



HAL
open science

Sum rules and large deviations for spectral matrix measures

Fabrice Gamboa, Jan Nagel, Alain Rouault

► **To cite this version:**

Fabrice Gamboa, Jan Nagel, Alain Rouault. Sum rules and large deviations for spectral matrix measures. 2016. hal-01264137v1

HAL Id: hal-01264137

<https://hal.science/hal-01264137v1>

Preprint submitted on 28 Jan 2016 (v1), last revised 20 Sep 2016 (v2)

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

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

Sum rules and large deviations for spectral matrix measures

Fabrice Gamboa, Jan Nagel and Alain Rouault

January 28, 2016

Abstract

A sum rule is a relationship between the reversed Kullback-Leibler divergence of a positive measure on \mathbb{R} and some non-linear functional built on spectral elements related to this measure (see for example [13] or [20]). In this paper, using only probabilistic tools of large deviations, we extend the sum rules obtained in [9] to the case of an Hermitian matrix valued measure. We recover earlier result of Damanik et al ([3]) for the case of the semi-circular law and obtain new sum rules for Hermitian matrix measures in the Pastur-Marchenko case.

1 Introduction

The main mathematical objects considered in this paper rely on the notion of Hermitian matrix measure (that we shorten in matrix measure). To begin with, let us recall what is a matrix measure. For $p, q \in \mathbb{N}_*$, let denote by $\mathbb{M}_{p,q}$ the set of all complex matrices of size $p \times q$. Further, let \mathbb{H}_p be the set of Hermitian matrices of size $p \times p$. A matrix measure $\Sigma = (\Sigma_{i,j})_{i,j}$ of size p on \mathbb{R} is a matrix of signed complex measures, such that $\Sigma(A) = (\Sigma_{i,j}(A))_{i,j} \in \mathbb{H}_p$ for any Borel set $A \subset \mathbb{R}$. Further, if for any A , $\Sigma(A)$ is a nonnegative matrix we say that Σ is nonnegative. We denote by $\mathcal{M}_p(T)$ the set of $p \times p$ nonnegative matrix measures with support in $T \subset \mathbb{R}$. Further, $\mathcal{M}_{p,1}(T)$ is the subset of $\mathcal{M}_p(T)$ of normalized measures Σ satisfying $\Sigma(T) = I_p$. For $p = 1$, a member of $\mathcal{M}_{p,1}(T)$ is called a scalar measure (classical probability measure). An obvious example of matrix measure is the spectral matrix measure. It is nothing more than the projections of the spectral resolution of a Hermitian matrix. More precisely, a Hermitian matrix X of size N may be written as UDU^* where $D = \text{diag}(\lambda_i)$ contains the eigenvalues of X and U is the matrix formed by an orthonormal basis of eigenvectors. Then, for $p \leq N$, there exists a unique spectral matrix measure $\Sigma_p^X \in \mathcal{M}_p(\mathbb{R})$ supported by the spectrum of X , such that for all

$k > 0$ and $v, w \in \mathbb{C}^p$

$$(1.1) \quad (v^*, 0_{1, N-p}) X^k \begin{pmatrix} w \\ 0_{N-p, 1} \end{pmatrix} = \int x^k d(v^* \Sigma_p^X w)(x),$$

where for $r, s \in \mathbb{N}_*$, $0_{r,s}$ is the null matrix of $\mathbb{M}_{r,s}$. For $j = 1, \dots, N$, let $u^{j,p} := (U_{i,j})_{i=1, \dots, p}$ be the j^{th} truncated column of U . Then, obviously

$$(1.2) \quad \Sigma_p^X(dx) = \sum_{j=1}^N u_j u_j^* \delta_{\lambda_j}(dx),$$

where δ_a is the Dirac measure in a . That is, $\{\lambda_1, \dots, \lambda_N\}$ is the support of Σ_p^X and $u_1 u_1^*, \dots, u_N u_N^*$ are the weights. Furthermore, as U is unitary we have $\sum_j u_j u_j^* = I_p$ so that $\Sigma_p^X \in \mathcal{M}_{p,1}(T)$.

If we assume further that $N = pn$ for some positive integer n and that all the eigenvalues of X have a single multiplicity then it is possible to build a block tridiagonal matrix

$$(1.3) \quad J_n = \begin{pmatrix} B_1 & A_1 & & & \\ A_1^* & B_2 & \ddots & & \\ & \ddots & \ddots & A_{n-1} & \\ & & A_{n-1}^* & B_n & \end{pmatrix}.$$

such that $\Sigma_p^X = \Sigma_p^{J_n}$. Here, all the blocks of J_n are matrices of $\mathbb{M}_{p,p}$. The case $p = 1$ is the most classical and relies on the construction of the orthogonal polynomials with respect to Σ_1^X (see for example [18]). The general case is more complicated and requires more technical tools from the theory of matrix orthogonal polynomials (see [19]). In Section 2.1, we will recall the construction of such tridiagonal representations. Notice that the case $p = 1$ is also useful in the context of random matrix theory. In this case, the tridiagonal representation (1.3) of the classical ensembles of random matrices have very nice properties [8]. First, of course, X and J_n share the same spectrum. Further, for the Gaussian unitary ensemble the coefficient appearing in the tridiagonal matrix are independent with simple distributions. The diagonal terms have Gaussian distribution while the subdiagonal ones have χ^2 distributions [8]. We will give in Lemmas 5.1 and 6.1 results in the same spirit in the general block case both for the Gaussian and Laguerre ensembles. In this paper, we assume that X is random and has quite general distribution. It is sampled in \mathbb{H}_N with a unitarily invariant distribution of the form

$$(1.4) \quad Z_N^{-1} e^{-\text{tr}NV(X)} dX,$$

for V a confining potential. Its eigenvalues behave as a Coulomb gas with potential V (see formula (3.3)) and the matrix of eigenvectors follows independently the Haar distribution in the set of unitary matrices.

One of our main result is that, as N goes to infinity, the random spectral matrix measure Σ_p^X converges to some equilibrium matrix measure (depending on the potential) at exponential rate. Indeed, we show that this random object satisfies a large deviations principle (LDP). To be self contained, we recall in Section 3.2 the definition and useful facts on LDP. Furthermore, the rate function of this LDP involves a matrix extension of the reversed Kullback-Leibler information with respect to the equilibrium matrix measure (see equation (2.13) and Theorem 4.2) and a contribution of the outlying eigenvalues. In the Gaussian and Laguerre unitary ensembles, as pointed out before, the coefficients encoding the Jacobi matrix J_n have very nice properties (independence and known distributions). These properties allow to compute the rate function of the LDP by the way of the blocks involved in the Jacobi representation. Further, the uniqueness of a rate function leads to our two mains Theorems 2.3 and 2.5 that give relationships between the reversed Kullback-Leibler information with respect to the equilibrium matrix measure and non linear functional built on the blocks involved in the Jacobi matrix. These formulas are sum rules (see [4], [3], [18], [19], [20] and references therein). Theorem 2.3 is proved in [3] by strong analysis tools. We recover this result by using only probabilistic arguments. For $p = 1$, it has been proved earlier in [13]. Up to our knowledge, our Theorem 2.5 is new and is the matrix extension of the one we have obtained in [9] for $p = 1$.

Our paper is organized as follows. In Section 2, we first give definitions and tools to handle both the block and the matrix measure frames. Then, we state our main results concerning the sum rule for spectral matrix measures. Section 3 deals with the asymptotic properties of the spectral random matrix measure. We state in Section 3.3 a general LDP for random spectral matrix measure drawn by using a general potential. The most technical proofs are postponed to Sections 4, 5 and 6.

2 Matrix orthogonal polynomials and block Jacobi operators

2.1 Construction

We will need to work with polynomials with coefficients in $\mathbb{M}_{p,p}$. They will be *orthogonal* with respect to some matrix measure on \mathbb{R} . We recall some useful facts and refer to [4] for more on the subject. Let us give some more notations. Let $p \in \mathbb{N}_*$ and $\Sigma \in \mathcal{M}_p(\mathbb{R})$ be a compactly supported matrix measure. Further, let F and G be measurable matrix valued functions $F, G : \mathbb{R} \rightarrow \mathbb{M}_{p,p}$.

We define two $p \times p$ matrices by setting

$$\begin{aligned}\langle\langle F, G \rangle\rangle_R &= \int_{\mathbb{R}} F(z)^* d\Sigma(z) G(z), \\ \langle\langle F, G \rangle\rangle_L &= \int_{\mathbb{R}} G(z) d\Sigma(z) F(z)^* .\end{aligned}$$

A sequence of functions on \mathbb{R} , (φ_j) with values in \mathbb{H}_p is called right-orthonormal if

$$\langle\langle \varphi_i, \varphi_j \rangle\rangle_R = \delta_{ij} I_p .$$

The orthogonal polynomial recursion is built as follows. First, assume that Σ is non-trivial, that is,

$$(2.1) \quad \text{tr} \langle\langle P, P \rangle\rangle_R > 0$$

for any non zero matrix polynomial P . See Lemma 2.1 of [4] for equivalent characterizations of non-triviality. We define the right monic matrix orthogonal polynomials P_n^R by applying Gram-Schmidt procedure to $\{I_p, xI_p, x^2I_p, \dots\}$. In other words, P_n^R is the unique matrix polynomial $P_n^R(x) = x^n I_p +$ lower order terms such that $\langle\langle x^k I_p, P_n^R \rangle\rangle_R = 0$ for $k = 0, \dots, n-1$. For non-trivial matrix measures, this is possible for any $n \geq 0$ and this sequence satisfies the recurrence relation:

$$(2.2) \quad xP_n^R(x) = P_{n+1}^R(x) + P_n^R(x)u_n^R + P_{n-1}^R v_n^R$$

If we set

$$\gamma_n := \langle\langle P_n^R, P_n^R \rangle\rangle_R$$

then γ_n is positive definite and we have

$$v_n^R = \gamma_{n-1}^{-1} \gamma_n .$$

To get normalized orthogonal polynomials p_n^R we set

$$(2.3) \quad p_0 = I_p \quad , \quad p_n^R = P_n^R \kappa_n^R$$

where the sequence of $p \times p$ matrix (κ_n^R) has to satisfy, for every n

$$(2.4) \quad \gamma_n = (\kappa_n^R (\kappa_n^R)^*)^{-1} .$$

This constraint opens several choices for κ_n^R (see Section 2.1.5 of [4]). Let us leave the choice open, setting

$$(2.5) \quad \kappa_n^R = \gamma_n^{-1/2} \sigma_n$$

with σ_n unitary.

Remark 2.1 We can define similarly the sequence of monic polynomials P_n^L and the sequence of left-orthonormal polynomials p_n^L in the same way. We have

$$P_n^R = (P_n^L)^*$$

and

$$\langle\langle P_n^L, P_n^L \rangle\rangle_L = \gamma_n$$

and the recurrence relation:

$$(2.6) \quad xP_n^L(x) = P_{n+1}^L(x) + u_n^L P_n^L(x) + v_n^L P_{n-1}^L$$

with

$$v_n^L = \gamma_n \gamma_{n-1}^{-1}.$$

The above condition (2.5) is replaced by $p_n^L = \tau_n \gamma_n^{-1/2} P_n^L$.

To formulate the recursion in terms of orthonormal polynomials, we use (2.3) and get

$$(2.7) \quad xp_n^R = p_{n+1}^R (\kappa_{n+1}^R)^{-1} \kappa_n^R + p_n^R (\kappa_n^R)^{-1} u_n^R \kappa_n^R + p_{n-1}^R (\kappa_{n-1}^R)^{-1} v_n^R \kappa_n^R$$

i.e.

$$(2.8) \quad xp_n^R = p_{n+1}^R A_{n+1}^* + p_n^R B_{n+1} + p_{n-1}^R A_n$$

with

$$(2.9) \quad A_n = (\kappa_{n-1}^R)^{-1} v_n^R \kappa_n^R = \sigma_{n-1}^* \gamma_{n-1}^{1/2} v_n^R \gamma_n^{-1/2} \sigma_n = \sigma_{n-1}^* \gamma_{n-1}^{-1/2} \gamma_n^{1/2} \sigma_n$$

$$(2.10) \quad B_{n+1} = (\kappa_n^R)^{-1} u_n^R \kappa_n^R = \sigma_n^* \gamma_n^{1/2} u_n^R \gamma_n^{-1/2} \sigma_n$$

Notice that (2.9) yields

$$(2.11) \quad A_n A_n^* = \sigma_{n-1}^* \gamma_{n-1}^{-1/2} \gamma_n \gamma_{n-1}^{-1/2} \sigma_{n-1}$$

In other terms the map $f \mapsto (x \mapsto xf(x))$ defined on the space of matrix polynomials is a right homomorphism and is represented in the (right-module) basis $\{p_0^R, p_1^R, \dots\}$ by the matrix

$$(2.12) \quad J = \begin{pmatrix} B_1 & A_1 & & \\ A_1^* & B_2 & \ddots & \\ & \ddots & \ddots & \end{pmatrix}$$

with B_k Hermitian and A_k non-singular. Moreover, if $\sigma_0 = I_p$, the measure Σ is again the spectral matrix measure of the matrix J defined as in (1.1) (Theorem 2.11 of [4]). Let us remark that although to each Σ corresponds a whole equivalence class of Jacobi coefficients given by the different σ_n , there is exactly one representative such that all A_k are Hermitian positive definite (Theorem 2.8 in [4]).

