Spectral Measures of Spiked Random Matrices

Nathan Noiry

Abstract

We study two spiked models of random matrices under general frameworks corresponding respectively to additive deformation of random symmetric matrices and multiplicative perturbation of random covariance matrices. In both cases, the limiting spectral measure in the direction of an eigenvector of the perturbation leads to old and new results on the coordinates of eigenvectors.

MSC 2010 Classification: 60B20.
Keywords: Spiked random matrices, spectral measures, BBP phase transition, overlaps.

1 Introduction

The study of deformed models of random matrices has been the subject of tremendous amount of works in the last decades. In this paper, we study two of them. The first one corresponds to an additive perturbation of a symmetric random matrix (or Wigner matrix):

$$W_n = \frac{1}{\sqrt{n}} X_n + A_n$$

where

- X_n is a random symmetric matrix of size $n \times n$, whose entries are, up to symmetry, i.i.d. centered and reduced;
- A_n is a deterministic symmetric matrix of size $n \times n$ (or random, independent of X_n).

The second one is a multiplicative deformation of a random covariance matrix (or Wishart matrix):

$$S_n = \frac{1}{n} \Sigma_n^{1/2} X_n X_n^T \Sigma_n^{1/2},$$

where

- X_n is a random matrix of size $n \times n$, $m/n \to \alpha \in (0, \infty)$, whose entries are i.i.d. centered and reduced;
- Σ_n is a deterministic symmetric matrix (or random, independent of X_n) having non-negative eigenvalues.

The spectra of these models have been well studied. Let $\text{Spec}(W_n)$ (resp. $\text{Spec}(S_n)$) be the set of eigenvalues of W_n (resp. S_n) counted with multiplicities. The empirical spectral measures of W_n and S_n are:

$$\mu_{W_n} = \frac{1}{n} \sum_{\lambda \in \text{Spec}(W_n)} \delta_\lambda \quad \text{and} \quad \mu_{S_n} = \frac{1}{n} \sum_{\lambda \in \text{Spec}(S_n)} \delta_\lambda.$$

Under mild assumptions, they converge respectively towards probability measures $\mu_{sc} \boxplus \mu_A$ and $\mu_\alpha \boxtimes \mu_\Sigma$ properly defined in Subsections 2.1 and 3.1.

In both models, we define an outlier as an eigenvalue that does not lie in the neighborhood of the support of the limiting spectrum. Many works identified necessary and sufficient conditions on the spectrum of A_n (resp. Σ_n) for the appearance of outlier in the spectrum of W_n (resp. S_n). The seminal paper is due to Baik, Ben Arous and Péché [4], who identified a phase transition for the existence of an outlier in the covariance setting with $\Sigma_n = \mathrm{Diag}(\theta, 1, \ldots, 1)$, $\theta \geq 1$. The most recent results can be found in [5] and we refer to the survey [12] for an extensive bibliography. In all previous approaches, two main techniques were used: a clever identity on determinants (first remarked in [8]), and a precise analysis of the empirical spectral measure.

The goal of this paper is to bring into focus the possible use of the spectral measures in the study of deformed models of random matrices. Let us introduce them. Let θ be an eigenvalue of A_n (or Σ_n) which we consider as atypical in that it may be responsible for the existence of an outlier. Denote $v_1^{(n)}$ the associated eigenvector. We will call θ a spike of W_n (resp. S_n) and $v_1^{(n)}$ the direction of the spike. The spectral measures in the direction of the spike are respectively defined by:

$$\mu_{(W_n, v_1^{(n)})} := \sum_{\lambda \in \mathrm{Spec}(W_n)} \left| \langle \phi_\lambda, v_1^{(n)} \rangle \right|^2 \delta_\lambda \quad \text{and} \quad \mu_{(S_n, v_1^{(n)})} := \sum_{\lambda \in \mathrm{Spec}(S_n)} \left| \langle \phi_\lambda, v_1^{(n)} \rangle \right|^2 \delta_\lambda,$$

where ϕ_λ is a normalized eigenvector associated to eigenvalue λ. Note that unlike empirical spectral measures, these probability measures contain information on the eigenvectors of W_n and S_n. Following the well-known observation that outliers have associated eigenvectors which are localized in the direction of the spike, their influence should be present in $\mu_{(W_n, v_1^{(n)})}$ and $\mu_{(S_n, v_1^{(n)})}$ at a macroscopic level. Moreover, they can be easily studied as their Stieltjes transforms are given by the generalized entries of the resolvent ($\langle v_1^{(n)}, (W_n - z)^{-1} v_1^{(n)} \rangle$ and $\langle v_1^{(n)}, (S_n - z)^{-1} v_1^{(n)} \rangle$), for which many results already exist. In particular, as stated in Corollaries 1 and 4, $\mu_{(W_n, v_1^{(n)})}$ and $\mu_{(S_n, v_1^{(n)})}$ converge weakly towards deterministic probability measures denoted $\mu_{sc,A,\theta}$ and $\mu_{\alpha,\Sigma,\theta}$. We are going to present two applications of the spectral measures.

The first one recovers a classical result concerning the value of an outlier and the norm of its associated eigenvector projection in the direction of the spike.

The second one is concerned with the behavior of the projection of non-outlier eigenvectors in the direction of the spike. Namely, in the setting of an additive perturbation, if $f_{sc,A}$ and $f_{sc,A,\theta}$ are the respective densities of $\mu_{sc} \boxplus \mu_A$ and $\mu_{sc,A,\theta}$ and if x is in the support of $\mu_{sc} \boxplus \mu_A$, we prove the following convergence in probability

$$\frac{n}{\left| \{ \lambda \in \mathrm{Spec}(W_n),\ |\lambda - x| \leq \varepsilon_n \} \right|} \sum_{\lambda \in \mathrm{Spec}(W_n),\, |\lambda - x| \leq \varepsilon_n} \left| \left\langle \phi_\lambda, v_1^{(n)} \right\rangle \right|^2 \underset{n \to +\infty}{\longrightarrow} \frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)}, \tag{1}$$

for any sequence $n^{-1/2} \ll \varepsilon_n \ll 1$. In other words, the left-hand side, which is an average in the vicinity of x of the square-projections of eigenvectors in the direction of the spike, converges towards a deterministic profile. A similar result holds in the covariance setting. These are the content of Theorems 2 and 5. Our proof is inspired by the work of Benaych-Georges, Enriquez and Michaïl [6] and uses local laws estimates recently obtained by Knowles and Yin in [15]. When θ belongs to the support of the asymptotic spectrum of A_n (resp. Σ_n), Theorems 2 and 5 are *microscopic* confirmations of the results of Allez and Bouchaud [1] (in the Wigner setting) and Ledoit and Péché [17] (in the Wishart setting), who derived the asymptotic behavior of the overlaps $|\langle \phi_\lambda, v_\gamma \rangle|^2$ by taking the average over eigenvectors ϕ_λ associated to eigenvalues of W_n (resp. S_n) belonging to a *macroscopic* part of $\mathrm{Supp}\,(\mu_{sc} \boxplus \mu_A)$ (resp. $\mathrm{Supp}\,(\mu_\alpha \boxtimes \mu_\Sigma)$) and the average over eigenvectors v_γ of A_n (resp. Σ_n) belonging to a *macroscopic* part of the asymptotic spectrum of A_n (resp. Σ_n). When θ does not belong to the support of the asymptotic spectrum of A_n (resp. S_n), such a macroscopic average is not available whereas the spectral measure approach still works.

Interestingly, in rank one perturbation cases, that is when A_n (resp. Σ_n) has only one nonzero (resp. non one) eigenvalue θ, all the computations are explicit. This is because $\mu_{sc} \boxplus \mu_A = \mu_{sc}$ and $\mu_\alpha \boxtimes \mu_\Sigma = \mu_\alpha$ are explicit, respectively equal to the semicircle law and to the Marchenko-Pastur law with parameter α:

$$\mu_{sc}(\mathrm{dx}) = \frac{\sqrt{4-x^2}}{2\pi}\mathbf{1}_{|x|\le 2}\mathrm{d}x$$
$$\mu_\alpha(\mathrm{d}x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi x}\mathbf{1}_{(a,b)}(x)\mathrm{d}x + \mathbf{1}_{\alpha<1}(1-\alpha)\delta_0(\mathrm{d}x),$$

where $a, b = (1 \pm \sqrt{\alpha})^2$. In particular, we obtain the following formulas for the spectral measures in the direction of the spike:

$$\mu_{sc,\theta}(\mathrm{d}x) = \frac{\sqrt{4-x^2}}{2\pi(\theta^2+1-\theta x)}\mathbf{1}_{|x|\le 2}\mathrm{d}x + \mathbf{1}_{|\theta|>1}\left(1-\frac{1}{\theta^2}\right)\delta_{\theta+\frac{1}{\theta}}(\mathrm{d}x)$$
$$\mu_{\alpha,\theta}(\mathrm{d}x) = \frac{\theta\sqrt{(b-x)(x-a)}}{2\pi x\,(x(1-\theta)+\theta(\alpha\theta-\alpha+1))}\mathbf{1}_{(a,b)}(x)\mathrm{d}x + c_{\alpha,\theta}\delta_0(\mathrm{d}x) + d_{\alpha,\theta}\mathbf{1}_{|\theta-1|>\frac{1}{\sqrt{\alpha}}}\delta_{x_{\alpha,\theta}}(\mathrm{d}x),$$

where $c_{\alpha,\theta}$, $d_{\alpha,\theta}$ and $x_{\alpha,\theta}$ are explicit constants, see Propositions 2 and 4. Hence, by (1), the limiting profiles for the averaged square projections of non-outlier eigenvectors are also explicit. We give numerical simulations that agree with our predictions, see Subsections 2.2 and 3.2. In the covariance setting, Bloemendal, Knowles, Yau and Yin proved that individual square projections of non-outlier eigenvectors that are associated to eigenvalues in the vicinity of the edge (of b) converge towards a chi squared random variable with given variance (see [10, Theorem 2.20]). Although it requires an averaging step, our result completes the picture as it is concerned with eigenvectors associated to any fixed location of the bulk of the spectrum. We believe that the convergence still holds for smaller averaging windows and provide numerical simulations supporting this conjecture at the end of Section 5.

