Electronic Journal of Probability

Vol. 9 (2004), Paper no. 7, pages 177–208.

Journal URL
http://www.math.washington.edu/~ejpecp/

DIFFERENTIAL OPERATORS AND SPECTRAL DISTRIBUTIONS OF INVARIANT ENSEMBLES FROM THE CLASSICAL ORTHOGONAL POLYNOMIALS. THE CONTINUOUS CASE

M. Ledoux

Institut de Mathématiques, Université Paul-Sabatier
31062 Toulouse, France
Email: ledoux@math.ups-tlse.fr

Abstract. Following the investigation by U. Haagerup and S. Thorbjørnsen, we present a simple differential approach to the limit theorems for empirical spectral distributions of complex random matrices from the Gaussian, Laguerre and Jacobi Unitary Ensembles. In the framework of abstract Markov diffusion operators, we derive by the integration by parts formula differential equations for Laplace transforms and recurrence equations for moments of eigenfunction measures. In particular, a new description of the equilibrium measures as adapted mixtures of the universal arcsine law with an independent uniform distribution is emphasized. The moment recurrence relations are used to describe sharp, non asymptotic, small deviation inequalities on the largest eigenvalues at the rate given by the Tracy-Widom asymptotics.

Submitted to EJP on February 3, 2003. Final version accepted on February 15, 2004.

1. Introduction

Limiting distributions of spectral measures of random matrices have been studied extensively in mathematical physics, since the pioneering work by E. Wigner (cf. [Me], [De], [Fo]), as well as in the context of multivariate analysis in statistics (cf. [Bai]).

To illustrate, as an introduction, some of the classical results, consider first, for each integer $N \geq 1$, $X = X^N$ a selfadjoint centered Gaussian random $N \times N$ matrix with variance σ^2. By this, we mean that X is a $N \times N$ Hermitian matrix such that the entries above the diagonal are independent complex (real on the diagonal) Gaussian random variables with mean zero and variance σ^2. Equivalently, X is distributed according to

$$\mathbb{P}(dX) = \frac{1}{C} \exp\big(- \operatorname{Tr}(X^2)/2\sigma^2\big) dX \tag{1}$$

where dX is Lebesgue measure on the space of Hermitian $N \times N$ matrices, and X is then said to be element of the Gaussian Unitary Ensemble (GUE). For such an Hermitian random matrix X, denote by $\lambda_1^N, \ldots, \lambda_N^N$ the eigenvalues of $X = X^N$. Denote furthermore by $\widehat{\mu}_\sigma^N$ the mean (empirical) spectral measure $\mathbb{E}\big(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^N}\big)$ of $X = X^N$. It is a classical result due to E. Wigner [Wi] that the mean density $\widehat{\mu}_\sigma^N$ converges weakly, as $\sigma^2 \sim \frac{1}{4N}$, $N \to \infty$, to the semicircle law (on $(-1, +1)$).

The second main example arose in the context of the so-called complex Wishart distributions. Let G be a complex $M \times N$ random matrix the entries of which are independent complex Gaussian random variables with mean zero and variance σ^2, and set $Y = Y^N = G^*G$. The law of Y defines similarly a unitary invariant probability measure on the space of Hermitian matrices called the Laguerre Unitary Ensemble (LUE) (cf. [Fo]). Denote by $\lambda_1^N, \ldots, \lambda_N^N$ the eigenvalues of the Hermitian matrix $Y = Y^N$ and by $\widehat{\mu}_\sigma^N$ the mean spectral measure $\mathbb{E}\big(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^N}\big)$. It is a classical result due to V. Marchenko and L. Pastur [M-P] (see also [G-S], [Wa], [Jon] for the real case in the context of sample covariance matrices) that the mean density $\widehat{\mu}_\sigma^N$ converges weakly, as $\sigma^2 \sim \frac{1}{4N}$ and $M = M(N) \sim cN$, $N \to \infty$, $c > 0$, to the so-called Marchenko-Pastur distribution (with parameter $c > 0$), or free Poisson distribution.

Recently, U. Haagerup and S. Thorbjørnsen [H-T] gave an entirely analytical treatment of these results on asymptotic eigenvalue distributions. This analysis is made possible by the determinantal representation of the eigenvalue distribution as a Coulomb gas and the use of orthogonal polynomials (cf. [Me], [De], [Fo], [H-T]...). For example, in the GUE case, by unitary invariance of the ensemble (1) and the Jacobian formula, the distribution of the eigenvalues $(\lambda_1^N, \ldots, \lambda_N^N)$ of $X = X^N$ is given by

$$\frac{1}{C}\,\Delta_N(x)^2 \prod_{i=1}^N d\mu(x_i/\sigma), \quad x = (x_1, \ldots, x_N) \in \mathbb{R}^N, \tag{2}$$

where

$$\Delta_N(x) = \prod_{1 \leq i < j \leq N} (x_j - x_i)$$

is the Vandermonde determinant, $d\mu(x) = \mathrm{e}^{-x^2/2}\frac{dx}{\sqrt{2\pi}}$ the standard normal distribution on $\mathbb{R}$ and C the normalization factor. Denote by P_ℓ, $\ell \in \mathbb{N}$, the normalized Hermite polynomials with respect to μ. Since, for each ℓ, P_ℓ is a polynomial function of degree ℓ, up to a constant depending on N, the Vandermonde determinant $\Delta_N(x)$ is easily seen to be equal to

$$\det\big(P_{\ell-1}(x_k)\big)_{1\leq k,\ell\leq N}\,.$$

On the basis of this observation and the orthogonality properties of P_ℓ, the marginals of the eigenvalue vector $(\lambda_1^N, \dots, \lambda_N^N)$ (the so-called correlation functions) may be represented as determinants of the (Hermite) kernel $K_N(x,y) = \sum_{\ell=0}^{N-1} P_\ell(x)P_\ell(y)$. In particular, the mean spectral measure $\widehat{\mu}_\sigma^N$ of the Gaussian matrix X is given, for every bounded measurable real-valued function f on $\mathbb{R}$, by

$$\int f d\widehat{\mu}_\sigma^N = \int f(\sigma x)\,\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2(x)d\mu(x). \tag{3}$$

Coulomb gas of the type (2) may be associated to a number of orthogonal polynomial ensembles for various choices of the underlying probability measure μ and its associated orthogonal polynomials. This concerns unitary random matrix ensembles (cf. [De], [Fo]...) as well as various random growth models associated to discrete orthogonal polynomial ensembles (cf. in particular [Joha1], [Joha2]). In this work, we only consider random matrix models associated to the classical orthogonal polynomial of the continuous variable (for the discrete case, see [Le2]) for which determinantal representations of the correlation functions are available. For example, the mean spectral measure $\widehat{\mu}_\sigma^N$ of the Wishart matrix Y may be represented in the same way by

$$\int f d\widehat{\mu}_\sigma^N = \Big(1-\frac{M}{N}\Big)^+ f(0) + \int f(\sigma^2 x)\,\frac{1}{N}\sum_{\ell=0}^{M\wedge N-1} P_\ell^2(x)d\mu(x) \tag{4}$$

for every bounded measurable function f on $\mathbb{R}_+$, where P_ℓ, $\ell \in \mathbb{N}$, are now the normalized orthogonal (Laguerre) polynomials for the Gamma distribution $d\mu(x) = \Gamma(\gamma+1)^{-1}x^\gamma \mathrm{e}^{-x}dx$ on $(0,+\infty)$ with parameter $\gamma = |M-N|$. (The Dirac mass at 0 comes from the fact that when $M < N$, $N-M$ eigenvalues of Y are necessarily 0.)

Based on these orthogonal polynomial representations, U. Haagerup and S. Thorbjørnsen deduce in [H-T] the asymptotic distributions of the complex Gaussian and Wishart spectral measures from explicit formulas for the Laplace transforms of the mean spectral measures in terms of confluent hypergeometric functions. They also determine recurrence equations for moments and describe the almost sure asymptotic behavior of the largest and smallest eigenvalues of these models.

