ALEA, Lat. Am. J. Probab. Math. Stat. **19**, 769–797 (2022)
DOI: 10.30757/ALEA.v19-30

Asymptotics of k dimensional spherical integrals and applications

Alice Guionnet and Jonathan Husson

Université de Lyon, ENSL, CNRS, France
E-mail address: Alice.Guionnet@ens-lyon.fr

University of Michigan, Ann Arbor
E-mail address: jhusson@umich.edu

Abstract. In this article, we prove that k-dimensional spherical integrals are asymptotically equivalent to the product of 1-dimensional spherical integrals. This allows us to generalize several large deviations principles known before only in a one-dimensional case. For example, we prove the universality of the large deviations principle for the law of k extreme eigenvalues of Wigner matrices (resp. Wishart matrices, resp. matrices with general variance profiles) with sharp sub-Gaussian entries, as well as derive large deviations principles for the distribution of extreme eigenvalues of Gaussian Wigner and Wishart matrices with a finite dimensional perturbation.

1. Introduction

Spherical integrals are integrals over the unitary or orthogonal group which can be seen as natural Fourier (or Laplace transforms) over matrices. As such, they play a central role in random matrix theory. They can for instance be used to express the density of the distribution of random matrices see Coquereaux et al. (2020) and Zuber (2018). In the unitary case (and more generally when one integrates over a compact, connected, semisimple Lie group), Harish-Chandra (1956) and Itzykson and Zuber (1980) derived formulas for such integrals. However, these formulas do not allow to estimate in general their asymptotics as the dimension goes to infinity because they are given in terms of a determinant, so a signed sum of diverging terms. It is however crucial to estimate such asymptotics in random matrix theory to prove laws of large numbers for some matrix models or large deviations principles, see Guionnet and Husson (2020); Husson (2020); Belinschi et al. (2020); Collins et al. (2009); Guionnet and Maurel-Segala (2006). These asymptotics also permit to see the R-transform as the limit of spherical integrals, and thus as a natural Laplace transform in the space of matrices as shown in Guionnet and Maïda (2005). This representation was recently generalized to the S-transform in Potters and Mergny (2020), see also Benaych-Georges (2011) for the rectangular R-transform. Spherical integrals depend generically on two Hermitian matrices; in the sequel, one will have full dimension N whereas the other will have dimension k smaller or equal

Received by the editors February 26th, 2021; accepted January 7th, 2022.
2010 *Mathematics Subject Classification.* 60B20,60F10.
Key words and phrases. Random matrices, Spherical integrals, Large deviations.
This work was supported in part by ERC Project LDRAM : ERC-2019-ADG Project 884584.

to N. In the case of a one dimensional spherical integral ($k = 1$) which amounts to computing the Laplace transform of the matrix evaluated in the direction of a uniformly distributed vector on the sphere, these asymptotics were derived in Guionnet and Maïda (2005) (see also Gorin and Panova (2015)). The case where the matrix has a full rank ($k = N$) was addressed in Guionnet and Zeitouni (2002). In the case where the exponent is small enough, and the spherical integral is k-dimensional, with k much smaller than the dimension, the spherical integrals were shown to be equivalent to a product of one dimensional spherical integrals when k is finite in Guionnet and Maïda (2005), or going to infinity in a mesoscopic regime where k grows like a power of the dimension Collins and Śniady (2007); Huang (2019). In this article, we show that when k is finite, this property remains true for all ranges of parameters. In the small parameter range, the limit of spherical integrals only depends on the limit of the empirical measure of the eigenvalues and not of the outliers, whereas it was shown in Guionnet and Maïda (2005, Theorem 6) that for large parameters the limit of one dimensional spherical integrals also depends on the largest eigenvalue. For general parameters, we prove that the limit of k-dimensional spherical integrals is equivalent to the product of one dimensional integrals which are evaluated at the successive largest eigenvalues. For instance, as foreseen in Maïda et al. (2007), the limit of a 2-dimensional spherical integral depends on the two largest outliers in the large parameters regime, and not only on the top one. As a consequence, k-dimensional spherical integrals allow us to study the universality of the large deviations for the joint distribution of k extreme eigenvalues of Wigner matrices with sharp sub-Gaussian entries, hence generalizing the results of Guionnet and Husson (2020) to finitely many extreme eigenvalues. Similarly, we extend to finitely many extreme eigenvalues the universality of large deviations for Wishart matrices obtained in Guionnet and Husson (2020) and for matrices with general variance profile studied in Husson (2020) with sharp sub-Gaussian tails. We also prove large deviations principles for extreme eigenvalues of Gaussian Wigner and Wishart matrices with a finite dimensional perturbation. This generalizes the one-dimensional case derived in Maïda (2007). The large deviations rate functions of these large deviations principles simply decompose as the sum of the one dimensional rate functions.

The approach of this paper differs from the arguments used in Guionnet and Maïda (2005) in the one-dimensional case, which relied heavily on the representation of the uniform law on the sphere in terms of Gaussian variables. Instead, it is based on considering first spherical integrals of matrices with finitely many different eigenvalues. In this case, the uniform law on the sphere can be easily described by Beta-distributions allowing to use Laplace's principle and rate functions can be more simply described as maxima over real numbers, see Section 3. We then generalize our results to matrices with a continuous spectrum by density, see Section 4. Applications to large deviations principles for the law of the extreme eigenvalues of random matrices are given in Section 5.

2. Statement of the results

We denote $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$ the set of non-negative real numbers. We consider a $N \times N$ Hermitian matrix $\mathbf{X}_N$ such that the empirical measure of its eigenvalues

$$\hat{\mu}_{\mathbf{X}_N} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i}$$

converges towards a probability measure μ with support with rightmost point r_μ and leftmost point l_μ which are assumed to be finite. Let k, ℓ be two integer numbers. Let $\lambda_1^N \geq \lambda_2^N \geq \cdots \geq \lambda_k^N \geq r_\mu$ be the k largest outliers of $\mathbf{X}_N$, $\lambda_{-1}^N \leq \cdots \leq \lambda_{-\ell}^N \leq l_\mu$ be the ℓ smallest outliers of $\mathbf{X}_N$. Here, each eigenvalue has multiplicity one (but they can be equal). Assume that there exists $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > r_\mu > l_\mu > \lambda_{-\ell} \geq \lambda_{-\ell+1} \geq \cdots \geq \lambda_{-1}$ so that

$$\lim_{N\to\infty} \lambda_i^N = \lambda_i > r_\mu \text{ for } i \in [1, k], \lim_{N\to\infty} \lambda_{-i}^N = \lambda_{-i} < l_\mu \text{ for } i \in [1, \ell]$$

The main result of our paper is the following asymptotics for the $k+\ell$ dimensional spherical integral of $\mathbf{X}_N$. Denote by $(e_i)_{\substack{-\ell\leq i\leq k\\ i\neq 0}}$ a family of $k+\ell$ orthonormal eigenvectors following the uniform law on the sphere with radius one, taken with complex coordinates if $\beta = 2$ and real coordinates if $\beta = 1$.

Proposition 2.1. *Let $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_k \geq 0 \geq \theta_{-\ell} \geq \cdots \geq \theta_{-1}$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{\substack{i=-l\\ i\neq 0}}^{k} \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]$$

$$= \frac{\beta}{2}\Big(\sum_{i=1}^{k} J(\mu,\theta_i,\lambda_i) + \sum_{i=1}^{\ell} J(\mu,\theta_{-i},\lambda_{-i})\Big).$$

Here, $J(\mu,0,\lambda) = 0$. For $\theta > 0$, $J(\mu,\theta,\lambda) = K\Big(\mu,\theta,\lambda,v(\mu,\theta,\lambda)\Big)$ with

$$K(\mu,\theta,\lambda,v) = \theta\lambda + (v-\lambda)G_\mu(v) - \log\theta - \int \log|v-x|d\mu(x) - 1$$

and

$$v(\mu,\theta,\lambda) = \begin{cases} \lambda \text{ if } G_\mu(\lambda) \leq \theta, \\ G_\mu^{-1}(\theta) \text{ if } G_\mu(\lambda) > \theta. \end{cases}.$$

G_μ denotes the Cauchy-Stieltjes transform of μ given, for z outside the support of μ, by $G_\mu(z) = \int(z-x)^{-1}d\mu(x)$.

For $\theta < 0$, if we denote by μ_- the push-forward of μ by $x \mapsto -x$, $\mu_-(x \in .) = \mu(-x \in .)$, $J(\mu,\theta,\lambda) = J(\mu_-,-\theta,-\lambda)$.

This key proposition is proved in Section 3 in the case of a spectrum with finitely many different points and in Section 4 in the general case. The convergence of finite rank spherical integral was already derived in Gorin and Panova (2015, Theorem 3.7 and Corollary 3.11) but in cases where there are no outliers. As a first application, we generalize the universality of the large deviations of the largest eigenvalue for Wigner matrices with sharp sub-Gaussian tails proved by Guionnet and Husson (2020) to the k-th extreme eigenvalues. We consider a Wigner matrix $\mathbf{X}_N$ with entries $\left(\frac{X_{ij}}{\sqrt{N}}\right)_{1\leq i,j\leq N}$ where $(X_{ij})_{i\leq j}$ are independent centered variables such that

$$\mathbb{E}[|X_{ij}|^2] = 1, \quad i<j \text{ and } \mathbb{E}[|X_{ii}|^2] = 2^{1_{\beta=1}} \tag{2.1}$$

where $\beta = 1$ if the entries are real, and $\beta = 2$ if they are complex. In the complex case we assume that the real and the imaginary part of $X_{ij}, 1 \leq i < j \leq N$, are independent. We moreover assume that the X_{ij} have sharp sub-Gaussian tails in the sense that

$$\mathbb{E}[\exp(\Re(aX_{ij}))] \leq \exp(\frac{1}{2}\mathbb{E}[\Re(\bar{a}X_{ij})^2]) \tag{2.2}$$

for any complex number a. Note that in the case where the X_{ij} are real, it is enough to take a real and the real part can be removed. We finally make the following concentration assumption.

Assumption 2.2. *We say that $\mathbf{X}_N$ concentrates if the spectral radius of $\mathbf{X}_N$, $||\mathbf{X}_N||$, as well as the empirical measure $\hat{\mu}_{\mathbf{X}_N}$ of its eigenvalues satisfy the following properties. First, $\|\mathbf{X}_N\|$ is exponentially tight at the scale N:*

$$\lim_{K\to+\infty} \limsup_{N\to+\infty} \frac{1}{N} \log \mathbb{P}(||\mathbf{X}_N|| > K) = -\infty. \tag{2.3}$$

Moreover, the empirical distribution of the eigenvalues $\hat{\mu}_{\mathbf{X}_N}$ concentrates at the scale N: There exists a probability measure π such that

$$\limsup_{N\to+\infty} \frac{1}{N} \log \mathbb{P}\Big(d(\hat{\mu}_{\mathbf{X}_N}, \pi) > \varepsilon(N)\Big) = -\infty, \tag{2.4}$$

for some $\varepsilon(N)$ goes to zero as N goes to infinity. Here, d is a distance compatible with the weak topology.

In our previous paper Guionnet and Husson (2020) we took $\varepsilon(N) = N^{-\kappa}$. This hypothesis was needed to ensure the continuity of spherical integrals according to Maïda (2007). However, part of the consequences of our new approach is that spherical integrals enjoy better continuity properties, see the Appendix 6. Assumption 2.2 is satisfied by all the matrix models we shall consider below (Wigner, Wishart and variance profile) as soon as the entries satisfy log-Sobolev inequality with uniformly bounded constant or are bounded (see Anderson et al. (2010) and the Appendix in Guionnet and Husson (2020)). Examples of entries satisfying all our hypotheses (including (2.2)) are Rademacher variables or uniform variables. We prove the following universality of the large deviations for the extreme eigenvalues of $\mathbf{X}_N$:

Theorem 2.3. *Let $\mathbf{X}_N = (\frac{X_{ij}}{\sqrt{N}})_{i,j}$ be a $N \times N$ Hermitian matrix where $(X_{i,j})_{i\le j}$ are centered independent entries satisfying (2.1) and (2.2), as well as such that $\mathbf{X}_N$ satisfies Assumption 2.2. Let $\lambda_1^N \ge \lambda_2^N \ge \cdots \ge \lambda_N^N$ be the N eigenvalues of $\mathbf{X}_N$. Let k, ℓ be fixed integer numbers. Then the law of $\bar{\lambda}^N = (\lambda_1^N, \lambda_2^N, \ldots, \lambda_k^N, \lambda_{N-\ell+1}^N, \ldots, \lambda_N^N)$ satisfies a large deviations principle in the scale N and with good rate function $I(x_1, \ldots, x_k, x_{-\ell}, \ldots, x_{-1})$ which is infinite unless $\bar{x} = (x_1, \ldots, x_k, x_{-\ell}, \ldots, x_{-1})$ satisfies*

$$x_1 \ge x_2 \ge \cdots \ge x_k \ge 2 \ge -2 \ge x_{-\ell} \ge x_{-\ell+1} \ge \cdots \ge x_{-1}$$

and is then given by

$$I(x_1, \ldots, x_k, x_{-\ell}, \ldots, x_{-1}) = \frac{\beta}{2}\Big(\sum_{i=1}^{k} \int_2^{x_i} \sqrt{t^2-4}dt + \sum_{i=1}^{\ell} \int_2^{-x_{-i}} \sqrt{t^2-4}dt\Big).$$

First note that if $\mathbf{X}_N$ satisfies Assumption 2.2, we must take π to be equal to the semi-circle law, defined by

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbb{1}_{|x|\le 2}dx.$$

Indeed, we know the semi-circle law is the almost sure limit of the empirical measure of the eigenvalues since Wigner (1955), see also Anderson et al. (2010). The proof of this theorem is given in Section 5.1. This result is well known in the Gaussian case for the case $k = 1$, see Anderson et al. (2010, Section 2.6.2) and Ben Arous et al. (2001). The case of more general k but Gaussian entries is a straightforward generalization, see e.g. Biroli and Guionnet (2020). The case of sharp sub-Gaussian entries and $k = 1$ was proven in Guionnet and Husson (2020, Theorem 1.4 and Theorem 1.5). This result can also be generalized to Wishart matrices. We consider $\mathbf{G}_{L,M}$ a $L \times M$ random matrix and set $N = L + M$. We define the Wishart matrix $\mathbf{W}_{L,M} = \frac{1}{M}\mathbf{G}_{L,M}\mathbf{G}^*_{L,M}$. If L/M goes to $\alpha \le 1$, it is well known that the spectral measure of $\mathbf{W}_{M,L}$ converges towards the Pastur-Marchenko distribution

$$d\pi_\alpha(x) = \frac{1}{2\pi\alpha x}\sqrt{(\lambda_+ - x)(x-\lambda_-)}dx\,.$$

where $\lambda_\pm = (1 \pm \sqrt{\alpha})^2$. Then we have the following :

Theorem 2.4. *Let $\mathbf{G}_{L,M} = (X_{ij})_{\substack{1\le i\le L\\ 1\le j\le M}}$ be a $L\times M$ matrix where $(X_{i,j})_{i,j}$ are centered independent entries satisfying (2.1) and (2.2), as well as such that $\mathbf{W}_{L,M}$ satisfies Assumption 2.2 (with $\pi = \pi_\alpha$). Let $k \ge 0$ and $\lambda_1^N \ge \ldots \ge \lambda_k^N$ be the k largest eigenvalues of $\mathbf{W}_{L,M}$. Assume that there exists*

$\alpha \in (0,1]$ *so that* $\lim_{N\to\infty} L/M = \alpha$*. Then* $(\lambda_1^N, ..., \lambda_k^N)$ *satisfies a large deviations principle in the scale* N *with good rate function* $J(x_1, ..., x_k)$ *which is infinite unless* $x_1 \geq ... \geq x_k \geq \lambda_+$ *and :*

$$J(x_1, ..., x_k) = \frac{\beta}{4(1+\alpha)} \sum_{i=1}^{k} \int_{\lambda_+}^{x_i} \frac{\sqrt{(y-\lambda_-)(y-\lambda_+)}}{y} dy$$

As in the Wigner case, as soon as the entries satisfy a log-Sobolev inequality or are compactly supported, the empirical measure converges towards the Pastur-Marchenko distribution with probability larger than any exponential and the norm of Wishart matrices stays bounded, yielding a property similar to Assumption 2.2 for $\mathbf{W}_{L,M}$. Note that the article Guionnet and Husson (2020) initially had the stronger assumption that $L/M - \alpha = O(N^{-\kappa})$ for some $\kappa > 0$ but there again, using the new continuity properties of this paper, see Theorem 6.1, we can relax this hypothesis. This theorem is proved in Section 5.2.