Let us finally say a few words about previous use of spectral measures in the literature. Benaych-Georges, Enriquez and Michaïl, in [6], obtained informations on the eigenvectors of a diagonal deterministic matrix perturbed by a random symmetric matrix. In a series of works (the most recent being [14]), Gamboa, Nagel and Rouault studied spectral measures of some classical ensembles of random matrix theory and their connections with sum rules. More related to our setting, in [2], Bai, Miao and Pan studied the spectral measure at the first vector of the canonical basis e_1 in general covariance cases, in the absence of spike.

We emphasize that our method could apply to other deformed models such as multiplicative perturbation of Wigner matrices or information plus noise matrices, but we chose to restrict our scope so that the present paper remains short and comprehensive.

Notations and organization of the paper. The random matrices that we study are built from an i.i.d. collection of real random variables $(X_{ij}^{(n)})$, $n \geq 1$, $1 \leq i, j \leq n$. Let X be a generic random variable with the same law. We suppose that $\mathbf{E}[X] = 0$, $\mathbf{E}[X^2] = 1$ and that X has moments of all order. Notice that the complex case could also be treated, replacing each transposed matrix A^T by its transposed-conjugate A^*, and making the hypothesis that $\mathbf{E}[|X|^2] = 1$.

For a complex $z \in \mathbf{C}$, we will denote $\Re(z)$ and $\Im(z)$ the real part and imaginary part of z.

For a probability measure ν, we always denote

$$s_\nu(z) := \int_{\mathbf{R}} \frac{\mathrm{d}\nu(x)}{x-z}$$

its Stieltjes transform that maps the upper half-plane to itself.

In Subsections 2.1 and 3.1 we give general results concerning additive perturbation of a Wigner matrix and multiplicative perturbation of a Wishart matrix. Subsections 2.2 and 3.2 provide explicit computations for rank-one deformation cases. The proofs are done in Sections 4 and 5.

2 Additive perturbation of a Wigner matrix

2.1 The general framework

In this section we consider for each $n \geq 1$ the following Wigner matrix:

$$X_n := \begin{bmatrix} X_{11}^{(n)} & X_{12}^{(n)} & \cdots & X_{1n}^{(n)} \\ X_{12}^{(n)} & X_{22}^{(n)} & \cdots & X_{2n}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1n}^{(n)} & X_{2n}^{(n)} & \cdots & X_{nn}^{(n)} \end{bmatrix}.$$

We also consider A_n a deterministic matrix (or random matrix independent of X_n) whose eigenvalues are $\gamma_1^{(n)} = \theta, \gamma_2^{(n)}, \dots, \gamma_n^{(n)}$, with associated eigenvectors $v_1^{(n)}, \dots, v_n^{(n)}$. We suppose that there exists a probability measure μ_A such that

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\gamma_i^{(n)}} \underset{n\to+\infty}{\longrightarrow} \mu_A \tag{2}$$

in the sense of weak convergence. We study the following additive perturbation model:

$$W_n := \frac{1}{\sqrt{n}} X_n + A_n.$$

We will denote $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ the eigenvalues of W_n and $\phi_1^{(n)}, \dots, \phi_n^{(n)}$ the associated normalized eigenvectors. Under assumption (2), it is known that the empirical spectral measure μ_{W_n} converges towards a deterministic probability measure which is the free convolution $\mu_{sc} \boxplus \mu_A$ between the semicircle law $\mu_{sc}(\mathrm{d}x) = (2\pi)^{-1}\sqrt{4-x^2}\mathbf{1}_{|x|\leq 2}\mathrm{d}x$ and μ_A. Its Stieltjes transform is characterized by

$$s_{\mu_{sc}\boxplus\mu_A}(z) = \int_{\mathbf{R}} \frac{d\mu_A(\lambda)}{\lambda - s_{\mu_{sc}\boxplus\mu_A}(z) - z}. \tag{3}$$

This has first been shown by Pastur in [18]. The parameter θ is considered as a spike which may create an outlier in the spectrum, that is an eigenvalue that does not lie in $\mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$.

Following the heuristic that an outlier in the spectrum creates a localized eigenvector, we study the spectral measure in the direction of the spike:

$$\mu_{(W_n, v_1^{(n)})} = \sum_{i=1}^{n} |\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2 \delta_{\lambda_i^{(n)}}.$$

The Stieltjes transform of $\mu_{(W_n, v_1^{(n)})}$ is given by $\langle v_1^{(n)}, (W_n - z)^{-1} v_1^{(n)} \rangle$ which is sometimes called a generalized entry of the resolvent and has already been studied in the literature. The most recent result is the local law recently obtained by Knowles and Yin [16]. It consists in a uniform estimation of $\langle v, (W_n - z)^{-1} w \rangle$ for any vectors v and w and for any complex z in a domain of the upper half plane that is allowed to approach the real axis as n tends to infinity. Since it is one ingredient of the proof of the forthcoming Theorem 2, we provide a precise statement.

Theorem 1. [16, Theorem 12.2] *Let*

$$G_n(z) = (W_n - z)^{-1} \quad \textit{and} \quad \Pi(z) = \left(A_n - z - s_{\mu_{sc}\boxplus\mu_A}(z)\right)^{-1},$$

Writing $z = E + i\eta$, let us introduce

$$\psi(z) := \sqrt{\frac{\Im\left(s_{\mu_{sc}\boxplus\mu_A}(z)\right)}{n\eta}} + \frac{1}{n\eta},$$

and, for any fixed $\tau > 0$:

$$\mathcal{D}_n^{(\tau)} := \left\{z \in \mathbf{C},\, |E| \le \tau^{-1},\, n^{-1+\tau} \le \eta \le \tau^{-1}\right\}.$$

Then, for any $\tau > 0$, uniformly in all vectors v, w and uniformly for in any $z \in \mathcal{D}_n^{(\tau)}$, for all $\varepsilon > 0$, there exists $D > 0$ such that

$$\mathbf{P}\left(|\langle v, G_n(z)w\rangle - \langle v, \Pi_n(z)w\rangle| \ge n^{\varepsilon}\psi(z)\|v\|\,\|w\|\right) \le \frac{1}{n^D}.$$

The non-local counterpart of Theorem 1 is the pointwise convergence of $\langle v, (W_n - z)^{-1}w\rangle$ in the domain $\Im(z) > 0$ and yields the following Corollary

Corollary 1. *The spectral measure $\mu_{(W_n, v_1^{(n)})}$ converges in probability towards a deterministic probability measure $\mu_{sc,A,\theta}$ whose Stieltjes transform is given by*

$$s_{\mu_{sc,A,\theta}}(z) = \frac{1}{\theta - s_{\mu_{sc}\boxplus\mu_A}(z) - z}. \tag{4}$$

Remark 1. *Equation (4) could allow to retrieve the limit of $\frac{1}{n}\mathrm{Tr}\left((W_n - z)^{-1}g(A_n)\right)$, obtained in [1] by Allez and Bouchaud, for any measurable function g. Indeed, this quantity can be rewritten*

$$\frac{1}{n}\sum_{i,j=1}^{n}\frac{|\langle\phi_i^{(n)}, v_j^{(n)}\rangle|^2}{\lambda_i^{(n)} - z}g\left(\gamma_j^{(n)}\right),$$

which is nothing but the average of the Stieltjes transform of the pushforward of the spectral measures of W_n by g. In particular, it converges to a non-degenerate limit only when the support of g is contained into a macroscopic part of $\mathrm{Supp}(\mu_A)$*, due to the renormalization by n. When g is non-null on a microscopic part of* $\mathrm{Supp}(\mu_A)$*, the study of the spectral measures allows to obtain the limit of* $\mathrm{Tr}\left((W_n - z)^{-1}g(A_n)\right)$ *whereas* $\frac{1}{n}\mathrm{Tr}\left((W_n - z)^{-1}g(A_n)\right)$ *brings no information as it converges to zero.*

Note that such a macroscopic result can be obtained using more simple arguments than the local law of Theorem 1. See for example [11, Proposition 6.2].

We provide two applications of the asymptotic behavior of the spectral measure of W_n in the direction of $v_1^{(n)}$.

The first one is concerned with outliers and the projection in the direction of the spike of their associated eigenvectors and relies on the following observation: unlike the empirical spectral measure which contains information on outliers only at the order $1/n$, the spectral measure in the direction of the spike already contains it at a *macroscopic* order. Let us introduce $w(x) = x + s_{\mu_{sc,A,\theta}}(x)$ which is well defined on $\mathbf{R} \setminus \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$. If there exists x such that $w(x) = \theta$, it is easy to deduce the existence of outliers for W_n as explained in the following Corollary. Although it is already-known in random matrix theory, our approach is new and relatively simple.

