



The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices

Florent Benaych-Georges, Raj Rao

► To cite this version:

Florent Benaych-Georges, Raj Rao. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. 2009. hal-00423593v2

HAL Id: hal-00423593

<https://hal.science/hal-00423593v2>

Preprint submitted on 28 Apr 2010 (v2), last revised 27 Dec 2010 (v3)

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

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

THE EIGENVALUES AND EIGENVECTORS OF FINITE, LOW RANK PERTURBATIONS OF LARGE RANDOM MATRICES

FLORENT BENAYCH-GEORGES AND RAJ RAO NADAKUDITI

ABSTRACT. We consider the eigenvalues and eigenvectors of finite, low rank perturbations of random matrices. Specifically, we prove almost sure convergence of the extreme eigenvalues and appropriate projections of the corresponding eigenvectors of the perturbed matrix for additive and multiplicative perturbation models.

The limiting non-random value is shown to depend explicitly on the limiting spectral measure and the assumed perturbation model via integral transforms that correspond to very well known objects in free probability theory that linearize non-commutative free additive and multiplicative convolution. Furthermore, we uncover a remarkable phase transition phenomenon whereby the large matrix limit of the extreme eigenvalues of the perturbed matrix differs from that of the original matrix if and only if the eigenvalues of the perturbing matrix are above a certain critical threshold.

This critical threshold is intimately related to the same aforementioned integral transforms and our proof techniques bring this connection and the origin of the phase transition into exceedingly sharp focus. Consequently, our results extend the class of ‘spiked’ random matrix models about which such predictions can be made well beyond the Gaussian, Wishart and Jacobi random ensembles found in the literature. We examine the impact of this eigenvalue phase transition on the associated eigenvectors and provide a sufficient condition when an analogous phase transition in the eigenvectors is observed. Various extensions of our results to the problem of non extreme eigenvalues are discussed.

1. INTRODUCTION

Let X_n be an $n \times n$ symmetric (or Hermitian) matrix with eigenvalues $\lambda_1(X_n), \dots, \lambda_n(X_n)$ and P_n be an $n \times n$ symmetric (or Hermitian) matrix with rank $r \leq n$ and eigenvalues $\theta_1, \dots, \theta_r$. A fundamental question in matrix analysis is the following [14, 2]:

How are the eigenvalues and eigenvectors of $X_n + P_n$ related to the eigenvalues and eigenvectors of X_n and P_n ?

Date: April 28, 2010.

2000 Mathematics Subject Classification. 15A52, 46L54, 60F99.

Key words and phrases. Random matrices, Haar measure, principal components analysis, informational limit, free probability, phase transition, random eigenvalues, random eigenvectors, random perturbation, sample covariance matrices.

F.B.G’s work was partially supported by the *Agence Nationale de la Recherche* grant ANR-08-BLAN-0311-03. R.R.N’s research was partially supported by an Office of Naval Research postdoctoral fellowship award and grant N00014-07-1-0269. R.R.N thanks Arthur Baggeroer for his feedback, support and encouragement. We thank Alan Edelman for feedback and encouragement and for facilitating this collaboration by hosting F.B.G’s stay at M.I.T. We gratefully acknowledge the Singapore-MIT alliance for funding F.B.G’s stay.

When X_n and P_n are diagonalized by the same eigenvectors, we have $\lambda_i(X_n + P_n) = \lambda_j(X_n) + \lambda_k(P_n)$ for appropriate choice of indices $i, j, k \in \{1, \dots, n\}$. In the general setting, however, the answer is much more complicated because the eigenvalues and eigenvectors of their sum depend on the relationship between the eigenspaces of the individual matrices.

In this scenario, one can use Weyl's interlacing inequalities [23] to obtain coarse bounds for the eigenvalues of the sum in terms of the eigenvalues of X_n . When the norm of P_n is small relative to the norm of X_n , tools from perturbation theory (see [23, Chapter 6] or [37]) can be employed to improve the characterization of the bounded set in which the eigenvalues of the sum must lie. Exploiting any special structure in the matrices allows us to refine these bounds [25] but this is pretty much as far as the theory goes. Instead of exact answers we have a system of messy, coupled bounds. The eigenvector story is even more convoluted.

Surprisingly, adding some randomness to the eigenspaces permits further analytical progress. Specifically, if the eigenspaces are assumed to be “in generic position with respect to each other”, then analytical elegance returns in the limit of large matrices.

In place of eigenvalue bounds we have simple, exact answers that are to be interpreted probabilistically. These results bring into sharp focus a remarkable phase transition phenomenon of the kind illustrated in Figure 1 for the eigenvalues and eigenvectors of $X_n + P_n$ and $X_n(I_n + P_n)$. A precise statement of the results may be found in Section 2.

Examining the structure of the analytical expression for θ_c and ρ in Figure 1 reveals a common underlying theme in the additive and multiplicative perturbation settings. The critical values θ_c and ρ in Figure 1 are related to integral transforms of the limiting spectral measure μ_X of X_n . It turns out that the integral transforms that emerge in the respective additive and multiplicative cases are deeply related to very well known objects in free probability theory [38, 22] that linearize free additive and multiplicative [38] convolutions respectively. In a forthcoming paper [12], we consider the analogue of problem for the extreme singular values of finite rank deformations of rectangular random matrices. There too, a phase transition occurs at a threshold determined by an integral transform which plays an analogous role in the computation of the rectangular additive free convolution [7, 8, 9]. The emergence of these transforms in the context of the study of the extreme/isolated eigenvalue behavior should be of independent interest to free probabilists. This justifies our anointment of the study of finite, low rank perturbations of large random matrices as “spiked” free probability theory. In this framework, regular free probability theory would correspond the study of full rank perturbations of large random matrices.

The development of spiked free probability theory is the main contribution of this paper. In doing so, we extend the results found in the literature for the eigenvalue phase transition in such finite, low rank perturbation models well beyond the Gaussian [4, 5, 32, 24, 19, 15, 6], Wishart [18, 31, 29] and Jacobi settings [28]. In our situation, the distribution μ_X in Figure 1 can be *any* probability measure. Consequently, the aforementioned results in the literature can be rederived rather simply using the formulas in Section 2 by substituting μ_X with the semi-circle measure [39] (for Gaussian matrices), the Marčenko-Pastur measure [27] (for Wishart matrices) or the free Jacobi measure (for Jacobi matrices [16]). Concrete computations are presented in Section 3.

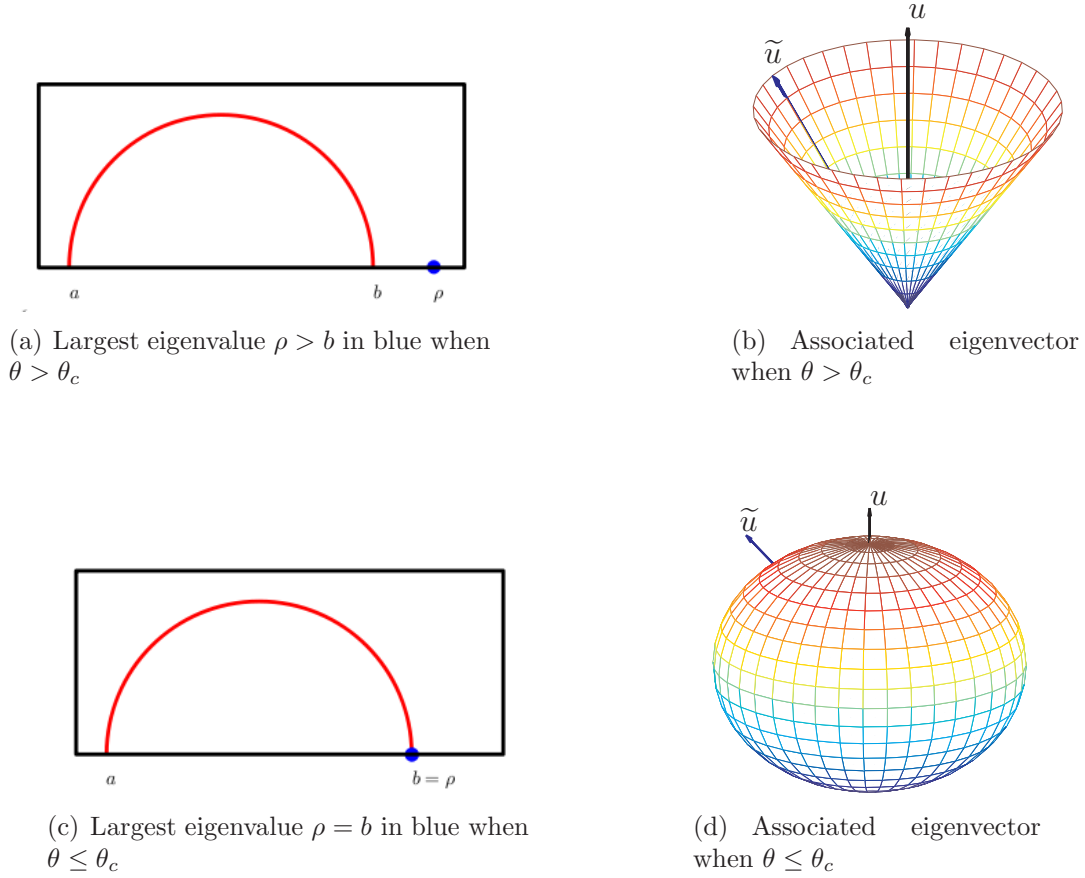


FIGURE 1. Assume that the limiting spectral measure of X_n is μ_X with largest eigenvalue b . Consider the matrix $P_n := \theta uu^*$ with rank $r = 1$ and largest eigenvalue $\theta (> 0$ say). The vector u is an $n \times 1$ vector chosen uniformly at random from the unit n -sphere. The largest eigenvalue of $X_n + P_n$ will differ from b if and only if θ is greater than some critical value θ_c . In this event, the largest eigenvalue will be concentrated around ρ with high probability as in (a). The associated eigenvector \tilde{u} will, with high probability, lie on a cone around u as in (b). When $\theta \leq \theta_c$, a phase transition occurs so that with high probability, the largest eigenvalue of the sum will equal b as in (c) and the corresponding eigenvector will be uniformly distributed on the unit sphere as in (d).

