1l

Almost sure localization of the eigenvalues in a Gaussian information plus noise model – Application to the spiked models

Philippe Loubaton   Pascal Vallet *

Abstract

Let $\mathbf{\Sigma}_N$ be a $M \times N$ random matrix defined by $\mathbf{\Sigma}_N = \mathbf{B}_N + \sigma \mathbf{W}_N$ where $\mathbf{B}_N$ is a uniformly bounded deterministic matrix and where $\mathbf{W}_N$ is an independent identically distributed complex Gaussian matrix with zero mean and variance $\frac{1}{N}$ entries. The purpose of this paper is to study the almost sure location of the eigenvalues $\hat{\lambda}_{1,N} \geq \ldots \geq \hat{\lambda}_{M,N}$ of the Gram matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ when M and N converge to $+\infty$ such that the ratio $c_N = \frac{M}{N}$ converges towards a constant $c > 0$. The results are used in order to derive, using an alternative approach, known results concerning the behaviour of the largest eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ when the rank of $\mathbf{B}_N$ remains fixed and M, N tend to $+\infty$.

Key words: random matrix theory, gaussian information plus noise model, localization of the eigenvalues, spiked models.

AMS 2010 Subject Classification: Primary 15B52, 60F15.

---

1 Introduction

The addressed problem and the results Let $\mathbf{\Sigma}_N$ be a $M \times N$ complex-valued matrix defined by

$$\mathbf{\Sigma}_N = \mathbf{B}_N + \sigma \mathbf{W}_N \tag{1}$$

where $\mathbf{B}_N$ is a $M \times N$ deterministic matrix such that $\sup_N \|\mathbf{B}_N\| < +\infty$, and where $\mathbf{W}_N = [\mathbf{W}_N]_{i,j}$ is a $M \times N$ complex Gaussian random matrix with independent identically distributed (i.i.d) entries such that $\mathbb{E}\left[[\mathbf{W}_N]_{i,j}\right] = 0$, $\left|\mathbb{E}\left[[\mathbf{W}_N]_{i,j}\right]\right|^2 = \frac{1}{N}$, $\mathrm{Re}\left([\mathbf{W}_N]_{i,j}\right)$ and $\mathrm{Im}\left([\mathbf{W}_N]_{i,j}\right)$ are i.i.d zero mean real Gaussian random variables. Model (1) is referred in the literature to as the information plus noise model (see e.g Dozier-Silverstein [14]). In this paper, we assume that $\mathrm{Rank}(\mathbf{B}_N) = K(N) = K < M$ because this assumption is verified in a number of practical situations, in particular in the context of the spiked models addressed here.

The purpose of this paper is to study the almost sure location of the eigenvalues $\hat{\lambda}_{1,N} \geq \ldots \geq \hat{\lambda}_{M,N}$ of the Gram matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ when M and N converge to $+\infty$ such that the ratio $c_N = \frac{M}{N}$ converges towards a constant $c > 0$ and to take benefit of the results to obtain, using a different approach than Benaych-Nadakuditi [7], the behaviour of the largest eigenvalues of the information plus noise spiked models for which the rank K of $\mathbf{B}_N$ remains constant when M and N increase to $+\infty$.

The empirical spectral measure (or eigenvalue distribution) $\hat{\mu}_N = \frac{1}{M}\sum_{m=1}^{M}\delta_{\hat{\lambda}_{m,N}}$ of matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ has the same asymptotic behaviour than a deterministic probability distribution μ_N (see e.g. Dozier-Silverstein [14, Th.1.1] or Girko [16, Th.7.4]) whose support $\mathscr{S}_N$ is the union of disjoint compact intervals called in the following the clusters of $\mathscr{S}_N$. The boundary points of each cluster coincide with the positive extrema of a certain rational function depending on the empirical spectral measure of matrix $\mathbf{B}_N \mathbf{B}_N^*$, σ^2 and on the ratio $c_N = \frac{M}{N}$ (see [28], Thereom 2). Each cluster $\mathscr{I}$ of $\mathscr{S}_N$ appears to be naturally associated to another interval containing a group of consecutive eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ ([28]). It is shown in [28] that the property proved in Bai-Silverstein [2] holds in the context of model (1). Roughly speaking, it means that for an interval $[a, b]$ located outside $\mathscr{S}_N$ for N large enough, no eigenvalue of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ belong to $[a, b]$ almost surely, for all large N.

In this paper, we establish the analog of the property called in Bai-Silverstein [3] "exact separation": almost surely, for N large enough, the number of eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ less than a (resp. greater than b) coincides with the number of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ associated to the clusters included into $[0, a]$ (resp. included into $[b, \infty)$). Note that these results also hold in the case where $K = M$, not treated in this paper. Indeed, the analysis of the support $\mathscr{S}_N$ provided in [28] can be extended when $\mathbf{B}_N \mathbf{B}_N^*$ is full rank. Once the characterization of the support is established, the probabilistic part of the proof of the above mentioned exact separation result eigenvalues can be used verbatim.

We also use the separation result to study the case where $\mathrm{Rank}(\mathbf{B}_N) = K$ is independent of N. It is assumed that for each $k = 1, \ldots, K$, the non zero eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ satisfy $\lim_{N \to +\infty} \lambda_{k,N} = \lambda_k$. The support $\mathscr{S}_N$ of μ_N is first characterized in this case, and using the above results related to the almost sure location of the $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$, it is proved that if $\lambda_k > \sigma^2 \sqrt{c}$, then,

$$\hat{\lambda}_{k,N} \to \frac{(\sigma^2 + \lambda_k)(\sigma^2 c + \lambda_k)}{\lambda_k}, \tag{2}$$

and that if $\lambda_k \leq \sigma^2 \sqrt{c}$, then,

$$\hat{\lambda}_{k,N} \to \sigma^2 (1 + \sqrt{c})^2. \tag{3}$$

This behaviour was first established in [7] using a different approach.

Motivations Our work has been originally motivated by the context of array processing in which the signals transmitted by $K < M$ sources are received by an array equiped with M sensors. Under certain assumptions, the M-dimensional vector $\mathbf{y}(n)$ received on the sensor array at time n can be written as

$$\mathbf{y}(n) = \sum_{k=1}^{K} \mathbf{d}_k s_k(n) + \mathbf{v}(n), \qquad (4)$$

where each time series $(s_k(n))_{n \in \mathbb{Z}}$ represents a non observable deterministic signal corresponding to source k and where $\mathbf{d}_k$ is an unknown deterministic M-dimensional vector depending on the direction of arrival of the k-th source. $(\mathbf{v}(n))_{n \in \mathbb{Z}}$ is an additive complex white Gaussian noise such that $\mathbb{E}[\mathbf{v}(n)\mathbf{v}(n)^*] = \sigma^2 \mathbf{I}_M$. It is clear that (4) is equivalent to (1) if we put $\mathbf{\Sigma}_N = N^{-1/2}[\mathbf{y}(1), \ldots, \mathbf{y}(N)]$, $\mathbf{W}_N = N^{-1/2}\sigma^{-1}[\mathbf{v}(1), \ldots, \mathbf{v}(N)]$ and $\mathbf{B}_N = N^{-1/2}\mathbf{D}[\mathbf{s}(1), \ldots, \mathbf{s}(N)]$, with $\mathbf{s}(n) = [s_1(n), \ldots, s_K(n)]^T$ and $\mathbf{D} = [\mathbf{d}_1, \ldots, \mathbf{d}_K]$.

Model (4) poses important statistical problems such as detection of the number of sources K or estimation of the direction of arrivals of the K sources. A number of estimation schemes based on the eigenvalues and the eigenvectors of matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ were developed, and analysed if $N \to +\infty$ while M remains fixed. If however M and N are of the same order of magnitude, the above technics may fail, and it is therefore quite relevant to study these statistical problems in the asymptotic regime $M, N \to +\infty$ in such a way that $\frac{M}{N} \to c$, $c \in (0, +\infty)$. The number of sources may be constant or scale up with the dimensions M and N. For this, the first step is to evaluate the behaviour of the eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$.

About the literature *Concerning the zero-mean correlated model.* The problems addressed in this paper were studied extensively in the context of the popular zero-mean correlated model defined by

$$\mathbf{\Sigma}_N = \mathbf{H}_N \mathbf{W}_N, \qquad (5)$$

where $\mathbf{H}_N$ is a deterministic $M \times M$ matrix and where $\mathbf{W}_N$ is a random matrix with possibly non Gaussian zero mean variance $\frac{1}{N}$ i.i.d entries. The most complete results concerning the almost sure localization of the eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ are due to Bai-Silverstein [2, 3] and were established in the non Gaussian case. Spiked models were first proposed by Johnstone [20] in the context of (5) (matrix $\mathbf{H}_N$ is a diagonal matrix defined as a finite rank perturbation of the identity matrix). Later, Baik et al. [4] studied, in the complex Gaussian case, the almost sure convergence of the largest eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ and established central limit theorems. The analysis of [4] uses extensively the explicit form of the joint probability distribution of the entries of $\mathbf{\Sigma}_N$. Using the results of [2, 3] as well as the characterization of the support of the limiting distribution μ_N of the empirical eigenvalue distribution $\hat{\mu}_N$ (see Silverstein-Choi [27]), Baik-Silverstein [5] addressed the non Gaussian case, and showed the almost sure convergence of certain eigenvalues of $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$. Mestre considered in [21] the case where $\mathbf{H}_N \mathbf{H}_N^*$ has a finite number of different positive eigenvalues having multiplicities converging to $+\infty$, and showed how to estimate the eigenvalues of $\mathbf{H}_N \mathbf{H}_N^*$ as well as their associated eigenspace. Similar ideas were also developed in [22] in order to address the source localization problem in the context of large sensor arrays when the source signals are i.i.d. sequences. The analysis of Mestre [22, 21] is based on the results of [2, 3] as well as on the observation that it is possible to exhibit contours depending on the Stieljes transform of μ_N, and enclosing each eigenvalue of $\mathbf{H}_N \mathbf{H}_N^*$. Paul studied in [26] the behaviour of the eigenvectors associated to the greatest

eigenvalues of a Gaussian spiked model (almost sure convergence and central limit theorems). Bai and Yao showed in [1] that certain eigenvalues of a non Gaussian spiked model satisfy a central limit theorem. We finally note that the above results on zero-mean spiked models have been used in the context of source localization (see [19, 23]).

Concerning the information plus noise model. Except our paper [28] devoted to the source localization of deterministic sources, the almost sure location of the eigenvalues of matrix $\Sigma_N \Sigma_N^*$ was not studied previously. In [28], we however followed partly the work of Capitaine et al. [9], devoted to finite rank deformed Gaussian (or satisfying a Poincaré inequality) Wigner matrices, which was inspired by previous results of Haagerup and Thorbjornsen [17]. See also the recent paper [10] in which the rank of the deformation may scale with the size of the matrix. We used in [28] the same approach to prove that for N large enough, no eigenvalue of $\Sigma_N \Sigma_N^*$ is outside the support $\mathscr{S}_N$ of μ_N. In [28], under the assumption that the eigenvalue 0 of $\mathbf{B}_N \mathbf{B}_N^*$ is "far enough" from the others, we established a partial result showing that the $M - K$ smallest eigenvalues of $\Sigma_N \Sigma_N^*$ are almost surely separated from the others. In the present paper, we prove a general exact separation property extending the result of [5] to the complex Gaussian information plus noise model.

The almost sure behaviour (2), (3), of the largest eigenvalues of information plus noise spiked models appears to be a consequence of the general results of [6, 7] devoted to the analysis of certain random models with additive and/or multiplicative finite rank perturbation. (2) and (3) are therefore not new, but the technics of [7] completely differ from the approach used of the present paper which can be seen as an extension to the information plus noise model of the paper [5].

Organization of the paper In section 2, we review some results of [13] and [28] concerning the support $\mathscr{S}_N$ of μ_N as well as some useful background material. As [28] assumed $c_N < 1$, we address the case $c_N = 1$ and prove some extra results concerning the behaviour of the Stieltjes transform of μ_N around 0. In section 3, we prove the analog of the exact separation of [3]. [9] generalized the approach of [3] to prove this property in the finite rank deformed Wigner model. We however show that it is still possible to use again the ideas of [17]. We establish that it is sufficient to prove that the mass (w.r.t. μ_N) of any interval $\mathscr{I}$ of $\mathscr{S}_N$ is equal to the proportion of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ associated to $\mathscr{I}$. For this, we evaluate an integral along a certain contour enclosing the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ associated to $\mathscr{I}$. This contour is the analog of the contour introduced by [21] in the context of model (5) and was extensively used in [28]. Section 4 addresses the behaviour of the largest eigenvalues of an information plus noise spiked model. We analyse the support $\mathscr{S}_N$ of μ_N, which appears equivalent to evaluate the positive extrema of a certain rational function. Using results concerning perturbed third order polynomial equations, it is shown that if $\lambda_k \neq \sigma^2 \sqrt{c}$ for $k = 1, \ldots, K$, the intervals of $\mathscr{S}_N$ are $[\sigma^2 (1 - \sqrt{c_N})^2 + \mathcal{O}(1/M), \sigma^2 (1 + \sqrt{c_N})^2 + \mathcal{O}(1/M)]$ and $[\lambda_{k,N}^{-1}(\lambda_{k,N} + \sigma^2 c_N)(\lambda_{k,N} + \sigma^2) - \mathcal{O}^+(M^{-1/2}), \lambda_{k,N}^{-1}(\lambda_{k,N} + \sigma^2 c_N)(\lambda_{k,N} + \sigma^2) + \mathcal{O}^+(M^{-1/2})]$, where k is any index for which $\lambda_{k,N} > \sigma^2 \sqrt{c}$, and where $\mathcal{O}^+(M^{-1/2})$ represents a positive $\mathcal{O}(M^{-1/2})$ term. The results of section 3 imply immediately (2) and (3) when $\lambda_k \neq \sigma^2 \sqrt{c}$ for $k = 1, \ldots, K$. If one the $(\lambda_k)_{k=1,\ldots,K}$ is equal to $\sigma^2 \sqrt{c}$, we use an argument similar to Baik-Silverstein [5], which relies on an eigenvalue perturbation technic.