In this paper, we actually push forward the investigation by U. Haagerup and S. Thorbjørnsen. With respect to their work, we however avoid any appeal to series expansions and rather only concentrate on the differential aspects of confluent hypergeometric functions. We present the methodology in the abstract framework of Markov diffusion generators, that provides a convenient language to derive both the basic differential equations on Laplace transforms and recurrence formulas for moments from the integration by parts formula. The setting moreover allows us to consider in the same framework and with no more efforts, spectral measures described by Jacobi polynomials, associated to Beta random matrices. Precisely, if G_1 and G_2 are independent complex, respectively $M_1 \times N$ and $M_2 \times N$ with $M_1 + M_2 \geq N$, random matrices the entries of which are independent complex Gaussian random variables with mean zero and variance 1, set $Y_1 = G_1^* G_1$, $Y_2 = G_2^* G_2$ and

$$Z = Z^N = \mathrm{Id} - 2Y_2(Y_1 + Y_2)^{-1} = (Y_1 - Y_2)(Y_1 + Y_2)^{-1}.$$

The law of Z defines a unitary invariant probability distribution, element of the Jacobi Unitary Ensemble (JUE). Denote again by $\lambda_1^N, \ldots, \lambda_N^N$ the eigenvalues of the Hermitian matrix $Z = Z^N$. Denote furthermore by $\widehat{\mu}^N$ the mean spectral measure $\mathbb{E}\big(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^N}\big)$ of $Z = Z^N$. Then, using the Coulomb gas description of the eigenvalues (cf. [T-W2], [Fo]), the mean spectral measure $\widehat{\mu}^N$ of $Z = Z^N$ may be represented, for every bounded measurable function f on $(-1, +1)$, by

$$\int f d\widehat{\mu}^N = \Big(1 - \frac{M_1}{N}\Big)^+ f(-1) + \int f(x)\, \frac{1}{N} \sum_{\ell=0}^{L-1} P_\ell^2(x) d\mu(x) + \Big(1 - \frac{M_2}{N}\Big)^+ f(+1) \quad (5)$$

with $L = M_1 \wedge N + M_2 \wedge N - N$, where P_ℓ, $\ell \in \mathbb{N}$, are the normalized orthogonal (Jacobi) polynomials for the Beta, or Jacobi, distribution

$$d\mu(x) = \frac{\Gamma(\alpha + \beta + 2)}{2^{\alpha+\beta+1}\Gamma(\alpha + 1)\Gamma(\beta + 1)}\, (1 - x)^\alpha (1 + x)^\beta dx$$

on $(-1, +1)$ with parameters $\alpha = |M_1 - N|$ and $\beta = |M_2 - N|$. Equilibrium measures of the mean density $\widehat{\mu}^N$ as $M_1 = M_1(N) \sim aN$ and $M_2 = M_2(N) \sim bN$, $N \to \infty$, $a, b > 0$, are described by P. Forrester in the symmetric case, see [Fo] and the references therein, and B. Collins [Co] for another particular family using asymptotics of Jacobi polynomials [C-I]. These limits may be predicted by the free probability calculus and the result of D. Voiculescu [Vo] and have been obtained this way in full generality by M. Capitaine and M. Casalis [C-C]. In our setting, their construction will follow the one of the Marchenko-Pastur distribution (and seems to be connected with free multiplicative products of Bernoulli laws [Co]).

It is worthwhile mentioning that the Gaussian Unitary random matrix Ensemble may be seen as a limiting case of the Laguerre Unitary random matrix Ensemble, and similarly the latter is a limiting case of the Jacobi Unitary random matrix Ensemble. Namely, with $\sigma = 1$,

$$M_1 Y_2 (Y_1 + Y_2)^{-1} \to Y_2$$

almost surely as $M_1 \to \infty$, whereas

$$\sqrt{M}\Big(\frac{Y}{M} - \mathrm{Id}\Big) \to X$$

almost surely as $M \to \infty$.

The main results of this work are presented in Section 3 where the framework of abstract Markov operators is used to derive identities for a general class of eigenfunction measures. In this setting, Hermite, Laguerre and Jacobi operators are treated in complete analogy in Section 4. On the basis of the general abstract identities developed in Section 3, we actually deal first with the eigenfunction measures $P_N^2 d\mu$ for which the universal limiting arcsine distribution is emphasized in accordance with the classical theory in the compact case developed by A. Maté, P. Nevai and V. Totik [M-N-T]. The method of [M-N-T], based on the recurrence relations for orthogonal polynomials, may certainly be adapted similarly to the non-compact case (and orthogonal polynomials with varying coeffícents). We however develop here a strategy using differential equations and moment recurrence relations that is well-suited to the small deviation inequalities on the largest eigenvalues presented in Section 5. The limiting distributions for the Cesaro averages $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$ (corresponding to the mean spectral measures) then appear as mixtures of affine transformations of the arcsine law with an independent uniform distribution, leading thus to a new picture of the Marchenko-Pastur law and the equilibrium law of Beta random matrices. These equilibrium measures also appear as limits of empirical measures on the zeroes of the associated orthogonal polynomials [K-VA].

¿From the general identites of Section 3, we obtain in Section 5 recurrence formulas for moments of the mean spectral measures of the three ensembles GUE, LUE and JUE. These recurrence equations are used to derive in a simple way small deviation inequalities on the largest eigenvalues for matrices with fixed size of the GUE, LUE and JUE, in concordance with the Tracy-Widom asymptotics [T-W1]. For example, it is known that the largest eigenvalue $\lambda_{\max}^N$ of the GUE random matrix $X = X^N$ with $\sigma^2 = \frac{1}{4N}$ converges almost surely to the right endpoint 1 of the support of the limiting spectral distribution (semicircle law). C. Tracy and H. Widom [T-W1] described the fluctuations of $\lambda_{\max}^N$ at the rate $N^{2/3}$. They showed that (some multiple of)

$$N^{2/3}(\lambda_{\max}^N - 1)$$

converges weakly to the Tracy-Widom distribution F constructed as a Fredholm determinant of the Airy kernel (as a limit in this regime of the Hermite kernel using delicate Plancherel-Rotach orthogonal polynomial asymptotics), and that may be described as

$$F(s) = \exp\bigg(-\int_s^\infty (x-s)u(x)^2 dx\bigg), \quad s \in \mathbb{R}, \tag{6}$$

where $u(x)$ is the unique solution of the Painlevé equation $u'' = 2u^3 + xu$ with the asymptotics $u(x) \sim \mathrm{Ai}\,(x)$ as $x \to \infty$. Similar results have been established for the largest eigenvalues of the Laguerre [Joha1], [Joho] and recently the Jacobi [Co] (cf. also [Fo]) Unitary Ensembles. For discrete orthogonal polynomial ensembles, see [Joha1], [Joha2].

Using the moment equations for the spectral distribution, we actually show that for every N and $0 < \varepsilon \leq 1$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1+\varepsilon\}\big) \leq C\,\mathrm{e}^{-N\varepsilon^{3/2}/C}$$

for some numerical constant $C > 0$, and similarly for the largest and (smallest – soft edge) eigenvalues of the LUE and soft edge of the symmetric JUE. These small deviation inequalities thus agree with the fluctuation rate $N^{2/3}$ (choose ε of the order of $N^{-2/3}$) and the asymptotics of the Tracy-Widom distribution

$$C'\,\mathrm{e}^{-s^{3/2}/C'} \leq 1 - F(s) \leq C\,\mathrm{e}^{-s^{3/2}/C}$$

for s large. They also related to the large deviation principle of [BA-D-G]. We obtain these results from bounds on the p-th moments of the trace by a simple induction procedure on the recurrence relations. The argument stays at a rather mild level, however easily tractable. In particular, with respect to the Tracy-Widom asymptotics, only the first correlation function is used and no sharp asymptotics of orthogonal polynomials is required. Non asymptotic small deviation inequalities on the largest eigenvalues of the LUE (as a limit of the discrete Meixner orthogonal polynomial ensemble) may also be shown to follow from earlier bounds that K. Johansson [Joha1] obtained using a delicate large deviation theorem together with a superadditivity argument (cf. [Le2]). In the limit from the LUE to the GUE, these bounds also cover the GUE case. An approach via hypercontractivity is attempted in [Le1]. In [Au], bounds over integral operators are used to this task.