The development of the eigenvector aspect of the story is another important contribution that we would like to highlight. Generally speaking, the eigenvector question has received much less much attention in random matrix theory and in free probability theory despite impressive breakthroughs [13]. A notable exception is the recent body of work on the eigenvectors of spiked Wishart matrices [31, 24, 29] which corresponds to μ_X being the Marčenko-Pastur measure. In this paper, we extend their results for multiplicative models of the kind $X_n(I + P_n)$ to the setting where μ_X is an arbitrary probability measure and obtain new results for the eigenvectors for additive models.

Our proofs rely heavily on the derivation of master equation representations of the eigenvalues and eigenvectors of the perturbed matrix and the subsequent application of concentration inequalities for random vectors uniformly distributed on high dimensional unit spheres (such as the ones appearing in [20, 21]) to these implicit master equation representations. Consequently, our technique is simpler, more general and brings into sharper focus the source of the phase transition phenomenon than other proofs found in the literature. The underlying methods can and have been adapted to study the extreme singular values and singular vectors of deformations of rectangular random matrices, as well as the fluctuations [10] and the large deviations [11] of our model.

The paper is organized as follows. In Section 2, we state the main results and present the integral transforms alluded to above. Section 3 presents some examples. A high-level sketch of the proof is presented in Section 4. Exact master equation representations of the eigenvalues and the eigenvectors of the perturbed matrices are derived in Section 5 and utilized in Sections 6 to 8 to prove the main results. Technical results needed in these proofs have been relegated to the Appendix.

2. MAIN RESULTS

2.1. Definitions. Let X_n be an $n \times n$ symmetric (or Hermitian) random matrix with eigenvalues $\lambda_1(X_n) \geq \dots \geq \lambda_n(X_n)$. We denote by μ_{X_n} the empirical distribution on the set of its eigenvalues, *i.e.*, the probability measure defined as

$$\mu_{X_n} = \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j(X_n)}.$$

Assume that the probability measure μ_{X_n} converges almost surely weakly, as n tends to infinity, to a non-random, compactly supported probability measure μ_X . Let the smallest and largest eigenvalues of X_n converge almost surely to a and b which are, respectively, the infimum and supremum of the support of μ_X .

Let us fix a finite family $\theta_1 \geq \dots \geq \theta_r$ of non-zero real numbers (which are independent of n and non random). For every n , let P_n be an $n \times n$ symmetric (or Hermitian) random matrix independent of X_n , having rank r with the r non-zero eigenvalues equal to $\theta_1, \dots, \theta_r$. We introduce $s \in \{0, \dots, r\}$ such that $\theta_1 \geq \dots \geq \theta_s > 0 > \theta_{s+1} \geq \dots \geq \theta_r$.

We assume that X_n or P_n or both are invariant, in law, by conjugation by any orthogonal (or unitary) matrix.

For an arbitrary $n \times n$ Hermitian matrix M , we denote the ordered eigenvalues of M by $\lambda_1(M) \geq \dots \geq \lambda_n(M)$. Let $\xrightarrow{\text{a.s.}}$ denote almost sure convergence. Lastly, for a subspace F of a Euclidian space E and a vector $x \in E$, we denote the norm of the orthogonal projection of x onto F by $\langle x, F \rangle$.

2.2. Extreme eigenvalues and eigenvectors under additive perturbations. Consider the rank r additive perturbation of the random matrix X_n given by

$$\tilde{X}_n = X_n + P_n.$$

Theorem 2.1 (Eigenvalue phase transition). *The extreme eigenvalues of \tilde{X}_n exhibit the following behavior as $n \rightarrow \infty$. We have that for each $i \geq 1$,*

$$\lambda_i(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} G_{\mu_X}^{-1}(1/\theta_i) & \text{if } 1 \leq i \leq s \text{ and } 1/\theta_i < G_{\mu_X}(b^+), \\ b & \text{otherwise,} \end{cases}$$

while for each $i \geq 0$,

$$\lambda_{n-i}(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} G_{\mu_X}^{-1}(1/\theta_{r-i}) & \text{if } s < r-i \text{ and } 1/\theta_i > G_{\mu_X}(a^-), \\ a & \text{otherwise,} \end{cases}$$

where

$$G_{\mu_X}(z) = \int \frac{1}{z-t} d\mu_X(t) \quad \text{for } z \notin \text{supp } \mu_X,$$

is the Cauchy transform of μ_X and the exponent -1 refers to the inverse with respect to the composition \circ .

Theorem 2.2 (Norm of eigenvector projection). *Consider indices $i_0 \in \{1, \dots, r\}$ such that $1/\theta_{i_0} \in (G_{\mu_X}(a^-), G_{\mu_X}(b^+))$. For each n , define*

$$z_n := \begin{cases} \lambda_{i_0}(\tilde{X}_n) & \text{if } \theta_{i_0} > 0, \\ \lambda_{n-r+i_0}(\tilde{X}_n) & \text{if } \theta_{i_0} < 0, \end{cases}$$

and let x_n be a unit eigenvector of \tilde{X}_n associated with the eigenvalue z_n . Then we have

(a)

$$\langle x_n, \ker(\theta_{i_0} I_n - P_n) \rangle^2 \xrightarrow{a.s.} \frac{-1}{\theta_{i_0}^2 G'_{\mu_X}(\rho)}$$

where $\rho = G_{\mu_X}^{-1}(1/\theta_{i_0})$ is the limit of z_n ;

(b)

$$\langle x_n, \oplus_{j \neq i_0} \ker(\theta_j I_n - P_n) \rangle \xrightarrow{a.s.} 0,$$

as $n \rightarrow \infty$.

Theorem 2.3 (Eigenvector phase transition). *When $r = 1$, let the sole non-zero eigenvalue of P_n be denoted by θ . Suppose that*

$$\frac{1}{\theta} \notin (G_{\mu_X}(a^-), G_{\mu_X}(b^+)), \quad \text{and} \quad \begin{cases} G'_{\mu_X}(b^+) = -\infty & \text{if } \theta > 0, \\ G'_{\mu_X}(a^-) = -\infty & \text{if } \theta < 0. \end{cases}$$

For each n , let x_n be a unit eigenvector of \tilde{X}_n associated with either the largest or smallest eigenvalue depending on whether $\theta > 0$ or $\theta < 0$, respectively. Then we have

$$\langle x_n, \text{ran}(P_n) \rangle \xrightarrow{a.s.} 0,$$

as $n \rightarrow \infty$.

Remark 2.4 (Necessity of eigenvalue repulsion for the eigenvector phase transition). Under additional hypotheses on the manner in which the spectral measure of X_n tends to μ_X as n tends to infinity, Theorem 2.2 can be generalized to any eigenvalue with limit ρ in $\{a, b\}$ such that $G'_{\mu_X}(\rho)$ is finite and Theorem 2.3 can be generalized for any value of r . The specific hypothesis has to do with requiring the spacings between the $\lambda_i(X_n)$'s to be

more “random matrix like” and exhibit repulsion instead of being “independent sample like” with possible clumping.

2.3. Extreme eigenvalues and eigenvectors under multiplicative perturbations.

We maintain the same hypotheses as before so that the limiting probability measure μ_X , the index s and the rank r matrix P_n are defined as in Section 2.1. In addition, we assume that for every n , X_n is a non-negative definite matrix and that the limiting probability measure μ_X is not the Dirac mass at zero.

Consider the rank r multiplicative perturbation of the random matrix X_n given by

$$\tilde{X}_n = X_n(I_n + P_n).$$

Theorem 2.5 (Eigenvalue phase transition). *The extreme eigenvalues of \tilde{X}_n exhibit the following behavior as $n \rightarrow \infty$. We have that for each $i \geq 1$*

$$\lambda_i(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} T_{\mu_X}^{-1}(1/\theta_i) & \text{if } 1 \leq i \leq s \text{ and } 1/\theta_i < T_{\mu_X}(b^+), \\ b & \text{otherwise,} \end{cases}$$

while for each $i \geq 0$,

$$\lambda_{n-r+i}(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} T_{\mu_X}^{-1}(1/\theta_i) & \text{if } s < r - i \text{ and } 1/\theta_i > T_{\mu_X}(a^-), \\ a & \text{otherwise,} \end{cases}$$

where

$$T_{\mu_X}(z) = \int \frac{t}{z-t} d\mu_X(t) \quad \text{for } z \notin \text{supp } \mu_X,$$

is the T -transform of μ_X .