Model and assumptions We now summarize the model and assumptions which will be used in the paper, and introduce some definitions. Let $M, N, K \in \mathbb{N}^*$ such that $1 \leq K < M$, $K = K(N)$ and $M = M(N)$, functions of N with $c_N = \frac{M}{N} \to c > 0$ as $N \to \infty$. We consider a $M \times N$ random matrix

Σ_N defined as

$$\Sigma_N = \mathbf{B}_N + \sigma \mathbf{W}_N,$$

where $\sigma > 0$ and $\mathbf{B}_N$ and $\mathbf{W}_N$ satisfy the two following assumptions.

Assumption A-1: *Matrix $\mathbf{B}_N$ is deterministic and satisfies $\sup_N \|\mathbf{B}_N\| < +\infty$.*

Assumption A-2: *The entries of matrix $\mathbf{W}_N$ are i.i.d and follow a standard complex normal distribution $\mathscr{CN}(0, \frac{1}{N})$.*

Note that the Gaussian assumption **A-2** will be only required in section 3. All the results in section 2 concerning the convergence of the spectral distribution of $\Sigma_N \Sigma_N^*$ are also valid in the non Gaussian case. In the following, we study the context where

Assumption A-3: *$\mathbf{B}_N \mathbf{B}_N^*$ is rank deficient, and the non zero eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$ have multiplicity 1.*

The assumption on the multiplicities of the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ is not really necessary, but it allows to simplify the notations. We denote by K the rank of $\mathbf{B}_N \mathbf{B}_N^*$ (K may depend on N), and by $\lambda_{1,N} > \lambda_{2,N} > \ldots > \lambda_{K,N} > \lambda_{K+1,N} = \ldots = \lambda_{M,N} = 0$ its eigenvalues. We also assume that

Assumption A-4: *$c_N = \frac{M}{N} \leq 1$ for each N.*

This of course implies that $c \leq 1$. Assuming $c_N \leq 1$ does not introduce any restriction because if $c_N > 1$, the eigenvalues of $\Sigma_N \Sigma_N^*$ are 0 with multiplicity $M - N$ as well as the eigenvalues of matrix $\Sigma_N^* \Sigma_N$. The location of this set of eigenvalues can of course be deduced from the results related to $c_N < 1$.

In this paper, $\mathscr{C}_c^\infty(\mathbb{R}, \mathbb{R})$ will denote the set of infinitely differentiable functions with compact support, defined from $\mathbb{R}$ to $\mathbb{R}$. If $\mathscr{A} \subset \mathbb{R}$, $\partial \mathscr{A}$ and $\text{Int}(\mathscr{A})$ represent the boundary and the interior of $\mathscr{A}$ respectively.

We finally recall the definition and useful well known properties of the Stieltjes transform, a fundamental tool for the study of the eigenvalues of random matrices. Let μ be a positive finite measure on $\mathbb{R}$. We define its Stieltjes transform Ψ_μ as the function

$$\Psi_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z} \quad \forall z \in \mathbb{C} \backslash \text{supp}(\mu),$$

where $\text{supp}(\mu)$ represents the support of measure μ. We have the following well-known properties

Property 1. Ψ_μ *satisfies*

1. Ψ_μ *is holomorphic on $\mathbb{C} \backslash \text{supp}(\mu)$.*

2. $z \in \mathbb{C}^+$ *implies $\Psi_\mu(z) \in \mathbb{C}^+$.*

3. *If $\mu(\mathbb{R}_*^-) = 0$, then $z\Psi_\mu(z) \in \mathbb{C}^+$ if $z \in \mathbb{C}^+$.*

2 Characterization of the support $\mathscr{S}_N$ of measure μ_N

In this section, we recall some known results of [13] and [28] related to the support $\mathscr{S}_N$ of measure μ_N. As we assumed in [28] that $c_N < 1$, we also provide, when it is necessary, some details on the specific case $c_N = 1$.

2.1 Convergence of the empirical spectral measure $\hat{\mu}_N$ of $\Sigma_N\Sigma_N^*$ torward μ_N

We recall that $\hat{\mu}_N$ is defined by $\hat{\mu}_N = M^{-1}\sum_{i=1}^M \delta_{\hat{\lambda}_{i,N}}$. Its Stieltjes transform $\hat{m}_N$ is given, for all $z \in \mathbb{C}\backslash\{\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M,N}\}$, by

$$\hat{m}_N(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\hat{\mu}_N(\lambda)}{\lambda - z}.$$

The following result, concerning the convergence of $\hat{m}_N(z)$ can be found in [14, Th.1.1], [16, Th.7.4] (see also [18, Th.2.5] for a more general model).

Theorem 1. *It exists a deterministic probability measure μ_N, such that $\hat{\mu}_N - \mu_N \xrightarrow{\mathscr{D}} 0$ as $N \to \infty$ with probability one. Equivalently, the Stieltjes transform m_N of μ_N satisfies $\hat{m}_N(z) - m_N(z) \to 0$ almost surely $\forall z \in \mathbb{C}\backslash\mathbb{R}^+$. Moreover, $\forall z \in \mathbb{C}\backslash\mathbb{R}^+$, $m_N(z)$ is the unique solution of the equation,*

$$m_N(z) = \frac{1}{M}\mathrm{Tr}\left[-z(1 + \sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1 - c_N)\mathbf{I}_M + \frac{\mathbf{B}_N\mathbf{B}_N^*}{1 + \sigma^2 c_N m_N(z)}\right]^{-1} \tag{6}$$

satisfying $\mathrm{Im}(zm_N(z)) > 0$ for $z \in \mathbb{C}^+$.

The behaviour of the Stieltjes transform m_N around the real axis is fundamental to evaluate the support $\mathscr{S}_N$ of μ_N. The following theorem is essentially due to [13].

Theorem 2. 1. *If $c_N < 1$, the limit of $m_N(z)$, as $z \in \mathbb{C}^+$ converges to x, exists for each $x \in \mathbb{R}$ and is still denoted by $m_N(x)$. If $c_N = 1$, the limit exists for $x \neq 0$. The function $x \to m_N(x)$ is continuous on $\mathbb{R}$ if $c_N < 1$ and on $\mathbb{R}^*$ if $c_N = 1$. It is also continuously differentiable on $\mathbb{R}\backslash\partial\mathscr{S}_N$.*

2. *If $c_N < 1$, then $\mathrm{Re}(1 + \sigma^2 c_N m_N(z)) \geq 1/2$ for each $z \in \mathbb{C}^+ \cup \mathbb{R}$, and if $c_N = 1$, this inequality holds on $\mathbb{C}^+ \cup \mathbb{R}^*$.*

3. *$m_N(x)$ is a solution of (6) for $x \in \mathbb{R}\backslash\partial\mathscr{S}_N$.*

4. *Measure μ_N is absolutely continuous and its density is given by $f_{\mu_N}(x) = \pi^{-1}\mathrm{Im}(m_N(x))$.*

The statements of this theorem are essentially contained in [13, Th.2.5] (see also [28] for more details), except item 2 because it is shown in [13, Lem.2.1] that $\mathrm{Re}(1 + \sigma^2 c_N m_N(z)) \geq 0$. We therefore prove item 2 in the Appendix A.

We note that as m_N is a Stietljes transform, it also satisfies $m_N(z^*) = m_N(z)^*$. Therefore, it holds that $m_N(z) \to m_N(x)^*$ as $z \in \mathbb{C}^- \to x$, for $x \in \mathbb{R}$ if $c_N < 1$ and for $x \in \mathbb{R}^*$ if $c_N = 1$.

In the following, we denote by f_N, ϕ_N and w_N the functions defined by

$$f_N(w) = \frac{1}{M} \mathrm{Tr}\left(\mathbf{B}_N \mathbf{B}_N^* - w \mathbf{I}_M\right)^{-1},$$

$$\phi_N(w) = w\left(1 - \sigma^2 c_N f_N(w)\right)^2 + \sigma^2 (1 - c_N)\left(1 - \sigma^2 c_N f_N(w)\right), \tag{7}$$

$$w_N(z) = z(1 + \sigma^2 c_N m_N(z))^2 - \sigma^2(1 - c_N)(1 + \sigma^2 c_N m_N(z)).$$

Functions w_N and ϕ_N are of crucial importance because, as shown in [28], the interior of $\mathscr{S}_N$ is given by $\mathrm{Int}(\mathscr{S}_N) = \{x > 0, \mathrm{Im}(w_N(x)) > 0\}$ and for each $x \in \mathbb{R} \backslash \partial \mathscr{S}_N$, $w_N(x)$ is a solution of the equation $\phi_N(w) = x$. The characterization of $\mathscr{S}_N$ proposed in [28], based on a reformulation of the results in [13, Th.3.2, Th.3.3], consists in identifying $w_N(x)$ out of the set of solutions of $\phi_N(w) = x$.

We also note that (6) is equivalent to

$$\frac{m_N(z)}{1 + \sigma^2 c_N m_N(z)} = f_N(w_N(z)), \tag{8}$$

and that the identity

$$\frac{1}{1 + \sigma^2 c_N m_N(z)} = 1 - \sigma^2 c_N f_N(w_N(z)) \tag{9}$$

holds for $z \in \mathbb{C}^+ \cup \mathbb{R}$ if $c_N < 1$, or for $z \in \mathbb{C}^+ \cup \mathbb{R}^*$ if $c_N = 1$.

2.2 Properties of ϕ_N and w_N, and characterization of $\mathscr{S}_N$

In this paragraph, we recall the main properties of functions ϕ_N and w_N, as well the structure of $\mathscr{S}_N$. Lemmas 1, 2 as well as theorem 3 are proved in [28, Prop.3, Th.2] for $c_N < 1$, but the derivations for $c_N = 1$ are similar, except items 6 and 8 of lemma 2.

Lemma 1.   1. *The function ϕ_N admits $2Q_N$ non-negative local extrema counting multiplicities (with $1 \leq Q_N \leq K+1$) whose preimages are denoted $w_{1,N}^- < 0 < w_{1,N}^+ \leq w_{2,N}^- \ldots \leq w_{Q_N,N}^- < w_{Q_N,N}^+$.*

  2. *Define $x_{q,N}^- = \phi_N(w_{q,N}^-)$ and $x_{q,N}^+ = \phi_N(w_{q,N}^+)$ for $q = 1 \ldots Q_N$. Then,*

  $$x_{1,N}^- < x_{1,N}^+ \leq x_{2,N}^- < \ldots \leq x_{Q_N,N}^- < x_{Q_N,N}^+,$$

  and $x_{1,N}^- > 0$ if $c_N < 1$ while $x_{1,N}^- = 0$ if $c_N = 1$.

  3. *For $q = 1, \ldots, Q_N$, each interval $(w_{q,N}^-, w_{q,N}^+)$ contains at least one element of the set of eigenvalues $\{\lambda_{1,N}, \ldots, \lambda_{K,N}, 0\}$ and each eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$ belongs to one of these intervals.*

  4. *ϕ_N is increasing on the intervals $(-\infty, w_{1,N}^-], [w_{1,N}^+, w_{2,N}^-], \ldots, [w_{Q_N-1,N}^+, w_{Q_N,N}^-], [w_{Q_N,N}^+, +\infty)$, and moreover*

  $$\phi_N\left((-\infty, w_{1,N}^-]\right) = (-\infty, x_{1,N}^-],$$

  $$\phi_N\left(\left[w_{q,N}^+, w_{q+1,N}^-\right]\right) = \left[x_{q,N}^+, x_{q+1,N}^-\right] \quad \text{for each } q = 1, \ldots, Q_N - 1,$$

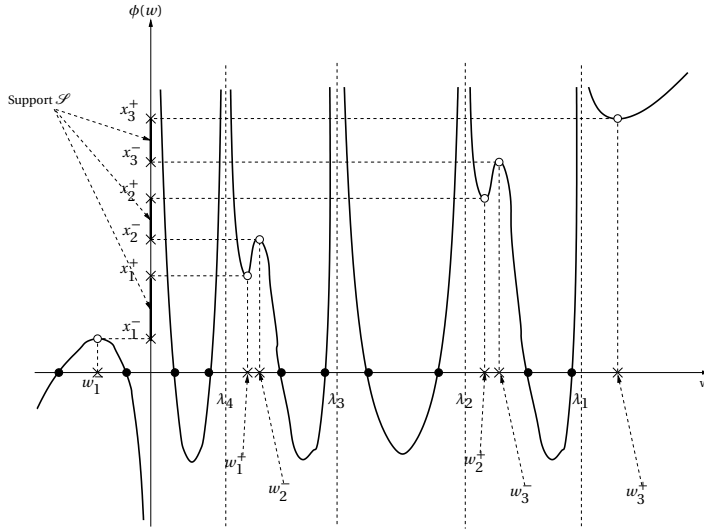 $$\phi_N\left([w_{Q_N,N}^+, +\infty)\right) = [x_{Q_N,N}^+, +\infty).$$

Figure 1: Function ϕ for $K = 4$ and $c < 1$. Here $Q = 3$.