Starting with a finite dimensional Jacobi matrix J_n as in (1.3), the spectral matrix measure of J_n is supported by at most n points and is in particular not non-trivial. However, we may still define p_1^R, \dots, p_{n-1}^R by the recursion (2.8). As long as the matrices A_k are invertible, these polynomials are orthonormal with respect to the spectral measure of J_n and (2.1) holds for all matrix polynomials up to degree $n - 1$.

If Σ is a *quasi scalar* measure, that is if $\Sigma = I_p \sigma$ with σ a scalar measure and if Π is a positive matrix measure with Lebesgue decomposition

$$\Pi(dx) = h(x)\sigma(dx) + \Pi^s(dx),$$

we define

$$(2.13) \quad \mathcal{K}(\Sigma|\Pi) := - \int \log \det h(x) \sigma(dx).$$

Remark 2.2 *It is possible to rewrite the above quantity in the flavour of Kullback-Leibler information (or relative entropy) with the notation of [14] or [16], i.e.*

$$\mathcal{K}(\Sigma|\Pi) = \int \log \det \frac{d\Sigma(x)}{d\Pi(x)} d\sigma(x),$$

if Σ is strongly absolutely continuous with respect to Π , and infinity otherwise. See Corollary 8 in [11].

2.2 Measures on $[0, \infty)$

In [7], it is proved that if a non-trivial matrix measure Σ has a support included in $[0, \infty)$ then there exists a sequence $(\zeta_n)_n$ of non-singular elements of $\mathbb{M}_{p,p}$ such that

$$(2.14) \quad u_n^R = \zeta_{2n+1} + \zeta_{2n} \quad , \quad v_n^R = \zeta_{2n-1}\zeta_{2n}$$

with $\zeta_0 = 0_{p,p}$ and moreover,

$$(2.15) \quad \zeta_n = h_{n-1}^{-1} h_n, \quad (n \geq 1);$$

with $h_0 = I_p$ and for $n \geq 1$, h_n is an Hermitian positive definite matrix. Note that this implies $\gamma_n = \zeta_1 \dots \zeta_{2n} = h_{2n}$. We then have the relations for the orthonormal polynomials

$$(2.16) \quad B_{n+1} = \sigma_n^* \gamma_n^{1/2} (\zeta_{2n+1} + \zeta_{2n}) \gamma_n^{-1/2} \sigma_n$$

$$(2.17) \quad A_n = \sigma_{n-1}^* \gamma_{n-1}^{1/2} \zeta_{2n-1} \zeta_{2n} \gamma_n^{-1/2} \sigma_n.$$

In the scalar case, this yields

$$(2.18) \quad B_{n+1} = \zeta_{2n+1} + \zeta_{2n}, \quad (A_n)^2 = \zeta_{2n-1} \zeta_{2n}.$$

In the matrix case, we may set

$$(2.19) \quad Z_{2n+1} = \sigma_n^* \gamma_n^{1/2} \zeta_{2n+1} \gamma_n^{-1/2} \sigma_n, \quad Z_{2n} = \sigma_n^* \gamma_n^{1/2} \zeta_{2n} \gamma_n^{-1/2} \sigma_n.$$

To highlight a further decomposition, we set

$$C_n = \sigma_n^* h_{2n}^{1/2} h_{2n-1}^{-1/2}, \quad D_{n+1} = \sigma_n^* h_{2n}^{-1/2} h_{2n+1}^{1/2}.$$

With these definitions and (2.15) we see that

$$(2.20) \quad Z_{2n+1} = D_{n+1} D_{n+1}^*, \quad Z_{2n} = C_n C_n^*$$

are in fact Hermitian positive definite. For the recursion coefficients we get the following matrix analogues of (2.18), $B_1 = D_1 D_1^*$,

$$(2.21) \quad B_{n+1} = D_{n+1} D_{n+1}^* + C_n C_n^* = Z_{2n+1} + Z_{2n}, \quad A_n = D_n C_n^*.$$

In other words, the Jacobi operator J can in fact be decomposed as $J = X X^*$, where X is the bidiagonal matrix

$$(2.22) \quad X = \begin{pmatrix} D_1 & 0 & & \\ C_1 & D_2 & \ddots & \\ & \ddots & \ddots & \ddots \end{pmatrix}.$$

2.3 Results part I: Sum rules

For $\alpha^- < \alpha^+$, let $\mathcal{S}_p = \mathcal{S}_p(\alpha^-, \alpha^+)$ be the set of all bounded nonnegative measures $\Sigma \in \mathcal{M}_p(\mathbb{R})$ with

- (i) $\text{supp}(\Sigma) = K \cup \{\lambda_i^-\}_{i=1}^{N^-} \cup \{\lambda_i^+\}_{i=1}^{N^+}$, where $K \subset I = [\alpha^-, \alpha^+]$, $N^-, N^+ \in \mathbb{N} \cup \{\infty\}$ and

$$\lambda_1^- < \lambda_2^- < \dots < \alpha^- \quad \text{and} \quad \lambda_1^+ > \lambda_2^+ > \dots > \alpha^+.$$

(ii) If N^- (resp. N^+) is infinite, then λ_j^- converges towards α^- (resp. λ_j^+ converges to α^+).

Such a measure Σ can be written as

$$(2.23) \quad \Sigma = \Sigma|_I + \sum_{i=1}^{N^+} \Gamma_i^+ \delta_{\lambda_i^+} + \sum_{i=1}^{N^-} \Gamma_i^- \delta_{\lambda_i^-},$$

for some nonnegative Hermitian matrices $\Gamma_1^+, \dots, \Gamma_{N^+}^+, \Gamma_1^-, \dots, \Gamma_{N^-}^-$. Further, we define $\mathcal{S}_{p,1} = \mathcal{S}_{p,1}(\alpha^-, \alpha^+) := \{\Sigma \in \mathcal{S}_p(\alpha^-, \alpha^+) \mid \Sigma(\mathbb{R}) = I_p\}$.

2.3.1 The Hermite case revisited

In the scalar frame ($p = 1$), the sum rule gives two different expressions for the divergence between a probability measure and the semicircle distribution (see [13] and [20]). In the more general case $p \geq 2$, it gives two forms for the divergence to the operator version of the semicircle law

$$\text{SC}(dx) = \frac{\sqrt{4-x^2}}{2\pi} 1_{[-2,2]}(x) dx$$

supported by $[\alpha^-, \alpha^+] = [-2, 2]$. We refer to [3] and [19] for this matrix sum rule. We denote by $\Sigma_{\text{SC}} = \text{SC} \cdot I_p$ the matrix version of the semicircle law. The semi-infinite block Jacobi operator of Σ_{SC} has entries

$$B_k^{\text{SC}} = 0_{p,p}, \quad A_k^{\text{SC}} = I_p,$$

for all $k \geq 1$. The spectral side of the sum rule involves a contribution of outlying eigenvalues, for which we define

$$\mathcal{F}_H^+(x) := \begin{cases} \int_2^x \sqrt{t^2 - 4} dt = \frac{x}{2} \sqrt{x^2 - 4} - 2 \log \left(\frac{x + \sqrt{x^2 - 4}}{2} \right) & \text{if } x \geq 2 \\ \infty & \text{otherwise.} \end{cases}$$

and $\mathcal{F}_H^-(x) = \mathcal{F}_H^+(-x)$. Let G be the very popular function (Cramér transform of the exponential distribution)

$$G(x) = x - 1 - \log x.$$

We adopt the convention of the functional calculus, so that for X a $p \times p$ nonnegative Hermitian matrix, we have

$$(2.24) \quad \text{tr} G(X) = \text{tr} X - \log \det X - p.$$

The following remarkable equality holds. This result has been first proven by [3]. In Section 3.4, we give a probabilistic proof. Indeed, we show that this sum rule is a consequence of two large deviation results.

Theorem 2.3 Let $\Sigma \in \mathcal{M}_{p,1}(\mathbb{R})$ be a spectral measure with Jacobi operator (2.12). If $\Sigma \in \mathcal{S}_{p,1}(-2, 2)$, then

$$\mathcal{K}(\Sigma_{SC} | \Sigma) + \sum_{k=1}^{N^+} \mathcal{F}_H^+(\lambda_k^+) + \sum_{k=1}^{N^-} \mathcal{F}_H^-(\lambda_k^-) = \sum_{k=1}^{\infty} \frac{1}{2} \text{tr} B_k^2 + \text{tr} G(A_k A_k^*),$$

where both sides may be infinite simultaneously. If $\Sigma \notin \mathcal{S}_{p,1}(-2, 2)$, the right hand side equals $+\infty$.

We remark that since $\text{tr} G(\sigma A A^* \sigma^*) = \text{tr} G(A A^*)$ for any unitary σ , the value of the right hand side in Theorem 2.3 is independent of the choice of unitary matrices σ_n .

Let us restate the sum rule as in the notation of [20]. For a spectral measure Σ , we define the m -function

$$m(z) = \int \frac{1}{x - z} d\Sigma(x), \quad (z \in \mathbb{C} \setminus \mathbb{R}).$$

For z in the interior of the unit disk the function $M(z) = -m(z + z^{-1})$ can be defined and for almost all $\theta \in [0, 2\pi]$, the limit $M(e^{i\theta}) = \lim_{r \rightarrow 1} M(re^{i\theta})$ exists and is neither vanishing nor infinite. Finally, let

$$(2.25) \quad \mathcal{Q}(\Sigma) = \frac{1}{\pi} \int_0^{2\pi} \log \left(\frac{\sin^p \theta}{\det(\text{Im } M(e^{i\theta}))} \right) \sin^2 \theta d\theta.$$

Then the following statement is the combination of Theorem 4.6.1 and Theorem 4.6.3 of [20].

Theorem 2.4 Let A_n, B_n be the block entries of the block Jacobi operator J and let Σ denote the spectral matrix measure of J . Then

$$\sum_{n=1}^{\infty} \text{tr}((A_n A_n^*)^{1/2} - I)^2 + \text{tr} B_n^2 < \infty$$

if and only if

(a) The essential support of J satisfies

$$\sigma_{\text{ess}}(J) \subset [-2, 2],$$

(b) The eigenvalues $\{\lambda_j\}_{j=1}^{\infty} \notin \sigma_{\text{ess}}(J_{\infty})$ satisfy

$$\sum_{k=1}^{\infty} (|\lambda_k| - 2)^{3/2} < \infty$$

(c) The density $f(x)$ of the absolute continuous part of Σ with respect to the Lebesgue measure multiplied with the identity obeys

$$\int_{-2}^2 \sqrt{4-x^2} \log \det f(x) dx > -\infty$$

In this case, we have

$$(2.26) \quad \sum_{n=1}^{\infty} \left(\frac{1}{2} \text{tr}(B_n^2) + \text{tr}(G(A_n A_n^*)) \right) = \mathcal{Q}(\Sigma) + \sum_{\lambda \notin \sigma_{\text{ess}}(J_{\infty})} F(\lambda).$$

Here, F is defined to be equal to \mathcal{F}_H^+ on $[2, \infty)$ and equal to \mathcal{F}_H^- on $(-\infty, 2]$.