In the companion paper [Le2], the analytic approach presented here is developed, with similar results, for discrete orthogonal polynomial ensembles as deeply investigated recently by K. Johansson [Joha1], [Joha2].

2. Abstract Markov operator framework

As announced, to unify the various examples we will investigate next, we consider more generally the setting of abstract Markov generators (see [Bak], [F-O-T]). The general conclusions obtained here might be of independent interest. Let thus μ be a probability measure on some measurable space $(E, \mathcal{E})$. Let L be a Markov generator invariant and symmetric with respect to μ. To work with more freedom, we assume that we are given an algebra $\mathcal{F}$ of functions, dense in the domain of L and contained in all L^p spaces (with respect to μ).

Given functions f, g in $\mathcal{F}$, define the carré du champ operator Γ by

$$\Gamma(f,g) = \frac{1}{2}\,\mathrm{L}(fg) - f\mathrm{L}g - g\mathrm{L}f.$$

Since L is invariant, the integration by parts formula indicates that for any f, g in $\mathcal{F}$,

$$\int f(-\mathrm{L}g)d\mu = \int \Gamma(f,g)d\mu. \tag{7}$$

Our main assumption in this work is that L is a diffusion operator, that is, Γ is a derivation in the sense that, for every smooth function ϕ on $\mathbb{R}^k$, any $f \in \mathcal{F}$ and any finite family $F = (f_1, \ldots, f_k)$ in $\mathcal{F}$,

$$\Gamma\big(f, \phi(F)\big) = \sum_{i=1}^{k} \partial_i \phi(F)\, \Gamma(f, f_i). \tag{8}$$

This setting conveniently includes the three basic one-dimensional diffusion operators we will consider in this work.

a) The first one is the Hermite or Ornstein-Uhlenbeck operator. Here, $E = \mathbb{R}$ and

$$\mathrm{L}f = f'' - xf'$$

is the Hermite operator with invariant measure the standard Gaussian measure $d\mu(x) = \mathrm{e}^{-x^2/2} \frac{dx}{\sqrt{2\pi}}$. The carré du champ operator is given, on smooth functions, by $\Gamma(f, g) = f'g'$. (One may choose for $\mathcal{F}$ the class of smooth functions with polynomial growth). In addition, the (normalized in $\mathrm{L}^2(\mu)$) Hermite polynomials P_N, $N \in \mathbb{N}$, that form an orthogonal basis of $\mathrm{L}^2(\mu)$, are eigenfunctions of $-\mathrm{L}$ with respective eigenvalues N, $N \in \mathbb{N}$.

b) Next we turn to the Laguerre operator. Consider the second order differential operator L acting on smooth functions f on $E = (0, +\infty)$ as

$$\mathrm{L}f = xf'' + (\gamma + 1 - x)f'$$

where $\gamma > -1$. The Laguerre operator L has invariant (Gamma) probability distribution $d\mu(x) = d\mu^\gamma(x) = \Gamma(\gamma + 1)^{-1} x^\gamma \mathrm{e}^{-x} dx$ (on $E = (0, +\infty)$). The carré du champ operator is given, on smooth functions, by $\Gamma(f, g) = xf'g'$. (One may choose for $\mathcal{F}$ the class of smooth functions on $E = (0, +\infty)$ with polynomial growth.) The (normalized in $\mathrm{L}^2(\mu)$) Laguerre polynomials P_N, $N \in \mathbb{N}$, that form an orthogonal basis of $\mathrm{L}^2(\mu)$, are eigenfunctions of $-\mathrm{L}$ with eigenvalues N, $N \in \mathbb{N}$.

c) Our third example concerns the Jacobi operator

$$\mathrm{L}f = (1 - x^2)f'' + \big[\beta - \alpha - (\alpha + \beta + 2)x\big]f',$$

$\alpha, \beta > -1$, acting on smooth functions f on $E = (-1, +1)$. The operator L has invariant (Beta or Jacobi) probability distribution

$$d\mu(x) = d\mu^{\alpha,\beta}(x) = \frac{\Gamma(\alpha + \beta + 2)}{2^{\alpha+\beta+1}\Gamma(\alpha + 1)\Gamma(\beta + 1)}\, (1 - x)^\alpha (1 + x)^\beta dx$$

(on $E = (-1, +1)$). The carré du champ operator is given, on smooth functions, by $\Gamma(f, g) = (1 - x^2)f'g'$. The (normalized in $\mathrm{L}^2(\mu)$) Jacobi polynomials P_N, $N \in \mathbb{N}$, that form an orthogonal basis of $\mathrm{L}^2(\mu)$, are eigenfunctions of $-\mathrm{L}$ with eigenvalues $\rho_N = N(\alpha + \beta + N + 1)$, $N \in \mathbb{N}$.

When $\alpha = \beta = \frac{\kappa}{2} - 1$, $\kappa > 0$, we speak of the ultraspheric or symmetric Jacobi operator

$$\mathrm{L}f = (1 - x^2)f'' - \kappa x f'$$

and distribution $d\mu_\kappa(x) = 2^{1-\kappa}\Gamma(\kappa)\Gamma(\frac{\kappa}{2})^{-2}(1-x^2)^{\frac{\kappa}{2}-1}dx$ of parameter (or dimension) κ. When κ is an integer, the ultraspheric measure μ_k is the projection on a diameter of the uniform measure on the κ-dimensional sphere.

Ultraspherical measures will be the basic limiting distributions in this setting, and particular values of κ are $\kappa = 1$ corresponding to the arcsine law, $\kappa = 2$ corresponding to the uniform distribution, and $\kappa = 3$ corresponding to the semicircle law. To this end, it is worthwhile describing the Laplace, or Fourier, transforms of the Jacobi distributions. Namely, let

$$J(\lambda) = \int \mathrm{e}^{\lambda x} d\mu, \quad \lambda \in \mathbb{C},$$

be the (complex) Laplace transform of $\mu = \mu^{\alpha,\beta}$. Since $-\mathrm{L}x = (\alpha+\beta+2)x + \alpha - \beta$, by the integration by parts formula,

$$(\alpha+\beta+2)J' + (\alpha-\beta)J = \int (-\mathrm{L}x)\,\mathrm{e}^{\lambda x} d\mu = \lambda \int (1-x^2)\,\mathrm{e}^{\lambda x} d\mu.$$

In particular thus

$$\lambda J'' + (\alpha+\beta+2)J' - \big[\lambda - (\alpha-\beta)\big]J = 0.$$

If J_κ is the Laplace transform of the symmetric Jacobi distribution μ_κ with parameter $\kappa > 0$, then J_κ solves the differential equation

$$\lambda J_\kappa'' + \kappa J_\kappa' - \lambda J_\kappa = 0. \tag{9}$$

Note also that for every $\kappa > 0$,

$$J_\kappa - J_\kappa'' = \mathrm{const.}\, J_{\kappa+2}, \quad \text{hence} \quad J_\kappa' = \mathrm{const.}\, \lambda J_{\kappa+2} \tag{10}$$

by (9). It will also be useful later to observe from (9) that the function $\widetilde{J}(\lambda) = \mathrm{e}^{v\lambda} J_\kappa(u\lambda)$, $u, v \in \mathbb{R}$, $u \neq 0$, satisfies the differential equation

$$\lambda \widetilde{J}'' + (\kappa - 2v\lambda)\widetilde{J}' - \big[\kappa v - (v^2 - u^2)\lambda\big]\widetilde{J} = 0. \tag{11}$$

The differential equation (9) may be expressed equivalently on the moments of μ_κ. Namely, if

$$\chi_p = \int x^{2p} d\mu_\kappa, \quad p \in \mathbb{N},$$

(the odd moments are clearly all zero by symmetry), then we have the recurrence equation

$$\chi_p = \frac{2p-1}{2p-1+\kappa}\,\chi_{p-1}, \quad p \geq 1. \tag{12}$$

It is worthwhile mentioning, following [Ma], that on the real line, up to affine transformations, the three preceding examples are the only Markov diffusion operators that may be diagonalized in a basis of orthogonal polynomials. Moreover, it is not difficult to check as in the introduction (cf. [Sz]) that the Hermite and Laguerre operators may be obtained in the limit from the Jacobi operators using a proper scaling.