In figure 1, we give a typical representation of function ϕ_N. We are now in position to recall the characterization of $\mathscr{S}_N$ presented in [28, Th.2].

Theorem 3. *The support $\mathscr{S}_N$ is given by*

$$\mathscr{S}_N = \bigcup_{q=1}^{Q_N} \left[x_{q,N}^-, x_{q,N}^+ \right],$$

with $x_{1,N}^- = 0$ if $c_N = 1$ and $x_{1,N}^- > 0$ if $c_N < 1$.

The intervals $([x_{q,N}^-, x_{q,N}^+])_{q=1,\dots,Q_N}$ will be called the clusters of $\mathscr{S}_N$. Cluster $[x_{q,N}^-, x_{q,N}^+]$ corresponds to the interval $[w_{q,N}^-, w_{q,N}^+]$ in the sense that $x_{q,N}^- = \phi_N(w_{q,N}^-)$ and $x_{q,N}^+ = \phi_N(w_{q,N}^+)$. Finally, we shall say that an eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N \mathbf{B}_N^*$ is associated to cluster $[x_{q,N}^-, x_{q,N}^+]$ if $\lambda_{k,N} \in (w_{q,N}^-, w_{q,N}^+)$.

In the same way as in theorem 2, we set $w_N(x) = \lim_{z \in \mathbb{C}^+, z \to x} w_N(z)$ for $x \in \mathbb{R}$ if $c_N < 1$ and for $x \in \mathbb{R}^*$ if $c_N = 1$. We notice that $\lim_{z \in \mathbb{C}^-, z \to x} w_N(z) = w_N(x)^*$. Function $x \to w_N(x)$ satisfies the following properties.

Lemma 2. *The following properties hold*

1. *$x \to w_N(x)$ is continuous on $\mathbb{R}$ if $c_N < 1$ and on $\mathbb{R}^*$ if $c_N = 1$, and is continuously differentiable on $\mathbb{R} \backslash \partial \mathscr{S}_N$.*

2. *w_N is real and increasing on $\mathbb{R} \backslash \mathscr{S}_N$.*

3. *$1 - \sigma^2 c_N f_N(w_N(x)) \neq 0$ for $x \in \mathbb{R} \backslash \partial \mathscr{S}_N$.*

4. *$x \in \text{Int}(\mathscr{S}_N)$ if and only if $w_N(x) \in \mathbb{C}^+$.*

5. *For $x \in \mathbb{R} \backslash \partial \mathscr{S}_N$, $w_N(x)$ is a solution of the equation $\phi_N(w) = x$. If $x \in \text{Int}(\mathscr{S}_N)$, $w_N(x)$ is the unique solution belonging to $\mathbb{C}^+$ and if $x \in \mathscr{S}_N^c$, $w_N(x)$ is the unique solution satisfying $\phi_N'(w_N(x)) > 0$ and $1 - \sigma^2 c_N f_N(w_N(x)) > 0$.*

1941

6. Function $x \to w_N(x)$ is continuous at $x = x_{1,N}^- = 0$ for $c_N = 1$.

7. For $q = 1, \ldots, Q_N$, $w_N(x_{q,N}^-) = w_{q,N}^-$ and $w_N(x_{q,N}^+) = w_{q,N}^+$.

8. Let $q = 1, \ldots, Q_N$. Then, there exists a constant $C > 0$ and neighborhoods $\mathcal{V}(x_{q,N}^-)$, $\mathcal{V}(x_{q,N}^+)$ of respectively $x_{q,N}^-$ and $x_{q,N}^+$ such that,

$$|w_N'(x)| \le C \left| x - x_{q,N}^- \right|^{-1/2} \quad \forall x \in \mathcal{V}(x_{q,N}^-) \cap \mathbb{R} \backslash \{x_{q,N}^-\}, \tag{10}$$

$$|w_N'(x)| \le C \left| x - x_{q,N}^+ \right|^{-1/2} \quad \forall x \in \mathcal{V}(x_{q,N}^+) \cap \mathbb{R} \backslash \{x_{q,N}^+\}, . \tag{11}$$

The lemma was proved in [28, Prop.2, Lem.3] in the case $c_N < 1$. The proofs extend easily to $c_N = 1$, except items 6 and 8 for $q = 1$. These 2 statements are proved in the Appendix B.

We finish this section by showing that the following result holds.

Corollary 1. We have

$$\sup_N x_{Q_N,N}^+ < \infty,$$

i.e. $\cup_N \mathscr{S}_N$ is a bounded set.

Proof: We define λ_{max} by $\lambda_{max} = \sup_N \|\mathbf{B}_N\|^2$. It follows that for $w > \lambda_{max}$,

$$\sup_N |f_N(w)| \le \frac{1}{|\lambda_{max} - w|},$$

$$\sup_N |f_N'(w)| \le \frac{1}{|\lambda_{max} - w|^2},$$

$$\sup_N |w f_N'(w)| \le \frac{w}{|\lambda_{max} - w|^2},$$

and since $\phi_N'(w) = (1 - \sigma^2 c_N f_N(w))^2 - 2\sigma^2 c_N w f_N'(w)(1 - \sigma^2 c_N f_N(w)) - \sigma^4 c_N(1 - c_N) f_N'(w)$ converges towards 1 when $w \to +\infty$, we deduce that for $\epsilon > 0$, $\exists w_\epsilon > \lambda_{max}$ such that $\forall w > w_\epsilon$, $\phi_N'(w) > \epsilon$ for all N. Since $\phi_N'(w_{Q_N,N}^+) = 0$, this implies that

$$\sup_N w_{Q_N,N}^+ \le w_\epsilon < +\infty.$$

Moreover, using $w_{Q_N,N}^+ = w_N(x_{Q_N,N}^+) = x_{Q_N,N}^+(1 + \sigma^2 c_N m_N(x_{Q_N,N}^+))^2 - \sigma^2(1 - c_N)(1 + \sigma^2 c_N m_N(x_{Q_N,N}^+))$, and item 2 of theorem 2, we get that

$$x_{Q_N,N}^+ \le \frac{w_\epsilon}{(1 + \sigma^2 c_N m_N(x_{Q_N,N}^+))^2} + \frac{\sigma^2(1 - c_N)}{1 + \sigma^2 c_N m_N(x_{Q_N,N}^+)} < 4w_\epsilon + 2\sigma^2.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3 Almost sure location of the sample eigenvalues.

We first recall the following result of [28, Th.3], which states the almost sure absence of eigenvalue of $\Sigma_N \Sigma_N^*$ outside the support $\mathscr{S}_N$ of μ_N for all large N. This property is well-known in the context of zero mean non Gaussian correlated matrices (see [2]). We note that the proof of theorem [28, Th.3] uses extensively that $\mathbf{W}_N$ is Gaussian (assumption **A-2**).

Theorem 4. *Let $a, b \in \mathbb{R}$, $\epsilon > 0$ and $N_0 \in \mathbb{N}$ such that $(a - \epsilon, b + \epsilon) \cap \mathscr{S}_N = \emptyset$ for each $N > N_0$. Then, with probability one, no eigenvalue of $\Sigma_N \Sigma_N^*$ belongs to $[a, b]$ for N large enough.*

We remark that theorem 4 extends to semi-infinite intervals $[b, +\infty)$ because, as $\|\mathbf{W}_N \mathbf{W}_N^*\| \to (1 + \sqrt{c})^2$ almost surely, then it holds that $\hat{\lambda}_{1,N} = \|\Sigma_N \Sigma_N^*\| \leq \sup_N \|\mathbf{B}_N \mathbf{B}_N^*\| + 2\sigma^2 (1 + \sqrt{c})^2$ almost surely for N large enough.

In order to interpret this result, assume that for each $N > N_1 \geq N_0$, the number of clusters of $\mathscr{S}_N$ does not depend on N (denote Q the number of clusters), and that for each $q = 1, \ldots, Q$, the sequences $(x_{q,N}^-)_{N > N_1}$ and $(x_{q,N}^+)_{N > N_1}$ converge torwards limits x_q^- and x_q^+ satisfying $x_1^- \leq x_1^+ < x_2^- \leq x_2^+ < \ldots < x_Q^- \leq x_Q^+$. In this context, theorem 4 implies that almost surely, for each $\epsilon > 0$, each eigenvalue belongs to one of the intervals $[x_q^- - \epsilon, x_q^+ + \epsilon]$ for N large enough.

We now establish the following property, also well-know in the literature and referred to as "exact separation" (see e.g. [3] in the context of non Gaussian correlated zero mean random matrices).

Theorem 5. *Let $a, b \in \mathbb{R}$, $\epsilon > 0$, $N_0 \in \mathbb{N}$ such that $(a - \epsilon, b + \epsilon) \cap \mathscr{S}_N = \emptyset$ for $N > N_0$. Then, with probability one,*

$$\mathrm{card}\{k : \hat{\lambda}_{k,N} < a\} = \mathrm{card}\{k : \lambda_{k,N} < w_N(a)\} \tag{12}$$
$$\mathrm{card}\{k : \hat{\lambda}_{k,N} > b\} = \mathrm{card}\{k : \lambda_{k,N} > w_N(b)\}$$

for N large enough.

Under the above simplified assumptions, this result means that almost surely for N large enough, the number of sample eigenvalues that belong to each interval $[x_q^- - \epsilon, x_q^+ + \epsilon]$ coincides with the number of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ that are associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$. To prove theorem 5, we use the same technic as in [28], where a less general result is presented in the case $c < 1$.

3.1 Preliminary results

We first need to state preliminary useful lemmas. The first lemma is elementary and is related to the solutions of the equation $1 - \sigma^2 c_N f_N(w) = 0$.

Lemma 3. *The equation $1 - \sigma^2 c_N f_N(w) = 0$ admits $K + 1$ real solutions $z_{0,N} < 0 < z_{1,N} < \lambda_{1,N} < \ldots < z_{K,N} < \lambda_{K,N}$. If $c_N < 1$, $z_{0,N} < w_{1,N}^-$ while if $c_N = 1$, $z_{0,N} = w_{1,N}^-$. Moreover, for each $k = 1, \ldots, K$, each solution $z_{k,N}$ belongs to the interval $(w_{q,N}^-, w_{q,N}^+)$ containing eigenvalue $\lambda_{k,N}$, with $q \in \{1, \ldots, Q_N\}$.*

The next two lemmas are fundamental, and were proved by Haagerup-Thorbjornsen in [17] in the Wigner case models (see also [8]). Lemma 4 and 5 are established in [28, Prop. 4, Lem. 2 and proof of Th.3]. Note that, unlike section 2, the Gaussian assumption is required here. We give here some insights on the proof of these two lemmas for the reader's convenience.

Lemma 4. *Let $\psi \in \mathscr{C}_c^{\infty}(\mathbb{R}, \mathbb{R})$, independent of N, then*

$$\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\psi\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right)\right] - \int_{\mathscr{S}_N}\psi(\lambda)\mathrm{d}\mu_N(\lambda) = \mathscr{O}\left(\frac{1}{N^2}\right).$$

Proof: Using the integration by part formula (see e.g. [24], [25]) and the Poincaré inequality for Gaussian random vectors [12], it is proved in [28, Prop.4] that

$$\mathbb{E}\left[\hat{m}_N(z)\right] = m_N(z) + \frac{\chi_N(z)}{N^2}, \tag{13}$$

where χ_N is holomorphic on $\mathbb{C}\backslash\mathbb{R}$ and satisfies

$$\left|\chi_N(z)\right| \le P_1\left(|z|\right)P_2\left(\left|\operatorname{Im}(z)^{-1}\right|\right), \tag{14}$$

with P_1, P_2 two polynomials with positive coefficients independent of N, z. The Stieltjes inversion formula gives

$$\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\psi\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right)\right] = \frac{1}{\pi}\lim_{y\downarrow 0}\operatorname{Im}\left(\int_{\mathbb{R}}\psi(x)\mathbb{E}\left[\hat{m}_N(x+iy)\right]\mathrm{d}x\right), \tag{15}$$

as well as $\int_{\mathbb{R}}\psi(\lambda)\mathrm{d}\mu_N(\lambda) = \pi^{-1}\lim_{y\downarrow 0}\operatorname{Im}\left(\int_{\mathbb{R}}\psi(x)m_N(x+iy)\mathrm{d}x\right)$. The polynomial bound (14) implies the bound $\limsup_{y\downarrow 0}\int_{\mathbb{R}}\psi(x)\left|\chi_N(x+iy)\right|\mathrm{d}x \le C < \infty$, with $C > 0$ independent of N (a result shown in [8, Sec.3.3] using the ideas of [17]). Plugging (13) into (15), we obtain the desired result. $\qquad\square$

Lemma 5 is not explicitly stated in [28], but it can be proved easily using the derivation of [28, eq. (37)].

Lemma 5. *Let $\psi \in \mathscr{C}_c^{\infty}(\mathbb{R}, \mathbb{R})$, independent of N and constant on each cluster of $\mathscr{S}_N$ for N large enough. Then, we have*

$$\operatorname{Var}\left[\frac{1}{M}\operatorname{Tr}\psi\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right)\right] = \mathscr{O}\left(\frac{1}{N^4}\right).$$

Proof: We only give a sketch of proof for the reader's convenience. Using the Poincaré inequality for gaussian random vectors, we obtain

$$\operatorname{Var}\left[\frac{1}{M}\operatorname{Tr}\psi\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right)\right] \le \frac{C}{N^2}\mathbb{E}\left[\frac{1}{M}\operatorname{Tr}\psi'\left(\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right)^2\boldsymbol{\Sigma}_N\boldsymbol{\Sigma}_N^*\right] = \frac{C}{N^2}\left(\int_{\mathbb{R}}\lambda\psi'(\lambda)^2\mathrm{d}\mu_N(\lambda) + \mathscr{O}\left(\frac{1}{N^2}\right)\right),$$

where the last equality follows from the application of lemma 4 to the function $\lambda \mapsto \lambda\psi'(\lambda)^2$. The conclusion follows from the observation that this function vanishes on $\mathscr{S}_N$ for all large N. $\qquad\square$

We are now in position to prove theorem 5.