Corollary 2. *Suppose that there exists $x_\theta \notin \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ such that $w(x_\theta) = \theta$. Then, x_θ is an outlier of W_n. More precisely, set $\delta > 0$ such that $[x_\theta - \delta, x_\theta + \delta] \cap \mathrm{Supp}(\mu_{sc} \boxplus \mu_A) = \emptyset$ and define k_n to be the number of eigenvalues of W_n inside $[x_\theta - \delta, x_\theta + \delta]$. There exists $1 \le i_n \le n$ such that these eigenvalues satisfy*

$$x_\theta + \delta \ge \lambda_{i_n+1}^{(n)} \ge \lambda_{i_n+2}^{(n)} \ge \cdots \ge \lambda_{i_n+k_n}^{(n)} \ge x_\theta - \delta.$$

Then, $k_n \ge 1$ for n sufficiently large and:

1. *Both $\lambda_{i_n+1}^{(n)}$ and $\lambda_{i_n+k_n}^{(n)}$ converge in probability towards x_θ;*
2. *$\sum_{p=1}^{k_n} |\langle\phi_{i_n+p}^{(n)}, v_1^{(n)}\rangle|^2$ converges in probability towards $\frac{1}{w'(x_\theta)}$.*

Proof. Let $x_\theta \notin \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ be such that $w(x_\theta) = \theta$. The value of $\mu_{sc,A,\theta}(\{x_\theta\})$ is given by the residue of $s_{\mu_{sc,A,\theta}}$ at x_θ:

$$(x_\theta - z)s_{\mu_{sc,A,\theta}}(z) = \frac{x_\theta - z}{w(x_\theta) - w(z)} \underset{z\to x^+}{\longrightarrow} \frac{1}{w'(x_\theta)} > 0.$$

Since $\mu_{(W_n, v_1^{(n)})}$ converges towards $\mu_{sc,A,\theta}$ by Proposition 1, the Corollary is proved. □

In particular, when W_n is known to be a rank-one perturbation of a matrix W_n' whose empirical spectral measure converges towards $\mu_{sc} \boxplus \mu_A$ and which contains no outlier, the interlacing property implies that $k_n = 1$ for all n sufficiently large in Corollary 2 that is there exists a unique outlier which converges towards x_θ and the square projection in the direction of the spike of its associated eigenvector converges towards $1/w'(x_\theta)$.

Before stating our the second result, which represents the main novelty of this paper, and is also an illustration of the use of the spectral measure, we need the following observation:

Proposition 1. *$\mu_{sc,A,\theta}$ is absolutely continuous with respect to the Lebesgue measure on* $\mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$.

Proof. In [9], Biane proved that $\mu_{sc} \boxplus \mu_A$ is absolutely continuous with respect to the Lebesgue measure. Therefore, the inverse formula

$$\frac{\mathrm{d}\mu_{\mathrm{sc},A,\theta}(x)}{\mathrm{d}x} = \frac{1}{\pi}\lim_{t\to 0^+} \Im(s_{\mu_{sc,A,\theta}}(x+it)) \tag{5}$$

and Equation (4) imply that $\mu_{sc,A,\theta}$ is also absolutely continuous with respect to the Lebesgue measure at any $x \in \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$. □

We will denote $f_{sc,A}$ and $f_{sc,A,\theta}$ the respective densities of $\mu_{sc} \boxplus \mu_A$ and $\mu_{sc,A,\theta}$ on $\mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ (these are well-defined quantities by Proposition 1). It turns out that the averaged square-projections of the non-outlier eigenvectors associated to eigenvalues in the vicinity of $x \in \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ converges towards the ratio of these two densities.

Theorem 2. *Let $x \in \mathrm{Supp}(\mu_{sc} \boxplus \mu_A)$ be such that $f_{sc,A}(x) > 0$. Let ε_n be a sequence that satisfies $1/\sqrt{n} \ll \varepsilon_n \ll 1$. Then, for every $\delta > 0$, if $\mathcal{I}_{\varepsilon_n}^{(n)}(x) = \left\{1 \le i \le n : |\lambda_i^{(n)} - x| \le \varepsilon_n\right\}$:*

$$\mathbf{P}\left(\left|\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)} \left|\langle \phi_i^{(n)}, v_1^{(n)}\rangle\right|^2 - \frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)}\right| > \delta\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

By taking g the indicator of an interval contained in $\mathrm{Supp}(\mu_A)$ into the statistic introduced in Remark 1, Allez and Bouchaud [1] obtained the asymptotic behavior of the overlaps $|\langle \phi_i^{(n)}, v_j^{(n)}\rangle|^2$ after taking average over eigenvectors $\phi_i^{(n)}$'s (resp. $v_j^{(n)}$'s) with associated eigenvalues $\lambda_i^{(n)}$'s belonging to a *macroscopic* proportion of $\mathrm{Supp}(\mu_{sc} \boxplus \mathrm{Supp}(\mu_A))$ (resp. μ_A). When $\theta \in \mathrm{Supp}(\mu_A)$, Theorem 2 confirms their result at a *microscopic* scale. Indeed, denoting respectively a and b the real and imaginary parts of $\frac{1}{\pi}\lim_{t\to 0^+} \Im\left(s_{\mu_{sc}\boxplus\mu_A}(x+it)\right)$, one can rewrite, using the inverse formula (5):

$$\frac{f_{sc,A,\theta}(x)}{f_{sc,A}(x)} = \frac{b}{(a-\theta-x)^2 + b^2}.$$

When $\theta \notin \mathrm{Supp}(\mu_A)$, the approach of [1] provides no information on the overlap because it only gives access to $n^{-1}s_{\mu_{(W_n, v_1^{(n)})}}(z)$ which converges to zero, whereas the spectral measure approach still works.

2.2 The rank-one perturbation

In the special case where $\gamma_2^{(n)} = \cdots = \gamma_n^{(n)} = 0$ for all $n \ge 1$, W_n is a rank-one perturbation of a classical Wigner matrix. The limiting spectrum of the perturbation is $\mu_A = \delta_0$ and almost surely, μ_{W_n} weakly converges towards the semicircle distribution. In this setting, we provide explicit computations. Proposition 1 has now the more explicit formulation:

Proposition 2. *In probability, $\mu_{(W_n, v_1^{(n)})}$ converges towards:*

$$\mu_{sc,\theta}(\mathrm{d}x) := \frac{\sqrt{4-x^2}}{2\pi(\theta^2+1-\theta x)}\mathbf{1}_{|x|\leq 2}\mathrm{d}x + \mathbf{1}_{|\theta|>1}\left(1-\frac{1}{\theta^2}\right)\delta_{\theta+\frac{1}{\theta}}(\mathrm{d}x).$$

Remark that W_n is a rank-one perturbation of $n^{-1/2}X_n$. Therefore, since $\lambda_1(n^{-1/2}X_n) \to 2$ and $\lambda_n(n^{-1/2}X_n) \to -2$ in probability (see [13]), W_n has a single outlier whose location is given by the atom of $\mu_{sc,\theta}$ and whose associated eigenvector has a square projection in the direction of the spike given by the mass of this atom.

Corollary 3. *The following holds:*

1. *If $\theta > 1$, then, in probability, $\lambda_1(W_n) \underset{n\to+\infty}{\longrightarrow} \theta + \frac{1}{\theta} > \ 2$ and $|\langle \phi_1^{(n)}, v_1^{(n)}\rangle| \underset{n\to+\infty}{\longrightarrow} \sqrt{1-\frac{1}{\theta^2}}$.*
2. *If $\theta < -1$, then, in probability, $\lambda_n(W_n) \underset{n\to+\infty}{\longrightarrow} \theta + \frac{1}{\theta} < -2$ and $|\langle \phi_n^{(n)}, v_1^{(n)}\rangle| \underset{n\to+\infty}{\longrightarrow} \sqrt{1-\frac{1}{\theta^2}}$.*

Finally, the averaged square-projections have also an explicit form, which is just the inverse of a linear function in that case:

Theorem 3. *Let $x \in (-2, 2)$. Let ε_n be a sequence that satisfies $1/\sqrt{n} \ll \varepsilon_n \ll 1$. Then, for every $\delta > 0$,*

$$\mathbf{P}\left(\left|\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)}\left|\langle\phi_i^{(n)}, v_1^{(n)}\rangle\right|^2 - \frac{1}{\theta^2-\theta x+1}\right| > \delta\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

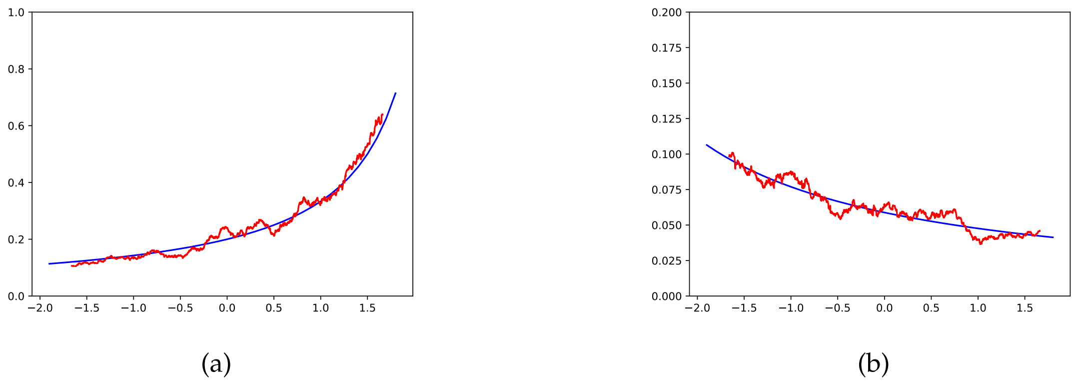


Figure 1: In red: simulations of the average squared-projections around all locations $x \in (-2, 2)$ where we took average over intervals of typical size $n^{0,1}/\sqrt{n}$ for a single matrix $W_n = n^{-1/2}X_n$, where X_n has gaussian entries and is of size 3000×3000. In case 2a $\theta = 2$ and in case 2b $\theta = -4$. In blue: theoretical predictions.

