow{\text{as}} I_s$. By the independence of Z_n and $(\Delta_n, \widetilde{W}_n)$, we have $\mathbb{E}[\operatorname{Tr} Z_n^* F_{i,n} F_{i,n}^* Z_n \,|\, (\Delta_n \widetilde{W}_n)] = s c_n^{-1} \operatorname{Tr} R_{i,n}^* \widetilde{Q}_n(\rho_i)(n^{-1} Y_n^* Y_n) \widetilde{Q}_n(\rho_i) R_{i,n}$ whose limit superior is bounded with probability one. Hence $Z_n^* F_{i,n}$ is tight, proving that the replacement can be done.

By deriving the variances of the elements of $N^{-1/2} Z_n^* \Upsilon_{i,n} Z_n$ with respect to the

law of Z_n, and by recalling that $\limsup_n \|\Upsilon_{i,n}\|$ is bounded with probability one, we obtain that these elements are also tight. It results that we can replace L_n with

$$\underline{L}_n = \left(Z_{i,n}^* F_{i,n},\ \frac{Z_{i,n}^* \Upsilon_{i,n} Z_{i,n}}{\sqrt{N}}, \sqrt{n} R_{i,n}^* (\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n} \right)_{i=1}^{p}.$$

For $i = 1, \ldots, p$, let A_i and B_i be deterministic Hermitian $j_i \times j_i$ matrices and let C_i be deterministic complex $j_i \times j_i$ matrices, all independent of n. The lemma will be established if we prove that

$$\begin{aligned} &\mathbb{E}\left\{ \exp\Big(\imath\sqrt{n} t \sum_{i=1}^{p} \operatorname{Tr} A_i R_{i,n}^* (\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n} \Big) \right. \\ &\quad \left. \times \mathbb{E}\Big[\exp\Big(\imath t \sum_{i=1}^{p} N^{-1/2} \operatorname{Tr} B_i Z_{i,n}^* \Upsilon_{i,n} Z_{i,n} + \Re(\operatorname{Tr} C_i Z_{i,n}^* F_{i,n}) \Big) \,\Big|\, (\Delta_n, \widetilde{W}_n) \Big] \right\} \\ &\xrightarrow[n\to\infty]{} \prod_{i=1}^{p} \exp(-t^2(\tilde{\boldsymbol{\sigma}}_i^2 \operatorname{Tr} A_i^2 + \boldsymbol{\sigma}_i^2 \operatorname{Tr} B_i^2 + \frac{1}{2}\boldsymbol{\varsigma}_i^2 \operatorname{Tr} C_i C_i^*)/2). \end{aligned} \tag{11}$$

In addition to the boundedness of $\|\Upsilon_{i,n}\|$ w.p. one, we have $\operatorname{Tr} \Upsilon_{i,n} = 0$, and

$$\begin{aligned} \frac{1}{N} \operatorname{Tr} \Upsilon_{i,n}^2 &= \frac{1}{c_n N} \sum_{\ell=1}^{N} \left((\lambda_\ell^n - \rho_i)^{-1} - \alpha_n(\rho_i) \right)^2 \\ &\xrightarrow[n\to\infty]{\text{a.s.}} c^{-1}(\mathbf{m}'(\rho_i) - \mathbf{m}(\rho_i)^2) = c^{-1}\mathbf{m}(\rho_i)^2 \left(\boldsymbol{\Delta}(\rho_i)^{-1} - 1 \right) = \boldsymbol{\sigma}_i^2. \end{aligned}$$

Moreover, using Lemma 5.2 in conjunction with Assumption 6, we obtain

$$F_{i,n}^* F_{i,n} = c_n^{-1} R_{i,n}^* \left(\widetilde{Q}_n(\rho_i) \frac{1}{n} Y_n^* Y_n \widetilde{Q}_n(\rho_i) \right) R_{i,n} \xrightarrow[n\to\infty]{\mathcal{P}} \boldsymbol{\varsigma}_i^2 I_{j_i}.$$

From any sequence of integers increasing to infinity, there exists a subsequence along which this convergence holds in the almost sure sense. Applying Lemma 5.1, we get that the inner expectation at the left hand side of (11) converges almost surely along this subsequence towards $\prod_{i=1}^{p} \exp(-t^2(\boldsymbol{\sigma}_i^2 \operatorname{Tr} B_i^2 + \frac{1}{2}\boldsymbol{\varsigma}_i^2 \operatorname{Tr} C_i C_i^*)/2)$. Using in addition Lemma 5.3 along with the dominated convergence theorem, we obtain that Convergence (11) holds true along this subsequence. Since the original sequence is arbitrary, we obtain the required result. □

The remainder of the proof of Theorem 2.4 is an adaptation of the approach of [7].

Lemma 5.5. *For a given $x \in \mathbb{R}$ and a given $i \in \{1, \ldots, p\}$, let $y = \rho_i + n^{-1/2}x$, and let*

$$\widehat{S}_n(y) = \begin{bmatrix} \sqrt{y} U_n^* Q_n(y) U_n & I_r + n^{-1/2} U_n^* Y_n \widetilde{Q}_n(y) R_n \\ I_r + n^{-1/2} R_n^* \widetilde{Q}_n(y) Y_n^* U_n & \sqrt{y} R_n^* \widetilde{Q}_n(y) R_n \end{bmatrix}$$

Let

$$\chi_n^{(i)}(x) = n^{\frac{j_i}{2}} \left[\det \hat{S}_n(y) - \prod_{k\neq i}^{t} \left[\omega_k^2 g(\rho_i) - 1\right]^{j_k} \right.$$
$$\times \det \left(\frac{\sqrt{n} U_{i,n}^* \left(Q_n(\rho_i) - m_n(\rho_i) I_N\right) U_{i,n}}{\mathbf{m}(\rho_i)} + \rho_i \mathbf{m}(\rho_i) \sqrt{n} R_{i,n}^* (\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n} \right.$$
$$\left. \left. - 2\Re \left[U_{i,n}^* Y_n \widetilde{Q}_n(\rho_i) R_{i,n} \right] - \sqrt{n}(H_{i,n}(\rho_i) + I_{j_i}) - x H_{i,n}'(\rho_i) \right) \right] \tag{12}$$

where $\Re(M) = (M + M^*)/2$ *for a square matrix* M*. Then*

$$(\chi_n^{(i)}(x_1), \ldots, \chi_n^{(i)}(x_p)) \xrightarrow[n\to\infty]{\mathcal{P}} 0$$

for every finite sequence $\{x_1, \ldots, x_p\}$*.*

Proof. We show the result for $i = 1$, the same procedure being valid for the other values of i. The notation $X_n = o_P(1)$ means that the random variable X_n converges to zero in probability, while $X_n = \mathcal{O}_P(n^{-\ell})$ means that $n^\ell X_n$ is tight. Write $U = [U_1, \bar{U}_1]$ and $R = [R_1, \bar{R}_1]$ where $\bar{U}_1 = [U_2, \ldots, U_t]$ and $\bar{R}_1 = [R_2, \ldots, R_t]$. Writing