3.2 End of the proof

We first prove (12) and assume that $a > 0$ because (12) is obvious if $a \leq 0$. We consider $\eta < \epsilon$ and assume without restriction that $0 < \eta < a$. We consider a function $\psi_a \in \mathscr{C}_c^\infty(\mathbb{R}, \mathbb{R})$, independent of N, such that $\psi_a \in [0, 1]$ and

$$\psi_a(\lambda) = \begin{cases} 1 & \forall \lambda \in [0, a - \eta] \\ 0 & \forall \lambda \geq a. \end{cases}$$

By lemma 4, we have

$$\mathbb{E}\left[\frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) \right] - \int_{\mathbb{R}^+} \psi_a(\lambda) \mathrm{d}\mu_N(\lambda) = \mathcal{O}\left(\frac{1}{N^2} \right),$$

or equivalently

$$\mathbb{E}\left[\frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) \right] = \mu_N \left([0, a - \eta] \right) + \mathcal{O}\left(\frac{1}{N^2} \right).$$

Lemma 5 also implies that

$$\operatorname{Var}\left[\frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) \right] = \mathcal{O}\left(\frac{1}{N^4} \right).$$

Therefore, Markov inequality leads to

$$\mathbb{P}\left(\left| \frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) - \mu_N \left([0, a - \eta] \right) \right| > \frac{1}{N^{4/3}} \right)$$

$$\leq N^{8/3} \operatorname{Var}\left[\frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) \right] + N^{8/3} \left| \mathbb{E}\left[\frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) - \mu_N \left([0, a - \eta] \right) \right] \right|^2$$

$$= \mathcal{O}\left(\frac{1}{N^{4/3}} \right),$$

which implies that with probability one,

$$\frac{1}{M} \operatorname{Tr} \psi_a \left(\Sigma_N \Sigma_N^* \right) = \mu_N \left([0, a - \eta] \right) + \mathcal{O}\left(\frac{1}{N^{4/3}} \right). \tag{16}$$

The remainder of the proof is dedicated to the evaluation of $\mu_N([0, a - \eta])$. Let $I_N = \max\{q : x_{q,N}^+ < a\}$. It is clear that $\mu_N([0, a - \eta]) = \sum_{q=1}^{I_N} \mu_N([x_{q,N}^-, x_{q,N}^+])$ because $\mu_N((a - \eta, a)) = 0$. By theorem 2, μ_N is absolutely continuous with density $\pi^{-1} \operatorname{Im}(m_N(x))$. Therefore, it holds that

$$\mu_N([x_{q,N}^-, x_{q,N}^+]) = \frac{1}{\pi} \operatorname{Im}\left(\int_{x_{q,N}^-}^{x_{q,N}^+} m_N(x) \mathrm{d}x \right). \tag{17}$$

In order to evaluate the righthandside of (17), we use the contour integral approach introduced in [28]. For this, we consider the curve $\mathscr{C}_{q,N}$ defined by

$$\mathscr{C}_{q,N} = \left\{ w_N(x) : x \in [x_{q,N}^-, x_{q,N}^+] \right\} \cup \left\{ w_N(x)^* : x \in [x_{q,N}^-, x_{q,N}^+] \right\}.$$

We notice that $x \to w_N(x)$ (resp. $x \to w_N(x)^*$) is a one-to-one correspondance from $(x_{q,N}^-, x_{q,N}^+)$ onto $\{w_N(x), x \in (x_{q,N}^-, x_{q,N}^+)\}$ (resp. $\{w_N(x)^*, x \in (x_{q,N}^-, x_{q,N}^+)\}$) because if $w_N(x) = w_N(y)$, then $\phi_N(w_N(x)) = x = \phi_N(w_N(y)) = y$ (see lemma 2, item 5).

It follows from lemma 2 items 1, 4 and 7 that $\mathscr{C}_{q,N}$ is a closed continuous contour enclosing the interval $(w_{q,N}^-, w_{q,N}^+)$. $\mathscr{C}_{q,N}$ is differentiable at each point except at $w_{q,N}^-$ and $w_{q,N}^+$ (see item 8 of lemma 2). However, (10) and (11) imply that $|w_N'|$ is summable on $[x_{q,N}^-, x_{q,N}^+]$. Therefore, for each function g continuous in a neighborhood of $\mathscr{C}_{q,N}$, satisfying $(g(w))^* = g(w^*)$, it is still possible to define the contour integral $\oint_{\mathscr{C}_{q,N}^-} g(w)dw$ by

$$\oint_{\mathscr{C}_{q,N}^-} g(w)dw = 2i\mathrm{Im}\left(\int_{x_{q,N}^-}^{x_{q,N}^+} g(w_N(x))w_N'(x)dx\right).$$

The notation $\mathscr{C}_{q,N}^-$ means that the contour $\mathscr{C}_{q,N}$ is clockwise oriented. Although $\mathscr{C}_{q,N}$ is not differentiable, the main results related to contour integrals of meromorphic functions remain valid. In particular, it holds that

$$\mathrm{Ind}_{\mathscr{C}_{q,N}^-}(\xi) = \frac{1}{2\pi i}\int_{\mathscr{C}_{q,N}^-} \frac{d\lambda}{\xi - \lambda} = \begin{cases} 1 & \text{if } \xi \in \left(w_{q,N}^-, w_{q,N}^+\right) \\ 0 & \text{if } \xi \notin \left[w_{q,N}^-, w_{q,N}^+\right] \end{cases}$$

In order to evaluate the righthandside of (17) using a contour integral, we remark that

$$m_N(x) = \frac{f_N(w_N(x))}{1 - \sigma_N^2 c_N f_N(w_N(x))} \quad \forall x \in \mathbb{R}\backslash\partial\mathscr{S}_N$$

(see (8) and item 3 of lemma 2). Moreover, by item 5 of lemma 2, we have $w_N'(x)\phi_N'(w_N(x)) = 1$ on $(x_{q,N}^-, x_{q,N}^+)$. Therefore, we have

$$\mu_N([x_{q,N}^-, x_{q,N}^+]) = \frac{1}{\pi}\mathrm{Im}\left(\int_{x_{q,N}^-}^{x_{q,N}^+} g_N(w_N(x))w_N'(x)dx\right), \tag{18}$$

where $g_N(w)$ is the rational function defined by

$$g_N(w) = \frac{f_N(w)\phi_N'(w)}{1 - \sigma_N^2 c_N f_N(w)} = f_N(w)\frac{(1 - \sigma^2 c_N f_N(w))^2 - 2\sigma_N^2 c_N w f_N'(w)(1 - \sigma_N^2 c_N f_N(w)) - \sigma_N^4 c_N(1 - c_N)f_N'(w)}{1 - \sigma_N^2 c_N f_N(w)}.$$

In order to justify the existence of the integral at the righthandside of (18), we prove that $g_N(w)$ is continuous in a neighborhood of $\mathscr{C}_{q,N}$. We first note that the poles of $g_N(w)$ coincide with the eigenvalues of $\mathbf{B}_N\mathbf{B}_N^*$ and the zeros $(z_{k,N})_{k=0,\ldots,K}$ of $1 - \sigma_N^2 c_N f_N(w)$. As $w_N(x)$ is not real on $(x_{q,N}^-, x_{q,N}^+)$, $x \to g_N(w_N(x))$ is continuous on $(x_{q,N}^-, x_{q,N}^+)$. It remains to check the continuity at $x_{q,N}^-$ and $x_{q,N}^+$. If $c_N < 1$, $w_{q,N}^- = w_N(x_{q,N}^-)$ and $w_{q,N}^+ = w_N(x_{q,N}^+)$ do not coincide with one the poles of $g_N(w)$. If $c_N = 1$ and $q = 1$, this property still holds true except for $w_{1,N}^- = w_N(x_{1,N}^-) = w_N(0)$ because $z_{0,N} = w_{1,N}^-$ (see lemma 3). However, if $c_N = 1$, the solutions of $1 - \sigma_N^2 c_N f_N(w)$ are not poles of g_N due to a pole zero cancellation.

Therefore, it is clear that $\mu_N([x_{q,N}^-, x_{q,N}^+])$ can also be written as

$$\mu_N([x_{q,N}^-, x_{q,N}^+]) = \frac{1}{2\pi i} \oint_{\mathscr{C}_{q,N}^-} g_N(\lambda) d\lambda.$$

The integral can be evaluated using residue theorem and we give here the main steps of calculation. Define $\mathscr{I}_q = \{k \in \{1, 2, \ldots, K\} : \lambda_{k,N} \in (w_{q,N}^-, w_{q,N}^+)\}$ and $L_q = \mathrm{card}(\mathscr{I}_q) > 0$ ($L_q > 0$ from lemma 1 item 3). Assume $c_N < 1$. Since $\mathscr{C}_{q,N}$ only encloses $(w_{q,N}^-, w_{q,N}^+)$, we will have residues at the following points:

- for $q = 1$: residues at $z_{0,N}$, 0 and $z_{k,N}, \lambda_{k,N}$ for $k \in \mathscr{I}_1$.

- for $q \geq 2$: residues at $z_{k,N}, \lambda_{k,N}$ for $k \in \mathscr{I}_q$.

If $c_N = 1$, the zeros of $1 - \sigma^2 c_N f_N(w)$ are not poles of $g_N(w)$:

- for $q = 1$: residues at 0 and $\lambda_{k,N}$ for $k \in \mathscr{I}_1$.

- for $q \geq 2$: residues at $\lambda_{k,N}$ for $k \in \mathscr{I}_q$.

We just consider the case $c_N < 1$ in the following (the calculations are similar for $c_N = 1$ and are therefore omitted). We consider the decomposition $g_N(\lambda) = g_{1,N}(\lambda) + g_{2,N}(\lambda) + g_{3,N}(\lambda)$, with

$$g_{1,N}(\lambda) = f_N(\lambda) \left(1 - \sigma^2 c_N f_N(\lambda)\right),$$
$$g_{2,N}(\lambda) = -2\sigma^2 c_N \lambda f_N(\lambda) f_N'(\lambda),$$
$$g_{3,N}(\lambda) = -\sigma^4 c_N (1 - c_N) \frac{f_N(\lambda) f_N'(\lambda)}{1 - \sigma^2 c_N f_N(\lambda)}.$$

These three functions admit poles at $0, \left(\lambda_{k,N}\right)_{k=1,\ldots,K}$, and $g_{3,N}$ has moreover poles at $(z_{k,N})_{k=0,\ldots,K}$. After tedious but straightforward calculations, we finally find that for $k \in \{1, 2, \ldots, K\}$,

$$\mathrm{Res}\left(g_{1,N}, \lambda_{k,N}\right) = -\frac{1}{M} + \frac{2\sigma^2 c_N}{M^2} \sum_{l \neq k} \frac{1}{\lambda_{l,N} - \lambda_{k,N}},$$

$$\mathrm{Res}\left(g_{2,N}, \lambda_{k,N}\right) = -\frac{2\sigma^2 c_N}{M^2} \sum_{l \neq k} \frac{1}{\lambda_{l,N} - \lambda_{k,N}},$$

$$\mathrm{Res}\left(g_{3,N}, \lambda_{k,N}\right) = -\frac{1 - c_N}{c_N}.$$

For the residues at 0, we get

$$\mathrm{Res}\left(g_{1,N}, 0\right) = -\frac{M - K}{M} + 2\sigma^2 c_N \frac{M - K}{M} \frac{1}{M} \sum_{l=1}^{K} \frac{1}{\lambda_{l,N}},$$

$$\mathrm{Res}\left(g_{2,N}, 0\right) = -2\sigma^2 c_N \frac{M - K}{M} \frac{1}{M} \sum_{l=1}^{K} \frac{1}{\lambda_{l,N}},$$

$$\mathrm{Res}\left(g_{3,N}, 0\right) = -\frac{1 - c_N}{c_N}.$$

Finally, the residues at $z_{k,N}$ for $k = 0, \ldots, K$ are given by $\text{Res}(g_{3,N}, z_{k,N}) = \frac{1-c_N}{c_N}$. Using these evaluations, we obtain immediately that if $q \geq 2$, then,

$$\mu_N([x_{q,N}^-, x_{q,N}^+]) = -\sum_{k \in \mathscr{I}_q} \left[\text{Res}\left(g_{1,N}, \lambda_{k,N}\right) + \text{Res}\left(g_{2,N}, \lambda_{k,N}\right) + \text{Res}\left(g_{3,N}, \lambda_{k,N}\right) + \text{Res}\left(g_{3,N}, z_{k,N}\right) \right] = \frac{L_q}{M}.$$

This coincides with the ratio of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ associated to the cluster $[x_{q,N}^-, x_{q,N}^+]$ (i.e. the eigenvalues $\lambda_{k,N}$ for $k \in \mathscr{I}_q$). If $q = 1$,

$$\begin{aligned}
\mu_N([x_{1,N}^-, x_{1,N}^+]) &= -\sum_{k \in \mathscr{I}_1} \left[\text{Res}\left(g_{1,N}, \lambda_{k,N}\right) + \text{Res}\left(g_{2,N}, \lambda_{k,N}\right) + \text{Res}\left(g_{3,N}, \lambda_{k,N}\right) + \text{Res}\left(g_{3,N}, z_{k,N}\right) \right] \\
&\quad - \left[\text{Res}\left(g_{1,N}, 0\right) + \text{Res}\left(g_{2,N}, 0\right) + \text{Res}\left(g_{3,N}, 0\right) + \text{Res}\left(g_{3,N}, z_{0,N},\right) \right] \\
&= \frac{L_1}{M} + \frac{M-K}{M},
\end{aligned}$$

which also coincides with the ratio of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ associated to the cluster $[x_{1,N}^-, x_{1,N}^+]$ (the $\lambda_{k,N}$ for $k \in \mathscr{I}_1$ and 0 with multiplicity $M - K$).