3. Differential equations

In the preceding abstract framework, we now describe a simple procedure, relying on the integration by parts formula (7), to establish differential equations on measures given by $pqd\mu$ where p and q are eigenfunctions of $-\mathrm{L}$, with respective eigenvalues ρ_p and ρ_q (> 0).

We assume that we are given a function h of $\mathcal{F}$ such that

$$-\mathrm{L}h = A(h) \tag{13}$$

for some real-valued function A on $\mathbb{R}$. Assume moreover that for all $f, g \in \mathcal{F}$,

$$\Gamma(h,f)\Gamma(h,g) = B(h)\Gamma(f,g) \tag{14}$$

where $B : \mathbb{R} \to \mathbb{R}$ is smooth enough (or only μ-almost everywhere). We also agree that $\Gamma(h,h) = B(h)$.

Furthermore, we assume that h, p, q are compatible for L in the sense that for real constants d_p, d_q, and functions $D_p, D_q : \mathbb{R} \to \mathbb{R}$,

$$-\mathrm{L}\Gamma(h,p) = d_p\Gamma(h,p) + D_p(h)p \quad \text{and} \quad -\mathrm{L}\Gamma(h,q) = d_q\Gamma(h,q) + D_q(h)q. \tag{15}$$

The main result in this setting is a differential identity on the operator F given by

$$F(\theta) = \int \theta(h)pqd\mu$$

for smooth functions $\theta : \mathbb{R} \to \mathbb{R}$. (We assume throughout the argument that all the integrals we will deal with are well-defined.)

Proposition 3.1. *Let $\theta : \mathbb{R} \to \mathbb{R}$ be smooth. Under the preceding assumptions on p, q and h,*

$$F\big(\mathcal{R}(\theta'')\big) - 2(\rho_p + \rho_q)F(B\theta'') + 2F\big(\mathcal{R}\big((B\theta'')'\big)\big) - sF(\theta) - tF\big(\mathcal{R}(\theta')\big) - F(D\theta') = 0$$

where $\mathcal{R}$ is the differential operator defined, on smooth functions $\phi : \mathbb{R} \to \mathbb{R}$, by $\mathcal{R}(\phi) = A\phi - B\phi'$, $D = D_p + D_q$ and

$$s = \frac{1}{2}(\rho_p - \rho_q)(d_p - d_q + \rho_p - \rho_q) \quad \text{and} \quad t = \frac{1}{2}(d_p + d_q - \rho_p - \rho_q).$$

Proof. First note that, by the integration by parts formula (7) and diffusion property (8), and assumptions (13), (14), for any smooth $\phi : \mathbb{R} \to \mathbb{R}$ and any $f \in \mathcal{F}$,

$$\int \mathcal{R}(\phi)(h) f d\mu = \int \phi(h)\Gamma(h, f) d\mu.$$

As a consequence, by (8) again,

$$F\big(\mathcal{R}(\theta)\big) = \int \theta(h)\big[\Gamma(h,p)q + \Gamma(h,q)p\big] d\mu \tag{16}$$

and

$$F\big(\mathcal{R}^2(\theta)\big) = 2\int B(h)\theta(h)\Gamma(p,q)d\mu + \int \theta(h)\big[\Gamma\big(h,\Gamma(h,p)\big)q + \Gamma\big(h,\Gamma(h,q)\big)p\big] d\mu \tag{17}$$

where we used (14).

Since $-\mathrm{L}p = \rho_p p$, by the integration by parts formula for L again,

$$\rho_p F(\theta) = \int \theta'(h)\Gamma(h,q) p d\mu + \int \theta(h)\Gamma(p,q) d\mu$$

and similarly, using $-\mathrm{L}q = \rho_q q$,

$$\rho_q F(\theta) = \int \theta'(h)\Gamma(h,p) q d\mu + \int \theta(h)\Gamma(p,q) d\mu.$$

In particular, together with (16) for θ',

$$2\int \theta(h)\Gamma(p,q) d\mu = (\rho_p + \rho_q)F(\theta) - F\big(\mathcal{R}(\theta')\big). \tag{18}$$

We also have that

$$2\int \theta'(h)\Gamma(h,p) q d\mu = (\rho_p - \rho_q)F(\theta) + F\big(\mathcal{R}(\theta')\big) \tag{19}$$

together with the same formula exchanging the roles of p and q.

Using (15), we now have that

$$\begin{aligned} &\int \theta'(h)\Gamma\big(h,\Gamma(h,p)\big) q d\mu \\ &\qquad = d_p \int \theta(h)\Gamma(h,p) q d\mu + F(D_p\theta) - \int \theta(h)\Gamma\big(q,\Gamma(h,p)\big) d\mu \end{aligned} \tag{20}$$

and similarly with q. To handle the last term on the right-hand side of (20), we use again the eigenfunction property of q to get that

$$\int \theta(h)\Gamma\big(q,\Gamma(h,p)\big) d\mu = \rho_q \int \theta(h)\Gamma(h,p) q d\mu - \int B(h)\theta'(h)\Gamma(p,q) d\mu.$$

Therefore, (20) yields

$$\int \theta'(h)\Gamma\big(h, \Gamma(h,p)\big) q d\mu \\ = (d_p - \rho_q) \int \theta(h)\Gamma(h,p) q d\mu + F(D_p\theta) + \int B(h)\theta'(h)\Gamma(p,q) d\mu. \tag{21}$$

We now summarize the preceding: Applying (17) to θ'', and then (21) to θ' (with the same formula exchanging the roles of p and q), we get that

$$F\big(\mathcal{R}^2(\theta'')\big) = 4 \int B(h)\theta''(h)\Gamma(p,q) d\mu + (d_p - \rho_q) \int \theta'(h)\Gamma(h,p) q d\mu \\ + (d_q - \rho_p) \int \theta'(h)\Gamma(h,q) p d\mu + F(D\theta').$$

The conclusion follows by applying (18) to $B\theta''$ and (19). ☐

Corollary 3.2. *In the setting of Proposition 3.1,*

$$F(T_4\theta^{(4)}) + F(T_3\theta''') + F(T_2\theta'') + F(T_1\theta') + F(T_0\theta) = 0$$

where the real-valued functions T_4, T_3, T_2, T_1, T_0 are given by

$$\begin{aligned} T_4 &= -B^2, \\ T_3 &= -3BB', \\ T_2 &= A^2 - A'B + 2AB' + tB - 2(\rho_p + \rho_q)B - 2BB'', \\ T_1 &= -tA - D, \\ T_0 &= -s. \end{aligned}$$

In case the functions A, B, D are polynomials, Corollary 3.2 may be used directly to deduce a differential equation for the Laplace of Fourier transforms of $pqd\mu$ as well as recurrence formulas for its moments. Indeed, when $\theta(x) = \mathrm{e}^{\lambda x}$, $x \in \mathbb{R}$ and λ is a parameter in some open domain in $\mathbb{C}$, we get the following consequence. For a given (complex) polynomial

$$P(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$$

of the variable λ, denote by

$$\mathcal{P}(\phi) = a_n\phi^{(n)} + \cdots + a_1\phi' + a_0\phi$$

the corresponding differential operator acting on smooth (complex) functions ϕ.