$$A = \begin{bmatrix} \sqrt{y} U_1^* Q(y) U_1 & \sqrt{y} U_1^* Q(y) \bar{U}_1 \\ \sqrt{y} \bar{U}_1^* Q(y) U_1 & \sqrt{y} \bar{U}_1^* Q(y) \bar{U}_1 \end{bmatrix} := \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix},$$
$$B = \begin{bmatrix} I_{j_1} + n^{-1/2} U_1^* Y \widetilde{Q}(y) R_1 & n^{-1/2} U_1^* Y \widetilde{Q}(y) \bar{R}_1 \\ n^{-1/2} \bar{U}_1^* Y \widetilde{Q}(y) R_1 & I_{r-j_1} + n^{-1/2} \bar{U}_1^* Y \widetilde{Q}(y) \bar{R}_1 \end{bmatrix} := \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \text{ and}$$
$$C = \begin{bmatrix} \sqrt{y} R_1^* \widetilde{Q}(y) R_1 & \sqrt{y} R_1^* \widetilde{Q}(y) \bar{R}_1 \\ \sqrt{y} \bar{R}_1^* \widetilde{Q}(y) R_1 & \sqrt{y} \bar{R}_1^* \widetilde{Q}(y) \bar{R}_1 \end{bmatrix} := \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^* & C_{22} \end{bmatrix},$$

we have

$$\det \widehat{S} = \det \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} = \det \begin{bmatrix} A_{11} & B_{11} & A_{12} & B_{12} \\ B_{11}^* & C_{11} & B_{21}^* & C_{12} \\ A_{12}^* & B_{21} & A_{22} & B_{22} \\ B_{12}^* & C_{12}^* & B_{22}^* & C_{22} \end{bmatrix} := \det \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{bmatrix}$$

after a row and column permutation. Hence $n^{j_1/2} \det \widehat{S} = \det M_{22} \times n^{j_1/2} \det(M_{11} - M_{12} M_{22}^{-1} M_{12}^*)$. Write $\mathbf{\Omega} = \operatorname{diag}(\omega_1^2 I_{j_1}, \mathbf{\Omega}_2)$. From the first order analysis we get that

$$M_{22} \xrightarrow[n\to\infty]{\text{a.s.}} \begin{bmatrix} \sqrt{\rho_1} \mathbf{m}(\rho_1) I_{r-j_1} & I_{r-j_1} \\ I_{r-j_1} & \sqrt{\rho_1} (c\mathbf{m}(\rho_1) - \rho_1^{-1}(1-c)) \mathbf{\Omega}_2 \end{bmatrix}$$

which is invertible since $\det M_{22} \overset{\text{as}}{\to} \prod_{k>1} (\omega_k^2 g(\rho_1) - 1)^{j_k} \neq 0$. Moreover, $\|M_{12}\| = \mathcal{O}_P(n^{-1/2})$. To see this, consider for instance the term $\sqrt{n} C_{12} = \sqrt{ny} R_1^* (\widetilde{Q} - \widetilde{T}) \bar{R}_1 + \sqrt{ny} R_1^* \widetilde{T} \bar{R}_1$. The first term is tight by Lemma 5.4 while the second is bounded by Assumption 7. The other terms are treated similarly. It results that $\|M_{12} M_{22}^{-1} M_{12}^*\| = \mathcal{O}_P(1/n)$.
In addition, $\det(y^{-1/2} C_{11}) \overset{\text{as}}{\to} [\omega_1^2 (c\mathbf{m}(\rho_1) - \rho_1^{-1}(1-c))]^{j_1} = (\rho_1 \mathbf{m}(\rho_1))^{-j_1}$ by the

definition of ρ_1. From these observations we get that

$$n^{j_1/2}\det \widehat{S} = \left(\prod_{k>1}(\omega_k^2 g(\rho_1)-1)^{j_k}+o_P(1)\right)\left((\rho_1\mathbf{m}(\rho_1))^{-j_1}+o_P(1)\right) \\ \times \det\left(\sqrt{ny}A_{11}-\sqrt{ny}B_{11}C_{11}^{-1}B_{11}^*+\mathcal{O}_P(n^{-1/2})\right).$$

Now we make the expansion

$$\begin{aligned}
&\sqrt{ny}A_{11}-\sqrt{ny}B_{11}C_{11}^{-1}B_{11}^* \qquad (13)\\
&= y\sqrt{n}U_1^*(Q(y)-m_n(y)I_N)U_1 + y\sqrt{n}m_n(y)I_{j_1}\\
&+\sqrt{n}\left(I_{j_1}+U_1^*\frac{Y\widetilde{Q}(y)}{\sqrt{n}}R_1\right)(R_1^*\widetilde{Q}(y)R_1)^{-1}(R_1^*(\widetilde{Q}(y)-\widetilde{T}(y))R_1)(R_1^*\widetilde{T}(y)R_1)^{-1}\\
&\times\left(I_{j_1}+R_1^*\frac{\widetilde{Q}(y)Y^*}{\sqrt{n}}U_1\right)\\
&-\sqrt{n}\left(I_{j_1}+U_1^*\frac{Y\widetilde{Q}(y)}{\sqrt{n}}R_1\right)(R_1^*\widetilde{T}(y)R_1)^{-1}\left(I_{j_1}+R_1^*\frac{\widetilde{Q}(y)Y^*}{\sqrt{n}}U_1\right). \qquad (14)
\end{aligned}$$

To go further, remark that

$$\begin{aligned}
&y\sqrt{n}m_n(y)I_{j_1}-\sqrt{n}(R_1^*\widetilde{T}(y)R_1)^{-1}\\
&\quad=(R_1^*\widetilde{T}(y)R_1)^{-1}\left[\sqrt{n}\left(ym_n(y)R_1^*\widetilde{T}(y)R_1-\rho_1 m_n(\rho_1)R_1^*\widetilde{T}(\rho_1)R_1\right)\right.\\
&\quad\left.+\sqrt{n}\left(\rho_1 m_n(\rho_1)R_1^*\widetilde{T}(\rho_1)R_1-I_{j_1}\right)\right]\\
&\quad=\rho_1\mathbf{m}(\rho_1)\left[-xH_1'(\rho_1)-\sqrt{n}(H_1(\rho_1)+I_{j_1})\right]+o(1)
\end{aligned}$$

where we recall that $y=\rho_1+xn^{-\frac{1}{2}}\to\rho_1$ and that $R_1^*\widetilde{T}(y)R_1\to(\rho_1\mathbf{m}(\rho_1))^{-1}I_{j_1}$. Recall from Lemma 5.4 that $\sqrt{n}U_1^*(Q-m_nI)U_1$, $\sqrt{n}R_1^*(\widetilde{Q}-\widetilde{T})R_1$, and $U_1^*Y\widetilde{Q}R_1$ are tight. Keeping the non negligible terms, we can write (13) under the form

$$\begin{aligned}
&\sqrt{ny}A_{11}-\sqrt{ny}B_{11}C_{11}^{-1}B_{11}^*\\
&=\rho_1\sqrt{n}U_1^*(Q(\rho_1)-m_n(\rho_1)I_N)U_1+(\rho_1\mathbf{m}(\rho_1))^2\sqrt{n}R_1^*(\widetilde{Q}(\rho_1)-\widetilde{T}(\rho_1)I_n)R_1\\
&-2\rho_1\mathbf{m}(\rho_1)\Re\left[U_1^*Y\widetilde{Q}(\rho_1)R_1\right]-\rho_1\mathbf{m}(\rho_1)\left(xH_1'(\rho_1)+\sqrt{n}(H_1(\rho_1)+I_{j_1})\right)+o_P(1).
\end{aligned}$$