Notice that the integral $\mathcal{Q}(\Sigma)$ appearing in (2.26) can be interpreted as the relative entropy. For a measure Σ supported on $[-2, 2]$, the (inverse) Szegő mapping pushes forward Σ to a symmetric measure Σ_{S_z} on the unit circle such that for all measurable and symmetric f ,

$$\int f(\theta) d\Sigma_{S_z}(\theta) = \int f\left(\arccos\left(\frac{x}{2}\right)\right) d\Sigma(x).$$

Let $f(x)$ denote the density of Σ with respect to the matrix semicircle distribution $\Sigma_{SC} = \text{SC} \cdot I_p$, then a straightforward generalization of the arguments in Lemma 3.5.1 in the book of [20] to the matrix case show

$$\text{Im } M(e^{i\theta}) = \text{Im } m(2 \cos \theta) = \pi f(2 \cos \theta)$$

for $\theta \in [0, \pi]$ and $M(e^{-i\theta}) = -M(e^{i\theta})$. Then, using the symmetry to write $\mathcal{Q}(\Sigma)$ as an integral over $[0, \pi]$ and setting $x = 2 \cos \theta$, we obtain

$$\begin{aligned} \mathcal{Q}(\Sigma) &= \frac{2}{\pi} \int_{-2}^2 \log \left(\frac{2^{-p}(4-x^2)^{-p/2}}{\det \pi f(x)} \right) \frac{1}{4} \sqrt{4-x^2} dx \\ &= \int_{-2}^2 \log \det \left(\frac{1}{2\pi} \sqrt{4-x^2} f(x)^{-1} \right) \frac{1}{2\pi} \sqrt{4-x^2} dx \\ &= \int \log \det \left(\frac{d\Sigma_{SC}}{d\Sigma} \right) d\text{SC} = \mathcal{K}(\Sigma_{SC} | \Sigma). \end{aligned}$$

2.3.2 Our new sum rule: the Laguerre case

In the Laguerre case, the central measure is the matrix Marchenko-Pastur law with scalar version

$$\text{MP}(\tau)(dx) = \frac{\sqrt{(\tau^+ - x)(x - \tau^-)}}{2\pi\tau x} 1_{(\tau^-, \tau^+)}(x) dx,$$

where $\tau \in (0, 1]$, $\alpha^\pm = \tau^\pm = (1 \pm \sqrt{\tau})^2$ and we set $\Sigma_{\text{MP}(\tau)} = \text{MP}(\tau) \cdot I_p$. For the new Laguerre sum rule, we have to replace \mathcal{F}_H^\pm by

$$\mathcal{F}_L^+(x) = \begin{cases} \int_{\tau^+}^x \frac{\sqrt{(t - \tau^-)(t - \tau^+)}}{t\tau} dt & \text{if } x \geq \tau^+, \\ \infty & \text{otherwise.} \end{cases}$$

and

$$\mathcal{F}_L^-(x) = \begin{cases} \int_x^{\tau^-} \frac{\sqrt{(\tau^- - t)(\tau^+ - t)}}{t\tau} dt & \text{if } x \leq \tau^-, \\ \infty & \text{otherwise.} \end{cases}$$

One of our main results is Theorem 2.5. Up to our knowledge, this result is new. The proof is again in Section 3.4.

Theorem 2.5 *Assume the Jacobi matrix J is nonnegative definite and let Σ be the spectral matrix measure associated with J . Then for any $\tau \in (0, 1]$, if $\Sigma \in \mathcal{S}_{p,1}(\tau^-, \tau^+)$,*

$$\mathcal{K}(\Sigma_{\text{MP}(\tau)} | \Sigma) + \sum_{n=1}^{N^+} \mathcal{F}_L^+(\lambda_n^+) + \sum_{n=1}^{N^-} \mathcal{F}_L^-(\lambda_n^-) = \sum_{k=1}^{\infty} \tau^{-1} \text{tr} G(\zeta_{2k-1}) + \text{tr} G(\tau^{-1} \zeta_{2k})$$

where both sides may be infinite simultaneously and ζ_k is defined as in (2.14). If $\Sigma \notin \mathcal{S}_{p,1}(\tau^-, \tau^+)$, the right hand side equals $+\infty$.

Since the matrices ζ_k can be decomposed as in (2.15), they are in fact similar to a Hermitian matrix

$$\zeta_k = h_{k-1}^{-1/2} \left[h_{k-1}^{-1/2} h_k h_{k-1}^{-1/2} \right] h_{k-1}^{1/2},$$

hence the sum on the right hand side in Theorem 2.5 is real valued.

Similar to the matrix gem , Theorem 2.4, we can formulate equivalent conditions on the matrices ζ_k and the spectral measure, which characterize finiteness of the sum. The following corollary is the matrix counterpart of Corollary 2.4 in [9]. It follows immediately from Theorem 2.5, since $\mathcal{F}_L^-(0) = \infty$ and

$$\mathcal{F}_L^\pm(\tau^\pm \pm h) = \frac{4}{3\tau^{3/4}(1 \pm \sqrt{\tau})^2} h^{3/2} + o(h^{3/2}) \quad (h \rightarrow 0^+)$$

and, for H similar to a Hermitian matrix,

$$\text{tr} G(I_p + H) = \frac{1}{2} \text{tr} H^2 + o(\|H\|) \quad (\|H\| \rightarrow 0).$$

Here, $\|\cdot\|$ is any matrix norm.

Corollary 2.6 *Assume the Jacobi matrix J is nonnegative definite and let Σ be the spectral matrix measure of J . Then*

$$(2.27) \quad \sum_{k=1}^{\infty} [\operatorname{tr}(\zeta_{2k-1} - I_p)^2 + \operatorname{tr}(\zeta_{2k} - \tau I_p)^2] < \infty$$

if and only if

1. $\Sigma \in \mathcal{S}_{p,1}(\tau^-, \tau^+)$
2. $\sum_{i=1}^{N^+} (\lambda_i^+ - \tau^+)^{3/2} + \sum_{i=1}^{N^-} (\tau^- - \lambda_i^-)^{3/2} < \infty$ and if $N^- > 0$, then $\lambda_1^- > 0$.
3. the spectral matrix measure Σ of J with Lebesgue decomposition $d\Sigma(x) = f(x)I_p dx + d\Sigma_s(x)$ satisfies

$$\int_{\tau^-}^{\tau^+} \frac{\sqrt{(\tau^+ - x)(x - \tau^-)}}{x} \log \det(f(x)) dx > -\infty.$$

3 Randomization and large deviations

3.1 Matrix random models

The results of the previous section rely on two classical distributions of random Hermitian matrices: the Gaussian (or Hermite) and the Laguerre (or Wishart) ensemble. A random variable X taking values in \mathbb{H}_N is distributed according to the Gaussian unitary ensemble GUE_N , if all real diagonal entries are standard normal distributed and the off-diagonal variables are complex standard normal distributed and all entries on and above the diagonal are independent. Recall that a complex random variable is said to be standard normal distributed if the real and imaginary part are independent normal random variables with mean 0 and variance 1/2. The density of the eigenvalues $\lambda = (\lambda_1, \dots, \lambda_N)$ of the Gaussian ensemble is (see for example [1])

$$(3.1) \quad g_G(\lambda) = c_G \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N e^{-\lambda_i^2/2}.$$

In analogy to the scalar χ^2 distribution, the Laguerre ensemble is the distribution of the square of Gaussian matrices: suppose G denotes a $N \times \gamma$ matrix with independent complex standard normal entries, then GG^* is distributed according to the Laguerre ensemble $\text{LUE}_N(\gamma)$ with parameter γ . If $\gamma \geq N$, the eigenvalues have the density (see for example [1])

$$(3.2) \quad g_L(\lambda) = c_L^\gamma \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N \lambda_i^{\gamma-N} e^{-\lambda_i} 1_{\{\lambda_i > 0\}}.$$

It is a well-known consequence of the invariance under unitary conjugation, that in the classical ensembles (1.4), the array of random eigenvalues and the random eigenvector (unitary) matrix are independent. Further, this latter matrix is Haar distributed ([5]). This implies the following equality in distribution for the weights (see Proposition 3.1 in [11]).

Lemma 3.1 *Let U be Haar distributed on the set of $N \times N$ unitary matrices and define $u_j = (U_{i,j})_{1 \leq i \leq p} \in \mathbb{C}^p$. Let further v_1, \dots, v_N be independent complex standard normal distributed vectors in \mathbb{C}^p . Then $(u_1 u_1^*, \dots, u_N u_N^*)$ has the same distribution as*

$$(H^{-1/2} v_1 v_1^* H^{-1/2}, \dots, H^{-1/2} v_N v_N^* H^{-1/2}),$$

where $H = \sum_{k=1}^n v_k v_k^*$.

In the following, we will denote the joint distribution of the matrix weights appearing in Lemma 3.1 by $\mathbb{D}_{p,N}$. Notice that this distribution is a kind of matrix Dirichlet law. Our first large deviation principle will hold for a general class of $p \times p$ matrix measures. We draw the random eigenvalues $\lambda_1, \dots, \lambda_N$ with the absolute continuous distribution \mathbb{P}_N^V where this law has Lebesgue density

$$(3.3) \quad \frac{d\mathbb{P}_N^V(\lambda)}{d\lambda} = \frac{1}{Z_V^N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N e^{-NV(\lambda_i)}.$$

We suppose that the potential V is continuous and real valued on the interval (b^-, b^+) ($-\infty \leq b^- < b^+ \leq +\infty$), infinite outside of $[b^-, b^+]$ and $\lim_{x \rightarrow b^\pm} V(x) = V(b^\pm)$ with possible limit $V(b^\pm) = +\infty$. Under the assumption

$$(A1) \text{ Confinement: } \quad \liminf_{x \rightarrow b^\pm} \frac{V(x)}{\log|x|} > 2,$$

the empirical distribution $\mu_n^{(N)}$ of eigenvalues $\lambda_1, \dots, \lambda_N$ has a limit μ_V in probability, which is the unique minimizer of

$$(3.4) \quad \mu \mapsto \mathcal{E}(\mu) := \int V(x) d\mu(x) - \iint \log|x-y| d\mu(x) d\mu(y).$$

μ_V has compact support (see [12] or [1]). Indeed, this is a consequence of the LDP for the empirical spectral measure. We need two additional assumptions on μ_V :

(A2) One-cut regime: the support of μ_V is a single interval $[\alpha^-, \alpha^+] \subset [b^-, b^+]$ ($\alpha^- < \alpha^+$).

(A3) Control (of large deviations): the effective potential

$$(3.5) \quad \mathcal{J}_V(x) := V(x) - 2 \int \log|x - \xi| d\mu_V(\xi)$$

achieves its global minimum value on $(b^-, b^+) \setminus (\alpha^-, \alpha^+)$ only on the boundary of this set.

In the Hermite case, we have $V(x) = \frac{1}{2}x^2$ and the equilibrium measure μ_V is the semicircle law. In the Laguerre case, we may set $V(x) = \tau^{-1}x - (\tau^{-1} - 1) \log(x)$ for $\tau \in (0, 1]$ and $V(x) = +\infty$ for negative x . In this case, μ_V is the Marchenko-Pastur law $\text{MP}(\tau)$. In both the Hermite and the Laguerre case, the assumption (A1), (A2) and (A3) are satisfied. We need one more definition related to outlying eigenvalues:

$$(3.6) \quad \mathcal{F}_V^+(x) = \begin{cases} \mathcal{J}_V(x) - \inf_{\xi \in \mathbb{R}} \mathcal{J}_V(\xi) & \text{if } \alpha^+ \leq x \leq b^+, \\ \infty & \text{otherwise,} \end{cases}$$

$$(3.7) \quad \mathcal{F}_V^-(x) = \begin{cases} \mathcal{J}_V(x) - \inf_{\xi \in \mathbb{R}} \mathcal{J}_V(\xi) & \text{if } b^- \leq x \leq \alpha^-, \\ \infty & \text{otherwise.} \end{cases}$$

One may check that in the Hermite case, $\mathcal{F}_V^\pm = \mathcal{F}_H^\pm$ and in the Laguerre case, $\mathcal{F}_V^\pm = \mathcal{F}_L^\pm$, where \mathcal{F}_H^\pm and \mathcal{F}_L^\pm have been defined in the previous section.

3.2 Large deviations: some useful lemmas

In order to be self-contained, let us recall the definition of large deviations:

Definition 3.2 *Let E be a topological Hausdorff space and let $\mathcal{I} : E \rightarrow [0, \infty]$ be a lower semi-continuous function. We say that a sequence $(P_n)_n$ of probability measures on $(E, \mathcal{B}(E))$ satisfies a large deviation principle (LDP) with rate function \mathcal{I} and speed a_n if:*

(i) *For all closed sets $F \subset E$:*

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log P_n(F) \leq - \inf_{x \in F} \mathcal{I}(x)$$

(ii) *For all open sets $O \subset E$:*

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \log P_n(O) \geq - \inf_{x \in O} \mathcal{I}(x)$$

The rate function \mathcal{I} is good if its level sets $\{x \in E \mid \mathcal{I}(x) \leq a\}$ are compact for all $a \geq 0$. We say that a sequence of E -valued random variables satisfies a LDP if their distributions satisfy a LDP.

We will frequently use the following principle (see e.g. [6], Chapter 4 p. 126).