Theorem 2.6 (Norm of eigenvector projection). *Consider indices $i_0 \in \{1, \dots, r\}$ such that $1/\theta_{i_0} \in (T_{\mu_X}(a^-), T_{\mu_X}(b^+))$. For each n , define*

$$z_n := \begin{cases} \lambda_{i_0}(\tilde{X}_n) & \text{if } \theta_{i_0} > 0, \\ \lambda_{n-r+i_0}(\tilde{X}_n) & \text{if } \theta_{i_0} < 0, \end{cases}$$

and let x_n be a unit eigenvector of \tilde{X}_n associated with the eigenvalue z_n . Then we have

a)

$$\langle x_n, \ker(\theta_{i_0} I_n - P_n) \rangle^2 \xrightarrow{a.s.} \frac{-1}{\theta_{i_0}^2 \rho T'_{\mu_X}(\rho) + \theta_{i_0}},$$

where $\rho = T_{\mu_X}^{-1}(1/\theta_{i_0})$ is the limit of z_n ;

b)

$$\langle x_n, \oplus_{j \neq i_0} \ker(\theta_j I_n - P_n) \rangle \xrightarrow{a.s.} 0,$$

as $n \rightarrow \infty$.

Theorem 2.7 (Eigenvector phase transition). *When $r = 1$, let the sole non-zero eigenvalue of P_n be denoted by θ . Suppose that*

$$\frac{1}{\theta} \notin (T_{\mu_X}(a^-), T_{\mu_X}(b^+)), \quad \text{and} \quad \begin{cases} T'_{\mu_X}(b^+) = -\infty & \text{if } \theta > 0, \\ T'_{\mu_X}(a^-) = -\infty & \text{if } \theta < 0. \end{cases}$$

For each n , let x_n be an eigenvector of \tilde{X}_n associated with either the largest or smallest eigenvalue depending on whether $\theta > 0$ or $\theta < 0$, respectively. Then, we have

$$\langle x_n, \text{ran}(P_n) \rangle \xrightarrow{\text{a.s.}} 0$$

as $n \rightarrow \infty$.

The analogue of Remark 2.4 also applies here.

Remark 2.8 (Eigenvalues and eigenvectors of a similarity transformation of X_n). Consider the matrix $S_n = (I_n + P_n)^{1/2} X_n (I_n + P_n)^{1/2}$. The matrix S_n and $\tilde{X}_n = X_n (I_n + P_n)$ share the same eigenvalues and consequently the same limiting eigenvalue behavior in Theorem 2.5. Additionally, if x_n is a unit norm eigenvector of \tilde{X}_n then $z_n = (I_n + P_n)^{1/2} x_n$ is an eigenvector of S_n and $y_n = z_n / \|z_n\|$ satisfies

$$\langle y_n, \ker(\theta_{i_0} I_n - P_n) \rangle^2 = \frac{(\theta_{i_0} + 1) \langle x_n, \ker(\theta_{i_0} I_n - P_n) \rangle^2}{\theta_{i_0} \langle x_n, \ker(\theta_{i_0} I_n - P_n) \rangle^2 + 1}.$$

It follows that we have the same phase transition and that when $1/\theta_{i_0} \in (T_{\mu_X}(a^-), T_{\mu_X}(b^+))$,

$$\langle y_n, \ker(\theta_{i_0} I_n - P_n) \rangle^2 \xrightarrow{\text{a.s.}} -\frac{\theta_{i_0} + 1}{\theta_{i_0} T'_{\mu_X}(\rho)} \quad \text{and} \quad \langle y_n, \oplus_{j \neq i_0} \ker(\theta_j I_n - P_n) \rangle \xrightarrow{\text{a.s.}} 0,$$

so that we have proved the analogue of Theorems 2.6 and 2.7 for the eigenvectors of S_n .

2.4. The Cauchy and T transforms in free probability theory.

2.4.1. The Cauchy transform and its relation to additive free convolution. The Cauchy transform of a compactly supported probability measure μ on the real line is defined as:

$$G_\mu(z) = \int \frac{d\mu(t)}{z - t} \quad \text{for } z \notin \text{supp } \mu_X.$$

If $[a, b]$ denotes the convex hull of the support of μ , then

$$G_\mu(a^-) := \lim_{z \uparrow a} G_\mu(z) \quad \text{and} \quad G_\mu(b^+) := \lim_{z \downarrow b} G_\mu(z)$$

exist in respectively $[-\infty, 0)$ and $(0, +\infty]$ and G_μ realizes decreasing homeomorphisms from $(-\infty, a)$ onto $(G_\mu(a^-), 0)$ and from $(b, +\infty)$ onto $(0, G_\mu(b^+))$.

In this paper, we shall denote by G_μ^{-1} the inverses of these homeomorphisms, even though G_μ sometimes defines other homeomorphisms on the holes of the support of μ .

The R -transform $R_\mu(z) := G_\mu^{-1}(z) - 1/z$, is the analogue of the logarithm of the Fourier transform for free additive convolution. The free additive convolution of probability measures on the real line is denoted by the symbol \boxplus and can be characterized as follows. Let A_n and B_n be independent $n \times n$ symmetric (or Hermitian) random matrices that are invariant, in law, by conjugation by any orthogonal (or unitary) matrix. Suppose that, as n tends to infinity, $\mu_{A_n} \rightarrow \mu_A$ and $\mu_{B_n} \rightarrow \mu_B$. Then, free probability theory states that $\mu_{A_n + B_n} \rightarrow \mu_A \boxplus \mu_B$, a probability measure which can be characterized in terms of the R -transform as

$$R_{\mu_A \boxplus \mu_B}(z) = R_{\mu_A}(z) + R_{\mu_B}(z).$$

The connection between free additive convolution and G_μ^{-1} (via the R -transform) is what makes its manifestation in Theorem 2.1 natural.

2.4.2. *The T -transform and its relation to multiplicative free convolution.* In the case where $\mu \neq \delta_0$ and the support of μ is contained in $[0, +\infty)$, one also defines its T -transform

$$T_\mu(z) = \int \frac{t}{z-t} d\mu(t) \quad \text{for } z \notin \text{supp } \mu_X,$$

which realizes decreasing homeomorphisms from $(-\infty, a)$ onto $(T_\mu(a^-), 0)$ and from $(b, +\infty)$ onto $(0, T_\mu(b^+))$.

In this paper, we shall denote by T_μ^{-1} the inverses of these homeomorphisms, even though T_μ shall sometimes define other homeomorphisms on the holes of the support of μ .

The S -transform $S_\mu(z) := (1+z)/(zT_\mu^{-1}(z))$ is the analogue of the Fourier transform for free multiplicative convolution \boxtimes , the binary operation on the set of probability measures on $[0, +\infty)$ defined by the fact that, with the above hypotheses on A_n, B_n , we have $\mu_{A_n B_n} \longrightarrow \mu_A \boxtimes \mu_B$, and characterized by the fact that $S_{\mu_A \boxtimes \mu_B} = S_{\mu_A} S_{\mu_B}$. The connection between free multiplicative convolution and T_μ^{-1} (via the S -transform) is what makes its manifestation in Theorem 2.5 natural.

2.5. Extensions.

Remark 2.9 (Phase transition in non-extreme eigenvalues). Theorem 2.1 can easily be adapted to describe the phase transition in the eigenvalues of $X_n + P_n$ which fall in the “holes” of the support of μ_X . Consider $c < d$ such that almost surely, for n large enough, X_n has no eigenvalue in (c, d) (which implies that G_{μ_X} induces a decreasing homeomorphism, that we shall denote by $G_{\mu_X, (c, d)}$, from (c, d) onto $(G_{\mu_X}(d^-), G_{\mu_X}(c^+))$). Then almost surely, for n large enough, $X_n + P_n$ has no eigenvalue in (c, d) , except if some of the $1/\theta_i$'s are in $(G_{\mu_X}(d^-), G_{\mu_X}(c^+))$, in which case for each such index i , one eigenvalue of \tilde{X}_n has limit $G_{\mu_X, (c, d)}^{-1}(1/\theta_i)$ as n tends to infinity.

Remark 2.10 (Isolated eigenvalues of X_n outside the support of μ_X). When X_n itself has isolated eigenvalues in the sense that the limit a of its smallest eigenvalue and/or the limit b of its largest eigenvalue are out of the support of μ_X , the phase transition occurs at the same values as in Theorem 2.1. Indeed, everything, in the proof, carries through, except (6). It follows that for each $i \in \{1, \dots, s\}$,

$$\begin{cases} \lim_{n \rightarrow \infty} \lambda_i(X_n + P_n) = G_{\mu_X}^{-1}(1/\theta_i) & \text{if } 1/\theta_i < G_{\mu_X}(b^+), \\ \limsup_{n \rightarrow \infty} \lambda_i(X_n + P_n) \leq b & \text{otherwise,} \end{cases}$$

and that for each $i \in \{s+1, \dots, r\}$,

$$\begin{cases} \lim_{n \rightarrow \infty} \lambda_{n-r+i}(X_n + P_n) = G_{\mu_X}^{-1}(1/\theta_i) & \text{if } 1/\theta_i > G_{\mu_X}(a^-), \\ \liminf_{n \rightarrow \infty} \lambda_{n-r+i}(X_n + P_n) \geq a & \text{otherwise.} \end{cases}$$

Remark 2.11 (Other perturbations of X_n). The previous remark forms the basis for an iterative application of our theorems to other perturbational models, such as $\tilde{X} = \sqrt{X}(I + P)\sqrt{X} + Q$ for example. Another way to deal with such perturbations is to first derive the corresponding *master equations* representations that describe how the eigenvalues and eigenvectors of \tilde{X} are related to the eigenvalues and eigenvectors of X and the perturbing matrices, along the lines of Proposition 5.1 for additive or multiplicative perturbations of Hermitian matrices.

