Alea **3**, 181–230 (2007)

Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles

Alain Rouault

LMV Bâtiment Fermat, Université Versailles-Saint-Quentin, F-78035 Versailles, France
E-mail address: `Alain.Rouault@math.uvsq.fr`

Abstract. We consider properties of determinants of some random symmetric matrices issued from multivariate statistics: Wishart/Laguerre ensemble (sample covariance matrices), Uniform Gram ensemble (sample correlation matrices) and Jacobi ensemble (MANOVA). If n is the size of the sample, $r \leq n$ the number of variates and $X_{n,r}$ such a matrix, a generalization of the Bartlett-type theorems gives a decomposition of $\det X_{n,r}$ into a product of r independent Gamma or Beta random variables. For n fixed, we study the evolution as r grows, and then take the limit of large r and n with $r/n = t \leq 1$. We derive limit theorems for the sequence of *processes with independent increments* $\{n^{-1} \log \det X_{n,\lfloor nt \rfloor}, t \in [0,T]\}_n$ for $T \leq 1$: convergence in probability, invariance principle, large deviations. Since the logarithm of the determinant is a linear statistic of the empirical spectral distribution, we connect the results for marginals (fixed t) with those obtained by the spectral method. Actually, all the results hold true for Coulomb gases or β-models, if we define the determinant as the product of charges. The classical matrix models (real, complex, and quaternionic) correspond to the particular values $\beta = 1, 2, 4$ of the Dyson parameter.

1. Introduction

Random determinants of symmetric matrices are of constant use in random geometry to compute volumes of parallelotopes (see Nielsen (1999), Mathai (1999)) and in multivariate statistics to build tests (see Muirhead (1982), Anderson (2003)). Twenty years after the book of Girko (1988), recent developments in Random Matrix Theory add a new interest to the study of their asymptotic behavior and invite to a new insight.

Received by the editors June 11, 2007; accepted October 2, 2007.

2000 *Mathematics Subject Classification.* Primary 15A52, 15A15; Secondary 60F10, 60F17, 62H10.

Key words and phrases. Random matrices, Wishart ensemble, Laguerre ensemble, Jacobi ensemble, Gram ensemble, Hadamard ratio, determinant, invariance principle, large deviations.

Let $B = [b_1, \dots, b_r]$ be the $n \times r$ matrix with r column vectors $b_1, \dots, b_r$ of $\mathbb{R}^n$. If B' denotes its transpose, the determinant of the $r \times r$ Gram matrix $B'B$ satisfies the well known Hadamard inequality:

$$\det B'B \leq \|b_1\|^2 \cdots \|b_r\|^2 \tag{1.1}$$

with equality if and only if $b_1, \dots, b_r$ are orthogonal (Hadamard, 1893). It means that the volume (or r-content) of the parallelotope built from $b_1, \dots, b_r$ is maximal when the vectors are orthogonal. The quantity

$$h(B) = \frac{\det B'B}{\|b_1\|^2 \cdots \|b_r\|^2}$$

is usually called the Hadamard ratio. If we replace sequentially b_i by its projection $\widehat{b}_i$ on the orthogonal of the subspace spanned by $b_1, \dots, b_{i-1}$ (Gram-Schmidt orthogonalization), we have

$$\det B'B = \prod_{i=1}^{r} \|\widehat{b}_i\|^2 .$$

Motivated by basis reduction problems, Schnorr (1986) defined the orthogonality defect as the quantity $1/\sqrt{h(B)}$ (see also Akhavi (2002) and references therein). Abbott and Mulders (2001) and Dixon (1984) are concerned with the tightness of the bound $h(B) \leq 1$ when B is random and $n = r$. For these authors, the random vectors b_i are sampled independently and uniformly on the unit sphere

$$\mathbb{S}^n_{\mathbb{R}} = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\} .$$

It is known that then the variables $\|\widehat{b}_i\|^2$ are independent and Beta distributed with varying parameters. When the entries of the matrix B are independent and $\mathcal{N}(0, 1)$ the variables $\|\widehat{b}_i\|^2$ are independent and Gamma distributed with varying parameters (Bartlett, 1933).

Writing $B_{n,r}$ instead of B to emphasize dimensions and $X_{n,r} = B'_{n,r}B_{n,r}$, we are interested in this paper in the asymptotic behavior of $\det X_{n,r}$ when n and r both tends to infinity in the regime $r/n \to c \in [0, 1]$. Since the construction of the $\widehat{b}_i$ is recursive, it is possible (for fixed n) to consider the whole sequence of variables $\{\det X_{n,r}, r = 1, \dots, n\}$ at the same time. It corresponds to the decomposition of the determinant of a $r \times r$ symmetric positive matrix A as

$$\det A = \prod_{j=1}^{r} \frac{\det A^{[j]}}{\det A^{[j-1]}} ,$$

where $A^{[j]}$ is the $j \times j$ upper-left corner of A with the convention $\det A^{[0]} = 1$. When using this approach we will refer to it as the *decomposition method*. This method is also valid when entries of the matrix are complex, considering the Hermitian conjugate $B^\star$ and then $B^\star B$, and also when the entries are real quaternions, considering the dual $B^\dagger$ and then $B^\dagger B$. In these three cases, and for the above models of random matrices, Bartlett-type theorems give the determinant as a product of independent variables, with Gamma or Beta distributions. Passing to logarithms, it is then possible to consider a triangular array of variables and a process with independent increments $\{n^{-1} \log \det X_{n,\lfloor nt \rfloor}, t \in [0, T]\}_n$ for $T \leq 1$ indexed by the "time" $t = r/n$. Thanks to the additive structure of the log det, we obtained limit theorems: convergence in probability, invariance principle and large deviations. The same is true for random matrices following the Jacobi (or MANOVA) distribution.

Actually, the whole construction is possible in the so-called β-models, which are an extension of the above ones, which correspond to the three-fold way $\beta = 1, 2, 4$ of Dyson. For other values of β they are not defined as matrix models but Coulomb gases models, in which the eigenvalues are replaced by charges and determinants by products of charges. It has be shown recently that they correspond also to models of tri-diagonal random matrices (see Dumitriu and Edelman (2003), Killip and Nenciu (2004), Edelman and Sutton (2007)).

Of course, for r fixed, there is also another underlying structure of product: the determinant as the product of eigenvalues. We may use the asymptotic behavior of empirical spectral distributions, i.e. convergence to the Marčenko-Pastur distribution in the Wishart/Laguerre case and to the generalized McKay distribution in the Jacobi case. However, this structure is not "dynamic": if you change r, the whole set of eigenvalues is changing. When using this approach we will refer to it as the *spectral method*.

The structure of the article is as follows. In Section 2 we set the framework. We begin with the matrix models (Wishart-Laguerre, Uniform Gram and Jacobi), and proceed with the β-models and processes of determinants. The main results of this paper are in Section 3: laws of large numbers and fluctuations, large deviations and variational problems. The comparison of results obtained by the two methods (decomposition and spectral) deserves interest and is the topic of Section 4. Some extensions to other models are given in Section 4.4. Sections 5, 6 and 7 are devoted to the proofs. In Appendix 1 we gather some details on Binet's formula for the Gamma function which are of constant use in this paper, and Appendix 2 gives identification of the McKay distribution.

2. Notation and known facts

In this long section, we present our different models whose common feature is to introduce processes of random determinants with independent multiplicative factors. The distribution of these factors are recorded in Proposition 2.1 for real matrix models, and settled in formulae (2.2), (2.3) and (2.4) for the (other) β-models.

Throughout, $|A|$ stands for $\det A$, and I_n for the $n \times n$ identity matrix. If X, Y are real random variables and μ a distribution on $\mathbb{R}$, we write

$$X \stackrel{(d)}{=} Y \quad (\text{resp. } X \stackrel{(d)}{=} \mu)$$

if X and Y have the same distribution (resp. if the distribution of X is μ).

2.1. *Real matrix models and Bartlett-type theorems.* In the basic model, we consider independent random vectors $b_i, i \geq 1$ with the same distribution ν_n in $\mathbb{R}^n$. The most important example is the Gaussian one with $\nu_n = \mathcal{N}(0, I_n)$. If $B = [b_1, \dots, b_r]$, all the entries of B are independent $\mathcal{N}(0,1)$ and the distribution of $W = B'B$ is denoted by $W_r(n, \mathbb{R})$ and called the Wishart ensemble. For $r \leq n$, its density on the space $\mathcal{S}_r$ of symmetric positive matrix is

$$\frac{1}{2^{rn/2}\Gamma_r(n/2)}|W|^{(n-r-1)/2} \exp\Big(-\frac{1}{2}\mathrm{tr}\, W\Big)$$

where Γ_r is the multivariate Gamma function

$$\Gamma_r(\alpha) = \pi^{r(r-1)/4}\Gamma(\alpha)\Gamma\Big(\alpha - \frac{1}{2}\Big)\cdots\Gamma\Big(\alpha - \frac{r-1}{2}\Big).$$

It is the matrix variate extension of the Gamma distribution. Recall that for $a, c > 0$, the Gamma(a, c) distribution has density

$$\frac{c^a}{\Gamma(a)}\, x^{a-1}e^{-cx} \quad (x > 0).$$

For $r > n$, the matrix is singular.

Motivated by Hadamard inequality (1.1), we may choose ν_n to be the uniform distribution on the unit sphere $\mathbb{S}^n_{\mathbb{R}}$. The corresponding ensemble for B is called Uniform Spherical Ensemble by Tsaig and Donoho (2006). The matrix ensemble for $B'B$ is called the Gram ensemble by De Cock, Fannes and Spincemaille (1999), since $B'B$ is the Gram matrix built from the b_i's. To emphasize the distribution, we call it Uniform Gram ensemble. The diagonal entries are one and for $r \le n$, the joint density of the non-diagonal entries $(r_{ij},\ 1 \le i < j \le r)$ of the matrix $G = B'B$ is

$$\frac{\left[\Gamma\left(n/2\right)\right]^r}{\Gamma_r\left(n/2\right)}\, |G|^{(n-r-1)/2} \quad (-1 < r_{ij} < 1) \tag{2.1}$$

(see Gupta and Nagar (2000) Theorem 3.3.24 p.107, Mathai (1997) Ex. 1.25 p.58).

We now introduce Jacobi ensembles. For $n_1, n_2 \ge 1$ and $r \le n := n_1 + n_2$, we can decompose every $(n_1 + n_2) \times r$ matrix M in two blocks

$$M = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$$

with M_1 of type $n_1 \times r$ and M_2 of type $n_2 \times r$. If the entries of M are independent $\mathcal{N}(0, 1)$, then $W_1 := M_1'M_1$ and $W_2 := M_2'M_2$ are independent Wishart matrices of distribution $W_r(n_1, \mathbb{R})$ and $W_r(n_2, \mathbb{R})$, respectively. It is well known that $W_1 + W_2$ is $W_r(n_1 + n_2, \mathbb{R})$ distributed and a.s. invertible. Denote by $(W_1 + W_2)^{1/2}$ the symmetric positive square root of $(W_1 + W_2)$. The $r \times r$ matrix

$$\mathcal{X} := (W_1 + W_2)^{-1/2}W_1(W_1 + W_2)^{-1/2}$$

has a distribution denoted by $J_r(n_1, n_2, \mathbb{R})$ and called the Jacobi ensemble.

If T is upper triangular with positive diagonal entries and $W_1 + W_2 = T'T$ (Cholesky decomposition) then

$$\mathcal{Z} = (T')^{-1}W_1T^{-1}$$

is also $J_r(n_1, n_2, \mathbb{R})$ distributed, (see Olkin and Rubin (1964), Muirhead (1982) p.108).

Another occurrence of the Jacobi ensemble is interesting (see Doumerc (2005), Collins (2005)). If M is as above, its singular value decomposition is

$$M = UDV \quad , \quad D = \begin{pmatrix} \Delta \\ 0 \end{pmatrix}$$

with D of type $n \times r$, with Δ diagonal with nonnegative entries, with $U \in \mathcal{O}(n)$ and $V \in \mathcal{O}(r)$ (the orthogonal groups). Although U and V are not uniquely determined, one can choose them according to the Haar distribution on their respective group and such that U, V, Δ are independent. Then $M'M = V'\Delta^2V$ and

$$(W_1 + W_2)^{1/2} = (M'M)^{1/2} = V'\Delta V.$$

Let $Y_r = U^{[n_1,r]}$ be the $n_1 \times r$ upper-left corner of U. Since $M_1 = Y_r \Delta V$ we have

$$M_1' M_1 = V' \Delta Y_r' Y_r \Delta V = (MM^*)^{1/2} (V' Y_r' Y_r V)(MM^*)^{1/2}$$

and then $\mathcal{X} = (Y_r V)'(Y_r V) \overset{(d)}{=} Y_r' Y_r$. In other words,

$$\mathcal{Y} := \left(U^{[n_1,r]}\right)' U^{[n_1,r]}$$

is also $J_r(n_1, n_2, \mathbb{R})$ distributed.

If $r \le \min(n_1, n_2)$, the distribution $J_r(n_1, n_2, \mathbb{R})$ has a density on $\mathcal{S}_r$ which is

$$\frac{1}{\beta_r\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} |\mathcal{Z}|^{\frac{n_1 - r - 1}{2}} \; |I_r - \mathcal{Z}|^{\frac{n_2 - r - 1}{2}} \; \mathbf{1}_{0 < \mathcal{Z} < I_r}\,, \tag{2.2}$$

where

$$\beta_r(a,b) = \frac{\Gamma_r(a)\,\Gamma_r(b)}{\Gamma_r(a+b)}\,,$$

(see for example Muirhead (1982) Theorem 3.3.1). It is the matrix variate extension of the Beta distribution. Recall that for $a > 0,\ b > 0$, the Beta(a,b) distribution has density

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, x^{a-1}(1-x)^{b-1} \quad (x > 0)\,. \tag{2.3}$$

Until now, we had r fixed. Our purpose is now to consider all values of r simultaneously to give a "sample path" study of determinants.

For an $n \times n$ matrix $B = [b_1, \dots, b_n]$, we have for $r \le n$

$$(B'B)^{[r]} = \left(B^{[n,r]}\right)' B^{[n,r]}\,,$$

and for every $j \le n$, the quantity

$$\rho_{j,n} := \frac{|(B'B)^{[j]}|}{|(B'B)^{[j-1]}|} \tag{2.4}$$

is a measurable function of $(b_1, \dots, b_j)$ and

$$|(B'B)^{[r]}| = \prod_{j=1}^{r} \rho_{j,n}\,. \tag{2.5}$$

The same occurs with $\widetilde{b}_i := b_i/\|b_i\|$ instead of b_i $(i = 1, \dots, n)$ and $\widetilde{B} := [\widetilde{b}_1, \dots, \widetilde{b}_n]$ instead of B. Note that $\widetilde{\rho}_{1,n} = 1$ and

$$\widetilde{\rho}_{j,n} = \frac{|\widetilde{W}^{[j]}|}{|\widetilde{W}^{[j-1]}|} = \frac{|W^{[j]}|}{|W^{[j-1]}| W_{jj}} = \frac{\rho_{j,n}}{\|b_j\|^2},\ \ j = 2, \dots, n\,, \tag{2.6}$$

so that

$$|(\widetilde{B}'\widetilde{B})^{[r]}| = \prod_{j=1}^{r} \widetilde{\rho}_{j,n}\,. \tag{2.7}$$

The Wishart case and the Uniform Gram case corresponds to (2.5) and (2.7) respectively, for $r = 1, \dots, n$.

In the Jacobi case, $r \in \{1, \dots, n_1\}$. If $M = [b_1, \dots, b_{n_1}]$, and if T, W_1, $\mathcal{Z}$ are defined as above with n_1 instead of r, then

$$\mathcal{Z}^{[r]} = \left(\left(T^{[r]}\right)'\right)^{-1} W_1^{[r]} \left(T^{[r]}\right)^{-1}.$$

For every j, the quantity

$$\rho^{\mathcal{Z}}_{j,n_1,n_2} := \frac{|\mathcal{Z}^{[j]}|}{|\mathcal{Z}^{[j-1]}|}$$

is a measurable function of $(b_1,\dots,b_j)$ and

$$|\mathcal{Z}^{[r]}| = \prod_{j=1}^{r} \rho^{\mathcal{Z}}_{j,n_1,n_2}\,.$$

It can be noticed that

$$\rho^{\mathcal{Z}}_{j,n_1,n_2} = \frac{|W_1^{[j]}|}{|W_1^{[j]}+W_2^{[j]}|} \times \frac{|W_1^{[j-1]}+W_2^{[j-1]}|}{|W_1^{[j-1]}|}\,.$$

Besides, the construction with the symmetric square root is different. If

$$\rho^{\mathcal{X}}_{j,n_1,n_2} := \frac{|\mathcal{X}^{[j]}|}{|\mathcal{X}^{[j-1]}|}$$

we have

$$\mathcal{X}^{[r]} \neq \left(W_1^{[r]}+W_2^{[r]}\right)^{-1/2} W_1^{[r]} \left(W_1^{[r]}+W_2^{[r]}\right)^{-1/2}.$$

(Take $n_1 = n_2 = 2$, $W_1 = I_2$, $W_2 = \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix}$ and $r = 1$ then

$$\left(W_1^{[1]}+W_2^{[1]}\right)^{-1/2} W_1^{[r]} \left(W_1^{[r]}+W_2^{[r]}\right)^{-1/2} = 2/(4-s^2)\,,$$

and $\mathcal{X}^{[1]} = 1/2$). Moreover we cannot say that $\rho^{\mathcal{X}}_{j,n_1,n_2}$ is measurable with respect to $b_1,\dots,b_j$.

Consider the construction from contraction of Haar matrices. Since

$$\left(\left(U^{[n_1]}\right)' U^{[n_1]}\right)^{[n_1,r]} = \left(U^{[n_1,r]}\right)' U^{[n_1,r]}\,,$$

we see that the quantity

$$\rho^{\mathcal{Y}}_{n_1,n_2,j} := \frac{|\mathcal{Y}^{[j]}|}{|\mathcal{Y}^{[j-1]}|}$$

depends only on the j first columns of the matrix U, and

$$|\mathcal{Y}^{[r]}| = \prod_{j=1}^{r} \rho^{\mathcal{Y}}_{n_1,n_2,j}\,.$$

It is possible to introduce a probability space on which all Uniform Gram and Wishart matrices are defined for all values of n simultaneously. It is enough to consider the infinite product space generated by a double infinite sequence of independent $\mathcal{N}(0,1)$ variables $\{b_{i,j}\}_{i,j=1}^{\infty}$, and for every n to perform the above constructions with $b_i = (b_{1i},\dots,b_{n,i})'$. To embed the Jacobi matrices in this framework, we have to restrict ourselves to the $\mathcal{X}$-type and $\mathcal{Z}$-type ones; however, only the $\mathcal{Z}$ one gives a natural meaning to the dynamic study.

The starting point of our study of random determinants is the following proposition which gathers known results about the factors entering into the above decompositions.

Proposition 2.1. 1) *(Bartlett) The random variables* $\rho_{j,n}$, $j = 1, \dots, n$ *are independent and*

$$\rho_{j,n} \overset{(d)}{=} Gamma\Big(\frac{n-(j-1)}{2}, \frac{1}{2}\Big),$$

2) *The random variables* $\widetilde{\rho}_{j,n}$, $j = 2, \dots, n$ *are independent and*

$$\widetilde{\rho}_{j,n} \overset{(d)}{=} Beta\Big(\frac{n-j+1}{2}, \frac{j-1}{2}\Big).$$

3) *For* $J = \mathcal{X}$ *(resp.* $\mathcal{Y}$, $\mathcal{Z}$*), the random variables* ρ^J_{j,n_1,n_2} , $j = 1, \dots, n_1$ *are independent and*

$$\rho^J_{j,n_1,n_2} \overset{(d)}{=} Beta\Big(\frac{n_1-j+1}{2}, \frac{n_2}{2}\Big).$$

The first claim is known as the celebrated Bartlett decomposition (stated with χ^2 distributions) (Bartlett, 1933). It is quoted in many books and articles, in particular Anderson (2003) pp.170-172, Muirhead (1982) Theorem 3.2.14 p.99, Kshirsagar (1972), Gupta and Nagar (2000) Theorem 3.3.4 p.91 and ex. 3.8 p.127. The second claim may be found in Anderson (2003) Theorem 9.3.3. In the third claim, we first note that it is enough to get the proof for $\mathcal{Z}$ since the three random matrices have the same distribution. It is a consequence of a result quoted in Anderson (2003), due to Kshirsagar, is proved in Muirhead (1982) Theorem 3.3.1 p.110 under the assumption $r \le n_1, n_2$ and in Rao (1973) p.541 under the only assumption $r \le n_1$. Actually (see Muirhead (1982) ex. 3.24 and Anderson (2003) Theorem 8.4.1), some proofs use probabilistic arguments (as Rao (1973) and Anderson (2003)), Jacobian arguments (as in Gupta and Nagar (2000) Theorem 5.3.24 p.181), or Mellin transform arguments (as in Mathai (1999) Theorem 2).

Note that for determinants in other ensembles, Mehta (2004), in his Section 15.4 and his Chapter 26, gives Mellin transforms and density functions.

2.2. *Distribution of eigenvalues and β-models.* In the study of stationary processes, random matrices of the Wishart type with complex entries play an important role (Goodman (1963)). In some papers, quaternionic entries are considered (Kabe (1984), Andersson et al. (1983), Hanlon et al. (1992)). The above constructions and results can be extended to the complex and quaternionic cases by replacing the transpose by the adjoint and the dual, respectively and by replacing the factor $1/2$ in the parameters of distributions by the factors 1 and 2, respectively. We do not give details but jump to a general framework.

Popularized by physicists, the modern point of view consists in introducing a parameter $\beta > 0$ taking value 1 in the real case, 2 in the complex case and 4 in the quaternionic case, this parameter playing the role of an inverse temperature. The distribution of eigenvalues becomes a distribution of charges and the model is a Coulomb gas or β-model, see for instance Forrester (2007) Chap. 2. Because of the connections with orthogonal polynomials in the complex case, the extended families are called β-Laguerre ensemble (or just Laguerre ensemble) instead of Wishart ensemble and β-Jacobi ensemble (or just Jacobi ensemble). As mentioned in Section 1, they correspond also to models of tri-diagonal random matrices (see Dumitriu and Edelman (2003), Killip and Nenciu (2004), Edelman and Sutton (2007)).