Plugging this expression at the right hand side of the expression of $n^{j_1/2}\det\widehat{S}$ and observing that $H_{1,n}'(\rho_1)\to-\omega_1^2 g(\rho_1)'I_{j_1}$ concludes the proof. □

For $i=1,\dots,p$, take $x_1(i)>y_1(i)>x_2(i)>y_2(i)>\ldots>y_{j_i}(i)$ fixed sequences of real numbers. Call $J_n=(\sqrt{n}(\hat{\lambda}^n_{k(i)+\ell}-\rho_i),\ i=1,\dots,p,\ \ell=1,\dots,j_i)$, with $k(i)=\sum_{m=1}^{i-1}j_m$. Let also C be the rectangle $C=[x_1(1),y_1(1)]\times\ldots\times[x_p(j_p),y_p(j_p)]$. Then, for all large n, we have

$$\mathbb{P}\left(J_n\in C\right)=\mathbb{P}\left(\left\{\det\hat{S}_n\left(\rho_i+\frac{x_\ell(i)}{\sqrt{n}}\right)\det\hat{S}_n\left(\rho_i+\frac{y_\ell(i)}{\sqrt{n}}\right)<0\right\}\right)$$

since $\det\hat{S}_n(t)$ changes sign around $t=\hat{\lambda}^n_{k(i)+\ell}$, and only there (with probability one, for all large n).

From Lemma 5.5, we see that, for growing n, the probability for the product of the determinants above to be negative for all i and ℓ approaches the probability

$$\mathbb{P}\left(\left\{\det A_{x_\ell(i)} \det A_{y_\ell(i)} < 0,\ i = 1, \ldots, p,\ \ell = 1, \ldots, j_i\right\}\right)$$

where A_x is the matrix

$$A_x = \frac{\sqrt{n} U_{i,n}^* \left(Q_n(\rho_i) - m_n(\rho_i) I_N\right) U_{i,n}}{\mathbf{m}(\rho_i)} + \frac{\rho_i \sqrt{n} R_{i,n}^* (\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n}}{\omega_i^2 (c + c\rho_i \mathbf{m}(\rho_i) - 1)}$$
$$- 2\Re\left[U_{i,n}^* Y \tilde{Q}_n(\rho_i) R_{i,n}\right] - \sqrt{n}(H_{i,n}(\rho_i) + I_{j_i}) - x H'_{i,n}(\rho_i).$$

This last probability is equal to $\mathbb{P}(\bar{J}_n \in C)$, where $\bar{J}_n$ is the vector obtained by stacking the p vectors of decreasingly ordered eigenvalues of the matrices

$$B_i = [H'_{i,n}(\rho_i)]^{-1} \left(\frac{\sqrt{n} U_{i,n}^* \left(Q_n(\rho_i) - m_n(\rho_i) I_N\right) U_{i,n}}{\mathbf{m}(\rho_i)} \right.$$
$$\left. + \frac{\rho_i \sqrt{n} R_{i,n}^* (\widetilde{Q}_n(\rho_i) - \widetilde{T}_n(\rho_i)) R_{i,n}}{\omega_i^2 (c + c\rho_i \mathbf{m}(\rho_i) - 1)} - 2\Re\left[U_{i,n}^* Y \tilde{Q}_n(\rho_i) R_{i,n}\right] - \sqrt{n}(H_{i,n}(\rho_i) + I_{j_i}) \right).$$

From Lemma 5.4, $\{B_1, \ldots, B_t\}$ asymptotically behave as scaled non-zero mean GUE matrices. Precisely, denoting $\bar{B}_i = H'_{i,n}(\rho_i) B_i + \sqrt{n}(H_{i,n}(\rho_i) + I_{j_i})$, from Lemma 5.4 and for all a, b,

$$\mathbb{E}\left[|(\bar{B}_i)_{ab}|^2\right]$$
$$\to \frac{\boldsymbol{\sigma}_i^2}{\mathbf{m}(\rho_i)^2} + \frac{\rho_i^2 \tilde{\boldsymbol{\sigma}}_i^2}{\omega_i^4 (c + c\rho_i \mathbf{m}(\rho_i) - 1)^2} + 2\boldsymbol{\varsigma}^2$$
$$= \frac{\boldsymbol{\sigma}_i^2}{\mathbf{m}(\rho_i)^2} + \rho_i^2 \mathbf{m}(\rho_i)^2 \tilde{\boldsymbol{\sigma}}_i^2 + 2\boldsymbol{\varsigma}^2$$
$$= \frac{\mathbf{m}^2(\rho_i)}{\boldsymbol{\Delta}(\rho_i)} \left[\int \frac{t^2 \nu(dt)}{(1 + c\mathbf{m}(\rho_i) t)^2} + c\omega_i^4 \left(\int \frac{\mathbf{m}(\rho_i) t \nu(dt)}{(1 + c\mathbf{m}(\rho_i) t)^2} \right)^2 + \int \frac{2\omega_i^2 t \nu(dt)}{(1 + c\mathbf{m}(\rho_i) t)^2} \right].$$

This concludes the proof of Theorem 2.4.

Appendix A. Proofs of Lemmas 5.1 to 5.3

A.1. **Proof of Lemma 5.1.** Given a $s \times s$ deterministic Hermitian matrix A and a $s \times s$ deterministic complex matrix B, let $\Gamma_N = N^{-1/2} \operatorname{Tr} A Z_N^* \Upsilon_N Z_N + \Re(\operatorname{Tr} B Z_N^* F_N)$ where $\Re(M) = (M + M^*)/2$ for any square matrix M. We shall show that for any $t \in \mathbb{R}$,

$$\varphi_N(t) := \mathbb{E}[\exp(\imath t \Gamma_N)] \xrightarrow[N \to \infty]{} \exp\Big(-t^2 \frac{\boldsymbol{\sigma}^2 \operatorname{Tr} A^2 + \boldsymbol{\varsigma}^2 \operatorname{Tr} B B^*/2}{2}\Big) := \exp\Big(-\frac{t^2 \boldsymbol{v}^2}{2}\Big).$$

The result will follow by invoking the Cramér-Wold device. To establish this convergence, we show that the derivative $\varphi'_N(t)$ satisfies $\varphi'_N(t) = -t\boldsymbol{v}^2 \varphi_N(t) + \varepsilon_N(t)$ where $\varepsilon_N(t) \to 0$ as $N \to \infty$ uniformly on any compact interval of $\mathbb{R}$. That being true, the function $\psi_N(t) = \varphi_N(t) \exp(t^2 \boldsymbol{v}^2/2)$ satisfies $\psi_N(t) = 1 + \int_0^t \varepsilon_N(u) \exp(u^2 \boldsymbol{v}^2/2) du \to 1$ which proves the lemma.