3 Multiplicative perturbation of a Wishart matrix

3.1 The general framework

Let $m = m(n)$ be a sequence of integers such that $m/n \to \alpha > 0$ as $n \to +\infty$. For all $n \geq 0$, we consider the following random rectangular matrix:

$$X_n := \begin{bmatrix} X_{11}^{(n)} & X_{12}^{(n)} & \cdots & X_{1m}^{(n)} \\ X_{21}^{(n)} & X_{22}^{(n)} & \cdots & X_{2m}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1}^{(n)} & X_{n2}^{(n)} & \cdots & X_{nm}^{(n)} \end{bmatrix}.$$

Let also Σ_n be a general covariance matrix of size $n \times n$, with eigenvalues given by $\gamma_1^{(n)} = \theta, \gamma_2^{(n)}, \ldots, \gamma_n^{(n)}$ and associated eigenvectors $v_1^{(n)}, \ldots, v_n^{(n)}$. We suppose that there exists a probability measure μ_Σ such that

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{\gamma_i^{(n)}} \underset{n \to +\infty}{\longrightarrow} \mu_\Sigma$$

in the sense of weak convergence. We study the following multiplicative perturbation model:

$$S_n := \frac{1}{n} \Sigma_n^{1/2} X_n X_n^T \Sigma_n^{1/2}.$$

The matrix S_n can be considered as the sampled covariance matrix of m i.i.d. vectors in $\mathbf{R}^n$ having covariance matrix Σ_n. Let $\lambda_1^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ be the eigenvalues of S_n and $\phi_1^{(n)}, \ldots, \phi_n^{(n)}$ the associated normalized eigenvectors.

The empirical spectral distribution of S_n converges towards the free product $\mu_\alpha \boxtimes \mu_\Sigma$ whose Stieltjes transform is characterized by:

$$s_{\mu_\alpha \boxtimes \mu_\Sigma}(z) = \int_{\mathbf{R}} \frac{\mathrm{d}\mu_\Sigma(t)}{t(\alpha - 1 - z s_{\mu_\alpha \boxtimes \mu_\Sigma}(z)) - z}.$$

This is a consequence of the work of Silverstein [19].

We think of θ as a spike, that is an atypical eigenvalue compared to the sequence $\gamma_i^{(n)}$, $2 \leq i \leq n$. We are interested in its influence on the apparition of outliers for S_n. To that purpose, we study the spectral measure in the direction of the spike:

$$\mu_{(S_n, v_1^{(n)})} := \sum_{i=1}^{n} |\langle \phi_i^{(n)}, v_1^{(n)} \rangle|^2 \delta_{\lambda_i^{(n)}}.$$

The Stieltjes transform of $\mu_{(S_n, v_1^{(n)})}$ is given by $\langle v_1^{(n)}, (S_n - z)^{-1} v_1^{(n)} \rangle$ and has already been studied in the literature. As in the Wigner case, the most recent result is the local law obtained by Knowles and Yin [16]. It consists in a uniform estimation of $\langle v, (S_n - z)^{-1} w \rangle$ for any vectors v and w and for any complex z in a domain of the upper half plane that is allowed to approach the real axis as n tends to infinity. Since it is an ingredient of the proof of Theorem 5, we provide a precise statement.

Theorem 4. [16, Theorem 12.2] *Let $G_n(z) = (S_n - z)^{-1}$. Writing $z = E + i\eta$, let us introduce*

$$\psi(z) := \sqrt{\frac{\Im\left(s_{\mu_\alpha \boxtimes \mu_\Sigma}(z)\right)}{n\eta} + \frac{1}{n\eta}},$$

and, for any fixed $\tau > 0$:

$$\mathcal{D}_n^{(\tau)} := \left\{ z \in \mathbf{C},\ |z| \geq \tau,\ |E| \leq \tau^{-1},\ n^{-1+\tau} \leq \eta \leq \tau^{-1} \right\}.$$

Then, for any $\tau > 0$, uniformly in all vectors v, w and uniformly for in any $z \in \mathcal{D}_n^{(\tau)}$, for all $\varepsilon > 0$, there exists $D > 0$ such that

$$\mathbf{P}\left(\left| \langle v, \Sigma_n^{-1/2} \left(G_n(z) - z(1 + s_{\mu_\alpha \boxtimes \mu_\Sigma}(z)\Sigma_n) \right) \Sigma_n^{-1/2} w \rangle \right| \geq n^\varepsilon \psi(z) \|v\| \, \|w\| \right) \leq \frac{1}{n^D}. \tag{6}$$

The non-local counterpart of Theorem 1 is the pointwise convergence of $\langle v, (S_n - z)^{-1} w \rangle$ in the domain $\Im(z) > 0$.

Corollary 4. *In probability, $\mu_{(S_n, v_1^{(n)})}$ weakly converges towards a probability measure $\mu_{\alpha,\Sigma,\theta}$ whose Stieltjes transform is given by:*

$$s_{\mu_{\alpha,\Sigma,\theta}}(z) = \frac{1}{\theta(\alpha - 1) - z\theta s_{\mu_\alpha \boxtimes \mu_\Sigma}(z) - z}. \tag{7}$$

Remark 2. *In principle, (7) allows to retrieve the limit of $\frac{1}{n}\mathrm{Tr}\left((S_n - z)^{-1} g(\Sigma_n)\right)$, obtained in [17] by Ledoit and Péché, for any measurable function g. Indeed, this quantity can be rewritten*

$$\frac{1}{n} \sum_{i,j=1}^{n} \frac{|\langle \phi_i^{(n)}, v_j^{(n)} \rangle|^2}{\lambda_i^{(n)} - z} g\left(\gamma_j^{(n)}\right),$$

which is nothing but the average of the Stieltjes transform of the pushforward of the spectral measures of S_n by g. In particular, it converges to a non-degenerate limit only when the support of g is contained into a macroscopic part of $\mathrm{Supp}(\mu_\Sigma)$, due to the renormalization by n. When g is non-null on a microscopic part of $\mathrm{Supp}(\mu_\Sigma)$, the study of the spectral measures yields the limit of $\mathrm{Tr}\left((S_n - z)^{-1} g(\Sigma_n)\right)$ whereas $\frac{1}{n}\mathrm{Tr}\left((S_n - z)^{-1} g(\Sigma_n)\right)$ brings no information as it converges to zero.

Of course, such a macroscopic result can be obtained using more simple arguments than the local law of Theorem 4. We refer for example to [11, Proposition 6.2].

In what follows, we provide two applications of the asymptotic behavior of the spectral measure of S_n in the direction of $v_1^{(n)}$.

The first one recovers a classical result on outliers of S_n and the projection in the direction of the spike of their associated eigenvectors. The proof we propose is simple and based on the following observation: unlike the empirical spectral measure which contains information on outliers only at the order $1/n$, the spectral measure in the direction of the spike already contains it at a *macroscopic* order. Before stating our result, let us introduce $F(x) = \alpha x - x - s_{\mu_\alpha \boxtimes \mu_\Sigma}(1/x)$ which is well defined for every $x > 0$ such that $x^{-1} \in \mathbf{R} \setminus \mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma)$. If there exists $x > 0$ such that $1/F(1/x) = \theta$, we easily obtain the existence of an outlier as explained in the following Corollary.

Corollary 5. *Suppose that there exists $x_{\alpha,\theta} \notin \mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma)$ such that $1/F(1/x_{\alpha,\theta}) = \theta$. Then, $x_{\alpha,\theta}$ is an outlier of S_n. More precisely, set $\delta > 0$ such that $[x_{\alpha,\theta} - \delta, x_{\alpha,\theta} + \delta] \cap \mathrm{Supp}(\mu_{sc} \boxplus \mu_A) = \emptyset$ and define k_n to be the number of eigenvalues of S_n inside $[x_{\alpha,\theta} - \delta, x_{\alpha,\theta} + \delta]$. There exists $1 \leq i_n \leq n$ such that these eigenvalues satisfy*

$$x_{\alpha,\theta} + \delta \geq \lambda_{i_n+1}^{(n)} \geq \lambda_{i_n+2}^{(n)} \geq \cdots \geq \lambda_{i_n+k_n}^{(n)} \geq x_{\alpha,\theta} - \delta.$$

Then, $k_n \geq 1$ for n sufficiently large and:

1. *Both $\lambda_{i_n+1}^{(n)}$ and $\lambda_{i_n+k_n}^{(n)}$ converge in probability towards $x_{\alpha,\theta}$;*
2. *$\sum_{p=1}^{k_n} |\langle \phi_{i_n+p}^{(n)}, v_1^{(n)} \rangle|^2$ converges in probability towards $\frac{x_{\alpha,\theta} F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})}$.*

Proof. Let $x_{\alpha,\theta} > 0$ be such that $x_{\alpha,\theta} \notin \mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma)$ and $1/F(1/x_{\alpha,\theta}) = \theta$. The value of $\mu_{\alpha,\Sigma,\theta}(\{x_{\alpha,\theta}\})$ is given by the residue of $s_{\mu_{\alpha,\Sigma,\theta}}$ at $x_{\alpha,\theta}$:

$$(x_{\alpha,\theta} - z)s_{\mu_{\alpha,\Sigma,\theta}}(z) = \frac{(x_{\alpha,\theta} - z)F(1/x_{\alpha,\theta})}{z(F(1/z) - F(1/x_{\alpha,\theta}))} \underset{z\to x_{\alpha,\theta}^+}{\longrightarrow} \frac{x_{\alpha,\theta}F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})} > 0.$$

Since $\mu_{(S_n,v_1^{(n)})}$ converges towards $\mu_{\alpha,\Sigma,\theta}$, the Corollary is easily deduced. □

A particular case is when S_n is a rank-one perturbation of a matrix S'_n which has no outlier and whose empirical spectral measure converges towards $\mu_\alpha \boxtimes \mu_\Sigma$. In that setting, the interlacing property implies that $k_n = 1$ for all n sufficiently large in Corollary 5, meaning that S_n has only one outlier which converges towards $x_{\alpha,\theta}$ and whose associated eigenvector has a square projection in the direction of the spike which converges towards $\frac{x_{\alpha,\theta}F(1/x_{\alpha,\theta})}{F'(1/x_{\alpha,\theta})}$.