Corollary 3.3. *Let*

$$\varphi(\lambda) = \int \mathrm{e}^{\lambda h} pq d\mu$$

assumed to be well-defined for λ in some open domain of $\mathbb{C}$. Then,

$$\lambda^4 \mathcal{T}_4(\varphi) + \lambda^3 \mathcal{T}_3(\varphi) + \lambda^2 \mathcal{T}_2(\varphi) + \lambda \mathcal{T}_1(\varphi) + \mathcal{T}_0(\varphi) = 0.$$

When $\theta(x) = x^k$, $x \in \mathbb{R}$ and $k \in \mathbb{N}$, we may deduce from Corollary 3.2 recurrence equations for moments (which will be used specifically in Section 5). In the next sections, we examine the Hermite, Laguerre and Jacobi diffusion operators with respect to their orthogonal polynomials.

4. Limiting spectral distributions of GUE, LUE and JUE

In this section, we test the efficiency of the preceding abstract results on the three basic examples, Gaussian, Laguerre and Jacobi, random matrix ensembles. We actually discuss first eigenfunction measures of the type $P_N^2 d\mu$ for which, under appropriate normalizations, the arcsine law appears as a universal limiting distribution (cf. [M-N-T]). The Cesaro averages $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$ associated to the mean spectral measures of the random matrix models are then obtained as appropriate mixtures of the arcsine law with an independent uniform distribution. This approach leads to a new look at the classical equilibrium measures of the spectral measures (such as the Marchenko-Pastur distribution). We are indebted to T. Kurtz, and O. Zeitouni, for enlightning remarks leading to this presentation.

We refer below to [Sz], [Ch]..., or [K-K], for the basic identities and recurrence formulas for the classical orthogonal polynomials.

a) The Hermite case. Let

$$\varphi(\lambda) = \int e^{\lambda x} P_N^2 d\mu, \quad \lambda \in \mathbb{C},$$

where μ is the standard normal distribution on $\mathbb{R}$ and P_N the N-th (normalized) Hermite polynomial for μ. We apply the general conclusions of Section 3. Choose $h = x$ so that $A = x$, $B = 1$, and $p = q = P_N$ so that $\rho_p = \rho_q = N$. Since $\Gamma(h,p) = P_N' = \sqrt{N}\, P_{N-1}$, we also have that $d_p = d_q = N-1$ and $D = 0$. Hence $s = 0$, $t = -1$, and $T_4 = -1$, $T_3 = 0$, $T_2 = x^2 - 4N - 2$, $T_1 = x$ and $T_0 = 0$. Therefore, by Corollary 3.3,

$$-\lambda^4 \varphi + \lambda^2 \big[\varphi'' - (4N+2)\varphi\big] + \lambda \varphi' = 0,$$

that is

$$\lambda \varphi'' + \varphi' - \lambda(\lambda^2 + 4N + 2)\varphi = 0.$$

Changing λ into $\sigma\lambda$, $\sigma > 0$, $\varphi_\sigma(\lambda) = \varphi(\sigma\lambda)$ solves the differential equation

$$\lambda \varphi_\sigma'' + \varphi_\sigma' - \sigma^2 \lambda(\sigma^2 \lambda^2 + 4N + 2)\varphi_\sigma = 0.$$

Whenever φ_σ and its derivatives converge as $\sigma^2 \sim \frac{1}{4N}$, $N \to \infty$, the limiting function Φ will thus satisfy the differential equation

$$\lambda\Phi'' + \Phi' - \lambda\Phi = 0 \tag{22}$$

which is precisely the differential equation satisfied by the Laplace transform J_1 of the arcsine law (cf. (9)). Precisely, let V_N, $N \in \mathbb{N}$, be random variables with distributions $P_N^2 d\mu$, and denote by ξ a random variable with the arcsine distribution on $(-1,+1)$. Using the recurrence equation

$$xP_N = \sqrt{N+1}\,P_{N+1} + \sqrt{N}\,P_{N-1}$$

of the (normalized) Hermite polynomials, it is easily checked that, for example, $\sup_N \mathbb{E}(V_N^4/N^2) < \infty$. Extracting a weakly convergent subsequence, it follows by uniform integrability that, along the imaginary axis, the Fourier transform φ_σ of σV_N ($\sigma^2 \sim \frac{1}{4N}$), as well as its first and second derivatives φ'_σ and φ''_σ, converge pointwise as $N \to \infty$ to Φ, Φ' and Φ'' respectively. By (22), $\Phi = J_1$ so that $V_N/2\sqrt{N}$ converges weakly to ξ.

As we saw it in the introduction, to investigate spectral measures of random matrices and their equilibrium distributions, we are actually interested in the measures $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$. Towards this goal, observe that if $f : \mathbb{R} \to \mathbb{R}$ is bounded and continuous,

$$\begin{aligned}\int f\Big(\frac{x}{2\sqrt{N}}\Big)\frac{1}{N}\sum_{\ell=1}^{N-1} P_\ell^2 d\mu &= \frac{1}{N}\sum_{\ell=1}^{N-1} f\bigg(\sqrt{\frac{\ell}{N}}\cdot\frac{x}{2\sqrt{\ell}}\bigg)P_\ell^2 d\mu \\ &= \int_{1/N}^{1} \mathbb{E}\Bigg(f\bigg(\sqrt{U_N(t)}\cdot\frac{V_{NU_N(t)}}{2\sqrt{NU_N(t)}}\bigg)\Bigg)dt\end{aligned}$$

where $U_N(t) = \ell/N$ for $\ell/N < t \le (\ell+1)/N$, $\ell = 0,1,\dots,N-1$ ($U_N(0) = 0$). Since $U_N(t) \to t$, $t \in [0,1]$, and $V_N/2\sqrt{N}$ converges weakly to ξ, it follows that

$$\lim_{N\to\infty}\int f\Big(\frac{x}{2\sqrt{N}}\Big)\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu = \int_0^1 \mathbb{E}\Big(f\big(\sqrt{t}\,\xi\big)\Big)dt = \mathbb{E}\Big(f\big(\sqrt{U}\,\xi\big)\Big)$$

where U is uniform on $[0,1]$ and independent from ξ. Together with (3), we thus conclude to the following formulation.

Proposition 4.1. *Let $\widehat{\mu}_\sigma^N$ be the mean spectral measure of the random matrix X^N from the GUE with parameter σ. Whenever$\sigma^2 \sim \frac{1}{4N}$, then $\widehat{\mu}_\sigma^N$ converges wea kly to the law of $\sqrt{U}\,\xi$ with ξ an arcsine random variable on $(-1,+1)$ and U uniform on $[0,1]$ and independent from ξ.*

Of course, $\sqrt{U}\,\xi$ has the semicircle law, in accordance with the classical Wigner theorem. This may be checked directly, or on the moment identities (12). Alternatively, and more in the differential context of this paper, one may observe from Lemma 5.1 below that whenever ψ is the Laplace transform of the measure $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$, then $2N\lambda\psi = \varphi' - \lambda\varphi$. In the limit under the scaling $\sigma^2 \sim \frac{1}{4N}$, $\lambda\Psi = 2\Phi' = 2J_1'$ so that, by (10), $\Psi = J_3$ the Laplace transform of the semicircular law.

b) The Laguerre case. Let

$$\varphi(\lambda) = \int e^{\lambda x} P_N^2 d\mu, \quad \mathrm{Re}(\lambda) < 1,$$

where $\mu = \mu^\gamma$ is the Gamma distribution with parameter $\gamma > -1$ on $(0, +\infty)$ and $P_N = P_N^\gamma$ the N-th (normalized) Laguerre polynomial for μ^γ. As in the Hermite case, we apply the general conclusions of Section 3. Choose $h = x$ so that $A = x - (\gamma + 1)$, $B = x$, and $p = q = P_N$ so that $\rho_p = \rho_q = N$. Since

$$\Gamma(h, p) = xP_N' = -\sqrt{N}\,\sqrt{\gamma + N}\,P_{N-1} + NP_N,$$

we also have that $d_p = d_q = N - 1$ and $D = 2N$. Hence $s = 0$, $t = -1$, and $T_4 = -x^2$, $T_3 = -3x$, $T_2 = x^2 - 2(\gamma + 1 + 2N)x + \gamma^2 - 1$, $T_1 = x - (\gamma + 1 + 2N)$ and $T_0 = 0$. Therefore, by Corollary 3.3,

$$-\lambda^4\varphi'' - 3\lambda^3\varphi' + \lambda^2\big[\varphi'' - 2(\gamma+1+2N)\varphi' + (\gamma^2 - 1)\varphi\big] + \lambda\big[\varphi' - (\gamma+1+2N)\varphi\big] = 0,$$

that is

$$\lambda(1-\lambda^2)\varphi'' + \big[1 - 3\lambda^2 - 2(\gamma+1+2N)\lambda\big]\varphi' - \big[(\gamma+1+2N) - (\gamma^2-1)\lambda\big]\varphi = 0.$$