By the IP formula, we get

$$\begin{aligned}\varphi'(t) &= \imath\mathbb{E}[\Gamma\exp(\imath t\Gamma)]\\ &= \imath\mathbb{E}\Big[\Big(\sum_{i,j=1}^{s}\sum_{k,\ell=1}^{N}\frac{A_{ij}Z_{kj}^*\Upsilon_{k\ell}Z_{\ell i}}{\sqrt{N}}+\sum_{i,j=1}^{s}\sum_{k=1}^{N}\frac{B_{ij}Z_{kj}^*F_{ki}+F_{ki}^*Z_{kj}B_{ij}^*}{2}\Big)\\ &\qquad\qquad\times\exp(\imath t\Gamma)\Big]\\ &= \imath\mathbb{E}\Big[\sum_{i,j,k,\ell}\frac{A_{ij}\Upsilon_{k\ell}}{\sqrt{N}}\frac{\partial(Z_{\ell i}\exp(\imath t\Gamma))}{\partial Z_{kj}}\\ &\qquad\qquad+\frac{1}{2}\sum_{i,j,k}B_{ij}F_{ki}\frac{\partial\exp(\imath t\Gamma)}{\partial Z_{kj}}+F_{ki}^*B_{ij}^*\frac{\partial\exp(\imath t\Gamma)}{\partial Z_{kj}^*}\Big].\end{aligned}$$

We obtain after a small calculation

$$\begin{aligned}\frac{\partial\exp(\imath t\Gamma)}{\partial Z_{kj}} &= \imath t\Big(\frac{[AZ^*\Upsilon]_{jk}}{\sqrt{N}}+\frac{1}{2}[B^*F^*]_{jk}\Big)\exp(\imath t\Gamma),\\ \frac{\partial\exp(\imath t\Gamma)}{\partial Z_{kj}^*} &= \imath t\Big(\frac{[\Upsilon ZA]_{kj}}{\sqrt{N}}+\frac{1}{2}[FB]_{kj}\Big)\exp(\imath t\Gamma)\end{aligned}$$

which leads to

$$\begin{aligned}\varphi'(t) &= -t\mathbb{E}[N^{-1}\operatorname{Tr}A^2Z^*\Upsilon^2Z\,\exp(\imath t\Gamma)]-(t/2)\operatorname{Tr}(BB^*F^*F)\,\varphi(t)\\ &\quad+\imath N^{-1/2}\operatorname{Tr}A\operatorname{Tr}\Upsilon\,\varphi(t)\\ &\quad-t\mathbb{E}[N^{-1/2}\operatorname{Tr}AB^*F^*\Upsilon Z\,\exp(\imath t\Gamma)]-(t/2)\mathbb{E}[N^{-1/2}\operatorname{Tr}Z^*\Upsilon FBA\,\exp(\imath t\Gamma)].\end{aligned}$$

Let us consider the first term at the right hand side of this equation. We have $\mathbb{E}[N^{-1}\operatorname{Tr}A^2Z^*\Upsilon^2Z]=N^{-1}\operatorname{Tr}A^2\operatorname{Tr}\Upsilon^2$. Applying the Poincaré-Nash inequality, we obtain after some calculations that $\mathbb{V}\mathrm{ar}(N^{-1}\operatorname{Tr}A^2Z^*\Upsilon^2Z)\le 2N^{-2}\operatorname{Tr}A^4\operatorname{Tr}\Upsilon^4=\mathcal{O}(N^{-1})$ since $\|\Upsilon\|$ is bounded. It results that $\mathbb{E}[N^{-1}\operatorname{Tr}A^2Z^*\Upsilon^2Z\exp(\imath t\Gamma)]=N^{-1}\operatorname{Tr}A^2\operatorname{Tr}\Upsilon^2\,\varphi(t)+\mathcal{O}(N^{-1/2})$ by Cauchy-Schwarz inequality. The third term is zero by hypothesis. Finally, $N^{-1}\mathbb{E}|\operatorname{Tr}Z^*\Upsilon FBA|^2=N^{-1}\operatorname{Tr}\Upsilon^2FBA^2B^*F^*\le N^{-1}\|\Upsilon\|^2\operatorname{Tr}FBA^2B^*F^*=\mathcal{O}(N^{-1})$. Hence, the last two terms are $\mathcal{O}(N^{-1/2})$ by Cauchy-Schwarz inequality, which proves Lemma 5.1.

A.2. **An intermediate result.** The following lemma will be needed in the proof of Lemma 5.2:

Lemma A.1. *For $x,y\in\operatorname{supp}(\mu)^c$,*

$$\begin{aligned}\mathbb{E}\left[\phi_n\frac{1}{n}\operatorname{Tr}\widetilde{Q}_n(x)D\widetilde{Q}_n(y)D\right] &= \frac{\frac{1}{n}\operatorname{Tr}D\widetilde{T}_n(x)D\widetilde{T}_n(y)}{1-c_nxm_n(x)ym_n(y)\frac{1}{n}\operatorname{Tr}D\widetilde{T}_n(x)D\widetilde{T}_n(y)}+\mathcal{O}(n^{-1})\\ \mathbb{E}\left[\phi_n\tilde{u}_n^*\widetilde{Q}_n(x)D\widetilde{Q}_n(y)\tilde{v}_n\right] &= \frac{\tilde{u}_n^*\widetilde{T}_n(x)D\widetilde{T}_n(y)\tilde{v}_n}{1-c_nxm_n(x)ym_n(y)\frac{1}{n}\operatorname{Tr}D\widetilde{T}_n(x)D\widetilde{T}_n(y)}+\mathcal{O}(n^{-1}).\end{aligned}$$

Proof. We denote here $\widetilde{Q}_x = \widetilde{Q}(x)$ and drop all unnecessary indices. Using the IP formula, we obtain

$$\begin{aligned}\mathbb{E}\left[\phi Y_{ia}^* Y_{ij} \widetilde{Q}_{x,jp} d_p \widetilde{Q}_{y,pq}\right] &= \frac{d_p d_j}{n}\Big(\delta(a-j)\mathbb{E}\left[\phi \widetilde{Q}_{x,jp}\widetilde{Q}_{y,pq}\right] \\ &-\frac{1}{n}\mathbb{E}\left[\phi \widetilde{Q}_{x,jj}[Y\widetilde{Q}_x]_{ip} Y_{ia}^* \widetilde{Q}_{y,pq}\right] - \frac{1}{n}\mathbb{E}\left[\phi Y_{ia}^* \widetilde{Q}_{x,jp}\widetilde{Q}_{y,pj}[Y\widetilde{Q}_y]_{iq}\right] \\ &+\mathbb{E}\left[\frac{1}{n}[\mathrm{adj}(\psi)\psi' Y]_{ij} Y_{ia}^* \widetilde{Q}_{x,jp}\widetilde{Q}_{y,pq}\right]\Big).\end{aligned}$$

Making the sum over i, p, and j, this is

$$\begin{aligned}&\frac{1}{n}\mathbb{E}\left[\phi[Y^*Y\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right] = \frac{1}{n^2}\mathbb{E}\left[[Y^*\,\mathrm{adj}(\psi)\psi' Y D\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right] + c_n d_a \mathbb{E}\left[\phi[\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right] \\ &-\frac{1}{n}\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_x[Y^*Y\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right] - \frac{1}{n}\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D[Y^*Y\widetilde{Q}_y]_{aq}\right].\end{aligned}$$