This result can be further extended to Wigner matrices with variance profiles. Those matrices are built by letting $\mathbf{X}_N^{\sigma}(i,j) = \sigma_N(i,j)\frac{X_{i,j}}{\sqrt{N}}$ where :

- either there exists $p \in \mathbb{N}$, $\alpha_1(N), ..., \alpha_p(N) > 0$ such that $\sum_1^p \alpha_i(N) = N$ and $\lim \alpha_i(N)/N = \alpha_i > 0$, and $(\sigma_{i,j})_{i,j} \in M_{p,p}(\mathbb{R}^+)$, $\sigma = \sigma^T$, such that :

  $$\sigma_N(i,j) = \sum_{k,l=1}^{p} \sigma_{k,l} \mathbb{1}_{I_N^k \times I_N^l}(i,j)$$

  where $I_N^1 = [\![1, \alpha_1(N)]\!]$ and $I_N^{i+1} = [\![\sum_{j=1}^{i} \alpha_j(N) + 1, \sum_{j=1}^{i+1} \alpha_j(N)]\!]$. This case will be called the piecewise constant case with parameters σ and α.
- or $\sigma_N(i,j) = \sigma(i/N, j/N)$ where σ is a continuous symmetric positive function of $[0,1]^2$. This case will be called the continuous case.

We will also make the following assumption on the variance profiles :

Assumption 2.5.

- *In the piecewise constant case, we assume that the quadratic form* $\psi \mapsto \sum_{i,j}^{p} \sigma_{i,j}^2 \psi_i \psi_j$ *is negative on the subspace* $Vect(1, ..., 1)^{\perp}$.
- *In the continuous case, we assume that the function* $\psi \mapsto \int \sigma^2(x,y) d\psi(x) d\psi(y)$ *is concave on the set* $\mathcal{P}([0,1])$ *of probability measures on* $[0,1]$.

When this Assumption, as well as Assumption 2.2 and (2.2), are verified, the almost sure convergence of the empirical measure of the eigenvalues towards a limiting profile μ_σ is guaranteed and one of the authors of this article proved in Husson (2020) that the largest eigenvalue of $\mathbf{X}_N^{\sigma}$ satisfies a large deviations principle with a good rate function $J_\sigma^{(1)}$. It is defined as follows. Let r_σ be the rightmost point of the support of the limit μ_σ of the empirical measure of $\mathbf{X}_N$. Then $J_\sigma^{(1)}$ is defined for $x \geq r_\sigma$ as :

$$J_\sigma^{(1)}(x) = \frac{\beta}{2} \sup_{\theta \geq 0} [J(\theta, x, \mu_\sigma) - F_\sigma(\theta)] \tag{2.5}$$

In the piecewise constant case, F_σ is defined for $\theta \geq 0$ by:

$$F_\sigma(\theta) = \max_{\substack{\sum_{i=1}^{p} \psi_i = 1 \\ \psi \in (\mathbb{R}^+)^p}} \Big[\frac{\theta^2}{2} \sum_{1 \leq i,j \leq p} \sigma_{i,j}^2 \psi_i \psi_j - \sum_{i=1}^{p} \alpha_i \log(\psi_i/\alpha_i) \Big]$$

In the continuous one, F_σ is defined for $\theta \geq 0$ by:

$$F_\sigma(\theta) = \max_{\mu \in \mathcal{P}[0,1]} \Big[\frac{\theta^2}{2} \int \sigma(x,y) d\mu(x) d\mu(y) - I(Leb||\mu) \Big]$$

where I is the Kullback-Leibler divergence and Leb the Lebesgue measure on $[0,1]$. For x below the leftmost point l_σ of the support of μ_σ, we set $J_\sigma^{(1)}(x) = J_\sigma^{(1)}(-x)$. In this article we generalize this result to the k-th largest eigenvalues and prove the following theorem :

Theorem 2.6. *Let $\mathbf{X}_N^\sigma = (\frac{\sigma_N(i,j)X_{ij}}{\sqrt{N}})_{i,j}$ be a $N \times N$ Hermitian matrix where $(X_{i,j})_{i\le j}$ are centered independent entries satisfying (2.1) and (2.2), as well as such that $\mathbf{X}_N^\sigma$ satisfies Assumption 2.2 with $\pi = \mu_\sigma$ and σ verifies Assumption 2.5. Let $k \ge 0$ and $\lambda_1^N \ge ... \ge \lambda_k^N$ be the k largest eigenvalues of $\mathbf{X}_N^\sigma$. Then $(\lambda_1^N, ..., \lambda_k^N)$ satisfies a large deviations principle in the scale N with good rate function $J_\sigma^{(k)}(x_1, ..., x_k)$ which is infinite unless $x_1 \ge ... \ge x_k \ge r_\sigma$ and in this case equals:*

$$J_\sigma^{(k)}(x_1, ..., x_k) = \sum_{i=1}^{k} J_\sigma^{(1)}(x_i)$$

where $J_\sigma^{(1)}$ is the large deviations rate function given in (2.5).

This result was proved in the case $k = 1, \ell = 0$ in Guionnet and Maïda (2005, Theorem 6). We outline the proof of Theorem 2.6 in Section 5.3.

Let us now consider $\mathbf{X}_N$ to be a GOE/GUE matrix, that is a $N \times N$ Hermitian matrix with centered real/complex Gaussian entries satisfying (2.1). Let ℓ and k be two integer numbers and let $(e_1, \ldots e_k, e_{-1}, \ldots, e_{-\ell})$ be orthonormal vectors following the uniform law on the sphere. Maïda (2007) showed that the largest eigenvalue of a Gaussian Wigner matrix perturbed by a rank one matrix satisfy a large deviations principle. In this article we generalize this result to the k-th largest eigenvalues and ℓ smallest eigenvalues when the Gaussian matrix is perturbed by a finite rank matrix with k non-negative eigenvalues and ℓ non-positive eigenvalues.

Proposition 2.7. *Let $\mathbf{X}_N$ be a GUE ($\beta = 2$) or GOE ($\beta = 1$) matrix. Let ℓ, k be two finite integer numbers. Let $\theta_1 \ge \theta_2 \ge \cdots \ge \theta_k \ge 0 \ge \theta_{-\ell} \ge \cdots \ge \theta_{-1}$ and define*

$$\mathbf{X}_N^\theta = \mathbf{X}_N + \sum_{\substack{-\ell \le i \le k \\ i \ne 0}} \theta_i e_i e_i^* .$$

Let $\lambda_1^{N,\theta} \ge \lambda_2^{N,\theta} \ge \cdots \ge \lambda_N^{N,\theta}$ be the eigenvalues of $\mathbf{X}_N^\theta$. Then, the distribution of $(\lambda_1^{N,\theta}, \ldots, \lambda_k^{N,\theta}, \lambda_{N-\ell+1}^{N,\theta}, \ldots, \lambda_N^{N,\theta})$ satisfies a large deviations principle in the scale N and with good rate function which is infinite unless

$$x_1 \ge x_2 \ge \cdots \ge x_k \ge 2 \ge -2 \ge x_{-\ell} \ge \cdots \ge x_{-1}$$

and is given then by $\beta \sum_{\substack{-\ell \le i \le k \\ i \ne 0}} I_{\theta_i}(x_i)$. Here, with $I(x) = \frac{1}{4}x^2 - \int \log|x - y| d\sigma(y)$, we have set

$$I_\theta(x) = I(x) - \frac{1}{2}J(\sigma, \theta, x) - \inf_y (I(y) - \frac{1}{2}J(\sigma, \theta, y)) .$$

This Proposition is proved in Section 5.4. A similar result holds for finite rank perturbation of Gaussian Wishart matrices. Indeed, let us consider a $L \times M$ matrix $\mathbf{G}_{L,M}$ with i.i.d standard Gaussian matrices with covariance 1, set $N = M + L$, and assume without loss of generality that $M \ge L$. We consider the Wishart matrix

$$\mathbf{W}_N^\gamma = \frac{1}{M} \Sigma_L^{1/2} \mathbf{G}_{L,M} \mathbf{G}_{L,M}^* \Sigma_L^{1/2}$$

where Σ_L is a $L \times L$ covariance matrix given by $I_L + \sum_{i=1}^k \gamma_i e_i e_i^*$ for some fixed $\gamma_i > -1$. Here $(e_i, 1 \le i \le k)$ are k orthonormal vectors. It is well known that when L/M goes to $\alpha \in [0,1]$, the empirical measure of the eigenvalues of $\mathbf{W}_N^\gamma$ goes to the Pastur-Marchenko distribution.

Large deviations for the extreme eigenvalues in the case $\gamma_i = 0$ are well known, and similar to the Gaussian case, see Anderson et al. (2010); Dean and Majumdar (2006). The rate function governing the large deviations in the scale N for the smallest eigenvalue is infinite outside $[0, \lambda_-]$ and is given for $y \in [0, \lambda_-]$ by $I_\alpha(y)$ defined by

$$I_\alpha(y) = \frac{\beta}{4(1+\alpha)}(y - (1-\alpha)\log y - 2\alpha \int \log|y-t| d\pi_\alpha(y) - C)$$
$$= \frac{\beta}{4(1+\alpha)} \int_y^{\lambda_-} \frac{1}{t}\sqrt{(t-\lambda_+)(t-\lambda_-)} dt$$

where C is the infimum of $y - (1-\alpha)\log y - 2\alpha \int \log|y-t| d\pi_\alpha(y)$. The same result holds for the largest eigenvalue where the rate function is infinite on $[0, \lambda_+)$ and otherwise given by I_α which for $y \in [\lambda_+, \infty)$ equals to

$$I_\alpha(y) = \frac{\beta}{4(1+\alpha)} \int_{\lambda_+}^{y} \frac{1}{t}\sqrt{(t-\lambda_+)(t-\lambda_-)} dt\,.$$

We have the following analogue to Proposition 2.7.

Proposition 2.8. *Let ℓ, k be two finite integer numbers. Let $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_k \geq 0 \geq \gamma_{-\ell} \geq \cdots \geq \gamma_{-1} > -1$. Let $\lambda_1^{N,\gamma} \geq \lambda_2^{N,\gamma} \geq \cdots \geq \lambda_M^{N,\gamma}$ be the eigenvalues of $\mathbf{W}_N^\gamma$ in decreasing order. Then, the law of $(\lambda_1^{N,\gamma}, \ldots, \lambda_k^{N,\gamma}, \lambda_{M-\ell+1}^{N,\gamma}, \ldots, \lambda_M^{N,\gamma})$ satisfies a large deviations principle in the scale N and with good rate function which is infinite unless*

$$x_1 \geq x_2 \geq \cdots \geq x_k \geq \lambda_+ \geq \lambda_- \geq x_{-\ell} \geq \cdots \geq x_{-1} \geq 0$$

and is given otherwise by $\sum_{\substack{i=-\ell \\ i \neq 0}}^{k} I_{\gamma_i,\alpha}(x_i)$ where we have set

$$I_{\gamma,\alpha}(x) = I_\alpha(x) - \frac{\beta}{2} J(\pi_\alpha, \frac{\gamma}{1-\gamma}, x) - \inf_y (I_\alpha(y) - \frac{\beta}{2} J(\pi_\alpha, \frac{\gamma}{1-\gamma}, y))$$

This result is proved in Section 5.5. We finally notice that since our results hold for any number of eigenvalues, they capture as well the large deviations for the point processes of the outliers. For instance, if we let $A_i = [a_i, b_i]$ be intervals above the bulk, $b_i < a_{i+1} < b_{i+1}$, if we denote I the large deviation rate function for any of the above models, the probability that there are n_i outliers in the set A_i has probability of order $\exp\{-N \sum_i n_i \inf_{A_i} I\}$.

Acknowledgements We are very grateful to Mylène Maïda and Marc Potters for preliminary discussions about the questions addressed in this paper, as well as to the anonymous referees for their comments which allowed to improve the presentation of our results.

3. Limiting spherical integral in the discrete case

Throughout this section, we restrict ourselves to the case where $\mathbf{X}_N$ has finitely many different eigenvalues :

$$\mathbf{X}_N = diag\Big(\underbrace{\eta_{-m_1+1}, \ldots, \eta_{-m_1+1}}_{N_{-m_1+1}}, \underbrace{\eta_{-m_1+2}, \ldots, \eta_{-m_1+2}}_{N_{-m_1+2}}, \ldots, \underbrace{\eta_{p+m_2}, \ldots, \eta_{p+m_2}}_{N_{p+m_2}}\Big), \tag{3.1}$$

where $(\eta_i)_{1-m_1 \leq i \leq p+m_2}$ is a $p+m_1+m_2$-tuple of real numbers such that $\eta_{-m_1+1} < \ldots < \eta_1 < \cdots < \eta_p < \eta_{p+1} < \cdots < \eta_{p+m_2}$. Moreover, η_i has multiplicity N_i where $(N_i)_{1-m_1 \leq i \leq p+m_2}$ is a $p+m_1+m_2$-tuple of integer numbers such that $\sum_{i=-m_1+1}^{p+m_2} N_i = N$. We assume that N_i/N goes to a positive limit α_i for $i \in \{1, p\}$ and to zero for $i \in \{1-m_1, \ldots, 0\} \cup \{p+1, \ldots, p+m_2\}$, the later representing the outliers of $\mathbf{X}_N$. m_1, m_2, p are independent of N (with the convention that if $\alpha_i = 0$, $\alpha_i \log \alpha_i = 0$ and $\log 0 = -\infty$). In the previous notations, the eigenvalues of $\mathbf{X}_N$ are given by $\lambda_1^N \geq \lambda_2^N \geq \cdots \geq \lambda_N^N$ with $\lambda_i^N = \eta_{p+m_2}$ for $i \in I_{p+m_2} = [1, N_{p+m_2}]$ and for $i \geq N_{p+m_2}+1$,

$$\lambda_i^N = \eta_j, i \in I_j = [N_{p+m_2} + \cdots + N_{j+1} + 1, N_{p+m_2} + \cdots + N_j] \tag{3.2}$$

Remark 3.1. We notice that if the sequences N_i are fixed, the spherical integrals are $\beta/2$-Lipschitz in the $p+m_1+m_2$-tuple $(\eta_i)_{-m_1+1\leq i\leq p+m_2}$ with the norm $||.||_\infty$.

3.1. *Limiting 1-d spherical integral.* We start by giving a new proof of Guionnet and Maïda (2005, Theorem 6) giving the asymptotics of spherical integrals in the one dimensional case, in the case of matrices with $p+m_1+m_2$ different eigenvalues with multiplicity as above. This proof will be extended to the higher dimensional setting in the next subsection.