Contraction principle. *Suppose that $(P_n)_n$ satisfies an LDP on $(E, \mathcal{B}(E))$ with good rate function \mathcal{I} and speed a_n . Let f be a continuous mapping from E to another topological Hausdorff space F . Then $P_n \circ f^{-1}$ satisfies a LDP on $(F, \mathcal{B}(F))$ with speed a_n and good rate function*

$$\mathcal{I}'(y) = \inf_{\{x \in E | f(x) = y\}} \mathcal{I}(x), \quad (y \in F).$$

To prove our main large deviation principle, we will use the Baldi Theorem recalled below (see [2]). To apply it in our setting, we remark that the topological dual of $\mathcal{M}_p(T)$ is the space $\mathcal{C}_p(T)$ of bounded measurable functions $f : T \rightarrow \mathbb{H}_p$ with the pairing

$$\langle \Sigma, f \rangle = \text{tr} \int f d\Sigma.$$

Theorem 3.3 *Let F be a topological vector space with dual F^* . Suppose the sequence of random variables $(X_n)_n$ is exponentially tight in F and that the limit*

$$\Lambda(Y) := \lim_{n \rightarrow \infty} \frac{1}{n} \log E [\exp \{n \langle X_n, Y \rangle\}], \quad Y \in F^*$$

exists, is finite in a neighborhood of the origin and lower semicontinuous. If the set where the Legendre transform Λ^ of Λ is strictly convex is dense in $\{X : \Lambda^*(X) < \infty\}$, then (X_n) satisfies the LDP with speed n and rate function Λ^* .*

3.3 Results part II: Large deviations

3.3.1 Large deviation principles

Our first LDP holds for $p \times p$ matrix measures

$$\Sigma^{(N)} = \sum_{k=1}^N W_k \delta_{\lambda_k},$$

when the distribution of the support points is \mathbb{P}_N^V and the distribution of weights is $\mathbb{D}_{p,N}$ as in the case of classical ensembles. As explained in the introduction, this is precisely the distribution of the spectral measure of an $N \times N$ matrix X_N , chosen according to the density (1.4). Recall that

under assumption (A1), the empirical measure of the eigenvalues converges to an equilibrium measure μ_V , supported by $[\alpha^-, \alpha^+]$. The rate function of our large deviation principle involved the reference matrix measure

$$\Sigma_V = I_p \cdot \mu_V.$$

We recall that \mathcal{F}_V^\pm has been defined in (3.6) and (3.7). The following theorem is the matrix counterpart of Theorem 3.1 in [9]. Note that in the scalar case, we have an additional parameter $\beta > 0$, corresponding to the inverse temperature of the log-gas. In the matrix case, we need to fix $\beta = 2$ (for complex matrices) due to the nature of the matrix spaces.

Theorem 3.4 *Assume that the potential V satisfies the assumptions (A1), (A2) and (A3). Then the sequence of spectral measures $\Sigma^{(N)}$ under $\mathbb{P}_N^V \otimes \mathbb{D}_{p,N}$ satisfies the LDP with speed N and rate function*

$$\mathcal{I}_V(\Sigma) = \mathcal{K}(\Sigma_V | \Sigma) + \sum_{k=1}^{N^+} \mathcal{F}_V^+(\lambda_k^+) + \sum_{k=1}^{N^-} \mathcal{F}_V^-(\lambda_k^-)$$

if $\Sigma \in \mathcal{S}_{p,1}(\alpha^-, \alpha^+)$ and $\mathcal{I}_V(\Sigma) = \infty$ otherwise.

Remark 3.5 *A natural generalization of Theorem 3.4 holds for potentials $V = V_N$ depending on N , provided that V_N converges to a deterministic potential V in an appropriate sense. For example, it holds if we suppose that $V_N : \mathbb{R} \rightarrow (-\infty, +\infty]$ is a sequence of potentials converging to V uniformly on the level sets $\{V \leq M\}$, where V satisfies assumptions (A1), (A2) and (A3) and such that $V_N(x) \geq V(x)$.*

In the special cases of Hermite and Laguerre ensemble we can prove further LDPs for the spectral measure, independently of Theorem 3.4. They are in the subset $\mathcal{M}_{p,1,c}$ of compactly supported normalized matrix measures. Since we need a specific block structure, we assume $N = np$.

Theorem 3.6 *Let $\Sigma^{(n)}$ be the spectral measure of $\frac{1}{\sqrt{np}}X_n$. Assume that X_n is distributed according to the Hermite ensemble GUE_N ($N = np$). Then the sequence $(\Sigma^{(n)})_n$ satisfies the LDP in $\mathcal{M}_{p,1,c}(\mathbb{R})$ with speed pn and rate function*

$$\mathcal{I}_H(\Sigma) = \sum_{k=1}^{\infty} \left[\frac{1}{2} \text{tr} B_k^2 + \text{tr} G(A_k A_k^*) \right]$$

where B_k, A_k are the recursion coefficients of Σ as in (2.8).

Theorem 3.7 *Let Y_n be distributed according to the Laguerre ensemble $\text{LUE}_N(p\gamma_n)$, with $(N = np)$, $\gamma_n \geq n$ an integer sequence such that $\frac{n}{\gamma_n} \rightarrow \tau \in (0, 1]$ and let $\Sigma^{(n)}$ be the spectral measure of $\frac{1}{p\gamma_n}Y_n$ with a decomposition of recursion coefficients as in Section 2. Then the sequence $(\Sigma^{(n)})_n$ satisfies the LDP in $\mathcal{M}_{p,1,c}([0, \infty))$ with speed pn and rate function*

$$\mathcal{I}_L(\Sigma) = \sum_{k=1}^{\infty} [\tau^{-1}G(\zeta_{2k-1}) + G(\tau^{-1}\zeta_{2k})],$$

with ζ_k as in (2.15). If Σ is a trivial measure, we have $\mathcal{I}_L(\Sigma) = \infty$.

In order to prove LDPs for the spectral measures in terms of the recursion coefficients we need the following results for matrices of fixed size. The first and third are straightforward generalizations of the scalar case, the second one can be found in [10], with small modifications to allow a general sequence of parameters.

Lemma 3.8 (i) *If $X \sim \text{GUE}_p$ with p fixed, then the sequence $(\frac{1}{\sqrt{n}}X)_n$ satisfies the LDP with speed n and good rate function*

$$\mathcal{I}_1(X) = \frac{1}{2} \text{tr } X^2.$$

(ii) *Let $Y_n \sim \text{LUE}_p(\gamma_n)$ with a positive sequence $(\gamma_n)_n$ such that $\frac{\gamma_n}{n} \rightarrow \gamma > 0$, then the sequence $(\frac{1}{n}Y_n)_n$ satisfies the LDP with speed n and good rate function*

$$\mathcal{I}_2(Y) = \gamma \text{tr } G(\gamma^{-1}Y)$$

if Y is Hermitian and nonnegative and $\mathcal{I}_2(Y) = \infty$ otherwise.

(iii) *Let $Z \sim \text{LUE}_p(1)$ with p fixed, that is, $Z = vv^*$ when v is a vector of independent complex standard normal random variables. Then the sequence $(\frac{1}{n}Z)_n$ satisfies the LDP with speed n and good rate function*

$$\mathcal{I}_3(Z) = \text{tr } Z$$

if Z is Hermitian and nonnegative and $\mathcal{I}_3(Z) = \infty$ otherwise.

3.4 From large deviations to sum rules

Proof of Theorem 2.3:

Consider the matrix measure $\Sigma^{(n)}$ with $N = np$ support points with density

$$\frac{1}{Z_V^N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N e^{-N\lambda_i^2/2}$$

and weight distribution $\mathbb{D}_{p,N}$ independent of the support points. By Theorem 3.4, the sequence $(\Sigma^{(n)})_n$ satisfies the LDP with speed np and rate function \mathcal{I}_V , where μ_V is the semicircle law and furthermore $\mathcal{F}_V^\pm = \mathcal{F}_H^\pm$. That is, the rate function is precisely the left hand side of the equation in Theorem 2.3. On the other hand, $\Sigma^{(n)}$ is also the spectral measure of the random matrix $\frac{1}{\sqrt{pn}}X_n$, where $X_n \sim \text{GUE}_{pn}$. By Theorem 3.6, the sequence $(\Sigma^{(n)})_n$ satisfies also the LDP in the space of compactly supported measures with speed np and rate function \mathcal{I}_H , the right hand side of the equation in Theorem 2.3. Since a large deviation rate function is unique, we must have $\mathcal{I}_V(\Sigma) = \mathcal{I}_H(\Sigma)$ for any compactly supported $\Sigma \in \mathcal{M}_{p,1,c}$. If Σ is not compactly supported, it suffices to remark that the recursion coefficients cannot satisfy $\sup_n (\|A_n\| + \|B_n\|) < \infty$, as otherwise J would be a bounded operator. But $\text{tr}B^2 + \text{tr}G(A)$ diverges as $\|A\| \rightarrow \infty$ or $\|B\| \rightarrow \infty$ and so the right hand side in Theorem 2.3 equals $+\infty$. \square

Proof of Theorem 2.5:

Fix $\tau \in (0, 1]$ and let $V(x) = \tau^{-1}x - (\tau^{-1} - 1)\log(x)$ for $x \geq 0$ and $V(x) = +\infty$ if $x < 0$. From Theorem 3.4 we get that under the distribution $\mathbb{P}_N^V \otimes \mathbb{D}_{p,N}$ the sequence $(\Sigma^{(n)})_n$ satisfies the LDP with speed $N = np$ and rate function \mathcal{I}_V . In this case, the equilibrium measure is the Marchenko-Pastur law $\text{MP}(\tau)$ multiplied by I_p . Further, we have $\mathcal{F}_V^\pm = \mathcal{F}_L^\pm$. So that, \mathcal{I}_V is nothing more than the left hand side of the sum rule in Theorem 2.5. We would like to combine this result with the LDP in Theorem 3.7, but since this requires integer parameters, we need to modify the potential slightly. Define $\gamma_n = \lceil n\tau^{-1} \rceil$ and consider the eigenvalue distribution with density

$$(3.8) \quad \frac{1}{Z_V^N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N \lambda_i^{p\gamma_n - pn} e^{-p\gamma_n \lambda_i} 1_{[0,\infty)}(\lambda_i)$$

This is the eigenvalue distribution of the matrix $\frac{1}{p\gamma_n}Y_n$, when $Y_n \sim \text{LUE}_{pn}(p\gamma_n)$. By Theorem 3.7, the spectral measure of this matrix satisfies the LDP with speed pn and rate function \mathcal{I}_L which is the right hand side of the sum rule in Theorem 2.5. We may as well write the eigenvalue

density (3.8) as

$$(3.9) \quad \frac{1}{Z_V^N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N e^{-pnV_n(\lambda_i)},$$

where

$$(3.10) \quad V_n(x) = \frac{\lceil n\tau^{-1} \rceil}{n} x - \left(\frac{\lceil n\tau^{-1} \rceil}{n} - 1 \right) \log(x)$$

for nonnegative x . Then $V_n(x) \geq V(x)$ for all x and on the sets $\{x \mid V(x) \leq M\}$, ($M > 0$). The potentials V_n converge uniformly to V . Note that the point 0 is included in the level sets of V only if $\tau = 1$. Therefore, by Remark 3.5, the spectral measure with support point density (3.9) satisfies the same LDP as under the density \mathbb{P}_{pn}^V and then with rate function \mathcal{I}_V . This yields $\mathcal{I}_V(\Sigma) = \mathcal{I}_H(\Sigma)$ for any compactly supported measure Σ . The extension to measures with non-compact support follows as in the proof of Theorem 2.3. \square

4 Spectral LDP for a general potential

This section is devoted to the proof of Theorem 3.4. The proof follows the track of the one developed for the scalar case in [9]. We will often refer to this paper for more details. The main idea is to apply the projective method and study a family of matrix measures restricted to the support $I = [\alpha^-, \alpha^+]$ of the equilibrium measure and a fixed number of extremal eigenvalues. For $\Sigma \in \mathcal{S}_p$ with

$$(4.1) \quad \Sigma = \Sigma|_I + \sum_{i=1}^{N^+} \Gamma_i^+ \delta_{\lambda_i^+} + \sum_{i=1}^{N^-} \Gamma_i^- \delta_{\lambda_i^-}$$

we define the j -th projector π_j by

$$\pi_j(\Sigma) = \Sigma|_I + \sum_{i=1}^{N^+ \wedge j} \Gamma_i^+ \delta_{\lambda_i^+} + \sum_{i=1}^{N^- \wedge j} \Gamma_i^- \delta_{\lambda_i^-},$$

that is, all but the j -th largest and smallest eigenvalues outside of $I = [\alpha^-, \alpha^+]$ are omitted. Note that π_j is not continuous in the weak topology. For this reason we need to change our topology on \mathcal{S}_p by identifying Σ as in (4.1) with the vector

$$(4.2) \quad (\Sigma|_I, (\lambda_i^+)_{i \geq 1}, (\lambda_i^-)_{i \geq 1}, (\Gamma_i^+)_{i \geq 1}, (\Gamma_i^-)_{i \geq 1})$$

with $\lambda_i^+ = \alpha^+$ and $\Gamma_i^+ = 0$ if $i > N^+$ and $\lambda_i^- = \alpha^-$ and $\Gamma_i^- = 0$ if $i > N^-$. Then we say that $\Sigma^{(n)}$ converges to Σ if:

$$(4.3) \quad \begin{aligned} & \Sigma_{|I}^{(n)} \xrightarrow[n \rightarrow \infty]{} \Sigma_{|I} \text{ weakly and for every } i \geq 1 \\ & (\lambda_i^+(\Sigma^{(n)}), \lambda_i^-(\Sigma^{(n)}), \Gamma_i^+(\Sigma^{(n)}), \Gamma_i^-(\Sigma^{(n)})) \xrightarrow[n \rightarrow \infty]{} (\lambda_i^+(\Sigma), \lambda_i^-(\Sigma), \Gamma_i^+(\Sigma), \Gamma_i^-(\Sigma)). \end{aligned}$$

Analogously to Lemma 4.5 in [9], one can show that on the smaller set $\mathcal{S}_{p,1}$ of normalized measures, this topology is (strictly) stronger than the weak topology.