Before stating our second result, which is concerned with the projection of non-outlier eigenvectors onto the direction of the spike, we need the following Proposition.

Proposition 3. *$\mu_{\alpha,\Sigma,\theta}$ is absolutely continuous with respect to the Lebesgue measure on* $\mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma) \setminus \{0\}$.

Proof. Let $x \in \mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma) \setminus \{0\}$. The work of Choï and Silverstein [20] ensures that $\mu_\alpha \boxtimes \mu_\Sigma$ is absolutely continuous with respect to the Lebesgue measure at x. Then, the inverse formula for the Stieltjes transform:

$$\frac{\mathrm{d}\mu_{\mathrm{sc},\mathrm{A},\theta}(x)}{\mathrm{d}x} = \frac{1}{\pi}\lim_{t\to 0^+} \Im(s_{\mu_{\mathrm{sc},A,\theta}}(x+it)), \tag{8}$$

combined with Equation (7), implies the Proposition. □

Let $f_{\alpha,\Sigma}$ and $f_{\alpha,\Sigma,\theta}$ be the respective densities of $\mu_\alpha \boxtimes \mu_\Sigma$ and $\mu_{\alpha,\Sigma,\theta}$ on $\mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma)$. It turns out that the averaged square projections of the non-outlier eigenvectors associated to eigenvalues in the vicinity of $x \in \mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma)$ converges towards the ratio of these two densities.

Theorem 5. *Let $x \in \mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma) \setminus \{0\}$ be such that $f_{\alpha,\Sigma}(x) > 0$. Let ε_n be a sequence that satisfies $1/\sqrt{n} \ll \varepsilon_n \ll 1$. Then, for every $\delta > 0$, if $\mathcal{I}^{(n)}_{\varepsilon_n}(x) = \left\{1 \le i \le n : |\lambda_i^{(n)} - x| \le \varepsilon_n\right\}$:*

$$\mathbf{P}\left(\left|\frac{n}{|\mathcal{I}^{(n)}_{\varepsilon_n}(x)|}\sum_{i\in\mathcal{I}^{(n)}_{\varepsilon_n}(x)} \left|\langle\phi_i^{(n)}, v_1^{(n)}\rangle\right|^2 - \frac{f_{\alpha,\Sigma,\theta}(x)}{f_{\alpha,\Sigma}(x)}\right| > \delta\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

As in the Wigner case, Theorem 5 can be seen as a generalization of a result of [17], where the authors obtained the asymptotic behavior of the overlaps $|\langle\phi_i^{(n)}, v_j^{(n)}\rangle|^2$ after taking average over eigenvectors $\phi_i^{(n)}$'s (resp. $v_j^{(n)}$'s) with associated eigenvalues $\lambda_i^{(n)}$'s belonging to a *macroscopic* proportion of $\mathrm{Supp}(\mu_\alpha \boxtimes \mu_\Sigma)$ (resp. μ_Σ), by taking taking functions of the type $g = \mathbf{1}_{[\gamma,+\infty)}$ in the statistics introduced in Remark 2. Indeed, when $\theta \in \mathrm{Supp}(\mu_\Sigma)$, Theorem 5 is a *microscopic* version of the result of [17, Theorem 3]. To obtain their formula it suffices to remark that, if a and b are the real and imaginary parts of $1 - \alpha - zs_{\mu_\alpha\boxtimes\mu_\Sigma}(z)$, one can rewrite, using Equation (8):

$$\frac{f_{\alpha,\Sigma,\theta}(x)}{f_{\alpha,\Sigma}(x)} = \frac{x\theta}{(a\theta - x)^2 + \theta^2 b^2},$$

When $\theta \notin \mathrm{Supp}(\mu_\Sigma)$, the techniques of [17] provide no information on the overlaps as it only gives access to $n^{-1}s_{\mu_{(S_n,v_1^{(n)})}}(z)$, which converges to zero, whereas the spectral measure approach still works.

3.2 Rank-one perturbation of the Marchenko-Pastur law

In the peculiar case where $\gamma_2^{(n)} = \cdots = \gamma_n^{(n)} = 1$ for all $n \ge 1$, S_n is a rank-one perturbation of a classical Wishart matrix. The limiting spectrum of the perturbation is $\mu_\Sigma = \delta_1$ and almost surely, μ_{S_n} weakly converges towards the Marchenko-Pastur law μ_α. All previous results have now a more explicit formulation. First, the limit of the spectral measure in the direction of the spike can be identified.

Proposition 4. *In probability, $\mu_{(S_n, v_1^{(n)})}$ converges towards:*

$$\mu_{\alpha,\theta}(\mathrm{d}x) = \frac{\theta\sqrt{(b-x)(x-a)}}{2\pi x(x(1-\theta)+\theta(\alpha\theta-\alpha+1))}\mathbf{1}_{(a,b)}(x)\mathrm{d}x + c_\alpha \mathbf{1}_{\alpha<1}\delta_0(\mathrm{d}x) + d_{\alpha,\theta}\mathbf{1}_{|\theta-1|>\frac{1}{\sqrt{\alpha}}}\delta_{x_{\alpha,\theta}}(\mathrm{d}x),$$

where $c_\alpha = \frac{1-\alpha}{\alpha(\theta-1)+1}$, $d_{\alpha,\theta} = \frac{1-\frac{1}{\alpha(\theta-1)^2}}{1+\frac{1}{\alpha(\theta-1)}}$ and $x_{\alpha,\theta} = \frac{\theta(\alpha\theta-\alpha+1)}{\theta-1}$.

A consequence of [3] is that $n^{-1}X_nX_n^T$ has no outlier. Since S_n is a rank one perturbation of this matrix, the discussion following Corollary 5 implies that S_n has a single outlier.

Corollary 6. *The following holds:*

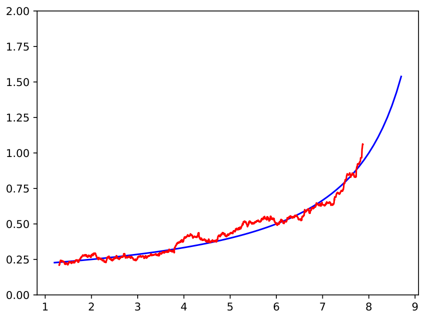
1. *If $\theta > 1 + 1/\sqrt{\alpha}$, then, in probability, $\lambda_1(S_n) \underset{n\to+\infty}{\longrightarrow} x_{\alpha,\theta} > b$ and $|\langle \phi_1^{(n)}, v_1^{(n)}\rangle| \underset{n\to+\infty}{\longrightarrow} \sqrt{d_{\alpha,\theta}}$.*
2. *If $\theta < 1 - 1/\sqrt{\alpha}$, then, in probability, $\lambda_n(S_n) \underset{n\to+\infty}{\longrightarrow} x_{\alpha,\theta} < a$ and $|\langle \phi_n^{(n)}, v_1^{(n)}\rangle| \underset{n\to+\infty}{\longrightarrow} \sqrt{d_{\alpha,\theta}}$.*

The ratio of the density of $\mu_{\alpha,\theta}$ and μ_α is explicit and we obtain the following Theorem.

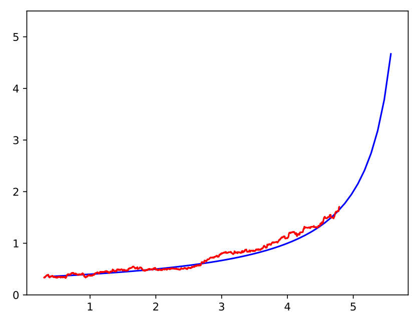
Theorem 6. *Let $x \in (a,b)$. Let ε_n be a sequence that satisfies $1/\sqrt{n} \ll \varepsilon_n \ll 1$. Then, for every $\delta > 0$,*

$$\mathbf{P}\left(\left|\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x)|}\sum_{i\in\mathcal{I}_{\varepsilon_n}^{(n)}(x)}\left|\langle\phi_i^{(n)}, v_1^{(n)}\rangle\right|^2 - \frac{\theta}{x(1-\theta)+\theta(\alpha\theta-\alpha+1)}\right| > \delta\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

When x tends to b, the limiting profile becomes $\theta/(1+\sqrt{\alpha}(1-\theta)^2)$, in accordance with [10, Theorem 2.20] where the authors obtain a convergence of individual square-projections onto the direction of the spike towards chi-squared random variables with expectation $\theta/(1+\sqrt{\alpha}(1-\theta)^2)$. A natural question would be to study an analog convergence in law in the bulk of the spectrum (for any $x \in (a,b)$). We do not pursue this issue here.



(a)



(b)

Figure 2: In red: simulations of the average squared projections around all locations $x \in (a,b)$ where we took average over interval of typical size $n^{0,1}/\sqrt{n}$, for a single matrix $S_n = n^{-1}\mathrm{Diag}(\sqrt{\theta}, 1, \dots, 1)X_nX_n^T\mathrm{Diag}(\sqrt{\theta}, 1, \dots, 1)$ where X_n is gaussian rectangular of size 2000×8000 ($\alpha = 4$) in case 2a and of size 2000×4000 ($\alpha = 2$) in case 2b. In each case $\theta = 2$. In blue: theoretical predictions.