Changing λ into $\sigma^2\lambda$, $\sigma > 0$, $\varphi_{\sigma^2}(\lambda) = \varphi(\sigma^2\lambda)$ solves the differential equation

$$\lambda(1-\sigma^4\lambda^2)\varphi_{\sigma^2}'' + \big[1-3\sigma^4\lambda^2-2(\gamma+2N+1)\sigma^2\lambda\big]\varphi_{\sigma^2}' - \sigma^2\big[(\gamma+1+2N)-(\gamma^2-1)\sigma^2\lambda\big]\varphi_{\sigma^2} = 0.$$

Whenever φ_{σ^2} and its derivatives converge as $\sigma^2 \sim \frac{1}{4N}$ and $\gamma = \gamma_N \sim c'N$, $N \to \infty$, $c' \geq 0$, the limiting function Φ satisfies the differential equation

$$\lambda\Phi'' + \Big[1 - \frac{c'+2}{2}\,\lambda\Big]\Phi' - \Big[\frac{c'+2}{4} - \frac{c'^2}{16}\,\lambda\Big]\Phi = 0.$$

In other words, by (11), $\Phi(\lambda) = e^{v\lambda} J_1(u\lambda)$ with

$$v = \frac{c'+2}{4} \quad \text{and} \quad u^2 = v^2 - \frac{c'^2}{16} = \frac{c'+1}{4}\,. \tag{23}$$

Precisely, let $V_N = V_N^\gamma$, $N \in \mathbb{N}$, be random variables with distributions $P_N^2 d\mu$. Recall we denote by ξ a random variable with the arcsine distribution on $(-1, +1)$ and Laplace transform J_1. Using the recurrence equation

$$xP_N = -\sqrt{N+1}\,\sqrt{\gamma+N+1}\,P_{N+1} + (\gamma+2N+1)P_N - \sqrt{N}\,\sqrt{\gamma+N}\,P_{N-1}$$

of the (normalized) Laguerre polynomials of parameter γ, it is easily checked that, for example, $\sup_N \mathbb{E}((V_N^\gamma)^4/N^4) < \infty$, $\gamma = \gamma_N \sim c'N$, $c' \geq 0$. Extracting a weakly convergent subsequence, it follows by uniform integrability that, along the imaginary axis, the Fourier transform φ_{σ^2} of $\sigma^2 V_N$ ($\sigma^2 \sim \frac{1}{4N}$, $\gamma = \gamma_N \sim c'N$, $c' \geq 0$), as well as its first and second derivatives φ'_{σ^2} and φ''_{σ^2}, converge pointwise as $N \to \infty$ to Φ, Φ' and Φ'' respectively. As we have seen, $\Phi(\lambda) = \mathrm{e}^{v\lambda} J_1(u\lambda)$ so that $V_N/4N$ converges weakly to $u\xi + v$ where u and v are given by (23).

To investigate spectral measures of random matrices and their equilibrium distributions, we are actually interested in the measures $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$. We proceed as in the Hermite case, with however the additional dependence of $\mu = \mu^\gamma$ and $P_\ell = P_\ell^\gamma$ on the varying parameter $\gamma = \gamma_N \sim c'N$, $N \to \infty$, $c' \geq 0$. For $f : \mathbb{R} \to \mathbb{R}$ bounded and continuous, write

$$\begin{aligned}\int f\Big(\frac{x}{4N}\Big)\frac{1}{N}\sum_{\ell=1}^{N-1}(P_\ell^{\gamma_N})^2 d\mu^{\gamma_N} &= \frac{1}{N}\sum_{\ell=1}^{N-1}\int f\Big(\frac{\ell}{N}\cdot\frac{x}{4\ell}\Big)(P_\ell^{\gamma_N})^2 d\mu^{\gamma_N} dt \\ &= \int_{1/N}^1 \mathbb{E}\bigg(f\bigg(U_N(t)\cdot\frac{V_{NU_N(t)}^{\gamma_N}}{4NU_N(t)}\bigg)\bigg)dt\end{aligned}$$

where $U_N(t) = \ell/N$ for $\ell/N < t \leq (\ell+1)/N$, $\ell = 0, 1, \ldots, N-1$ ($U_N(0) = 0$). Since $U_N(t) \to t$, $t \in [0,1]$, and $V_N^{\gamma_N}/4N$ converges weakly to

$$\frac{1}{2}\sqrt{c'+1}\,\xi + \frac{1}{4}(c'+2)$$

as $\gamma_N/N \to c'$, it follows that

$$\begin{aligned}\lim_{N\to\infty}\int f\Big(\frac{x}{4N}\Big)\frac{1}{N}\sum_{\ell=0}^{N-1}(P_\ell^{\gamma_N})^2 d\mu^{\gamma_N} &= \mathbb{E}\bigg(f\bigg(U\bigg[\frac{1}{2}\sqrt{\frac{c'}{U}+1}\,\xi + \frac{1}{4}\Big(\frac{c'}{U}+2\Big)\bigg]\bigg)\bigg) \\ &= \mathbb{E}\bigg(f\Big(\frac{1}{2}\sqrt{U(c'+U)}\,\xi + \frac{1}{4}(c'+2U)\Big)\bigg)\end{aligned}$$

where U is uniform on $[0,1]$ and independent from ξ.

We obtain in this way convergence of the mean spectral measure $\widehat{\mu}_\sigma^N$, as $\sigma^2 \sim \frac{1}{4N}$ and $M = M(N) \sim cN$, $N \to \infty$, $c > 0$, of the Hermitian Wishart matrix $Y = Y^N = G^*G$ presented in the introduction. By appropriate operations on (4), we conclude to the following statement.

Proposition 4.2. *Let $\widehat{\mu}_\sigma^N$ be the mean spectral measure of the random matrix Y^N from the LUE with parameter σ. Whenever $\sigma^2 \sim \frac{1}{4N}$ and $M = M(N) \sim cN$, $N \to \infty$, $c > 0$, then $\widehat{\mu}_\sigma^N$ converges weakly to*

$$(1-c)^+\delta_0 + (c \wedge 1)\nu$$

where ν is the law of

$$(c \wedge 1)\Big(\frac{1}{2}\sqrt{U(c'+U)}\,\xi + \frac{1}{4}\,(c'+2U)\Big),$$

$c' = |c-1|/(c \wedge 1)$, *with ξ an arcsine random variable on $(-1,+1)$ and U uniform on $[0,1]$ and independent from ξ.*