Remark 2.12 (Random matrices with Haar-like eigenvectors). Let G be an $n \times m$ Gaussian random matrix with independent real (or complex) entries that are normally distributed with mean 0 and variance 1. The eigenvectors U of the matrix $X = GG^*/m$, will be Haar distributed. When G is a Gaussian-like matrix in the sense that its entries are i.i.d. with mean zero and variance one, then upon placing adequate restrictions on the higher order moments, we label the eigenvectors of X as being Haar-like. Following the development in [34, 35, 36], when U is Haar-like, then for non-random unit norm vector x_n , the vector U^*x_n will be close to uniformly distributed on the unit real (or complex) sphere. Since our proofs rely heavily on the properties of unit norm vectors uniformly distributed on the n -sphere, they can be easily adapted to the setting where the unit norm vectors are close to uniformly distributed. Hence, we assert without proof, the applicability of our results to the setting where X_n or P_n or both have independent Haar-like distributed eigenvectors.

Remark 2.13 (Setting where eigenvalues of P_n are not fixed). Suppose that P_n is a random matrix independent of X_n , with exactly r non-zero eigenvalues given by $\theta_1^n, \dots, \theta_r^n$. Let $\theta_i^n \xrightarrow{\text{a.s.}} \theta_i$ as $n \rightarrow \infty$. Using [23, Cor. 6.3.8] as in Section 6.2, one can easily see that our results will also apply in this case.

The analogues of Remarks 2.9, 2.10, 2.12 and 2.13 for the multiplicative setting also hold here. In particular, Wishart matrices with $c > 1$ (cf Section 3.2) gives an illustration of the case where there is a hole in the support of μ_X .

3. EXAMPLES

We now illustrate our results with some concrete computations. The key to applying our results lies in being able to compute the Cauchy or T transforms of the probability measure μ_X and their associated functional inverses. In what follows, we focus on settings where the transforms and their inverses can be expressed in closed form. In settings where the transforms are algebraic so that they can be represented as solutions of polynomial equations, the techniques and software developed in [33] can be utilized. In more complicated settings, one will have to resort to numerical techniques.

3.1. Additive perturbation of a random Gaussian matrix. Let X_n be an $n \times n$ symmetric (or Hermitian) matrix with independent, zero mean, normally distributed entries with variance σ^2/n on the diagonal and $\sigma^2/(2n)$ on the off diagonal. It is known that the spectral measure of X_n converges to the famous semi-circle distribution with density

$$d\mu_X(x) = \frac{\sqrt{4\sigma^2 - x^2}}{2\sigma^2\pi} dx \quad \text{for } x \in [-2\sigma, 2\sigma].$$

It is known that the extreme eigenvalues converge to the bounds of this support [1]. Associated with the spectral measure, we have

$$G_{\mu_X}(z) = \frac{z - \text{sgn}(z)\sqrt{z^2 - 4\sigma^2}}{2\sigma^2}, \quad \text{for } z \in (-\infty, -2\sigma) \cup (2\sigma, +\infty),$$

$$G_{\mu_X}(\pm 2\sigma) = \pm\sigma \text{ and } G_{\mu_X}^{-1}(1/\theta) = \theta + \frac{\sigma^2}{\theta}.$$

Thus for a P_n with r non-zero eigenvalues $\theta_1 \geq \dots \geq \theta_s > 0 > \theta_{s+1} \geq \dots \geq \theta_r$, for any fixed $i \geq 1$, by Theorem 2.1, we have

$$\lambda_i(X_n + P_n) \xrightarrow{\text{a.s.}} \begin{cases} \theta_i + \frac{\sigma^2}{\theta_i} & \text{if } 1 \leq i \leq s \text{ and } \theta_i > \sigma \\ 2\sigma & \text{otherwise,} \end{cases}$$

as $n \rightarrow \infty$. This result has already been established in [19] for the symmetric case and in [32] for the Hermitian case. Remark 2.12 explains why our results hold for Wigner matrices of the sort considered in [32, 19]. Now, onto the eigenvectors.

In the setting where $r = 1$ and $\theta := \theta_1 > 0$, let x_n be an eigenvector of $X_n + P_n$ associated with its largest eigenvalue. By Theorems 2.2 and 2.3, we have

$$\langle x_n, \ker(\theta I_n - P_n) \rangle^2 \xrightarrow{\text{a.s.}} \begin{cases} 1 - \frac{\sigma^2}{\theta^2} & \text{if } \theta \geq \sigma, \\ 0 & \text{if } \theta < \sigma. \end{cases}$$

3.2. Multiplicative perturbation of a random Wishart matrix. Let G_n be an $n \times m$ real (or complex) matrix with independent, zero mean, normally distributed entries with variance 1. Let $X_n = G_n G_n^* / m$. It is known [27] that, as $n, m \rightarrow \infty$ with $n/m \rightarrow c > 0$, the spectral measure of X_n converges to the famous Marčenko-Pastur distribution with density

$$d\mu_X(x) := \frac{1}{2\pi c x} \sqrt{(b-x)(x-a)} \mathbb{1}_{[a,b]}(x) dx + \max\left(0, 1 - \frac{1}{c}\right) \delta_0,$$

where $a = (1 - \sqrt{c})^2$ and $b = (1 + \sqrt{c})^2$. It is known [3] that the extreme eigenvalues converge to the bounds of this support.

Associated with this spectral measure, we have

$$\begin{aligned} T_{\mu_X}^{-1} &= \frac{(z+1)(cz+1)}{z}, \\ T_{\mu_X}(z) &= \frac{z - c - 1 - \operatorname{sgn}(z-a)\sqrt{(z-a)(z-b)}}{2c}, \\ T_{\mu_X}(b^+) &= 1/\sqrt{c}, \quad T_{\mu_X}(a^-) = -1/\sqrt{c}. \end{aligned}$$

When $c > 1$, there is an atom at zero so that the smallest eigenvalue of X_n equals zero. For simplicity, let us consider the setting when $c < 1$ so that the extreme eigenvalues of X_n tend almost surely to a and b . Thus for P_n with r non-zero eigenvalues $\theta_1 \geq \dots \geq \theta_s > 0 > \theta_{s+1} \geq \dots \geq \theta_r$, with $l_i := \theta_i + 1$, for any fixed $i \geq 1$ and $c < 1$, by Theorem 2.5, we have

$$\lambda_i(X_n(I_n + P_n)) \xrightarrow{\text{a.s.}} \begin{cases} l_i \left(1 + \frac{c}{l_i - 1}\right) & \text{if } 1 \leq i \leq s \text{ and } |l_i - 1| > \sqrt{c} \\ b & \text{otherwise,} \end{cases}$$

as $n \rightarrow \infty$. An analogous result for the smallest eigenvalue may be similarly derived by making the appropriate substitution for a in Theorem 2.5. Consider the matrix $S_n = (I_n + P_n)^{1/2} X_n (I_n + P_n)^{1/2}$. The matrix S_n may be interpreted as a Wishart distributed sample covariance matrix with “spiked” covariance $I_n + P_n$. By Remark 2.8, the above result applies for the eigenvalues of S_n as well. This result for the largest eigenvalue of spiked sample covariance matrices was established in [4, 31] and for the extreme eigenvalues in [5]. Now, onto the eigenvectors.

In the setting where $r = 1$, let us denote $l = \theta_1 + 1$, and let be x_n a unit norm eigenvector of $X_n(I + P_n)$ associated with its largest (or smallest, according to whether $l > 1$ or $l < 1$) eigenvalue. By Theorem 2.7, we have

$$\langle x_n, \ker(lI_n - \Sigma) \rangle^2 \xrightarrow{\text{a.s.}} \begin{cases} \frac{(l-1)^{2-c}}{(l-1)^{[c(l+1)+l-1]}} & \text{if } |l-1| \geq \sqrt{c}, \\ 0 & \text{if } |l-1| < \sqrt{c}. \end{cases}$$

Let y_n be a unit eigenvector of $S_n = (I_n + P_n)^{1/2} X_n (I_n + P_n)^{1/2}$ associated with its largest (or smallest, according to whether $l > 1$ or $l < 1$) eigenvalue. Then, by Theorem 2.7 and Remark 2.8, we have

$$\langle y_n, \ker(lI_n - \Sigma) \rangle^2 \xrightarrow{\text{a.s.}} \begin{cases} \frac{1 - \frac{c}{(l-1)^2}}{1 + \frac{c}{l-1}} & \text{if } |l-1| \geq \sqrt{c}, \\ 0 & \text{if } |l-1| < \sqrt{c}. \end{cases}$$

The result has been established in [31] for the largest eigenvector under particular hypotheses. We generalize it to the case where $\Sigma \leq I_n$.

4. SKETCH OF THE PROOFS

We now provide a sketch of the proofs. We focus on Theorems 2.1, 2.2 and 2.3 which describe the the phase transition in the extreme eigenvalues and associated eigenvectors of $X + P$ (the index n in X_n and P_n has been suppressed for brevity). An analogous argument applies for the multiplicative perturbation setting.

Consider the setting where $r = 1$, so that $P = \theta uu^*$, with u being a unit norm column vector. Since either X or P is assumed to be invariant, in law, under unitary (or orthogonal) conjugation, one can, without loss of generality, suppose that $X = \text{diag}(\lambda_1, \dots, \lambda_n)$ and that u is uniformly distributed on the unit n -sphere.

4.1. Largest eigenvalue phase transition. For any $z \in \mathbb{C}$ such that z is not an eigenvalue of X , we have

$$z - (X + P) = (z - X) \times (I - (z - X)^{-1}P),$$

so that z is an eigenvalue of $X + P$ if and only if 1 is an eigenvalue of $(z - X)^{-1}P$. But $(z - X)^{-1}P = (z - X)^{-1}\theta uu^*$ has rank one, so its only non-zero eigenvalue will equal its trace, which in turn is equal to $\theta G_{\mu_n}(z)$, where μ_n is a “weighted” spectral measure of X , defined by

$$\mu_n = \sum_{k=1}^n |u_k|^2 \delta_{\lambda_k} \quad (\text{the } u_k \text{'s are the coordinates of } u).$$

Thus any z out of the spectrum of X is an eigenvalue of $X + P$ if and only if

$$\sum_{k=1}^n \frac{|u_k|^2}{z - \lambda_k} =: G_{\mu_n}(z) = \frac{1}{\theta}. \quad (1)$$

Equation (1) describes the relationship between the eigenvalues of $X + P$ and the eigenvalues of X and the dependence on the coordinates of the vector u (via the measure μ_n).