Proposition 3.2. *Let $\theta \geq 0$ and $\mathbf{X}_N$ be given by (3.1). Then, if e follows the uniform law on the sphere $\mathbb{S}^{N-1}$ of radius one, we have*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}\theta\langle e, \mathbf{X}_N e\rangle\Big)\Big] = \frac{\beta}{2} \sup_{\substack{\gamma_i\geq 0\\ \sum_{i=1-m_1}^{p+m_2}\gamma_i=1}} \Big\{\theta \sum_{i=-m_1+1}^{p+m_2} \eta_i\gamma_i + \sum_{i=1}^{p}\alpha_i \log\frac{\gamma_i}{\alpha_i}\Big\}$$

Proof: We have the following formula :

$$\langle e, \mathbf{X}_N e\rangle = \sum_{i=-m_1+1}^{p+m_2} \eta_i\gamma_i^N$$

where we have denoted $\gamma_j^N = \sum_{i\in I_j}|u_i|^2$ with $u_i = \langle v_i, e\rangle$ if v_i is the eigenvector for the i-th eigenvalue of $\mathbf{X}_N$. In other words, γ_j^N is the ℓ^2-norm of the projection of e onto the eigenspace of η_j. The vector γ^N follows a Dirichlet law of parameters $\frac{\beta}{2}(N_{1-m_1},\dots,N_{p+m_2})$, that is the distribution on $\Sigma = \{x\in[0,1]^{m_1+m_2+p} : \sum_{i=1-m_1}^{m_2+p} x_i = 1\}$ given by $\gamma_{1-m_1}^N = 1 - \sum_{i=2-m_1}^{p+m_2}\gamma_i^N$ and

$$d\mathbb{P}_{\mathbf{N}}^N(\gamma) = \frac{1}{Z_\alpha^N}\mathbf{1}_{\sum_{i=2-m_1}^{p+m_2}\gamma_i\leq 1}(1-\sum_{i=2-m_1}^{p+m_2}\gamma_i)^{\frac{\beta}{2}N_{1-m_1}-1}\prod_{j=2-m_1}^{p+m_2}\gamma_j^{\frac{\beta}{2}N_i-1}\mathbf{1}_{\gamma_j\geq 0}d\gamma_j\,. \tag{3.3}$$

From this explicit formula of the density, we deduce the following large deviations principle

Theorem 3.3. *Assume that N_i/N converges towards α_i for all i, with $\alpha_i = 0$ for $i\notin[1,p]$. Then, the law of γ^N satisfies a large deviations principle with scale N and good rate function I_α given for $x\in\Sigma$ by*

$$I_\alpha(x_{1-m_1},...,x_{p+m_2}) = -\frac{\beta}{2}\sum_{i=1}^{p}\alpha_i\log\frac{x_i}{\alpha_i}\,.$$

The proof is a direct consequence of Laplace's method. We deduce Proposition 3.2 from Theorem 3.3 by Varadhan's lemma. □

Lemma 3.4. *Let $(\alpha_i)_{1\leq i\leq p}\in(\mathbb{R}^+)^p$ such that $\sum_{i=1}^p\alpha_i = 1$. For $\theta\geq 0$ and $\eta = (\eta_{1-m_1} < \cdots < \eta_{p+m_2})$, the function J given by*

$$J(\theta,\eta) := \sup_{\substack{\gamma_i\geq 0\\ \sum\gamma_i=1}} \Big\{\theta\sum_{i=-m_1+1}^{p+m_2}\eta_i\gamma_i + \sum_{i=1}^{p}\alpha_i\log\frac{\gamma_i}{\alpha_i}\Big\}$$

only depends on η_{p+m_2}, θ and $\mu = \sum_{i=1}^p \alpha_i\delta_{\eta_i}$. For $\theta = 0$, $J(0,\eta)$ vanishes and for $\theta > 0$, it is given by:

$$J(\theta,\eta) = J(\mu,\theta,\eta_{p+m_2}) = K\Big(\mu,\theta,\eta_{p+m_2}, v(\mu,\theta,\eta_{p+m_2})\Big)$$

with

$$K(\mu,\theta,\lambda,v) = \theta\lambda + (v-\lambda)G_\mu(v) - \log|\theta| - \int\log|v-x|d\mu(x) - 1$$

and

$$v(\mu,\theta,\lambda)=\begin{cases}\lambda \text{ if } G_\mu(\lambda)\le\theta,\\ G_\mu^{-1}(\theta) \qquad\qquad \text{if } G_\mu(\lambda)>\theta.\end{cases}.$$

Proof: $J(\theta,\eta)$ is the supremum over $\Sigma=\{x\in(\mathbb{R}^+)^{p+m_1+m_2}:\sum_{i=1-m_1}^{p+m_2}x_i=1\}$ of the function

$$I_{\theta,\eta}^{p+m_2}(\gamma):=\theta\sum_{i=-m_1+1}^{p+m_2}\eta_i\gamma_i+\sum_{i=1}^{p}\alpha_i\log\frac{\gamma_i}{\alpha_i}\,. \tag{3.4}$$

Observe that $I_{\theta,\eta}^{p+m_2}$ is continuous except when the γ_i vanishes but then is equal to $-\infty$. Hence, since Σ is compact the supremum is achieved. The entropic term in $I_{\theta,\eta}^{p+m_2}$ does not depend on $(\gamma_i, i<1 \text{ or } i>p)$, and the first term increases when we take these terms equal to zero except γ_{m_2+p}. Hence, the maximum is taken at $\gamma_i=0$ for $i<1$ or $i\in[p+1,p+m_2-1]$. Then, putting $\gamma_{p+m_2}=1-\sum_{i=1}^p\gamma_i$ we see that we need to maximize

$$I_{\theta,\eta}(\gamma)=\theta\eta_{p+m_2}+\Big\{\theta\sum_{i=1}^p(\eta_i-\eta_{p+m_2})\gamma_i+\sum_{i=1}^p\alpha_i\log\frac{\gamma_i}{\alpha_i}\Big\}$$

over $(\gamma_i)_{1\le i\le p}\in(\mathbb{R}^+)^p,\sum_{i=1}^p\gamma_1\le 1$. Note here that we can assume without loss of generality that $\gamma_i>0, 1\le i\le p$. We see that the critical point of $I_{\theta,\eta}$ is

$$\gamma_i^*=\frac{\alpha_i}{\theta(\eta_{p+m_2}-\eta_i)},1\le i\le p,\ \gamma_{p+m_2}^*=1-\sum_{i=1}^p\gamma_i^*=1-\frac{1}{\theta}G_\mu(\eta_{p+m_2})$$

provided $\sum_{i=1}^p\gamma_i^*=\frac{1}{\theta}G_\mu(\eta_{p+m_2})\le 1$. For $\theta<G_\mu(\eta_{p+m_2})$, the supremum is achieved at

$$\gamma_i^{**}=\frac{\alpha_i}{\theta(G_\mu^{-1}(\theta)-\eta_i)},1\le i\le p,$$

so that $\gamma_{p+m_2}^{**}=0$. This gives the announced formula. □

3.2. *Limiting 2-d spherical integral.* We next consider the 2-dimensional case with non negative parameters θ_1 and θ_2 and show that it only depends on the two largest eigenvalues of $\mathbf{X}_N$ as follows. Let (e,f) denote two orthonormal vectors following the uniform law in the sphere. We then have the following proposition.

Proposition 3.5. *Let $\theta_1\ge\theta_2\ge 0$. Then, if $N_{p+m_2}=1$,*

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}(\theta_1\langle e,\mathbf{X}_Ne\rangle+\theta_2\langle f,\mathbf{X}_Nf\rangle)\Big)\Big]$$
$$=\frac{\beta}{2}(J(\mu,\theta_1,\eta_{p+m_2})+J(\mu,\theta_2,\eta_{p+m_2-1}))\,.$$

If $N_{m+p_2}\ge 2$,

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}(\theta_1\langle e,\mathbf{X}_Ne\rangle+\theta_2\langle f,\mathbf{X}_Nf\rangle)\Big)\Big]$$
$$=\frac{\beta}{2}(J(\mu,\theta_1,\eta_{p+m_2})+J(\mu,\theta_2,\eta_{m+p_2}))\,.$$

Proof: We first assume that $N_{p+m_2}=1$. We can write :

$$\mathbb{E}\Big[\exp\Big(N\frac{\beta}{2}\Big(\theta_1\langle e,\mathbf{X}_Ne\rangle+\theta_2\langle f,\mathbf{X}_Nf\rangle\Big)\Big)\Big]$$
$$=\mathbb{E}\Big[\exp\Big(N\frac{\beta}{2}\theta_1\langle e,\mathbf{X}_Ne\rangle\Big)\mathbb{E}\Big[\exp\Big(N\frac{\beta}{2}\theta_2\langle f,\mathbf{X}_N^{(e)}f\rangle\Big)\Big|e\Big]\Big]$$

where $\mathbf{X}_N^{(e)} = P_{e^\perp}\mathbf{X}_N P_{e^\perp}$ if $P_{e^\perp} = I - ee^*$ is the orthogonal projection onto the ortho-complement of e. We can see $\mathbf{X}_N^{(e)}$ as a $(N-1)\times(N-1)$ matrix living in $Vect(e)^\perp$. Its largest eigenvalue χ belongs to $[\eta_{p+m_2-1}, \eta_{p+m_2}]$. Writing that the corresponding eigenvector $v \in Vect(e)^\perp$ satisfies $\mathbf{X}_N^{(e)} v = \chi v$ and $\langle e, v\rangle = 0$, we find that

$$\chi = \chi(\gamma^N(e)) \tag{3.5}$$

where

$$\gamma_i^N(e) = \sum_{j\in I_i} |\langle v_j, e\rangle|^2 \tag{3.6}$$

with v_j the jth eigenvector of $\mathbf{X}_N$ and we remind that the I_i are defined in (3.4). $\chi(.)$ is the function on Σ with values in $[\eta_{p+m_2-1}, \eta_{p+m_2}]$ given by the solution of the equation

$$\sum_{i=1-m_1}^{p+m_2} \frac{x_i}{\chi(x) - \eta_i} = 0 \tag{3.7}$$

if there is a solution in this interval. If there is no solution (which can happen only if $x_{p+m_2} = 0$ or $x_{p+m_2-1} = 0$) then $\chi(x) = \eta_{p+m_2}$ if the rational function is positive on this interval and $\chi(x) = \eta_{p+m_2-1}$ if it is negative. Note that χ is a continuous function on Σ.

Moreover, the spectral measure of $\mathbf{X}_N^{(e)}$ converges towards μ, the limiting spectral measure of $\mathbf{X}_N$ by Weyl's interlacing property. Therefore, on the event where $\gamma^N(e)$ converges towards some $\kappa \in \Sigma$, and since the empirical measure of $\mathbf{X}_N^{(e)}$ converges toward the same limit that the empirical measure of $\mathbf{X}_N$, we have

$$\lim_{N\to\infty} \frac{1}{N-1} \log \mathbb{E}[e^{\theta_2 N \frac{\beta}{2}\langle f, \mathbf{X}_N f\rangle} | e] = \frac{\beta}{2} J(\mu, \theta_2, \chi(\kappa)).$$

Moreover, the right hand side depends continuously on κ (since J is continuous in χ and χ in κ). We can also easily see that the convergence above is uniform in κ. We now can apply the fact that the law of $\gamma^N(e) = (\gamma_i^N(e))_{i=1-m_1}^{p+m_2}$ follows a large deviations principle, see Theorem 3.3, to conclude by Varadhan's lemma that

$$\lim \frac{1}{N} \log \mathbb{E}[\exp\Big(N\frac{\beta}{2}\Big(\theta_1\langle e, \mathbf{X}_N e\rangle + \theta_2\langle f, \mathbf{X}_N f\rangle\Big)\Big)]$$
$$= \frac{\beta}{2} \sup_{\gamma\in\Sigma}\Big(J(\mu,\theta_2,\chi(\gamma)) + \sum_{i=1}^{p} \alpha_i \log\frac{\gamma_i}{\alpha_i} + \sum_{i=1-m_1}^{p+m_2} \theta_1 \eta_i \gamma_i\Big)$$

Since J is bounded and due to the continuity in γ of the function we optimize, we can change the domain of the supremum to $(\mathbb{R}^{+,*})^{p+m_1+m_2}$ where $\mathbb{R}^{+,*} = (0,\infty)$. We next complete the proof by computing the right hand side and showing it equals the sum of the two limiting spherical integrals as stated.

We first denote by $\tilde{\gamma}_i := \gamma_i|\chi - \eta_i|^{-1}$ with $\chi = \chi(\gamma)$. By definition we have $\tilde{\gamma}_i > 0$, $\chi \in (\eta_{p+m_2-1}, \eta_{p+m_2})$ and (3.7) holds so that

$$\tilde{\gamma}_{p+m_2} = \sum_{i=1-m_1}^{p+m_2-1} \tilde{\gamma}_i, \qquad \sum_{i=1-m_1}^{p+m_2-1} (\chi - \eta_i)\tilde{\gamma}_i + (\eta_{p+m_2} - \chi)\tilde{\gamma}_{p+m_2} = 1$$

This simplifies into the condition

$$\tilde{\gamma}_{p+m_2} = \sum_{i=1-m_1}^{p+m_2-1} \tilde{\gamma}_i, \qquad \eta_{p+m_2}\tilde{\gamma}_{p+m_2} - \sum_{i=1-m_1}^{p+m_2-1} \eta_i \tilde{\gamma}_i = 1 \tag{3.8}$$

which is independent of χ. We thus first take the supremum over $\chi \in [\eta_{p+m_2-1}, \eta_{p+m_2}]$ of

$$
\begin{aligned}
I(\chi,\tilde\gamma) &= J(\mu,\theta_2,\chi) + \sum_{i=1}^{p}\alpha_i \log[|\eta_i-\chi|\frac{\tilde\gamma_i}{\alpha_i}] - \theta_1(\chi-\eta_{p+m_2})\eta_{p+m_2}\tilde\gamma_{p+m_2} \\
&\quad + \sum_{i=1-m_1}^{p+m_2-1}\theta_1(\chi-\eta_i)\eta_i\tilde\gamma_i \\
&= H(\chi) + \sum_{i=1}^{p}\alpha_i\log\frac{\tilde\gamma_i}{\alpha_i} + \theta_1\eta_{p+m_2}^2\tilde\gamma_{p+m_2} - \sum_{i=1-m_1}^{p+m_2-1}\theta_1\eta_i^2\tilde\gamma_i
\end{aligned}
$$

with

$$
H(\chi) = J(\mu,\theta_2,\chi) + \sum_{i=1}^{p}\alpha_i\log|\eta_i-\chi| - \chi\theta_1. \tag{3.9}
$$

Recall the formula for J from Lemma 3.4. When $\theta_2 \le G_\mu(\chi)$, that is $\chi \le G_\mu^{-1}(\theta_2)$, J does not depend on χ and the function H increases till $G_\mu^{-1}(\theta_1)$ and decreases afterwards. When $\theta_2 \ge G_\mu(\chi)$, that is $\chi \ge G^{-1}(\theta_2)$, Lemma 3.4 gives

$$
\begin{aligned}
H(\chi) &= \theta_2\chi - \sum_{i=1}^{p}\alpha_i\log(\chi-\eta_i) - \log\theta_2 - 1 + \sum_{i=1}^{p}\alpha_i\log|\eta_i-\chi| - \chi\theta_1 \\
&= \chi(\theta_2-\theta_1) - \log\theta_2 - 1
\end{aligned}
$$

which is decreasing since $\theta_1 > \theta_2$. Therefore, H increases till $G_\mu^{-1}(\theta_1)$ and decreases afterwards. As a consequence,

$$
\max_{\chi\in[\eta_{p+m_2-1},\eta_{p+m_2}]} H(\chi) = \begin{cases} H(\eta_{p+m_2-1}) \text{ if } G_\mu^{-1}(\theta_1) \le \eta_{p+m_2-1}, \\ H(G_\mu^{-1}(\theta_1)) \text{ if } G_\mu^{-1}(\theta_1) \in [\eta_{p+m_2-1}, \eta_{m+p_2}], \\ H(\eta_{p+m_2}) \text{ if } G_\mu^{-1}(\theta_1) > \eta_{p+m_2}. \end{cases} . \tag{3.10}
$$

Let us also optimize on $\tilde\gamma$ satisfying (3.8) the function

$$
L(\tilde\gamma) = \sum_{i=1}^{p}\alpha_i\log[\frac{\tilde\gamma_i}{\alpha_i}] + \theta_1\eta_{p+m_2}^2\tilde\gamma_{p+m_2} - \sum_{i=1-m_1}^{p+m_2-1}\theta_1\eta_i^2\tilde\gamma_i\,.
$$

Replacing $\tilde\gamma_{p+m_2}$ by $\sum_{i=1-m_1}^{p+m_2-1}\tilde\gamma_i$ we get

$$
L(\tilde\gamma) = \sum_{i=1}^{p}\alpha_i\log[\frac{\tilde\gamma_i}{\alpha_i}] + \theta_1\sum_{i=1-m_1}^{p+m_2-1}(\eta_{p+m_2}^2-\eta_i^2)\tilde\gamma_i
$$

with by (3.8), $\sum(\eta_{p+m_2}-\eta_i)\tilde\gamma_i = 1$. We may again do the change of variables $\bar\gamma_i = (\eta_{p+m_2}-\eta_i)\tilde\gamma_i$ which are non-negative and with sum equal to one by (3.8). We get by (3.8)

$$
\begin{aligned}
L(\tilde\gamma) &= \sum_{i=1}^{p}\alpha_i\log[\frac{\bar\gamma_i}{\alpha_i(\eta_{p+m_2}-\eta_i)}] + \theta_1\sum_{i=1-m_1}^{p+m_2-1}(\eta_{p+m_2}+\eta_i)\bar\gamma_i \\
&= \sum_{i=1}^{p}\alpha_i\log[\frac{\bar\gamma_i}{\alpha_i(\eta_{p+m_2}-\eta_i)}] + \theta_1\sum_{i=1-m_1}^{p+m_2-1}\eta_i\bar\gamma_i + \theta_1\eta_{p+m_2} \\
&= I_{\theta_1,\eta}^{p+m_2-1}(\bar\gamma) + \sum_{i=1}^{p}\alpha_i\log[\frac{1}{(\eta_{p+m_2}-\eta_i)}] + \theta_1\eta_{p+m_2}
\end{aligned}
$$

where $I_{\theta_1,\eta}^{p+m_2-1}$ is defined as in (3.4) with largest outlier η_{p+m_2-1}. Its maximum is $J(\mu,\theta_1,\eta_{p+m_2-1})$. We thus get

$$\max L = \sum_{i=1}^{p} \alpha_i \log[\frac{1}{(\eta_{p+m_2}-\eta_i)}] + \theta_1\eta_{p+m_2} + J(\mu,\theta_1,\eta_{p+m_2-1})$$

We finally compute $\max I(\chi,\tilde{\gamma}) = \max L(\tilde{\gamma}) + \max H(\chi)$.