For $\beta = 1, 2$ and 4, there are two ways of reaching the law of determinants:

a) from the joint distribution of eigenvalues (spectral method).

b) from the distribution of matrices, and using the decomposition method.

The spectral method is easily extended to β-models, considering products of charges. We will see in this section that actually a product decomposition holds true also.

Throughout, we use the symbol β' for $\beta/2$ to simplify displays.

2.2.1. *Laguerre.* When $\beta = 1, 2, 4$ the joint probability density of the eigenvalues $\lambda_j, j = 1, \ldots, r$ of W on the orthant $\lambda_j > 0,\ j = 1, \ldots, r$ is

$$\frac{1}{Z_r^{L,\beta}(n)} \prod_{j=1}^{r} \left(\lambda_j^{\beta'(n-r+1)-1} e^{-\beta'\lambda_j}\right) \prod_{1\le j<k\le r} |\lambda_k - \lambda_j|^{2\beta'} \quad , \tag{2.8}$$

and the normalizing constant is

$$Z_r^{L,\beta}(n) = \left(\frac{1}{\beta'}\right)^{\beta' rn} \prod_{j=1}^{r} \frac{\Gamma\left(1+\beta' j\right)\Gamma\left(\beta'(n-j+1)\right)}{\Gamma\left(1+\beta'\right)}\,.$$

(It is the inverse of the Laguerre form of the Selberg integral, see Mehta (2004) formula 17.6.5 or Hiai and Petz (2000) p.118).

When $\beta > 0$ is not $1, 2, 4$, formula (2.8) still gives a density on $(0, \infty)^r$. We also denote the product $\prod_{j=1}^r \lambda_j$ by $|W|$. Its Mellin transform is

$$\mathbb{E}|W|^{\beta' s} = \frac{Z_r^{L,\beta}(n+s)}{Z_r^{L,\beta}(n)} = \left(\frac{1}{\beta'}\right)^{\beta' rs} \prod_{k=1}^{r} \frac{\Gamma\left(\beta'(n-k+1+s)\right)}{\Gamma\left(\beta'(n-k+1)\right)}\,.$$

Recalling that if $X \overset{(d)}{=} \text{Gamma}(a, 1/2)$ then

$$\mathbb{E}X^\mu = 2^\mu \frac{\Gamma(\mu+a)}{\Gamma(a)} \quad (\mu > -a)\,,$$

we deduce the following proposition from the uniqueness of Mellin transform.

Proposition 2.2. *We have*

$$|W| \overset{(d)}{=} \prod_{j=1}^{r} \rho_{j,n}^{L,\beta}\,,$$

where the variables $\rho_{j,n}^{L,\beta}, j = 1, \ldots, r$ *are independent and*

$$\rho_{j,n}^{L,\beta} \overset{(d)}{=} \textit{Gamma}\left(\beta'(n-j+1), 1/2\right)\,. \tag{2.9}$$

We point out that our point of view is not compatible with the construction by (Dumitriu and Edelman (2003)) of matrix models for the (general) β-Laguerre ensemble. Actually, they define a random $r \times r$ matrix $B^{(r)}$ where only diagonal and subdiagonal terms are nonzero, independent and satisfy (for n fixed):

$$\begin{aligned} B_{ii}^{(r)} &\overset{(d)}{=} \sqrt{\text{Gamma}\left(\beta'(n-i+1), 1/2\right)} \quad (1 \le i \le r)\,, \\ B_{i,i-1}^{(r)} &\overset{(d)}{=} \sqrt{\text{Gamma}\left(\beta'(r-i+1), 1/2\right)} \quad (2 \le i \le r)\,. \end{aligned}$$

They prove that the distribution of eigenvalues of $B^{(r)}\left(B^{(r)}\right)'$ is precisely (2.8). Of course we recover the determinant as a product of elements with the good distribution, but the problem is that we cannot consider all r simultaneously in their framework, since

$$\left(B^{(r)}\left(B^{(r)}\right)'\right)^{[r-1]} \neq B^{(r-1)}\left(B^{(r-1)}\right)'.$$

2.2.2. *Uniform Gram.* It is useful in the study of correlations. A correlation matrix is a positive definite matrix with diagonal entries equal to one. Here, there is no explicit expression for the law of eigenvalues. However, the expression

$$\frac{1}{Z_r^{G,\beta}(n)}\,|G|^{\beta'(n-r+1)-1}$$

with

$$Z_r^{G,\beta}(n) = \pi^{\beta' r(r-1)}\prod_{j=1}^{r}\frac{\Gamma\left(\beta'(n-j+1)\right)}{\Gamma\left(\beta' n\right)}$$

is a density on the space of symmetric (resp. Hermitian, resp. self-dual) positive matrices with diagonal entries equal to one, and it coincides with the distribution of correlation matrix in the real (see (2.1)), complex and quaternion case, for the appropriate values of β. This yields (Gupta and Nagar (2000) ex. 3.26 p.130) the Mellin transform

$$\mathbb{E}|G|^{\beta' s} = \frac{Z_r^{G,\beta}(n+s)}{Z_r^{G,\beta}(n)} = \prod_{j=1}^{r}\frac{\Gamma\left(\beta'(n-j+1+s)\right)\Gamma\left(\beta' n\right)}{\Gamma\left(\beta'(n-j+1)\right)\Gamma\left(\beta'(n+s)\right)}\,.$$

From (2.3), it is clear that if $X \overset{(d)}{=} \mathrm{Beta}(a,b)$ then

$$\mathbb{E}X^{\mu} = \frac{\Gamma(a+\mu)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+\mu)} \quad (\mu > -a)\,. \tag{2.10}$$

Again the uniqueness of the Mellin transform leads to the proposition.

Proposition 2.3. *We have*

$$|G| \overset{(d)}{=} \prod_{j=2}^{r}\rho_{j,n}^{G,\beta}$$

where the variables $\rho_{j,n}^{G,\beta}, j = 2,\ldots,r$ *are independent and*

$$\rho_{j,n}^{G,\beta} \overset{(d)}{=} \mathit{Beta}\left(\beta'(n-j+1), \beta'(j-1)\right)\,. \tag{2.11}$$

The above product is meaningful for every $\beta > 0$, and then for $\beta \neq 1,2,4$ we define $|G|$ as a random variable with the designated distribution.

2.2.3. *Jacobi.* If $\mathcal{Z}$ is distributed as in (2.2), the joint density of eigenvalues on the set $(0 < \lambda_j < 1\;,\; j = 1,\ldots,r)$ is given by

$$\frac{1}{Z_r\left(n_1,n_2\right)}\prod_{i=1}^{r}\lambda_i^{\frac{n_1-r-1}{2}}\left(1-\lambda_i\right)^{\frac{n_2-r-1}{2}}\prod_{1\le i<j\le r}|\lambda_j-\lambda_i|\,,$$

where $Z_r\left(n_1,n_2\right)$ is a normalizing constant (see Muirhead (1982) Theorem 3.3.4).

For $n_2 < r < n_1$, W_2 is singular and $\mathcal{Z}$ has 1 as an eigenvalue of multiplicity $r - n_2$, so that the distribution of $\mathcal{Z}$ has no density. Nevertheless we may study its determinant. Indeed, the matrix $I_r - \mathcal{Z}$ has 0 as an eigenvalue of multiplicity

$r - n_2$. Actually the density of the law of the non-zero eigenvalues of this matrix is known (see Srivastava (2003) and Diaz-Garcia and Gutierrez Jaimez (1997)), so that the non-one eigenvalues of $\mathcal{Z}$ have the joint density

$$\frac{1}{\widetilde{Z}_r\left(n_1, n_2\right)} \prod_{i=1}^{n_2} \lambda_i^{\frac{n_1-r-1}{2}}\left(1-\lambda_i\right)^{\frac{r-n_2-1}{2}} \prod_{1 \leq i<j \leq n_2}\left|\lambda_j-\lambda_i\right|,$$

where the normalizing constant is $\widetilde{Z}_r\left(n_1, n_2\right)=Z_{n_2}\left(n_1+n_2-r, r\right)$.

We now consider matrices with elements in $\mathbb{C}$ or $\mathbb{H}$. When $r \leq \min(n_1, n_2)$, the distribution of $\mathcal{Z}$ has a density proportional to

$$|\mathcal{Z}|^{\beta'(n_1-r+1)-1}\left|I_r-\mathcal{Z}\right|^{\beta'(n_2-r+1)-1} \mathbf{1}_{0<\mathcal{Z}<I_r} .$$

where $\beta'=1$ or 2. The joint density of the eigenvalues of $\mathcal{Z}$ is (on $[0,1]^r$):

$$f_{r, n_1, n_2}^{\beta}\left(\lambda_1, \ldots, \lambda_r\right)= \tag{2.12}$$

$$\frac{1}{Z_r^{(J, \beta)}\left(n_1, n_2\right)} \prod_{i=1}^r \lambda_i^{\beta'(n_1-r+1)-1}\left(1-\lambda_i\right)^{\beta'(n_2-r+1)-1} \prod_{1 \leq i<j \leq r}\left|\lambda_j-\lambda_i\right|^{2 \beta'},$$

where

$$Z_r^{(J, \beta)}(n_1, n_2)=\prod_{j=1}^r \frac{\Gamma\left(1+\beta' j\right) \Gamma\left(\beta'(n_1+j-r)\right) \Gamma\left(\beta'(n_2+j-r)\right)}{\Gamma\left(1+\beta'\right) \Gamma\left(\beta'(n_1+n_2+j-r)\right)}, \tag{2.13}$$

is the value of the Selberg integral (see Mehta (2004) formula 17.1.3 or Hiai and Petz (2000) p.118). In the singular case ($n_2 \leq r \leq n_1$), the density of the non-one eigenvalues is $f_{n_2, n_1+n_2-r, r}^{\beta}(\lambda_1, \ldots, \lambda_{n_2})$.

We consider an extension of the above models. For every $\beta>0$, we define a family of distribution densities $\mathbf{f}_{r, n_1, n_2}^{\beta}$ on $[0,1]^{\min(n_2, r)}$:

$$\mathbf{f}_{r, n_1, n_2}^{\beta}= \begin{cases} f_{r, n_1, n_2}^{\beta} & \text{if } r \leq \min(n_1, n_2) \\ f_{n_2, n_1+n_2-r, r}^{\beta} & \text{if } n_2 \leq r \leq n_1 .\end{cases} \tag{2.14}$$

We set by convention

$$|\mathcal{Z}|=\prod_{i=1}^{\min(n_2, r)} \lambda_i$$

in all cases, and we call it the determinant, even if we do not define any matrix.

For $r \leq n_1, n_2$, using (2.12) and (2.13) we obtain

$$\begin{aligned} \mathbb{E}\left(|\mathcal{Z}|^{\beta' s}\right) & =\frac{Z_r^{J, \beta}\left(n_1+s, n_2\right)}{Z_r^{J, \beta}\left(n_1, n_2\right)} \\ & =\prod_{j=1}^r \frac{\Gamma\left(\beta'(n_1+n_2+j-r)\right) \Gamma\left(\beta'(n_1+j-r+s)\right)}{\Gamma\left(\beta'(n_1+j-r)\right) \Gamma\left(\beta'(n_1+n_2+j-r+s)\right)} . \end{aligned} \tag{2.15}$$

If $n_2<r \leq n_1$, starting directly from (2.14) and (2.13) we have

$$\begin{aligned} \mathbb{E}\left(|\mathcal{Z}|^{\beta' s}\right) & =\frac{Z_{n_2}^{J, \beta}\left(n_1+n_2-r+s, r\right)}{Z_{n_2}^{J, \beta}\left(n_1+n_2-r, r\right)} \\ & =\prod_{j=1}^{n_2} \frac{\Gamma\left(\beta'(n_1+j)\right) \Gamma\left(\beta'(n_1+j-r+s)\right)}{\Gamma\left(\beta'(n_1+j-r)\right) \Gamma\left(\beta'(n_1+j+s)\right)} . \end{aligned} \tag{2.16}$$

Multiplying up and down by $\prod_{k=n_2+1}^{r} \Gamma\left(\beta'(n_1+k-r)\right) \Gamma\left(\beta'(n_1+k-r+s)\right)$ yields the right hand side of (2.15). In view of (2.10) and the unicity of the Mellin transform, we have the proposition.

Proposition 2.4. *For $r \le n_1$,*

$$|\mathcal{Z}| \overset{(d)}{=} \prod_{j=1}^{r} \rho^{\beta,J}_{j,n_1,n_2}\,,$$

where $\rho^{\beta,J}_{j,n_1,n_2}$, $j = 1,\dots,r$ are independent and

$$\rho^{\beta,J}_{j,n_1,n_2} \overset{(d)}{=} \mathit{Beta}\left(\beta'(n_1-j+1), \beta' n_2\right) . \tag{2.17}$$

2.3. *Processes.* In the three ensembles defined above, we have met arrays of independent variables with remarkable distributions. In Section 2.1, we have discussed the interest of studying all values of r simultaneously in the matrix cases ($\beta = 1, 2, 4$). Since the structure remains the same in the β-ensembles, it is meaningful to consider the processes (indexed by r) of partial sums. A now classical asymptotic regime is $n, r \to \infty$ with fixed ratio in the Laguerre and Uniform Gram case, and $n_1, n_2, r \to \infty$ with fixed ratios in the Jacobi case. It means that we consider the asymptotic behavior determinants in a dynamic (or pathwise) way.

For the Laguerre ensemble, we define

$$\log \Delta^{L,\beta}_{n,p} := \sum_{k=1}^{p} \log \frac{\rho^{L,\beta}_{k,n}}{\beta n} \qquad (p \le n) \tag{2.18}$$

and the process

$$\Delta^{L,\beta}_{n}(t) := \Delta^{L,\beta}_{n,\lfloor nt \rfloor}, \quad t \in [0,1]\,. \tag{2.19}$$

For the Uniform Gram ensemble, we define

$$\log \Delta^{G,\beta}_{n,p} := \sum_{k=1}^{p} \log \rho^{G,\beta}_{k,n} \qquad (p \le n) \tag{2.20}$$

and the process

$$\Delta^{G,\beta}_{n}(t) := \Delta^{G,\beta}_{n,\lfloor nt \rfloor}, \quad t \in [0,1]\,. \tag{2.21}$$

For the Jacobi ensemble, we fix τ_1 and $\tau_2 > 0$, set $n_1 = \lfloor n\tau_1 \rfloor, n_2 = \lfloor n\tau_2 \rfloor$, and define

$$\log \Delta^{J,\beta}_{n,p} := \sum_{k=1}^{p} \log \rho^{J,\beta}_{k,n_1,n_2} \qquad (p \le n_1) \tag{2.22}$$

and the process

$$\Delta^{J,\beta}_{n}(t) := \Delta^{J,\beta}_{n,\lfloor nt \rfloor}, \quad t \in [0,\tau_1]\,. \tag{2.23}$$

There are some connections between the above processes. For instance, in the real matrix ensemble ($\beta = 1$) we saw in (2.6) that

$$\rho^{L,1}_{j,n} = \rho^{G,1}_{j,n}\ \|b_j\|^2\,.$$

From elementary properties of the $\mathcal{N}(0, I_n)$ distribution, we know also that $\|b_j\|^2$ is independent of $\rho^{G,1}_{j,n}$ and Gamma$(n/2, 1/2)$ distributed. To see these connections in the general case, we use the so-called "beta-gamma" algebra that will be really helpful in the sequel. Details can be found in Chaumont and Yor (2003) pp.93-94. In the following relation, $\gamma(c)$ denotes a variable with distribution Gamma$(c, 1)$, and $\beta(a, b)$ denotes a variable with distribution Beta(a, b). The relation is

$$\big(\gamma(a), \gamma(b)\big) \quad \stackrel{(d)}{=} \quad \big(\beta(a, b)\gamma(a+b), (1-\beta(a, b))\gamma(a+b)\big)\,, \tag{2.24}$$

where, on the left hand side the variables $\gamma(a)$ and $\gamma(b)$ are independent and on the right hand side the variables $\beta(a, b)$ and $\gamma(a+b)$ are independent. It entails in particular

$$\frac{\gamma(a)}{\gamma(a)+\gamma(b)} \quad \stackrel{(d)}{=} \quad \beta(a, b)\,. \tag{2.25}$$

Note that this relation can be extended to the matrix variate level.

From definitions (2.18) and (2.20) and owing to the equalities in distribution (2.9) and (2.11), we have then

$$\log \Delta_n^{L,\beta} \stackrel{(d)}{=} \log \Delta_n^{G,\beta} + S_n\,, \tag{2.26}$$

where S_n is independent of $\log \Delta_n^{G,\beta}$, and specified by

$$S_n(t) = \sum_{k=1}^{\lfloor nt \rfloor} \log \varepsilon_k^{(n)} \quad,\ t \in [0, 1] \tag{2.27}$$

where $\varepsilon_k^{(n)}$, $k = 1, \dots, n$ are independent and satisfy $\varepsilon_k^{(n)} \stackrel{(d)}{=} \text{Gamma}\,(\beta' n, \beta' n)$. In the sequel, we begin with the Uniform Gram ensemble and then deduce the corresponding results for the Laguerre ensemble.

For the Jacobi ensemble, we use definitions (2.18) and (2.22) and equalities in distribution (2.9) and (2.17). We get, by another application of (2.24)

$$\log \Delta_{n_1,r}^{L,\beta} \stackrel{(d)}{=} \log \Delta_{n,r}^{J,\beta} + \log \Delta_{n_1+n_2,r}^{L,\beta} - r \log \frac{n_1}{n_1+n_2}\,, \tag{2.28}$$

where this equality holds for all indices $r = 1, \dots, n_1$ simultaneously, and the two processes $\log \Delta_n^{J,\beta}$ and $\log \Delta_{n_1+n_2}^{L,\beta}$ are independent. It allows to deduce asymptotic results for the Jacobi ensemble from those of the Laguerre ensemble.

3. Main results

In this section, we state a law of large numbers and fluctuations for processes and marginals in our three models (Section 3.1), and then the corresponding large deviations (Section 3.2).

Let $D_T = \{v \in \mathbb{D}([0, T]) : v(0) = 0\}$ denote the set of càdlàg functions on $[0, T]$ and $D = \{v \in \mathbb{D}([0, 1)) : v(0) = 0\}$ the set of càdlàg functions on $[0, T]$ and $[0, 1)$, respectively, starting from 0.

We use often the following entropy function

$$\mathcal{J}(u) = \begin{cases} u \log u - u + 1 & \text{if } u > 0 \\ 1 & \text{if } u = 0 \\ +\infty & \text{if } u < 0 \end{cases} \tag{3.1}$$

and its primitive:

$$F(t) = \int_0^t \mathcal{J}(u)\, du = \frac{t^2}{2}\log t - \frac{3t^2}{4} + t, \quad (t \geq 0)\,. \tag{3.2}$$

We use also the function defined in Hiai and Petz (2006), for $s, t \geq 0$:

$$\begin{aligned} B(s,t) \quad &:= \quad \frac{(1+s)^2}{2}\log(1+s) - \frac{s^2}{2}\log s + \frac{(1+t)^2}{2}\log(1+t) - \frac{t^2}{2}\log t \\ &- \quad \frac{(2+s+t)^2}{2}\log(2+s+t) + \frac{(1+s+t)^2}{2}\log(1+s+t)\,, \end{aligned} \tag{3.3}$$

which may also be written as

$$B(s,t) = F(1+s) - F(s) + F(1+t) - F(t) - F(2+s+t) + F(1+s+t) - \frac{7}{4}\,.$$

3.1. *Law of large numbers and fluctuations.*

3.1.1. *Uniform Gram ensemble.* Define a drift and a diffusion coefficient by

$$\mathtt{d}^{G,\beta}(t) := \frac{1}{\beta} + \left(\frac{1}{2} - \frac{1}{\beta}\right)\frac{1}{1-t} \quad , \quad \sigma^{G,\beta}(t) := \sqrt{\frac{2t}{\beta(1-t)}} \qquad (t<1)\,. \tag{3.4}$$

Theorem 3.1. (1) *As $n \to \infty$,*

$$\lim_n \sup_{p \leq n} \left| \frac{1}{n}\mathbb{E}\,\log \Delta_{n,p}^{G,\beta} + \mathcal{J}\left(1 - \frac{p}{n}\right)\right| = 0\,. \tag{3.5}$$

(2) *For every $t \in [0,1)$, as $n \to \infty$,*

$$\mathbb{E}\,\log \Delta_n^{G,\beta}(t) + n\mathcal{J}\left(1 - \frac{\lfloor nt \rfloor}{n}\right) \to \int_0^t \mathtt{d}^{G,\beta}(s)\, ds \tag{3.6}$$

and

$$\mathbb{E}\,\log \Delta_n^{G,\beta}(1) + n + \left(\frac{1}{\beta} - \frac{1}{2}\right)\log n \to K_\beta^1\,, \tag{3.7}$$

where

$$K_\beta^1 := \frac{1}{2}\log(2\pi) + \frac{1-\gamma}{\beta} - \int_0^\infty \frac{sf(s)}{e^{\beta s/2}-1}\, ds\,, \tag{3.8}$$

and $\gamma = -\Gamma'(1)$ is the Euler constant.

(3) *For every $t \in [0,1)$, as $n \to \infty$,*

$$Var\ \log \Delta_n^{G,\beta}(t) \quad \to \quad \int_0^t \left(\sigma^{G,\beta}(s)\right)^2\, ds \tag{3.9}$$

$$Var\ \log \Delta_n^{G,\beta}(1) - \frac{2}{\beta}\log n \quad \to \quad K_\beta^2\,, \tag{3.10}$$

where

$$K_\beta^2 := \frac{2(\gamma-1)}{\beta} + \int_0^\infty \frac{s(sf(s)+\frac{1}{2})}{e^{\beta s/2}-1}\, ds\,. \tag{3.11}$$

(4) *As* $n \to \infty$,

$$\lim_n \sup_{t\in[0,1]} \left|\frac{\log \Delta_n^{G,\beta}(t)}{n} + \mathcal{J}(1-t)\right| = 0\,. \tag{3.12}$$

in probability.