Let for j fixed and $n > 2j$

$$\lambda^+(j) = (\lambda_1^+, \dots, \lambda_j^+), \quad \lambda^-(j) = (\lambda_1^-, \dots, \lambda_j^-).$$

Then the following joint LDP holds for the largest and/or smallest eigenvalues, where we write $\mathbb{R}^{\uparrow j}$ (resp. $\mathbb{R}^{\downarrow j}$) for the subset of \mathbb{R}^j of all vectors with non-decreasing (resp. non-increasing) entries and, slightly abusing the notation, write α^\pm for the vector $(\alpha^\pm, \dots, \alpha^\pm) \in \mathbb{R}^j$.

Theorem 4.1 *Let j be a fixed integer and the potential V such that (A1), (A2) and the control condition (A3) are satisfied.*

1. *If $b^- < \alpha^-$ and $\alpha^+ < b^+$, then the law of $(\lambda^+(j), \lambda^-(j))$ under \mathbb{P}_N^V satisfies the LDP in \mathbb{R}^{2j} with speed N and rate function*

$$\mathcal{I}_{\lambda^\pm}(x^+, x^-) := \begin{cases} \sum_{k=1}^j \mathcal{F}_V^+(x_k^+) + \sum_{k=1}^j \mathcal{F}_V^-(x_k^-) & \text{if } (x_1^+, \dots, x_j^+) \in \mathbb{R}^{\downarrow j} \text{ and } (x_1^-, \dots, x_j^-) \in \mathbb{R}^{\uparrow j} \\ \infty & \text{otherwise.} \end{cases}$$

2. *If $b^- = \alpha^-$, but $\alpha^+ < b^+$, the law of $\lambda^+(j)$ satisfies the LDP with speed N and rate function*

$$\mathcal{I}_{\lambda^+}(x^+) = \mathcal{I}_{\lambda^\pm}(x^+, \alpha^-) = \begin{cases} \sum_{k=1}^j \mathcal{F}_V^+(x_k^+) & \text{if } (x_1^+, \dots, x_j^+) \in \mathbb{R}^{\downarrow j} \\ \infty & \text{otherwise.} \end{cases}$$

3. *If $b^- < \alpha^-$, but $\alpha^+ = b^+$, the law of $\lambda^-(j)$ satisfies the LDP with speed N and rate function*

$$\mathcal{I}_{\lambda^-}(x^-) = \mathcal{I}_{\lambda^\pm}(\alpha^+, x^-) = \begin{cases} \sum_{k=1}^j \mathcal{F}_V^-(x_k^-) & \text{if } (x_1^-, \dots, x_j^-) \in \mathbb{R}^{\uparrow j} \\ \infty & \text{otherwise.} \end{cases}$$

4.1 LDP for the restricted measure and extremal eigenvalues

Suppose now that the distribution of $\Sigma^{(N)}$ is as in Theorem 3.4 and the assumptions (A1), (A2) and (A3) are satisfied. By Lemma 3.1, we may decouple the weights and consider the (non-normalized) measure

$$(4.4) \quad \tilde{\Sigma}^{(N)} = \frac{1}{N} \sum_{k=1}^N v_k v_k^* \delta_{\lambda_k},$$

where the entries of $v_1, \dots, v_N \in \mathbb{C}^p$ are independent complex standard normal distributed random vectors and we denote the joint distribution of the v_k 's by $\mathbb{G}_{p,N}$. The original distribution can then be recovered as the pushforward under

$$\tilde{\Sigma} \mapsto \tilde{\Sigma}(\mathbb{R})^{-1/2} \cdot \tilde{\Sigma} \cdot \tilde{\Sigma}(\mathbb{R})^{-1/2}.$$

Let $I(j) := I \setminus \{\lambda_1^+, \lambda_1^-, \dots, \lambda_j^+, \lambda_j^-\}$ denote the interval I minus the j largest and smallest eigenvalues. Analogously, let $I^+(j) := I \setminus \{\lambda_1^+, \dots, \lambda_j^+\}$ and $I^-(j) := I \setminus \{\lambda_1^-, \dots, \lambda_j^-\}$. Then we write $\tilde{\Sigma}_{I(j)}^{(N)}$ for the restriction of $\tilde{\Sigma}^{(N)}$ to $I(j)$. We use the analogous notation for the restrictions to $I^+(j), I^-(j)$ and I . The main result in this section is a joint LDP for the restricted measure and the collection of extremal eigenvalues.

Theorem 4.2 *Suppose that the law of eigenvalues and weights is given by $\mathbb{P}_N^V \otimes \mathbb{G}_{p,N}$ and consider $\tilde{\Sigma}^{(N)}$ as a random element in \mathcal{S}_p with topology (4.3).*

1. *If $b^- < \alpha^- < \alpha^+ < b^+$, then for any fixed $j \in \mathbb{N}$, the sequence $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda^+(j), \lambda^-(j))$ satisfies the joint LDP with speed N and rate function*

$$\mathcal{I}(\Sigma, x^+, x^-) = \mathcal{K}(\Sigma_V | \Sigma) + \text{tr} \Sigma(\mathbb{R}) - p + \mathcal{I}_{\lambda^\pm}(x^+, x^-)$$

2. *If $b^- = \alpha^-$, but $\alpha^+ < b^+$ (or $b^+ = \alpha^+$, but $\alpha^- > b^-$), then, with the same notation as in Theorem 4.1, $(\tilde{\Sigma}_{I^+(j)}^{(N)}, \lambda^+(j))$ (or $(\tilde{\Sigma}_{I^-(j)}^{(N)}, \lambda^-(j))$ respectively,) satisfies the LDP with speed N and rate function*

$$\mathcal{I}^+(\Sigma, x^+) = \mathcal{I}(\Sigma, x^+, \alpha^-) \quad (\text{or } \mathcal{I}^-(\Sigma, x^-) = \mathcal{I}(\Sigma, \alpha^+, x^-) \text{ respectively}).$$

Proof: We show here only the first part of the theorem, for the other cases just omit the largest or smallest eigenvalues. To begin with, for $M > \max\{|\alpha^+|, |\alpha^-|\}$, let $\lambda_{M,i}^+(j)$ (resp. $\lambda_{M,i}^-(j)$) be the collection of truncated eigenvalues

$$\lambda_{M,i}^+ = \min\{\lambda_i^+, M\} \quad (\text{resp. } \lambda_{M,i}^- = \max\{\lambda_i^-, -M\}),$$

for $i = 1, \dots, j$. To further simplify the notations, set $\lambda_M^\pm(j) := (\lambda_{M,1}^+, \dots, \lambda_{M,j}^+, \lambda_{M,1}^-, \dots, \lambda_{M,j}^-)$. Let f be a continuous function from \mathbb{R} to \mathbb{H}_p such that $\|f\|_\infty < 1$. The sequence $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda_M^\pm(j))$ is exponentially tight, since, with the compact set

$$K_{\gamma,M} = \{(\Sigma, \lambda) \in \mathcal{M}_p(I) \times \mathbb{R}^{2j} \mid \Sigma(I) \leq \gamma \cdot I_p, \lambda \in [-M, M]^{2j}\},$$

we have

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log P \left((\tilde{\Sigma}_{I(j)}^{(N)}, \lambda_M^\pm(j)) \notin K_{\gamma,M} \right) \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log P \left(\frac{1}{N} \sum_{k=1}^N v_k v_k^* > \gamma I_p \right) \leq -c\gamma,$$

where we used the fact that $\sum_{k=1}^N v_k v_k^*$ follows the $\text{LUE}_p(N)$ distribution. For $s^\pm \in \mathbb{R}^{2j}$, we consider the joint moment generating function

$$\mathcal{G}_N(f, s^\pm) = \mathbb{E} \left[\exp \left\{ N \left(\text{tr} \int f d\tilde{\Sigma}_{I(j)}^{(N)} + \langle s^\pm, \lambda_M^\pm(j) \rangle \right) \right\} \right].$$

Since the weights $v_k v_k^*$ of $\tilde{\Sigma}_{I(j)}^{(N)}$ are independent, we may first integrate with respect to the v_k 's, so that

$$(4.5) \quad \mathcal{G}_N(f, s^\pm) = \mathbb{E} \left[\exp \left(N \langle s^\pm, \lambda_M^\pm(j) \rangle \right) \prod_{i: \lambda_i \in I(j)} \mathbb{E} [\exp \{ \text{tr}(f(\lambda_i) v_k v_k^*) \} \mid \lambda_1, \dots, \lambda_n] \right]$$

$$(4.6) \quad = \mathbb{E} \left[\exp \left(N \langle s^\pm, \lambda_M^\pm(j) \rangle \right) \prod_{i: \lambda_i \in I(j)} \mathbb{E} [\exp \{ v_k^* f(\lambda_i) v_k \} \mid \lambda_1, \dots, \lambda_n] \right].$$

Now, it is clear that for v a standard normal complex vector in \mathbb{C}^p and $A \in \mathbb{H}_p$ such that $\|A\| < 1$, we have

$$(4.7) \quad \log \mathbb{E} [\exp (v^* A v)] = -\log \det(I_p - A) =: L(A),$$

so that (4.5) becomes

$$\mathcal{G}_N(f, s^\pm) = \mathbb{E} \left[\exp \left\{ N \left(\mu_{\mathbf{u}, I(j)}^{(N)}(L \circ f) + \langle s^\pm, \lambda_M^\pm(j) \rangle \right) \right\} \right],$$

where $\mu_{\mathbf{u}, I(j)}^{(N)}$ is the restriction of the (scalar) empirical eigenvalue distribution to $I(j)$. It remains to calculate the expectation with respect to \mathbb{P}_N^V .

By Theorem 4.1, the extremal eigenvalues $\lambda^\pm(j)$ of the spectral measure satisfy the LDP with speed N rate function $\mathcal{I}_{\lambda^\pm}$. By the contraction principle (see [6] p.126), the truncated eigenvalues satisfy the LDP with rate function

$$\mathcal{I}_{M, \lambda^\pm}(x^\pm) = \begin{cases} \mathcal{I}_{\lambda^\pm}(x^+, x^-) & \text{if } x^\pm = (x^+, x^-) \in [-M, M]^{2j}, \\ \infty & \text{otherwise.} \end{cases}$$

Since the truncated eigenvalues are bounded, we can conclude from Varadhan's Integral Lemma ([6] p. 137)

$$(4.8) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} [\exp \{N \langle s^\pm, \lambda_M^\pm(j) \rangle\}] = \mathcal{I}_{M, \lambda^\pm}^*(s^\pm),$$

where

$$\mathcal{I}_{M, \lambda^\pm}^*(s^\pm) = \sup_{x^\pm \in \mathbb{R}^{2j}} \{ \langle s^\pm, x^\pm \rangle - \mathcal{I}_{M, \lambda^\pm}(x^\pm) \}$$

is the convex dual of $\mathcal{I}_{M, \lambda^\pm}$. Since $\mu_{\text{u}}^{(N)}$ satisfies a LDP with speed N^2 , but we consider only the slower scale at speed N , we may replace it by the limit measure μ_V at a negligible cost. For the exact estimates, we may follow along the lines of [9] to conclude

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{G}_N(f, s^\pm) &= \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} [\exp \{N (\mu_V(L \circ f) + \langle s^\pm, \lambda_M^\pm(j) \rangle)\}] \\ &= \mu_V(L \circ f) + \mathcal{I}_{M, \lambda^\pm}^*(s^\pm) =: \mathcal{G}(f, s^\pm). \end{aligned}$$

The convex dual of \mathcal{G} is given by

$$\begin{aligned} \mathcal{G}^*(\Sigma, \lambda^\pm) &= \sup_{f \in C_p(I), s^\pm \in \mathbb{R}^{2j}} \left(\text{tr} \int f d\Sigma + \langle \lambda^\pm, s^\pm \rangle - \mathcal{G}(f, s^\pm) \right) \\ &= \sup_{f \in C_p(I)} \left(\text{tr} \int f d\Sigma - \mu_V(L \circ f) \right) + \sup_{s^\pm \in \mathbb{R}^{2j}} (\langle \lambda^\pm, s^\pm \rangle - \mathcal{I}_{M, \lambda^\pm}^*(s^\pm)) \\ &= \Lambda^*(\Sigma) + \mathcal{I}_{M, \lambda^\pm}(\lambda^\pm), \end{aligned}$$

where Λ^* is the convex dual of $\mu_V(L \circ \cdot)$. Theorem 5 of [17] identifies Λ^* as

$$(4.9) \quad \Lambda^*(\Sigma) = \int L^*(h) d\mu_V + \int r \left(\frac{d\Sigma^S}{d\theta} \right) d\theta,$$

where:

- L^* is the convex dual of L
- r its recession function
- the Lebesgue-decomposition of Σ with respect to μ_V is

$$d\Sigma(x) = h(x) d\mu_V(x) + d\Sigma^S(x)$$

- θ is any scalar measure such that Σ^S is absolutely continuous with respect to θ .