This is where randomization simplifies analysis. Since u is a random vector with uniform distribution on the unit n -sphere, we have that for large n , $|u_k|^2 \approx \frac{1}{n}$ with high probability. Consequently, we have $\mu_n \approx \mu_X$ so that $G_{\mu_n}(z) \approx G_{\mu_X}(z)$. Inverting equation (1) after substituting these approximations yields the location of the largest eigenvalue to be $G_{\mu_X}^{-1}(1/\theta)$ as in Theorem 2.1.

The phase transition for the extreme eigenvalues emerges because under our assumption that the limiting probability measure μ_X is compactly supported on $[a, b]$, the Cauchy transform G_{μ_X} is defined *outside* $[a, b]$ and unlike what happens for G_{μ_n} , we do not always have $G_{\mu_X}(b^+) = +\infty$. Consequently, when $1/\theta < G_{\mu_X}(b^+)$, we have that $\lambda_1(\tilde{X}) \approx G_{\mu_X}^{-1}(1/\theta)$ as before. However, when $1/\theta \geq G_{\mu_X}(b^+)$ then the phase transition manifests and $\lambda_1(\tilde{X}) \approx \lambda_1(X) = b$.

An extension of these arguments for fixed $r > 1$ yields the general result and constitutes the most transparent justification, as sought by the authors in [4], for the emergence of this phase transition phenomenon in such perturbed random matrix models. We rely on concentration inequalities to make the arguments rigorous.

4.2. Eigenvectors phase transition. Let x be a unit eigenvector of $X + P$ associated with the eigenvalue z that satisfies (1). From the relationship $(X + P)x = zx$, we deduce that, for $P = \theta uu^*$,

$$(z - X)x = Px = \theta uu^*x = (\theta u^*x)u \quad (\text{because } u^*x \text{ is a scalar}),$$

implying that x is proportional to $(z - X)^{-1}u$. Hence, since x has norm one,

$$x = \frac{(z - X)^{-1}u}{\sqrt{u^*(z - X)^{-2}u}} \quad (2)$$

and

$$\langle x, \ker(\theta I - P) \rangle^2 = |u^*x|^2 = \frac{(u^*(z - X)^{-1}u)^2}{u^*(z - X)^{-2}u} = \frac{G_{\mu_n}(z)^2}{\int \frac{d\mu_n(t)}{(z-t)^2}} = \frac{1}{\theta^2 \int \frac{d\mu_n(t)}{(z-t)^2}}. \quad (3)$$

Equation (2) describes the relationship between the eigenvectors of $X + P$ and the eigenvalues of X and the dependence on the coordinates of the vector u (via the measure μ_n).

Here too, randomization simplifies analysis since for large n , we have $\mu_n \approx \mu_X$. Consequently,

$$\int \frac{d\mu_n(t)}{(z-t)^2} \approx \int \frac{d\mu_X(t)}{(\rho-t)^2} = -G'_{\mu_X}(\rho),$$

so that when $1/\theta < G_{\mu_X}(b^+)$, which implies that $\rho > b$, we have

$$\langle x, \ker(\theta I - P) \rangle^2 \xrightarrow{\text{a.s.}} \frac{1}{\theta^2 \int \frac{d\mu_X(t)}{(\rho-t)^2}} = \frac{-1}{\theta^2 G'_{\mu_X}(\rho)} > 0,$$

whereas when $1/\theta \geq G_{\mu_X}(b^+)$ and G_{μ_X} has infinite derivative at b , we have

$$\langle x, \ker(\theta I - P) \rangle^2 \xrightarrow{\text{a.s.}} 0.$$

An extension of these arguments for fixed $r > 1$ yields the general result and brings into sharp focus the connection between the eigenvalue phase transition and the associated

eigenvector phase transition. As before, concentration inequalities allow us to make these arguments rigorous.

5. THE EXACT MASTER EQUATIONS FOR THE EIGENVALUES AND EIGENVECTORS OF $X + P$ OR $X(I + P)$

In this section, we provide the r -dimensional analogues of the master equations (1) and (2) employed in our high-level sketch of the proof.

Proposition 5.1. *Let us fix some positive integers $1 \leq r \leq n$. Let $X = \text{diag}(\lambda_1, \dots, \lambda_n)$ be a diagonal $n \times n$ matrix and $P = V\Theta V^*$, with $\Theta = \text{diag}(\theta_1, \dots, \theta_r)$ an $r \times r$ diagonal matrix and V an $n \times r$ matrix which columns are orthonormal (i.e. such that $V^*V = I_r$).*

a) Then any $z \notin \{\lambda_1, \dots, \lambda_n\}$ is an eigenvalue of $\tilde{X} := X + P$ if and only if the $r \times r$ matrix

$$I_r - V^*(z - X)^{-1}V\Theta \quad (4)$$

*is not invertible. In this case, for all $x \in \ker(z - \tilde{X})$, V^*x belongs to the kernel of the above matrix and*

$$x = (z - X)^{-1}V\Theta V^*x. \quad (5)$$

b) Let us denote $V = [v_{k,l}]_{k,l=1}^{n,r}$. Then for all $i, j = 1, \dots, r$, the (i, j) -th entry of the matrix of (4) is

$$\mathbb{1}_{i=j} - \theta_j G_{\mu_{i,j}}(z),$$

where $\mu_{i,j}$ is the complex measure defined by

$$\mu_{i,j} = \sum_{k=1}^n \bar{v}_{k,i} v_{k,j} \delta_{\lambda_k}$$

and $G_{\mu_{i,j}}$ is the Cauchy transform of $\mu_{i,j}$.

c) If one replaces $X + P$ by $X(I + P)$, everything stays true as long as one replaces $(z - X)^{-1}$ by $(z - X)^{-1}X$ in (4) and in (5) and the Cauchy transform by the T -transform in b).

Proof. Part b) follows from a straightforward computation of the (i, j) -th entry of $V^*(z - X)^{-1}V\Theta$. Part a) is proved, for example, in [2, Th. 2.3]. Part c) can be proved in the same way. \square

6. PROOF OF THEOREM 2.1

6.1. First step: Setting where the θ_i 's are pairwise distinct. Note that the empirical spectral measure of \tilde{X}_n tends almost surely to $\mu_X \boxplus \delta_0 = \mu_X$, which support contains a and b , thus for all $i \geq 1$ fixed,

$$\liminf_{n \rightarrow \infty} \lambda_i(\tilde{X}_n) \geq b \quad \text{and} \quad \limsup_{n \rightarrow \infty} \lambda_{n+1-i}(\tilde{X}_n) \leq a. \quad (6)$$

Let us denote the eigenvalues of X_n by $\lambda_1^n \geq \dots \geq \lambda_n^n$ and introduce orthogonal (or unitary) $n \times n$ matrices U_X^n, U_P^n such that

$$X_n = U_X^n \text{diag}(\lambda_1^n, \dots, \lambda_n^n) U_X^{n*}, \quad P_n = U_P^n \text{diag}(\theta_1, \dots, \theta_r, 0, \dots, 0) U_P^{n*}.$$

The spectrum of $X_n + P_n$ is the one of

$$\text{diag}(\lambda_1^n, \dots, \lambda_n^n) + \underbrace{U_X^{n*} U_P^n}_{\text{denoted by } U_n} \text{diag}(\theta_1, \dots, \theta_r, 0, \dots, 0) U_P^{n*} U_X^n,$$

and since either X_n or P_n is invariant in law by conjugation by orthogonal (or unitary) matrices, U_n is Haar-distributed and independent of $\lambda^n := (\lambda_1^n, \dots, \lambda_n^n)$ (see the first paragraph of the proof of [22, Th. 4.3.5] for more details). Let us denote the entries of U_n by $[u_{i,j}^n]_{i,j=1}^n$.

Since $\lambda_1(X_n) \xrightarrow{\text{a.s.}} b$ and $\lambda_n(X_n) \xrightarrow{\text{a.s.}} a$, we can focus on the eigenvalues of $X_n + P_n$ which are out of $[\lambda_n(X_n), \lambda_1(X_n)]$. By Proposition 5.1, these eigenvalues are the numbers z out of $[\lambda_n(X_n), \lambda_1(X_n)]$ such that the $r \times r$ matrix

$$M(n, z) := I_r - [\theta_j G_{\mu_{i,j}^n}(z)]_{i,j=1}^r$$

is not invertible, where for all $i, j = 1, \dots, r$, $\mu_{i,j}^n$ is the random complex measure defined by

$$\mu_{i,j}^n = \sum_{k=1}^n \overline{u_{k,i}^n} u_{k,j}^n \delta_{\lambda_k^n}. \quad (7)$$

By Proposition 9.3, for all $i \neq j$, the random complex measure $\mu_{i,j}^n$ converges almost surely weakly to zero and for all i , $\mu_{i,i}^n$ converges almost surely weakly to μ_X . Thus, by the second statement of Lemma 9.1, for all $\eta > 0$, the matrix-valued function $M(n, \cdot)$ converges to

$$M_{G_{\mu_X}}(\cdot) := \text{diag}(1 - \theta_1 G_{\mu_X}(\cdot), \dots, 1 - \theta_r G_{\mu_X}(\cdot))$$

uniformly on $\{z \in \mathbb{C}; d(z, [a, b]) \geq \eta\}$, almost surely. Now, the conclusion follows exactly from Lemma 9.4, which hypotheses are satisfied (hypotheses a), b), c) follow from the definition of G_{μ_X} , hypothesis d) follows from Proposition 5.1 and from the fact that $X_n + P_n$ is Hermitian and hypothesis e) has been checked above).