For $\beta = 1$, formulae (3.7) and (3.10) are due to Abbott and Mulders (2001) (see their lemmas 4.2 and 4.4), using a variant of the decomposition method.

Theorem 3.2. (1) *Let for* $n \geq 1$

$$\eta_n^{G,\beta}(t) := \log \Delta_n^{G,\beta}(t) + n\mathcal{J}\left(1 - \frac{\lfloor nt \rfloor}{n}\right) \quad , \quad t \in [0,1)\,.$$

Then as $n \to \infty$

$$\left(\eta_n^{G,\beta}(t);\ t \in [0,1)\right) \Rightarrow \left(X_t^{G,\beta};\ t \in [0,1)\right), \tag{3.13}$$

where $X^{G,\beta}$ *is the (Gaussian) diffusion solution of the stochastic differential equation:*

$$dX_t^{G,\beta} = \mathtt{d}^{G,\beta}(t)\, dt + \sigma^{G,\beta}(t)\ d\mathbf{B}_t\,, \tag{3.14}$$

with $X_0^{G,\beta} = 0$, $\mathbf{B}$ *is a standard Brownian motion and* $\Rightarrow$ *stands for the weak convergence of distributions in* D *endowed with the Skorokhod topology.*

(2) *Let*

$$\widehat{\eta}_n^{G,\beta} = \frac{\log \Delta_n^{G,\beta}(1) + n + \left(\frac{1}{\beta} - \frac{1}{2}\right)\log n}{\sqrt{\frac{2}{\beta}\log n}}\quad .$$

Then as $n \to \infty$, $\widehat{\eta}_n^{G,\beta} \Rightarrow N$ *where* N *is* $\mathcal{N}(0,1)$ *and independent of* $\mathbf{B}$, *(and* $\Rightarrow$ *stands for the weak convergence of distributions in* $\mathbb{R}$*).*

3.1.2. *Laguerre ensemble.* Define a drift and a diffusion coefficient by

$$\mathtt{d}^{L,\beta}(t) := \left(\frac{1}{2} - \frac{1}{\beta}\right)\frac{1}{1-t} \quad , \quad \sigma^{L,\beta}(t) := \sqrt{\frac{2}{\beta(1-t)}} \quad (t < 1)\,. \tag{3.15}$$

Theorem 3.3. (1) *As* $n \to \infty$,

$$\lim_n \sup_{p\leq n} \left|\frac{1}{n}\,\mathbb{E}\log \Delta_{n,p}^{L,\beta} + \mathcal{J}\left(1 - \frac{p}{n}\right)\right| = 0\,. \tag{3.16}$$

(2) *For every* $t \in [0,1)$, *as* $n \to \infty$,

$$\mathbb{E}\ \log \Delta_n^{L,\beta}(t) + n\mathcal{J}\left(1 - \frac{\lfloor nt \rfloor}{n}\right) \to \int_0^t \mathtt{d}^{L,\beta}(s)\ ds\,, \tag{3.17}$$

and

$$\mathbb{E}\ \log \Delta_n^{L,\beta}(1) + n + \left(\frac{1}{\beta} - \frac{1}{2}\right)\log n \to K_\beta^1 - \frac{1}{\beta}\,, \tag{3.18}$$

(3) *For every* $t \in [0,1)$*, as* $n \to \infty$*,*

$$\begin{aligned} Var \log \Delta_n^{L,\beta}(t) &\to \int_0^t \left(\sigma^{L,\beta}(s)\right)^2 \, ds && (3.19) \\ Var \log \Delta_n^{L,\beta}(1) - \frac{2}{\beta}\log n &\to K_\beta^2 + \frac{2}{\beta}\,. && (3.20) \end{aligned}$$

(4) *As* $n \to \infty$*,*

$$\sup_{t\in[0,1]} \left| \frac{1}{n} \log \Delta_n^{L,\beta}(t) + \mathcal{J}(1-t) \right| \to 0 \tag{3.21}$$

in probability.

Remark 3.4. In the Uniform Gram and Laguerre ensembles, when all the variables are defined on the same space (i.e. $\beta = 1, 2, 4$), an application of the Borel-Cantelli lemma leads to almost sure convergence.

Theorem 3.5. *Let*

$$\begin{aligned} \eta_n^{L,\beta}(t) &:= \log \Delta_n^{L,\beta}(t) + n\mathcal{J}\left(1 - \frac{\lfloor nt \rfloor}{n}\right) \quad , \quad t \in [0,1)\,, \\ \widehat{\eta}_n^{L,\beta} &= \frac{\log \Delta_n^{L,\beta}(1) + n + \left(\frac{1}{\beta} - \frac{1}{2}\right)\log n}{\sqrt{\frac{2}{\beta}\log n}} \quad . \end{aligned}$$

Then as $n \to \infty$

$$\begin{aligned} \left(\eta_n^{L,\beta}(t);\ t \in [0,1)\right) &\Rightarrow \left(X_t^{L,\beta}, \quad t \in [0,1)\right) && (3.22) \\ \widehat{\eta}_n^{L,\beta} &\Rightarrow N \end{aligned}$$

where $X^{L,\beta}$ *is the Gaussian diffusion solution of the stochastic differential equation:*

$$dX_t^{L,\beta} = \mathsf{d}^{L,\beta}(t)\, dt + \sigma^{L,\beta}(t)\ d\mathbf{B}_t\,, \tag{3.23}$$

with $X_0^{L,\beta} = 0$*, where* $\mathbf{B}$ *is a standard Brownian motion and* N *is* $\mathcal{N}(0,1)$ *and independent of* $\mathbf{B}$*.*

The convergence of $\eta_n^{L,1}(t)$, for fixed t and of $\widehat{\eta}_n^{L,1}$ were proved by Jonsson (1982) Theorem 5.1a. Recently and independently the convergence of $\widehat{\eta}_n^{L,1}$ was proved in Theorem 4 of Rempała and Wesołowski (2005).

3.1.3. *Jacobi ensemble.* In this part we use new auxiliary functions. Let

$$\mathcal{E}(x,y,z) = x\log x - (x+y)\log(x+y) + (x+y-z)\log(x+y-z) - (x-z)\log(x-z)$$

or using $\mathcal{J}$ defined in (3.1)

$$\mathcal{E}(x,y,z) = \mathcal{J}(x) - \mathcal{J}(x-z) - \mathcal{J}(x+y) + \mathcal{J}(x+y-z)\,. \tag{3.24}$$

The partial derivative of $\mathcal{E}$ with respect to x is:

$$\mathcal{E}_1(x,y,z) := \frac{\partial}{\partial x}\,\mathcal{E}(x,y,z) = \log\frac{x(x+y-z)}{(x-z)(x+y)}\,. \tag{3.25}$$

Let for $0 \le t < \tau_1$

$$\sigma^2(t) := \frac{\partial}{\partial t}\,\mathcal{E}_1(\tau_1,\tau_2,t) = \frac{\tau_2}{(\tau_1 - t)(\tau_1+\tau_2-t)}\,. \tag{3.26}$$

Again we define drift and diffusion coefficients:

$$\mathtt{d}^{J,\beta}(t) = \left(\frac{1}{2} - \frac{1}{\beta}\right)\sigma^2(t) \quad , \quad \sigma^{J,\beta}(t) = \sqrt{\frac{2}{\beta}}\,\sigma(t)\,.$$

Theorem 3.6. *(1) As $n \to \infty$,*

$$\sup_{t\in[0,\tau_1]}\left|\frac{1}{n}\mathbb{E}\log\Delta_n^{J,\beta}(t) - \mathcal{E}\left(\tau_1,\tau_2,t\right)\right| \to 0\,. \tag{3.27}$$

(2) *For every $t \in [0,\tau_1)$, as $n \to \infty$,*

$$\mathbb{E}\log\Delta_n^{J,\beta}(t) - \mathcal{E}(\lfloor\tau_1 n\rfloor, \lfloor\tau_2 n\rfloor, \lfloor tn\rfloor) \longrightarrow \int_0^t \mathtt{d}_J(s)\ ds\,, \tag{3.28}$$

and[1]

$$\mathbb{E}\log\Delta_n^{J,\beta}(\tau_1) - \mathcal{E}(\lfloor\tau_1 n\rfloor, \lfloor\tau_2 n\rfloor, \lfloor\tau_1 n\rfloor) + \left(\frac{1}{\beta} - \frac{1}{2}\right)\log n \longrightarrow \left(\frac{1}{2} - \frac{1}{\beta}\right)\log\frac{\tau_1\tau_2}{\tau_1+\tau_2} + K_\beta^1 - \frac{1}{\beta}\,. \tag{3.29}$$

.

(3) *For every $t \in [0,\tau_1)$, as $n \to \infty$,*

$$Var\ \log\Delta_n^{J,\beta}(t) \to \int_0^t \left(\sigma^{J,\beta}(s)\right)^2\ ds\,, \tag{3.30}$$

and[1]

$$Var\ \log\Delta_n^{J,\beta}(\tau_1) - \frac{2}{\beta}\log n \longrightarrow \frac{2}{\beta}\log\left(\frac{\tau_1\tau_2}{\tau_1+\tau_2}\right) + K_\beta^2 + \frac{2}{\beta}\,. \tag{3.31}$$

(4) *As $n \to \infty$,*

$$\sup_{t\in[0,\tau_1]}\left|\frac{1}{n}\log\Delta_n^{J,\beta}(t) - \mathcal{E}(\tau_1,\tau_2,t)\right| \to 0 \tag{3.32}$$

in probability.

Remark 3.7. For $\beta = 1,2,4$, when all variables are on the same probability space, the convergence in (4) may be strengthened to almost sure convergence.

Theorem 3.8. *Let for $n \geq 1$*

$$\begin{aligned}
\eta_n^{J,\beta}(t) &:= \log\Delta_n^{J,\beta}(t) - \mathcal{E}(\lfloor\tau_1 n\rfloor, \lfloor\tau_2 n\rfloor, \lfloor tn\rfloor) \quad , \quad t \in [0,\tau_1)\,,\\
\widehat{\eta}_n^{J,\beta} &:= \frac{\log\Delta_n^{J,\beta}(\tau_1) - n\mathcal{E}(\tau_1,\tau_2,\tau_1) + \left(\frac{1}{2} - \frac{1}{\beta}\right)\log n}{\sqrt{\frac{2}{\beta}\log n}}\,.
\end{aligned}$$

Then as $n \to \infty$

$$\begin{aligned}
\left(\eta_n^{J,\beta}(t);\ t \in [0,\tau_1)\right) &\Rightarrow \left(X_t^J;\ t \in [0,\tau_1)\right)\,, \qquad (3.33)\\
\widehat{\eta}_n^{J,\beta} &\Rightarrow N
\end{aligned}$$

where $X^{J,\beta}$ is the (Gaussian) diffusion solution of the stochastic differential equation:

$$dX_t^J = \mathtt{d}^{J,\beta}(t)dt + \sigma^{J,\beta}(t)\ d\mathbf{B}_t\,, \tag{3.34}$$

---

[1] where K_β^1 (resp. K_β^2) was defined in (3.8) (resp. (3.11))

with $X_0^{J,\beta} = 0$*,* $\mathbf{B}$ *is a standard Brownian motion and* N *is a standard normal variable independent of* $\mathbf{B}$*.*

3.2. *Large deviations.* Throughout this section, we use the notation of Dembo and Zeitouni (1998). In particular we write LDP for Large Deviation Principle. The reader may have some interest in consulting Dette and Gamboa (2007) where a similar method is used for a different model, but here we use a slightly different topology to be able to catch the marginals in T.

For $T < 1$, let M_T be the set of signed measures on $[0, T]$ and let $\mathrm{M}_<$ be the set of measures whose support is a compact subset of $[0, 1)$. We provide D with the weakened topology $\sigma(D, \mathrm{M}_<)$. So, D is the projective limit of the family, indexed by $T < 1$ of topological spaces $(D_T, \sigma(D_T, \mathrm{M}_T))$.

Let V_ℓ (resp. V_r) be the space of left (resp. right) continuous $\mathbb{R}$-valued functions with bounded variations. We put a superscript T to specify the functions on $[0, T]$. There is a bijective correspondence between V_r^T and M_T:

- for any $v \in V_r^T$, there exists a unique $\mu \in \mathrm{M}_T$ such that $v = \mu([0, \cdot])$; we denote it by $\dot{v}$,

- for any $\mu \in \mathrm{M}_T$, $v = \mu([0, \cdot])$ stands in V_r.

For $v \in D$, let $\dot{v} = \dot{v}_a + \dot{v}_s$ be the Lebesgue decomposition of the measure $\dot{v}$ in absolutely continuous and singular parts with respect to the Lebesgue measure and let μ be any bounded positive measure dominating $\dot{v}_s$.

For $A \subset [0, 1)$ and $v \in D$ let

$$I_A(v) = \int_A L_a\Big(t, \frac{d\dot{v}_a}{dt}(t)\Big)dt + \int_A L_s\Big(t, \frac{d\dot{v}_s}{d\mu}(t)\Big)d\mu(t) \quad if \quad v \in V_r\,, \tag{3.35}$$

and $I_A(v) = \infty$ if $v \in D \setminus V_r$, where functions $L_a(t, x)$ and $L_s(t, x)$ will be defined later for each of the ensembles of interest.

3.2.1. *Uniform Gram ensemble.* For the following statement, we need some notation. Let $\mathbf{H}$ be the entropy function:

$$\mathbf{H}(x|p) = x \log \frac{x}{p} + (1 - x) \log \frac{1 - x}{1 - p}\,,$$

and put

$$\begin{aligned} L_a^G(t, y) &= \mathbf{H}(1 - t|e^y)\ \delta(y|(-\infty, 0))\,, \\ L_s^G(t, y) &= -(1 - t)y\ \delta(y|(-\infty, 0))\,, \end{aligned} \tag{3.36}$$

where we set $\delta(y|A) = 0$ if $y \in A$ and $= \infty$ if $y \notin A$.

Theorem 3.9. *The sequence* $\{n^{-1} \log \Delta_n^{G,\beta}(t),\ t \in [0, 1)\}_n$ *satisfies a LDP in* $(D, \sigma(D, \mathrm{M}_<))$ *in the scale* $2\beta^{-1}n^{-2}$ *with good rate function* $I_{[0,1)}^G$*.*

That means, roughly speaking, that

$$\mathbb{P}(\log \Delta_n^{G,\beta} \simeq nv) \approx e^{-\frac{\beta n^2}{2} I_{[0,1)}^G(v)}\,.$$

The proof, in Section 6.1, needs several steps. Let $\Theta_n^G = n^{-1} \log \Delta_n^{G,\beta}$, so that

$$\dot{\Theta}_n^G = \frac{1}{n} \sum_{j=1}^n \Big(\log \rho_{n,j}^{G,\beta}\Big)\, \delta_{j/n}\,. \tag{3.37}$$

First we show that $\{\dot\Theta_n^G\}$ satisfies a LDP in M_T equipped with the topology $\sigma(\mathrm{M}_T, V_\ell)$. Then we carry the LDP to $(D_T, \sigma(D_T, \mathrm{M}_T))$ with good rate function :

$$I^G_{[0,T]}(v) = \int_{[0,T]} L_a^G\Big(t, \frac{d\dot v_a}{dt}(t)\Big)dt + \int_{[0,T]} L_s^G\Big(t, \frac{d\dot v_s}{d\mu}(t)\Big)d\mu(t)\,. \tag{3.38}$$

To end the proof we apply the Dawson-Gärtner theorem on projective limits (Dembo and Zeitouni (1998) Theorem 4.6.1, see also Léonard (2000) Proposition A2).

Note that $I^G_{[0,T]}(v)$ vanishes only when v satisfies (essentially)

$$\frac{d\dot v_a}{dt}(t) = \log(1-t) \quad , \quad \frac{d\dot v_s}{d\mu}(t) = 0\,, \tag{3.39}$$

i.e. for $v(t) = -\mathcal{J}(1-t)$, which is consistent with the result (3.12).

The LDP for marginals is given in the following theorem, where a rate function with affine part appears.

Theorem 3.10. *For every $T < 1$, the sequence $\left\{n^{-1}\log\Delta_n^{G,\beta}(T)\right\}_n$ satisfies a LDP in $\mathbb{R}$ in the scale $2\beta^{-1}n^{-2}$ with good rate function denoted by*

$$I_T^G(\xi) = \inf\{I^G_{[0,T]}(v)\ ;\ v(T) = \xi\}\,. \tag{3.40}$$

(1) *If $\xi \in [-T, 0)$ the equation*

$$\mathcal{J}(1+\theta) - \mathcal{J}(1-T+\theta) - T\log(1+\theta) = \xi\,, \tag{3.41}$$

has a unique solution, and we have

$$\begin{aligned} I_T^G(\xi) = \theta\xi \quad &+ \quad T\mathcal{J}(1+\theta) \\ &+ \quad (F(1) - F(1-T) - F(1+\theta) + F(1-T+\theta))\,. \end{aligned} \tag{3.42}$$

(2) *If $\xi < -T$, we have*

$$I_T^G(\xi) = I_T^G(-T) - (1-T)(\xi+T)\,. \tag{3.43}$$

(3) *If $\xi \geq 0$, $I_T^G(\xi) = \infty$.*

3.2.2. *Laguerre ensemble.* Let

$$\begin{aligned} L_a^L(t,y) &= (e^y - 1) - (1-t)y + \mathcal{J}(1-t) \\ L_s^L(t,y) &= -(1-t)y\ \delta(y|(-\infty, 0))\,. \end{aligned} \tag{3.44}$$

Theorem 3.11. *The sequence $\{n^{-1}\log\Delta_n^{L,\beta}(t), t \in [0,1)\}_n$ satisfies a LDP in $(D, \sigma(D, \mathrm{M}_<))$, in the scale $2\beta^{-1}n^{-2}$ with good rate function $I^L_{[0,1)}$.*

That means, roughly speaking, that

$$\mathbb{P}(\log\Delta_n^L \simeq nv) \approx e^{-\frac{\beta n^2}{2} I^L_{[0,1)}(v)}\,.$$

The proof uses the above result for the Uniform Gram process and the beta-gamma algebra. Note that $I^L_{[0,T]}(v)$ vanishes only when v satisfies (3.39) (again) i.e. for $v(t) = -\mathcal{J}(1-t)$, which is consistent with the result (3.21).

The LDP for marginals is given in the following theorem.

Theorem 3.12. *For every $T < 1$, the sequence $\{n^{-1} \log \Delta_n^{L,\beta}(T)\}_n$ satisfies a LDP in $\mathbb{R}$ in the scale $2\beta^{-1}n^{-2}$ with good rate function denoted by I_T^L.*

$$I_T^L(\xi) = \inf\{I_{[0,T]}^L(v) \ ; \ v(T) = \xi\}\,. \tag{3.45}$$

(1) *If $\xi \geq \xi_T := \mathcal{J}(T) - 1$ the equation*

$$\mathcal{J}(1+\theta) - \mathcal{J}(1-T+\theta) = \xi\,. \tag{3.46}$$

has a unique solution, and we have

$$I_T^L(\xi) = \theta\xi + F(1) - F(1-T) - F(1+\theta) + F(1-T+\theta)\,. \tag{3.47}$$

(2) *If $\xi < \xi_T$, we have*

$$I_T^L(\xi) = I_T^L(\xi_T) + (1-T)(\xi_T - \xi)\,. \tag{3.48}$$

3.2.3. *Jacobi ensemble.* Let, for $t < \tau_1$,

$$\begin{aligned} L_a^J(t,y) &= (\tau_1+\tau_2-t)\,\mathbf{H}\Big(\frac{\tau_1 - t}{\tau_1+\tau_2-t}\ \Big|\ e^y\Big) \\ L_s^J(t,y) &= -(\tau_1-t)y \text{ if } y<0\,. \end{aligned} \tag{3.49}$$

Theorem 3.13. *The sequence $\{n^{-1}\log \Delta_n^{J,\beta}(t),\ t \in [0,\tau_1)\}_n$ satisfies a LDP in $(D, \sigma(D, \mathrm{M}_<))$ in the scale $2\beta^{-1}n^{-2}$ with good rate function $I_{[0,\tau_1)}$.*

That means, roughly speaking, that

$$\mathbb{P}(\log \Delta_n^{J,\beta} \simeq nv) \approx e^{-\frac{\beta n^2}{2} I_{[0,1)}^J(v)}\,.$$

Note that $I_{[0,T]}^J(v)$ vanishes only when v satisfies (essentially)

$$\frac{d\dot v_a}{dt}(t) = \log\frac{\tau_1 - t}{\tau_1+\tau_2-t} \quad , \quad \frac{d\dot v_s}{d\mu}(t) = 0\,,$$

i.e. for $v(t) = \mathcal{E}(\tau_1,\tau_2,t)$, which is consistent with the result (3.32).

The LDP for marginals is given in the following theorem.

Theorem 3.14. *Let $T \in [0,\tau_1)$, and $\xi_T^J = \mathcal{J}(\tau_2) + \mathcal{J}(T) - \mathcal{J}(T+\tau_2) - 1$.*

(1) *The sequence $\{n^{-1}\log\Delta_n^{J,\beta}(T)\}_n$ satisfies a LDP in $\mathbb{R}$ in the scale $2\beta^{-1}n^{-2}$ with good rate function I_T^J where*

$$I_T^J(\xi) := \inf\{I_{[0,T]}^J(v) \ ; \ v(T) = \xi\}\,. \tag{3.50}$$

(2) *If $\xi \in [\xi_T^J, 0)$, the equation*

$$\mathcal{E}(\theta+\tau_1,\tau_2,T) = \xi \tag{3.51}$$

has a unique solution $\theta \geq T - \tau_1$, and we have

$$\begin{aligned} I_T^J(\xi) = \theta\xi \quad &- \quad [F(\theta+\tau_1) - F(\theta+\tau_1-T)] - [F(\tau_1+\tau_2) - F(\tau_1+\tau_2-t)] \\ &+ \quad [F(\tau_1) - F(\tau_1 - T)] \\ &+ \quad [F(\theta+\tau_1+\tau_2) - F(\theta+\tau_1+\tau_2-T)]\,. \end{aligned} \tag{3.52}$$

(3) *If $\xi < \xi_T^J$, we have*

$$I_T^J(\xi) = I_T^J(\xi_T^J) + (\xi_T^J - \xi)(\tau_1 - T)\,. \tag{3.53}$$

(4) *If $\xi \geq 0$, then $I_T^J(\xi) = \infty$.*

4. Connections with the spectral method

The logarithm of the determinant of a non singular matrix is a linear statistic of the empirical distribution of its eigenvalues, so that we may compare the above result with those obtained by this spectral approach.