Therefore, using (16), we get that

$$\text{Tr}\,\psi_a\left(\Sigma_N \Sigma_N^*\right) - \left(\sum_{q=1}^{I_N} L_q + (M-K) \right) = \mathcal{O}\left(\frac{1}{N^{1/3}} \right). \tag{19}$$

But almost surely, for N large enough, $\text{Tr}\,\psi_a(\Sigma_N \Sigma_N^*)$ is exactly the number of eigenvalues contained in $[0, a]$ because no eigenvalue of $\Sigma_N \Sigma_N^*$ belong to $[a - \eta, a]$ (use theorem 4 with $a - \eta$ in place of a). The left handside of (19) is thus an integer. Since this integer decreases at rate $N^{-1/3}$, it is equal to zero for N large enough. (12) follows from the observation that $\sum_{q=1}^{I_N} L_q + M - K$ is equal to the number of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ that are less than $w_N(a)$.

To evaluate the number of eigenvalues in interval $(b, +\infty)$, we use that no eigenvalue belongs to $[a, b]$ (theorem 4). Therefore,

$$\text{card}\{k : \hat{\lambda}_{k,N} > b\} = M - \text{card}\{k : \hat{\lambda}_{k,N} < a\}.$$

(12) implies that

$$\text{card}\{k : \hat{\lambda}_{k,N} > b\} = M - \sum_{q=1}^{I_N} L_q - (M-K),$$

which coincides with the number of eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ in interval $(w_N(b), +\infty)$. This concludes the proof of theorem 5.

4   Applications to the spiked models

In this section, we use the above results in order to evaluate the behaviour of the largest eigenvalues of the information plus noise spiked models. In the remainder of this section, we assume that

Assumption A-5: *K does not depend on N and for all $k = 1, \ldots, K$, the positive sequence $(\lambda_{k,N})$ writes*

$$\lambda_{k,N} = \lambda_k + \varepsilon_{k,N},$$

with $\lim_{N \to +\infty} \varepsilon_{k,N} = 0$ and $\lambda_i \neq \lambda_j$ for $i \neq j$.

We define $K_s = \max\{k : \lambda_k > \sigma^2 \sqrt{c}\}$ and the function $\psi(\lambda) = \frac{(\sigma^2 + \lambda)(\sigma^2 c + \lambda)}{\lambda}$. In the following, we characterize the support $\mathscr{S}_N$ of measure μ_N and use the above results on the almost sure location of the sample eigenvalues in order to prove the theorem

Theorem 6. *We have with probability one,*

$$\hat{\lambda}_{k,N} \xrightarrow[N \to \infty]{} \begin{cases} \psi(\lambda_k) & \text{if} \quad k \leq K_s \\ \sigma^2 (1 + \sqrt{c})^2 & k \in \{K_s + 1, \ldots, K\} \end{cases}$$

We note that theorem 6 was already proved in the recent paper [7] using a different approach.

4.1 Preliminary results on perturbed equations

We first state two useful lemmas related to the solutions of perturbed equations. They can be interpreted as extensions of lemmas 3.2 and 3.3 of [5]. In the following, we denote respectively by $\mathscr{D}_o(z, r)$, $\mathscr{D}_c(z, r)$ and $\mathscr{C}(z, r)$ the open disk, closed disk and circle of radius $r > 0$ with center z. Moreover, in this paragraph, the notation $o(1)$ denotes a term that converges towards 0 when the variable ϵ converges towards 0. The first result is a straightforward modification of [5, lemma 3.2]. Its proof is thus omitted.

Lemma 6. *For each $\epsilon > 0$, we consider $h_\epsilon(z) = h(z) + \chi_\epsilon(z)$ with h, χ_ϵ two holomorphic functions in a disk $\mathscr{D}_o(z_0, r_0)$. We assume that $\sup_{z \in \mathscr{D}_o(z_0, r_0)} |\chi_\epsilon(z)| = o(1)$. We consider $z_{0,\epsilon} = z_0 + \delta_\epsilon$ with $\delta_\epsilon = o(1)$. Then, $\exists \epsilon_0 > 0$ and $r > 0$ such that for each $0 < \epsilon \leq \epsilon_0$, $z_{0,\epsilon} \in \mathscr{D}_o(z_0, r)$ and the equation*

$$z - z_{0,\epsilon} - \epsilon h_\epsilon(z) = 0,$$

admits a unique solution in $\mathscr{D}_o(z_0, r)$ given by

$$z_\epsilon = z_{0,\epsilon} + \epsilon h(z_0) + o(\epsilon).$$

Moreover, if we assume that $z_0 \in \mathbb{R}$, $h(z) \in \mathbb{R}$ for $z \in \mathbb{R}$, and that for ϵ small enough, $z_{0,\epsilon} \in \mathbb{R}$, $h_\epsilon(z) \in \mathbb{R}$ for $z \in \mathbb{R}$, then $z_\epsilon \in \mathbb{R}$.

The second result is an extension of [5, Lem.3.3] to certain third degree equations. The proof is given the Appendix C.

Lemma 7. *For each $\epsilon > 0$ and $i = 1, 2$, we consider $h_{i,\epsilon}(z) = h_i(z) + \chi_{i,\epsilon}(z)$ with $h_i, \chi_{i,\epsilon}$ holomorphic functions in a disk $\mathscr{D}_o(z_0, r_0)$. We assume that $h_1(z_0) \neq 0$ and that $\sup_{z \in \mathscr{D}_o(z_0, r_0)} |\chi_{i,\epsilon}(z)| = o(1)$ for $i = 1, 2$. We consider $z_{0,\epsilon} = z_0 + \delta_\epsilon$ with $\delta_\epsilon = o(1)$. Then, $\exists \epsilon_0 > 0$ and $r > 0$ such that $z_{0,\epsilon} \subset \mathscr{D}_o(z_0, r)$ $\forall \epsilon \in (0, \epsilon_0)$ and the equation*

$$\left(z - z_{0,\epsilon}\right)^3 - \epsilon \left(z - z_{0,\epsilon}\right) h_{1,\epsilon}(z) + \epsilon^2 h_{2,\epsilon}(z) = 0$$

has 3 solutions in $\mathscr{D}_o(z_0, r)$ given by

$$z_\epsilon^- = z_{0,\epsilon} - \sqrt{\epsilon}\sqrt{h_1(z_0)} + o(\sqrt{\epsilon})$$

$$z_\epsilon^+ = z_{0,\epsilon} + \sqrt{\epsilon}\sqrt{h_1(z_0)} + o(\sqrt{\epsilon})$$

$$z_\epsilon = z_{0,\epsilon} + \epsilon\frac{h_2(z_0)}{h_1(z_0)} + o(\epsilon),$$

where $\sqrt{\cdot}$ is an arbitrary branch of the square root, analytic in a neighborhood of $h_1(z_0)$. Moreover, if we assume that $z_0 \in \mathbb{R}$, $h_i(z) \in \mathbb{R}$ for $z \in \mathbb{R}$ and that for ϵ small enough that $z_{0,\epsilon} \in \mathbb{R}$, $h_{i,\epsilon}(z) \in \mathbb{R}$ for $z \in \mathbb{R}$, then z_ϵ is real. Moreover, if $h_1(z_0) > 0$ then z_ϵ^-, z_ϵ^+ and z_ϵ are real while z_ϵ^-, z_ϵ^+ are non real if $h_1(z_0) < 0$.

4.2 Characterization of $\mathscr{S}_N$ and limits of the eigenvalues if $\lambda_k \neq \sigma^2\sqrt{c}$

In this paragraph, we identify the clusters of the support $\mathscr{S}_N$, and evaluate the points $x_{q,N}^-, x_{q,N}^+$ for $q = 1, \dots, Q_N$. From theorem 3, these points coincide with the positive extrema of function ϕ_N (defined in (7)), and eventually $x_{1,N}^- = 0$ if $c_N = 1$. Therefore, we first evaluate the real zeros of $\phi_N'(w) = (1 - \sigma^2 c_N f_N(w))^2 - 2\sigma^2 c_N w f_N'(w)(1 - \sigma^2 c_N f_N(w)) - \sigma^4 c_N(1 - c_N)f_N'(w)$. Straightforward calculations give

$$\phi_N'(w) = \frac{1}{w^2 \prod_{k=1}^{K}(\lambda_{k,N} - w)^3}\left[\gamma_{1,N}(w) + \frac{1}{M}\gamma_{2,N}(w) + \frac{1}{M^2}\gamma_{3,N}(w)\right],$$

with

$$\gamma_{1,N}(w) = (w^2 - \sigma^4 c_N)\prod_{k=1}^{K}(\lambda_{k,N} - w)^3,$$

$$\gamma_{2,N}(w) = -2\sigma^2 c_N \prod_{k=1}^{K}(\lambda_{k,N} - w)\sum_{j=1}^{K}\left[\lambda_{j,N}\left(w^2 + \sigma^2(1 + c_N)w - \frac{\sigma^2(1 + c_N)\lambda_{j,N}}{2}\right)\prod_{\substack{l=1 \\ l \neq j}}^{K}(\lambda_{l,N} - w)^2\right],$$

$$\gamma_{3,N}(w) = \sigma^4 c_N^2\left(\sum_{k=1}^{K}\lambda_{k,N}\prod_{\substack{l=1 \\ l \neq k}}^{K}(\lambda_{l,N} - w)\right)\left(\sum_{k=1}^{K}\lambda_{k,N}(3w - \lambda_{k,N})\prod_{\substack{l=1 \\ l \neq k}}^{K}(\lambda_{l,N} - w)^2\right).$$

Therefore, $\phi_N'(w) = 0$ if and only if

$$\gamma_{1,N}(w) + \frac{1}{M}\gamma_{2,N}(w) + \frac{1}{M^2}\gamma_{3,N}(w) = 0. \tag{20}$$

We assume $c < 1$, which implies that $c_N < 1$ for N large enough. The calculations are essentially the same if $c = 1$. We first observe that the zeros of ϕ_N are included into a compact interval $\mathscr{I}$ independent of N (see the proof of Corollary 1). Next, we claim that for each $\alpha > 0$, it exists $\beta > 0$ and $N_0 \in \mathbb{N}$ such that

$$\left|\gamma_{1,N}(w) + \frac{1}{M}\gamma_{2,N}(w) + \frac{1}{M^2}\gamma_{3,N}(w)\right| > \beta,$$

if $N > N_0$ and $|w - \sigma^2 \sqrt{c}| > \alpha, |w + \sigma^2 \sqrt{c}| > \alpha, |w - \lambda_k| > \alpha, k = 1, \ldots, K$ and $w \in \mathscr{I}$. This follows immediately from the inequality

$$\left| \gamma_{1,N}(w) + \frac{1}{M} \gamma_{2,N}(w) + \frac{1}{M^2} \gamma_{3,N}(w) \right| \geq |\gamma_{1,N}(w)| - \frac{1}{M} \gamma_{2,max} - \frac{1}{M^2} \gamma_{3,max},$$

where $\gamma_{i,max} = \max_{w \in \mathscr{I}} |\gamma_{i,N}(w)|$ for $i = 2, 3$. This shows that the solutions of eq. (20) are located around the points $\sigma^2 \sqrt{c}, -\sigma^2 \sqrt{c}, \lambda_k, k = 1, \ldots, K$.

In a disk $\mathscr{D}_o(\sigma^2 \sqrt{c}, r)$, (20) is equivalent to

$$w - \sigma^2 \sqrt{c_N} + \frac{1}{M} \frac{w - \sigma^2 \sqrt{c_N}}{\gamma_{1,N}(w)} \left(\gamma_{2,N}(w) + \frac{1}{M} \gamma_{3,N}(w) \right) = 0. \tag{21}$$

We use lemma 6 with $\epsilon = M^{-1}$, $z_0 = \sigma^2 \sqrt{c}$, $z_{0,\epsilon} = \sigma^2 \sqrt{c_N}$, and the functions

$$h_\epsilon(w) = -\frac{(w - \sigma^2 \sqrt{c_N})}{\gamma_{1,N}(w)} \left[\gamma_{2,N}(w) + \frac{1}{M} \gamma_{3,N}(w) \right]$$

and $h(w) = \lim_{M \to +\infty} h_\epsilon(w)$. $h(w)$ is obtained by replacing c_N and the $(\lambda_{k,N})_{k=1,\ldots,K}$ by c and the $(\lambda_k)_{k=1,\ldots,K}$ in the expression of h_ϵ. Lemma 6 implies that it exists r for which equation (21), or equivalently equation (20), has a unique solution in $\mathscr{D}_o(\sigma^2 \sqrt{c}, r)$ for M large enough. This solution is given by $\sigma^2 \sqrt{c_N} + \mathcal{O}(\frac{1}{M})$. It is easy to check that

$$\phi_N \left(\sigma^2 \sqrt{c_N} + \mathcal{O}\left(M^{-1} \right) \right) = \sigma^2 (1 + \sqrt{c_N})^2 + \mathcal{O}\left(\frac{1}{M} \right).$$

This quantity is positive, thus showing that $\sigma^2 \sqrt{c_N} + \mathcal{O}(M^{-1})$ is the pre-image of a positive extremum of ϕ_N. Exchanging $\sigma^2 \sqrt{c}$ with $-\sigma^2 \sqrt{c}$, we obtain similarly that it exists a neighborhood of $-\sigma^2 \sqrt{c}$ in which equation (20) has a unique solution given by $-\sigma^2 \sqrt{c_N} + \mathcal{O}(\frac{1}{M})$. Moreover,

$$\phi_N \left(-\sigma^2 \sqrt{c_N} + \mathcal{O}\left(M^{-1} \right) \right) = \sigma^2 (1 - \sqrt{c_N})^2 + \mathcal{O}\left(\frac{1}{M} \right),$$

so that $-\sigma^2 \sqrt{c_N} + \mathcal{O}(\frac{1}{M})$ is also the pre-image of a positive extremum of ϕ_N.