6.2. Second step: Extension to the general case. We now treat the case where the θ_i 's are not pairwise distinct. We concentrate on the largest eigenvalues. By (6) and Weyl's interlacing inequalities, the case where $i_0 > s$ is clear. So let us suppose that $i_0 \leq s$. We denote, for $\theta > 0$,

$$\rho_\theta = \begin{cases} G_{\mu_X}^{-1}(1/\theta) & \text{if } 1/\theta < G_{\mu_X}(b^+), \\ b & \text{in the other case.} \end{cases}$$

There is $\eta > 0$ such that $|\rho_\theta - \rho_{\theta_{i_0}}| \leq \varepsilon$ whenever $|\theta - \theta_{i_0}| \leq \eta$. Consider pairwise distinct real numbers $\theta'_1 > \dots > \theta'_r$ such that for all $i = 1, \dots, r$, $\theta_i \theta'_i > 0$ and

$$\sum_{i=1}^r (\theta'_i - \theta_i)^2 \leq \min(\eta^2, \varepsilon^2).$$

It implies that $|\rho_{\theta'_{i_0}} - \rho_{\theta_{i_0}}| \leq \varepsilon$. Employing the notation in Section 6.1, for each k and each n , we define

$$P'_n = U_P^n \text{diag}(\theta'_1, \dots, \theta'_r, 0, \dots, 0) U_P^{n*}.$$

Note that by [23, Cor. 6.3.8], we have, for all n ,

$$\sum_{j=1}^n (\lambda_j(X_n + P'_n) - \lambda_j(X_n + P_n))^2 \leq \text{Tr}(P'_n - P_n)^2 = \sum_{i=1}^r (\theta'_i - \theta_i)^2 \leq \varepsilon^2.$$

Since the theorem can be applied to $X_n + P'_n$, it follows that almost surely, for n large enough,

$$|\lambda_{i_0}(X_n + P'_n) - \rho_{\theta'_{i_0}}| \leq \varepsilon.$$

By the triangular inequality, almost surely, for n large enough,

$$|\lambda_{i_0}(X_n + P'_n) - \rho_{\theta_{i_0}}| \leq 3\varepsilon.$$

□

7. PROOFS OF THEOREMS 2.2 AND 2.3 FOR THE EXTREME EIGENVECTORS OF $X_n + P_n$

7.1. Proof of Theorem 2.2. Let us write, for each n ,

$$X_n = U_X^n \text{diag}(\lambda_1^n, \dots, \lambda_n^n) U_X^{n*}, \quad P_n = U_P^n \text{diag}(\theta_1, \dots, \theta_r, 0, \dots, 0) U_P^{n*}.$$

The eigenvectors of $X_n + P_n$ are U_X^n times the ones of

$$\text{diag}(\lambda_1^n, \dots, \lambda_n^n) + U_X^{n*} U_P^n \text{diag}(\theta_1, \dots, \theta_r, 0, \dots, 0) U_P^{n*} U_X^n,$$

hence it suffices to prove the result in the case where $X_n = \text{diag}(\lambda_1^n, \dots, \lambda_n^n)$ and $P_n = U_n \text{diag}(\theta_1, \dots, \theta_r, 0, \dots, 0) U_n^*$, with U_n Haar-distributed. We denote the entries of U_n by $[u_{i,j}^n]_{i,j=1}^n$, its columns by C_1^n, \dots, C_n^n and the $n \times r$ matrix whose columns are respectively C_1^n, \dots, C_r^n by V_n .

Let r_0 be the number of i 's such that $\theta_i = \theta_{i_0}$. Up to a reindex of the θ_i 's (which is then no longer decreasing, but it is not a problem here), one can suppose that $i_0 = 1$, $\theta_1 = \dots = \theta_{r_0}$. For each n , $\ker(\theta_1 I_n - P_n)$ is then the linear span of the r_0 first columns of U_n , namely $C_1^n, \dots, C_{r_0}^n$. Since these columns are orthonormal, it suffices to prove that as n tends to infinity,

$$\sum_{j=1}^{r_0} |\langle C_j^n, x_n \rangle|^2 \quad \text{tends almost surely to} \quad \frac{-1}{\theta_1^2 G'_{\mu_X}(\rho)} = \frac{1}{\theta_1^2 \int \frac{d\mu_X(t)}{(\rho-t)^2}} \quad (8)$$

and

$$\sum_{j=r_0+1}^r |\langle C_j^n, x_n \rangle|^2 \quad \text{tends almost surely to} \quad 0. \quad (9)$$

Let us introduce, for each n , for all z out of the spectrum of X_n , the $r \times r$ random matrix

$$M(n, z) := I_r - [\theta_j G_{\mu_{i,j}^n}(z)]_{i,j=1}^r,$$

where, for all $i, j = 1, \dots, r$, $\mu_{i,j}^n$ is the random complex measure defined by (7). As in the proof of Theorem 2.1 (end of the first step), for all $\eta > 0$, the random matrix-valued function $M(n, \cdot)$ converges to

$$M_{G_{\mu_X}}(\cdot) := \text{diag}(1 - \theta_1 G_{\mu_X}(\cdot), \dots, 1 - \theta_r G_{\mu_X}(\cdot))$$

uniformly on $\{z \in \mathbb{C}; d(z, [a, b]) \geq \eta\}$, almost surely. Since z_n tends almost surely to $G_{\mu_X}^{-1}(1/\theta_1)$, which is out of $[a, b]$, it follows that $M(n, z_n)$ tends almost surely to

$$\text{diag}(\underbrace{0, \dots, 0}_{r_0 \text{ zeros}}, 1 - \frac{\theta_{r_0+1}}{\theta_1}, \dots, 1 - \frac{\theta_r}{\theta_1}). \quad (10)$$

Moreover, by Proposition 5.1 a), for n large enough such that z_n is not an eigenvalue of X_n , the vector

$$V_n^* x_n = \begin{bmatrix} \langle C_1^n, x_n \rangle \\ \vdots \\ \langle C_r^n, x_n \rangle \end{bmatrix}$$

is a vector of the kernel of the $r \times r$ matrix $M(n, z_n)$ with norm ≤ 1 . It follows that any limit point of $V_n^* x_n$ belongs to the kernel of the matrix of (10). Hence (9) holds.

Let us now prove (8). By Proposition 5.1 a) again, one has, for all n ,

$$\begin{aligned} x_n &= (z_n - X_n)^{-1} V_n \text{diag}(\theta_1, \dots, \theta_r) V_n^* x_n \\ &= (z_n - X_n)^{-1} \sum_{j=1}^r \theta_j \langle C_j^n, x_n \rangle C_j^n, \\ &= \underbrace{(z_n - X_n)^{-1} \sum_{j=1}^{r_0} \theta_j \langle C_j^n, x_n \rangle C_j^n}_{\text{denoted by } x'_n} + \underbrace{(z_n - X_n)^{-1} \sum_{j=r_0+1}^r \theta_j \langle C_j^n, x_n \rangle C_j^n}_{\text{denoted by } x''_n}. \end{aligned}$$

We have $z_n \xrightarrow{\text{a.s.}} \rho \notin [a, b]$, hence the sequence $(z_n - X_n)^{-1}$ is bounded for the operator norm, thus by (9), $\|x''_n\| \xrightarrow{\text{a.s.}} 0$. As a consequence, since $\|x_n\| = 1$, $\|x'_n\| \xrightarrow{\text{a.s.}} 1$. Since $\theta_1 = \dots = \theta_{r_0}$,

$$\|x'_n\|^2 = \theta_1^2 \sum_{i,j=1}^{r_0} \overline{\langle C_i^n, x_n \rangle} \langle C_j^n, x_n \rangle \underbrace{C_i^{n*} (z_n - X_n)^{-2} C_j^n}_{= \int \frac{1}{(z_n - t)^2} d\mu_{i,j}^n(t)} \quad (11)$$

By Proposition 9.3, for all $i \neq j$, $\mu_{i,j}^n$ converges almost surely weakly to zero and for all i , $\mu_{i,i}^n$ converges almost surely weakly to μ_X . Thus, since z_n converges almost surely to $\rho \notin [a, b]$, for all $i, j = 1, \dots, r_0$,

$$\int \frac{d\mu_{i,j}^n(t)}{(z_n - t)^2} \xrightarrow{\text{a.s.}} \mathbb{1}_{i=j} \int \frac{d\mu(t)}{(\rho - t)^2}.$$

By (11) and the fact that $\|x'_n\| \xrightarrow{\text{a.s.}} 1$, (8) follows. \square

7.2. Proof of Theorem 2.3. Let us suppose that $\theta > 0$ (the other case can be treated in the same way). Note first that because $G'_{\mu_X}(b^+) = -\infty$, we have $\int \frac{d\mu_X(t)}{(b-t)^2} = +\infty$.