4.1. *Laguerre/Wishart.* We start with

$$\frac{1}{n}\log \Delta_{n,r}^{L,\beta} = \frac{r}{n}\int (\log x)\ d\mu_{n,r}(x)$$

where

$$\mu_{n,r} = \frac{1}{r}\sum_{k=1}^{r} \delta_{\lambda_k} \tag{4.1}$$

is the so-called empirical spectral distribution (ESD). For $c > 0$ and $\sigma > 0$, let $\pi^c_{\sigma^2}$ be the distribution on $\mathbb{R}$ defined by

$$\pi^c_{\sigma^2}(dx) = \left(1 - c^{-1}\right)_+ \delta_0(dx) + \frac{\left((x-\sigma^2 a(c))(\sigma^2 b(c) - x)\right)_+^{1/2}}{2\pi\sigma^2 c x}\ dx\,, \tag{4.2}$$

where δ_0 is the Dirac mass in 0, $x_+ = \max(x, 0)$ and

$$a(c) = (1-\sqrt{c})^2\ ,\quad b(c) = (1+\sqrt{c})^2\,. \tag{4.3}$$

It is called the Marčenko-Pastur distribution with ratio index c and scale index σ^2 (Bai (1999) p.621).

It is well known (Marčenko and Pastur (1967), Bai (1999) Section 2.1.2 for the cases $\beta = 1$ and $\beta = 2$) that as $n, r \to \infty$ with $r/n \to T \in (0, \infty)$, the family of ESD $(\mu_{n,r})$ converges a.s. weakly to π_1^T. If we replace the common law $\mathcal{N}(0,1)$ by $\mathcal{N}(0,\sigma^2)$ then the limiting distribution is $\pi^T_{\sigma^2}$.

To conclude that

$$\lim_n \int (\log x)\ d\mu_{n,r}(x) = \int (\log x)\, d\pi_1^T(x)\,, \tag{4.4}$$

an additional control is necessary, since $x \mapsto \log x$ is not bounded.

Actually, the largest and the smallest eigenvalue converge a.s. to $b(T) < \infty$ and $a(T) > 0$, respectively. For comments on these results and references, one may consult Bai (1999) Sections 2.1.2 and 2.2.2., (see also Johnstone (2001)). In our context, this implies easily that a.s.

$$\frac{1}{n}\log \Delta_{n,r}^{L,\beta} = \frac{r}{n}\int (\log x)\ d\mu_{n,r}(x) \to T\int (\log x)\ d\pi_1^T(x) \tag{4.5}$$

Moreover, it is known (Jonsson (1982) p.31 and Bai and Silverstein (2004) p.596-597) that:

$$\begin{aligned} T\int (\log x)\ d\pi_1^T(x) &= \int_{a(T)}^{b(T)} \frac{\log x}{2\pi x}\ \sqrt{(x-a(T))(b(T)-x)}\ dx \\ &= (T-1)\log(1-T) - T = -\mathcal{J}(1-T) \end{aligned} \tag{4.6}$$

which implies that claim (4.5) is consistent with (3.21).

Recently, Bai and Silverstein (2004) proved a CLT for linear statistics of sample covariance matrices (non necessarily Gaussian), with the meaningful example of determinants. They consider the real and complex case, and their results (Theorem

1.1 ii) and iii) are consistent with the marginal version of (3.22). It is likely that $\beta = 4$ can also be handled under their assumptions.

Let us end with the large deviations. Hiai and Petz (1998), (see also Hiai and Petz (2000) Section 5.5) proved[1] that if $n \to \infty$ and $r/n \to T < 1$, then $\{\mu_{n,r}\}$ satisfies a LDP in $\mathcal{M}_1([0,\infty))$ - the set of probability measures on $[0,\infty)$ endowed with the weak topology - in the scale $2\beta^{-1}n^{-2}$ with some explicit good rate function I_T^{spL} given below in (4.8, 4.9, 4.10). If the contraction $\mu \mapsto \int(\log x)\ d\mu(x)$ were continuous, we would claim that $\{n^{-1}\log\Delta^L_{n,\lfloor nT\rfloor}\}_n$ satisfies a LDP in $\mathbb{R}$ in the same scale, with good rate function

$$\widetilde{I}_T^L(\xi) = \inf\left\{\widetilde{I}_T^{spL}(\mu)\ ;\ T\int(\log x)\ d\mu(x) = \xi\right\}. \tag{4.7}$$

Actually,

$$I_T^{spL}(\mu) = -T^2\Sigma(\mu) + T\int(x-(1-T)\log x)\,d\mu(x) + 2B(T) \tag{4.8}$$

where

$$\Sigma(\mu) := \iint \log|x-y|\ d\mu(x)d\mu(y) \tag{4.9}$$

is the so-called logarithmic entropy and for $T \in (0,1)$

$$2B(T) = -\frac{1}{2}\left(3T - T^2\log T + (1-T)^2\log(1-T)\right). \tag{4.10}$$

We do not know if the contraction $\mu \mapsto \int(\log x)\ d\mu(x)$ does work, although not continuous. However we will prove the following result, where for $u \in \mathbb{R}$ we put

$$\mathcal{A}(u) = \{\mu : \int(\log x)\ d\mu(x) = u\}. \tag{4.11}$$

Proposition 4.1. *For $\xi \geq \xi_T$ and θ solution of (3.46), let $\sigma^2 = 1+\theta$. Then the infimum of $I_T^{spL}(\mu)$ over $\mathcal{A}(\xi/T)$ is uniquely achieved for $\pi_{\sigma^2}^{T/\sigma^2}$ and*

$$I_T^L(\xi) = I_T^{spL}(\pi_{\sigma^2}^{T/\sigma^2}) = \inf\{I_T^{spL}(\mu);\ \mu \in \mathcal{A}(\xi/T)\}. \tag{4.12}$$

Remark 4.2.
(1) The endpoint is $\xi_T = \mathcal{J}(T) - 1$, with $\sigma^2 = T$.
(2) For $\xi < \xi_T$ we do not know what happens. We can imagine that the infimum in (4.12) has a solution in some extended space.

4.2. *Uniform Gram.* Let $\widetilde{\lambda}_k, k = 1,\dots,r$ be the eigenvalues of G in the Uniform Gram ensemble, and set

$$\widetilde{\mu}_{n,r} = \frac{1}{r}\sum_{k=1}^{r}\delta_{\widetilde{\lambda}_k}. \tag{4.13}$$

For $\beta = 1$, De Cock et al. (1999) proved that, as $n \to \infty$ and $r/n \to T \in (0,\infty)$, the family $(\widetilde{\mu}_{n,r})$ converges a.s. to π_1^T. More recently, Jiang (2004) proved that the same result holds true in a complex Gram ensemble not necessarily uniform. Again, like in Section 4.1, we may write

$$\frac{1}{n}\log\Delta_{n,r}^{G,\beta} = \frac{r}{n}\int(\log x)\ d\widetilde{\mu}_{n,r}(x)$$

[1] Their β is our β'.

and use the weak convergence of $\widetilde{\mu}_{n,r}$ towards π_1^T. Jiang (2004) proved that the largest and the smallest eigenvalue converge a.s. as $r/n \to T < 1$ to $b(T) < \infty$ and $a(T) > 0$ respectively. So, we have

$$\lim_n \int (\log x)\, d\widetilde{\mu}_{n,r}(x) = \int (\log x)\, d\pi_1^T(x)\,. \tag{4.14}$$

In view of (4.6), this coincides with the result (3.12).

No result on fluctuations or large deviations seems to be known on $\widetilde{\mu}_{n,r}$.

4.3. *Jacobi.* In the matrix models ($\beta = 1, 2$ or 4), take $r \le n_1$ and let $\lambda_k, k = 1, \dots, r$ be the eigenvalues of $\mathcal{Z}$. The ESD is

$$\nu_{n_1,n_2,r} = \frac{1}{r}\sum_{k=1}^r \delta_{\lambda_k}\,.$$

When $n_2 \le r \le n_1$ we have

$$\nu_{n_1,n_2,r} = \frac{n_2}{r}\mu_{n_1,n_2,r} + \left(1 - \frac{n_2}{r}\right)\delta_1\,,$$

where $\mu_{n_1,n_2,r}$ is the ESD built with eigenvalues different from 1. We can write in all cases

$$\log \Delta_{n,r}^{J,\beta} = \min(r, n_2) \int (\log x)\, \mu_{n_1,n_2,r}(dx)\,. \tag{4.15}$$

It is then possible to carry asymptotic results of this empirical distribution to $\log \Delta_{n,r}^{J,\beta}$.

Capitaine and Casalis (2004) studied the complex case in the regime $n_1/r \to u',\ n_2/r \to v'$ with $u' + v' \ge 1$. They prove[2] that $\mathbb{E}\nu_{n_1,n_2,r}$ converges (in moments hence) in distribution. To give the expression of the limiting distribution, which we denote $\mathrm{CC}_{u',v'}$ and to compare with known results in some other contexts with coherent notation, we will use in the following, four functions:
for $(b, c) \in (0, 1) \times (0, 1)$ we put

$$\sigma_\pm(b, c) = \frac{1}{2}\left[1 + \sqrt{bc} \pm \sqrt{(1-b)(1-c)}\right]\,, \tag{4.16}$$

and for $(x, y) \in (0, 1) \times (0, 1)$

$$\begin{aligned} a_\pm(x, y) &= (1 - x - y + 2xy) \pm 2\sqrt{x(1-x)y(1-y)} \\ &= \left(\sqrt{(1-x)(1-y)} \pm \sqrt{xy}\right)^2\,. \end{aligned} \tag{4.17}$$

The mappings $\sigma_\pm$ and $a_\pm$ are inverse in the following sense:

$$\{(b, c) : 0 < b < c < 1\} \underset{(a_-,a_+)}{\overset{(\sigma_-,\sigma_+)}{\rightleftarrows}} \{(x, y) : 0 < x < y < 1 \text{ and } x + y > 1\} \tag{4.18}$$

For $0 < a_- < a_+ < 1$, let π_{a_-,a_+} be the distribution on $\mathbb{R}$ defined by

$$\pi_{a_-,a_+}(dx) = C_{a_-,a_+}\frac{\sqrt{(x - a_-)(a_+ - x)}}{2\pi x(1-x)}\,\mathbf{1}_{[a_-,a_+]}(x)\ dx\,, \tag{4.19}$$

[2] They use the notation α and β but we change not to confuse with β already defined.

where C_{a_-,a_+} is the normalization constant. Since we found some mistakes in the literature, let us compute explicitly the constant C_{a_-,a_+}. From the obvious decomposition

$$\frac{1}{x(1-x)} = \frac{1}{x} + \frac{1}{1-x}$$

we get

$$(C_{a_-,a_+})^{-1} = I(a_-,a_+) + I(1-a_+,1-a_-)$$

where, for $0 < u < v$

$$I(u,v) = \int_u^v \frac{\sqrt{(x-u)(v-x)}}{2\pi x}\,dx$$

This last integral could be calculated by elementary method, but it is shorter to connect it with the Marčenko-Pastur distribution. Taking

$$\sigma^2 = \frac{\sqrt{v}+\sqrt{u}}{4} \quad , \quad \sqrt{c} = \frac{\sqrt{v}-\sqrt{u}}{\sqrt{v}+\sqrt{u}}$$

the simple fact that $\pi^c_{\sigma^2}$, given in (4.2), is a probability distribution yields

$$I(u,v) = \frac{\left(\sqrt{v}-\sqrt{u}\right)^2}{4} \quad (0 < u < v)\,.$$

Finally, we get:

$$(C_{a_-,a_+})^{-1} = \frac{1}{2}\left[1 - \sqrt{a_-a_+} - \sqrt{(1-a_-)(1-a_+)}\right]\,. \tag{4.20}$$

The distribution $\mathrm{CC}_{u',v'}$ is then (recall $u'+v' \geq 1$) :

$$\begin{aligned}\mathrm{CC}_{u',v'} := (1-u')^+\delta_0 \quad &+ \quad (1-v')^+\delta_1 \\ &+ \quad \left[1-(1-u')^+ - (1-v')^+\right]\pi_{a_-,a_+}\,,\end{aligned} \tag{4.21}$$

where

$$(a_-,a_+) = a_\pm\left(\frac{u'}{u'+v'}, 1-\frac{1}{u'+v'}\right)\,. \tag{4.22}$$

Remark 4.3. The case ($v' < 1$) corresponds to $r > n_2$, the second matrix W_2 is singular and the case ($v' \geq 1$) corresponds to $r \leq n_2$, the second matrix is non-singular.

For particular values of the parameters and up to an affine change to make the distribution symmetric, the distribution π_{a_-,a_+} was introduced by Kesten (1959) as limit distribution for random walks on some classical groups. It was (independently) introduced by McKay (1981) as a limit distribution in a graph problem. It is sometimes called the generalized McKay distribution. Some important connections are in Section 9.

For the LLN, the same remarks as above are relevant. Recall the notation

$$r \leq n_1\ ,\ n \to \infty\ ,\ \frac{r}{n} \to T\ ,\ \frac{n_1}{n} \to \tau_1\ ,\ \frac{n_2}{n} \to \tau_2\ ,\ u' = \frac{\tau_1}{T}\ ,\ v' = \frac{\tau_2}{T}\,.$$

The weak convergence of the ESD (Capitaine and Casalis (2004)) and the control on the extremal eigenvalues (Ledoux (2004), Collins (2005) and references therein), yield, if $u' \geq 1$

$$\lim_n \frac{1}{r}\log\Delta_n^{J,\beta}(T) = \int(\log x)\,\mathrm{CC}_{u',v'}(dx) = \min(v',1)\int(\log x)\,\pi_{a_-,a_+}(dx) \tag{4.23}$$

where $a_\pm$ are in (4.22). Nevertheless a computation of this integral by elementary methods is not so easy. After some attempts, we choose to consider the above result as an indirect way to compute this integral and we obtain the following result.

Proposition 4.4. *For* $0 < a_- < a_+ < 1$,

$$\int (\log x)\ \pi_{a_-,a_+}(dx) = \tag{4.24}$$
$$= \frac{\sigma_+ \log \sigma_+ + \sigma_- \log \sigma_- - (\sigma_+ + \sigma_- - 1)\log(\sigma_+ + \sigma_- - 1)}{1 - \sigma_+}$$

where $\sigma_\pm$ *are specified by (4.16).*

Proof: From (3.32),

$$\lim_n \frac{1}{r} \log \Delta_n^{J,\beta}(T) = \frac{1}{T} \lim_n \frac{1}{n} \log \Delta_n^{J,\beta}(T) = \frac{1}{T}\,\mathcal{E}(\tau_1, \tau_2, T) = \mathcal{E}(u', v', 1)\,,$$

where for the last equality we noticed that $\mathcal{E}$ is homogeneous. With the help of (4.23) we get

$$\min(v', 1) \int (\log x)\ \pi_{a_-,a_+}(dx) = \mathcal{E}(u', v', 1)\,. \tag{4.25}$$

From (4.18) we see that if $u' \geq 1$ then

$$\{\sigma_-, \sigma_+\} = \left\{ \frac{u'}{u' + v'}, \frac{u' + v' - 1}{u' + v'} \right\}$$

We have two cases. When $v' > 1$

$$\sigma_- = \frac{u'}{u' + v'} \quad , \quad \sigma_+ = \frac{u' + v' - 1}{u' + v'}\,,$$

so that (4.25) yields

$$\int (\log x)\ \pi_{a_-,a_+}(dx) = \mathcal{E}(u', v', 1) = \mathcal{E}\Big(\frac{\sigma_-}{1 - \sigma_+}, \frac{1 - \sigma_-}{1 - \sigma_+}, 1\Big) \tag{4.26}$$

When $v' < 1$

$$\sigma_+ = \frac{u'}{u' + v'} \quad , \quad \sigma_- = \frac{u' + v' - 1}{u' + v'}\,,$$

so that (4.25) yields

$$\int (\log x) \pi_{a_-,a_+}(dx) = \frac{1}{v'}\mathcal{E}(u', v', 1) = \frac{1 - \sigma_-}{1 - \sigma_+}\ \mathcal{E}\Big(\frac{\sigma_+}{1 - \sigma_-}, \frac{1 - \sigma_+}{1 - \sigma_-}, 1\Big) \tag{4.27}$$

and together (4.26-4.27) provide (4.24). This ends the proof. □

Let us end with the large deviations. In the complex case ($\beta = 2$), Hiai and Petz (2006) proved that if $n \to \infty, n_1/n \to \tau_1, n_2/n \to \tau_2 > \tau_1, r/n \to T < \tau_1$, then $\{\mu_{n_1,n_2,r}\}_n$ satisfies a LDP in $\mathcal{M}_1([0,1])$ - the set of probability measures on $[0,1]$ endowed with the weak topology - in the scale n^{-2}, with the good rate function

$$\begin{aligned} I_T^{spJ}(\mu) \quad &:= \quad -T^2\Sigma(\mu) - T \int_0^1 \left((\tau_1 - T)\log x + (\tau_2 - T)\log(1 - x)\right)\, d\mu(x) \\ &+ \quad T^2 B\Big(\frac{\tau_1 - T}{T}, \frac{\tau_2 - T}{T}\Big)\,, \end{aligned} \tag{4.28}$$

where B is defined in (3.3) (it is the limiting free energy). A computation similar to Hiai and Petz (2006) p.10 gives the same result for general β.

Proposition 4.5. *If $T < \tau_1 \leq \tau_2$, the family $\{\mu_{\lfloor n\tau_1\rfloor,\lfloor n\tau_2\rfloor,\lfloor nT\rfloor}\}$ satisfies a LDP in $\mathcal{M}_1([0,1])$ in the scale $2\beta^{-1}n^{-2}$ and good rate function I_T^{spJ}.*

If the contraction $\mu \mapsto \int(\log x)\, d\mu(x)$ from the set $\mathcal{M}_1([0,1])$ to $\mathbb{R}$ were continuous, we would claim that $\{n^{-1}\log\Delta_n^{J,\beta}(T)\}_n$ satisfies a LDP in $\mathbb{R}$ with good rate function $\widetilde{I}_T^J$ where

$$\widetilde{I}_T^J(\xi) = \inf\left\{I_T^{spJ}(\mu)\ ;\ \mu \in \mathcal{A}(\xi T^{-1})\right\} \tag{4.29}$$

with $\mathcal{A}(u)$ as defined in (4.11).

Like in the Laguerre case we will prove the following result.

Proposition 4.6. *Let $T < \min(\tau_1, \tau_2)$, $\xi \in [\xi_T^J, 0)$ and θ solution of (3.51). Then the infimum of $I_T^{spJ}(\mu)$ over $\mathcal{A}(\xi T^{-1})$ is uniquely achieved at $\mu = \pi_{\tilde{a}_-,\tilde{a}_+}$ where*

$$(\tilde{a}_-, \tilde{a}_+) = a_\pm(\tilde{s}_-, \tilde{s}_+)$$

with

$$\tilde{s}_- = \frac{\tau_1 + \theta}{\tau_1 + \tau_2 + \theta}\ ,\ \tilde{s}_+ = \frac{\tau_1 + \theta + \tau_2 - t}{\tau_1 + \tau_2 + \theta}\,, \tag{4.30}$$

and

$$I_T^J(\xi) = I_T^{spJ}(\pi_{\tilde{a}_-,\tilde{a}_+}) = \inf\{I_T^{spJ}(\mu);\ \mu \in \mathcal{A}(\xi T^{-1})\}\,. \tag{4.31}$$

Remark 4.7. The endpoint is ξ_T^J, which corresponds to $\theta = T - \tau_1$, i.e.

$$\tilde{a}_- = 0\ \ ,\ \ \tilde{a}_+ = \frac{4\tau_2 T}{(\tau_2 + T)^2}\,.$$

For $\xi < \xi_T^J$ we do not know what happens. We can imagine that the infimum in (4.29) has a solution in some extended space.

Remark 4.8. In the range $\tau_2 \leq T < \tau_1$ we have a similar result, exchanging $\tilde{s}_-$ and $\tilde{s}_+$ in (4.30). We omit the details.

4.4. *Extensions.* We already mentioned that in the Wishart and Gram models, limiting results exist for marginals when we leave the Gaussian/Uniform world, in particular for fluctuations in Bai and Silverstein (2004).

The Bartlett decomposition is not possible in the general case. Nevertheless, a product formula for the determinant is well known (see for example Lemma 3.1 p.9 and formula 4.3 p.15 in Friedland et al. (2004)), but nothing can be said about the distribution of the components of the product in general.

Nevertheless, if the columns (or the rows) of the matrix B are i.i.d. and isotropic, the previous results extend easily.

Begin with the "column" case. The beta-gamma algebra allowed us to pass from the Uniform Gram ensemble to the Wishart ensemble. The polar decomposition allows to obtain similar results as for the Wishart ensemble under convenient assumptions on the radial distribution. Let $\varepsilon_n = \log\|b_1\|^2 - \log\mathbb{E}\|b_1\|^2$ (remember that we omit the dimension index n). To get convergence and fluctuations it is enough to assume

$$n\mathbb{E}\varepsilon_n \to a_1\ ,\ n\mathrm{Var}\,\varepsilon_n \to a_2\ ,\ n\mathbb{E}(\varepsilon_n - \mathbb{E}\varepsilon_n)^4 \to 0\,. \tag{4.32}$$

To get large deviations, it would be sufficient to assume that, for some convenient functions φ, the quantity $n^{-2}\sum_{k=1}^n\ \log\mathbb{E}\exp\left(n\varphi(k/n)\varepsilon_n\right)$ has a limit. Akhavi

(2002) uses the uniform distribution in the unit *ball*, so that the distribution of $\|b_1\|^2$ is Beta$(n/2,1)$ and (4.32) is satisfied with $a_1=-2,\ a_2=0$. The contribution of the radial part is then roughly "deterministic" since $\mathbb{E}\|b_1\|^2$ is bounded.