Using the relation $\frac{1}{n}Y^*Y\widetilde{Q}_x = x\widetilde{Q}_x + I_n$ and appropriately gathering the terms on each side gives

$$\begin{aligned}&\mathbb{E}\left[\phi[\widetilde{Q}_x D\widetilde{Q}_y]_{aq}(x - c_n\ d_a + x\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_x)\right] \\ &= -\mathbb{E}\left[\phi[D\widetilde{Q}_y]_{aq}(1+\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_x)\right] - \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D(\delta(a-q)+y[\widetilde{Q}_y]_{aq})\right] \\ &+\mathbb{E}\left[\frac{1}{n^2}[Y^*\,\mathrm{adj}(\psi)\psi' Y D\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right]. \qquad (15)\end{aligned}$$

Introducing the term $\tilde{\beta}_x = \phi\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_x$ and $\hat{\tilde{\beta}}_x = \tilde{\beta}_x - \phi\mathbb{E}[\tilde{\beta}_x]$, we have

$$\begin{aligned}&\mathbb{E}\left[\phi[\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right](x - c_n\ d_a + x\mathbb{E}[\tilde{\beta}_x]) \\ &= -\mathbb{E}\left[\phi[D\widetilde{Q}_y]_{aq}\right](1+\mathbb{E}[\beta_x]) - \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D\right](\delta(a-q)+y\mathbb{E}[[\widetilde{Q}_y]_{aq}]) \\ &-\mathbb{E}\left[[D\widetilde{Q}_y]_{aq}\hat{\tilde{\beta}}_x\right] - \mathbb{E}\left[[\widetilde{Q}_x D\widetilde{Q}_y]_{aq} x\hat{\tilde{\beta}}_x\right] - \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D y([\widetilde{Q}_y]_{aq} - \mathbb{E}[\widetilde{Q}_y]_{aq})\right] \\ &+\mathbb{E}\left[\frac{1}{n^2}[Y^*\,\mathrm{adj}(\psi)\psi' Y D\widetilde{Q}_x D\widetilde{Q}_y]_{aq}\right]. \qquad (16)\end{aligned}$$

At this point, we can prove both results for the trace and for the quadratic form. We start by dividing each side by $x - c_n\ d_a + x\mathbb{E}[\tilde{\beta}_x]$. We begin with the trace result. Multiplying the resulting left- and right-hand sides by d_a, summing over $a = q$ and normalizing by $1/n$, we obtain

$$\begin{aligned}\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_x D\right] &= -(1+\mathbb{E}[\tilde{\beta}_x])\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_y D A_x\right] \\ &\quad -\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D\right]\left(y\mathbb{E}[\phi\frac{1}{n}\operatorname{Tr} D A_x\widetilde{Q}_y] + \frac{1}{n}\operatorname{Tr} D A_x\right) + \varepsilon_n\end{aligned}$$

where we denoted $A_x = (x(1+\mathbb{E}[\tilde{\beta}_x])I_n - c_n D)^{-1}$ and where

$$\begin{aligned}\varepsilon_n = &\mathbb{E}\left[\operatorname{Tr}\frac{Y^*\operatorname{adj}(\psi)\psi' Y}{n^3} D\widetilde{Q}_x D\widetilde{Q}_y D A_x\right] - \mathbb{E}\left[\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_y D A_x \hat{\tilde{\beta}}_x\right]\\ &- \mathbb{E}\left[\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D A_x \hat{\tilde{\beta}}_x\right]\\ &- \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D y\left(\frac{1}{n}\operatorname{Tr}\widetilde{Q}_y D A_x - \mathbb{E}(\frac{1}{n}\operatorname{Tr}\widetilde{Q}_y D A_x)\right)\right]. \end{aligned} \tag{17}$$

From Lemma 4.5, $\mathbb{E}[\tilde{\beta}_x] = \tilde{\delta}_x + \mathcal{O}(n^{-2})$, where we denoted $\tilde{\delta}_x = \frac{1}{n}\operatorname{Tr} D\widetilde{T}_x$. Also, it is easily observed that

$$\left(I_n(1+\tilde{\delta}_x)x - c_n D\right)^{-1} = -\frac{1}{1+\tilde{\delta}_x}\widetilde{T}_x \tag{18}$$

with $\widetilde{T}_x = \widetilde{T}(x)$. Therefore, along with Lemma 4.5, we now have

$$\begin{aligned}&\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_x D\right]\\ &= \frac{1}{n}\operatorname{Tr} D\widetilde{T}_x D\widetilde{T}_y + \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_x D\right]\frac{y\frac{1}{n}\operatorname{Tr} D\widetilde{T}_x\widetilde{T}_y + \frac{1}{n}\operatorname{Tr} D\widetilde{T}_x}{1+\tilde{\delta}_x} + \varepsilon_n + \mathcal{O}(n^{-2}).\end{aligned}$$

Using now the fact that $y\widetilde{T}_y + I_n = c_n\frac{1}{1+\tilde{\delta}_y}D\widetilde{T}_y$, we conclude

$$\mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_x D\right] = \frac{\frac{1}{n}\operatorname{Tr} D\widetilde{T}_x D\widetilde{T}_y}{1 - c_n(1+\tilde{\delta}_x)^{-1}(1+\tilde{\delta}_y)^{-1}\frac{1}{n}\operatorname{Tr} D\widetilde{T}_x D\widetilde{T}_y} + \varepsilon_n + \mathcal{O}(n^{-2}).$$

It therefore remains to prove that $\varepsilon_n = \mathcal{O}(n^{-1})$. Due to the presence of ϕ in the expression of $\hat{\tilde{\beta}}_x$, and using Lemma 4.3 and Cauchy-Schwarz inequality, one can see that the last three terms in the expression of ε_n are $\mathcal{O}(n^{-1})$. As for the first term, it is treated in a similar manner as in the proof of Lemma 4.6, and is $\mathcal{O}(n^{-2})$.

In order to prove the result on the quadratic form, we start again from (16). Dividing each side again by $x - c_n\, d_a + x\mathbb{E}[\tilde{\beta}_x]$, introducing $[\tilde{u}]_a$, $[\tilde{v}]_q$, and summing over the indices, we obtain

$$\begin{aligned}&\mathbb{E}\left[\phi\tilde{u}^*\widetilde{Q}_x D\widetilde{Q}_y\tilde{v}\right]\\ &= -\mathbb{E}\left[\phi\tilde{u}^* A_x D\widetilde{Q}_y\tilde{v}\right] - \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D\right]\left(\tilde{u}^* A_x(y\mathbb{E}[\phi\widetilde{Q}_y] + I_n)\tilde{v}\right) + \varepsilon_n'\end{aligned} \tag{19}$$

where ε_n' is very similar to ε_n and is shown to be $\mathcal{O}(n^{-1})$ with the same line of arguments. Using Lemma 4.3, (18), and the previous result on $\mathbb{E}[\phi\frac{1}{n}\operatorname{Tr}\widetilde{Q}_x D\widetilde{Q}_y D]$, we finally obtain

$$\begin{aligned}&\mathbb{E}\left[\phi\tilde{u}^*\widetilde{Q}_x D\widetilde{Q}_y\tilde{v}\right]\\ &= \tilde{u}^*\widetilde{T}_x D\widetilde{T}_y\tilde{v}\left(1 + \frac{c_n(1+\tilde{\delta}_x)^{-1}(1+\tilde{\delta}_y)^{-1}\frac{1}{n}\operatorname{Tr} D\widetilde{T}_x D\widetilde{T}_y}{1 - c_n(1+\tilde{\delta}_x)^{-1}(1+\tilde{\delta}_y)^{-1}\frac{1}{n}\operatorname{Tr} D\widetilde{T}_x D\widetilde{T}_y}\right) + \mathcal{O}(n^{-1}).\end{aligned}$$

from which

$$\mathbb{E}\left[\phi\tilde{u}^*\widetilde{Q}_x D\widetilde{Q}_y\tilde{v}\right] = \frac{\tilde{u}^*\widetilde{T}_x D\widetilde{T}_y\tilde{v}}{1 - c_n(1+\tilde{\delta}_x)^{-1}(1+\tilde{\delta}_y)^{-1}\frac{1}{n}\operatorname{Tr} D\widetilde{T}_x D\widetilde{T}_y} + \mathcal{O}(n^{-1}).$$