As in the proof of Theorem 2.2, it suffices to prove the result in the case where $X_n = \text{diag}(\lambda_1^n, \dots, \lambda_n^n)$ and $P_n = \theta u^n u^{n*}$, with u^n a column vector uniformly distributed on the unit real (or complex) n -sphere. We denote by u_1^n, \dots, u_n^n the coordinates of the vector

u^n and define, for each n , the random probability measure

$$\mu^n = \sum_{k=1}^n |u_k^n|^2 \delta_{\lambda_k^n}.$$

Note that by the $r = 1$ case of Proposition 5.1 b), the eigenvalues of $X_n + P_n$ which do not belong to $\{\lambda_1^n, \dots, \lambda_n^n\}$ are the solutions of $G_{\mu^n}(z) = \frac{1}{\theta}$. Since G_{μ^n} decreases from $+\infty$ to 0 as z runs through $(\lambda_1^n, +\infty)$, we have $\lambda_1(X_n + P_n) > \lambda_1^n$.

Reproducing the arguments leading to (3) in Section (4.2), we get

$$\langle x_n, \ker(\theta I_n - P_n) \rangle^2 = \frac{1}{\theta^2 \int \frac{d\mu^n(t)}{(z_n - t)^2}}.$$

Thus it suffices to prove that, as n tends to infinity, $\int \frac{d\mu^n(t)}{(z_n - t)^2} \xrightarrow{\text{a.s.}} +\infty$. By hypothesis, we have $\frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k^n} \xrightarrow{\text{a.s.}} \mu_X$, so that $\frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k^n + b - z_n} \xrightarrow{\text{a.s.}} \mu_X$ (indeed, by Theorem 2.1, $z_n \xrightarrow{\text{a.s.}} b$). Hence, by (13),

$$\tilde{\mu}^n := \sum_{k=1}^n |u_k^n|^2 \delta_{\lambda_k^n + b - z_n} \xrightarrow{\text{a.s.}} \mu_X.$$

It implies that almost surely,

$$\liminf_{n \rightarrow +\infty} \int \frac{d\mu^n(t)}{(z_n - t)^2} = \liminf_{n \rightarrow +\infty} \int \frac{d\tilde{\mu}^n(t)}{(b - t)^2} \geq \int \frac{d\mu_X(t)}{(b - t)^2} = +\infty,$$

which concludes the proof of the Theorem. \square

8. PROOFS OF THEOREMS 2.5–2.7

These proofs can easily be obtained adapting to the square multiplicative case the proofs of the square additive case.

9. APPENDIX: TECHNICAL PRELIMINARIES NEEDED FOR THE PROOFS

9.1. Convergence of weighted spectral measures.

9.1.1. *A few facts about the weak convergence of complex measures.* Recall that a sequence (μ_n) of complex measures on \mathbb{R} is said to *converge weakly* to a complex measure μ on \mathbb{R} if, for any continuous bounded function f on \mathbb{R} ,

$$\int f(t) d\mu_n(t) \xrightarrow{n \rightarrow \infty} \int f(t) d\mu(t). \quad (12)$$

The following lemma shall be useful. It is well known for probability measures, but since we did not find any reference on its “complex measures version”, we give a proof. Recall that a sequence (μ_n) of complex measures on \mathbb{R} is said to be *tight* if

$$\lim_{R \rightarrow +\infty} \sup_n |\mu_n|(\{t \in \mathbb{R}; |t| \geq R\}) = 0.$$

Lemma 9.1. *Let D be a dense subset of the set of continuous functions on \mathbb{R} tending to zero at infinity endowed with the topology of the uniform convergence. Consider a tight sequence (μ_n) of complex measures on \mathbb{R} such that $(|\mu_n|(\mathbb{R}))$ is bounded and (12) holds for any function f in D . Then (μ_n) converges weakly to μ . Moreover, the convergence of (12) is uniform on any set of uniformly bounded and uniformly Lipschitz functions.*

Proof. Firstly, note that using the boundedness of $(|\mu_n|(\mathbb{R}))$, one extends easily (12) to any continuous function tending to zero at infinity. It follows that for any continuous bounded function f and any continuous function g tending to zero at infinity, we have

$$\limsup_{n \rightarrow \infty} \left| \int f d(\mu - \mu_n) \right| \leq \sup_n \int |f(1 - g)| d(|\mu| + |\mu_n|),$$

which, with a good choice of g (relying on the tightness hypothesis), can be made as small as we want. Thus the first statement is proved. The last part of the Lemma is a classical application of Ascoli's Theorem. \square

9.1.2. Convergence of weighted spectral measures.

Lemma 9.2. *Let, for each n , $U^n = (u_1^n, \dots, u_n^n)$, $V^n = (v_1^n, \dots, v_n^n)$ be the two first rows of a uniform random orthogonal (resp. unitary) matrix. Let also $x^n = (x_1^n, \dots, x_n^n)$ be a random family of real numbers independent of (U^n, V^n) such that almost surely, $\sup_{n,k} |x_k^n| < \infty$.*

a) *Then as n tends to infinity,*

$$\overline{u}_1^n v_1^n x_1^n + \dots + \overline{u}_n^n v_n^n x_n^n \xrightarrow{a.s.} 0.$$

b) *Suppose that $\frac{1}{n}(x_1^n + \dots + x_n^n)$ tends almost surely to a deterministic limit l . Then*

$$|u_1^n|^2 x_1^n + \dots + |u_n^n|^2 x_n^n \xrightarrow{a.s.} l.$$

Proof. Firstly, by conditioning, one can suppose the x^n 's to be deterministic.

Let us first prove a). Let us define the random variable $Z_n = \overline{u}_1^n v_1^n x_1^n + \dots + \overline{u}_n^n v_n^n x_n^n$. It suffices to prove that $\mathbb{E}[Z_n^4] = O(n^{-2})$. We have

$$\mathbb{E}[Z_n^4] = \sum_{i,j,k,l=1}^n x_i^n x_j^n x_k^n x_l^n \mathbb{E}[u_i^n u_j^n u_k^n u_l^n v_i^n v_j^n v_k^n v_l^n].$$

Note that by definition of the Haar measure, $\mathbb{E}[u_i^n u_j^n u_k^n u_l^n v_i^n v_j^n v_k^n v_l^n] = 0$ whenever among i, j, k, l , one is different of all others. It follows that

$$\begin{aligned} \mathbb{E}[Z_n^4] &\leq 3 \sum_{i,j} x_i^{n^2} x_j^{n^2} \mathbb{E}[u_i^{n^2} u_j^{n^2} v_i^{n^2} v_j^{n^2}] \\ &\leq 3 \sum_{i,j} x_i^{n^2} x_j^{n^2} \underbrace{(\mathbb{E}[u_i^{n^8}] \mathbb{E}[u_j^{n^8}] \mathbb{E}[v_i^{n^8}] \mathbb{E}[v_j^{n^8}])^{\frac{1}{4}}}_{=\mathbb{E}[u_1^{n^8}]} \end{aligned}$$

Since, by [17], $\mathbb{E}[u_1^{n^8}] = O(n^{-4})$, the conclusion holds.

To prove b), the strategy is quite different. Only U^n is involved, thus since a uniform random vector on the unit sphere of \mathbb{C}^n is a uniform random vector on the unit sphere of \mathbb{R}^{2n} , it suffices to treat the real case, which follows from a well-known concentration

result [26, Th. 2.3 and Prop. 1.8]: there are positive constants C, c such that for n large enough, for any 1-Lipschitz function g_n on the unit sphere of \mathbb{R}^n , for all $\varepsilon > 0$,

$$\mathbb{P} \{ |g_n(U^n) - \mathbb{E}[g_n(U^n)]| \geq \varepsilon \} \leq Ce^{-c\varepsilon^2},$$

which implies that if $\mathbb{E}[g_n(U^n)]$ converges, as n goes to infinity, to a finite limit, then $g_n(U^n)$ converges almost surely to the same limit. \square

Proposition 9.3. *Let, for each n , $U^n = (u_1^n, \dots, u_n^n)$, $V^n = (v_1^n, \dots, v_n^n)$ be the two first columns of a uniform random orthogonal (resp. unitary) matrix. Let also $\lambda^n = (\lambda_1^n, \dots, \lambda_n^n)$ be a random family of real numbers independent of (U^n, V^n) such that almost surely, $\sup_{n,k} |\lambda_k^n| < \infty$. We suppose that there exists a deterministic probability measure μ on \mathbb{R} such that almost surely, as n tends to infinity, $\frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k^n}$ converges weakly to μ .*

Then as n tends to infinity,

$$\mu_{U^n} := \sum_{k=1}^n |u_k^n|^2 \delta_{\lambda_k^n} \text{ converges almost surely weakly to } \mu, \quad (13)$$

$$\mu_{U^n, V^n} := \sum_{k=1}^n \bar{u}_k^n v_k^n \delta_{\lambda_k^n} \text{ converges almost surely weakly to } 0. \quad (14)$$

Proof. We use Lemma 9.1. Note first that almost surely, since $\sup_{n,k} |\lambda_k^n| < \infty$, both sequences are tight. Moreover, we have

$$|\mu_{U^n, V^n}| = \sum_{k=1}^n |u_k^n v_k^n| \delta_{\lambda_k^n},$$

thus, by the Cauchy-Schwartz inequality, $|\mu_{U^n, V^n}|(\mathbb{R}) \leq 1$. The set of continuous functions on the real line tending to zero at infinity admits a countable dense subset, so it suffices to prove that for any fixed such function f , the convergences of (13) and (14) hold almost surely when applied to f . This follows easily from Lemma 9.2. \square

9.2. A technical lemma. We shall need the following result. Note that nothing, in its hypotheses, is random. We define, for $z \in \mathbb{C}$ and E a closed subset of \mathbb{R} , $d(z, E) = \min_{x \in E} |z - x|$.