4 Identification of the limiting laws in rank-one perturbation cases

In this section we prove Proposition 2 and 4. We use the following branch of the complex square-root:

$$\sqrt{z} = \operatorname{sign}(\Im(z)) \frac{|z| + z}{\sqrt{2(|z| + \Re(z))}}.$$

Proof of Proposition 2. The Stieltjes transform of the semicircle law is given by:

$$s_{\mu_{sc}}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}.$$

Therefore, using Equation (4) and the fact that μ_A is in this case δ_0:

$$\begin{aligned} s_{\mu_{sc,\theta}}(z) &= \frac{2}{2\theta - z - \sqrt{z^2 - 4}} \\ &= \frac{\sqrt{z^2 - 4} + 2\theta - z}{2(\theta^2 - \theta z + 1)}. \end{aligned}$$

The absolutely continuous part of $\mu_{sc,\theta}$ is given by

$$\frac{\mathrm{d}\mu_{sc,\theta}(x)}{\mathrm{d}x} = \frac{1}{\pi} \lim_{z \to x^+} \Im(s_{\mu_{sc,\theta}}(z)) = \frac{\sqrt{4 - x^2}}{2\pi(\theta^2 + 1 - \theta x)} \mathbf{1}_{|x| \leq 2} \mathrm{d}x.$$

The atom at $\theta + 1/\theta$ is given by the corresponding residue of $s_{\mu_{sc,\theta}}$:

$$-\lim_{z \to (\theta + 1/\theta)^+} (z - \theta - 1/\theta) s_{\mu_{sc,\theta}}(z).$$

By our choice of square-root, $\lim_{z \to \theta + 1/\theta} \sqrt{z^2 - 4} = |\theta - 1|/\theta$ and one easily deduces:

$$\mu_{sc,\theta}(\{\theta + 1/\theta\}) = \begin{cases} 0 & \text{if } |\theta| \leq 1 \\ 1 - \frac{1}{\theta^2} & \text{if } |\theta| > 1. \end{cases}$$

□

Proof of Proposition 4. Recall the expression of the Stieltjes transform of the Marchenko-Pastur law $\mu_\alpha = \mu_{\alpha,1}$:

$$s_{\mu_\alpha}(z) = \frac{\alpha - z - 1 + \sqrt{(z - b)(z - a)}}{2z}.$$

Substituting in Equation (7), we get

$$\begin{aligned} s_{\mu_{\alpha,\theta}}(z) &= \frac{-2}{2z - 2\theta(\alpha - 1) + \theta(\alpha - 1) - \theta z + \theta \sqrt{(z - b)(z - a)}} \\ &= \frac{-2}{2z - \theta(\alpha - 1) - \theta z + \theta \sqrt{(z - b)(z - a)}} \\ &= \frac{-2\left(2z - \theta(\alpha - 1) - \theta z - \theta \sqrt{(z - b)(z - a)}\right)}{((2 - \theta)z - \theta(\alpha - 1))^2 - \theta^2(z - b)(z - a)} \\ &= \frac{\theta \sqrt{(z - b)(z - a)} + z(\theta - 2) + \theta(\alpha - 1)}{2z\,(z(1 - \theta) + \theta(\alpha\theta - \alpha + 1))}. \end{aligned} \tag{9}$$

This expression will allow us to obtain an explicit formula for $\mu_{\alpha,\theta}$, through classical inversion results.

The absolutely continuous part of $\mu_{\alpha,\theta}$ is given by:

$$\frac{\mathrm{d}\mu_{\alpha,\theta}}{\mathrm{d}x}(x) = \frac{1}{\pi}\lim_{z\to x^+} \Im(s_{\mu_{\alpha,\theta}}(z)) = \frac{\theta\sqrt{(b-x)(x-a)}}{2\pi x\,(x(1-\theta)+(\alpha\theta-\alpha+1))}\mathbf{1}_{(a,b)}(x).$$

The atom of $\mu_{\alpha,\theta}$ at zero is given by:

$$-\lim_{\varepsilon\to 0^+} i\varepsilon s_{\mu_{\alpha,\theta}}(i\varepsilon) = -\lim_{\varepsilon\to 0^+}\frac{\sqrt{(i\varepsilon-b)(i\varepsilon-a)}+(\alpha-1)}{2(\alpha\theta-\alpha+1)}.$$

By our choice of square-root that preserves the upper-half plane, $\sqrt{(i\varepsilon-b)(i\varepsilon-a)} \to -|ab| = -|\alpha-1|$ as $\varepsilon\to 0^+$. Therefore:

$$\mu_{\alpha,\theta}(\{0\}) = \frac{|\alpha-1|-(\alpha-1)}{2(\alpha\theta-\alpha+1)} = \begin{cases} 0 & \text{if } \alpha\geq 1 \\ \frac{1-\alpha}{\alpha(\theta-1)+1} & \text{if } \alpha<1. \end{cases}$$

Finally, let us compute the atom at $x_\theta = \frac{\theta(\alpha\theta-\alpha+1)}{\theta-1}$. It is given by:

$$-\lim_{z\to x_\theta^+}(z-x_\theta)s_{\mu_{\alpha,\theta}}(z) = \frac{\sqrt{(x_\theta-b)(x_\theta-a)}+x_\theta(\theta-2)+\theta(\alpha-1)}{2(\theta-1)x_\theta}.$$

We use the following relations which are easily verified.

- $x_\theta - b = \frac{\alpha\left((\theta-1)-\frac{1}{\sqrt{\alpha}}\right)^2}{\theta-1}$,
- $x_\theta - a = \frac{\alpha\left((\theta-1)+\frac{1}{\sqrt{\alpha}}\right)^2}{\theta-1}$.

As for the computation of the atom at zero:

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\sqrt{(x_\theta+i\varepsilon-b)(x_\theta+i\varepsilon-a)} &= \mathrm{sign}(\theta-1)\frac{\alpha}{|\theta-1|}\left|(\theta-1)^2-\frac{1}{\alpha}\right| \\ &= \frac{\alpha}{\theta-1}\left|(\theta-1)^2-\frac{1}{\alpha}\right|.\end{aligned}$$

Besides:

$$\begin{aligned} x_\theta(\theta-2)+\theta(\alpha-1) &= \frac{\theta(\alpha\theta-\alpha+1)(\theta-2)+\theta(\alpha-1)(\theta-1)}{\theta-1} \\ &= \frac{\theta}{\theta-1}\left(\alpha\theta^2-2\alpha\theta+\alpha-1\right) \\ &= \frac{\alpha\theta}{\theta-1}\left((\theta-1)^2-\frac{1}{\alpha}\right). \end{aligned}$$

One obtains:

$$\mu_{\alpha,\theta}(\{x_{\alpha,\theta}\}) = \alpha\theta\frac{\left|(\theta-1)^2-\frac{1}{\alpha}\right|+\left((\theta-1)^2-\frac{1}{\alpha}\right)}{2\theta(\theta-1)(\alpha\theta-\alpha+1)} = \begin{cases} 0 & \text{if } |\theta-1|\leq 1/\sqrt{\alpha} \\ \frac{\alpha\left((\theta-1)^2-\frac{1}{\alpha}\right)}{(\theta-1)(\alpha\theta-\alpha+1)} & \text{if } |\theta-1|>1/\sqrt{\alpha}. \end{cases}$$

□

5 Convergence of the averaged square projections

We only focus on the proof of Theorem 5 concerning the convergence of averaged square-projections into the direction of the spike in the Wishart setting. The proof of Theorem 2, which concerns the

Wigner setting, would follow the same reasoning, the only difference being the use of the local law of Theorem 1 instead of the local law of Theorem 4.

Let us explain the heuristic behind Theorem 5. For any $x_0 \in \mathbf{R}$, we will denote $I_{\varepsilon_n}(x_0) := [x_0 - \varepsilon_n, x_0 + \varepsilon_n]$. Recall that $f_{\alpha,\Sigma}$ and $f_{\alpha,\Sigma,\theta}$ are the respective densities of $\mu_\alpha \boxtimes \mu_\Sigma$ and $\mu_{\alpha,\Sigma,\theta}$ on Supp $(\mu_\alpha \boxtimes \mu_\Sigma)$. Then,

$$\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n, v_1^{(n)})}(x) \approx \int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{\alpha,\theta}(x) + o_1(1) \approx 2\varepsilon_n f_{\alpha,\Sigma,\theta}(x_0) + o_1(1).$$

On the other hand, if $\mu_{S_n} = n^{-1}\sum_{1\leq i\leq n} \delta_{\lambda_i^{(n)}}$ denotes the empirical spectral measure of S_n:

$$\begin{aligned}\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n, v_1^{(n)})}(x) = \sum_{i\in \mathcal{I}_{\varepsilon_n}^{(n)}(x_0)} \left|\langle \phi_i, e_1\rangle\right|^2 &= \left(\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)|}\sum_{i\in \mathcal{I}_{\varepsilon_n}^{(n)}(x_0)} \left|\langle \phi_i^{(n)}, v_1^{(n)}\rangle\right|^2\right) \times \int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{S_n}(x) \\ &\approx \left(\frac{n}{|\mathcal{I}_{\varepsilon_n}^{(n)}(x_0)|}\sum_{i\in \mathcal{I}_{\varepsilon_n}^{(n)}(x_0)} \left|\langle \phi_i, e_1\rangle\right|^2\right) \times (2\varepsilon_n f_{\alpha,\Sigma}(x_0) + o_2(1)),\end{aligned}$$

where we recall that $\mathcal{I}_{\varepsilon_n}^{(n)}(x_0) = \{1 \leq i \leq n : |\lambda_i^{(n)} - x_0| \leq \varepsilon_n\}$.