We conclude with the remark $xm_n(x) = -(1+\tilde{\delta}_x)^{-1}$. □

A.3. **Proof of Lemma 5.2.** The line of proof closely follows the proof of Lemma A.1. We provide here its main steps. By the IP formula, we have

$$\begin{aligned}\mathbb{E}[\phi\widetilde{Q}_{pk}Y^*_{\ell k}Y_{\ell m}\widetilde{Q}_{mr}] = &-\frac{d_m}{n}\mathbb{E}[\phi\widetilde{Q}_{pk}Y^*_{\ell k}\widetilde{Q}_{mm}[Y\widetilde{Q}]_{\ell r}] + \delta(k-m)d_m\mathbb{E}[\phi\widetilde{Q}_{pk}\widetilde{Q}_{mr}]\\ &- \frac{d_m}{n}\mathbb{E}[\phi Y^*_{\ell k}\widetilde{Q}_{mr}\widetilde{Q}_{pm}[Y\widetilde{Q}]_{\ell k}]\\ &+ \frac{d_m}{n}\mathbb{E}[\widetilde{Q}_{pk}Y^*_{\ell k}\widetilde{Q}_{mr}[\mathrm{adj}(\psi)\psi' Y]_{\ell m}]\end{aligned}$$

Taking the sum over m, we obtain

$$\begin{aligned}\mathbb{E}[\phi\widetilde{Q}_{pk}Y^*_{\ell k}[Y\widetilde{Q}]_{\ell r}] = &\frac{d_k}{1+\mathbb{E}[\tilde{\beta}]}\mathbb{E}[\phi\widetilde{Q}_{pk}\widetilde{Q}_{kr}] - \frac{1}{1+\mathbb{E}\tilde{\beta}}\frac{1}{n}\mathbb{E}[\phi Y^*_{\ell k}[\widetilde{Q}D\widetilde{Q}]_{pr}[Y\widetilde{Q}]_{\ell k}]\\ &+ \frac{1}{1+\mathbb{E}[\tilde{\beta}]}\frac{1}{n}\mathbb{E}[\widetilde{Q}_{pk}Y^*_{\ell k}[\mathrm{adj}(\psi)\psi' Y D\widetilde{Q}]_{\ell r}] - \mathbb{E}[\hat{\tilde{\beta}}\widetilde{Q}_{pk}Y^*_{\ell k}[Y\widetilde{Q}]_{\ell r}]\end{aligned}$$

where $\tilde{\beta}(x) = \phi\frac{1}{n}\operatorname{Tr} D\widetilde{Q}(x)$ and $\hat{\tilde{\beta}}(x) = \tilde{\beta}(x) - \phi\mathbb{E}[\tilde{\beta}(x)]$ as in the proof of Lemma A.1. Taking the sum over ℓ then over k, we obtain

$$\begin{aligned}\mathbb{E}[\phi[\widetilde{Q}\frac{Y^*Y}{n}\widetilde{Q}]_{pr}] = &c_n\frac{1}{1+\mathbb{E}[\tilde{\beta}]}\mathbb{E}[\phi[\widetilde{Q}D\widetilde{Q}]_{pr}] - \frac{1}{1+\mathbb{E}[\tilde{\beta}]}\mathbb{E}[\phi[\widetilde{Q}D\widetilde{Q}]_{pr}\frac{1}{n}\operatorname{Tr}(\frac{Y^*Y}{n}\widetilde{Q})]\\ &+ \frac{1}{1+\mathbb{E}[\tilde{\beta}]}\frac{1}{n}\mathbb{E}[[\widetilde{Q}\frac{Y^*\,\mathrm{adj}(\psi)\psi' Y}{n}D\widetilde{Q}]_{pr}] - \mathbb{E}[\hat{\tilde{\beta}}[\widetilde{Q}\frac{Y^*Y}{n}\widetilde{Q}]_{pr}]\end{aligned}$$

Observing that $(1+\mathbb{E}[\tilde{\beta}(x)])^{-1} = -xm_n(x) + \mathcal{O}(n^{-2})$ and making the usual approximations, we get

$$\mathbb{E}[\phi\tilde{u}^*\widetilde{Q}\frac{Y^*Y}{n}\widetilde{Q}\tilde{u}] = \Big(xm_n(x)\frac{1}{n}\operatorname{Tr}(\mathbb{E}[\phi\frac{Y^*Y}{n}\widetilde{Q}]) - c_n xm_n(x)\Big)\mathbb{E}[\phi\tilde{u}^*\widetilde{Q}D\widetilde{Q}\tilde{u}] + \mathcal{O}(n^{-1})$$

Observing that $n^{-1}\operatorname{Tr}(\mathbb{E}[\phi(n^{-1}Y^*Y)\widetilde{Q}(x)] = Nn^{-1}xm_n(x) + Nn^{-1} + \mathcal{O}(n^{-2})$ and invoking Lemma A.1, we obtain the desired result.

A.4. **Proof of Lemma 5.3.** As in the previous proofs, we discard unnecessary indices. We also denote $\widetilde{Q}_i = \widetilde{Q}(\rho_i)$. For readability, we also write $\tilde{M}_i = M_{i,n}A_i$ and use the shortcut notation $\Gamma = \sqrt{n}\sum_{i=1}^p \operatorname{Tr} M_i^*\widetilde{Q}_i\tilde{M}_i$. We focus first on the term in ρ_1. The line of proof closely follows that of Lemma 5.1 with the exception that we need to introduce the regularization function ϕ to ensure the existence of all the quantities under study. That is, with $\varphi_N(t) = \mathbb{E}[\exp(\imath t\phi\Gamma)]$, we only need to show that $\varphi_N'(t) = -t\tilde{\sigma}_n^2\varphi_N(t) + \mathcal{O}(1/\sqrt{n})$. Using $|\varphi_N(t)| \le 1$ and Lemma 4.4, $|\mathbb{E}[\exp(\imath t\Gamma)] - \varphi_N(t)| \le 1 - \mathbb{E}[\phi] \to 0$ as $N \to \infty$, from which the result unfolds.