Now of course, the law $\nu_{c'}$ of the random variable $\frac{1}{2}\sqrt{U(c'+U)}\,\xi + \frac{1}{4}\,(c'+2U)$ is the classical Marchenko-Pastur distribution [M-P] given by

$$d\nu_{c'}(x) = \text{const.}\ x^{-1}\sqrt{\big(x-(v-u)\big)\big((v+u)-x\big)}\,dx \tag{24}$$

on $(v-u, v+u) \subset (0,+\infty)$ where u and v are described by (23) so that

$$v \pm u = \frac{1}{4}\,\big(\sqrt{c'+1} \pm 1\big)^2.$$

This may be checked directly, although at the expense of some tedious details. Alternatively, more in the context of this paper, one may observe from Lemma 5.3 below that whenever ψ is the Laplace transform of the measure $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$, then

$$2N(\varphi + \lambda\psi') = (1-\lambda)\varphi' - (\gamma+1)\varphi.$$

In the limit under the scaling $\sigma^2 \sim \frac{1}{4N}$, $\gamma = \gamma_N \sim c'N$, $N \to \infty$, $c' \geq 0$,

$$2\lambda\Psi' = \Phi' - \frac{c'+4}{2}\,\Phi.$$

Since $\Phi = \mathrm{e}^{v\lambda}J_1(u\lambda)$ (with u and v given by (23)), it must be, by (10), that $\Psi' = \text{const.}\ \mathrm{e}^{v\lambda}J_3(u\lambda)$. Now, assuming a priori that Ψ is the Laplace transform $\Psi(\lambda) = \int \mathrm{e}^{\lambda x} d\nu_{c'}$ of some (probability) distribution $\nu_{c'}$ on $(0,+\infty)$, we have that

$$\Psi'(\lambda) = \int x\,\mathrm{e}^{\lambda x} d\nu_{c'} = \text{const.} \int_{-1}^{+1} \mathrm{e}^{\lambda(ux+v)}\sqrt{1-x^2}\,dx.$$

Changing the variable,

$$\int x\,\mathrm{e}^{\lambda x} d\nu_{c'} = \text{const.} \int_{v-u}^{v+u} \mathrm{e}^{\lambda x}\sqrt{\big[x-(v-u)\big]\big[(v+u)-x\big]}\,dx.$$

In other words, $\nu_{c'}$ is the Marchenko-Pastur distribution (24). Note that when $c' = 0$, then $v - u = 0$ and ν_0 is simply the distribution of ζ^2 where ζ has the semicircle law.

c) The Jacobi case. Let

$$\varphi(\lambda) = \int \mathrm{e}^{\lambda x} P_N^2 d\mu, \quad \lambda \in \mathbb{C},$$

where $\mu = \mu^{\alpha,\beta}$ is the Beta distribution with parameters $\alpha, \beta > -1$ on $(-1,+1)$ and $P_N = P_N^{\alpha,\beta}$ the N-th (normalized) Jacobi polynomial for $\mu^{\alpha,\beta}$. As in the Hermite and Laguerre cases, we apply the general conclusions of Section 3. Choose $h = x$ so that $A = (\alpha+\beta+2)x + \alpha - \beta$, $B = 1 - x^2$, and $p = q = P_N$ so that $\rho_p = \rho_q = \rho_N = N(\alpha+\beta+N+1)$. Since $\Gamma(h,p) = (1-x^2)P_N'$ and

$$\begin{aligned}(\alpha+\beta+2N)(1-x^2)P_N' = &-N\big[\beta-\alpha+(\alpha+\beta+2N)x\big]P_N \\ &+ 2(\alpha+N)(\beta+N)\kappa_{N-1}\kappa_N^{-1}P_{N-1}\end{aligned}$$

(where κ_N is defined below), we also have that $d_p = d_q = \rho_N - (\alpha+\beta)$ and $D = -4\rho_N x$. Hence $s = 0$, $t = -(\alpha+\beta)$, and

$$\begin{aligned}T_4 &= -x^4 + 2x^2 - 1, \\ T_3 &= -6x^3 + 6x, \\ T_2 &= \big[(\alpha+\beta)(\alpha+\beta+2) + 4\rho_N - 6\big]x^2 + 2(\alpha^2-\beta^2)x \\ &\quad + (\alpha-\beta)^2 - 2(\alpha+\beta) - 4\rho_N + 2, \\ T_1 &= \big[(\alpha+\beta)(\alpha+\beta+2) + 4\rho_N\big]x + \alpha^2 - \beta^2, \\ T_0 &= 0.\end{aligned}$$

By Corollary 3.3,

$$\begin{aligned}&-\lambda^3\varphi^{(4)} - 6\lambda^2\varphi''' + \lambda\big[2\lambda^2 + (\alpha+\beta)(\alpha+\beta+2) + 4\rho_N - 6\big]\varphi'' \\ &+ \Big(6\lambda^2 + 2(\alpha^2-\beta^2)\lambda + \big[(\alpha+\beta)(\alpha+\beta+2) + 4\rho_N\big]\Big)\varphi' \\ &+ \Big(-\lambda^3 + \big[(\alpha-\beta)^2 - 2(\alpha+\beta) - 4\rho_N + 2\big]\lambda + (\alpha^2-\beta^2)\Big)\varphi = 0.\end{aligned}$$

When $\alpha = \alpha_N \sim a'N$, $\beta = \beta_N \sim b'N$, $a', b' \geq 0$, as $N \to \infty$, whenever φ and its derivatives converge, the limiting function Φ solves the differential equation

$$\begin{aligned}\lambda(a'+b'+2)^2\Phi'' &+ \big[(a'+b'+2)^2 - 2(b'^2-b'^2)\lambda\big]\Phi' \\ &- \Big((b'^2-a'^2) - \big[(a'-b')^2 - 4(a'+b'+1)\big]\lambda\Big)\Phi = 0.\end{aligned}$$

In other words, by (11), $\Phi(\lambda) = \mathrm{e}^{v\lambda}J_1(u\lambda)$ where

$$\begin{aligned}v &= \frac{b'^2 - a'^2}{(a'+b'+2)^2}\,, \\ u^2 &= v^2 - \frac{(a'-b')^2 - 4(a'+b'+1)}{(a'+b'+2)^2} = \frac{16(a'+b'+1)(a'b'+a'+b'+1)}{(a'+b'+2)^4}\,.\end{aligned} \tag{25}$$

Precisely, let $V_N = V_N^{\alpha,\beta}$, $N \in \mathbb{N}$, be random variables with distributions $P_N^2 d\mu$. Recall we denote by ξ a random variable with the arcsine distribution on $(-1,+1)$ and Laplace transform J_1. Using the recurrence equation

$$\begin{aligned}(\alpha+\beta+2N)(\alpha+\beta+2N+1)(\alpha+\beta+2N+2)xP_N \\ = 2(N+1)(\alpha+\beta+N+1)(\alpha+\beta+2N)\kappa_N^{-1}\kappa_{N+1}P_{N+1} \\ + (\alpha+\beta+2N+1)(\beta^2-\alpha^2)P_N \\ + 2(\alpha+N)(\beta+N)(\alpha+\beta+2N+2)\kappa_N^{-1}\kappa_{N-1}P_{N-1}\end{aligned}$$

where

$$\kappa_N^2 = \frac{\Gamma(\alpha+\beta+2)\Gamma(\alpha+N+1)\Gamma(\beta+N+1)}{\Gamma(\alpha+1)\Gamma(\beta+1)(\alpha+\beta+2N+1)\Gamma(N+1)\Gamma(\alpha+\beta+N+1)}$$

of the (normalized) Jacobi polynomials of parameter α, β, it is easily checked that, for example, $\sup_N \mathbb{E}((V_N^{\alpha,\beta})^4) < \infty$, $\alpha = \alpha_N \sim a'N$, $\beta = \beta_N \sim b'N$, $a', b' \geq 0$. Extracting a weakly convergent subsequence, it follows by uniform integrability that, along the imaginary axis, φ, φ', φ'', converge pointwise as $\alpha = \alpha_N \sim a'N$, $\beta = \beta_N \sim b'N$, $N \to \infty$, $a', b' \geq 0$, to Φ, Φ' and Φ'' respectively. As we have seen, $\Phi(\lambda) = \mathrm{e}^{v\lambda} J_1(u\lambda)$ so that V_N converges weakly to $u\xi + v$ where u and v are given by (25). (Note that when $a' = b' = 0$, $\Phi = J_1$ recovering the result of [M-N-T] in this particular case.)