Theorem 6 would be proved if the errors $o_1(1)$ and $o_2(1)$ were explicit and of a smaller order than ε_n. The understanding of these errors is precisely the purpose of the so-called local laws that have been recently developed in random matrix theory. In the Wishart setting, it is given by the local law of Knowles and Yin stated in Theorem 4. Taking $v = w = v_1^{(n)}$ is Equation (6) implies that for any $\varepsilon > 0$, there exists $D > 0$ such that

$$\mathbf{P}\left(\left|s_{(S_n, v_1^{(n)})}(z) - s_{\mu_{\alpha,\Sigma,\theta}}(z)\right| \geq n^\varepsilon \psi(z)\right) \leq \frac{1}{n^D} \tag{10}$$

Note that Equation 10 is uniform in all $z \in \mathcal{D}_n^{(\tau)}$ where we recall that:

$$\mathcal{D}_n^{(\tau)} := \left\{z = E + i\eta \in \mathbf{C},\, |z| \geq \tau,\, |E| \leq \tau^{-1},\, n^{-1+\tau} \leq \eta \leq \tau^{-1}\right\},$$

where $\tau > 0$ is fixed.

Remark 3. *When $z \in \mathcal{D}_n^{(\tau)}$ is such that $f_{\alpha,\Sigma}(E) > 0$, then the imaginary part of $s_{\mu_\alpha\boxtimes\mu_\Sigma}(E + it)$ as $t > 0$ approaches the real axis. In that case, the error term $\psi(z)$ is of order:*

$$\psi(z) = \sqrt{\frac{\Im\left(s_{\mu_\alpha\boxtimes\mu_\Sigma}(z)\right)}{n\eta}} + \frac{1}{n\eta} = O\left((n\eta)^{-1/2}\right).$$

This estimate will be repeatedly used in the proof of Theorem 5 as it is concerned with points $x \in$ Supp $(\mu_\alpha \boxtimes \mu_\Sigma)$ such that $f_{\alpha,\Sigma}(x) > 0$.

The rest of this section consists in an approximation argument based on the Helffer-Sjöstrand formula, which allows to translate an estimate of the form (10) into an estimate on sufficiently regular functions integrated against $\mu_{(S_n, v_1^{(n)})}$. Although the argument is standard and can be found for the empirical spectral measure of a Wigner matrix in the survey of Benaych-Georges and Knowles [7], we choose to provide the details as it has not been done for the spectral measures. A similar argument is also present in the work of Benaych-Georges, Enriquez and Michaïl [6].

Proof of Theorem 6. Let $x_0 \in (a, b)$ and let ω_n and ε_n be two sequences such that $n^{-1/2} \ll \omega_n \ll \varepsilon_n \ll 1$. We will denote

$$I_{\varepsilon_n}(x_0) := [x_0 - \varepsilon_n, x_0 + \varepsilon_n].$$

We bound the indicator function of $I_{\varepsilon_n}(x_0)$ using two smooth regularization functions ϕ_n^- and ϕ_n^+ such that:

- the support of ϕ_n^- (resp. ϕ_n^+) is included in $[x_0 - \varepsilon_n + 2\omega_n, x_0 + \varepsilon - 2\omega_n]$ (resp. $[x_0 - \varepsilon_n - 2\omega_n, x_0 + \varepsilon + 2\omega_n]$);
- ϕ_n^- (resp. ϕ_n^+) is constant equal to 1 on $[x_0 - \varepsilon_n + \omega_n, x_0 + \varepsilon_n - \omega_n]$ (resp. $[x_0 - \varepsilon_n - \omega_n, x_0 + \varepsilon + \omega_n]$);
- $\|(\phi_n^-)'\|_\infty = \|(\phi_n^+)'\|_\infty = O(1/\omega_n)$ and $\|(\phi_n^-)''\|_\infty = \|(\phi_n^+)''\|_\infty = O(1/\omega_n^2)$.

By construction,

$$\int_{\mathbf{R}} \phi_n^-(\lambda)\mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda) \leq \int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda) \leq \int_{\mathbf{R}} \phi_n^+(\lambda)\mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda). \tag{11}$$

Let us write $\phi_n^+ = \phi_n$ and estimate the right-hand side, as the computations are the same for the left-hand side. Define the signed measure $\hat{\mu}_n := \mu_{(S_n,v_1^{(n)})} - \mu_{\alpha,\Sigma,\theta}$ and denote $\hat{s}_n$ its associated Stieltjes transform. We are going to prove that for any $\varepsilon > 0$, there exists $D > 0$ such that with probability at least $1 - n^{-D}$:

$$\left| \int_{\mathbf{R}} \phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda) \right| = O\left(\frac{n^\varepsilon}{n^{1/2}}\right). \tag{12}$$

Before turning to the proof of (12), we explain how it leads to Theorem 5.

By Equation (11), for any $\varepsilon > 0$, there exists $D > 0$ such that with probability at least $1 - n^{-D}$:

$$\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n,v_1^{(n)})}(x) = 2\varepsilon_n f_{\alpha,\Sigma,\theta}(x_0) + O\left(\frac{n^\varepsilon}{n^{1/2}}\right). \tag{13}$$

It is possible to obtain an analog estimation on the empirical spectral measure μ_{S_n}: for any $\varepsilon > 0$, there exists $D > 0$ such that with probability at least $1 - n^{-D}$:

$$\left| \int_{\mathbf{R}} \phi_n(\lambda) \left(\mathrm{d}\mu_{S_n} - \mathrm{d}\mu_\alpha \boxtimes \mu_\Sigma\right)(\lambda) \right| = O\left(\frac{n^\varepsilon}{n}\right).$$

The proof is the same as the proof of (12) that we provide below. Note that the $n^{1/2}$ of the denominator is replaced by an n, because at the level of empirical spectral measures, the error term in the local law is $\frac{1}{n\eta}$ instead of $\left(\frac{\Im\left(s_{\mu_\alpha \boxtimes \mu_\Sigma}(z)\right)}{n\eta}\right)^{-1/2}$ (see Theorem 3.6 of [16]). Therefore, for any $\varepsilon > 0$, there exists $D > 0$ such that with probability at least $1 - n^{-D}$:

$$\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{S_n}(x) = 2\varepsilon_n f_{\alpha,\Sigma}(x_0) + O\left(\frac{n^\varepsilon}{n}\right). \tag{14}$$

Putting estimates (13) and (14) together, Theorem (5) is proved because $\varepsilon_n \gg n^{-1/2}$ and

$$\frac{n}{\left|\mathcal{I}_{\varepsilon_n}(x_0)\right|} \sum_{i \in \mathcal{I}_{\varepsilon_n}(x_0)} \left|\langle \phi_i, e_1 \rangle\right|^2 = \frac{\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{(S_n,v_1^{(n)})}(\lambda)}{\int_{I_{\varepsilon_n}(x_0)} \mathrm{d}\mu_{S_n}(\lambda)}.$$

We now turn to the proof of (12). The main idea is to use the Helffer-Sjöstrand formula which states that for all $x \in \mathbf{R}$:

$$\phi_n(x) = \int_{\mathbf{C}} \frac{\overline{\partial}\left(\tilde{\phi}_n(z)\chi(z)\right)}{x - z} \mathrm{d}z, \tag{15}$$

where:

- χ is a smooth cutoff function that equals 1 on $[-1, 1]$ and 0 outside $[-2, 2]$;
- $\tilde{\phi}_n$ is the quasi-analytic extension of degree 1 of ϕ_n, defined by $\tilde{\phi}_n(x + iy) = \phi_n(x) + iy\phi_n'(x)$;
- $\overline{\partial} = \frac{1}{2}\left(\partial_n + i\partial_y\right)$.

The proof, based on Green's formula, can be found in [7] (Proposition C.1). Equation (15) leads to:

$$\int_{\mathbf{R}} \phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda) = \frac{i}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}} \phi_n''(x)y\chi(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y + \frac{i}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}} \left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y,$$

Remark that the right-hand-side is real so that

$$\int_{\mathbf{R}} \phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda) \leq \frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\leq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y \quad (16)$$

$$+\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\geq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y \quad (17)$$

$$+\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}} \left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y\right|. \quad (18)$$

We estimate all of the three terms of the right-hand side.

For the first term, we use the two following facts:

- $\Im\left(\hat{s}_n(z)\right) \leq \Im\left(s_{(S_n,v_1^{(n)})}(z)\right)$;
- $y \mapsto y\Im\left(s_{(S_n,v_1^{(n)})}(x+iy)\right)$ is non-decreasing.

Moreover, the integral of ϕ'' over x is of order $\|\phi'\|_\infty = O(1/\omega_n)$. Therefore, there exists some constant $C>0$ such that:

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\leq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right| \leq C\frac{1}{\omega_n}\int_{|y|\leq\omega_n}\omega_n \sup_{x\in\mathbf{R}}\left(\Im\left(s_{(S_n,v_1^{(n)})}(x+i\omega_n)\right)\right)\mathrm{d}y$$

$$\leq C\omega_n \sup_{x\in\mathbf{R}}\left(\Im\left(s_{(S_n,v_1^{(n)})}(x+i\omega_n)\right)\right).$$

But, for any probability measure ν on $\mathbf{R}$, if s_ν denotes its Stieltjes transform, then, for any $t>0$, $\sup_{x\in\mathbf{R}} s_\nu(x+it)$ is bounded as can be easily checked. Therefore:

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\leq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right| = O(\omega_n). \quad (19)$$

For the second term, remark first that

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\geq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right| \leq 2\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\geq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right| \quad (20)$$

because $\hat{s}_n(x-iy) = \overline{\hat{s}_n(x+iy)}$. Differentiating with respect to x and y and using that $\partial_x\Im(\hat{s}_n(x+iy)) = -\partial_y\Re(\hat{s}_n(x+iy))$, we obtain:

$$\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\geq\omega_n} \phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y$$

$$= \frac{-1}{2\pi}\int_{y\geq\omega_n} y\chi(y)\mathrm{d}y \times (-1)\int_{x\in\mathbf{R}} \phi_n'(x)\partial_x\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x$$

$$= \frac{-1}{2\pi}\int_{x\in\mathbf{R}} \phi_n'(x)\mathrm{d}x\left(\omega_n\Im\left(\hat{s}_n(x+i\omega_n)\right) - \int_{y>\omega_n}(y\chi'(y)+\chi(y))\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}y\right). \quad (21)$$