Using the IP formula, we first obtain

$$\begin{aligned}
&\mathbb{E}\left[\phi\left[\frac{Y^*Y}{n}\widetilde{Q}_1\right]_{pq} e^{\imath t\phi\Gamma}\right] \\
&= c_n\mathbb{E}\left[\phi[D\widetilde{Q}_1]_{pq}e^{\imath t\phi\Gamma}\right] - \mathbb{E}\left[\phi\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_1\left[\frac{Y^*Y}{n}\widetilde{Q}_1\right]_{pq} e^{\imath t\phi\Gamma}\right] \\
&- \mathbb{E}\left[\imath t e^{\imath t\phi\Gamma}\phi^2\frac{1}{\sqrt{n}}\sum_{j=1}^{p}\sum_{a=1}^{r}\left[(\tilde{M}_j)_a^*\widetilde{Q}_j D\widetilde{Q}_1\right]_q\left[\frac{Y^*Y}{n}\widetilde{Q}_j(M_j)_a\right]_p\right] + \varepsilon_{n,pq}
\end{aligned}$$

where

$$\varepsilon_{n,pq} = \mathbb{E}\left[\frac{1}{n}\left[\frac{Y^*\operatorname{adj}(\psi)\psi' Y}{n}D\widetilde{Q}_1\right]_{pq} e^{\imath t\phi\Gamma}\right] + \mathbb{E}\left[\phi\frac{1}{n}\left[\frac{Y^*\operatorname{adj}(\psi)\psi' Y}{n}D\widetilde{Q}_1\right]_{pq}\imath t\Gamma e^{\imath t\phi\Gamma}\right]$$

and where we denoted X_a the column a of matrix X, X_a^* being the row vector $(X_a)^*$.

With $\tilde{\beta}_j = \phi\frac{1}{n}\operatorname{Tr} D\widetilde{Q}_j$, $\hat{\tilde{\beta}}_j = \tilde{\beta}_j - \phi\mathbb{E}\left[\tilde{\beta}_j\right]$, and with the relation $n^{-1}Y^*Y\widetilde{Q}_1 = I_n + \rho_1\widetilde{Q}_1$, we obtain

$$\begin{aligned}
&\left(\rho_1(1+\mathbb{E}[\tilde{\beta}_1]) - c_n\ d_p\right)\mathbb{E}\left[\phi[\widetilde{Q}_1]_{pq}e^{\imath t\phi\Gamma}\right] = -\delta(p-q)(1+\mathbb{E}[\tilde{\beta}_1])\mathbb{E}\left[\phi e^{\imath t\phi\Gamma}\right] \\
&- \mathbb{E}\left[\imath t e^{\imath t\phi\Gamma}\phi^2\frac{1}{\sqrt{n}}\sum_{j=1}^{p}\sum_{a=1}^{r}\left[(\tilde{M}_j)_a^*\widetilde{Q}_j D\widetilde{Q}_1\right]_q\left[\frac{Y^*Y}{n}\widetilde{Q}_j(M_j)_a\right]_p\right] + \varepsilon'_{n,pq}
\end{aligned}$$

where

$$\varepsilon'_{n,pq} = \varepsilon_{n,pq} - \mathbb{E}\left[\hat{\tilde{\beta}}_1\left[\frac{Y^*Y}{n}\widetilde{Q}_1\right]_{pq} e^{\imath t\phi\Gamma}\right].$$

Dividing each side by $\rho_1(1+\mathbb{E}[\tilde{\beta}_1]) - c_n\ d_p$, then multiplying by $(\tilde{M}_1)_p$ and $(M_1)_q$, and summing over p, q gives

$$\begin{aligned}
\mathbb{E}[\phi\operatorname{Tr}(\tilde{M}_1^*\widetilde{Q}_1 M_1)e^{\imath t\phi\Gamma}] &= -(1+\mathbb{E}[\tilde{\beta}_1])\mathbb{E}[\phi e^{\imath t\Gamma}]\operatorname{Tr}\left(\tilde{M}_1^* A_{\rho_1} M_1\right) \\
&- \imath t\mathbb{E}\left[\phi^2 e^{\imath t\phi\Gamma}\frac{1}{\sqrt{n}}\sum_{j=1}^{p}\operatorname{Tr}\tilde{M}_1^* A_{\rho_1}\frac{Y^*Y}{n}\widetilde{Q}_j M_j\tilde{M}_j^*\widetilde{Q}_j D\widetilde{Q}_1\right] + \varepsilon'_n
\end{aligned}$$

with $A_{\rho_i} = (\rho_i(1+\mathbb{E}[\tilde{\beta}_i])I_n - c_n D)^{-1}$, and

$$\varepsilon'_n = \operatorname{Tr}\tilde{M}_1^* A_{\rho_1} E' M_1$$

with $(E')_{pq} = \varepsilon'_{pq}$. From (18), the identity $n^{-1}Y^*Y\widetilde{Q}_j = I_n + \rho_j\widetilde{Q}_j$, and Lemma 4.5, we finally obtain

$$\begin{aligned}
&\mathbb{E}\left[\phi\operatorname{Tr}(\tilde{M}_1^*\widetilde{Q}_1 M_1)e^{\imath t\phi\Gamma}\right] - \mathbb{E}[\phi e^{\imath t\phi\Gamma}]\operatorname{Tr}\tilde{M}_1\widetilde{T}_1 M_1 \\
&= \imath t\mathbb{E}[\phi e^{\imath t\phi\Gamma}]\frac{1}{\sqrt{n}}\sum_{j=1}^{p}\frac{\operatorname{Tr}\tilde{M}_1^*\frac{\widetilde{T}_1}{1+\tilde{\delta}_1}\frac{c_n D\widetilde{T}_j}{1+\tilde{\delta}_j}M_j\tilde{M}_j^*\widetilde{T}_j D\widetilde{T}_1 M_1}{1 - c_n(1+\tilde{\delta}_1)^{-1}(1+\tilde{\delta}_j)^{-1}\frac{1}{n}\operatorname{Tr} D\widetilde{T}_1 D\widetilde{T}_j} + \varepsilon'_n + \mathcal{O}(n^{-2})
\end{aligned}$$

with $\widetilde{T}_i = \widetilde{T}(\rho_i)$, from which

$$\mathbb{E}\left[\phi \operatorname{Tr}(\tilde{M}_1^* \widetilde{Q}_1 M_1) e^{\imath t\phi\Gamma}\right] - \mathbb{E}[\phi e^{\imath t\phi\Gamma}] \operatorname{Tr} \tilde{M}_1 \widetilde{T}_1 M_1$$
$$= \frac{\imath t \mathbb{E}[\phi e^{\imath t\phi\Gamma}]}{\sqrt{n}} \sum_{j=1}^{p} \frac{c_n \rho_1 m_n(\rho_1) \rho_j m_n(\rho_j) \operatorname{Tr} \tilde{M}_1^* \widetilde{T}_1 D\widetilde{T}_j M_j \tilde{M}_j^* \widetilde{T}_j D\widetilde{T}_1 M_1}{1 - c_n \rho_1 m_n(\rho_1) \rho_j m_n(\rho_j) \frac{1}{n} \operatorname{Tr} D\widetilde{T}_1 D\widetilde{T}_j} + \varepsilon_n' + \mathcal{O}(n^{-2}).$$