- For $G^{-1}(\theta_1) \le \eta_{p+m_2-1} \le \eta_{p+m_2}$, we check that $J(\mu,\theta_1,\eta_{p+m_2})$ equals

$$J(\mu,\theta_1,\eta_{p+m_2-1}) + \sum \alpha_i \log \frac{|\eta_i-\eta_{p+m_2-1}|}{|\eta_i-\eta_{p+m_2}|} + \theta_1(\eta_{p+m_2}-\eta_{p+m_2-1}) \tag{3.11}$$

  so that by (3.10) we find

$$\max I = J(\mu,\theta_1,\eta_{p+m_2}) + J(\mu,\theta_2,\eta_{p+m_2-1})\,.$$

- For $G_\mu^{-1}(\theta_1) \in [\eta_{p+m_2-1},\eta_{m+p_2}]$, $J(\mu,\theta_2,G_\mu^{-1}(\theta_1)) = J(\mu,\theta_2,G_\mu^{-1}(\theta_2)) = J(\mu,\theta_2,\eta_{p+m_2-1})$ since $\theta_1>\theta_2$ and $\eta_{p+m_2-1} < G_\mu^{-1}(\theta_1) < G_\mu^{-1}(\theta_2)$. Moreover as $\theta_1 > G_\mu(\eta_{p+m_2})$,

$$\max L = J(\mu,\theta_1,\eta_{p+m_2}) + J(\mu,\theta_1,\eta_{m_2+p-1}) + \log\theta_1 + 1$$

  which again does not depend on η_{m_2+p-1}. Hence

$$\begin{aligned}\max I = J(\mu,\theta_2,\eta_{p+m_2-1}) &+ \sum \alpha_i \log|\eta_i - G_\mu^{-1}(\theta_1)| - \theta_1 G_\mu^{-1}(\theta_1)\\ &+ J(\mu,\theta_1,\eta_{p+m_2}) + J(\mu,\theta_1,G^{-1}(\theta_1)) + \log\theta_1 + 1\\ &= J(\mu,\theta_2,\eta_{p+m_2-1}) + J(\mu,\theta_1,\eta_{p+m_2})\end{aligned}$$

- For $G_\mu^{-1}(\theta_1) > \eta_{m+p_2}$, we compute

$$\begin{aligned}\max I &= J(\mu,\theta_2,\eta_{p+m_2}) + J(\mu,\theta_1,\eta_{p+m_2-1})\\ &= J(\mu,\theta_2,\eta_{p+m_2-1}) + J(\mu,\theta_1,\eta_{p+m_2})\end{aligned}$$

  since $\theta_2 < \theta_1 < G_\mu(\eta_{p+m_2}) < G(\eta_{p+m_2-1})$ so that the above supremum does not depend on the outliers.

In the case where $N_{p+m_2} \ge 2$, we have $\chi = \eta_{p+m_2}$ by Weyl interlacing property and therefore it does not depend on the choice of γ. The result follows immediately after conditioning as in the proof above. □

A similar (but easier) argument shows that

Proposition 3.6. *Let $\theta_1 \ge 0 \ge \theta_2$. Then,*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}(\theta_1\langle e,\mathbf{X}_N e\rangle + \theta_2\langle f,\mathbf{X}_N f\rangle)\Big)\Big]$$

$$= \frac{\beta}{2}(J(\mu,\theta_1,\eta_{p+m_2}) + J(\mu,\theta_2,\eta_{1-m_1}))\,.$$

Here J is extended to negative values of θ_2 by putting

$$J(\mu,\theta_2,\eta_{1-m_1}) = J(\mu_-,-\theta_2,-\eta_{1-m_1})$$

if $\mu_-(x\in .) = \mu(-x\in .)$.

3.3. *Limiting k-d spherical integrals.* We now consider more general k-dimensional spherical integrals with $k \geq 2$. In the sequel we let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ be the limit of the k largest outliers of $\mathbf{X}_N$, and $\lambda_{-1} \leq \cdots \leq \lambda_{-\ell} \leq 0$ be the limit of the ℓ smallest outliers of $\mathbf{X}_N$. We assume they all have multiplicity one (but they can be equal, allowing for outliers with any finite multiplicity). With the previous notations, $\lambda_i = \eta_{p+m_2}$ for $i \in [1, N_{p+m_2}]$, $\lambda_i = \eta_{p+m_2-1}$ for $i \in [N_{p+m_2}+1, N_{p+m_2}+N_{p+m_2-1}]$.

Proposition 3.7. *Fix two integer numbers k and ℓ. Let $(e_1, \dots e_k, e_{-1}, \dots e_{-\ell})$ be $k+\ell$ orthonormal vectors following the uniform law in the sphere and assume that the sequence $\mathbf{X}_N$ has the form described at the beginning of this section. Let $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_k \geq 0 \geq \theta_{-\ell} \geq \cdots \geq \theta_{-1}$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{i=-\ell, i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]$$

$$= \frac{\beta}{2}\Big(\sum_{i=1}^{k} J(\mu, \theta_i, \lambda_i) + \sum_{i=1}^{\ell} J(\mu, \theta_{-i}, \lambda_{-i})\Big).$$

Proof: For the sake of simplicity we will assume the outliers $\lambda_1, ..., \lambda_k$ and $\lambda_{-1}, ..., \lambda_{-\ell}$ are distinct. The general case can be deduced by equi-continuity of the spherical integral. We will prove this proposition by induction over $k+\ell$. We know it is true for $k+\ell \leq 2$ by the previous section. By symmetry we can assume the proposition true for $(\ell, k-1)$ and it is enough to show it still holds for (ℓ, k). Thus we set $\eta_{p+m_2-i+1} = \lambda_i$ for $i \in [1, k]$, and $\eta_{-m_1+i} = \lambda_{-i}$ and we assume $N_{p+m_2-i} = 1$ for $i \in [1, k]$ and $N_{-m_1+i} = 1$ for $i \in [1, \ell]$. We proceed as in the 2-dimensional case by conditioning on the vector e_1 and so we have :

$$\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{i=-\ell, i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big] =$$

$$\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \theta_1 \langle e_1, \mathbf{X}_N e_1\rangle\Big) \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{\substack{i=-\ell, i\neq 0 \\ i\neq 1}}^{k} \theta_i \langle e_i, \mathbf{X}_N^{(e_1)} e_i\rangle\Big)\Big| e_1\Big]\Big]$$

As previously, the eigenvalues of $\mathbf{X}_N^{(e_1)}$ (seen as a $N-1 \times N-1$ matrix) are interlaced with those of $\mathbf{X}_N$. Thus if we denote χ_j the j-th largest eigenvalue of $\mathbf{X}_N^{(e_1)}$, $\chi_j = \chi_j(\gamma^N(e))$ where $\chi_j(x)$ is the unique solution in the interval $[\eta_{p+m_2-j}, \eta_{p+m_2-j+1}]$ of the equation :

$$\sum_{i=1-m_1}^{p+m_2} \frac{x_i}{\chi_j(x) - \eta_i} = 0 \tag{3.12}$$

for $j \in [1, k-1]$. $\gamma^N(e)$ is defined as in (3.6) The same equation holds for the ℓ smallest eigenvalues below the bulk : if we denote χ_{-j} the j-th smallest eigenvalue of $\mathbf{X}_N^{(e)}$, it is solution of the same equation in $[\eta_{-m_1+j}, \eta_{-m_1+j+1}]$. Observe that unless $\gamma_i^N(e)$ vanishes, $\chi_i(\gamma^N(e))$ can not be equal to η_i. So, if we denote for $i = -m_1+l+1, ..., p+m_2-k$, δ_i the solution of the same interlacing equation in $[\eta_i, \eta_{i+1}]$, up to diagonalisation, $\mathbf{X}_N^{(e_1)}$ has the following form :

$$\mathbf{X}_N^{(e_1)} = diag(\chi_{-1}, ..., \chi_{-l}, \underbrace{\eta_{-m_1+l+1}}_{N_{-m_1+l+1}-1}, \delta_{-m_1+l+1}, \underbrace{\eta_{-m_1+l+2}}_{N_{-m_1+l+2}-1}, ..., \underbrace{\eta_{p+m_2-k}}_{N_{p+m_2-k}-1}, \delta_{p+m_2-k}, \chi_{k-1}, ..., \chi_1)$$

where the δ_j and the χ_i are continuous functions of $\gamma(e)$. We deduce by induction and using the continuity in Remark 3.1 that on the event where $\gamma^N(e_1)$ converges toward κ :

$$\lim_{N\to\infty}\frac{1}{N-1}\log\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-\ell,i\neq 0,1}^{k}\theta_i\langle e_i,\mathbf{X}_N^{(e_1)}e_i\rangle\Big)\Big|e_1\Big]=\frac{\beta}{2}\sum_{i=-\ell,i\neq 0,1}^{k}J(\mu,\theta_i,\chi_i(\kappa))$$

Then, using again that γ^N satisfies a large deviations principle we can deduce from Varadhan's Lemma that :

$$\begin{aligned}&\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-\ell,i\neq 0}^{k}\theta_i\langle e_i,\mathbf{X}_N e_i\rangle\Big)\Big] \qquad (3.13)\\ &=\frac{\beta}{2}\sup_{\gamma\in(\mathbb{R}^+)^{p+m_1+m_2},\sum\gamma_i=1}\Big\{\sum_{i=-\ell}^{-1}J(\mu,\theta_i,\chi_i(\gamma))+\sum_{i=2}^{k}J(\mu,\theta_i,\chi_{i-1}(\gamma))\\ &\qquad+\sum_{i=1}^{p}\alpha_i\log\frac{\gamma_i}{\alpha_i}+\sum_{i=1-m_1}^{p+m_2}\theta_1\eta_i\gamma_i\Big\}\end{aligned}$$

By continuity, taking this supremum only on the set of γ summing up to 1 and such that $\gamma_i>0$ does not change its value. Notice that for such γ we have for all i and j $\chi_i(\gamma)\neq\eta_j$. We set $I_{-j}=]\eta_{-m_1+j},\eta_{-m_1+j+1}[$ for $j=1,...,\ell$ and $I_j=]\eta_{m_2+p-j},\eta_{m_2+p-j+1}[$ for $j=1,...,k-1$ and we define :

$$D=\Big\{(\chi,\gamma)\in\prod_{j=-\ell,j\neq 0}^{k-1}I_j\times(\mathbb{R}^{+,*})^{m_1+m_2+p}:\sum_i\gamma_i=1,\sum_i\frac{\gamma_i}{\chi_j-\eta_i}=0,\forall j\in[-\ell,k-1]\setminus\{0\}\Big\}$$

Therefore we have :

$$\begin{aligned}&\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-\ell,i\neq 0}^{k}\theta_i\langle e_i,\mathbf{X}_N e_i\rangle\Big)\Big] \qquad (3.14)\\ &=\frac{\beta}{2}\sup_{(\chi,\gamma)\in D}\Big\{\sum_{i=-\ell}^{-1}J(\mu,\theta_i,\chi_i)+\sum_{i=2}^{k}J(\mu,\theta_i,\chi_{i-1})+\sum_{i=1}^{p}\alpha_i\log\frac{\gamma_i}{\alpha_i}+\sum_{i=1-m_1}^{p+m_2}\theta_1\eta_i\gamma_i\Big\}\end{aligned}$$

We next show that the above right hand side matches the announced formula. For $i=-m_1+\ell,...,m_2+p-k$, we set :

$$\bar{\gamma}_i=\frac{\prod_{j=1}^{\ell}(\eta_{-m_1+j}-\eta_i)\prod_{j=1}^{k-1}(\eta_{m_2+p-j+1}-\eta_i)}{\prod_{j=1}^{\ell}(\chi_{-j}-\eta_i)\prod_{j=1}^{k-1}(\chi_j-\eta_i)}\gamma_i$$

We have that if $\gamma_i>0$ for all i then $\bar{\gamma}_i>0$ for all i and $\bar{\gamma}_i$ vanishes at the outliers. We want to prove that this definition provides a one to one correspondance between the set D of parameters (χ,γ) and the set $\bar{D}$ parameters $(\chi,\bar{\gamma})$ defined as follows :

$$\bar{D}=\Big\{(\chi,\bar{\gamma})\in\prod_{j=-\ell,j\neq 0}^{k-1}I_j\times(\mathbb{R}^{+,*})^{m_1+m_2+p-k-\ell+1},\sum_i\bar{\gamma}_i=1\Big\}$$

Note that $\bar{\gamma}$ lives a priori in a set of $k+\ell-1$ dimension smaller but γ was satisfying $k+\ell-1$ additional equations. First, let us prove that if $(\chi,\gamma)\in D$ the $\bar{\gamma}_i's$ sum up to 1. We let for a real number X, F to be the rational function

$$F(X)=\frac{\prod_{j=1}^{\ell}(\eta_{-m_1+j}-X)\prod_{j=1}^{k-1}(\eta_{m_2+p-j+1}-X)}{\prod_{j=1}^{\ell}(\chi_{-j}-X)\prod_{j=1}^{k-1}(\chi_j-X)}$$

so that $\bar{\gamma}_i=F(\eta_i)\gamma_i$ for $i\in[-m_1+\ell+1,m_2+p-k+1]$ and $F(\eta_i)=0$ for the other values of i. Let us decompose F in partial fractions : as it goes to one at infinity, we find

$$F(X) = 1 + \sum_{j=-\ell, j\neq 0}^{k-1} \frac{a_j}{\chi_j - X}$$

for some real numbers a_j. Then, since $F(\eta_i) = 0$ for $i \neq -m_1 + 1 + \ell, ..., m_2 + p - k + 1$ we have :

$$\begin{aligned} \sum_{i=-m_1+\ell+1}^{m_2+p-k+1} \bar{\gamma}_i &= \sum_{i=-m_1+1}^{m_2+p} F(\eta_i)\gamma_i \\ &= \sum_{i=-m_1+1}^{m_2+p} \gamma_i + \sum_{j=-\ell, j\neq 0}^{k-1} a_j \sum_{i=-m_1+1}^{m_2+p} \frac{\gamma_i}{\chi_j - \eta_i} \\ &= 1 \end{aligned}$$

where we used the interlacing relations. Therefore, since when χ is fixed the function $\gamma \mapsto \bar{\gamma}$ is an affine map between the affine subspace E of $\mathbb{R}^{p+m_1+m_2}$ defined by the $k + \ell - 1$ interlacing relations and the condition of sum one and the affine subspace F of $\mathbb{R}^{p+m_1+m_2-k-\ell+1}$ defined by the condition of sum one. Since these spaces have the same dimension, to conclude we only need to prove that this map is injective and that for all $\gamma \in E$, $\bar{\gamma}_i > 0$ for all i implies $\gamma_i > 0$ for all i. To prove injectivity first notice that $\gamma_i = F(\eta_i)^{-1}\bar{\gamma}_i$ for $i \in [-m_1 + \ell + 1, m_2 + p - k + 1]$. We next show how to reconstruct γ_i for $i \in [-m_1 + \ell + 1, m_2 + p - k + 1]^c$. To this end, for $j = 1, ..., k-1$, we let $G_j(X) = \frac{F(X)}{(\eta_{p+m_2-j+1} - X)}$ and for $j = 1, ..., \ell$, we let $G_{-j}(X) = \frac{F(X)}{(\eta_{-m_1+j} - X)}$. Let us first reconstruct γ_i for $i \in [m_2 + p - k + 2, m_2 + p]$ (the case where $i \in [-m_1 + 1, -m_1 + l]$ is similar). Then again decomposing G_j in partial fractions, we have

$$G_j(X) = \sum_{j'=-\ell, j'\neq 0}^{k-1} \frac{b_{j'}}{\chi_{j'} - X}$$

for some real numbers b_j. Again by the interlacing relations

$$\sum_{i=-m_1+1}^{p+m_2} G_j(\eta_i)\gamma_i = 0\,.$$

But we can also write :

$$\sum_{i=-m_1+1}^{p+m_2} G_j(\eta_i)\gamma_i = \sum_{i=-m_1+\ell}^{p+m_2-k} \frac{\bar{\gamma}_i}{\eta_{p+m_2-j+1} - \eta_i} + G_j(\eta_{p+m_2-j+1})\gamma_{p+m_2-j+1}$$

so that we deduce for $j = 1, ..., k-1$

$$\gamma_{p+m_2-j+1} = -\Big(\sum_{i=-m_1=\ell}^{p+m_2-k+1} \frac{\bar{\gamma}_i}{\eta_{p+m_2-j+1} - \eta_i}\Big)/G_j(\eta_{p+m_2-j+1})\,.$$