We now consider $i \in \{1, \ldots, K\}$, and study the equation (20) in a neighborhood $\mathscr{D}_o(\lambda_i, r)$ of λ_i. In order to use lemma 7, we put $\epsilon = \frac{1}{M}, z_0 = \lambda_i, z_{0,\epsilon} = \lambda_{i,N}$. It is easily seen that in $\mathscr{D}_o(\lambda_i, r)$, eq. (20) is equivalent to

$$(w - \lambda_{i,N})^3 - \frac{1}{M}(w - \lambda_{i,N}) h_{1,\epsilon}(w) + \frac{1}{M^2} h_{2,\epsilon}(w) = 0,$$

where

$$h_{1,\epsilon}(w) = \frac{2\sigma^2 c_N \sum_{k=1}^{N} \left[\lambda_{k,N} \left(w^2 + \sigma^2(1 + c_N)w - \frac{\sigma^2(1 + c_N)\lambda_{k,N}}{2} \right) \prod_{\substack{l=1 \\ l \neq k}}^{K} (\lambda_{l,N} - w)^2 \right]}{(w^2 - \sigma^4 c_N) \prod_{\substack{k=1 \\ k \neq i}}^{K} (\lambda_{k,N} - w)^2},$$

$$h_{2,\epsilon}(w) = -\frac{\gamma_{3,N}(w)}{(w^2 - \sigma^4 c_N) \prod_{k \neq i}^{K} (\lambda_{k,N} - w)^3}.$$

We denote by $h_1(w)$ and $h_2(w)$ the limits of $h_{1,\epsilon}(w)$ and $h_{2,\epsilon}(w)$ when $\epsilon \to 0$, i.e. the functions obtained by replacing c_N and the $(\lambda_{k,N})_{k=1,\dots,K}$ by c and the $(\lambda_k)_{k=1,\dots,K}$ respectively in the expressions of $h_{1,\epsilon}, h_{2,\epsilon}$. After some algebra, we obtain that

$$h_1(\lambda_i) = \frac{2\sigma^2 c \lambda_i^2 (\lambda_i + \frac{\sigma^2(1+c)}{2})}{\lambda_i^2 - \sigma^4 c},$$

while $h_2(\lambda_i)$ is equal to

$$h_2(\lambda_i) = -\frac{2\sigma^4 c^2 \lambda_i^3}{\lambda_i^2 - \sigma^4 c}.$$

Lemma 7 implies that it exists r such that

$$\lambda_{i,N} - \frac{1}{M} \frac{\sigma^2 c \lambda_i}{\lambda_i + \sigma^2 \frac{1+c}{2}} + o\left(\frac{1}{M}\right)$$

is solution of (20) contained in $\mathcal{D}_o(\lambda_i, r)$. It is however easy to check that

$$\phi_N\left(\lambda_{i,N} - \frac{1}{M} \frac{\sigma^2 c \lambda_i}{\lambda_i + \sigma^2 \frac{1+c}{2}} + o\left(\frac{1}{M}\right)\right) = -\frac{\sigma^4(1-c)^2}{2\lambda_i}\left(1 - \frac{c}{2}\right) < 0.$$

Therefore, the above extremum is negative, and its pre-image cannot be one the points $w_{q,N}^-, w_{q,N}^+$. Moreover, if $\lambda_i < \sigma^2 \sqrt{c}$, then $h_1(\lambda_i) < 0$ and (20) has no extra real solution in $\mathcal{D}_o(\lambda_i, r)$. If $\lambda_i > \sigma^2 \sqrt{c}$, then $h_1(\lambda_i) > 0$, and the quantities

$$\lambda_{i,N} - \frac{1}{\sqrt{M}}\sqrt{h_1(\lambda_i)} + o\left(\frac{1}{\sqrt{M}}\right) \quad \text{and} \quad \lambda_{i,N} + \frac{1}{\sqrt{M}}\sqrt{h_1(\lambda_i)} + o\left(\frac{1}{\sqrt{M}}\right)$$

are the 2 other real solutions of (20) contained in $\mathcal{D}_o(\lambda_i, r)$. After some algebra, we get that

$$\phi_N\left(\lambda_{i,N} - \frac{1}{\sqrt{M}}\sqrt{h_1(\lambda_i)} + o\left(\frac{1}{\sqrt{M}}\right)\right) = \frac{(\lambda_{i,N} + \sigma^2 c_N)(\lambda_{i,N} + \sigma^2)}{\lambda_{i,N}} - \frac{1}{\sqrt{M}} \frac{2\sqrt{h_1(\lambda_i)}(\lambda_i^2 - \sigma^4 c)}{\lambda_i^2} + o\left(\frac{1}{\sqrt{M}}\right),$$

$$\phi_N\left(\lambda_{i,N} + \frac{1}{\sqrt{M}}\sqrt{h_1(\lambda_i)} + o\left(\frac{1}{\sqrt{M}}\right)\right) = \frac{(\lambda_{i,N} + \sigma^2 c_N)(\lambda_{i,N} + \sigma^2)}{\lambda_{i,N}} + \frac{1}{\sqrt{M}} \frac{2\sqrt{h_1(\lambda_i)}(\lambda_i^2 - \sigma^4 c)}{\lambda_i^2} + o\left(\frac{1}{\sqrt{M}}\right),$$

are both positive. It is easy to check that if $k \leq K_s$, then, $\sigma^2 \sqrt{c_N} < \lambda_{k,N}$ for N large enough. The above discussion thus implies that $\mathscr{S}_N$ has $K_s + 1$ clusters, and that for $k \leq K_s$, then

$$x_{1,N}^- = \sigma^2 (1 - \sqrt{c_N})^2 + \mathcal{O}\left(\frac{1}{M}\right),$$

$$x_{1,N}^+ = \sigma^2 (1 + \sqrt{c_N})^2 + \mathcal{O}\left(\frac{1}{M}\right),$$

$$x_{K_s+2-k,N}^- = \frac{(\lambda_{k,N} + \sigma^2 c_N)(\lambda_{k,N} + \sigma^2)}{\lambda_{k,N}} - \frac{1}{\sqrt{M}} \frac{2\sqrt{h_1(\lambda_k)}(\lambda_k^2 - \sigma^4 c)}{\lambda_k^2} + o\left(\frac{1}{\sqrt{M}}\right),$$

$$x_{K_s+2-k,N}^+ = \frac{(\lambda_{k,N} + \sigma^2 c_N)(\lambda_{k,N} + \sigma^2)}{\lambda_{k,N}} + \frac{1}{\sqrt{M}} \frac{2\sqrt{h_1(\lambda_k)}(\lambda_k^2 - \sigma^4 c)}{\lambda_k^2} + o\left(\frac{1}{\sqrt{M}}\right).$$

1952

In order to complete the proof, we use theorem 5. Let $k \in \{1, \ldots, K_s\}$. From the previous analysis, the eigenvalue $\lambda_{k,N}$ is the unique eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$ associated with interval $(w_{q,N}^-, w_{q,N}^+)$ with $q = K_s - k + 2$, for N large enough. Moreover, the number of clusters of $\mathscr{S}_N$ is equal to $K_s + 1$ for N large enough and the sequences $x_{q,N}^-$ and $x_{q,N}^+$ converge towards limits equal to $\sigma^2(1 - \sqrt{c})^2$ and $\sigma^2(1 + \sqrt{c})^2$ for $q = 1$, and both coincide with $\psi(\lambda_{K_s+2-q})$ for $q \geq 2$. This implies that for each $\epsilon > 0$, almost surely for N large enough, then $\hat{\lambda}_{k,N} \in (\psi(\lambda_k) - \epsilon, \psi(\lambda_k) + \epsilon)$ for $k = 1, \ldots, K_s$ and that $\hat{\lambda}_{k,N} \in (\sigma^2(1 - \sqrt{c})^2 - \epsilon, \sigma^2(1 + \sqrt{c})^2 + \epsilon)$ for $k > K_s$. This shows that $\hat{\lambda}_{k,N} \to \psi(\lambda_k)$ for $k = 1, \ldots, K_s$.

We now prove the convergence of $\hat{\lambda}_{k,N}$ to $\sigma^2(1 + \sqrt{c})^2$ for $K_s < k \leq K$. Let $k_{\max} = K_s + 1$ (i.e the index of the largest eigenvalue associated with the first cluster $[x_{1,N}^-, x_{1,N}^+]$). We have already shown $\limsup_N \hat{\lambda}_{k_{\max},N} \leq \sigma^2(1 + \sqrt{c})^2$ almost surely. It remains to prove $\liminf_N \hat{\lambda}_{k_{\max},N} \geq \sigma^2(1 + \sqrt{c})^2$. Assume the converse is true. Then it exists $\epsilon > 0$ such that $\liminf_N \hat{\lambda}_{k_{\max},N} < \sigma^2(1 + \sqrt{c})^2 - \epsilon$. We can thus extract a subsequence $\hat{\lambda}_{k_{\max},\phi(N)}$ converging towards a limit less than $\sigma^2(1 + \sqrt{c})^2 - \epsilon$. Let $\hat{\mu}_{\phi(N)}$ be the empirical spectral measure associated with matrix $\Sigma_{\phi(N)} \Sigma_{\phi(N)}^*$. We deduce that

$$\hat{\mu}_{\phi(N)} \left((\sigma^2(1 + \sqrt{c})^2 - \epsilon, \sigma^2(1 + \sqrt{c})^2] \right) = 0 \quad \text{a.s for all large N.} \tag{22}$$

Theorem 1 implies that $\hat{\mu}_{\phi(N)}$ converges torwards the Marcenko-Pastur distribution, which contradicts (22). This proves that $\hat{\lambda}_{k_{\max},N} \to \sigma^2(1 + \sqrt{c})^2$ with probability one. We can prove similarly that $\hat{\lambda}_{k,N} \to \sigma^2(1 + \sqrt{c})^2$ a.s for $K_s + 1 < k \leq K$.

4.3   Characterization of $\mathscr{S}_N$ and limits of the eigenvalues if $\sigma^2\sqrt{c} \in \{\lambda_1, \ldots, \lambda_K\}$

In this section, we handle the case where one the $(\lambda_k)_{k=1,\ldots,K}$, say λ_j with $j \leq K$, is equal to $\sigma^2\sqrt{c}$. For this, we will use the Fan inequality (see [15, Th.2]). For a rectangular matrix $\mathbf{A}$, we will denote by $\kappa_k(\mathbf{A})$ its k-th largest singular value. With this notations, we have $\kappa_j(\mathbf{B}) = \sqrt{\lambda_j} = \sqrt{\sigma^2\sqrt{c}}$. We also denote by $\mathbf{u}_{j,N}$ and $\mathbf{v}_{j,N}$ the left and right singular vector of $\mathbf{B}_N$ associated with $\kappa_j(\mathbf{B}_N)$. Fan inequality gives, for $\epsilon > 0$,

$$\kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N) \leq \kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N + \epsilon\,\mathbf{u}_{j,N}\mathbf{v}_{j,N}^*) + \kappa_1(\epsilon\,\mathbf{u}_{j,N}\mathbf{v}_{j,N}^*),$$

$$\kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N + \epsilon\,\mathbf{u}_{j,N}\mathbf{v}_{j,N}^*) \leq \kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N) + \kappa_1(\epsilon\,\mathbf{u}_{j,N}\mathbf{v}_{j,N}^*).$$

From the results of the previous section, it is clear that, almost surely,

$$\kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N + \epsilon\,\mathbf{u}_{j,N}\mathbf{v}_{j,N}^*) = \sqrt{\psi\left(\left(\sqrt{\lambda_j} + \epsilon\right)^2 \right)} + o(1).$$

Therefore, we end up with

$$\sqrt{\psi\left(\left(\sqrt{\lambda_j} + \epsilon\right)^2 \right)} - \epsilon \leq \liminf_N \kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N) \leq \limsup_N \kappa_j(\mathbf{B}_N + \sigma\mathbf{W}_N) \leq \sqrt{\psi\left(\left(\sqrt{\lambda_j} + \epsilon\right)^2 \right)} + \epsilon.$$

Since $\psi(\lambda) \to \sigma^2(1 + \sqrt{c})^2$ when $\lambda \to \sigma^2\sqrt{c}$, this completes the results of theorem 6.