Lemma 9.4. *Let us fix a positive integer r , a family $\theta_1, \dots, \theta_r$ of pairwise distinct nonzero real numbers, two real numbers $a < b$, a function G which is analytic on $\mathbb{C} \setminus [a, b]$ and such that*

- a) G does not take any real value out of $\mathbb{R} \setminus [a, b]$,
- b) for all $z \in (-\infty, a) \cup (b, +\infty)$, $G(z) \in \mathbb{R}$ and $G'(z) < 0$,
- c) $G(z)$ tends to zero as z tends to $\pm\infty$ within \mathbb{R} .

Let us define, for $z \in \mathbb{C} \setminus [a, b]$, the $r \times r$ matrix

$$M_G(z) = \text{diag}(1 - \theta_1 G(z), \dots, 1 - \theta_r G(z)), \quad (15)$$

and denote by $z_1 > \dots > z_p$ the z 's such that $M_G(z)$ is not invertible ($p \in \{0, \dots, r\}$ is the number of i 's such that $G(a^-) < 1/\theta_i < G(b^+)$).

Let us also consider two sequences a_n, b_n with respective limits a, b and, for each n , a function $M(n, \cdot)$, defined on $\mathbb{C} \setminus [a_n, b_n]$, with values in the set of $r \times r$ complex matrices and which coefficients are analytic functions. We suppose that

- d) for all n , for all $z \in \mathbb{C} \setminus \mathbb{R}$, the matrix $M(n, z)$ is invertible,

- e) for all $\eta > 0$, $M(n, \cdot)$ converges, as n tends to infinity, to the function $M_G(\cdot)$, uniformly on $\{z \in \mathbb{C}; d(z, [a, b]) \geq \eta\}$.

Then there exists p real sequences $z_{n,1} > \dots > z_{n,p}$ converging respectively to z_1, \dots, z_p such that for all $\varepsilon \in (0, \min_i d(z_i, [a, b]))$, for n large enough, the z 's in $\mathbb{R} \setminus [a - \varepsilon, b + \varepsilon]$ such that $M(n, z)$ is not invertible are exactly $z_{n,1}, \dots, z_{n,p}$. Moreover, for n large enough, for each i , $M(n, z_{n,i})$ has rank $r - 1$.

Proof. Note firstly that by c), there exists $R > 0$ such that for $z \in \mathbb{R}$ such that $|z| \geq R$, $|G(z)| \leq \min_i \frac{1}{2|\theta_i|}$. By d) and e), it follows that for n large enough, all the z 's such that $M(n, z)$ is not invertible are in $[-R, R]$. To conclude, it suffices to prove that for all $c, d \in \mathbb{R} \setminus ([a, b] \cup \{z_1, \dots, z_p\})$ such that $c < d < a$ or $b < c < d$, we have:

- (H) the number $C_{c,d}(n)$ of z 's in (c, d) such that $\det M(n, z) = 0$ tends to the cardinality $C_{c,d}$ of the i 's in $\{1, \dots, p\}$ such that $c < z_i < d$.

(The assumption about the ranks following then from the fact that the set of matrices with rank at least $r - 1$ is open in the set of $r \times r$ matrices).

To prove (H), by additivity, one can suppose that c and d are close enough to have $C_{c,d} = 0$ or 1. Let us define γ to be the circle with diameter $[c, d]$. By a) and since $c, d \notin \{z_1, \dots, z_p\}$, $\det M_G(\cdot)$ does not vanish on γ , thus

$$C_{c,d} = \frac{1}{2i\pi} \int_{\gamma} \frac{\partial_z \det M_G(z)}{\det M_G(z)} dz = \lim_{n \rightarrow \infty} \frac{1}{2i\pi} \int_{\gamma} \frac{\partial_z \det M(n, z)}{\det M(n, z)} dz,$$

the last equality following from e). It follows that for n large enough, $C_{c,d}(n) = C_{c,d}$ (indeed, since $C_{c,d} = 0$ or 1, no ambiguity due to the orders of the zeros has to be taken into account here). \square

REFERENCES

- [1] G. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge studies in advanced mathematics, 118 (2009).
- [2] P. Arbenz, W. Gander, and G. H. Golub. Restricted rank modification of the symmetric eigenvalue problem: theoretical considerations. *Linear Algebra Appl.*, 104:75–95, 1988.
- [3] Z. D. Bai, and J. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*, Second Edition, Springer, New York, 2009.
- [4] J. Baik, G. Ben Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.
- [5] J. Baik, and J. W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.*, 97(6): 1382–1408, 2006.
- [6] K. E. Bassler, P. J. Forrester, and N. E. Frankel. Eigenvalue separation in some random matrix models. *J. Math. Phys.*, 50(3):033302, 24, 2009.
- [7] F. Benaych-Georges. Rectangular random matrices, related convolution. *Probab. Theory Related Fields*, 144(3-4):471–515, 2009.
- [8] F. Benaych-Georges. On a surprising relation between the Marchenko-Pastur law, rectangular and square free convolutions. To appear in *Annales de l'IHP Prob. Stats*. Available online at <http://arxiv.org/abs/0808.3938>, 2009.
- [9] F. Benaych-Georges Rectangular R -transform at the limit of rectangular spherical integrals, Available online at <http://arxiv.org/abs/0909.0178>, 2009.
- [10] F. Benaych-Georges, A. Guionnet and M. Maïda. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. In preparation.

- [11] F. Benaych-Georges, A. Guionnet and M. Maïda. Large deviations of the extreme eigenvalues of finite rank deformations of random matrices. In preparation.
- [12] F. Benaych-Georges and R. R. Nadakuditi. The extreme singular values and singular vectors of finite, low rank perturbations of large random rectangular matrices. In preparation.
- [13] P. Biane. Processes with free increments. *Math. Z.*, 227(1):143–174, 1998.
- [14] J. R. Bunch, C. P. Nielsen, and D. C. Sorensen. Rank-one modification of the symmetric eigenproblem. *Numer. Math.*, 31(1):31–48, 1978/79.
- [15] M. Capitaine, C. Donati-Martin, and D. Féral. The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. *Ann. Probab.*, 37(1):1–47, 2009.
- [16] B. Collins. Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields*, 133(3):315–344, 2005.
- [17] B. Collins and P. Śniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Comm. Math. Phys.*, 264(3):773–795, 2006.
- [18] N. El Karoui. Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.*, 35(2):663–714, 2007.
- [19] D. Féral and S. Péché. The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.*, 272(1):185–228, 2007.
- [20] A. Guionnet and M. Maïda. Character expansion method for the first order asymptotics of a matrix integral *Probab. Theory Related Fields*, 132(4):539–578, 2005.
- [21] A. Guionnet, M. Maïda. A Fourier view on the R -transform and related asymptotics of spherical integrals *J. Funct. Anal.* 222, 2, (2005), 435–490.
- [22] F. Hiai and D. Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.
- [23] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1985.
- [24] D. C. Hoyle and M. Rattray. Statistical mechanics of learning multiple orthogonal signals: asymptotic theory and fluctuation effects. *Phys. Rev. E (3)*, 75(1):016101, 13, 2007.
- [25] I. C. F. Ipsen and B. Nadler. Refined perturbation bounds for eigenvalues of Hermitian and non-Hermitian matrices. *SIAM J. Matrix Anal. Appl.*, 31(1):40–53, 2009.
- [26] M. Ledoux. The concentration of measure phenomenon. Providence, RI, AMS, 2001.
- [27] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.
- [28] R. R. Nadakuditi and J. W. Silverstein. Fundamental limit of sample generalized eigenvalue based detection of signals in noise using relatively few signal-bearing and noise-only samples. Available online at <http://arxiv.org/abs/0902.4250v1>, 2009.
- [29] B. Nadler. Finite sample approximation results for principal component analysis: a matrix perturbation approach. *Ann. Statist.*, 36(6):2791–2817, 2008.
- [30] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.
- [31] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica*, 17(4):1617–1642, 2007.
- [32] S. Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields*, 134(1):127–173, 2006.
- [33] N. R. Rao and A. Edelman. The polynomial method for random matrices. *Found. Comput. Math.*, 8(6), 649–702, 2008.
- [34] J. W. Silverstein Some limit theorems on the eigenvectors of large-dimensional sample covariance matrices *J. Multivariate Anal.*, 15(3), 295–324, 1984.
- [35] J. W. Silverstein On the eigenvectors of large-dimensional sample covariance matrices *J. Multivariate Anal.*, 30(1), 1–16, 1989.
- [36] J. W. Silverstein Weak convergence of random functions defined by the eigenvectors of sample covariance matrices *J. Multivariate Anal.*, 18(3), 1–16, 1990.
- [37] G. W. Stewart and J. G. Sun. *Matrix perturbation theory*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1990.

- [38] D. V. Voiculescu, K. J. Dykema, and A. Nica. *Free random variables*, volume 1 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1992. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups.
- [39] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

FLORENT BENAYCH-GEORGES, LPMA, UPMC UNIV PARIS 6, CASE COURIER 188, 4, PLACE JUSSIEU, 75252 PARIS CEDEX 05, FRANCE, AND CMAP, ÉCOLE POLYTECHNIQUE, ROUTE DE SACLAY, 91128 PALAISEAU CEDEX, FRANCE.

E-mail address: florent.benaych@upmc.fr

URL: <http://www.cmapx.polytechnique.fr/~benaych/>

RAJ RAO NADAKUDITI, DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, UNIVERSITY OF MICHIGAN, 1301 BEAL AVENUE, ANN ARBOR, MI 48109. USA.

E-mail address: rajnrao@eecs.umich.edu

URL: <http://www.eecs.umich.edu/~rajnrao/>