We begin by calculating L^* and r . By definition,

$$L^*(X) = \sup_{Y \in \mathbb{H}_p} \{\text{tr}(XY) - L(Y)\}.$$

The recession function is

$$r(X) = \sup \{\text{tr}(XW) \mid \|W\|_\infty < 1\}.$$

The function L is convex (as in the scalar case, apply Hölder's inequality in the definition (4.7)) and analytic. The supremum is then reached at a critical value. We denote by $\mathcal{D}[F(Y)]$ the Fréchet derivative of a function $F: \mathbb{H}_p \rightarrow \mathbb{R}$ at Y and look for Y such that

$$(4.10) \quad \mathcal{D}[\text{tr}(XY) - L(Y)](Z) = 0$$

for every Z . It is well known that, as functions of Y for X fixed, $\mathcal{D}[\text{tr}(XY)](Z) = \text{tr}(XZ)$ and $\mathcal{D}[\log \det Y](Z) = \text{tr}(Y^{-1}Z)$ so that (4.10) becomes, by the chain rule,

$$\text{tr}(XZ) - \text{tr}((I_p - Y)^{-1}Z) = 0$$

for every Z i.e. $X - (I - Y)^{-1} = 0$ hence $Y = I_p - X^{-1}$ and

$$(4.11) \quad L^*(X) = \text{tr}(X - I_p) + \log \det(X^{-1}) = \text{tr}X - p + \log \det(X^{-1}) = \text{tr}G(X).$$

If X has a negative eigenvalue, then $r(X) = \infty$. For X nonnegative definite, the supremum is attained for $W = I_p$, so that

$$(4.12) \quad r(X) = \text{tr}X$$

Gathering (4.11) and (4.12) and plugging into (4.9) we get

$$\begin{aligned} \Lambda^*(\Sigma) &= \text{tr} \int h d\mu_V - \int \log \det h d\mu_V - p + \text{tr} \int d\Sigma^S \\ &= \mathcal{K}(\Sigma_V | \Sigma) + \text{tr} \Sigma(I) - p. \end{aligned}$$

It remains to show that \mathcal{G}^* is strictly convex on a set of points that is dense in the set of all points where \mathcal{G}^* is finite, then Theorem 1.1 of [2] yields the LDP for $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda_M^\pm(j))$ with rate function \mathcal{G}^* . We already know that $\mathcal{I}_{M, \lambda^\pm}$ is convex, so it suffices to check Λ^* for convexity.

Λ^* is strictly convex at Σ if there exists $f \in C_p([\alpha^-, \alpha^+])$, called the exposing hyperplane, such that

$$(4.13) \quad \text{tr} \int f d\Sigma - \Lambda^*(\Sigma) > \text{tr} \int f d\zeta - \Lambda^*(\zeta)$$

for all $\zeta \neq \Sigma$. Let $\Sigma = h_\Sigma \mu_V$ be absolutely continuous with respect to μ_V with positive definite continuous density h_Σ and choose

$$f = I_p - h_\Sigma^{-1}.$$

Let $\zeta = h_\zeta \mu_V + \zeta^S$ the Lebesgue-decomposition of ζ . Recalling the representation (4.9) and (4.11), inequality (4.13) is satisfied as soon as

$$(4.14) \quad \int \log \det h_\Sigma d\mu_V - p > \int \log \det h_\zeta d\mu_V - \text{tr} \int h_\Sigma^{-1} d\zeta$$

Since $\text{tr} \int h_\Sigma^{-1} d\zeta \geq \text{tr} \int h_\Sigma^{-1} h_\zeta d\mu_V$, it is enough to prove

$$(4.15) \quad \int \log \det h_\Sigma d\mu_V - p > \int \log \det h_\zeta d\mu_V - \text{tr} \int h_\Sigma^{-1} h_\zeta d\mu_V$$

This inequality follows from

$$(4.16) \quad \log \det A - \log \det B > p - \text{tr}(A^{-1}B)$$

for Hermitian positive $A \neq B$. In order to prove (4.16), write

$$(4.17) \quad \log \det A - \log \det B = \sum_{i=1}^p \log(\lambda_i(A)\lambda_i(B)^{-1}) \geq p - \sum_{i=1}^p \lambda_i(A^{-1})\lambda_i(B).$$

with $\lambda_i(A), \lambda_i(B)$ the eigenvalues of A, B written in any order. If we choose to order the eigenvalues of A^{-1} in decreasing order (i.e. those of A increasing) and those of B in increasing order, it follows from the Hardy-Littlewood rearrangement inequality (see [15]) that

$$\sum_{i=1}^p \lambda_i(A^{-1})\lambda_i(B) \leq \text{tr}(A^{-1}B).$$

With this ordering of eigenvalues, (4.17) is strict unless A, B have the same eigenvalues. If all eigenvalues of A and B coincide, then the left hand side of (4.16) is 0, while the right hand side is $p - \text{tr}H$ with $\det H = 1$. The minimum value of $\text{tr}H$ is p which is achieved only for $H = I_p$, in which case $A = B$. We get that Λ^* is strictly convex at all points $\Sigma = h\mu_V$ with h positive definite and continuous.

It remains to show that the points at which the rate function is strictly convex are dense in $\mathcal{M}_p([\alpha^-, \alpha^+])$. For a given $\Sigma \in \mathcal{M}_p([\alpha^-, \alpha^+])$, we divide $[\alpha^-, \alpha^+]$ by dyadic points into intervals $I_{k,n} = [\alpha^- + (k-1)(\alpha^+ - \alpha^-)/2^n, \alpha^- + k(\alpha^+ - \alpha^-)/2^n]$ and put $\bar{I}_{k,n} = [\alpha^- + (k-1)(\alpha^+ - \alpha^-)/2^n +$

$2^{-2n}, \alpha^- + k(\alpha^+ - \alpha^-)/2^n - 2^{-2n}]$, i.e., $\bar{I}_{k,n}$ is constructed from $I_{k,n}$ by cutting off subintervals of length 2^{-2n} . Define \bar{h}_n on $\bar{I}_{k,n}$ as

$$\bar{h}_n|_{\bar{I}_{k,n}} \equiv ((1 - 2^{-2n})\Sigma(I_{k,n}) + 2^{-2n} \cdot I) \Sigma_V(I_{k,n})^{-1}$$

and let h_n be the continuous function on $[\alpha^-, \alpha^+]$ obtained by linear interpolation of the step function \bar{h}_n . Then h_n is positive definite and continuous on $[\alpha^-, \alpha^+]$ and as in the scalar case, $h_n \cdot \Sigma_V$ converges weakly to Σ . This concludes the proof of the LDP for $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda_M^\pm(j))$.

In order to extend the LDP to the untruncated eigenvalues, note that the LDP for $(\lambda^+(j), \lambda^-(j))$ implies the exponential tightness of the (unrestricted) extremal eigenvalues that is, for every $K > 0$ there exists a M such that

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log P(\lambda_1^+ > M \text{ or } \lambda_1^- < -M) \leq -K.$$

In particular,

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \log P(\lambda_M^\pm(j) \neq \lambda^\pm(j)) = -\infty,$$

so that as $M \rightarrow \infty$, the truncated eigenvalues are exponentially good approximation of the unrestricted ones. Moreover, $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda_M^\pm(j))$ are exponentially good approximations of $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda^\pm(j))$. By Theorem 4.2.16 in [6] $(\tilde{\Sigma}_{I(j)}^{(N)}, \lambda^\pm(j))$ satisfies the LDP with speed n and rate function

$$\begin{aligned} \mathcal{I}(\Sigma, x^\pm) &= \mathcal{K}(\Sigma_V | \Sigma) + \text{tr } \Sigma(I) - p + \mathcal{I}_{\lambda^\pm}(x^\pm) \\ &= \mathcal{K}(\Sigma_V | \Sigma) + \text{tr } \Sigma(I) - p + \sum_{i=1}^j \mathcal{F}^+(x_i^+) + \mathcal{F}^-(x_i^-), \end{aligned}$$

which ends the proof of Theorem 4.2. □

4.2 LDP for the projective family

Theorem 4.3 *For any fixed j , the sequence of projected spectral measures $\pi_j(\tilde{\Sigma}^{(N)})$ as elements of \mathcal{S}_p with topology (4.3) satisfies the LDP with speed N and rate function*

$$\tilde{\mathcal{I}}_j(\tilde{\Sigma}) = \mathcal{K}(\Sigma_V | \tilde{\Sigma}) + \text{tr } \tilde{\Sigma}(I) - p + \sum_{i=1}^{N^+ \wedge j} (\mathcal{F}_V^+(\lambda_i^+) + \text{tr } \Gamma_i^+) + \sum_{i=1}^{N^- \wedge j} (\mathcal{F}_V^-(\lambda_i^-) + \text{tr } \Gamma_i^-).$$

Proof: The proof is almost analogous to the proof of Theorem 4.3 in [9] and we omit the details for the sake of brevity. It is a combination of the LDP in Theorem 4.2 and the LDP of the weights $\frac{1}{N}\Gamma_k = \frac{1}{N}v_kv_k^*$ corresponding to the extreme eigenvalues. Indeed $\Gamma_i \sim LUE_p(1)$, so by Proposition 3.8 (iii), each individual weight $\frac{1}{N}\Gamma_k$ satisfies the LDP with speed N and rate function

$$\mathcal{I}_3(X) = \begin{cases} \text{tr } X & \text{if } X \geq 0, \\ \infty & \text{otherwise.} \end{cases}$$

The independence of the weights and an application of the contraction principle complete then the proof. \square

In order to come back to a normalized matrix measure in $\mathcal{S}_{p,1}$, we note that the LDP for $\pi_j(\tilde{\Sigma}^{(N)})$ also implies the joint LDP for

$$(\pi_j(\tilde{\Sigma}^{(N)}), \pi_j(\tilde{\Sigma}^{(N)})(\mathbb{R})),$$

with the rate function

$$\bar{\mathcal{I}}_j(\tilde{\Sigma}, Z) = \tilde{\mathcal{I}}_j(\tilde{\Sigma})$$

if $\tilde{\Sigma}(\mathbb{R}) = Z$ and $\bar{\mathcal{I}}_j(\tilde{\Sigma}, Z) = \infty$ otherwise. Keeping the weights along the way, we may apply the projective method (the Dawson-Gärtner theorem, p. 162 in the book of [6]) to the family of projections $(\pi_j(\tilde{\Sigma}^{(N)}), \pi_j(\tilde{\Sigma}^{(N)})(\mathbb{R}))_j$ and get a LDP for the pair $(\tilde{\Sigma}^{(N)}, \tilde{\Sigma}^{(N)}(\mathbb{R}))$ with rate function

$$\bar{\mathcal{I}}(\tilde{\Sigma}, Z) = \sup_j \bar{\mathcal{I}}_j(\tilde{\Sigma}, Z)$$

This rate function equals $+\infty$ unless $\tilde{\Sigma}(\mathbb{R}) = Z$ and in this case is given by

$$(4.18) \quad \bar{\mathcal{I}}(\tilde{\Sigma}, Z) = \mathcal{K}(\Sigma_V \mid \tilde{\Sigma}) + \text{tr} Z - p + \sum_{i=1}^{N^+} \mathcal{F}(\lambda_i^+) + \sum_{i=1}^{N^-} \mathcal{F}(\lambda_i^-).$$

We remark that normalizing the matrix measure is not possible unless we keep track of the total mass for any j , as the mapping $\tilde{\Sigma} \mapsto \tilde{\Sigma}(\mathbb{R})^{-1/2}\tilde{\Sigma}\tilde{\Sigma}(\mathbb{R})^{-1/2}$ is not continuous in the topology (4.3). However, we may now apply the continuous mapping $(\tilde{\Sigma}, Z) \mapsto Z^{-1/2}\tilde{\Sigma}Z^{-1/2}$ and obtain a LDP for the sequence of measures $\Sigma^{(N)}$ in $\mathcal{S}_{p,1}$. The rate function is

$$\mathcal{I}(\Sigma) = \inf_{\tilde{\Sigma}=Z^{1/2}\Sigma Z^{1/2}, Z>0} \bar{\mathcal{I}}(\tilde{\Sigma}) = \inf_{Z>0} \tilde{\mathcal{I}}(Z^{1/2}\Sigma Z^{1/2}).$$

By (4.18), we need to minimize over positive definite $Z \in \mathbb{H}_p$ the function

$$\begin{aligned}
& - \int \log \det \left(\frac{d(Z^{1/2} \Sigma Z^{1/2})}{d\mu_V} \right) d\mu_V + \operatorname{tr} Z - p \\
&= - \int \log \det \left(Z^{1/2} \frac{d\Sigma}{d\mu_V} Z^{1/2} \right) d\mu_V + \operatorname{tr} Z - p \\
&= - \int \log \det \left(\frac{d\Sigma}{d\mu_V} \right) d\mu_V - \log \det Z + \operatorname{tr} Z - p \\
&= - \int \log \det \left(\frac{d\Sigma}{d\mu_V} \right) d\mu_V + \mathcal{I}_2(Z)
\end{aligned}$$

The term $\mathcal{I}_2(Z)$ comes from Lemma 3.4 (ii) with $\gamma = 1$ and attains its minimal value 0 for $Z = I_p$. \square

We have obtained the LDP claimed in Theorem 3.4 on the subset $\mathcal{S}_{p,1}$ in the topology induced by (4.3). On $\mathcal{S}_{p,1}$ this is stronger than the weak topology and the rate function can be extended to $\mathcal{M}_{p,1}$ by setting $\mathcal{I}(\Sigma) = \infty$ for $\Sigma \notin \mathcal{S}_{p,1}$. This yields Theorem 3.4.