A Proof of item 2 of theorem 2

The proof is a direct consequence of Section 2 of [13] (formulas (2.1) to (2.5) of [13]). We first recall that if we denote by $g(z)$ and $G(z)$ the terms defined for $z \in \mathbb{C}^+$ by

$$g(z) = \frac{\sigma^2 c_N}{|1 + \sigma^2 c_N m_N(z)|^2} \frac{1}{M} \mathrm{Tr}\, \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z) \mathbf{T}_N(z)^*,$$

$$G(z) = \sigma^2 c_N \frac{1}{M} \mathrm{Tr}\, \mathbf{T}_N(z) \mathbf{T}_N(z)^*,$$

where $\mathbf{T}_N(z) = \left[-z(1 + \sigma^2 c_N m_N(z))\mathbf{I}_M + \sigma^2(1 - c_N)\mathbf{I}_M + \frac{\mathbf{B}_N \mathbf{B}_N^*}{1 + \sigma^2 c_N m_N(z)} \right]^{-1}$, then, it is shown in [13] that

$$0 < |z| G(z) < 1 - g(z) \tag{23}$$

for each $z \in \mathbb{C}^+$. If $z_1 = \mathrm{Re}(z)$ and $z_2 = \mathrm{Im}(z)$, (23) implies that $0 < 1 - g(z) - |z_1| G(z) \le 1 - g(z) + z_1 G(z)$. It is shown in [13] that

$$\mathrm{Re}(1 + \sigma^2 c_N m_N(z)) = \frac{1 + \sigma^2(1 - c_N) G(z) + \mathrm{Im}(1 + \sigma^2 c_N m_N(z)) z_2 G(z)}{1 - g(z) + z_1 G(z)},$$

for $z \in \mathbb{C}^+$. As $\mathrm{Im}(1 + \sigma^2 c_N m_N(z)) = \sigma^2 c_N \mathrm{Im}(m_N(z)) > 0$ on $\mathbb{C}^+$ (see item 3 of Property 1), we get that

$$\mathrm{Re}(1 + \sigma^2 c_N m_N(z)) > \frac{1}{1 - g(z) + z_1 G(z)} > \frac{1}{1 + z_1 G(z)}.$$

The inequality $|z_1| G(z) < 1$ implies that $\mathrm{Re}(1 + \sigma^2 c_N m_N(z)) > \frac{1}{2}$ for each $z \in \mathbb{C}^+$. This also implies that $\mathrm{Re}(1 + \sigma^2 c_N m_N(x)) \ge \frac{1}{2}$ for $x \in \mathbb{R}$ if $c_N < 1$ and for $x \in \mathbb{R}^*$ if $c_N = 1$.

B Proof of items 6 and 8 of lemma 2 when $q = 1$

In order to prove these 2 statements, we study the behaviour of $w_N(x)$ and of $w_N'(x)$ when $x \to 0, x < 0$ and $x \to 0, x > 0$.

We first look at the limit for $x < 0$. Lemmas 1 and 2 imply that w_N is the inverse of ϕ_N on interval $(-\infty, 0)$. $w_N(x)$ is a continuous increasing function on $(-\infty, 0)$ upperbounded by $w_{1,N}^-$; therefore, $\lim_{x \to 0, x < 0} w_N(x)$ exists, and is less than $w_{1,N}^-$. Taking the limit when $x \to 0, x < 0$ from both sides of the equation $\phi_N(w_N(x)) = x$ for $x \in (-\infty, 0)$, and using the continuity of ϕ_N on $(-\infty, 0]$, we obtain immediately that $\phi_N(\lim_{x \to 0, x < 0} w_N(x)) = 0$. This implies that $\lim_{x \to 0, x < 0} w_N(x) = w_{1,N}^-$. This shows that w_N is left continuous at $x = 0$. Since $w_N(x) = x(1 + \sigma^2 m_N(x))^2$, it follows that $1 + \sigma^2 m_N(x) = \mathcal{O}(|x|^{-1/2})$. As w_N is continuously differentiable on $(-\infty, 0)$, we can differentiate the relation $\phi_N(w_N(x)) = x$, and obtain that $\phi_N'(w_N(x)) w_N'(x) = 1$ for $x < 0$, or equivalently that $w_N'(x) = \frac{1}{\phi_N'(w_N(x))}$. In other words, it holds that

$$w_N'(x) = \frac{1}{[1 - \sigma^2 f_N(w_N(x))][1 - \sigma^2 f_N(w_N(x)) - 2\sigma^2 w_N(x) f_N'(w_N(x))]}. \tag{24}$$

We observe that $1 - \sigma^2 f_N(w_{1,N}^-) = 0$ so that

$$\lim_{x \uparrow 0} 1 - \sigma^2 f_N(w_N(x)) - 2\sigma^2 w_N(x) f_N'(w_N(x))) = -2\sigma^2 w_{1,N}^- f_N'(w_{1,N}^-) \neq 0 \tag{25}$$

Moreover, (9) implies that

$$\frac{1}{1 - \sigma^2 f_N(w_N(x))} = 1 + \sigma^2 m_N(x) \quad \text{for } x < 0,$$

which proves that $(1 - \sigma^2 f_N(w_N(x)))^{-1} = \mathcal{O}((-x)^{-1/2})$. (24) implies immediately that $w_N'(x) = \mathcal{O}\left(\frac{1}{\sqrt{-x}}\right)$.

We now study the behaviour of w_N and w_N' when $x \to 0$, $x > 0$. We first study $\sqrt{x} m_N(x)$ for $x \to 0, x > 0$. For this, we introduce the function $\psi(\xi, y)$ defined by

$$\psi(\xi, y) = 1 - \frac{1}{M} \mathrm{Tr} \left(\mathbf{B}_N \mathbf{B}_N^* \frac{\xi}{y + \sigma^2 \xi} - \xi(y + \sigma^2 \xi) \right)^{-1}.$$

The introduction of ψ is based on the observation that eq. (6) is equivalent to $\psi(\sqrt{x} m_N(x), \sqrt{x}) = 0$ for $x > 0$. We denote by ξ_0 the term $\xi_0 = i\sigma^{-2}\sqrt{|w_{1,N}^-|}$ and notice that $\psi(\xi_0, 0) = 0$. It is easily checked that ψ is holomorphic in a neighborhood of $(\xi_0, 0)$ and that $\frac{\partial \psi}{\partial \xi}(\xi_0, 0) \neq 0$. Therefore, from the implicit function theorem (the analytic version - see e.g Cartan [11, Prop.6]), it exists a unique function $\xi(y)$, holomorphic in a neighborhood $\mathscr{V}$ of 0 satisfying $\psi(\xi(y), y) = 0$ for $y \in \mathscr{V}$ and $\xi(0) = \xi_0$. As $\mathrm{Im}(\xi_0) > 0$, it is clear that it exists a neighborhood $\mathscr{V}'$ of 0 included in $\mathscr{V}$ such that $\mathrm{Im}(\xi(y)) > 0$ for each $y \in \mathscr{V}'$. We claim that for $\sqrt{x} \in \mathscr{V}' \cap \mathbb{R}^{+*}$, $\xi(\sqrt{x}) = \sqrt{x} m_N(x)$. For this, we notice that if $x \in (0, x_{1,N}^+)$, $m_N(x)$ is the unique solution of Eq. (6) for which $\mathrm{Im}(m_N(x)) > 0$. Indeed, from item 5 of lemma 2, for $x \in (0, x_{1,N}^+)$, $w_N(x)$ is the unique solution with positive imaginary part of equation $\phi_N(w) = x$. But, $m_N(x)$ is solution of (6) iff $w_N(x)$ is solution of $\phi_N(w) = x$. Moreover $m_N(x) \in \mathbb{C}^+$ iff $w_N(x) \in \mathbb{C}^+$, a property which is readily seen from the relation (8). The conclusion follows from the observation that $m_N(x)$ satisfies (6) iff $\sqrt{x} m_N(x)$ satisfies $\psi(\sqrt{x} m_N(x), \sqrt{x}) = 0$. This in turn shows that for each $\sqrt{x} \in \mathscr{V}' \cap \mathbb{R}^{+*}$, $\xi(\sqrt{x}) = \sqrt{x} m_N(x)$, or equivalently that $\xi(y) = y m_N(y^2)$ for $y \in \mathscr{V}' \cap \mathbb{R}^{+*}$. As $\xi(y)$ is holomorphic in $\mathscr{V}'$, $\xi(y) = \xi_0 + o(1)$ and $\xi'(y) = \xi_1 + o(1)$ for some coefficient ξ_1. Therefore, $y m_N(y^2) = \xi_0 + o(1)$ and $2y^2 m_N'(y^2) + m_N(y^2) = \xi_1 + o(1)$ for $y \in \mathscr{V}' \cap \mathbb{R}^{+*}$, or equivalently $\sqrt{x} m_N(x) = \xi_0 + o(1)$ and $2x m_N'(x) + m_N(x) = \xi_1 + o(1)$ for $x > 0$ small enough. As $w_N(x) = x(1 + \sigma^2 m_N(x))^2$, we get that

$$w_N'(x) = \left(1 + \sigma^2 m_N(x)\right)\left(1 + \sigma^2(m_N(x) + 2x m_N'(x))\right).$$

As $(m_N(x) + 2x m_N'(x))$ is a $\mathcal{O}(1)$ term, and as $m_N(x) = \frac{\xi_0}{\sqrt{x}} + o(\frac{1}{\sqrt{x}})$, we obtain that $|w_N'(x)| \leq \frac{C}{\sqrt{x}}$ for $x > 0$ small enough for some constant $C > 0$.

C Proof of lemma 7

We begin by choosing $r > 0$ and $\epsilon_1 > 0$ such that $r < r_0$, $z_{0,\epsilon} \in \mathscr{D}_c(z_0, r)$ and $\mathscr{D}_c(z_{0,\epsilon}, r) \subset \mathscr{D}_0(z_0, r_0)$, for each $0 < \epsilon < \epsilon_1$. Let $f_\epsilon(z) = (z - z_{0,\epsilon})^3 - \epsilon(z - z_{0,\epsilon})h_{1,\epsilon}(z) + \epsilon^2 h_{2,\epsilon}(z)$ and $g_\epsilon(z) = (z - z_{0,\epsilon})^3$. Moreover, define $\frac{K_i}{2} = \sup_{\mathscr{D}_c(z_0, r)} |h_i(z)|$ (for $i = 1, 2$).

As $\sup_{z\in\mathscr{D}_o(z_0,r_0)}|\chi_{i,\epsilon}(z)| = o(1)$, it exists $\epsilon_2 \le \epsilon_1$ such that $\sup_{\mathscr{D}_c(z_0,r)}|h_{i,\epsilon}(z)| \le K_i$ (for $i = 1, 2$) for each $\epsilon \le \epsilon_2$. For $z \in \mathscr{D}_c(z_0, r)$, it holds that

$$\left|f_\epsilon(z) - g_\epsilon(z)\right| \le \epsilon\left|z - z_{0,\epsilon}\right|\left|h_{1,\epsilon}(z)\right| + \epsilon^2\left|h_{2,\epsilon}(z)\right|.$$

As $z_{0,\epsilon} - z_0 = o(1)$, it exists $\epsilon_3 \le \epsilon_2$ such that, for each $\epsilon \le \epsilon_3$, $|z - z_{0,\epsilon}| < 2r$ on $\mathscr{D}_c(z_0, r)$. Hence, for each $\epsilon \le \epsilon_3$, it holds that $\left|f_\epsilon(z) - g_\epsilon(z)\right| \le 2\epsilon r K_1 + \epsilon^2 K_2$ on $\mathscr{D}_c(z_0, r)$. We now restrict z to $\mathscr{C}(z_0, r)$, the boundary of $\mathscr{D}_c(z_0, r)$. It exists $\epsilon_4 \le \epsilon_3$ for which $2\epsilon r K_1 + \epsilon^2 K_2 < \frac{r^3}{2} < r^3 = |z - z_0|^3$ holds on $\mathscr{C}(z_0, r)$ for each $\epsilon \le \epsilon_4$. Therefore, $\forall z \in \mathscr{C}(z_0, r)$, we have $|f_\epsilon(z) - g_\epsilon(z)| < |g_\epsilon(z)|$ for $\epsilon \le \epsilon_4$. It follows from Rouché's theorem that these values of ϵ, then f_ϵ and g_ϵ have the same number of zeros inside $\mathscr{D}_o(z_0, r)$. Thus, for $\epsilon \le \epsilon_4$, the equation

$$(z - z_{0,\epsilon})^3 - \epsilon(z - z_{0,\epsilon})h_{1,\epsilon}(z) + \epsilon^2 h_{2,\epsilon}(z) = 0 \tag{26}$$

has three solutions in $\mathscr{D}_o(z_0, r)$. Using the the same procedure to functions $f_\epsilon(z) = (z - z_{0,\epsilon})^2 - \epsilon h_{1,\epsilon}(z)$ and $g_\epsilon(z) = (z - z_{0,\epsilon})^2$, we deduce that if $\epsilon \le \epsilon_5 \le \epsilon_4$, the equation

$$(z - z_{0,\epsilon})^2 - \epsilon h_{1,\epsilon}(z) = 0 \tag{27}$$

has two solutions $\hat{z}_\epsilon^-, \hat{z}_\epsilon^+$ in $\mathscr{D}_o(z_0, r)$. We clearly have $|z_{0,\epsilon} - \hat{z}_\epsilon^-| = \mathcal{O}(\epsilon^{1/2})$ and $|z_0 - \hat{z}_\epsilon^-| = o(1)$. Therefore, $h_{1,\epsilon}(\hat{z}_\epsilon^-) - h_1(z_0) = o(1)$. As $h_1(z_0) \ne 0$, it exists $\epsilon_6 \le \epsilon_5$ and a neighborhood of $h_1(z_0)$, containing $h_{1,\epsilon}(\hat{z}_\epsilon^-), h_{1,\epsilon}(z_0)$ for each $\epsilon \le \epsilon_6$, in which a suitable branch of the square-root $\sqrt{\cdot}$ is analytic. We assume that solution $\hat{z}_\epsilon^-$ is given by $z_{0,\epsilon} - \hat{z}_\epsilon^- = -\sqrt{\epsilon}\sqrt{h_{1,\epsilon}(\hat{z}_\epsilon^-)}$. As $|h_1(z_0) - h_{1,\epsilon}(\hat{z}_\epsilon^-)| = o(1)$, we have $z_{0,\epsilon} - \hat{z}_\epsilon^- = -\sqrt{\epsilon}\sqrt{h_1(z_0)} + o(\sqrt{\epsilon})$. We obtain similarly that $z_{0,\epsilon} - \hat{z}_\epsilon^+ = \sqrt{\epsilon}\sqrt{h_1(z_0)} + o(\sqrt{\epsilon})$.