In the "row" case, we can use the results of the "column" case since the eigenvalues of BB' are (except 0 with multiplicity $n-r$) the same as those of $B'B$.

5. Proofs of Theorems of Section 3.1

5.1. *Proof of Theorem 3.1.* We will use Mellin transforms and their first two derivatives at $\theta=0$. From the decomposition (2.20) we have

$$\log \mathbb{E}|\Delta_{n,r}^{G,\beta}|^{\beta'\theta} = \sum_{k=1}^{r} \Lambda_{n,k}^{G,\beta}(\theta) \tag{5.1}$$

with

$$\Lambda_{n,k}^{G,\beta}(\theta) := \log \mathbb{E}\left[\rho_{n,k}^{G,\beta}\right]^{\beta'\theta} \tag{5.2}$$

and from (2.11)

$$\Lambda_{n,k}^{G,\beta}(\theta) = \ell\big(\beta'(n-k+1+\theta)\big) - \ell\big(\beta'(n-k+1)\big) + \ell(\beta' n) - \ell\left(\beta'(n+\theta)\right) \tag{5.3}$$

where we set

$$\ell(x) = \log\Gamma(x)\,.$$

We will use in the sequel some expansions of ℓ and of its derivative, the digamma function

$$\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}\,;$$

all these properties are based on Binet's first formula and are given in Section 8.

Proof of 1) and 2) Differentiating once, we get

$$\mathbb{E}\log\Delta_{n,r}^{G,\beta} = \sum_{j=1}^{r}\left[\Psi\left(\beta'(n-j+1)\right) - \Psi\left(\beta' n\right)\right]\,,$$

and from formula (8.5),

$$\mathbb{E}\log\Delta_{n,r}^{G,\beta} = \log\frac{(n)_r}{n^r} + \frac{1}{\beta}\left(H_{n-r}-H_n\right) + \frac{r}{\beta n} - \delta_{n,r}^{1}\,. \tag{5.4}$$

in which

1) $(p)_r = p(p-1)\cdots(p-r+1)$ is the falling factorial
2) $H_0=0$ and $H_p = 1+\frac{1}{2}+\cdots+\frac{1}{p}$ are the harmonic numbers
3) the delta term is

$$\delta_{n,r}^{1} = \int_0^\infty s f(s)\sum_{k=1}^{r}\left[e^{-\beta'(n-k+1)s} - e^{-ns}\right]\,ds\,. \tag{5.5}$$

Using now formula (8.1) twice, we have for $r<n$

$$\begin{aligned}\log\frac{(n)_r}{n^r} &= -\Big(n-r+\frac{1}{2}\Big)\log\left(1-\frac{r}{n}\right) - r - \int_0^\infty f(s)[e^{-s(n-r)}-e^{-sn}]ds\\ &= -n\mathcal{J}\left(1-\frac{r}{n}\right) - \frac{1}{2}\log\left(1-\frac{r}{n}\right) - \int_0^\infty f(s)[e^{-s(n-r)}-e^{-sn}]ds\,.\end{aligned}$$

For $r = n$ the Stirling formula gives

$$\log \frac{(n)_n}{n^n} = -n + \frac{1}{2}\log(2\pi n) + o(1)\,.$$

The harmonic contribution in (5.4) is

$$H_{n-r} - H_n = \log\left(1 - \frac{r}{n}\right) + o(1)$$

as soon as $n - r \to \infty$. For $r = n$, we have $H_0 - H_n = -\log n - \gamma + o(1)$. Applying the dominated convergence theorem and (8.4), we see that the delta contribution satisfies:

$$\sup_{r\le n} \delta^1_{n,r} = \delta^1_{n,n} \to \int_0^\infty \frac{sf(s)}{e^{\beta' s} - 1}\, ds\,,$$

and $\lim_n \delta^1_{n,\lfloor nt\rfloor} = 0$ for $t < 1$. Gathering all these estimates, and applying again the dominated convergence theorem, we get (for $n - r \to \infty$)

$$\mathbb{E}\log\Delta^{G,\beta}_{n,r} = -n\mathcal{J}\left(1 - \frac{r}{n}\right) + \frac{r}{\beta n} + \left(\frac{1}{\beta} - \frac{1}{2}\right)\log\left(1 - \frac{r}{n}\right) + o(1)\,,$$

and for $r = n$

$$\mathbb{E}\log\Delta^{G,\beta}_{n,n} = -n - \left(\frac{1}{\beta} - \frac{1}{2}\right)\log n + K^1_\beta + o(1)\,.$$

Moreover, for the supremum, we have

$$\begin{aligned}
\sup_{r\le n}\left|\mathbb{E}\log\Delta^{G,\beta}_{n,r} - \log\frac{(n)_r}{n^r}\right| &= O(\log n)\\
\sup_{r\le n}\left|\log\frac{(n)_r}{n^r} + n\mathcal{J}\left(1 - \frac{r}{n}\right)\right| &= O(\log n)
\end{aligned}$$

so that (3.5), (3.6) and (3.7) are proved.

3) Taking logarithms in (5.1) and differentiating twice, we get

$$\operatorname{Var}\log\Delta^{G,\beta}_{n,r} = \sum_{j=1}^{r}\Psi'\left(\beta'(n-j+1)\right) - \Psi'(\beta' n)$$

and owing to (8.9) and (8.8)

$$\operatorname{Var}\log\Delta^{G,\beta}_{n,r} = \frac{1}{\beta'}(H_n - H_{n-r}) - \frac{r}{\beta' n} + \delta^2_{n,r}\,,$$

where

$$\begin{aligned}
|\delta^2_{n,r}| &\le \sum_{n-r+1}^{n}\frac{2}{\beta'^2 j^2}\,,\\
\delta^2_{n,n} &= \int_0^\infty s\left(sf(s) + \frac{1}{2}\right)\sum_{k=1}^{n}[e^{-\beta'(n-k+1)s} - e^{-\beta' ns}]\, ds\,.
\end{aligned}$$

On the one hand $\lim \delta^2_{n,\lfloor nt\rfloor} = 0$ for $t < 1$ and we get (3.9). On the other hand, applying the dominated convergence theorem and (8.4) we have

$$\lim_n \delta^2_{n,n} = \int_0^\infty \frac{s\left(sf(s) + \frac{1}{2}\right)}{e^{\beta' s} - 1}\, ds\,.$$

and we get (3.10).

To prove 4), let us note that since $\mathcal{J}$ is uniformly continuous on $[0,1]$ we have

$$\lim_n \sup_{t\in[0,1]} \left|\mathcal{J}\left(1-\frac{\lfloor nt\rfloor}{n}\right) - \mathcal{J}(1-t)\right| = 0\,,$$

so that, owing to (3.5), it is enough to prove that in probability

$$\sup_{1\leq p\leq n} \left|\log \Delta_{p,n}^{G,\beta} - \mathbb{E}\log \Delta_{p,n}^{G,\beta}\right| = o(n)\,.$$

Actually this convergence is a consequence of Doob's inequality and of the variance estimate Var $n^{-1}\Delta_{n,n}^{G,\beta} = O(n^{-2}\log n)$ coming from (3.10). □

5.2. *Proof of Theorem 3.2.* First note that, thanks to the estimations of expectations in (3.6) and (3.7), we can reduce the problem to the centered process and centered variable:

$$\delta_n(t) := \log \Delta_n^{G,\beta}(t) - \mathbb{E}\,\log \Delta_n^{G,\beta}(t) \quad , \quad \widehat{\delta}_n = \delta_n(1)/\sqrt{(2/\beta)\log n}\,.$$

1) We have $\delta_n(t) = \sum_{k=1}^{\lfloor nt\rfloor} \eta_{n,k}$ where

$$\eta_{n,k} := (\log \rho_{k,n}^{G,\beta}) - \mathbb{E}(\log \rho_{k,n}^{G,\beta}), \quad k\leq n \tag{5.6}$$

is a row-wise independent arrow. To prove (3.13) it is enough to prove the convergence in distribution in $\mathbb{D}([0,T])$, for every $T<1$, of δ_n to a centered Gaussian process with independent increments, and variance $\int_0^t \left(\sigma^{G,\beta}(s)\right)^2\,ds$. To this purpose we apply a version of the Lindeberg-Lévy-Lyapunov criterion (see Dacunha-Castelle and Duflo (1986) Volume II Theorem 7.4.28, or Jacod and Shiryaev (1987, Chap. 3 c)). For $t<1$, from (3.9) it is enough to prove that

$$\lim_n \sum_{k=1}^{\lfloor nt\rfloor} \mathbb{E}\,(\eta_{n,k}^4) = 0\,. \tag{5.7}$$

From definitions (5.6) and (5.2) we have:

$$\beta'^4\mathbb{E}(\eta_{n,k}^4) = (\Lambda_{n,k}^{G,\beta})^{(4)}(0) + 3[(\Lambda_{n,k}^{G,\beta})^{(2)}(0)]^2\,. \tag{5.8}$$

On the one hand, differentiating four times in (5.3) yields

$$(\Lambda_{n,k}^{G,\beta})^{(4)}(0) = \beta'^4[\Psi^{(3)}(\beta'(n-k+1)) - \Psi^{(3)}(\beta' n)]$$

and using Binet estimates (8.8), (8.9) for $q=4$:

$$\left|\sum_{k=1}^{p}(\Lambda_{n,k}^{G,\beta})^{(4)}(0) - 6\beta'\sum_{k=1}^{p}\left[\frac{1}{(n-k+1)^3} - \frac{1}{n^3}\right]\right| \leq 6\sum_{k=1}^{p}\frac{1}{(n-k+1)^4}\,. \tag{5.9}$$

Fixing $0<t<1$ and taking $p=\lfloor nt\rfloor$ we get $\lim_n \sum_{k=1}^{\lfloor nt\rfloor}\left(\Lambda_{n,k}^{G,\beta}\right)^{(4)}(0) = 0\,.$

On the other hand,

$$\sum_{k=1}^{p}[(\Lambda_{n,k}^{G,\beta})''(0)]^2 \leq \left(\sup_{j\leq p}(\Lambda_{n,j}^{G,\beta})''(0)\right)\sum_{k=1}^{p}(\Lambda_{n,k}^{G,\beta})''(0)\,. \tag{5.10}$$

We already know, from (3.9) that

$$\beta'^{-2}\sum_{k=1}^{\lfloor nt\rfloor}(\Lambda_{k,n}^{G,\beta})''(0) = \mathrm{Var}\log\Delta_n^{G,\beta}(t) \to \int_0^t \left(\sigma^{G,\beta}(s)\right)^2\,ds\,.$$

Now since $(\Lambda_{n,k}^{G,\beta})''(0) = \beta'^2[\Psi'(\beta'(n-k+1)) - \Psi'(\beta' n)]$ and since Ψ' is non-increasing (see (8.8)) we obtain

$$\sup_{j\leq \lfloor nt\rfloor} (\Lambda_{j,n}^{G,\beta})''(0) \leq \beta'^2 \Psi'\left(\beta'(n-\lfloor nt\rfloor + 1)\right)\,,$$

and from (8.9) (again), this term tends to 0. We just checked (5.7), which proves that the sequence of processes $\{\delta_n(t), t\in[0,1)\}_n$ converges to a Gaussian centered process $\mathcal{W}$ with independent increments and the convenient variance. It is now straightforward to get equation (3.14).

2) When $t=1$ most of the sums studied above explode when n tends to infinity and we need a renormalization. In fact, for every n, the process $(\delta_n(t), t\in[0,1])$ has independent increments. For $t_1 < \cdots < t_r$, the variable $\delta_n(1)$, conditionally upon $\delta_n(t_1) = \varepsilon_1, \dots, \delta_n(t_r) = \varepsilon_r$, has the same distribution as $\varepsilon_r + \sum_{[nt_r]+1}^n \eta_{k,n}$. Formulae (3.9) and (3.10) yield

$$\sum_{[nt_r]+1}^{n} \mathbb{E}(\eta_{k,n}^2) = (2/\beta)\log n + O(1)\,. \tag{5.11}$$

In order to apply the Lindeberg-Lyapunov theorem to the triangular array of random variables $\widehat{\eta}_{k,n} = \eta_{k,n}/\sqrt{(2/\beta)\log n}$ with with $k = [nt_r]+1, \dots, n$, we want to prove

$$\lim_n \sum_{k=1}^n \mathbb{E}(\widehat{\eta}_{k,n}^4) = 0\,. \tag{5.12}$$

We start again with the decomposition (5.8). From the above estimate (5.9), the sum $\sum_{k=1}^n (\Lambda_{n,k}^{G,\beta})^{(4)}(0)$ is bounded. In (5.10), we have

$$\sum_{k=1}^n (\Lambda_{n,k}^{G,\beta})''(0) = \beta'^{-2}\mathrm{Var}\ \log\Delta_n^{G,\beta}(1)$$

which is equivalent to $2\log n$ (see (3.10)) and the supremum in (5.10) with $p=n$ is bounded. This yields

$$\sum_{k=1}^n \mathbb{E}(\widehat{\eta}_{k,n}^4) = \beta'^2(\log n)^{-2}\sum_{k=1}^n \mathbb{E}(\eta_{k,n}^4) = O((\log n)^{-1})$$

which proves (5.12). Then the distribution of $\sum_{[nt_r]+1}^n \widehat{\eta}_{k,n}$ converges to $\mathcal{N}(0,1)$, and the same is true for the conditional distribution of $\widehat{\delta}_n$ knowing $\delta_n(t_1) = \varepsilon_1, \dots, \delta_n(t_r) = \varepsilon_r$. Since the limiting distribution does not depend on $\varepsilon_1, \dots, \varepsilon_r$, we have proved that $\widehat{\delta}_n$ converges in distribution to a random variable which is $\mathcal{N}(0,1)$ and independent of $\mathcal{W}$. $\square$

5.3. *Proof of Theorems 3.3 and 3.5.* It is of course possible to follow the same schemes of proof. Actually we prefer, at least for the beginning, exploit the beta-gamma algebra and the fundamental relation (2.26). So, for instance

$$\mathbb{E}\left[\varepsilon_k^{(n)}\right]^{\beta'\theta} = \left(\frac{1}{\beta' n}\right)^{\beta'\theta} \frac{\Gamma(\beta'(n+\theta))}{\Gamma(\beta' n)}$$

hence

$$\log \mathbb{E}\left[\varepsilon_k^{(n)}\right]^{\beta'\theta} = \ell\left(\beta'(\theta+n)\right) - \ell\left(\beta' n\right) - \beta'\theta \log\left(\beta' n\right)\ , \tag{5.13}$$

which provides estimates for the expectation and the variance. Differentiating once and taking $\theta = 0$, we see that

$$\mathbb{E}\log\varepsilon_k^{(n)} = \Psi\left(\beta' n\right) - \log\left(\beta' n\right) = -\frac{1}{\beta n} + O\Big(\frac{1}{n^2}\Big)$$

(see (8.7)), which gives

$$\sup_{p\le n}\left|\mathbb{E}S_{n,p} + \frac{p}{n\beta}\right| = O\Big(\frac{1}{n}\Big)\,. \tag{5.14}$$

Besides, differentiating (5.13) twice and taking $\theta = 0$ again, we have

$$\mathrm{Var}\Big(\log\varepsilon_k^{(n)}\Big) = \Psi'\left(\beta' n\right) = \frac{2}{\beta n} + O\Big(\frac{1}{n^2}\Big)$$

(see (8.9)), which yields

$$\sup_{p\le n}\left|\mathrm{Var}\, S_{n,p} - \frac{2p}{\beta n}\right| = O\Big(\frac{1}{n}\Big)\,. \tag{5.15}$$

From (5.14) and (5.15) it is easy to check (via a fourth moment estimate) that S_n converges in distribution in $\mathbb{D}([0,1])$ to

$$\Big(-(t/\beta) + \sqrt{2/\beta}\,\widetilde{\mathbf{B}}_t,\ t\in[0,1]\Big)$$

where $\widetilde{\mathbf{B}}$ is a Brownian motion independent of $(\Delta_n^{G,\beta}, n\in\mathbb{N})$. Finally the family of processes $\Delta_n^{L,\beta} = \Delta_n^{G,\beta} + S_n$ converges in distribution towards

$$\Big(X_t^{G,\beta} - (t/\beta) + \sqrt{2/\beta}\,\widetilde{\mathbf{B}}_t,\ t\in[0,1)\Big)\,.$$

It is a Gaussian process, whose drift and variance coefficients are

$$\mathtt{d}^{G,\beta}(t) - \frac{1}{\beta} = \left(\frac{1}{2} - \frac{1}{\beta}\right)\frac{1}{1-t} = \mathtt{d}^{L,\beta}(t)\ \ ,\ \ \Big(\sigma^{G,\beta}(t)\Big)^2 + \frac{2}{\beta} = \Big(\sigma^{L,\beta}(t)\Big)^2\,.$$

which identify the process $X^{L,\beta}$.

Besides, we have

$$\widehat{\eta}_n^{L,\beta}(1) = \widehat{\eta}_n^{G,\beta}(1) + \frac{S_n(1)}{\sqrt{2\log n}}\,,$$

so that the convergence of $\widehat{\eta}_n^{L,\beta}(1)$ is clear. Moreover the independence properties seen in Theorem 3.2 remain true. $\square$

5.4. *Proof of Theorem 3.6.* Again, we could follow the same schemes as in the Gram section. Actually we take again the benefit of beta-gamma algebra. We delete the superscript β for the sake of simplicity.

Recall the equality in law (2.28)

$$\log\Delta_{n_1,r}^{L} \overset{(d)}{=} \log\Delta_{n,r}^{J} + \log\Delta_{n_1+n_2,r}^{L} - r\log\frac{n_1}{n_1+n_2}$$

with independence in the left hand side.

We deduce easily

$$\mathbb{E}\log\Delta_{n,r}^{J} = \mathbb{E}\log\Delta_{n_1,r}^{L} - \mathbb{E}\log\Delta_{n_1+n_2,r}^{L} + r\log\frac{n_1}{n_1+n_2}$$

and

$$\mathrm{Var}\,\log\Delta^J_{n,r} = \mathrm{Var}\,\log\Delta^L_{n_1,r} - \mathrm{Var}\,\log\Delta^L_{n_1+n_2,r}$$

The results are now straightforward and we let the proof to the reader. We just note that since $r/n_1 \to t/\tau_1$ and $r/(n_1+n_2) \to t/(\tau_1+\tau_2)$ then

$$\begin{aligned}\mathbb{E}\log\Delta^L_{n_1,r} + n_1\mathcal{J}\left(1-\frac{r}{n_1}\right) \quad &\to \quad \int_0^{t/\tau_1} \mathtt{d}^{L,\beta}(s)ds \\ \mathbb{E}\log\Delta^L_{n_1+n_2,r} + (n_1+n_2)\mathcal{J}\left(1-\frac{r}{n_1+n_2}\right) \quad &\to \quad \int_0^{t/(\tau_1+\tau_2)} \mathtt{d}^{L,\beta}(s)ds\end{aligned}$$

hence

$$\mathtt{d}^{J,\beta}(t) = \frac{1}{\tau_1}\mathtt{d}^L\left(\frac{t}{\tau_1}\right) - \frac{1}{\tau_1+\tau_2}\mathtt{d}^L\left(\frac{t}{\tau_1+\tau_2}\right).$$

In the same vein

$$\left(\sigma^J(t)\right)^2 = \frac{1}{\tau_1}\left(\sigma^L\left(\frac{t}{\tau_1}\right)\right)^2 - \frac{1}{\tau_1+\tau_2}\left(\sigma^L\left(\frac{t}{\tau_1+\tau_2}\right)\right)^2.$$

5.5. *Proof of Theorem 3.8.* Again, it is possible to follow the classical scheme. Instead let us look at the situation we are faced to. Put

$$\begin{aligned}&U_n = \log\Delta^L_{n_1,r} - \mathbb{E}\Delta^L_{n_1,r}\ ,\ V_n = \log\Delta^L_{n_1+n_2,r} - \mathbb{E}\log\Delta^L_{n_1+n_2,r}\,, \\ &W_n = \log\Delta^J_{n,r} - \mathbb{E}\log\Delta^J_{n,r}\,.\end{aligned} \tag{5.16}$$

so that $U_n = V_n + W_n$ with $U_n \Rightarrow U$ and $V_n \Rightarrow V$, where U and V are Gaussian processes with independent increments, and V_n and W_n are independent. Looking for instance at characteristic functions, it is clear that W_n converges in the sense of finite distributions to a Gaussian process with independent increments. Its drift and variance are the difference of the corresponding ones. Moreover, since $\{U_n\}_n$ and $\{V_n\}_n$ are tight, $\{U_n - V_n\}_n$ is tight.

6. Proofs of Theorems of Section 3.2

6.1. *Proof of Theorem 3.9.* Recall the notation $\Theta^G_n = n^{-1}\log\Delta^{G,\beta}_n$. As mentioned after the statement of the theorem, we are going to prove at first the LDP for the restriction of $\dot\Theta^G_n$ to $[0,T]$, viewed as an element of M_T, in the scale $\beta'^{-1}n^{-2}$ with rate function

$$\widetilde{I}^G_{[0,T]}(m) := \int_0^T L^G_a\Big(t, \frac{dm_a}{dt}(t)\Big)dt + \int_0^T L^G_s\Big(t, \frac{dm_s}{d\mu}(t)\Big)d\mu(t)\,. \tag{6.1}$$

Recall that V_ℓ is the set of functions from $[0,T]$ to $\mathbb{R}$ which are left continuous and have bounded variation. Let V^*_ℓ be its topological dual when V_ℓ is equipped with the uniform convergence topology. Actually $\dot\Theta^G_n \in \mathrm{M}_T$ may be identified with an element of V^*_ℓ (see Léonard (2000) Appendix B): owing to (3.37) its action on $\varphi \in V_\ell$ is given by

$$\langle\dot\Theta^G_n, \varphi\rangle := \frac{1}{n}\sum_{k=1}^{\lfloor nT\rfloor}\varphi(k/n)\log\rho^{G,\beta}_{n,k}\,.$$

The proof of Theorem 3.9 is based on the ideas of Baldi theorem (Dembo and Zeitouni (1998) p.157). The main tool is the normalized cumulant generated function (n.c.g.f.) which here takes the form

$$\mathcal{L}^{G,\beta}_{n,\lfloor nT\rfloor}(\varphi) := \frac{1}{\beta' n^2}\log\mathbb{E}\left[\exp\left(\beta' n^2\langle\dot\Theta^G_n,\varphi\rangle\right)\right]. \tag{6.2}$$

Owing to (6.1) we have

$$\mathcal{L}^{G,\beta}_{n,\lfloor nT\rfloor}(\varphi) = \frac{1}{\beta' n^2}\sum_{k=1}^{\lfloor nT\rfloor}\Lambda^{G,\beta}_{n,k}\left(n\varphi(k/n)\right) \tag{6.3}$$

and from (5.2) it is finite iff $\varphi(k/n) > -(n-k+1)/n$ for every $1\le k\le\lfloor nT\rfloor$.