It remains to show that $\varepsilon_n' = \mathcal{O}(n^{-1})$. We have explicitly

$$\varepsilon_n' = \mathbb{E}\left[\frac{1}{n} \operatorname{Tr}\left(\tilde{M}_1^* A_{\rho_1} \frac{Y^* \operatorname{adj}(\psi) \psi' Y}{n} D\widetilde{Q}_1 M_1\right)(1 + \phi\imath t\Gamma) e^{\imath t\phi\Gamma}\right]$$
$$- \mathbb{E}\left[\phi \hat{\tilde{\beta}}_1 \operatorname{Tr}\left(\tilde{M}_1^* A_{\rho_1} \frac{Y^* Y}{n} \widetilde{Q}_1 M_1\right) e^{\imath t\phi\Gamma}\right].$$

Using the fact that $|e^{\imath t\phi\Gamma}| = 1$ and the relation $n^{-1} Y^* Y \widetilde{Q}_1 = \rho_1 \widetilde{Q}_1 + I_n$, the second term is easily shown to be $\mathcal{O}(n^{-1})$ from the Cauchy-Scwharz inequality and Lemma 4.3. If it were not for the factor Γ, the convergence of the first term would unfold from similar arguments as in the proof of Lemma 4.6. We only need to show here that $\mathbb{E}[|\phi\Gamma|^2] = \mathcal{O}(1)$. But this follows immediately from Lemma 4.3 and Lemma 4.5.

The generalization to $\sum_i \mathbb{E}[\phi \operatorname{Tr}(\tilde{M}_i^* \widetilde{Q}_i M_i) e^{\imath t\phi\Gamma}]$ is then immediate and we have the expected result.

References

[1] Z. Bai and J.-F. Yao. Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(3):447–474, 2008.

[2] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3):611–677, 1999.

[3] 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(1):316–345, 1998.

[4] Z. D. Bai and Jack W. Silverstein. Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.*, 27(3):1536–1555, 1999.

[5] J. Baik, G. Ben Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.

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

[7] F. Benaych-Georges, A. Guionnet, and M. Maïda. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. *Electron. J. Prob.*, 16:1621–1662, 2011. Paper no. 60.

[8] F. Benaych-Georges and R. R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. in Math.*, 227(1):494–521, 2011.

[9] F. Benaych-Georges and R. R. Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *J. Multivariate Anal.*, 111:295–309, 2012.

[10] Vladimir Bolotnikov. On a general moment problem on the half axis. *Linear Algebra Appl.*, 255:57–112, 1997.

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

[12] M. Capitaine, C. Donati-Martin, and D. Féral. Central limit theorems for eigenvalues of deformations of Wigner matrices. *Ann. Inst. H. Poincaré Probab. Statist.*, 48(1):107–133, 2012.

[13] L. H. Y. Chen. An inequality for the multivariate normal distribution. *J. Multivariate Anal.*, 12(2):306–315, 1982.

[14] R. Couillet and W. Hachem. Fluctuations of spiked random matrix models and failure diagnosis in sensor networks. *IEEE Transactions on Information Theory*, 2012. In press.
[15] Fritz Gesztesy and Eduard Tsekanovskii. On matrix-valued Herglotz functions. *Math. Nachr.*, 218:61–138, 2000.
[16] V. L. Girko. *Theory of random determinants*, volume 45 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990.
[17] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2):711–775, 2005.
[18] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim, and L. Pastur. A new approach for mutual information analysis of large dimensional multi-antenna chennels. *IEEE Trans. Inform. Theory*, 54(9):3987–4004, 2008.
[19] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. Large information plus noise random matrix models and consistent subspace estimation in large sensor networks. *To be published in Random Matrices: Theory and Applications*, 2011. ArXiv preprint arXiv:1106.5119.
[20] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. A subspace estimator for fixed rank perturbations of large random matrices. *ArXiv preprint arXiv:1106.1497*, 2011.
[21] W. Hachem, P. Loubaton, J. Najim, and P. Vallet. On bilinear forms based on the resolvent of large random matrices. *To be published in Annales de l'IHP*, 2012.
[22] W. Hachem, Ph. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. *Ann. Appl. Probab.*, 17(3):875–930, 2007.
[23] I. M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29(2):295–327, 2001.
[24] A. M. Khorunzhy and L. A. Pastur. Limits of infinite interaction radius, dimensionality and the number of components for random operators with off-diagonal randomness. *Comm. Math. Phys.*, 153(3):605–646, 1993.
[25] M. G. Kreĭn and A. A. Nudel′man. *The Markov moment problem and extremal problems*. American Mathematical Society, Providence, R.I., 1977. Ideas and problems of P. L. Čebyšev and A. A. Markov and their further development, Translated from the Russian by D. Louvish, Translations of Mathematical Monographs, Vol. 50.
[26] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.
[27] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *IEEE Trans. Inform. Theory*, 54(11):5113–5129, 2008.
[28] X. Mestre. On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices. *IEEE Trans. Signal Process.*, 56(11):5353–5368, 2008.
[29] T. Næs, T. Isaksson, T. Fearn, and T. Davies. *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, UK, 2002.
[30] L. A. Pastur, A. M. Khorunzhiĭ, and V. Yu. Vasil′chuk. On an asymptotic property of the spectrum of the sum of one-dimensional independent random operators. *Dopov. Nats. Akad. Nauk Ukraïni*, (2):27–30, 1995.
[31] L. A. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.
[32] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica*, 17(4):1617–1642, 2007.
[33] S. Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields*, 134(1):127–173, 2006.
[34] M. Potters, J.-P. Bouchaud, and L. Laloux. Financial applications of random matrix theory: Old laces and new pieces. In *Proceedings of the Cracow conference on "Applications of Random Matrix Theory to Economy and Other Complex Systems"*, 2005. arXiv preprint: arXiv:physics/0507111.
[35] J. W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.*, 55(2):331–339, 1995.
[36] J. W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):175–192, 1995.
[37] Jack W. Silverstein and Sang-Il Choi. Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):295–309, 1995.

[38] D. Tse and P. Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.

[39] P. Vallet, P. Loubaton, and X. Mestre. Improved subspace estimation for multivariate observations of high dimension: The deterministic signals case. *Information Theory, IEEE Transactions on*, 58(2):1043 –1068, feb. 2012.

[40] J. Vinogradova, R. Couillet, and W. Hachem. Statistical inference in large antenna arrays under unknown noise pattern. *ArXiv preprint*, 2012.

François Chapon,
Université Paul Sabatier,
Institut de Mathématiques de Toulouse,
118 route de Narbonne,
31062 Toulouse Cedex 9,
France.
e-mail: francois.chapon@math.univ-toulouse.fr

Romain Couillet,
Supélec,
Plateau de Moulon,
91192 Gif-sur-Yvette Cedex,
France.
e-mail: romain.couillet@supelec.fr

Walid Hachem,
CNRS, Télécom Paristech
46, rue Barrault,
75634 Paris Cedex 13,
France.
e-mail: walid.hachem@telecom-paristech.fr

Xavier Mestre,
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC),
Parc Mediterrani de la Tecnologia (PMT) - Building B4,
Av. Carl Friedrich Gauss 7,
08860 - Castelldefels
Barcelona, Spain.
e-mail: xavier.mestre@cttc.cat