Considering again $\hat{z}_\epsilon^-$, it follows that it exists $\epsilon_7 \le \epsilon_6$ such that for each $\epsilon \le \epsilon_7$, it holds that

$$\left|z_{0,\epsilon} - \hat{z}_\epsilon^-\right| > \frac{\sqrt{\epsilon}\sqrt{h_1(z_0)}}{2} > \sqrt{\epsilon}\sqrt{r'}, \tag{28}$$

with $r' < \frac{|h_1(z_0)|}{4}$. For $\epsilon \le \epsilon_8 \le \epsilon_7$, we have $\sqrt{\epsilon r'} < r$ and for $z \in \mathscr{D}_c(z_{0,\epsilon}, \sqrt{\epsilon r'})$, we get

$$\left|(z - z_{0,\epsilon})^2 - \epsilon h_{1,\epsilon}(z)\right| > \epsilon|h_{1,\epsilon}(z)| - |z - z_{0,\epsilon}|^2 > \epsilon\left(|h_{1,\epsilon}(z)| - r'\right).$$

It is easy to check that for each $\epsilon \le \epsilon_9 \le \epsilon_8$, then $|h_{1,\epsilon}(z)| > \frac{|h_1(z_0)|}{2}$ for $z \in \mathscr{D}_c(z_{0,\epsilon}, \sqrt{\epsilon r'})$. Therefore,

$$\left|(z - z_{0,\epsilon})^2 - \epsilon h_{1,\epsilon}(z)\right| > \epsilon\left(\frac{|h_1(z_0)|}{2} - r'\right) > \epsilon r'. \tag{29}$$

The inequalities (28) and (29) prove that in $\mathscr{D}_c(z_{0,\epsilon}, \sqrt{\epsilon r'})$, the equation (27) has no solution and that the equation $(z - z_{0,\epsilon})^3 - \epsilon(z - z_{0,\epsilon})h_{1,\epsilon}(z) = 0$ has only one solution there.

We now study the number of solutions in $\mathscr{D}_c(z_{0,\epsilon}, \sqrt{\epsilon r'})$ of the equation (26). Consider

$$f_\epsilon(z) = (z - z_{0,\epsilon})^3 - \epsilon(z - z_{0,\epsilon})h_{1,\epsilon}(z) + \epsilon^2 h_{2,\epsilon}(z),$$
$$g_\epsilon(z) = (z - z_{0,\epsilon})^3 - \epsilon(z - z_{0,\epsilon})h_{1,\epsilon}(z).$$

We have $|f_\epsilon(z) - g_\epsilon(z)| = \epsilon^2 |h_{2,\epsilon}(z)|$. We consider $z \in \mathscr{C}(z_{0,\epsilon}, \sqrt{\epsilon r'})$. From (29), $|g_\epsilon(z)| > (\epsilon r')^{3/2}$. Therefore, for each $\epsilon \le \epsilon_{10} \le \epsilon_9$, it holds that $|g_\epsilon(z)| > \epsilon^2 |h_{2,\epsilon}(z)| = |f_\epsilon(z) - g_\epsilon(z)|$. Thus, from Rouché's theorem, the equation (26) has only one solution in $\mathscr{D}_o(z_{0,\epsilon}, \sqrt{\epsilon r'})$, denoted by z_ϵ. To obtain z_ϵ, we write

$$z_\epsilon - z_{0,\epsilon} = \frac{-\epsilon^2 h_{2,\epsilon}(z_\epsilon)}{(z_\epsilon - z_{0,\epsilon})^2 - \epsilon h_{1,\epsilon}(z_\epsilon)}.$$

Since $|(z - z_{0,\epsilon})^2 - \epsilon h_{1,\epsilon}(z)| > \epsilon r'$ on $\mathscr{D}_c(z_{0,\epsilon}, \sqrt{\epsilon r'})$ (see (29)), we get that

$$\left| z_\epsilon - z_{0,\epsilon} \right| \le \frac{\epsilon^2 K_2}{\epsilon r'} = \mathscr{O}(\epsilon).$$

But from equation (26), we also have $\epsilon(z_\epsilon - z_{0,\epsilon}) h_{1,\epsilon}(z_\epsilon) = (z_\epsilon - z_{0,\epsilon})^3 + \epsilon^2 h_{2,\epsilon}(z_\epsilon)$ which leads to

$$z_\epsilon - z_{0,\epsilon} = \epsilon \frac{h_{2,\epsilon}(z_\epsilon)}{h_{1,\epsilon}(z_\epsilon)} + \frac{(z_\epsilon - z_{0,\epsilon})^3}{\epsilon h_{1,\epsilon}(z_\epsilon)}.$$

It is clear that

$$\frac{h_{2,\epsilon}(z_\epsilon)}{h_{1,\epsilon}(z_\epsilon)} - \frac{h_2(z_0)}{h_1(z_0)} = o(1),$$

so that

$$z_\epsilon - z_{0,\epsilon} = \epsilon \frac{h_2(z_0)}{h_1(z_0)} + o(\epsilon).$$

We now evaluate the two remaining solutions of (26) located in the set $\mathscr{D}_o(z_0, r) \backslash \mathscr{D}_o(z_{0,\epsilon}, \sqrt{\epsilon r'})$, denoted $z_\epsilon^-, z_\epsilon^+$. As $|z_\epsilon^- - z_{0,\epsilon}| > \sqrt{r'\epsilon}$, we can write

$$\left(z_\epsilon^- - z_{0,\epsilon} \right)^2 = \epsilon h_{1,\epsilon}(z_\epsilon^-) - \epsilon^2 \frac{h_{2,\epsilon}(z_\epsilon^-)}{z_\epsilon^- - z_{0,\epsilon}} \tag{30}$$

This implies that $|z_\epsilon^- - z_{0,\epsilon}| = \mathscr{O}(\sqrt{\epsilon})$ and that $|z_\epsilon^- - z_0| = o(1)$. Taking a suitable branch of the square root, (30) implies that

$$z_\epsilon^- - z_{0,\epsilon} = -\sqrt{\epsilon h_{1,\epsilon}(z_\epsilon^-)} + o(\sqrt{\epsilon}) = -\sqrt{\epsilon h_1(z_0)} + o(\sqrt{\epsilon}).$$

We obtain similarly that $z_\epsilon^+ - z_{0,\epsilon} = \sqrt{\epsilon h_1(z_0)} + o(\sqrt{\epsilon})$.

We finally verify that if z_0 and $z_{0,\epsilon}$ belong to $\mathbb{R}$ for each ϵ, and that $h_i(z)$ and $h_{i,\epsilon}(z)$ belong to $\mathbb{R}$ for each ϵ if $z \in \mathbb{R}$ for $i = 1, 2$, then z_ϵ is real while $z_\epsilon^-, z_\epsilon^+$ are real if $h_1(z_0) > 0$.

If z_ϵ is not real, it is clear that z_ϵ^* is also solution of (26) because functions $h_{i,\epsilon}$ verifies $(h_{i,\epsilon}(z))^* = h_{i,\epsilon}(z^*)$. As $|z_\epsilon^* - z_{0,\epsilon}| = |z_\epsilon - z_{0,\epsilon}| = \mathscr{O}(\epsilon)$, and that (26) has a unique solution in the disk $\mathscr{D}_o(z_{0,\epsilon}, \sqrt{\epsilon r'})$, this implies that $z_\epsilon^* = z_\epsilon$. On the other hand, assume that $h_1(z_0) > 0$ and the $z_\epsilon^-, z_\epsilon^+$ are non-real. Then, z_ϵ^{+*} and z_ϵ^{-*} are also solution of (27). Since equation (27) has only two solutions outside the disk $\mathscr{D}_o(z_{0,\epsilon}, \sqrt{\epsilon r'})$, it follows that $\hat{z}_\epsilon^+$ and $\hat{z}_\epsilon^-$ are complex conjuguate. But as their real parts have opposite sign for ϵ small enough, this leads to a contradiction. Therefore $\hat{z}_\epsilon^+$ and $\hat{z}_\epsilon^-$ are real. We finally note that if $h_1(z_0) < 0$, then $\hat{z}_\epsilon^+$ and $\hat{z}_\epsilon^-$ are non real.

References

[1] Z. Bai and J. Yao. Central limit theorems for eigenvalues in a spiked population model. *Annales de l'Institut Henri Poincaré-Probabilités et Statistiques*, 44(3):447–474, 2008. MR2451053

[2] Z.D. Bai and J.W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Annals of Probability*, 26(1):316–345, 1998. MR1617051

[3] Z.D. Bai and J.W. Silverstein. Exact separation of eigenvalues of large dimensional sample covariance matrices. *Annals of Probability*, 27(3):1536–1555, 1999. MR1733159

[4] J. Baik, G. Ben Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Annals of Probability*, 33(3):1643–1697, 2005.

[5] J. Baik and J.W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis*, 97(6):1382–1408, 2006. MR2279680

[6] F. Benaych-Georges and R.R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 2011. MR2782201

[7] F. Benaych-Georges and R.R. Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *submitted*, 2011. arXiv:1103.2221. MR2782201

[8] M. Capitaine and C. Donati-Martin. Strong asymptotic freeness of Wigner and Wishart matrices. *Indiana Univ. Math. Journal*, 56:295–309, 2007. MR2317545

[9] M. Capitaine, C. Donati-Martin, and D. Féral. The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and non-universality of the fluctuations. *Annals of Probability*, 37(1):1–47, 2009. MR2489158

[10] M. Capitaine, C. Donati-Martin, D. Féral, and M. Février. Free convolution with a semi-circular distribution and eigenvalues of spiked deformations of wigner matrices. *submitted*, 2010. arXiv:1006.3684.

[11] H.P. Cartan. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs varibales complexes*. Hermann, 1961.

[12] L.H.Y. Chen. An inequality for the multivariate normal distribution. *Journal of Multivariate Analysis*, 12(2):306–315, 1982. MR0661566

[13] R.B. Dozier and J.W. Silverstein. Analysis of the limiting spectral distribution of large dimensional information plus noise type matrices. *Journal of Multivariate Analysis*, 98(6):1099–1122, 2007. MR2326242

[14] R.B. Dozier and J.W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information plus noise type matrices. *Journal of Multivariate Analysis*, 98(4):678–694, 2007. MR2322123

[15] K. Fan. Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Proceedings of the National Academy of Sciences of the United States of America*, 37(11):760, 1951. MR0045952

[16] V. L. Girko. *Theory of stochastic canonical equations. Vol. I,* volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.

[17] U. Haagerup and S. Thorbjornsen. A new application of random matrices: $\text{Ext}(C^*_{red}(F_2))$ is not a group. *Annals of Mathematics*, 162(2):711, 2005. MR2183281

[18] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. *Annals of Applied Probability*, 17(3):875–930, 2007. MR2326235

[19] B.A. Johnson, Y.I. Abramovich, and X. Mestre. MUSIC, G-MUSIC, and Maximum-Likelihood Performance Breakdown. *IEEE Transactions on Signal Processing*, 56(8):3944–3958, 2008. MR2517079

[20] I.M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Annals of Statistics*, 29(2):295–327, 2001. MR1863961

[21] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *IEEE Transactions on Information Theory*, 54(11):5113–5129, 2008. MR2589886

[22] X. Mestre and M.A. Lagunas. Modified subspace algorithms for DoA estimation with large arrays. *IEEE Transactions on Signal Processing*, 56(2):598, 2008. MR2445537

[23] B. Nadler. Nonparametric Detection of Signals by Information Theoretic Criteria: performance analysis and an improved estimator. *IEEE Transactions on Signal Processing*, 58(5):2746–2756, 2010. MR2789420

[24] E.A. Novikov. Functionals and the random-force method in turbulence theory(euler velocity field described by random forces method, using lagrangian representation of turbulence). *Soviet Physics-JETP*, 20:1290–1294, 1965. MR0191516

[25] L. Pastur. A simple approach to the global regime of Gaussian ensembles of random matrices. *Ukrainian Math. J.*, 57(6):936–966, 2005. MR2208456

[26] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, 17:1617–1642, 2007. MR2399865

[27] J.W. Silverstein and S. Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis*, 52(2):175–192, 1995.

[28] P. Vallet, P. Loubaton, and X. Mestre. Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case. *submitted*, 2010. arXiv: 1002.3234.