As a consequence, the map $\gamma \mapsto \bar{\gamma}$ is injective. Furthermore if $\bar{\gamma}_i > 0$ for all i, then $\gamma_j > 0$ since $G_j(\eta_j) < 0$. The same remains true for j negative. Therefore we have that the change of variables from $(\chi, \gamma) \in D$ to $(\chi, \bar{\gamma}) \in \bar{D}$ is one to one. But before changing variables, let us compare $\sum \eta_i\gamma_i$ and $\sum \eta_i\bar{\gamma}_i$. We use the following decomposition :

$$XF(X) = X + S + \sum_{j=-\ell, j\neq 0}^{k-1} \frac{c_j}{\chi_j - X}$$

for some real numbers c_j and where

$$S = \sum_{j=-\ell, j\neq 0}^{k-1} \chi_j - \sum_{j=1}^{\ell} \eta_{-m_1+j} - \sum_{j=1}^{k-1} \eta_{m_2+p-j+1}\,.$$

We deduce that

$$\sum_{-m_1+l}^{m_2+p-k+1} \eta_i \bar{\gamma}_i = \sum_{-m_1+1}^{m_2+p} \eta_i F(\eta_i)\gamma_i = \sum_{-m_1+1}^{m_2+p} \eta_i \gamma_i + S$$

where we used again the interlacing relationships and the fact that the γ_i's sum up to 1. Coming back to (3.13), we have to take the supremum of the following function I for $(\chi, \bar{\gamma}) \in \bar{D}$:

$$\begin{aligned} I(\bar{\gamma}, \chi) = & \sum_{j=1}^{k-1} \Big[J(\mu, \theta_{j+1}, \chi_j) + \sum_{i=1}^{p} \alpha_i \log|\chi_j - \eta_i| - \sum_{i=1}^{p} \alpha_i \log|\eta_{m_2+p-j+1} - \eta_i| \Big] \\ & + \sum_{j=-\ell}^{-1} \Big[J(\mu, \theta_j, \chi_j) + \sum_{i=1}^{p} \alpha_i \log|\chi_j - \eta_i| - \sum_{i=1}^{p} \alpha_i \log|\eta_{-m_1-j} - \eta_i| \Big] \\ & + \sum_{i=1}^{p} \alpha_i \log \frac{\bar{\gamma}_i}{\alpha_i} + \theta_1 \Big(\sum_{i=-m_1+1}^{p+m_2} \eta_i \bar{\gamma}_i - \sum_{j=-\ell, j\neq 0}^{k-1} \chi_j + \sum_{j=1}^{\ell} \eta_{-m_1+j} + \sum_{j=1}^{k-1} \eta_{m_2+p-j+1} \Big) \end{aligned}$$

Therefore we have :

$$\begin{aligned} I(\bar{\gamma}, \chi) = & \sum_{i=1}^{k-1} H(\chi_i, \theta_{i+1}) + \sum_{i=-\ell}^{-1} H(\chi_i, \theta_i) + \sum_{i=1}^{p} \alpha_i \log \frac{\bar{\gamma}_i}{\alpha_i} + \theta_1 \sum_{i=m_1-\ell}^{p+m_2-k+1} \eta_i \bar{\gamma}_i \\ & - \sum_{j=1}^{k-1} \sum_{i=1}^{p} \alpha_i \log|\eta_{m_2+p-j+1} - \eta_i| - \sum_{j=-\ell}^{-1} \sum_{i=1}^{p} \alpha_i \log|\eta_{-m_1-j} - \eta_i| \\ & + \theta_1 \sum_{j=1}^{\ell} \eta_{-m_1+j} + \theta_1 \sum_{j=1}^{k-1} \eta_{m_2+p-j+1} \end{aligned}$$

where we set :

$$H(\chi, \theta) = J(\mu, \theta, \chi) + \sum_{i=1}^{p} \alpha_i \log|\chi - \eta_i| - \chi\theta_1 .$$

The supremum over $\bar{\gamma}$ and χ are now decoupled and the χ_i belongs to $]\eta_{-m_1+i}, \eta_{-m_1+i+1}[$ if $i \in [-\ell, -1]$ and $]\eta_{p+m_2-i+1}, \eta_{p+m_2-i+2}[$ if $i \in [1, k]$. As in the two-dimensional case we can compute for $i = 2, ..., k$,

$$\sup_{\chi \in]\eta_{p+m_2-i+1}, \eta_{p+m_2-i+2}[} H(\chi, \theta_i) = \begin{cases} H(\eta_{p+m_2-i+1}, \theta_i) \text{ if } G_\mu^{-1}(\theta_1) \leq \eta_{p+m_2-i+1}, \\ H(G_\mu^{-1}(\theta_1), \theta_i) \text{ if } G_\mu^{-1}(\theta_1) \in [\eta_{p+m_2-i+1}, \eta_{m+p_2-i+2}], \\ H(\eta_{p+m_2-i+2}, \theta_i) \text{ if } G_\mu^{-1}(\theta_1) > \eta_{p+m_2-i+2}. \end{cases} .$$

Moreover, for $i = 1, ..., \ell$, $H(\chi, \theta_{-i})$ is a decreasing function of χ since θ_{-i} is negative and so

$$\sup_{\chi \in]\eta_{-m_1+i}, \eta_{-m_1+i+1}[} H(\chi, \theta_{-i}) = H(\eta_{-m_1+i}, \theta_{-i}) .$$

It remains to optimize the sum of the third and fourth term in $I(\bar{\gamma}, \chi)$. But this sum is equal to $I^{p+m_2-k+1}_{\theta_1, \eta_{p+m_2-k+1}}(\bar{\gamma})$, see (3.4). Thus, taking the supremum for $\bar{\gamma}_i > 0$ and $\sum \bar{\gamma}_i = 1$ gives $J(\mu, \theta_1, \eta_{p+m_2-k+1})$.

To conclude, we need to look at the position of $G_\mu^{-1}(\theta_1)$ relatively to the k largest outliers. Let us denote $H_i = \max H(., \theta_i)$ for $i < 0$ and $H_i = \max H(., \theta_{i+1})$ for $i > 0$. We have

$$\begin{aligned}\sup I = &\sum_{j=1}^{k-1}\Big(H_j - \sum \alpha_i \log|\eta_i - \eta_{p+m_2-j+1}| + \theta_1 \eta_{p+m_2-j+1}\Big) \\ &+ \sum_{j=1}^{\ell}\Big(H_{-j} - \sum \alpha_i \log|\eta_i - \eta_{-m_1+j}| + \theta_1 \eta_{-m_1+j}\Big) \\ &+ J(\mu, \theta_1, \eta_{p+m_2-k+1})\end{aligned}$$

Here we will treat the case where $G_\mu^{-1}(\theta_1) \in [\eta_{p+m_2-k+1}, \eta_{p+m_2}]$ which is the most complex one. First of all, since for $j = 1, ..., \ell$ $H_{-j} = H(\eta_{-m_1+j}, \theta_{-j})$, we have that in the second sum, the term of index j is indeed equal to $J(\mu, \theta_{-j}, \eta_{-m_1+j})$. If j' is the index such that, $G_\mu^{-1}(\theta_1) \in [\eta_{p+m_2-j'}, \eta_{p+m_2-j'+1}]$ then for $j < j'$, $H_j = H(\eta_{p+m_2-j}, \theta_{j+1})$ and the term of index j of the first sum is :

$$J(\mu, \theta_{j+1}, \eta_{p+m_2-j}) + \sum \alpha_i \log \frac{|\eta_{p+m_2-j} - \eta_i|}{|\eta_{p+m_2-j+1} - \eta_i|} + \theta_1(\eta_{p+m_2-j} - \eta_{p+m_2-j+1})$$

The term of index j' is equal to :

$$J(\mu, \theta_{j+1}, G_\mu^{-1}(\theta_1)) + \sum_i \alpha_i \log \frac{|G_\mu^{-1}(\theta_1) - \eta_i|}{|\eta_{p+m_1-j'+1} - \eta_i|} + \theta_1(\eta_{p+m_2-j'} - G_\mu^{-1}(\theta_1))$$

And the terms $j > j'$ are equal to $J(\mu, \theta_{j+1}, \eta_{p+m_2-j+1})$. Since $\theta_{j+1} \le \theta_1$, $G_\mu^{-1}(\theta_{j+1}) \ge G_\mu^{-1}(\theta_1)$, so we have that for $j > j'$

$$J(\mu, \theta_{j+1}, \eta_{p+m_2-j+1}) = J(\mu, \theta_{j+1}, \eta_{p+m_2-j}) \text{ and } J(\mu, \theta_{j'+1}, G_\mu^{-1}(\theta_1)) = J(\mu, \theta_{j'+1}, \eta_{p+m_2-j'})$$

Therefore the whole sum can be simplified as follows :

$$\begin{aligned}\max I = &\sum_{j=1}^{k-1} J(\mu, \theta_{j+1}, \eta_{p+m_2-j}) + \sum_{j=1}^{\ell} J(\mu, \theta_{-j}, \eta_{-m_1+j}) + \sum_i \alpha_i \log \frac{|G_\mu^{-1}(\theta_1) - \eta_i|}{|\eta_{p+m_2} - \eta_i|} \\ &+ \theta_1(\eta_{p+m_2} - G_\mu^{-1}(\theta_1)) + J(\mu, \theta_1, \eta_{p+m_2-k+1})\end{aligned}$$

Then we notice that $J(\mu, \theta_1, \eta_{p+m_2-k+1}) = J(\mu, \theta_1, G_\mu^{-1}(\theta_1))$ and conclude since :

$$J(\mu, \theta_1, \eta_{p+m_2}) = J(\mu, \theta_1, G_\mu^{-1}(\theta_1)) + \sum_i \alpha_i \log \frac{|G_\mu^{-1}(\theta_1) - \eta_i|}{|\eta_{p+m_2} - \eta_i|} + \theta_1(\eta_{p+m_2} - G_\mu^{-1}(\theta_1)).$$

□

4. Diffuse spectrum

We consider in this section the general case where $\mathbf{X}_N$ is a Hermitian matrix such that

$$\hat{\mu}_{\mathbf{X}_N} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}$$

converges towards a probability measure μ with support with rightmost point r_μ and leftmost point l_μ which are assumed to be finite. Let $\lambda_1^N \ge \lambda_2^N \ge \cdots \ge \lambda_k^N \ge r_\mu$ be the k largest outliers of $\mathbf{X}_N$, $\lambda_N^N \le \cdots \le \lambda_{N-\ell+1}^N \le l_\mu$ be the smallest outliers of $\mathbf{X}_N$. They all are assumed to have multiplicity one (but are possibly equal). Assume that there exists real numbers $\lambda_1 \ge \lambda_2 \cdots \ge \lambda_k \ge r_\mu > l_\mu \ge \lambda_{-\ell} \ge \lambda_{1-\ell} \ge \cdots \ge \lambda_{-1}$ such that

$$\lim_{N\to\infty} \lambda_i^N = \lambda_i \text{ for } i \in [1, k], \lim_{N\to\infty} \lambda_{N-i+1}^N = \lambda_{-i} \text{ for } i \in [1, \ell]\,. \tag{4.1}$$

We are going to prove Proposition 2.1.

Proposition 4.1. *Fix two integer numbers k, ℓ and assume that $(\mathbf{X}_N)$ is a sequence of matrices whose empirical measure $\hat{\mu}_{\mathbf{X}_N}$ converges weakly towards some compactly supported measure μ on $\mathbb{R}$ and the convergence of its outliers (4.1). Let $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_k \geq 0 \geq \theta_{-\ell} \geq \cdots \geq \theta_{-1}$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{i=-\ell, i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N e_i\rangle \Big)\Big]$$

$$= \frac{\beta}{2}\Big(\sum_{i=1}^{k} J(\mu, \theta_i, \lambda_i) + \sum_{i=1}^{\ell} J(\mu, \theta_{-i}, \lambda_{-i})\Big) .$$

Observe that if all the θ_i's are non-negative (resp. non-positive), we do not need to assume the convergence of the smallest (resp. largest) eigenvalues of $\mathbf{X}_N$ to derive the convergence of the spherical integral but our argument requires that they are uniformly bounded.

Proof: We first remark that we can assume $\mathbf{X}_N$ diagonal without loss of generality. In a first step, we assume that the partition function $F_\mu(x) = \mu((-\infty, x])$ of μ is continuous. We fix $\varepsilon > 0$. We let $\mathbf{X}_N^\varepsilon$ be the diagonal matrix with eigenvalues $k+\ell$ eigenvalues equal to $(\lambda_1 \geq \lambda_2 \geq \cdots \lambda_k \geq \lambda_{-\ell} \geq \cdots \lambda_{-1})$ and $N-k-\ell$ other eigenvalues given for $i \in [k+1, N-\ell]$ by

$$\lambda_i^{N,\varepsilon} = \lfloor \varepsilon^{-1}(\lambda_i^N - \lambda_{-\ell}) \rfloor \varepsilon + \lambda_{-\ell} .$$

Then the eigenvalue $\lambda_{-\ell}+j\varepsilon$ has multiplicity $n_j^{\varepsilon,N} = \hat{\mu}_{\mathbf{X}_N}([\lambda_{-\ell}+j\varepsilon, \lambda_{-\ell}+(j+1)\varepsilon))$ for $\mathbf{X}_N^\varepsilon$ (up to a finite correction bounded by k when j is close to N). $n_j^{\varepsilon,N}/N$ converges towards $\mu([\lambda_{-\ell}+j\varepsilon, \lambda_{-\ell}+(j+1)\varepsilon])$ since we assumed the partition function of μ to be continuous. Hence, $\mathbf{X}_N^\varepsilon$ satisfies the hypotheses of Proposition 3.7. Moreover, by definition, we know that for N large enough (so that the outliers are at distance smaller than ε from their limit) the spectral norm $\|\mathbf{X}_N - \mathbf{X}_N^\varepsilon\|$ of $\mathbf{X}_N - \mathbf{X}_N^\varepsilon$ is bounded from above by ε Therefore,

$$\Big| \frac{1}{N} \log \frac{\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-\ell,i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]}{\mathbb{E}\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-\ell,i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N^\varepsilon e_i\rangle\Big)\Big]} \Big| \leq \frac{\beta}{2} \sum |\theta_i| \varepsilon .$$

On the other hand, Proposition 3.7 implies

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{i=-\ell, i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N^\varepsilon e_i\rangle \Big)\Big] = \frac{\beta}{2}\Big(\sum_{i=-\ell, i\neq 0}^{k} J(\mu^\varepsilon, \theta_i, \lambda_i)\Big)$$

with $\mu^\varepsilon = \sum_j \mu([\lambda_{-\ell}+j\varepsilon, \lambda_{-\ell}+(j+1)\varepsilon])\delta_{\lambda_{-\ell}+j\varepsilon}$. By continuity of $\mu \to J(\mu, \theta_i, \lambda_i)$, see Maïda (2007) or the Appendix (Theorem 6.1), and the weak convergence of μ^ε towards μ, the conclusion follows.