In Section 6.1.1, we prove the convergence of this sequence of n.c.g.f. for a large class of functions φ. It will be sufficient, jointly to the variational formula given in Section 6.1.2 to get the upper bound for compact sets. Then Section 6.1.3 is devoted to exponential tightness, which allows to get the upper bound for closed sets. However, since the limiting n.c.g.f. is not defined everywhere, the lower bound (for open sets) is more delicate than in Baldi theorem. Actually a careful study of exposed points as in Gamboa et al. (1999) is managed in Section 6.1.4. We end the proof in Section 6.1.5.

6.1.1. *Convergence of the n.c.g.f.* Let, for $t\in[0,1]$ and $\theta > -(1-t)$

$$g^G(t,\theta) := \mathcal{J}(1-t+\theta) - \mathcal{J}(1-t) - \mathcal{J}(1+\theta). \tag{6.4}$$

Lemma 6.1. *If $\varphi\in V_\ell$ satisfies $\varphi(t)+1-t>0$ for every $t\in(0,T]$, then*

$$\lim_n \mathcal{L}^{G,\beta}_{n,\lfloor nT\rfloor}(\varphi) = \Lambda^G_{[0,T]}(\varphi) := \int_0^T g^G(t,\varphi(t))\, dt. \tag{6.5}$$

Proof: The key point is a convergence of Riemann sums. From (5.3) and (8.1) we have, for every $\theta > -\frac{n-k+1}{n}$,

$$\begin{aligned}\Lambda^{G,\beta}_{n,k}(n\theta) = \beta'(n-k+n\theta)\log\left(1-\frac{k}{n}+\theta+\frac{1}{n}\right) - \beta'(n-k)\log\left(1-\frac{k}{n}+\frac{1}{n}\right)\\ -\beta'(n-1+n\theta)\log(1+\theta) + R_{n,k}(\theta)\end{aligned}$$

where the quantity

$$\begin{aligned}R_{n,k}(\theta) = &\int_0^\infty f(s)e^{-\beta' s}\left[e^{-\beta'(n-k+n\theta)s} - e^{-\beta'(n-k)s}\right] ds\\ &-\int_0^\infty f(s)\left[e^{-\beta'(n-1+n\theta)s} - e^{-\beta' s}e^{-\beta'(n-1)s}\right] ds\end{aligned}$$

is bounded by $2\int_0^\infty e^{-\beta' s}f(s)\,ds$. If we set

$$\begin{aligned}\Phi_n(t) \;:=\; &(1-t+\varphi(t))\log\left(1-t+\varphi(t)+\frac{1}{n}\right)\\ &-(1-t)\log(1-t+\frac{1}{n}) - \left(1-\frac{1}{n}+\varphi(t)\right)\log(1+\varphi(t))\end{aligned}$$

then, making $\theta = \varphi(k/n)$, and taking the sum over k, we get from (6.3)

$$\frac{1}{\beta' n^2}\Big(\mathcal{L}^G_{n,\lfloor nt\rfloor}(\varphi) - \sum_2^{\lfloor nt\rfloor} R_{n,k}(\varphi(k/n))\Big) = \frac{1}{n}\sum_1^{\lfloor nt\rfloor}\Phi_n\left(\frac{k}{n}\right) =$$
$$= \int_{1/n}^{\lfloor nt\rfloor/n}\Phi_n\left(\frac{\lfloor ns\rfloor}{n}\right)ds + \frac{1}{n}\Phi_n\left(\frac{\lfloor nt\rfloor}{n}\right).$$

On the one hand, since φ is left continuous, $\lim_n \Phi_n\left(\frac{\lfloor nt\rfloor}{n}\right) = g(t,\varphi(t))$ for every $t \in [0,T]$. On the other hand the following double inequality holds true:

$$\begin{aligned}
\Phi_n(t) &\geq (1-t+\varphi(t))\log(1-t+\varphi(t)) - (1-t)\log(2-t)\\
&- (1+\varphi(t))\log(1+\varphi(t)) - |\log(1-t+\varphi(t))|\\
\Phi_n(t) &\leq (1-t+\varphi(t))\log(2-t+\varphi(t)) - (1-t)\log(1-t)\\
&- (1+\varphi(t))\log(1+\varphi(t)) + |\log(1-t+\varphi(t))|,
\end{aligned}$$

and with our assumptions on φ, these bounds are both integrable. This allows to apply the dominated convergence theorem which ends the proof of Lemma 6.7. □

If there exists $s < T$ such that $\varphi(s) < -(1-s)$ then for n large enough, $\mathcal{L}_{n,\lfloor nT\rfloor}(\varphi) = +\infty$ and we set $\Lambda^G_{[0,T]}(\varphi) = \infty$. In the other cases we do not know what happens, but as in Gamboa et al. (1999), we will study the exposed points. Before, we need another expression of the dual of $\Lambda^G_{[0,T]}$.

6.1.2. *Variational formula.* Define $\Lambda^G_{[0,T]}(\varphi) = +\infty$ if φ does not satisfy the assumption of Lemma 6.1. The dual of $\Lambda^G_{[0,T]}$ is then

$$\left(\Lambda^G_{[0,T]}\right)^\star(\nu) = \sup_{\varphi\in V_\ell}\left\{\langle\nu,\varphi\rangle - \Lambda^G_{[0,T]}(\varphi)\right\} \tag{6.6}$$

for $\nu \in V_\ell^*$. Mimicking the method of Léonard (2000) p. 112-113, we get

$$\left(\Lambda^G_{[0,T]}\right)^\star(\nu) = \sup_{\varphi\in\mathcal{C}}\left\{\langle\nu,\varphi\rangle - \Lambda^G_{[0,T]}(\varphi)\right\} \tag{6.7}$$

where $\mathcal{C}$ is the set of continuous functions from $[0,T]$ into $\mathbb{R}$ vanishing at 0. Then we apply Theorem 5 of Rockafellar (1971) and get

$$\left(\Lambda^G_{[0,T]}\right)^\star(\nu) = \int_0^T g^\star\Big(t,\frac{d\nu_a}{dt}\Big)\,dt + \int_0^T r\Big(t,\frac{d\nu_s}{d\mu}(t)\Big)\,d\mu(t)$$

where

$$g^\star(t,y) \;=\; \sup_\lambda\left\{\lambda y - g^G(t,\lambda)\delta(\lambda|(-1,\infty))\right\}, \tag{6.8}$$

and r is the recession function:

$$r(t,y) = \lim_{\kappa\to\infty}\frac{g^\star(t,\kappa y)}{\kappa}.$$

Actually, if $y < 0$, the supremum is achieved for

$$\lambda^G(t,y) := -\Big(1 - \frac{t}{1-e^y}\Big) \tag{6.9}$$

and we have

$$\begin{aligned} g^\star(t,y) &= \lambda^G(t,y)y - g^G\left(t, \lambda^G(t,y)\right) \\ &= -y(1-t) + (1-t)\log(1-t) + t\log t - t\log(1-e^y) \\ &= \mathbf{H}\left(1-t|e^y\right) . \end{aligned} \tag{6.10}$$

If $y \geq 0$, then $g^\star(t,y) = \infty$. The recession is now $r(t,y) = -(1-t)y$ if $y \leq 0$, and $r(t,y) = \infty$ si $y > 0$. As a result

$$g^\star(t,y) = L_a^G(t,y) \quad , \quad r(t,y) = L_s^G(t,y) \, . \tag{6.11}$$

So we proved the identification $\left(\Lambda^G_{[0,T]}\right)^\star = \widetilde{I}^G_{[0,T]}$ (recall (6.1)).

6.1.3. *Exponential tightness.* In this paragraph and in Section 6.2 we use the function defined for $\theta > -(1-T)$ by

$$L_T^G(\theta) := \int_0^T g^G(t,\theta)\, dt \, . \tag{6.12}$$

If V_ℓ^* is equipped with the topology $\sigma(V_\ell^*, V_\ell)$, the set

$$B_a := \{\mu \in V_\ell^* : |\mu|_{[0,T]} \leq a\}$$

is compact according to the Banach-Alaoglu theorem. Now $-\dot\Theta_n^G$ is a positive measure and its total mass is $-\dot\Theta_n^G([0,T]) = -\Xi_n(T)$. We have then

$$\mathbb{P}\Big(\dot\Theta_n^G \notin B_a\Big) = \mathbb{P}\Big(\Theta_n^G(T) < -a\Big) \, .$$

Now for $\theta < 0$

$$\mathbb{P}\left(\Theta_n^G(T) < -a\right) \leq e^{\beta'\theta n^2 a}\, \mathbb{E}\exp\{n^2\beta'\theta\Theta_n^G(T)\}$$

so that, taking logarithm and applying Lemma 6.1 we get, for $\theta \in (-(1-T), 0)$

$$\limsup_n \frac{1}{\beta' n^2} \log \mathbb{P}\left(\Theta_n^G(T) < -a\right) \leq \theta a + L_T^G(\theta) \, .$$

It remains to let $a \to \infty$ and we have proved the exponential tightness.

Note that the restriction $T < 1$ is crucial in the above proof.

6.1.4. *Exposed points.* Let $\mathcal{R}$ be the set of functions from $[0,T]$ into $\mathbb{R}$ which are positive, continuous and with bounded variation. Let $\mathcal{F}$ be the set of those $m \in V_\ell^*$ (identified with M_T as in Léonard (2000)) which are absolutely continuous and whose density ρ is such that $-\rho \in \mathcal{R}$. Let us prove that such a m is exposed, with exposing hyperplane $f_m(t) = \lambda(t, \rho(t))$ (recall (6.9)). Actually we follow the method of Gamboa et al. (1999). For fixed t, $g^\star(t,.)$ is strictly convex on $(-\infty, 0)$ so that, if $z \neq \rho(t)$, we have

$$g^\star(t,\rho(t)) - g^\star(t,z) < \lambda(t,\rho(t))(\rho(t) - z) \, .$$

Let $d\xi = \tilde{l}(t)dt + \xi^\perp$ the Lebesgue decomposition of some element $\xi \in \mathrm{M}_T$ such that $\widetilde{I}^G_{[0,T]}(\xi) < \infty$. Taking $z = \tilde{l}(t)$ and integrating, we get

$$\int_0^T g^\star(t,\rho(t)dt - \int_0^T g^\star(t,\tilde{l}(t))dt < \int_0^T \lambda(t,\rho(t))\rho(t)dt - \int_0^T \lambda(t,\rho(t))\tilde{l}(t)\; dt$$

and since $\int_0^T g^\star(t,\tilde{l}(t))dt = \int_0^T L_a^G(t,\tilde{l}(t))dt \le \widetilde{I}^G_{[0,T]}(\xi)$ this yields

$$\widetilde{I}^G_{[0,T]}(m) - \widetilde{I}^G_{[0,T]}(\xi) < \int_0^T f_m dm - \int_0^T f_m d\xi\,.$$

The following lemma says that this set of exposed points is rich enough.

Lemma 6.2. *Let $m \in V_r$ such that $\widetilde{I}^G_{[0,T]}(m) < \infty$. There exists a sequence of functions $l_n \in \mathcal{R}$ such that*

(1) $\lim_n l_n(t)dt = -m$ *in V_ℓ^* with the $\sigma(V_\ell^*, V_\ell)$ topology,*
(2) $\lim_n \widetilde{I}^G_{[0,T]}(-l_n(t)dt) = \widetilde{I}^G_{[0,T]}(m)$.

Proof: The method may be found in Gamboa et al. (1999) and in Dette and Gamboa (2007). The only difference is in the topology because we want to recover marginals. We will use the basic inequality which holds for every $\epsilon \le 0$:

$$L_a^G(t, v+\epsilon) \le L_a^G(t,v) - \epsilon(1-t) \tag{6.13}$$

Let $m = m_a + m_s$ such that $\widetilde{I}^G_{[0,T]}(m) < \infty$. From (3.38) and (3.36) it is clear that $-m_a$ and $-m_s$ must be positive measures.

First step We assume that $m = -l(t)dt - \eta$ with $l \in L^1([0,T];dt)$ and η a singular positive measure. One can find a sequence of nonnegative continuous functions h_n such that $h_n(t)dt \to \eta$ for the topology $\sigma(V_\ell^*, V_\ell)$. Indeed every function $\psi \in V_\ell$ may be written as a difference $\psi_1 - \psi_2$ of two increasing functions. There exists a unique (positive) measure α_1 such that $\psi_1(t) = \alpha_1([t,T])$ for every $t \in [0,T]$. Moreover, the function $g = \eta([0,\cdot]) \in V_r$ is non decreasing and may be approached by a sequence of continuously derivable and non decreasing functions (g_n) such that $g_n \le g$. Setting $h_n := g_n'$ and $\nu_n = h_n(t)dt$, the dominated convergence theorem gives

$$\langle \psi_1, \nu_n\rangle = \int_0^T \nu_n([0,t])\alpha_1(dt) \to \int_0^T \eta([0,t])\alpha_1(dt)\,.$$

With the same result for ψ_2 we get

$$\begin{aligned}\langle \psi, \nu_n\rangle &= \int_0^T \nu_n([0,t])\alpha_1(dt) - \int_0^T \nu_n([0,t])\alpha_2(dt)\\ &\to \int_0^T \eta([0,t])\alpha_1(dt) - \int_0^T \eta([0,t])\alpha_2(dt)\,.\end{aligned}$$

or $\lim_n\langle \psi, \nu_n\rangle = \langle \psi, \eta\rangle$. On the one hand, the lower semi-continuity of $\widetilde{I}^G_{[0,T]}$ yields

$$\liminf_n \widetilde{I}^G_{[0,T]}\left(-(l(t)+h_n(t))dt\right) \ge \widetilde{I}^G_{[0,T]}(m)\,.$$

On the other hand, integrating (6.13) yields

$$\begin{aligned}\widetilde{I}^G_{[0,T]}(-(l(t)+h_n(t))dt) &\le \int_0^T L_a^G(t,-l(t))dt + \int_0^T (1-t)h_n(t)dt\\ &\to \int_0^T L_a^G(t,-l(t))dt + \int_0^T (1-t)\eta(dt) = \widetilde{I}^G_{[0,T]}(m)\,.\end{aligned}$$

Second step Assume that $m = -l(t)dt$ with $l \in L^1([0,T];dt)$ and for every n, let us set $l_n = \max(l, 1/n)$. It is clear that as $n \to \infty$, then $l_n \downarrow l$. On the one hand the lower semi-continuity gives

$$\liminf_A \widetilde{I}^G_{[0,T]}(-l_n(t)dt) \geq I^G_{[0,T]}(-l(t)dt)\,.$$

On the other hand, by integration of inequality (6.13), since $l_n - l \leq 1/n$

$$I^G_{[0,T]}(-l_n(t)dt) \leq I^G_{[0,T]}(-l(t)dt) + \frac{1}{n}\,.$$

It is then possible to reduce the problem to the case of functions bounded below.

Third step Assume that $m = -l(t)dt$ with $l \in L^1([0,T];dt)$ and bounded below by $A > 0$. One can find a sequence of continuous functions (h_n) with bounded variation such that $h_n \geq A/2$ for every n and such that $h_n \to l$ a.e. and in $L^1([0,T],dt)$. We have $h_n(t)dt \to l(t)dt$ in $\sigma(V_\ell^*, V_\ell)$ and since $L^G_a(t,\cdot)$ is uniformly Lipschitz on $(-\infty, -A/2]$, say with constant κ, we get

$$|\widetilde{I}^G_{[0,T]}(-h_n(t)dt) - \widetilde{I}^G_{[0,T]}(-l(t)dt)| \leq \kappa \int_0^T |h_n(t) - l(t)|dt \to 0\,.$$

Actually, $h_n \in \mathcal{R}$ and $\varphi_n(t) := \lambda(t, -h_n(t))$ satisfies the assumption of Lemma 6.1 since

$$1 + \varphi_n(t) - t \geq \frac{t}{1 - e^{-A/2}}\,.$$

6.1.5. *End of the proof of Theorem 3.9.* The first step is the upper bound for compact sets. We use Theorem 4.5.3 b) in Dembo and Zeitouni (1998) and the following lemma.

Lemma 6.3. *For every $\delta > 0$ and $m \in V_\ell^*$, there exists φ_δ fulfilling the conditions of Lemma 6.1 and such that*

$$\int_0^T \varphi_\delta dm - \Lambda^G_T(\varphi_\delta) \geq \min\left[I^G_{[0,T]}(m) - \delta,\ \delta^{-1}\right]. \tag{6.14}$$

The second step is the upper bound for closed sets: we use the exponential tightness. The third step is the lower bound for open sets. The method is classical (see Dembo and Zeitouni (1998) Theorem 4.5.20 c)), owing to Lemma 6.2.

To prove Lemma 6.3, we start from the definition (6.6) or (6.7). One can find $\bar{\varphi}_\delta \in V_\ell$ satisfying (6.14). If $\bar{\varphi}_\delta$ does not check assumptions of the lemma we add $\varepsilon > 0$ to $\bar{\varphi}_\delta$ which allows to check them and satisfy (6.14) up to a change of δ. $\square$

6.2. *Proof of Theorem 3.10.* We use the contraction from the LDP for paths. Since the mapping $m \mapsto m([0,T])$ is continuous from D to $\mathbb{R}$, the family $\Theta^G_n(T)$ satisfies the LDP with good rate function specified by (3.40):

$$I^G_T(\xi) = \inf\{I^G_{[0,T]}(v)\ ;\ v(T) = \xi\}\,.$$

Since the process Θ^G_n takes its values in $(-\infty, 0]$ (remember Hadamard inequality), it is clear that $I^G_T(\xi) = \infty$ for $\xi > 0$.

Fixing $\xi \leq 0$, we can look for optimal v. Let $\theta > -(1-T)$ (playing the role of a Lagrange multiplier). By the duality property (6.8)

$$g^\star\Big(t, \frac{d\dot{v}_a}{dt}(t)\Big) \geq \theta \frac{d\dot{v}_a}{dt}(t) - g^G(t,\theta)\,.$$

Integrating and using (6.1), (6.11) and (6.5) we get

$$I^G_{[0,T]}(v) \geq \theta \dot v_a([0,T]) - L^G_T(\theta) - \int_0^T (1-t)\ d\dot v_s(t)\,. \tag{6.15}$$

For every v such that $v(T)=\xi$ it turns out that

$$I^G_{[0,T]}(v) \geq \theta\xi - L^G_T(\theta) - \int_0^T (1-t+\theta)\ d\dot v_s(t) \geq \theta\xi - L^G_T(\theta)\,. \tag{6.16}$$

Besides, from Section 6.1.2 the Euler-Lagrange equation is

$$\begin{aligned} \lambda^G(t,\phi'(t)) &= \theta \\ \phi(0) &= 0\,. \end{aligned}$$

This ordinary differential equation admits for unique solution in $\mathcal{C}^1([0,T])$

$$t \mapsto \phi(\theta;t) := \mathcal{J}(1+\theta) - \mathcal{J}(1-t+\theta) - t\log(1+\theta)\,.$$

Now, since

$$\frac{\partial}{\partial\theta}\phi(\theta;T) = -\left[\log\left(1-\frac{T}{1+\theta}\right) + \frac{T}{1+\theta}\right] > 0$$

we see that the mapping $\theta \mapsto \phi(\theta;T)$ is bijective from $[-(1-T),\infty)$ onto $[-T,0)$. Moreover, by duality

$$g^\star\Big(t, \frac{\partial}{\partial t}\phi(\theta,t)\Big) = \theta\frac{\partial}{\partial t}\phi(\theta,t) - g^G(t,\theta)\,.$$

There are three cases.

• If $\xi \in [-T,0)$, there exists a unique θ_ξ such that $\phi(\theta_\xi,T)=\xi$ (i.e. the relation (3.41) is satisfied). For $v^\xi := \phi(\theta_\xi\,,\cdot)$, we get from (6.1), (6.11) and (6.12) again

$$I^G_{[0,T]}(v^\xi) = \theta_\xi\xi - L^G_T(\theta_\xi)$$

so that v^ξ realizes the infimum in (3.40). A simple computation ends the proof of the first statement of Theorem 3.10.

Note that at the end point $\xi=-T$, we have

$$\theta_\xi = -(1-T)\ ,\ v^\xi(t) = \mathcal{J}(T) - \mathcal{J}(T-t) - t\log T\ ,\ (v^\xi)'(t) = \log(1-t/T)\,.$$

Finally

$$\begin{aligned} I^G_T(-T) &= 2T(1-T) + \big(F(1) - F(1-T) - F(T) + T^2\log T\big) \\ &= \frac{T(1-T)}{2} + \frac{T^2\log T}{2} - \frac{(1-T)^2\log(1-T)}{2} + \frac{3}{4}\,. \end{aligned}$$

• If $\xi < -T$, set $\varepsilon = -T-\xi$. Plugging $\theta = -(1-T)$ in (6.16) yields, for every v such that $v(T)=\xi$

$$I^G_{[0,T]}(v) \geq -(1-T)\xi - L^G_T\left(-(1-T)\right) = (1-T)\varepsilon + I^G_T(-T)\,,$$

and this lower bound is achieved by the measure $\widetilde v = (v^{-T})'(t)dt - \varepsilon\delta_T(t)$, since

$$\int_0^T L^G_a\left(t,(v^{-T})'(t)\right)dt = I^G_T(T) \quad , \quad \int_0^T (1-t)\,\varepsilon\, d\delta_T(t) = (1-T)\,\varepsilon\,.$$

• If $\xi = 0$, make $\xi = 0$ in (6.16). We get $I_T^G(0) \geq -L_T^G(\theta)$ for every $\theta \geq -(1-T)$. Now, from (6.4) and (6.12) we may write

$$-L_T^G(\theta) = \int_0^T (1-t)\log(1-t)\ dt + \int_0^T (1+\theta)\log\Big(1 - \frac{t}{1+\theta}\Big)\,dt + \int_0^T t\log(1-t+\theta)\,dt\,.$$

When θ tends to infinity the second term tends to zero and the third, which is larger than $(T^2/2)\log(1-T+\theta)$, tends to infinity. Finally $I_T^G(0) = \infty$.