The first term of (21) satisfies:

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\phi_n'(x)\omega_n\Im\left(\hat{s}_n(x+i\omega_n)\right)\mathrm{d}x\right|=O\left(\omega_n\sup_{x\in\mathbf{R}}\left(\Im\left(\hat{s}_n(x+i\omega_n)\right)\right)\right)=O(\omega_n)$$

Let $\varepsilon>0$. By the local law (10), there exists $D>0$ such that on an event of probability at least $1-n^{-D}$, $|\hat{s}_n(z)|=O(n^{\varepsilon}\psi(z))$. On this event, the other two terms of (21) are bounded as follows:

- $\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\geq\omega_n}\phi_n'(x)y\chi'(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}y\mathrm{d}x\right|=O\left(\int_{x\in\mathbf{R}}\int_{y=1}^{2}\phi_n'(x)y|\hat{s}_n(x+iy)|\mathrm{d}y\mathrm{d}x\right)=O\left(\frac{n^{\varepsilon}}{n^{1/2}}\right)$, because for $1\leq y\leq 2$, $\psi(x+iy)=O\left(\frac{1}{n^{1/2}}\right)$.
- $\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\geq\omega_n}\phi_n'(x)\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right|=O\left(\int_{y=\omega_n}^{2}\frac{n^{\varepsilon}}{n^{1/2}y^{1/2}}\mathrm{d}y\right)=O\left(\frac{n^{\varepsilon}}{n^{1/2}}\right)$.

Hence, in view of (20) and (21), we proved that for any $\varepsilon>0$, there exists $D>0$ such that with probability at least $1-n^{-D}$:

$$\left|\frac{-1}{2\pi}\int_{x\in\mathbf{R}}\int_{|y|\geq\omega_n}\phi_n''(x)y\chi(y)\Im\left(\hat{s}_n(x+iy)\right)\mathrm{d}x\mathrm{d}y\right|=O\left(\max\left\{\frac{n^{\varepsilon}}{n^{1/2}},\omega_n\right\}\right).\tag{22}$$

Finally, the third term is bounded as follows, using that χ' is supported on $[-2,-1]\cup[1,2]$:

$$\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}}\left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y\right|\\ \leq\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{1\leq|y|\leq 2}\left(\phi_n(x)+2|\phi_n'(x)|\right)|\chi'(y)|\cdot|\hat{s}_n(x+iy)|\mathrm{d}x\mathrm{d}y.$$

Since $\phi_n(x)$ and $\phi_n'(x)$ are equal to zero when $|x|$ is sufficiently large and since the integrals of these functions are of order one, the local law (10) implies that, for any $\varepsilon>0$, there exists $D>0$ such that with probability at least $1-n^{-D}$:

$$\left|\frac{1}{2\pi}\int_{x\in\mathbf{R}}\int_{y\in\mathbf{R}}\left(\phi_n(x)+iy\phi_n'(x)\right)\chi'(y)\hat{s}_n(x+iy)\mathrm{d}x\mathrm{d}y\right|=O\left(\frac{n^{\varepsilon}}{n^{1/2}}\right),\tag{23}$$

where we used that $\psi(x+iy)$ is of order $n^{-1/2}$ for $1\leq|y|\leq 2$.

Putting estimates (19), (22), (23) together and recalling Equation (18), we proved that for any $\varepsilon>0$, there exists $D>0$ such that with probability at least $1-n^{-D}$:

$$\left|\int_{\mathbf{R}}\phi_n(\lambda)\mathrm{d}\hat{\mu}_n(\lambda)\right|=O\left(\max\left\{\frac{n^{\varepsilon}}{n^{1/2}},\omega_n\right\}\right).$$

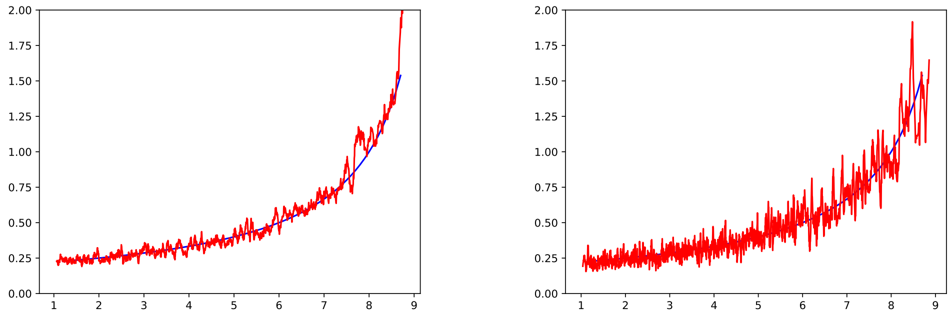
Since $\omega_n\gg n^{-1/2}$ was arbitrary, the estimation (12) on the integral of ϕ_n against $\hat{\mu}_n$ is proved. This ends the proof of Theorem 5. □

We underline that a weaker version of Theorem 5 (resp. Theorem 2) can be obtained as long as a uniform estimation of $s_{(S_n,v_1^{(n)})}(z)-s_{\mu_{\alpha,\Sigma,\theta}}(z)$ is available for z in a domain of the upper half-plane that is allowed to approach the real axis as n becomes larger. Indeed, if

$$\left|s_{(S_n,v_1^{(n)})}(z)-s_{\mu_{\alpha,\Sigma,\theta}}(z)\right|=O\left(\varepsilon_n\right),$$

then, the Helffer-Sjöstrand argument that we developed during the proof of Theorem 5 yields a convergence of the averaged-square projection onto the direction of the spike for averaging windows of size ε_n, as long as $\varepsilon\gg n^{-1/2}$. The limitation $\varepsilon_n\gg n^{-1/2}$ corresponds to the optimal rate in the local laws of Knowles and Yin (10).

One natural question would be to weaken the assumption on the size ε_n of the averaging window: do our Theorems 3, 6, 2 and 5 hold as soon as $\varepsilon_n = o(1)$? We believe that the answer is positive (see Figure 3 for simulations). Morevoever, the results of [10], which state the convergence in law of properly rescaled *individual* square projection of eigenvectors associated to eigenvalues in the vicinity of the edge, suggest the following natural question: does such a convergence also holds in the bulk of the spectrum?



(a) (b)

Figure 3: In red: simulations of the average squared projections around all locations $x \in (a, b)$ where we took average over interval of typical size $n^{0,3}$ for 3a and $n^{0,2}$ for 3b, for 10 independent matrices of the form $S_n = n^{-1}\mathrm{Diag}(\sqrt{\theta}, 1, \dots, 1)X_nX_n^T\mathrm{Diag}(\sqrt{\theta}, 1, \dots, 1)$ where X_n is gaussian rectangular of size 2000×8000 ($\alpha = 4$) in case 2a and of size 3000×12000 ($\alpha = 4$) in case 2b. In each case $\theta = 2$. In blue: theoretical predictions.

Acknowledgment I am very thankful to Nathanaël Enriquez for many suggestions about this work. Many thanks to Maxime Février for numerous insightful discussions and for his careful reading of an earlier version of this paper. I would also like to thank Laurent Ménard for his advices.

References

[1] Romain Allez and Jean-Philippe Bouchaud. Eigenvector dynamics under free addition. *Random Matrices Theory Appl.*, 3(3):1450010, 17, 2014.

[2] Z. D. Bai, B. Q. Miao, and G. M. Pan. On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.*, 35(4):1532–1572, 2007.

[3] Z. D. Bai and Jack 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] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.

[5] Serban T. Belinschi, Hari Bercovici, Mireille Capitaine, and Maxime Février. Outliers in the spectrum of large deformed unitarily invariant models. *Ann. Probab.*, 45(6A):3571–3625, 2017.

[6] Florent Benaych-Georges, Nathanaël Enriquez, and Alkéos Michaïl. Eigenvectors of a matrix under random perturbation. *arXiv preprint arXiv:1801.10512*, 2018.

[7] Florent Benaych-Georges and Antti Knowles. Local semicircle law for Wigner matrices. In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 1–90. Soc. Math. France, Paris, 2017.

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

[9] Philippe Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.*, 46(3):705–718, 1997.

[10] Alex Bloemendal, Antti Knowles, Horng-Tzer Yau, and Jun Yin. On the principal components of sample covariance matrices. *Probab. Theory Related Fields*, 164(1-2):459–552, 2016.

[11] M. Capitaine. Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices. *J. Theoret. Probab.*, 26(3):595–648, 2013.

[12] Mireille Capitaine and Catherine Donati-Martin. Spectrum of deformed random matrices and free probability. In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 151–190. Soc. Math. France, Paris, 2017.

[13] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[14] Fabrice Gamboa, Jan Nagel, and Alain Rouault. Sum rules and large deviations for spectral matrix measures in the jacobi ensemble. *arXiv preprint arXiv:1811.06311*, 2018.

[15] Antti Knowles and Jun Yin. The outliers of a deformed Wigner matrix. *Ann. Probab.*, 42(5):1980–2031, 2014.

[16] Antti Knowles and Jun Yin. Anisotropic local laws for random matrices. *Probab. Theory Related Fields*, 169(1-2):257–352, 2017.

[17] Olivier Ledoit and Sandrine Péché. Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Related Fields*, 151(1-2):233–264, 2011.

[18] L. A. Pastur. The spectrum of random matrices. *Teoret. Mat. Fiz.*, 10(1):102–112, 1972.

[19] Jack W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.*, 55(2):331–339, 1995.

[20] 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.

Nathan Noiry :
Laboratoire Modal'X,
UPL, Université Paris Nanterre,
F92000 Nanterre France