Finally, to remove the condition that μ has a continuous partition function we note that we can always add a small matrix to $\mathbf{X}_N$ and its contribution will go to zero as its norm goes to zero after N goes to infinity. We again assume $\mathbf{X}_N$ diagonal and replace it by the diagonal matrix with the same outliers and in the bulk the entries are added independent uniform variables with uniform distribution on $[0, \varepsilon]$. Again $\mathbf{X}_N^\varepsilon - \mathbf{X}_N$ has norm bounded by ε. Moreover, the spectral measure of $\mathbf{X}_N^\varepsilon$ converges towards $\mu * 1_{[0,\varepsilon]} du/\varepsilon$ whose partition function is continuous. Hence, we can apply our result to this new matrix and then let ε go to zero to conclude. □

5. Applications to large deviations for the extreme eigenvalues of random matrices

5.1. *Universality of the large deviations for the $k+\ell$ extreme eigenvalues of Wigner matrices with sharp sub-Gaussian entries.* In this section, we prove Theorem 2.3 and to simplify take without loss of generality $\ell = k$. The proof follows the ideas of Guionnet and Husson (2020) quite closely: we simply sketch the main arguments and changes. First note that it is enough to prove a weak large deviations principle thanks to our assumption which insures that exponential tightness holds. Moreover. let $\bar\lambda^N = (\lambda_1, \dots, \lambda_k, \lambda_{N-k+1}, \dots, \lambda_N)$ be the $2k$ extreme eigenvalues of $\mathbf{X}_N$. To get a weak large deviations upper bound, we proceed as in Guionnet and Husson (2020, Corollary 1.16) and we tilt the measure by spherical integrals as above : if $(e_i)_{-k\le i\le k}$ follows the uniform law on the set of $2k$ orthonormal vectors on the sphere ($\theta_0 = 0$ and $e_0 = 0$ is added to shorten the notations), θ_i are real numbers of the same sign than $i \in [-k, k]$ to be chosen later, we write

$$\begin{aligned}\mathbb{P}\Big(\|\bar\lambda^N - \bar x\|_2 \le \varepsilon\Big) &\le \mathbb{E}_{\mathbf{X}_N}\Big[1_{\|\bar\lambda^N-\bar x\|_2\le\varepsilon}\frac{\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k\theta_i\langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]}{\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k\theta_i\langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]}\Big]\\ &\le e^{-N\frac{\beta}{2}F(\bar x,\bar\theta)+o(\varepsilon)N}\mathbb{E}_{\mathbf{X}_N}\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k\theta_i\langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\end{aligned}$$

where

$$F(\bar x, \bar\theta) = \sum_{i=-k}^k J(\sigma, \theta_i, x_i)\,.$$

We used in the second line that by Theorem 6.2, the spherical integrals are uniformly continuous and are asymptotically given by $F(\bar x, \bar\theta)$, and our assumption that the spectral measure of $\mathbf{X}_N$ converges towards the semi-circle law σ faster than any exponential. Here $o(\varepsilon)$ goes to zero when ε does. We also used the bound

$$\frac{\mathbb{E}_{\mathbf{X}_N}\Big[1_{\|\bar\lambda^N-\bar x\|_2\le\varepsilon}\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k\theta_i\langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big]}{\mathbb{E}_{\mathbf{X}_N}\Big[\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k\theta_i\langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big]} \le 1 \tag{5.1}$$

We next compute the expectation of the spherical integral by using that our entries are sharp sub-Gaussian as in the proof of Guionnet and Husson (2020, Lemma 3.2) :

$$\begin{aligned}&\mathbb{E}_{\mathbf{X}_N}\Big[\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k\theta_i\langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big]\\ &\le \mathbb{E}_e[\exp\{\frac{\beta}{4}N\sum_{i=-k}^k\sum_{k\le j}2^{1_{k\neq j}}|\sum\theta_i e_i(k)e_i(j)|^2\}] = \exp\{\frac{\beta}{4}\sum_{j=-k}^k\theta_j^2\}\end{aligned}$$

We hence get the upper bound

$$\limsup_{\varepsilon\to0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big(\|\bar\lambda^N - \bar x\|_2 \le \varepsilon\Big) \le -\frac{\beta}{2}\sup_{\theta_i}\{\sum_{j=-k}^k\frac{\theta_j^2}{2} - F(\bar x, \bar\theta)\}$$

where we take the supremum over non-negative θ_i for $i \in [1, k]$ and non-positive θ_i's for $i \in [-k, -1]$. Finally we observe that the supremum decouples and recall from Guionnet and Husson (2020, Section 4.1) that the supremum over each θ_i of $\theta_i^2/2 - J(\sigma, \theta_i, x_i)$ gives $\int_2^{|x_i|}\sqrt{t^2-4}dt$. To get the lower bound, we need to show that there exists $\bar\theta = (\theta_{-k}, \dots, \theta_{-1}, \theta_1, \dots, \theta_k)$ such that (5.1) is almost an equality in the sense that for every $\varepsilon > 0$

$$\liminf_{\mathbb{N}\to\infty} \frac{1}{N} \log \frac{\mathbb{E}_{\mathbf{X}_N}\Big[1_{\|\bar\lambda^N - \bar x\|_2 \le \varepsilon} \mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big]}{\mathbb{E}_{\mathbf{X}_N}\Big[\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big]} \ge 0 \tag{5.2}$$

and

$$\liminf_{\mathbb{N}\to\infty} \frac{1}{N} \log \mathbb{E}_{\mathbf{X}_N}\Big[\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big] \ge \frac{\beta}{2}\sum_{j=-k}^k \theta_j^2 \,. \tag{5.3}$$

In both cases we use the fact that under the uniform measure, the vectors e_i are delocalised with overwhelming probability, namely if $V_N^\kappa = \cap_{1\le i\le k}\{\|e_i\|_\infty \le N^{-1/4-\kappa}\}$ then $\mathbb{P}(V_N^\kappa)$ goes to one for any $\kappa \in (0, 1/4)$. Therefore, to prove (5.3) we notice that

$$\begin{aligned}
&\mathbb{E}_{\mathbf{X}_N}\Big[\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]\Big] \\
&\qquad \ge \mathbb{E}_e\Big[1_{e\in V_N^\kappa} \prod_{i\le j} \mathbb{E}[\exp\{\frac{\beta}{2} N 2^{1_{i\ne j}} \sum_r \theta_r \Re(e_r(i)\bar e_r(j) X_{ij})\}]\Big] \\
&\qquad\qquad \ge \exp\{N\frac{\beta}{2}\sum_{r=-k}^k \theta_j^2 + O(N^{1-2\kappa})\}\mathbb{P}(V_N^\kappa)
\end{aligned}$$

where we used that $\sum_r \theta_r e_r(i)\bar e_r(j)$ is of order at most $N^{-1/2-2\kappa}$ on V_N^κ so that we can expand the Laplace transform of the entries around the origin. This proves (5.3). To prove (5.2) we notice that it is enough to show that for N large enough

$$\inf_{\bar e\in V_N^\varepsilon} \frac{\mathbb{E}_{\mathbf{X}_N}\Big[1_{\|\bar\lambda^N - \bar x\|_2 \le \varepsilon} \exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]}{\mathbb{E}_{\mathbf{X}_N}\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)\Big]} \ge \frac{1}{2}\,.$$

But under the law tilted by $\exp\Big(\frac{\beta N}{2}\sum_{i=-k}^k \theta_i \langle e_i, \mathbf{X}_N e_i\rangle\Big)$, $\mathbf{X}_N$ still has independent entries. We can compute their mean and covariance under the tilted law and using again that $\sum_r \theta_r e_r(i)\bar e_r(j)$ is of order at most $N^{-1/2-2\kappa}$, we see that its mean is $\sum \theta_i e_i e_i^*$ and its covariance is close to $1/N$. We deduce as in Guionnet and Husson (2020, Subsection 5.1) and the BBP transition Baik et al. (2005); Pizzo et al. (2013) that under this tilted law the outliers of $\mathbf{X}_N$ are given by $\theta_i + \theta_i^{-1}$: it is therefore sufficient to choose $\theta_i = \frac{1}{2}(x_i \pm \sqrt{x_i^2 - 4})$. We refer the reader to Guionnet and Husson (2020) for more details.

5.2. *Universality of the large deviations for the k largest eigenvalues of Wishart matrices with sharp sub-Gaussian entries.* We here prove Theorem 2.4 and, as in the previous subsection, we will only sketch the changes from the proof in Guionnet and Husson (2020). It is enough to study the largest eigenvalues of the linearized matrix $\mathbf{Y}_N$ of the matrix $N^{-1}\mathbf{G}_{L,M}\mathbf{G}_{L,M}^*$:

$$\mathbf{Y}_N = \begin{pmatrix} 0_{L\times L} & \frac{1}{\sqrt N}\mathbf{G}_{L,M} \\ \frac{1}{\sqrt N}\mathbf{G}_{L,M}^* & 0_{M\times M} \end{pmatrix}$$

Up to a factor $(N/L)^{1/2} = ((1+\alpha) + o(N^{-\kappa}))^{1/2}$, $\mathbf{Y}_N$ is the linearization of $\mathbf{W}_{L,M}$. The main difference with the proof for Wigner matrices will be that computing the asymptotics of the annealed spherical integral requires more skill as it depends on the large deviations for the scalar products of projections of vectors uniformly distributed on the sphere: we can not merely assume that they are delocalized since this could a priori change the large deviations weight. To be more precise, let Λ_N be the annealed spherical integral given for $\bar\theta = (\theta_1, \dots, \theta_k) \in (\mathbb{R}^+)^k$ by

$$\Lambda_N(\bar\theta) = \frac{1}{N}\log \mathbb{E}_{\mathbf{Y}_N}\Big[\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=1}^k \theta_i\langle e_i, \mathbf{Y}_N e_i\rangle\Big)\Big]\Big].$$

We shall prove that

$$\lim_{N\to\infty}\frac{1}{N}\log\Lambda_N(\bar\theta) = \Lambda(\bar\theta) = \sum_{i=1}^k \Lambda(\theta_i) \tag{5.4}$$

with, if $\alpha' = (1+\alpha)^{-1}$, and with α the limit of M/N,

$$\Lambda(\theta) = \sup_{a\in]0,1[}\Big(\theta^2 a(1-a) + \alpha'\log\frac{a}{\alpha'} + (1-\alpha')\log\frac{1-a}{1-\alpha'}\Big).$$

The above supremum is achieved at $x_{\theta,\alpha}$, as defined in Lemma 3.4 of Guionnet and Husson (2020). We first prove the upper bound in (5.4).

$$\begin{aligned} e^{N\Lambda_N(\bar\theta)} &= \mathbb{E}_e\mathbb{E}_{\mathbf{Y}}\Big[\Big[\exp\Big(\frac{\beta N}{2}\sum_{i=1}^k\theta_i\langle e_i,\mathbf{Y}_N e_i\rangle\Big)\Big]\Big] \\ &= \mathbb{E}_e\mathbb{E}_{\mathbf{Y}}\Big[\Big[\exp\Big(\frac{\beta}{2}\sum_{i=1}^k\sum_{\substack{l=1,\dots,L\\ m=L+1,\dots,N}}\theta_i\sqrt{N}\Re(e_i(l)\bar e_i(m)X_{l,m})\Big)\Big]\Big] \\ &\le \mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{4}\sum_{\substack{l=1,\dots,L\\ m=L+1,\dots,N}}\sum_{i,j=1}^k\theta_i\theta_j\Re(e_i(l)\bar e_i(m)\bar e_j(l)e_j(m))\Big)\Big]\end{aligned}$$

where we used that the entries are sharp sub-Gaussian. Now, let us call $e^{(1)}$ the vector of $\mathbb{C}^L$ whose coordinates are the L first coordinates of e and $e^{(2)}$ the vector of $\mathbb{C}^M$ whose coordinates are the M last of e. If we let $\psi^{(p)}_{l,m} = \langle e^{(p)}_l, e^{(p)}_m\rangle$, the upper bound gives :

$$\Lambda_N(\bar\theta) \le \frac{1}{N}\log\mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{2}\sum_{i,j=1}^k\theta_i\theta_j\psi^{(1)}_{i,j}\psi^{(2)}_{j,i}\Big)\Big]$$

but since the e_i are unitary and orthogonal , if we let $\Psi^{(p)} = (\psi^{(p)}_{i,j})_{1\le i,j\le k}$ we have $\Psi^{(1)}+\Psi^{(2)} = I_{2k}$ and so $\psi^{(1)}_{i,j}\psi^{(2)}_{j,i} = \psi^{(1)}_{i,j}(\mathbb{1}_{i=j} - \bar\psi^{(1)}_{i,j})$. Furthermore the $\Psi^{(1)}$ is an element of a Jacobi ensemble as the following lemma states :

Lemma 5.1. *The distribution of the matrix $\Psi^{(1)}$ when $N>k$ is given by the following density for the Lebesgue measure on the set of symmetric/Hermitian matrices :*

$$\frac{1}{Z}\det(\Psi^{(1)})^{\beta\frac{L-k+1}{2}-1}\det(I_k-\Psi^{(1)})^{\beta\frac{M-k+1}{2}-1}\mathbb{1}_{0\le\Psi^{(1)}\le I_k}d\Psi^{(1)}$$

Proof: Let U be a orthogonal/unitary $N\times N$ Haar matrix, U_1 its $L\times k$ top left block. Then $\Psi^{(1)}$ has the same law as $U_1^*U_1$. If we denote Π the matrix $diag(\underbrace{1,...,1}_{\text{L times}},0,...0)$ and Π' the matrix $diag(\underbrace{1,...,1}_{\text{k times}},0,...0)$, then $U_1^*U_1 = \Pi'U^*\Pi U\Pi'$. Then we can apply Collins (2005, Theorem 2.2) (up to adapt this theorem to the real case). □

Therefore, using Laplace's method, we see that the distribution of $\Psi^{(1)}$ satisfies a large deviations principle with rate function I :

$$I(M) = \begin{cases} -\frac{\beta}{2}\Big[\frac{1}{1+\alpha}\log\det(M) + \frac{\alpha}{1+\alpha}\log\det(I_k-M)\Big] - C \text{ if } 0\le M\le I_k,\\ +\infty \text{ otherwise.}\end{cases}$$

where C is such that $\min I = 0$. As a consequence, Varadhan's lemma implies that

$$\limsup_{N\to\infty} \Lambda_N(\bar\theta) \le \Lambda(\bar\theta)$$

where :

$$\Lambda(\bar\theta) = \sup_{0\le M\le I_k} [f(M) - I(M)]$$

with $f(M) = \frac{\beta}{4}\sum_{i,j=1}^k \theta_i\theta_j M_{i,j}(I_k - M)_{j,i}$. We notice by taking $M = \alpha' I_k$ that

$$Z \le -\frac{k\beta}{2}\Big(\alpha\log\alpha' + (1-\alpha')\log(1-\alpha')\Big)\,.$$

On the other hand, because $\det(A) \le \prod A_{ii}$ for any positive self-adjoint matrix A, for all matrix M such that $0 \le M \le I_k$:

$$I(M) \ge -\frac{\beta}{2}\Big[\sum_{i=1}^k \{\alpha'\log(M_{i,i}) + (1-\alpha')\log(1-M_{i,i})\}\Big] + Z$$

whereas $f(M) \ge \frac{\beta}{4}\sum_{i=1}^k \theta_i\theta_j M_{i,i}(I_k - M)_{i,i}$ since the off-diagonal terms are non-positive (because M is symmetric and the θ_i's non-negative). We deduce (with $M_{i,i} = a_i$) that

$$\Lambda(\bar\theta) \le \frac{\beta}{2}\sup_{(a_i)_{i=1}^k\in]0,1[^k}\sum_{i=1}^k \Big(\theta_i^2 a_i(1-a_i) + \alpha'\log\frac{a_i}{\alpha'} + (1-\alpha')\log\frac{1-a_i}{1-\alpha'}\Big) = \sum_{i=1}^k \Lambda(\theta_i)$$

To obtain the lower bound on $\liminf_N \Lambda_N(\bar\theta)$ as in Guionnet and Husson (2020), it is enough to find a sequence of events V_N^κ independent of $\Psi^{(1)}$ such that on these events $|e_i(l)| \le CN^{-1/4-\kappa}$ for some $\kappa > 0$ and all i and l since then we will be in the regime where the sharp sub-Gaussian bound is also a lower bound. Note here that $\Psi^{(2)}$ is determined by $\Psi^{(1)}$, so we only condition on $\Psi^{(1)}$. To do that let us denote U the $k\times L$ matrix with column vectors $(e_i^{(1)}, 1 \le i \le k)$. Then

$$U = (\Psi^{(1)})^{1/2} V$$

and conditionally to $\Psi^{(1)}$, $V = (v_1,\dots,v_k)$ follows the uniform law on the set of k orthonormal vectors on the sphere $\mathbb{S}_L$. We can then let $V_N^\kappa = \{\max_i \max_l |v_i(l)| \le N^{-1/4-\kappa}\}$. On this set, $\max_i\max_l |e_i(l)| \le CN^{-1/4-\kappa}$ so that

$$\begin{aligned}\Lambda_N(\bar\theta) &\ge \mathbb{E}_e\Big[1_{e\in V_N^\kappa}\exp\Big(\frac{\beta N}{4}(1+o(1))\sum_{\substack{l=1,\dots,L\\ m=L+1,\dots,N}}\sum_{i,j=1}^k \theta_i\theta_j \Re(e_i(l)\bar e_i(m)\bar e_j(l)e_j(m))\Big)\Big]\\ &= \mathbb{E}_e\Big[1_{e\in V_N^\kappa}\exp\Big(\frac{\beta N}{2}\sum_{i,j=1}^k \theta_i\theta_j\psi_{i,j}^{(1)}\psi_{j,i}^{(2)}\Big)\Big]\end{aligned}$$

where we expended the Laplace transform of the entries close to the origin. We finally notice that V_N^κ is independent of $\Psi^{(1)}$ and with probability going to one. We can therefore apply the large deviations principle to deduce that

$$\limsup_{N\to\infty}\Lambda_N(\bar\theta) \ge \sup_{0\le M\le I_k}[f(M) - I(M)]\,.$$