That ends the proof of the second statement of Theorem 3.10. $\square$

Remark 6.4. It is possible to try a direct method to get (3.42), (3.43) using Gärtner-Ellis' theorem (Dembo and Zeitouni (1998), Theorem 2.3.6). From Lemma 6.1 the limiting n.c.g.f. of $\Theta_n^G(T)$ is L_T^G which is analytic on $(-(1-T), \infty)$. When $\theta \downarrow -(1-T)$ we have $(L_T^G)'(\theta) \downarrow -T$. We meet a case of so-called non steepness. To proceed in that direction we could use the method of time dependent change of probability (see Dembo and Zeitouni (1995)). We will not give details here. Nevertheless, this approach allows to get one-sided large deviations in the critical case $T = 1$. Actually we get

$$\lim_n \frac{1}{\beta' n^2} \log \mathbb{P}(\Theta_n^G(1) \geq \xi) = -I_1^G(\xi)$$

for $\xi \geq -1$. The function I_1^G is obtained in the same way as in (3.41, 3.42). The value $\xi = -1$ corresponds to the limit of $\Theta_n^G(1)$. Note that the second (right) derivative of I_1^G at this point is zero (or equivalently $\lim (L_1^G)''(\theta) = \infty$ as $\theta \downarrow 0$), which is consistent with previous results on the variance. I do not know the rate of convergence to 0 of $\mathbb{P}(\Theta_n^G(1) \leq \xi)$ for $\xi < -1$.

6.3. *Proof of Theorem 3.11 and Theorem 3.12.* Again, the three routes are possible to tackle the problem of large deviations for determinant of Wishart matrices. A direct method would use the cumulant generating function from (5.13) and would meet computations similar to those seen in the Uniform Gram case. To avoid repetitions, we use the decomposition (2.26), drawing benefit from an auxiliary study of $S_{n,r}$.

Lemma 6.5. *The sequence $\{n^{-1}\, S_n(t), t \in [0,1)\}_n$ satisfies a LDP in the space $(D, \sigma(D, \mathrm{M}_<))$ in the scale $2\beta^{-1}n^{-2}$ with good rate function*

$$I_{[0,1)}^S(v) = \int_{[0,1)} L_a^S\Big(\frac{d\dot v_a}{dt}(t)\Big)dt + \int_{[0,1)} L_s^S\Big(\frac{d\dot v_s}{d\mu}(t)\Big)d\mu(t) \tag{6.17}$$

where

$$L_a^S(y) = (e^y - y - 1)\ \ ,\ L_s^S(y) = -y\delta(y|(-\infty,0))\,, \tag{6.18}$$

and μ is any measure dominating $d\dot v_s$.

We stress that the instantaneous rate functions are time homogeneous and then we may write $[0,1]$ instead of $[0,1)$.

6.3.1. *Proof of Lemma 6.5.* It is a route similar to the proof of Theorem 3.9 in Section 6.1 (see also Najim (2002)). We start from (2.27) so that

$$\frac{1}{n}\dot{S}_n = \sum_{j=1}^{n} \left(\log \varepsilon_k^{(n)}\right) \delta_{j/n}\,.$$

With the help of (5.13) this yields:

$$\log \mathbb{E} \exp\langle \beta' n \dot{S}_n, \gamma\rangle = \sum_{k=1}^{n} \left[\ell\left(\beta' n\Big(1+\gamma\Big(\frac{k}{n}\Big)\Big)\right) - \beta'\gamma\Big(\frac{k}{n}\Big)\log(\beta' n)\right] \\ -n\ell(\beta' n)$$

if $\gamma(s)+1>0$ for every $s\in[0,1]$. A little computation shows that the limiting n.c.g.f. is

$$\mathcal{L}^S(\gamma) = \int_0^1 \mathcal{J}(1+\gamma(t))dt\,, \tag{6.19}$$

which yields (6.18) by duality (see Rockafellar (1971) again). □

6.3.2. *Proof of Theorem 3.11.* Let $\Theta_n^L = n^{-1}\log \Delta_n^{L,\beta}$. We deduce from Lemma 6.5 and Theorem 3.9 that the sum $\dot{\Theta}_n^L = \dot{\Theta}_n^G + \frac{1}{n}\dot{S}_n$ satisfies a LDP in the same scale with good rate function $I^G_{[0,T]}\Box I^S_{[0,T]}$ where $\Box$ denotes the infimum convolution:

$$(f\Box g)(x) = \inf\{f(x_1)+g(x_2)\mid x_1+x_2=x\}\,.$$

The two characteristics of the rate function are then

$$\begin{aligned} L_a^L &= \inf_v\{L_a^G(v)+L_a^S(u-v)\} \\ L_s^L &= \inf_v\{L_s^G(v)+L_s^S(u-v)\}\,. \end{aligned}$$

which yield (3.44) by an explicit computation. □

Alternatively, it is possible to sum the two n.c.g.f. ((6.5) and (6.19)) and get the rate function by duality. We claim: if $\gamma(s)+1>0$ for every $s\in[0,1]$

$$\frac{1}{\beta' n^2}\log \mathbb{E}\exp\langle \beta' n^2 \dot{\Theta}_n^L, \gamma\rangle \to \int_0^T g^L(t,\gamma(t))\,dt\,, \tag{6.20}$$

where

$$g^L(t,\gamma) = g^G(t,\gamma) + \mathcal{J}(1+\gamma) = \mathcal{J}(1-t+\gamma) - \mathcal{J}(1-t)\,. \tag{6.21}$$

6.3.3. *Proof of Theorem 3.12.* We may either use the contraction $\Theta_n^L \mapsto \Theta_n^L(T)$ or establish a LDP for the marginal $S_n(T)$ and then perform an inf-convolution. We leave the details of the proof to the reader. We just give the expression of the optimal path when it exists.

For $\theta > -(1-T)$, the function

$$t \mapsto \phi(\theta;t) := \mathcal{J}(1+\theta) - \mathcal{J}(1-t+\theta)\,.$$

is in $\mathcal{C}^1([0,T])$ and the mapping $\theta\mapsto\phi(\theta;T)$ is bijective from $[-(1-T),\infty)$ onto $[\xi_T,\infty)$, where $\xi_T = \mathcal{J}(T)-1$. Fixing $\xi\ge\xi_T$, we can look for optimal v. There exists a unique θ_ξ such that $\phi(\theta_\xi,T)=\xi$. Then $v^\xi := \phi(\theta_\xi\,,\cdot)$ is the optimal path (v^ξ realizes the infimum in (3.45). Note that at the end point $\xi=\xi_T$, we have

$$\theta_\xi = -(1-T)\ ,\ v^\xi(t) = \mathcal{J}(T)-\mathcal{J}(T-t)\ ,\ (v^\xi)'(t) = \log(T-t)\,. \ \Box$$

Remark 6.6. It is possible to get (3.47), (3.48) using Gärtner-Ellis' theorem (Dembo and Zeitouni (1998), Theorem 2.3.6). We are in the same situation as in Remark 6.4. This approach allows to get one-sided large deviations in the critical case $T = 1$. Actually we get

$$\lim_n \frac{2}{\beta n^2} \log \mathbb{P}(\Theta_n^L(1) \geq \xi) = -I_1^L(\xi)$$

for $\xi \geq -1$. The value $\xi = -1$ corresponds to the limit of . Note that the second (right) derivative of I_1^L at this point is zero (or equivalently $\lim (L_1^L)''(\theta) = \infty$ as $\theta \downarrow 0$), which is consistent with previous results on the variance. I do not know the rate of convergence to 0 of $\mathbb{P}(\Theta_n^L(1) \leq \xi)$ for $\xi < -1$.

6.4. *Proof of Theorem 3.13 and Theorem 3.14.* We may try again to use the beta-gamma algebra, but we do not succeed to go until the end. Let, as in Section 5.5, U_n and V_n be the two Laguerre variables. From the exponential tightness of U_n and V_n, we deduce easily the exponential tightness of W_n. From Puhalskii (2001), the sequence W_n contains subsequences satisfying LDP. If for such a subsequence we call I^p the rate function, the independence gives

$$I^U = I^V \square I^p$$

This equation has many solutions and only one convex solution, which is

$$I^p = I^U \boxminus I^V$$

defined by

$$(f \boxminus g)(x) = \sup\{f(x_1) - g(x_2) \mid x_1 - x_2 = x\}$$

(Mazure and Volle (1991)). But we do not know a priori that I^p is convex.

The alternate (and classical) route needs the study of the n.c.g.f. For that point we use the beta-gamma trick. For the remaining we do not give details since it is similar to the above cases and again based on the ideas of Baldi theorem (Dembo and Zeitouni (1998)) and a variational formula.

6.4.1. *Convergence of the n.c.g.f.* Put $\Theta_n^J = n^{-1} \log \Delta_n^{J,\beta}$ so that

$$\dot{\Theta}_n^J = \frac{1}{n} \sum_{k=1}^{n_1} \left(\log \rho_{j,n}^{J,\beta} \right) \delta_{j/n} \,,$$

and put for $T \leq \tau_1$ and $\varphi \in V_\ell^T$:

$$\mathcal{L}_{n,\lfloor nT \rfloor}^J(\varphi) = \frac{2}{\beta n^2} \log \mathbb{E} \exp n \langle \dot{\Theta}_n^J, \varphi \rangle \,.$$

Lemma 6.7. *If $\varphi \in V_\ell^{\tau_1}$ satisfies $\varphi(s) + \tau_1 - s > 0$ for every $s \in (0, T]$, then*

$$\lim_n \mathcal{L}_{n,\lfloor nT \rfloor}^J(\varphi) = \Lambda_{[0,T]}^J(\varphi) := \int_0^T g^J(s, \varphi(s)) \, ds \,, \tag{6.22}$$

where, for $\theta + \tau_1 - s > 0$

$$g^J(s, \theta) = \mathcal{E}\left(\tau_1 - s + \theta, \tau_2, \theta\right) \,. \tag{6.23}$$

Proof: From (2.28) we have

$$\langle n\dot\Theta^J_n,\gamma\rangle+\langle(n_1+n_2)\dot\Theta^L_{n_1+n_2},\gamma((n_1+n_2)\cdot/n)\rangle=$$
$$\langle n_1\dot\Theta^L_{n_1},\gamma(n_1\cdot/n)\rangle+\log\frac{n_1}{n_1+n_2}\sum_{k=1}^{\lfloor nT\rfloor}\gamma(k/n)$$

and then, by independence,

$$\begin{aligned}\log\mathbb{E}\exp\langle\beta' n^2\dot\Theta^J_n,\gamma\rangle &= \log\mathbb{E}\exp\langle\beta' nn_1\dot\Theta^L_{n_1},\gamma(n_1\cdot/n)\rangle\\ &\quad-\log\mathbb{E}\exp\langle\beta' n(n_1+n_2)\dot\Theta^L_{n_1+n_2},\gamma((n_1+n_2)\cdot/n)\rangle\\ &\quad+n\log\frac{n_1}{n_1+n_2}\Big(\sum_{k=1}^{\lfloor nT\rfloor}\gamma(k/n)\Big).\end{aligned}$$

By a slight modification of (6.20) we have, for $p/n\to\tau$

$$\frac{1}{\beta' p^2}\log\mathbb{E}\exp\langle\beta' np\dot\Theta^L_r,\gamma(p\cdot/n)\rangle\to\frac{1}{\tau}\int_0^T g^L\Big(\frac{s}{\tau},\frac{\gamma(s)}{\tau}\Big)\,ds\,, \tag{6.24}$$

so that taking $\tau=\tau_1$ and $\tau=\tau_1+\tau_2$ and subtracting, we get

$$\frac{1}{\beta' n^2}\log\mathbb{E}\exp\langle\beta' n^2\dot\Theta^J_n,\varphi\rangle\to\int_0^T g^J\left(s,\gamma(s)\right)\,,$$

where

$$\begin{aligned}g^J\left(s,\theta\right)&=\tau_1 g^L\Big(\frac{s}{\tau_1},\frac{\theta}{\tau_1}\Big)-(\tau_1+\tau_2)g^L\Big(\frac{s}{\tau_1+\tau_2},\frac{\theta}{(\tau_1+\tau_2)}\Big)\\&\quad+\theta\,\log\frac{\tau_1}{\tau_1+\tau_2}\,,\end{aligned}$$

and this is equivalent to (6.23). □

6.4.2. *Duality.* Define $\Lambda^J_{[0,T]}(\varphi)=+\infty$ if φ does not satisfy the assumption of Lemma 6.7. The dual of $\Lambda^J_{[0,T]}$ is then

$$\left(\Lambda^J_{[0,T]}\right)^\star(\nu)=\sup_{\varphi\in V_\ell}\left\{<\nu,\varphi>-\Lambda^J_{[0,T]}(\varphi)\right\} \tag{6.25}$$

for $\nu\in V_\ell^*$. Mimicking the method of Léonard (2000) p. 112-113, we get

$$\left(\Lambda^J_{[0,T]}\right)^\star(\nu)=\sup_{\varphi\in\mathcal{C}}\left\{<\nu,\varphi>-\Lambda^J_{[0,T]}(\varphi)\right\} \tag{6.26}$$

where $\mathcal{C}$ is the set of continuous functions from $[0,T]$ into $\mathbb{R}$ vanishing at 0. Then we apply Theorem 5 of Rockafellar (1971). We get

$$\left(\Lambda^J_{[0,T]}\right)^\star(\nu)=\int_0^T\left(g^J\right)^\star\Big(t,\frac{d\nu_a}{dt}\Big)\,dt+\int_0^T r^J\Big(t,\frac{d\nu_s}{d\mu}(t)\Big)\,d\mu(t) \tag{6.27}$$

where

$$\left(g^J\right)^\star(s,y)=\sup_\lambda\left\{\lambda y-g^J(s,\lambda)\delta(\lambda|(-\tau_1,\infty))\right\}\,. \tag{6.28}$$

This supremum is achieved by

$$\lambda^J(s,y)=-(\tau_1-s)+\frac{\tau_2}{e^{-y}-1} \tag{6.29}$$

and we have

$$\left(g^J\right)^\star(s,y) = \lambda^J(s,y)y - g^J(s,\lambda^J(s,y)) \tag{6.30}$$

$$= (\tau_1+\tau_2-s)\,\mathbf{H}\Big(\frac{\tau_1-s}{\tau_1+\tau_2-s}\ \Big|\ e^y\Big)\,. \tag{6.31}$$

The recession is $r^J(s;y) = -(\tau_1 - s)y$ if $y < 0$.

6.4.3. *Proof of Theorem 3.10.* We use the contraction from the LDP for paths. Since the mapping $m \mapsto m([0,T])$ is continuous from D to $\mathbb{R}$, the family $\{\Theta_n^J(T)\}_n$ satisfies the LDP with good rate function given by (3.50). Since the process Θ_n takes its values in $(-\infty, 0]$, it is clear that $I_T^J(\xi) = \infty$ for $\xi > 0$. Fixing $\xi < 0$, we can look for optimal v, i.e. a path $(v(t), t \in [0,T])$ such that $v(T) = \xi$ and v achieves the infimum in (3.50). Fix $\theta \geq T - \tau_1$ (playing the role of a Lagrange multiplier). In view of (6.27), (6.28) and (6.29), it is clear that (in the generic case) the Euler-Lagrange equation is

$$\begin{aligned}\lambda^J(s,\phi'(s)) &= \theta\\ \phi(0) &= 0\,.\end{aligned}$$

This ordinary differential equation admits for unique solution in $\mathcal{C}^1([0,T])$

$$s \mapsto \phi^J(\theta;s) := \mathcal{E}(\theta+\tau_1,\tau_2,s)\,.$$

To know if the path ϕ^J may have ξ as its terminal value, look at

$$\begin{aligned}\mathcal{E}'(\theta+\tau_1,\tau_2,T) &= \frac{\partial}{\partial\tau}\mathcal{E}(\tau,\tau_2,T)|_{\tau=\theta+\tau_1}\\ &= \log\Big(1-\frac{T}{\theta+\tau_1+\tau_2}\Big) - \log\Big(1-\frac{T}{\theta+\tau_1}\Big)\,;\end{aligned}$$

since it is positive, we see that the mapping

$$\theta \longmapsto \mathcal{E}(\theta+\tau_1,\tau_2,T)$$

is continuous and increasing from $[T-\tau_1,\infty)$ onto $\mathcal{D}_T = [\xi_T^J,\ 0)$. If $\xi \in [\xi_T^J, 0)$, we call θ_ξ the unique solution of $\phi^J(\theta,T) = \xi$ or in other words,

$$\mathcal{E}(\theta_\xi+\tau_1,\tau_2,T) = \xi\,,$$

and we set $v^\xi := \phi^J(\theta_\xi\,,\cdot)$.

To end the proof, let us now consider some inequalities. The duality property (6.28) gives, for every v and t

$$\left(g^J\right)^\star\Big(t,\frac{d\dot v_a}{dt}(t)\Big) \geq \theta\frac{d\dot v_a}{dt}(t) - g^J(t,\theta)\,. \tag{6.32}$$

Setting

$$L_T^J(\theta) := \int_0^T g^J(t,\theta)\ dt\,, \tag{6.33}$$

integrating (6.32) and using (3.35), (3.49) and (6.33) we get

$$I_{[0,T]}^J(v) \geq \theta\dot v_a([0,T]) - L_T^J(\theta) - \int_0^T (\tau_1 - t)\ d\dot v_s(t)\,.$$

For every v such that $v(T) = \xi$ it turns out that

$$I^J_{[0,T]}(v) \geq \theta\xi - L^J_T(\theta) - \int_0^T (\tau_1 - T + \theta)\ d\dot{v}_s(t) \geq \theta\xi - L^J_T(\theta)\,. \tag{6.34}$$

There are three cases.

• If $\xi \in [\xi^J_T, 0)$, we get

$$I^J_{[0,T]}(v^\xi) \quad = \quad \theta_\xi \xi - L^J_T(\theta_\xi)$$

so that v^ξ realizes the infimum in (3.50). A simple computation leads to (3.42) which ends the proof of the first statement of Theorem 3.10.

Note that at the end point $\xi = \xi^J_T$, we have

$$\theta_\xi = (T - \tau_1)\ ,\ v^\xi(t) = \mathcal{E}(T, \tau_2, t)\ ,\ (v^\xi)'(t) = \log\frac{T - t}{\tau_1 + \tau_2 - t}\,.$$

• If $\xi < \xi^J_T$ set $\varepsilon = \xi^J_T - \xi$. Plugging $\theta = T - \tau_1$ in (6.34) yields, for every v such that $v(T) = \xi$

$$I^J_{[0,T]}(v) \geq (T - \tau_1)\xi^J_T - L^J_T(T - \tau_1) - \varepsilon(T - \tau_1) = I^J_T(\xi^J_T) - \varepsilon(T - \tau_1)\,,$$

and this lower bound is achieved by the measure $\tilde{v} = \left(v^{\xi^J_T}\right)'(t)dt - \varepsilon d\delta_T(t)$, since

$$\int_0^T L^J_a\Big(t, \left(v^{\xi^J_T}\right)'(t)\Big)dt = I^J_T(\xi^J_T) \quad , \quad \int_0^T (\tau_1 - t)\varepsilon d\delta_T(t) = (\tau_1 - T)\varepsilon\,.$$

• If $\xi = 0$, make $\xi = 0$ in (6.34). We get $I^J_{[0,T]}(0) \geq -L^J_T(\theta)$ for every $\theta \geq T - \tau_1$. Now, from (6.23) and (6.33), we may write (after some calculation)

$$\begin{aligned} -L^J_T(\theta) = \int_0^T -\mathcal{E}(\tau_1 - t + \theta, \tau_2, \theta)dt \quad &= \quad \int_0^T \int_0^\theta \log\Big(1 + \frac{\tau_2}{\tau_1 - t + s}\Big)ds \\ &\geq \quad T\int_0^\theta \log\Big(1 + \frac{\tau_2}{\tau_1 + s}\Big)\ ds\,, \end{aligned}$$

which tends to infinity as $\theta \to \infty$. We conclude $I^J_{[0,T]}(0) = \infty$. $\square$

7. Proofs of Propositions 4.1 and 4.6

7.1. *Proof of Proposition 4.1.* Let $\theta \in \mathbb{R}$ (Lagrangian multiplier). We begin with a minimization of

$$I^{spL}_T(\mu) - \theta T \int (\log x)\ d\mu(x) = T^2\left[-\Sigma(\mu) + 2\int q_{\lambda,s}(x)\ d\mu(x)\right] + 2B(T)$$

where

$$q_{\lambda,s}(x) = \lambda x - s\log x, \quad \lambda = \frac{1}{2T},\ \ s = \frac{1 - T + \theta}{2T}\,. \tag{7.1}$$

In the book of Saff and Totik (1997) p.43 example 5.4, it is stated that for $\lambda > 0$ and $2s + 1 > 0$ fixed, the infimum

$$\inf_\mu\ -\Sigma(\mu) + 2\int q_{\lambda,s}(x)\ d\mu(x)$$

is achieved by the unique extremal measure $\pi^c_{\sigma^2}$ with

$$\sigma^2 = \frac{2s + 1}{2\lambda}\ ,\ c = \frac{1}{2s + 1}\,.$$

We see from (7.1) that if $\theta > -1$ we can take:

$$\sigma^2 = 1 + \theta, \quad c = \frac{T}{\sigma^2} = \frac{T}{1+\theta}\,.$$

Now it remains to look for θ such that the constraint $\mu \in \mathcal{A}(\xi/T)$ is saturated. Since

$$\int (\log x)\ d\pi^c_{\sigma^2}(x) = \log \sigma^2 + \int (\log x)\ d\pi^c_1(x) dx\,,$$

and thanks to (4.6), we see that θ must satisfy

$$\xi = T \log \sigma^2 - T \frac{\mathcal{J}(1-c)}{c} = \mathcal{J}(1+\theta) - \mathcal{J}(1-T+\theta)\,,$$

which is exactly (3.46).