4.3 Proof of Remark 3.5

Let A be a measurable subset of $\mathcal{M}_{p,1}$ and set

$$A_N = \left\{ (\lambda, W) \in \mathbb{R}^N \times \mathbb{H}_p^N \mid \sum_{k=1}^N W_k \delta_{\lambda_k} \in A \right\}.$$

The LDP for $\Sigma^{(N)}$ with weight distribution $\mathbb{P}_N^{V_N}$ will follow from the LDP for weight distribution \mathbb{P}_N^V once we show

$$(4.19) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^{V_N} \otimes \mathbb{D}_{p,N})(A_N) \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^V \otimes \mathbb{D}_{p,N})(A_N)$$

and

$$(4.20) \quad \liminf_{N \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^{V_N} \otimes \mathbb{D}_{p,N})(A_N) \geq \liminf_{N \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^V \otimes \mathbb{D}_{p,N})(A_N).$$

In fact, this does not require A to be closed or open, respectively. For this, let

$$\Gamma_N^V(A_N) = \iint_{A_N} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^2 \prod_{i=1}^N e^{-NV(\lambda_i)} d\lambda d\mathbb{D}_{p,N}(W).$$

and define $\Gamma_N^{V_N}(A_N)$ analogously, with V replaced by V_N . Since $V_N(x) \geq V(x)$, we have

$$(4.21) \quad \Gamma_N^{V_N}(A_N) \leq \Gamma_N^V(A_N).$$

To get a converse inequality, let $K_{N,M}$ be the set of $(\lambda, W) \in \mathbb{R}^N \times \mathbb{H}_p^N$, where $V(\lambda_i) \leq M$ for all i . Then

$$\Gamma_N^{V_N}(A_N) \geq \Gamma_N^{V_N}(A_N \cap K_{N,M}) \geq \left(\inf_{x: V(x) \leq M} e^{V(x) - V_N(x)} \right)^N \Gamma_N^V(A_N \cap K_{N,M}).$$

Since by assumption $e^{V(x) - V_N(x)}$ converges to 1 uniformly on $\{x \mid V(x) \leq M\}$, this implies

$$(4.22) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \log \frac{\Gamma_N^{V_N}(A_N)}{\Gamma_N^V(A_N)} \geq \lim_{N \rightarrow \infty} \frac{1}{N} \log \frac{\Gamma_N^V(A_N \cap K_{N,M})}{\Gamma_N^V(A_N)}.$$

If we set now A to be the whole space $\mathcal{M}_{p,1}$, then $\Gamma_N^V(A_N) = Z_N^V$ and the right hand side of (4.22) is given by

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_N^V(\forall i : V(\lambda_i) \leq M).$$

By the LDP for the extreme eigenvalues, Theorem 4.1, this limit tends to 0 as $M \rightarrow \infty$. Together with (4.21), we have shown that for $A = \mathcal{M}_{p,1}$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \frac{\Gamma_N^{V_N}(A_N)}{\Gamma_N^V(A_N)} = \lim_{N \rightarrow \infty} \frac{1}{N} \log \frac{Z_N^{V_N}}{Z_N^V} = 0.$$

Since $(\mathbb{P}_N^{V_N} \otimes \mathbb{D}_{p,N})(A_N) = (Z_N^V)^{-1} \Gamma_N^V(A_N)$, the inequality (4.21) leads to the inequality (4.19) and the arguments for (4.22) yield

$$\liminf_{n \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^{V_N} \otimes \mathbb{D}_{p,N})(A_N) \geq \liminf_{N \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^V \otimes \mathbb{D}_{p,N})(A_N \cap K_{N,M})$$

for any $M \geq 0$. Letting $M \rightarrow \infty$, this implies inequality (4.20), as by the LDP for the extreme eigenvalues we have

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \log(\mathbb{P}_N^V \otimes \mathbb{D}_{p,N})(K_{N,M}^c) = -\infty.$$

□

5 Hermite block case

The starting point for the proof of Theorem 3.6 is the following block-tridiagonal representation of the Gaussian ensemble.

Lemma 5.1 *Let $D_k \sim \text{GUE}_p$ and C_k be Hermitian such that $C_k^2 \sim \text{LUE}_p(p(n-k))$ for $k = 1, \dots, n$ and let all these matrices be independent. Then the $p \times p$ matrix spectral measure of the block-tridiagonal matrix*

$$(5.1) \quad \mathcal{G}_n = \begin{pmatrix} D_1 & C_1 & & & \\ C_1 & D_2 & \cdots & & \\ & \cdots & \cdots & C_{n-1} & \\ & & C_{n-1} & D_n & \end{pmatrix}$$

has the same distribution as the spectral measure of the Hermite ensemble GUE_{pn} .

Proof: Starting from a matrix X_n distributed according to the Hermite ensemble GUE_{pn} , we can construct the tridiagonal matrix \mathcal{G}_n as

$$\mathcal{G}_n = TX_nT^*,$$

where T is unitary and leaves invariant the subspaces spanned by the first p unit vectors. Consequently, the spectral measure of X_n is also the spectral measure of \mathcal{G}_n . The transformation T is a composition of unitaries T_1, \dots, T_{n-1} analogous to the ones used by [8]. To illustrate the procedure, we construct the first transformation T_1 . By $x_{i,j}$ we denote the $p \times p$ subblock of X_n in position i, j , let $\bar{x}_1 = (x_{1,2}, \dots, x_{1,n})^*$ and $\bar{X} = (x_{i,j})_{2 \leq i, j \leq n}$. With this notation, X_n can be structured as

$$X_n = \begin{pmatrix} x_{1,1} & \bar{x}_1^* \\ \bar{x}_1 & \bar{X} \end{pmatrix}.$$

Note that the Gaussian distribution implies that all (square) blocks are almost surely invertible. Then, set

$$\begin{aligned} \xi &= [(x_{2,1}^*)^{-1}(\bar{x}_1^* \bar{x}_1)(x_{2,1})^{-1}]^{1/2} x_{2,1} \in \mathbb{M}_{p,p} \\ \Gamma &= (\xi^*, 0_{p,p}, \dots, 0_{p,p})^* \in \mathbb{M}_{(n-1)p,p} \end{aligned}$$

and define for $Z \in \mathbb{M}_{(n-1)p,p}$ the block-Householder reflection

$$H(Z) = I_{(n-1)p} - 2Z(Z^*Z)^{-1}Z^*.$$

If we set $Z = \Gamma - \bar{x}_1$ one may check that

$$\Gamma^* \Gamma = \xi^* \xi = \bar{x}_1^* \bar{x}_1, \quad \Gamma^* \bar{x}_1 = \xi^* x_{2,1} = x_{2,1}^* \xi = \bar{x}_1^* \Gamma, \quad Z^* Z = -2Z^* \bar{x}_1$$

and $H(Z)\bar{x}_1 = \Gamma$. We extend $H(Z)$ to an operator \hat{H} on \mathbb{C}^{np} leaving the first p unit vectors invariant, which yields

$$\hat{H}X_n\hat{H}^* = \begin{pmatrix} I_p & 0 \\ 0 & H(Z) \end{pmatrix} \begin{pmatrix} x_{1,1} & \bar{x}_1^* \\ \bar{x}_1 & \bar{X} \end{pmatrix} \begin{pmatrix} I_p & 0 \\ 0 & H(Z)^* \end{pmatrix} = \begin{pmatrix} x_{1,1} & \Gamma^* \\ \Gamma & H(Z)\bar{X}H(Z)^* \end{pmatrix}.$$

Finally, let $W \in \mathbb{M}_{(n-1)p, (n-1)p}$ be the unitary block-diagonal matrix with the blocks $((\xi^*\xi)^{1/2}\xi^{-1}, I_p, \dots, I_p)$ on the diagonal and extend W to an operator \hat{W} on \mathbb{C}^{np} as we did with $H(Z)$. Then $T_1 = \hat{W}\hat{H}$ is unitary, leaves the subspace spanned by the first p unit vectors invariant and

$$T_1X_nT_1^* = \begin{pmatrix} x_{1,1} & \tilde{\Gamma} \\ \tilde{\Gamma}^* & WH(Z)\bar{X}H(Z)^*W^* \end{pmatrix}.$$

with

$$\tilde{\Gamma} = ((\bar{x}_1^*\bar{x}_1)^{1/2}, 0_{p,p}, \dots, 0_{p,p}).$$

The block $\bar{x}_1^*\bar{x}_1$ is distributed according to $\text{LUE}_p(p(n-1))$ and since the definition of W and $H(Z)$ is independent of \bar{X} , the block $WH(Z)\bar{X}H(Z)^*W^*$ is again a matrix of the Gaussian ensemble $\text{GUE}_{p(n-1)}$. The assertion follows then from an iteration of these reflections. \square

Proof of Theorem 3.6:

By Lemma 5.1, the spectral measure $\Sigma^{(n)}$ is also the spectral measure of the rescaled matrix $\frac{1}{\sqrt{np}}\mathcal{G}_n$. If we consider each block entry of this matrix separately, we are up to a linear change of the speed in the setting of Lemma 3.8. Thus, for any fixed k , the block $D_k^{(n)} := D_k/\sqrt{np}$ of the matrix in (5.1) satisfies the LDP in \mathbb{H}_p with speed pn and rate function \mathcal{I}_1 (see Lemma 3.8). Similarly, if we let $C_k^{(n)} = C_k/\sqrt{np}$, then the squared block $(C_k^{(n)})^2$ satisfies the LDP with speed pn and rate function \mathcal{I}_2 or equivalently, $C_k^{(n)}$ in the block-tridiagonal matrix satisfies the LDP with speed np and rate

$$\mathcal{I}'_2(Y) = \inf_{C: C^{1/2}=Y} \mathcal{I}_2(C) = \mathcal{I}_2(Y^2)$$

if Y is nonnegative definite and $\mathcal{I}'_2(Y) = \infty$ otherwise. Since the block entries are independent, we get a joint LDP for any fixed collection $(D_1^{(n)}, C_1^{(n)}, \dots, D_k^{(n)})$ with rate given by the corresponding sum of rate functions \mathcal{I}_1 and \mathcal{I}'_2 . The Dawson-Gärtner Theorem (see [6]) yields a LDP in $\mathbb{H}_p^{\mathbb{N}}$ for the infinite sequence $(D_1^{(n)}, C_1^{(n)}, \dots)$ with rate

$$\mathcal{I}_{D,C}(D_1, C_1, \dots) = \sum_{k=1}^{\infty} \left(\frac{1}{2} \text{tr} D_k^2 + \text{tr} G(C_k^2) \right),$$

which is finite if and only if all matrices C_k are positive definite and the sum converges. In particular, the set where $\mathcal{I}_{D,C} < \infty$ is a subset of

$$\Xi = \left\{ (D_1, C_1, \dots) \in \mathbb{H}_p^{\mathbb{N}} \mid \sup_k \|D_k\| + \|C_k\| < \infty \right\}$$

and so the LDP holds in the subset Ξ . For sequences in Ξ , the Jacobi operator

$$\begin{pmatrix} D_1 & C_1 & & \\ C_1 & D_2 & \ddots & \\ & \ddots & \ddots & \end{pmatrix}$$

is bounded and we can define the (compactly supported) spectral measure of this operator. Let $\psi : \Xi \rightarrow \mathcal{M}_{p,1,c}$ be the mapping assigning to (D_1, C_1, \dots) this spectral measure. Then ψ is continuous and by Lemma 5.1

$$\psi((D_1^{(n)}, C_1^{(n)}, \dots, D_n^{(n)}, 0_{p,p}, \dots)) = \Sigma^{(n)}.$$

The contraction principle yields the LDP for the sequence $(\Sigma^{(n)})_n$ with rate function

$$\mathcal{I}_H(\Sigma) = \inf_{(D_1, C_1, \dots) : \psi((D_1, C_1, \dots)) = \Sigma} \sum_{k=1}^{\infty} \left[\frac{1}{2} \text{tr} D_k^2 + \text{tr} G(C_k^2) \right].$$

This function can only be finite if Σ is non-trivial. Indeed, we know from the remarks of Section 2.1, that in this case all the C_k are non-singular. As a matter of fact, there is a one-to-one correspondence between non trivial measures and sequences (B_1, A_1, \dots) of block Jacobi coefficients with A_k positive definite, with (B_1, A_1, \dots) being the unique representative in the equivalence class of Jacobi parameters such that all A_k are Hermitian. We obtain

$$\mathcal{I}_H(\Sigma) = \sum_{k=1}^{\infty} \left[\frac{1}{2} \text{tr} B_k^2 + \text{tr} G(A_k A_k^*) \right],$$

where B_k, A_k are the recursion coefficients of Σ . Note that by writing $A_k A_k^*$ instead of A_k^2 , the right hand side is defined also for non-Hermitian A_k and moreover, the value of the sum is constant on the whole equivalence class of Jacobi coefficients of Σ . This concludes the proof. \square

6 Laguerre block case

Lemma 6.1 *Let $m \geq n$ and for $k = 1, \dots, n$ let D_k and C_k for $k = 1, \dots, n$ be independent random matrices in \mathbb{H}_p such that*

$$C_k^2 \sim \text{LUE}_p(p(n-k)) \quad , \quad D_k^2 \sim \text{LUE}_p(p(m-k+1))$$

and define the block matrix

$$Z_n = \begin{pmatrix} D_1 & 0 & & \\ C_1 & D_2 & \cdots & \\ & \cdots & \cdots & 0 \\ & & C_{n-1} & D_n \end{pmatrix}.$$

Then the $p \times p$ spectral matrix measure of $L_n = Z_n Z_n^*$ has the same distribution as the spectral matrix measure of a $pn \times pn$ matrix, distributed according to the $\text{LUE}_{pn}(pm)$ ($m \geq n$).