We finally conclude by taking M diagonal that the above right hand side is bounded below by $\sum\Lambda(\theta_i)$, which completes the proof of (5.4). To deduce the large deviations principle for the k largest eigenvalues of Wishart matrices, we first obtain a large deviations upper bound by tilting the measure by the k-dimensional spherical integral. Because it factorizes as well as $\Lambda(\bar\theta)$ the upper bound has a rate function given by the sum of the rate functions for each outliers. To obtain the large deviations upper bound, we tilt again the measure by $\exp(\frac{\beta N}{2}\sum_{i=1}^k \theta_i\langle e_i, \mathbf{Y}_N e_i\rangle)$ with $e_i \in V_N^\epsilon$. Under this tilted measure, we have the following expectations :

$$\mathbb{E}^{(e,\theta)}[\mathbf{Y}_N] = \sum_{i=1}^k \theta_i\Big(e_i^{(1)}(e_i^{(2)})^* + e_i^{(2)}(e_i^{(1)})^*\Big)$$

(where we identify $\mathbb{C}^L$ and $\mathbb{C}^M$ respectively with $\mathbb{C}^L \times \{0\}^M$ and $\{0\}^L \times \mathbb{C}^M$). We can then write

$$\mathbf{Y}_N = \tilde{\mathbf{Y}}_N + \sum_{i=1}^k \theta_i\Big(e_i^{(1)}(e_i^{(2)})^* + e_i^{(2)}(e_i^{(1)})^*\Big) + o(1)$$

where $\tilde{\mathbf{Y}}_N$ has the same form as $\mathbf{Y}_N$ under the original measure. Then to identify the eigenvalues of $\mathbf{Y}_N$ outside the bulk of the limit measure we need to solve the following equation

$$\det\Big(I_N + (\tilde{Y}_N - z)^{-1}\sum_{i=1}^k \theta_i(e_i^{(1)}(e_i^{(2)})^* + e_i^{(2)}(e_i^{(1)})^*)\Big) = 0$$

Note that the above arguments also show that in $\mathbb{P}^{\bar{\theta}}$-probability $\Psi^{(1)}$ converges towards the diagonal matrix with entries $(x_{\theta_i,\alpha})_{1\le i\le k}$. We also have local laws for $(z - \tilde{Y}_N)^{-1}$ under $\mathbb{P}^{\bar{\theta}}$. Therefore, if we denote $\tilde{\lambda}_+ = \sqrt{(1+\alpha)^{-1}\lambda_+}$ (which is the rightmost point of the support of the limit measure of $\mathbf{Y}_N$), the left hand side converges uniformly on any band $\{z \in \mathbb{C} : \tilde{\lambda}_+ + \epsilon \le \Re z \le A, |\Im z| \le 1\}$ toward :

$$g : z \mapsto \prod_{i=1}^k \Big(1 - \theta^2 z^2 x_{\theta_i,\alpha}(1 - x_{\theta_i,\alpha})(1+\alpha)^2 G_{MP(\alpha)}((1+\alpha)z^2) G_{MP(1/\alpha)}((1+\alpha)z^2)\Big)$$

where $MP(\alpha)$ is the Marchenko-Pastur distribution of parameter α. Using the fact the these functions are holomorphic, we have the k largest eigenvalue converges toward $z_{\theta_1,\alpha} \ge z_{\theta_k,\alpha}$ where $z_{\theta,\alpha}$ is defined as the unique solution of

$$1 - \theta^2 x_{\theta,\alpha}(1 - x_{\theta,\alpha})(1+\alpha)^2 z^2 G_{MP(\alpha)}((1+\alpha)z^2) G_{MP(1/\alpha)}((1+\alpha)z^2) = 0$$

on $]\tilde{\lambda}_+, +\infty[$ (see Guionnet and Husson (2020) for details).

5.3. *Universality of the large deviations for the k largest eigenvalues of Hermitian matrices with variance profiles and sharp sub-Gaussian entries.* We consider in this section the setting of Theorem 2.6, which generalizes the previous subsection. We will proceed as in Husson (2020) and we will first deal with the piecewise constant case with the supplementary technical assumption that the variance profile is non-negative.

The main point is to prove the following estimate for the annealed spherical integral.

Lemma 5.2. *Let*

$$\Lambda_N^\sigma(\bar{\theta}) = \frac{1}{N}\log \mathbb{E}_{\mathbf{X},e}[\exp(N\sum_{i=1}^k \theta_i\langle e_i, \mathbf{X}_N^\sigma e_i\rangle)]$$

Then, let σ be piecewise constant and under the assumptions of Theorem 2.6, for all $\theta_i \in \mathbb{R}^+$

$$\lim_{N\to\infty} \Lambda_N^\sigma(\bar{\theta}) = \sum_{i=1}^k \Lambda^\sigma(\theta_i)$$

with, if $R_{ij} := \sigma_{ij}^2$ and $S := \{\psi \in (\mathbb{R}^+)^p : \psi(1) + ... + \psi(p) = 1\}$ and denoting for an element ψ of S, $\psi(i)$ its i-th coordinate in $\mathbb{R}^p$,

$$\Lambda^\sigma(\theta) = \frac{\beta}{2}\sup_{\psi\in S}\Big[\frac{\theta^2}{2}\langle\psi, R\psi\rangle + \sum_{i=1}^p \alpha_i \log\frac{\psi(i)}{\alpha_i}\Big]$$

Indeed, let us define for $e \in \mathbb{S}^{\beta N-1}$ and $j \in [1,p]$, $e^{(p)}$ the vector of $\mathbb{C}^{\alpha_i(N)}$ whose coordinates are the coordinates of e whose indices lie in I_N^i. We then define for $j=1,...,p$ the random matrix $\Psi^{(j)} = (\langle e_l^{(j)}, e_m^{(j)}\rangle)_{1\le l,m\le k}$. Following the same computations as before and using the sharp sub-Gaussian character of the entries, we have :

$$\Lambda_N^\sigma(\bar\theta) \le \mathbb{E}_e\Big[\exp\Big(\frac{\beta N}{4}\sum_{l,m=1}^{k}\sum_{i,j=1}^{p}\theta_l\theta_m\Psi_{l,m}^{(i)}\bar\Psi_{l,m}^{(j)}\sigma_{i,j}^2\Big)\Big]$$

Notice that the $\Psi^{(j)}$ are Gram matrices (hence self-adjoint and positive) and that their sum is I_k. There again we will use a slightly improved version of the Lemma 5.1 to determine the distribution of the $\Psi^{(j)}$:

Lemma 5.3. *The joint distribution of the matrices $\Psi^{(1)},...,\Psi^{(p-1)}$ when $\alpha_1(N),...,\alpha_p(N)>k$ is given by the following density for the Lebesgue measure on the set of symmetric/Hermitian matrices :*

$$\frac{1}{Z}\prod_{i=1}^{p-1}\Big(\mathbb{1}_{0\le\Psi^{(i)}}\det(\Psi^{(i)})^{\beta\frac{\alpha_i(N)-k+1}{2}-1}\Big)\det(I_k-\sum_{i=1}^{p-1}\Psi^{(i)})^{\beta\frac{\alpha_p(N)-k+1}{2}-1}\mathbb{1}_{\sum_{i=1}^{p-1}\Psi^{(i)}\le I_k}\prod_{i=1}^{p-1}d\Psi^{(i)}$$

Proof: Here we need an improved version of Collins (2005, Theorem 2.2) which states as follows. Let U be a $N\times N$ Haar-distributed orthogonal or unitary matrix, $n_0=0<n_1<...<n_p=N$ a p-uplet of integers and for $i\in[1,p]$, $\tilde\pi_i$ the orthogonal projection on the vector span of the columns of U with indices between $n_{i-1}+1$ and n_i. Let π be a constant projection of rank k. Then, if we identify $\pi S_N(\mathbb{R})\pi$ (respectively $\pi H_N(\mathbb{C})\pi$) to $S_k(\mathbb{R})$ (respectively $H_k(\mathbb{C})$), the joint distribution of $(M_1,...,M_{p-1}) = \pi\tilde\pi_1\pi,...,\pi\tilde\pi_{p-1}\pi$ has the following density on $S_N(\mathbb{R})^{p-1}$ (resp. $H_N(\mathbb{C})^{p-1}$) :

$$\frac{1}{Z}\prod_{i=1}^{p-1}\Big(\mathbb{1}_{0\le M_i}\det(M_i)^{\beta\frac{m_i-k+1}{2}-1}\Big)\det(I_k-\sum_{i=1}^{p-1}M_i)^{\beta\frac{m_p-k+1}{2}-1}\mathbb{1}_{\sum_{i=1}^{p-1}M_i\le I_k}\prod_{i=1}^{p-1}dM_i$$

where $m_i = n_i - n_{i-1}$.

The proof of this result is the same as the proof of Collins (2005, Theorem 2.2). The difference is that one needs to prove that $(\pi\tilde\pi_i\pi)_{1\le p-1}$ has the same law as $(\Sigma^{-1/2}X_i\Sigma^{-1/2})$ where the X_i are independent Gaussian Wishart of parameters (k,m_i) and $\Sigma = X_1+...+X_p$. Once we have this result, we take a Haar-distributed unitary matrix U and we denote U_i the $\alpha_i(N)\times k$ matrix extracted from U by taking its k first columns and its rows of indices in I_N^i. We denote $\Pi' = \mathrm{diag}(\underbrace{1,...,1}_{\text{k times}},0,..,0)$ and Π_i the diagonal matrix with entries equal to 1 for indices in I_N^i and 0 elsewhere. Then, since $(\Psi^{(i)})_{1\le p-1}$ has the same law as $(\Pi' U^*\Pi_i U\Pi')_{i\le p-1}$, we can use the previous theorem. □

We deduce from this explicit distribution of the $p-1$-uplet $(\Psi^{(i)})_{1\le i\le p}$ that it follows a large deviations principle with rate function :

$$I((M_i)_{1\le i\le p-1}) = \begin{cases} -\frac{\beta}{2}\Big[\sum_{i=1}^p\alpha_i\log\det(M_i)\Big]-C \text{ if } \forall i\in[1,p], 0\le M_i\le I_k, \text{ and } \sum_{i=1}^p M_i = I_k \\ +\infty \text{ otherwise.}\end{cases}$$

Then we have using Varadhan's lemma :

$$\limsup_N \Lambda_N^\sigma(\bar\theta) \le \Lambda^\sigma(\bar\theta)$$

where :

$$\Lambda^\sigma(\bar\theta) = \sup_{(M_i)_{1\le i\le p}}[f((M_i)) - I((M_i))]$$

with $f((M_i)) = \frac{\beta}{4}\sum_{i,j=1}^k \theta_i\theta_j \langle M^{i,j}, RM^{i,j}\rangle$ where R is the $p\times p$ matrix $(\sigma_{i,j}^2)$ and $M^{i,j}$ is the vector $(M_1(i,j),...,M_p(i,j))$. But, as before, if $d(M)$ represents the diagonal matrix with entries $(M_{ii})_{1\leq i\leq k}$ we have that

$$I((M_i)) \geq I((d(M_i))) \text{ and for } i\neq j \langle M^{i,j}, RM^{i,j}\rangle \leq 0$$

where the last inequality is due to Assumption 2.5 which implies $\sum_{i=1}^p M_i = I_k$ and therefore that for $i \neq j$ $\sum_{l=1}^p M^{i,j}(l) = 0$. Therefore we can again restrict the sup to diagonal matrices and it then decouples into

$$\Lambda^\sigma(\bar\theta) = \frac{\beta}{2}\sup_{(\psi_i)_{1\leq i\leq k}\in S^k}\Big[\sum_{j=1}^k \frac{\theta_i^2}{2}\langle\psi_j, R\psi_j\rangle + \sum_{j=1}^k\sum_{i=1}^p \alpha_i \log\frac{\psi_i(j)}{\alpha_i}\Big] = \sum_{j=1}^k \Lambda^\sigma(\theta_j)$$

where we remind that $S := \{\psi \in (\mathbb{R}^+)^p : \psi(1)+...+\psi(p) = 1\}$. In particular, since the function $\psi\mapsto\langle\psi, R\psi\rangle$ is concave on S thanks to Assumption 2.5, the function optimized is strictly concave and thus is maximum at a unique ψ. Furthermore ψ_j only depends on θ_j so that we will denote $\psi_j = \psi_{\theta_j}$. Using again the strict concavity and the implicit function theorem, we have that the function $\theta\mapsto\psi_\theta$ is analytic in θ. Furthermore, if we tilt our measure by $\mathbb{E}_{\mathbf{X}}[\exp(N\sum\theta_i\langle e_i, \mathbf{X}_N^\sigma e_i\rangle)]$, the $\Psi^{(i)}$'s follow a large deviations principle and converges respectively toward $\mathrm{diag}(\psi_{\theta_1}(i),...,\psi_{\theta_k}(i))$. For the lower bound we restrict the integral as in the preceding subsection to delocalized vectors with fixed Ψ and conclude similarly.

To prove the large deviations principle, we first observe that the large deviations upper bound is direct after a tilt by spherical integrals and decoupling of the annealed spherical integrals. For the large deviations lower bound, we tilt by $\exp(N\sum_{i=1}^k\theta_i\langle e_i, \mathbf{X}_N^\sigma e_i\rangle)$. Under this tilted measure $\mathbb{P}^{e,\bar\theta}$, we have the following expectation $\mathbb{E}^{e,\bar\theta}[\mathbf{X}_N^\sigma] = \sum_{i=1}^k\theta_i\sum_{l,m=1}^p\sigma_{l,m}^2 e_i^{(l)}(e_i^{(m)})^*$. Using the BBP transition phenomenon, the local law for $\mathbf{X}_N$ as in Husson (2020, Lemma 5.3) and the fact that the $\Psi^{(k)}$ converges in $\mathbb{P}^{\bar\theta}$ - probability, we have that the eigenvalues outside the bulk are asymptotically solution of the following equation in z :

$$\prod_{i=1}^k \det(I_p - \theta_i RD(\theta_i, z)) = 0$$

where $D(\theta, z)$ is defined as in Husson (2020, Section 5). To conclude, it suffices to prove that for any $z_1 > ... > z_k > r_\sigma$, there exists $\theta_1\geq...\geq\theta_k\geq 0$ such that z_i is the unique solution of $\det(I_p - \theta RD(\theta,z)) = 0$ on $]r_\mu, +\infty[$. We already know thanks again to the proof of the large deviations lower bound in Husson (2020) that there is for every z, a θ such that z is the largest solution. Let us prove that with Assumption 2.5, this solution is unique on $]r_\sigma, +\infty[$. First, one can notice that this assumption implies that the quadratic form whose matrix is R has signature $(1, p-1)$ and so it is also true for the quadratic form whose matrix is $\sqrt{D(\theta,z)}R\sqrt{D(\theta,z)}$. Therefore, if we denote $\rho(\theta, z)$ the largest eigenvalue of $\sqrt{D(\theta,z)}R\sqrt{D(\theta,z)}$, the equation $\det(I_p - \theta_i RD(\theta_i, z)) = 0$ is equivalent for $\theta > 0$ and $z > r_\sigma$ to $\theta\rho(\theta,z) = 1$. Since $z\mapsto\rho(\theta,z)$ is strictly decreasing, the result is then proved.