To compute $I_T^{spL}(\pi^c_{\sigma^2})$, we start from the definition (4.8):

$$I_T^{spL}(\pi^c_{\sigma^2}) = -T^2 \Sigma(\pi^c_{\sigma^2}) + T \int (x - (1-T)\log x)\ d\pi^c_{\sigma^2}(x) + 2B(T)\,,$$

and transform $\pi^c_{\sigma^2}$ to π^c_1 using the dilatation. In particular, (4.9) yields

$$\Sigma(\pi^c_{\sigma^2}) = \log \sigma^2 + \Sigma(\pi^c_1)$$

and $\Sigma(\pi^c_1)$ may be picked from formula (13) p.10 in Hiai and Petz (1998):

$$\Sigma(\pi^c_1) = -1 + \frac{1}{2}\left(c^{-1} + \log c\ + (c^{-1} - 1)^2 \log(1-c)\right)\,.$$

Besides we have easily $\int x\ d\pi^c_1(x) = 1$. After some tedious but elementary computations we get exactly the RHS of (3.47), which yields

$$I_T^{spL}(\pi^c_{\sigma^2}) = I_T^L(\xi)\,,$$

and ends the proof of (4.12). $\square$

7.2. *Proof of Proposition 4.6.* Let $\theta < T - \tau_1$ (Lagrangian multiplier). We begin with a minimization of

$$\begin{aligned} I_T^{spJ}(\mu) \quad &- \quad \theta T \int (\log x)\ d\mu(x) \qquad\qquad (7.2)\\ &= \quad T^2 \left[-\Sigma(\mu) - 2\zeta_1 \int (\log x)\, d\mu(x) - 2\zeta_2 \int \log(1-x)\, d\mu(x)\right] + C \end{aligned}$$

where

$$2\zeta_1 = \frac{\tau_1 + \theta - T}{T}\ ,\ 2\zeta_2 = \frac{\tau_2 - T}{T}\ \text{ and } C = T^2 B\Big(\frac{\tau_1 - T}{T}, \frac{\tau_2 - T}{T}\Big)\,.$$

We use the following lemma.

Lemma 7.1. *For $\zeta_1, \zeta_2 > 0$, the infimum of*

$$-\Sigma(\mu) - 2\zeta_1 \int (\log x)\, d\mu(x) - 2\zeta_2 \int \log(1-x)\, d\mu(x)$$

among the probability measures μ on $[0,1]$ is achieved by π_{a_-,a_+} where

$$(a_-, a_+) = \lambda_\pm(s_-, s_+)$$

with

$$s_- = \frac{1 + 2\zeta_1}{2(1 + \zeta_1 + \zeta_2)}\ ,\ s_+ = \frac{1 + 2\zeta_1 + 2\zeta_2}{2(1 + \zeta_1 + \zeta_2)}\,.$$

The infimum in (7.2) is achieved by $\pi_{\tilde\xi,\tilde\eta}$, where

$$(\tilde\xi,\tilde\eta) = \lambda_\pm(\tilde s_-,\tilde s_+)\,,\quad \tilde s_- = \frac{\tau_1+\theta}{\tau_1+\tau_2+\theta}\ ,\ \tilde s_+ = \frac{\tau_1+\theta+\tau_2-t}{\tau_1+\tau_2+\theta}\,.$$

It should be clear that

$$\Sigma(\pi_{\tilde\xi,\tilde\eta}) + \frac{\tau_1+\theta-T}{T}\int(\log x)\ d\pi_{\tilde\xi,\tilde\eta}(x) + \frac{\tau_2-T}{T}\int\log(1-x)\ d\pi_{\tilde\xi,\tilde\eta}(x) = B\Big(\frac{\tau_1+\theta-T}{T},\frac{\tau_2-T}{T}\Big)$$

and then, on $\mathcal{A}(\xi T^{-1})$ the infimum is uniquely realized in $\pi_{\tilde\xi,\tilde\eta}$ and its value is

$$\theta\xi + T^2\left[B\Big(\frac{\tau_1-T}{T},\frac{\tau_2-T}{T}\Big) - B\Big(\frac{\tau_1+\theta-T}{T},\frac{\tau_2-T}{T}\Big)\right]\,.$$

Finally a small computation leads to (3.52) and (4.31).

Proof of Lemma 7.1. In Saff and Totik (1997) p.241, it is proved that the infimum of

$$\int\!\!\int -\log|x-y|d\mu(x)d\mu(y) - 2\zeta_1\int\log(1-x)\,d\mu(x) - 2\zeta_2\int\log(1+x)\,d\mu(x)$$

among the probability measures μ on $[-1,+1]$ is achieved by

$$d\mu(y) = K(b_-,b_+)\frac{\sqrt{(y-b_-)(b_+-y)}}{2\pi(1-y^2)}\,\mathbf{1}_{[b_-,b_+]}(y)\ dy\,,$$

where $b_\pm = \theta_2^2-\theta_1^2\pm\sqrt{\Delta}$ with

$$\theta_i = \frac{\zeta_i}{1+\zeta_1+\zeta_2}\ ,\ i=1,2\quad ,\quad \Delta = \left[1-(\theta_1+\theta_2)^2\right]\left[1-(\theta_1-\theta_2)^2\right]\ ,$$

and $K(b_-,b_+)$ is a normalizing constant. With the push forward under the mapping $x\to(x+1)/2$, we get the result. □

8. Appendix 1 : Some properties of $\ell = \log\Gamma$ and Ψ

From Binet's formula (Abramowitz and Stegun (1972) pp. 258-259 or Erdélyi et al. (1981) p.21), we have

$$\ell(x) = \Big(x-\frac12\Big)\log x - x + 1 + \int_0^\infty f(s)[e^{-sx}-e^{-s}]\,ds \tag{8.1}$$

$$= \Big(x-\frac12\Big)\log x - x + \frac12\log(2\pi) + \int_0^\infty f(s)e^{-sx}\,ds\,. \tag{8.2}$$

where the function f is defined by

$$f(s) = \left[\frac12-\frac1s+\frac{1}{e^s-1}\right]\frac1s = 2\sum_{k=1}^\infty\frac{1}{s^2+4\pi^2k^2}\,, \tag{8.3}$$

and satisfies for every $s\ge 0$

$$0 < f(s)\le f(0) = 1/12\ ,\quad 0 < \Big(sf(s)+\frac12\Big) < 1\,. \tag{8.4}$$

By differentiation

$$-\Psi(x) = \frac{1}{2x} + \int_0^\infty sf(s)e^{-sx}\,ds = \int_0^\infty e^{-sx}\Big(sf(s)+\frac12\Big)ds\,. \tag{8.5}$$

As easy consequences, we have, for every $x > 0$

$$0 \quad < \quad x\Big(\log x - \Psi(x)\Big) \leq 1\,, \tag{8.6}$$

$$0 \quad < \quad \log x - \Psi(x) - \frac{1}{2x} \leq \frac{1}{12x^2}\,. \tag{8.7}$$

Differentiating again we see that for $q \geq 1$

$$\Psi^{(q)}(z) = (-1)^{q-1} q! z^{-q} + (-1)^{q-1} \int_0^\infty e^{-sz} s^q \Big(sf(s) + \frac{1}{2}\Big)\, ds \tag{8.8}$$

and then

$$|\Psi^{(q)}(z) - (-1)^{q-1} q! z^{-q}| \leq z^{-q-1} q!\,. \tag{8.9}$$

9. Appendix 2 : Identification of the McKay distribution

The reader is recalled that, for u' and v' positive numbers[3] such that $u' + v' > 1$, Capitaine and Casalis (2004) defined the probability measure

$$\mathrm{CC}_{u',v'} := (1-u')^+ \delta_0 + (1-v')^+ \delta_1 + \left[1 - (1-u')^+ - (1-v')^+\right] \pi_{a_-,a_+}\,,$$

where

$$(a_-, a_+) = a_\pm\Big(\frac{u'}{u'+v'}, 1 - \frac{1}{u'+v'}\Big)\,.$$

We present now three identifications of this distribution connected with free probability.

For $k \neq 0$, let D_k the dilatation operator by factor k. For $p \leq 1$, let $\mathtt{b}_p$ denote the Bernoulli distribution of parameter p. At last, let $\boxplus$ (resp. $\boxtimes$) denote the additive (resp. multiplicative) free convolution.

1) Rewriting the distribution with the notation of Demni (2006), we find four cases

- Situation I : $\min(u', v') \geq 1$, no Dirac mass,

$$\sigma_- = \frac{u'}{u'+v'}\ ,\ \sigma_+ = 1 - \frac{1}{u'+v'}\ ,\ u' = \frac{\sigma_-}{1-\sigma_+}\ ,\ v' = \frac{1-\sigma_-}{1-\sigma_+}$$

$$CC_{u',v'} = \pi_{a_-,a_+} = D_{1-\sigma_+}(\mathtt{b}_{\sigma_-})^{\boxplus \frac{1}{1-\sigma_+}}$$

- Situation II : $u' < 1 \leq v'$, one Dirac mass at 0

$$\sigma_- = \frac{1}{u'+v'}\ ,\ \sigma_+ = 1 - \frac{u'}{u'+v'}\ ,\ u' = \frac{1-\sigma_+}{\sigma_-}\ ,\ v' = \frac{\sigma_+}{\sigma_-}$$

$$\begin{aligned} CC_{u',v'} &= (1-u')\delta_0 + u'\pi_{a_-,a_+} \\ &= D_{\sigma_-}(\mathtt{b}_{1-\sigma_+})^{\boxplus \frac{1}{\sigma_-}} \end{aligned}$$

- Situation III : $v' < 1 \leq u'$, one Dirac mass at 1

$$\sigma_- = 1 - \frac{1}{u'+v'}\ ,\ \sigma_+ = \frac{u'}{u'+v'}\ ,\ u' = \frac{\sigma_+}{1-\sigma_-}\ ,\ v' = \frac{1-\sigma_+}{1-\sigma_-}$$

$$\begin{aligned} CC_{u',v'} &= (1-v')\delta_1 + v'\pi_{a_-,a_+} \\ &= D_{1-\sigma_-}(\mathtt{b}_{\sigma_+})^{\boxplus \frac{1}{1-\sigma_-}} \end{aligned}$$

[3] we use the symbol v' (hence u') not to confuse with β already defined.

• Situation IV : $\max(u', v') < 1$, two Dirac masses (at 0 and at 1)

$$\sigma_- = 1 - \frac{u'}{u'+v'}\ ,\ \sigma_+ = \frac{1}{u'+v'}\ ,\ u' = \frac{1-\sigma_-}{\sigma_+}\ ,\ v' = \frac{\sigma_-}{\sigma_+}$$

$$\begin{aligned} CC_{u',v'} &= (1-u')\delta_0 + (1-v')\delta_1 + (u'+v'-1)\pi_{a_-,a_+} \\ &= D_{\sigma_+}(\mathtt{b}_{1-\sigma_-})^{\boxplus\frac{1}{\sigma_+}}\ . \end{aligned}$$

2) There is a connection with the family of free Meixner law (Bozejko and Bryc (2005), Bryc and Ismail (2006), Bryc and Ismail (2005)). Indeed, computing the mean m and variance V of the distribution $CC_{u',v'}$, we get

Situation	m	V
I	σ_-	$\sigma_-(1-\sigma_-)(1-\sigma_+)$
II	$1-\sigma_+$	$\sigma_-\sigma_+(1-\sigma_+)$
III	σ_+	$(1-\sigma_-)\sigma_+(1-\sigma_+)$
IV	$1-\sigma_-$	$\sigma_-\sigma_+(1-\sigma_-)$

so that, in all cases

$$m = \frac{u'}{u'+v'}\ ,\ V = \frac{u'v'}{(u'+v')^3}\ .$$

We see that fixing $u'+v' = s^{-1}$, we get $V = s^2 m(1-m)$, and then up to an affine transformation we find the "free binomial type law" as in Bryc and Ismail (2005) example vi p.18 or Bozejko and Bryc (2005) example 6 p.8. It could also be seen starting from the above formulae using dilatations and free convolutions and comparing with formula (7) page 6 in Bozejko and Bryc (2005).

3) Finally, we quote the correspondence with the results of Collins (2005) who claimed that for $0 < p_- < p_+ < 1$

$$\mathtt{b}_{p_-} \boxtimes \mathtt{b}_{p_+} = (1-p_-)\delta_0 + (p_- + p_+ - 1)^+\delta_1 + C^{-1}_{a_-,a_+}\pi_{a_-,a_+}$$

where $a_\pm = a_\pm(1-p_-, p_+)$. In Hiai and Petz (2006), formula (2.8) the authors consider the same distribution.

• Situation II : $p_- + p_+ - 1 < 0$, $\sigma_- = p_+$, $\sigma_+ = 1 - p_-$

$$\begin{aligned} \mathtt{b}_{p_-} \boxtimes \mathtt{b}_{p_+} &= \sigma_+\delta_0 + (1-\sigma_+)\pi_{a_-,a_+} \\ &= \sigma_- CC_{u',v'} + (1-\sigma_-)\delta_0 \end{aligned}$$

• Situation IV : $p_- + p_+ - 1 > 0$, $\sigma_- = 1 - p_-$, $\sigma_+ = p_+$

$$\begin{aligned} \mathtt{b}_{p_-} \boxtimes \mathtt{b}_{p_+} &= \sigma_-\delta_0 + (\sigma_+ - \sigma_-)\delta_1 + (1-\sigma_+)\pi_{a_-,a_+} \\ &= \sigma_+ CC_{u',v'} + (1-\sigma_+)\delta_0\ . \quad \square \end{aligned}$$

Acknowledgement

I thank Catherine Donati-Martin for providing useful references. I am also grateful to the anonymous referee for a careful reading of the manuscript.

References

J. Abbott and T. Mulders. How tight is Hadamard bound? *Experiment. Math.* **10** (3), 331–336 (2001).

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, 9th edition (1972).

A. Akhavi. Random lattices, threshold phenomena and efficient reduction algorithms. *Theoretical Computer Science* **257**, 359–385 (2002).

T. W. Anderson. *An introduction to multivariate statistical analysis.* Wiley Series in Probability and Statistics. John Wiley, 3rd edition (2003).

S. A. Andersson, H. K. Brons and S. T. Jensen. Distribution of Eigenvalues in Multivariate Statistical Analysis. *The Annals of Statistics* **11** (2), 392–415 (1983).

Z. D. Bai. Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica* pages 611–677 (1999).

Z. D. Bai and J. W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probab.* **32** (1A), 553–605 (2004).

M. S. Bartlett. On the theory of statistical regression. *Proc. R. Soc. Edinb.* **53**, 260–283 (1933).

M. Bozejko and W. Bryc. On a class of free Lévy laws related to a regression problem (2005). Available in arXiv:math.OA/0410601.

W. Bryc and M. Ismail. Approximation operators, exponential, q-exponential and free exponential families (2005). Available in arXiv:math.ST/0512224.

W. Bryc and M. Ismail. Cauchy-Stieltjès kernel families (2006). Available in arXiv:math.PR/0601273.

M. Capitaine and M. Casalis. Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices. *Indiana Univ. Math. J.* **53** (2), 397–432 (2004).

L. Chaumont and M. Yor. Exercises in probability. In *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge (2003).

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

D. Dacunha-Castelle and M. Duflo. *Probability and Statistics.* Springer-Verlag, New York (1986).

M. De Cock, M. Fannes and P. Spincemaille. On quantum dynamics and statistics of vectors. *J. Phys. A: Math. Gen.* **32**, 6547–6571 (1999).

A. Dembo and O. Zeitouni. Large deviations via parameter dependent change of measure, and an application to the lower tail of Gaussian processes. In *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993)*, volume 36 of *Progr. Probab.*, pages 111–121. Birkhäuser, Basel (1995).

A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer, 2nd edition (1998).

N. Demni. Free Jacobi processes (2006). Available in arXiv:math.PR/0606218.

H. Dette and F. Gamboa. Asymptotic properties of the algebraic moment range process. *Acta Mathematica Hungarica* **116**, 247–264 (2007).

J. A. Diaz-Garcia and R. Gutierrez Jaimez. Proof of the conjectures of H. Uhlig on the singular multivariate Beta and the Jacobian of a certain matrix transformation. *Ann. Statist.* **25**, 2018–2023 (1997).

J.D. Dixon. How good is Hadamard's inequality for determinants? *Can. Math. Bull.* **27** (3), 260–264 (1984).

Y. Doumerc. *Matrices aléatoires, processus stochastiques et groupes de réflexion.* Ph.D. thesis, Université de Toulouse (2005).

I. Dumitriu and A. Edelman. Matrix models for beta-ensembles. *J. Math. Phys.* **43**, 5830–5847 (2003).

A. Edelman and B. D. Sutton. The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. *Foundations of Computational Mathematics* (2007).

A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi. *Higher transcendental functions*, volume I. Robert E. Krieger Publishing Co. Inc., Melbourne, Fla. (1981).

P. J. Forrester. Log-gases and random matrices (2007). Book available at http://www.ms.unimelb.edu.au/∼matpjf/matpjf.html.

S. Friedland, B. Rider and O. Zeitouni. Concentration of permanent estimators for certain large matrices. *The Annals of Applied Probab.* **14** (3), 1559–1576 (2004).

F. Gamboa, A. Rouault and M. Zani. A functional large deviation principle for quadratic forms of Gaussian stationary processes. *Stat. and Probab. Letters* **43**, 299–308 (1999).

V. L. Girko. *Theory of random determinants.* Kluwer Academic Publishers, Dordrecht (1988). Transl. from the Russian.

N. R. Goodman. Statistical analysis based on a certain multivariate complex gaussian distribution (An introduction). *Ann. Math. Stat.* **34**, 152–177 (1963).

A. K. Gupta and D. K. Nagar. *Matrix variate distributions.* Chapman & Hall (2000).

J. Hadamard. Résolution d'une question relative aux déterminants. *Bull. Sci. Math.* **17**, 240–246 (1893).

P. J. Hanlon, R. P. Stanley and J. R. Stembridge. Some combinatorical aspects of the spectra of normally distributed random matrices. *Contemporary Mathematics* **138**, 151–174 (1992).

F. Hiai and D. Petz. Eigenvalue density of the Wishart matrix and large deviations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **1** (4), 633–646 (1998).

F. Hiai and D. Petz. The semicircle law, free random variables and entropy. In *Mathematical Surveys and Monographs*, volume 77. Amer. Math. Soc., Providence (2000).

F. Hiai and D. Petz. Large deviations for functions of two random projection matrices. *Acta Sci. Math. (Szeged)* **72**, 581–609 (2006).

J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Springer (1987).

T. Jiang. The limiting distributions of eigenvalues of sample correlation matrices. *Sankhya* **66** (1), 35–48 (2004).

I. Johnstone. On the distribution of the largest eigenvalue in principal component analysis. *The Annals of Statistics* **29** (2), 295–327 (2001).

D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12**, 1–38 (1982).

D. G. Kabe. Classical statistical analysis based on a certain hypercomplex multivariate normal distribution. *Metrika* **31** (1), 63–76 (1984).

H. Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92**, 336–354 (1959).

R. Killip and I. Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.* **50**, 2665–2701 (2004).

A. M. Kshirsagar. *Multivariate Analysis.* Marcel Dekker (1972).

M. Ledoux. Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The continuous case. *Electron. J. Probab.* **9** (7), 177–208 (2004).

C. Léonard. Large deviations for Poisson random measures and processes with independent increments. *Stoch. Proc. and their Appl.* **85**, 93–121 (2000).

V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues of some sets of random matrices. *Math. USSR Sb.* **1**, 457–483 (1967).

A. M. Mathai. *Jacobians of matrix transformations and functions of matrix argument.* World scientific, Singapore (1997).

A. M. Mathai. Random p-content of a p-parallelotope in Euclidean n-space. *Adv. Appl. Prob.* **31**, 343–354 (1999).

M.-L. Mazure and M. Volle. Equations inf-convolutives et conjugaison de Moreau-Fenchel. *Ann. Fac. Sci. Toulouse, V. Sér., Math.* **12**, 103–126 (1991).

B. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.* **40**, 203–216 (1981).

M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition (2004).

R. J. Muirhead. *Aspects of multivariate statistical theory.* John Wiley (1982).

J. Najim. A Cramér type theorem for weighted random variables. *Electronic Journal of Probability* **7** (4), 1–32 (2002).

J. Nielsen. The distribution of volume reductions. *Adv. Appl. Prob.* **31**, 985–994 (1999).

I. Olkin and H. Rubin. Multivariate Beta distribution and independence properties of the Wishart distribution. *Annals of Math. Statist.* **35**, 261–269 (1964).

A. Puhalskii. Large deviations and idempotent probability. In *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*, volume 119. Chapman & Hall/CRC, Boca Raton, FL (2001).

C. R. Rao. *Linear statistical inference and its application.* Wiley (1973).

G. Rempała and J. Wesołowski. Asymptotics for products of independent sums with an application to Wishart determinants. *Stat. Probab. Lett.* **74** (2), 129–138 (2005).

R. T. Rockafellar. Integrals which are convex functionals, II. *Pacific J. Math.* **39** (2), 439–469 (1971).

E. B. Saff and V. Totik. *Logarithmic potentials with external fields.* Springer (1997).

C. P Schnorr. A hierarchy of polynomial time basis reduction algorithms. In *Theory of algorithms, Colloq. Pécs/Hung. 1984, Colloq. Math. Soc. János Bolyai*, volume 44, pages 375–386 (1986).

M. S. Srivastava. Singular Wishart and multivariate Beta distributions. *Ann. Statist.* **31** (5), 1537–1560 (2003).

Y. Tsaig and D. L. Donoho. Breakdown of equivalence between the minimal ℓ^1-norm solution and the sparsest solution. *Signal Processing* **86** (3), 533–548 (2006).