Proof: We use the construction of the Laguerre ensemble $L_n = W_n W_n^*$, with W_n a $pn \times pm$ matrix with independent complex Gaussian entries. Writing $w_{i,j}$ for the p -block of W_n in position i, j , let R be a $pm \times pm$ unitary matrix constructed analogously to the matrix $\hat{W}\hat{H}$ in the proof of Lemma 5.1, such that

$$W_n R = \begin{pmatrix} \tilde{w} \\ \widetilde{W} \end{pmatrix}$$

with

$$\tilde{w} = (w_{1,1}, \dots, w_{1,m}) R = \left(\left(\sum_{i=1}^m w_{1,i}^* w_{1,i} \right)^{1/2}, 0_{p,p}, \dots, 0_{p,p} \right).$$

The matrix R can be chosen independently of $w_{i,j}, i \geq 2$ such that the entries of \widetilde{W} are again independent complex Gaussian, independent of \tilde{w} . Similarly, write $z_{i,j}$ for the $p \times p$ subblock of $W_n R$ in position i, j and let L be a $p(n-1) \times p(n-1)$ unitary matrix such that

$$L(z_{2,1}^*, \dots, z_{n,1}^*)^* = \left(\left(\sum_{i=2}^n z_{1,i}^* z_{1,i} \right)^{1/2}, 0_{p,p}, \dots, 0_{p,p} \right)^*$$

If $\tilde{L} = I \oplus L$ is the extension of L to an operator on \mathbb{C}^{pm} , leaving the subspace of the first p unit vectors invariant, then

$$\tilde{L} W_n R = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ C_1 & & & \\ 0 & & L \widetilde{W} R & \\ \vdots & & & \end{pmatrix}.$$

The first blocks satisfy $D_1^2 \sim \text{LUE}_p(pm)$, $C_1^2 \sim \text{LUE}_p(p(n-1))$ and by the invariance of the Gaussian distribution, the entries of $L\widetilde{W}R$ are again Gaussian distributed. Since we started with independent entries, all blocks in $\widetilde{L}W_nR$ are independent. The conjugation by \widetilde{L} leaves the first p eigenvector rows invariant, so $\widetilde{L}L_n\widetilde{L}^* = \widetilde{L}W_nRR^*W_n^*\widetilde{L}^*$ has the same spectral matrix measure as L_n . This yields the first step in the reduction, the iterations are straightforward. \square

Proof of Theorem 3.7:

As in the proof of Theorem 3.6, we start by looking at the individual entries of the tridiagonal representation of Lemma 6.1, now multiplies by $\frac{1}{p\gamma_n}$. For any k , the rescaled block $\frac{1}{p\gamma_n}C_k^2$ satisfies by Lemma 3.8 the LDP with speed $p\gamma_n$ and rate \mathcal{I}_2 with $\gamma = \tau$. With the speed pn we would like to consider, $\frac{1}{p\gamma_n}C_k^2$ satisfies then the LDP with rate $\text{tr}G(\tau^{-1}\cdot)$ and, taking the square root, $C_k^{(n)} := \frac{1}{\sqrt{p\gamma_n}}C_k$ satisfies the LDP with speed pn and rate function

$$\mathcal{I}_C(C) = \text{tr} G(\tau^{-1}C^2)$$

for C positive definite and $\mathcal{I}'_2(C) = \infty$ otherwise. Similarly, if we let $D_k^{(n)} := \frac{1}{\sqrt{p\gamma_n}}D_k$, then $(D_k^{(n)})^2$ satisfies the LDP with speed $p\gamma_n$ and rate function \mathcal{I}_2 of Lemma 3.8 with $\gamma = 1$. If we consider the speed pn and the square root $D_k^{(n)}$, this changes the rate to

$$\mathcal{I}_D(D) = \tau^{-1}\text{tr} G(D^2)$$

for D positive definite and $\mathcal{I}_D(D) = \infty$ otherwise. By the independence of the matrices C_k, D_k , this yields the LDP for $(D_1^{(n)}, C_1^{(n)}, D_2^{(n)}, \dots)$ in the sequence space of Hermitian nonnegative definite matrices with speed pn and rate

$$(6.1) \quad \mathcal{I}_{D,C}(D_1, C_1, \dots) = \sum_{k=1}^{\infty} \tau^{-1}\text{tr} G(D_k^2) + \text{tr} G(\tau^{-1}C_k^2).$$

As in the proof of Theorem 3.6, this rate can only be finite if $\sup_k \|D_k\| + \|C_k\| < \infty$ and we may restrict the LDP to the subset Θ of such sequences of Hermitian nonnegative definite matrices. On Θ define the function $\varphi : \Theta \rightarrow \mathcal{M}_{p,1,c}([0, \infty))$ mapping (D_1, C_1, \dots) to the compactly supported spectral measure of the matrix ZZ^* , where

$$Z = \begin{pmatrix} D_1 & 0_{p,p} & & \\ C_1 & D_2 & \ddots & \\ & \ddots & \ddots & \end{pmatrix}.$$

Then φ is continuous and by Lemma 6.1, the measure

$$\varphi(D_1^{(n)}, C_1^{(n)}, \dots, D_n^{(n)}, 0_{p,p}, \dots) = \Sigma^{(n)}$$

is distributed as the spectral measure of $\frac{1}{p\gamma_n}Y_n$, when $Y_n \sim \text{LUE}_N(p\gamma_n)$. By the contraction principle, the sequence $(\Sigma^{(n)})_n$ satisfies then the LDP with speed pn and rate function

$$\mathcal{I}_L(\Sigma) = \inf_{(D_1, C_1, \dots): \varphi((D_1, C_1, \dots)) = \Sigma} \sum_{k=1}^{\infty} \tau^{-1} \text{tr} G(D_k^2) + \text{tr} G(\tau^{-1} C_k^2).$$

Next, we show that if Σ is a non-trivial matrix measure compactly supported on $[0, \infty)$, the infimum is not taken over the empty set, that is, there exists a sequence $(D_1, C_1, \dots) \in \Theta$ mapped to Σ . While the discussion in Section 2.2 showed that every nontrivial measure with compact support in $[0, \infty)$ is the spectral measure of a matrix XX^* with X bidiagonal as in (2.22), we have to show that the entries of X can be chosen to be Hermitian nonnegative. We are still free to choose the unitary matrices σ_n (although we have to fix $\sigma_1 = I_p$) in the definition of orthonormal polynomials and we let U be the block-diagonal matrix with $\sigma_1, \sigma_2, \dots$ on the diagonal. Moreover, let P denote a block-diagonal matrix with unitary $p \times p$ matrices τ_1, τ_2, \dots on the diagonal. Then our measure Σ is also the spectral matrix measure of the matrix $UXPP^*X^*U^* = (UXP)(UXP)^*$. The matrix UXP has the form

$$UXP = \begin{pmatrix} \sigma_1 D_1 \tau_1 & 0_{p,p} & & & \\ \sigma_2 C_1 \tau_1 & \sigma_2 D_2 \tau_2 & 0_{p,p} & & \\ & \sigma_3 C_2 \tau_2 & \sigma_3 D_3 \tau_3 & \ddots & \\ & & & \ddots & \ddots \end{pmatrix}.$$

For the first entry, $\sigma_1 = I_p$ and D_1 is always Hermitian positive definite, so we may set $\tau_1 = I_p$. Recall that for A a non-singular matrix, there exists a unique unitary σ such that $A\sigma$ is Hermitian positive definite, and if Σ is non-trivial, all D_k, C_k are non-singular. Therefore, we can recursively choose σ_{k+1} such that $\sigma_{k+1} C_k \tau_k$ is Hermitian positive definite and then τ_{k+1} such that $\sigma_{k+1} D_{k+1} \tau_{k+1}$ is positive definite. This yields a unique decomposition with positive definite blocks.

To simplify the rate function and finish the proof, we can use the fact that $\text{tr}(AB) = \text{tr}(BA)$ and $\det(AB) = \det(BA)$ to get

$$\text{tr} G(D_k^2) = \text{tr} G(D_k D_k^*) = \text{tr} G(Z_{2k-1}) = \text{tr} G(\zeta_{2k-1})$$

with Z_{2k-1} as in (2.19) and (2.20), and

$$\text{tr} G(\tau^{-1} C_k^2) = \text{tr} (\tau^{-1} C_k C_k^*) = \text{tr} (\tau^{-1} Z_{2k}) = \text{tr} (\tau^{-1} \zeta_{2k}).$$

So the value of the rate function does not depend on the unitary matrices σ_n and τ_n , but only on the matrices ζ_k , which in particular are uniquely determined by Σ . \square

References

- [1] G. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*. Cambridge University Press, Cambridge, 2010.
- [2] P. Baldi. Large deviations and stochastic homogenization. *Ann. Mat. Pura Appl., IV. Ser.*, 151:161–177, 1988.
- [3] D. Damanik, R. Killip, and B. Simon. Perturbations of orthogonal polynomials with periodic recursion coefficients. *Annals of mathematics*, 171(3):1931–2010, 2010.
- [4] D. Damanik, A. Pushnitski, and B. Simon. The analytic theory of matrix orthogonal polynomials. *Surveys in Approximation Theory*, 4:1–85, 2008.
- [5] A. P. Dawid. Spherical matrix distributions and a multivariate model. *J. R. Stat. Soc., Ser. B*, 39:254–261, 1977.
- [6] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, 1998.
- [7] H. Dette and W.J. Studden. Matrix measures, moment spaces and Favard’s theorem for the interval $[0,1]$ and $[0,\infty)$. *Linear Algebra and its Applications*, 345:169–193, 2002.
- [8] I. Dumitriu and A. Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.
- [9] F. Gamboa, J. Nagel, and A. Rouault. Sum rules via large deviations. *Journal of Functional Analysis*, 270(2):509 – 559, 2016.
- [10] F. Gamboa, J. Nagel, A. Rouault, and J. Wagener. Large deviations for random matricial moment problems. *J. Multivariate Anal.*, 106:17–35, 2012.
- [11] F. Gamboa and A. Rouault. Operator-valued spectral measures and large deviations. *Journal of Statistical Planning and Inference*, 154:72–86, 2014.
- [12] K. Johansson. On fluctuations of eigenvalues of random hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.

- [13] R. Killip and B. Simon. Sum rules for Jacobi matrices and their applications to spectral theory. *Ann. of Math. (2)*, 158(1):253–321, 2003.
- [14] V. Mandrekar and H. Salehi. On singularity and Lebesgue type decomposition for operator-valued measures. *Journal of Multivariate Analysis*, 1(2):167–185, 1971.
- [15] L. Mirsky. On the trace of matrix products. *Math. Nachr.*, 20:171–174, 1959.
- [16] J.B. Robertson and M. Rosenberg. The decomposition of matrix-valued measures. *Michigan Math. J*, 15:353–368, 1968.
- [17] R.T. Rockafellar. Integrals which are convex functionals, II. *Pacific J. Math.*, 39(2):439–469, 1971.
- [18] B. Simon. *Orthogonal polynomials on the unit circle. Part 1*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005. Classical theory.
- [19] B. Simon. *Orthogonal polynomials on the unit circle. Part 2*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005. Spectral theory.
- [20] B. Simon. *Szegő’s theorem and its descendants*. M. B. Porter Lectures. Princeton University Press, Princeton, NJ, 2011.

Fabrice Gamboa: Université Paul Sabatier, Institut de Mathématiques de Toulouse, 118 route de Narbonne, 31062 Toulouse Cedex 9, France, gamboa@math.univ-toulouse.fr

Jan Nagel: Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, 85748 Garching bei München, Germany, jan.nagel@tum.de

Alain Rouault: Université Versailles-Saint-Quentin, LMV UMR 8100, 45 Avenue des Etats-Unis, 78035-Versailles Cedex, France, alain.rouault@uvsq.fr