For the continuous case, we can as in Husson (2020, Section 6) approximate our continuous variance profiles by piecewise constant ones. This approximation step is in fact easier than in the more general case of Husson (2020) since if σ satisfy Assumption 2.5 then we can approximate $\mathbf{X}_N^\sigma$ by the $\mathbf{X}_N^{(p)}$ defined as follows :

$$\mathbf{X}_N^{(p)} = \sigma_N^{(p)}(i,j)\frac{X_{i,j}}{\sqrt{N}}$$

where $\sigma_N^{(p)}(i,j) = \sum_{k,l=1}^p \sigma_{k,l}^{(p)} \mathbb{1}_{I_N^k \times I_N^l}(i,j)$ if $I_N^1 = [0, N/p]$ and $I_N^i =]N(i-1)/p, Ni/p]$ for $i = 2...p$ and

$$\sigma_{i,j}^{(p)} = \sqrt{p^2 \int_{(i-1)/p}^{i/p} \int_{(j-1)/p}^{j/p} \sigma^2(x,y)dxdy}\,.$$

Since σ satisfy Assumption 2.5, it is easy to check that $\sigma^{(p)}$ also satisfies Assumption 2.5 for all p and therefore if we denote $\lambda_N^{(p),1}, ..., \lambda_N^{(p),k}$ its k largest eigenvalues, they satisfy a large deviations principle with rate function $I^{(p)}(x_1, ..., x_p) = \sum I^{(p)}(x_i)$. If we denote $\lambda_N^1, ..., \lambda_N^k$ the k largest eigenvalues of $\mathbf{X}_N$, we have for all $i = 1, ..., k$, $|\lambda_N^i - \lambda_N^{(p),i}| \leq ||\mathbf{X}_N - \mathbf{X}_N^{(p)}||$. Using Husson (2020, Lemma 6.6), we have that $||\mathbf{X}_N - \mathbf{X}_N^{(p)}||$ can be neglected at exponential scale once p is large enough. And using again Husson (2020, Lemma 6.4 and Lemma 6.5), we have that the rate function converges toward the sum of the rate functions for one eigenvalue. Therefore, $\lambda_N^1, ..., \lambda_N^k$ satisfy a large deviations principle and the rate function is the sum of the rate functions for one eigenvalue.

Remark 5.4. Contrary to the Wigner case where we can see that asymptotically the positive and negative eigenvalues deviate independently from one another, this is not the case for matrices with variance profiles. An example is the linearization of a Wishart matrix where the negative eigenvalues are always exactly the opposite of the positive ones.

5.4. *Large deviations for the k largest eigenvalues for the Gaussian ensembles with a k-dimensional perturbation.* We next prove Proposition 2.7. We first observe that the result is well known when $\bar{\theta} = 0$, see e.g. Theorem 2.3. We next remark that the joint law of the eigenvalues of $\mathbf{X}_N^\theta$ is given by

$$d\mathbb{P}_N^\theta(\lambda) = \frac{1}{Z_N}\Delta(\lambda)^\beta \int \exp\{-\frac{\beta}{4}N\mathrm{Tr}|UD(\lambda)U^* - \sum_{i=1}^k \theta_i e_i e_i^*|^2\}dU \prod_{1\leq i\leq N} d\lambda_i$$

where U follows the Haar measure on the unitary group (resp. the orthogonal group) when $\beta = 2$ (resp, $\beta = 1$). $\Delta(\lambda) = \prod_{i<j}|x_i - x_j|$ is the Vandermonde determinant and $D(\lambda)$ is a diagonal matrix with entries given by $\lambda = (\lambda_1, \ldots, \lambda_N)$. Expanding the integral under the unitary (or orthogonal) group, we find that

$$d\mathbb{P}_N^\theta(\lambda) = \frac{1}{\tilde{Z}_N}\mathbb{E}_e[e^{\frac{\beta}{2}N\sum_{i=1}^k \theta_i\langle e_i, D(\lambda)e_i\rangle}]d\mathbb{P}_N^0(\lambda)\,,$$

where $(e_1, \ldots, e_k)$ follows the uniform law on k orthonormal vectors in dimension N. Hence the density is exactly given by the spherical integral. Using that Assumption 2.2 holds under $\mathbb{P}_N^0$ (see e.g Guionnet and Zeitouni (2000)), we see that the empirical measure of λ is close to the semi-circle law with overwhelming probability. Assume that $\theta_1 \geq \theta_2 \cdots \geq \theta_p \geq 0 \geq \theta_{p+1} \cdots \geq \theta_k$. Then, on the set where the extreme eigenvalues $\lambda_N^N \geq \cdots \lambda_{N-p}^N$ and $\lambda_1^N \leq \cdots \leq \lambda_{k-p+1}^N$ are close to $x_1 \geq x_2 \geq \cdots \geq x_p \geq 2 \geq -2 \geq x_{-k+p} \geq \cdots \geq x_{-1}$, Theorem 4.1 and Varadhan's Lemma give the result.

5.5. *Large deviations for k extreme eigenvalues for Gaussian Wishart matrices with a k-dimensional perturbation.* The proof of Proposition 2.8 is similar to the previous one. Again the proof is based on the explicit joint law of $\lambda_1^{N,\gamma} \geq \lambda_2^{N,\gamma} \geq \cdots \lambda_M^{N,\gamma}$ given by the law on $(\mathbb{R}^+)^M$

$$d\mathbb{P}_{M,N}^{\bar{\gamma}}(d\lambda) = \frac{1}{Z_N}\Delta(\lambda)^\beta \int e^{-\frac{\beta}{2}N\mathrm{Tr}(UD(\lambda)U^*\Sigma^{-1})}dU \prod_{1\leq i\leq M} \lambda_i^{\frac{\beta}{2}(N-M+1)}d\lambda_i$$

Noticing that

$$\Sigma^{-1} = I + \sum_{i=1}^{k} \frac{\gamma_i}{1-\gamma_i} e_i e_i^*$$

we recognize again that the density with respect to the case $\gamma = 0$ is given by a spherical integral. The result follows as for the Wigner case.

6. Appendix

In this Appendix we investigate the continuity property of spherical integrals. First we need to prove the continuity of the deterministic limit itself :

Theorem 6.1. *Let d be a distance compatible with the weak topology on the set $\mathcal{P}(\mathbb{R})$ and $||.||$ any norm on $\mathbb{R}^{k+\ell}$, and for $M > 0$, $\mathcal{K}_M$ the subset of $E = \mathbb{R}^{k+\ell} \times (\mathbb{R}^+)^k \times (\mathbb{R}^-)^\ell \times \mathcal{P}(\mathbb{R})$ defined by*

$$\mathcal{K}_M := \{(\bar{\lambda}, \bar{\theta}, \mu) \in E | M \geq \theta_1 \geq ... \geq \theta_k \geq 0 \geq \theta_{-\ell} \geq ... \geq \theta{-1} \geq -M$$
$$M \geq \lambda_1 \geq ... \geq \lambda_k \geq r_\mu \geq l_\mu \geq \lambda_{-\ell} \geq ... \theta_{-1} \geq -M\}$$

where r_μ and l_μ are respectively the rightmost and the leftmost point of the support of μ. We endow $\mathcal{K}_M$ with the distance D given by $D((\bar{\lambda}, \bar{\theta}, \mu), (\bar{\lambda}', \bar{\theta}', \mu')) = d(\mu, \mu') + ||\bar{\lambda} - \bar{\lambda}'|| + ||\bar{\theta} - \bar{\theta}'||$. Then $\mathcal{K}_M$ is a compact set and the function J

$$J(\mu, \bar{\theta}, \bar{\lambda}) = \sum_{i=-\ell, \neq 0}^{k} J(\mu, \theta_i, \lambda_i)$$

is continuous on $\mathcal{K}_M$.

Proof: It is clear that we only need to prove the continuity of $(\theta, \lambda, \mu) \mapsto J(\mu, \theta, \lambda)$ where either $\theta \geq 0$ and $\lambda \geq r_\mu$ or $\theta \leq 0$ and $\lambda \leq l_\mu$. We assume without loss of generality that we are in the first case. Furthermore since $J(\mu, \theta, \lambda) = J(\theta * \mu, \theta\lambda, 1)$ we only need to prove the continuity for the first two arguments with the third being fixed equal to 1. Let us take a sequence (μ_n, λ_n) such that $\forall n \in \mathbb{N}, l_\mu \geq -M, r_{\mu_n} \leq \lambda_n$ and $\lim \lambda_n = \lambda$ and $\lim_n \mu_n = \mu$. First, since $|J(\mu_n, \lambda_n, 1) - J(\mu_n, \lambda + \epsilon, 1)| \leq |\lambda_n - \lambda| + \epsilon$ for n large enough so that $\lambda + \epsilon \geq r_{\mu_n}$, and $|J(\mu, \lambda, 1) - J(\mu_n, \lambda + \epsilon, 1)| \leq \epsilon$ we can restrict ourselves to proving $\lim J(\mu_n, \lambda + \epsilon, 1) = J(\mu, \lambda + \epsilon, 1)$. But, when we differentiate $J(\mu, \lambda, 1)$ on the variable λ, we find

$$\frac{\partial}{\partial \lambda} J(\mu, \lambda, 1) = \mathbb{1}_{[G_\mu^{-1}(1), +\infty[}(\lambda)(1 - G_\mu(\lambda))$$

On $[\lambda + \epsilon, +\infty[$, since $r_{\mu_n} \leq \lambda + \epsilon$ it is in fact easy to see that the weak convergence of μ_n imply the uniform convergence of $\partial / \partial \lambda J(\mu_n, \lambda, 1)$. The we conclude by choosing $\Lambda > \lambda + \epsilon$ so that $G_\mu(\Lambda) \leq 1/2$ so that $v(\mu_n, 1, \Lambda) = \Lambda$ for n large enough and then using the weak convergence and the fact that $x \mapsto \log(\Lambda - x)$ is bounded on $[-M, \lambda + \epsilon]$, we have that $J(\mu_n, 1, \Lambda)$ converges toward $J(\mu, 1, \Lambda)$. □

With this continuity and the compactness of $\mathcal{K}_M$, we can prove the following theorem of uniform continuity, which generalizes Maïda (2007) :

Theorem 6.2. *Let $k, \ell \in \mathbb{N}$, $\bar{\theta} \in (\mathbb{R}^-)^\ell \times (\mathbb{R}^+)^k$ and*

$$J_N(X_N, \bar{\theta}) = \frac{1}{N} \log \mathbb{E}\Big[\exp\Big(\frac{\beta N}{2} \sum_{i=-\ell, i\neq 0}^{k} \theta_i \langle e_i, \mathbf{X}_N e_i \rangle\Big)\Big]$$

Let us denote $\lambda_{-1}^N \leq ... \leq \lambda_{-\ell}^N$ the smallest outliers of $\mathbf{X}_N$ and $\lambda_1^N \geq ... \geq \lambda_k^N$ the largest outliers. We will denote this $k + \ell$-tuple $\bar{\lambda}_N$. Then for every $M > 0$ and $\epsilon > 0$, there is $N_0 \in \mathbb{N}$ so that for *every $N \geq N_0$, for any matrix $\mathbf{X}_N$ such that $(\bar{\lambda}_N, \bar{\theta}, \mu_{\mathbf{X}_N}) \in \mathcal{K}_M$*

$$|J_N(\mathbf{X}_N, \bar{\theta}) - J(\mu_{\mathbf{X}_N}, \bar{\lambda}_N, \bar{\theta})| \leq \epsilon$$

Proof: We first notice that in the proof of Proposition 4.1 we approximated $J_N(\mathbf{X}_N, \bar{\theta})$ by $J_N(\mathbf{X}_N^{\varepsilon}, \bar{\theta})$ with an error depending only on ε. Hence we may and shall replace in the above statement $\mathbf{X}_N$ by $\mathbf{X}_N^{\delta}$ for some small enough $\delta = \delta(\epsilon)$. $\mathbf{X}_N^{\delta}$ has the same extreme eigenvalues than $\mathbf{X}_N$ and otherwise eigenvalues $\lambda_{-\ell}^N + j\delta$ with multiplicity $\lfloor N\hat{\mu}^N([\lambda_{-\ell}^N + j\delta, \lambda_{-\ell}^N + (j+1)\delta]) \rfloor$. Therefore, we see that $J_N(\mathbf{X}_N, \bar{\theta})$ is a function of the extreme eigenvalues and the empirical measure, hence a function on $\mathcal{K}_N$. By the previous uniform approximation and the continuity of the limit, we deduce that it is uniformly continuous on $\mathcal{K}_N$, hence the result. □

References

Anderson, G. W., Guionnet, A., and Zeitouni, O. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2010). ISBN 978-0-521-19452-5. MR2760897.

Baik, J., Ben Arous, G., and Péché, S. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, **33** (5), 1643–1697 (2005). MR2165575.

Belinschi, S., Guionnet, A., and Huang, J. Large deviations principles via spherical integrals. *ArXiv Mathematics e-prints* (2020). To appear in *Prob. Math. Phys* (2022+). arXiv: 2004.07117.

Ben Arous, G., Dembo, A., and Guionnet, A. Aging of spherical spin glasses. *Probab. Theory Related Fields*, **120** (1), 1–67 (2001). MR1856194.

Benaych-Georges, F. Rectangular R-transform as the limit of rectangular spherical integrals. *J. Theoret. Probab.*, **24** (4), 969–987 (2011). MR2851240.

Biroli, G. and Guionnet, A. Large deviations for the largest eigenvalues and eigenvectors of spiked Gaussian random matrices. *Electron. Commun. Probab.*, **25**, Paper No. 70, 13 (2020). MR4158230.

Collins, B. Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields*, **133** (3), 315–344 (2005). MR2198015.

Collins, B., Guionnet, A., and Maurel-Segala, E. Asymptotics of unitary and orthogonal matrix integrals. *Adv. Math.*, **222** (1), 172–215 (2009). MR2531371.

Collins, B. and Śniady, P. New scaling of Itzykson-Zuber integrals. *Ann. Inst. H. Poincaré Probab. Statist.*, **43** (2), 139–146 (2007). MR2303115.

Coquereaux, R., McSwiggen, C., and Zuber, J.-B. On Horn's problem and its volume function. *Comm. Math. Phys.*, **376** (3), 2409–2439 (2020). MR4104554.

Dean, D. S. and Majumdar, S. N. Large deviations of extreme eigenvalues of random matrices. *Phys. Rev. Lett.*, **97** (16), 160201, 4 (2006). MR2274338.

Gorin, V. and Panova, G. Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. *Ann. Probab.*, **43** (6), 3052–3132 (2015). MR3433577.

Guionnet, A. and Husson, J. Large deviations for the largest eigenvalue of Rademacher matrices. *Ann. Probab.*, **48** (3), 1436–1465 (2020). MR4112720.

Guionnet, A. and Maïda, M. A Fourier view on the R-transform and related asymptotics of spherical integrals. *J. Funct. Anal.*, **222** (2), 435–490 (2005). MR2132396.

Guionnet, A. and Maurel-Segala, E. Combinatorial aspects of matrix models. *ALEA Lat. Am. J. Probab. Math. Stat.*, **1**, 241–279 (2006). MR2249657.

Guionnet, A. and Zeitouni, O. Concentration of the spectral measure for large matrices. *Electron. Comm. Probab.*, **5**, 119–136 (2000). MR1781846.

Guionnet, A. and Zeitouni, O. Large deviations asymptotics for spherical integrals. *J. Funct. Anal.*, **188** (2), 461–515 (2002). MR1883414.

Harish-Chandra. Invariant differential operators on a semisimple Lie algebra. *Proc. Nat. Acad. Sci. U.S.A.*, **42**, 252–253 (1956). MR80260.

Huang, J. Asymptotic expansion of spherical integral. *J. Theoret. Probab.*, **32** (2), 1051–1075 (2019). MR3959637.

Husson, J. Large deviations for the largest eigenvalue of matrices with variance profiles. *ArXiv Mathematics e-prints* (2020). arXiv: 2002.01010.

Itzykson, C. and Zuber, J. B. The planar approximation. II. *J. Math. Phys.*, **21** (3), 411–421 (1980). MR562985.

Maïda, M. Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.*, **12**, 1131–1150 (2007). MR2336602.

Maïda, M., Najim, J., and Péché, S. Large deviations for weighted empirical mean with outliers. *Stochastic Process. Appl.*, **117** (10), 1373–1403 (2007). MR2353032.

Pizzo, A., Renfrew, D., and Soshnikov, A. On finite rank deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, **49** (1), 64–94 (2013). MR3060148.

Potters, M. and Mergny, P. Asymptotic behavior of the multiplicative counterpart of the Harish-Chandra integral and the S-transform. *ArXiv Mathematics e-prints* (2020). arXiv: 2007.09421.

Wigner, E. P. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)*, **62**, 548–564 (1955). MR77805.

Zuber, J.-B. Horn's problem and Harish-Chandra's integrals. Probability density functions. *Ann. Inst. Henri Poincaré D*, **5** (3), 309–338 (2018). MR3835548.