Cesaro means of orthogonal polynomials are studied as in the Hermite and Laguerre cases so to yield the following conclusion. Recall the mean spectral measure $\widehat{\mu}^N$ of the Beta random matrix

$$Z = Z^N = \mathrm{Id} - 2Y_2(Y_1+Y_2)^{-1} = (Y_1 - Y_2)(Y_1+Y_2)^{-1}$$

presented in the introduction for $M_1 = M_1(N) \sim aN$, $a > 0$, $M_2 = M_2(N) \sim bN$, $b > 0$, $a+b \geq 1$, as $N \to \infty$. Together with (5), we conclude to the following statement.

Proposition 4.3. *Let $\widehat{\mu}^N$ be the mean spectral measure of the random matrix Z^N from the JUE. Whenever $M_1 = M_1(N) \sim aN$, $a > 0$, $M_2 = M_2(N) \sim bN$, $b > 0$, $a + b \geq 1$, as $N \to \infty$, then $\widehat{\mu}^N$ converges weakly to*

$$(1-a)^+\delta_{-1} + (a \wedge 1 + b \wedge 1 - 1)\nu_{a',b'} + (1-b)^+\delta_{+1}$$

where $\nu_{a',b'}$ is the law of

$$\frac{4\sqrt{U(a'+b'+U)}\,\sqrt{a'b' + U(a'+b'+U)}}{(a'+b'+2U)^2}\,\xi + \frac{b'^2 - a'^2}{(a'+b'+2U)^2}$$

$a' = |a-1|/(a \wedge 1 + b \wedge 1 - 1)$, $b' = |b-1|/(a \wedge 1 + b \wedge 1 - 1)$, with ξ an arcsine random variable on $(-1,+1)$ and U uniform on $[0,1]$ and independent from ξ.

The distribution $\nu_{a',b'}$ may be identified, however after some tedious calcultations, with the distribution put forward in [Co] and [C-C] (see also [K-VA]) given by

$$d\nu_{a',b'}(x) = \text{const.}\ (1-x^2)^{-1}\sqrt{\big(x-(v-u)\big)\big((v+u)-x\big)}\,dx \tag{26}$$

on $(v-u, v+u) \subset (-1,+1)$ where u and v are described in (25). In the context of the differential equations for Laplace transforms studied here, it might be observed from Lemma 5.6 below that the limiting Laplace transform Ψ of the measures $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$ as $\alpha = \alpha_N \sim a'N$, $\beta = \beta_N \sim b'N$, satisfies

$$2\lambda(\Psi - \Psi'') = (a'+b'+2)\Phi' + \frac{a'^2 - b'^2}{a'+b'+2}\,\Phi.$$

Since $\Phi(\lambda) = \mathrm{e}^{v\lambda} J_1(u\lambda)$, it must be by (10) that $\Psi - \Psi'' = \text{const. } \mathrm{e}^{v\lambda} J_3(u\lambda)$. Now, assuming a priori that Ψ is the Laplace transform $\Psi(\lambda) = \int \mathrm{e}^{\lambda x} d\nu_{a',b'}$ of some (probability) distribution $\nu_{a',b'}$ on $(-1,+1)$, we have that

$$\Psi(\lambda) - \Psi''(\lambda) = \int (1 - x^2)\, \mathrm{e}^{\lambda x} d\nu_{a',b'} = \text{const.} \int_{-1}^{+1} \mathrm{e}^{\lambda(ux+v)} \sqrt{1 - x^2}\, dx.$$

Changing the variable,

$$\int (1 - x^2)\, \mathrm{e}^{\lambda x} d\nu_{a',b'} = \text{const.} \int_{v-u}^{v+u} \mathrm{e}^{\lambda x} \sqrt{\big[x - (v-u)\big]\big[(v+u) - x\big]}\, dx$$

yielding (26). Note in particular that when $a' = b' = 0$, $\Phi = J_1$ and $\lambda(\Psi - \Psi'') = \Phi'$ so that, by (10), $\Psi = \Phi = J_1$.

5. Moment equations and small deviation inequalities

In this section, we draw from the general framework developed in Section 3 moment recurrence formulas for the spectral measures of Hermite, Laguerre and Jacobi random matrix ensembles to obtain sharp small deviation bounds, for fixed N, on the largest eigenvalues. As presented in the introduction, in the three basic examples under study, the largest eigenvalue $\lambda_{\max}^N$ of the random matrices X^N, Y^N and Z^N is known to converge almost surely to the right endpoint of the support of the corresponding limiting spectral measure. Universality of the Tracy-Widom arises in the fluctuation of $\lambda_{\max}^N$ at the common rate $N^{2/3}$ and has been established in [TW] for the GUE (cf. also [De]), in [Joha1] and [John] for the LUE, and recently in [Co] for the JUE. We present here precise non asymptotic deviation inequalities at the rate $N^{2/3}$. Classical measure concentration arguments (cf. [G-Z]) do not apply to reach the rate $N^{2/3}$. We use to this task a simple induction argument on the moment recurrence relations of the orthogonal polynomial ensembles

As in the preceding section, we investigate successively the Hermite, Laguerre and Jacobi Unitary Ensembles.

a) GUE random matrices. In order to describe the moment relations for the measure $\frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$, where μ is the standard normal distribution and P_ℓ, $\ell \in \mathbb{N}$, are the (normalized) orthogonal Hermite polynomials for μ, the following alternate formulation of the classical Christoffel-Darboux (cf. [Sz]) formula will be convenient.

Lemma 5.1. *For any smooth function g on $\mathbb{R}$, and each $N \geq 1$,*

$$\int g' \sum_{\ell=0}^{N-1} P_\ell^2 d\mu = \int g \sqrt{N}\, P_N P_{N-1} d\mu.$$

Proof. Let $\mathcal{L}$ be the first order operator

$$\mathcal{L} f = f' - xf$$

acting on smooth functions f on the real line $\mathbb{R}$. The integration by parts formula indicates that for smooth functions f and g,

$$\int g(-\mathcal{L}f)d\mu = \int g' f d\mu.$$

The recurrence relation for the (normalized) Hermite polynomials P_N, $N \in \mathbb{N}$, is given by

$$xP_N = \sqrt{N+1}\, P_{N+1} + \sqrt{N}\, P_{N-1}.$$

We also have that $P_N' = \sqrt{N}\, P_{N-1}$. Hence,

$$\begin{aligned}\mathcal{L}(P_N^2) &= P_N\big[2P_N' - xP_N\big] \\ &= \sqrt{N}\, P_N P_{N-1} - \sqrt{N+1}\, P_{N+1}P_N.\end{aligned}$$

Therefore,

$$(-\mathcal{L})\bigg(\sum_{\ell=0}^{N-1} P_\ell^2\bigg) = \sqrt{N}\, P_N P_{N-1}$$

from which the conclusion follows. □

We now address recurrence equations for moments. Set $a_k = \int x^k P_N P_{N-1} d\mu$, $k \in \mathbb{N}$. By Corollary 3.2 with $\theta(x) = x^k$, $p = P_N$, $q = P_{N-1}$,

$$-k(k-1)(k-2)(k-3)a_{k-4} + k(k-1)[a_k - 4Na_{k-2}] + ka_k - a_k = 0,$$

that is

$$(k+1)a_k - 4Nka_{k-2} - k(k-2)(k-3)a_{k-4} = 0, \quad k \geq 4.$$

While $a_0 = a_2 = 0$, note that the proof of Proposition 3.1 cannot produce the value of a_1. To this task, simply use the recurrence relation of the Hermite polynomials to see that $a_1 = \sqrt{N}$. Then $a_3 = 3N^{3/2}$. Now, by Lemma 5.1,

$$k \int x^{k-1} \sum_{\ell=0}^{N-1} P_\ell^2 d\mu = \int x^k \sqrt{N}\, P_N P_{N-1} d\mu = \sqrt{N}\, a_k.$$

Set now

$$b_p = b_p^N = \int (\sigma x)^{2p} \frac{1}{N} \sum_{\ell=0}^{N-1} P_\ell^2 d\mu, \quad p \in \mathbb{N},$$

so that

$$a_{2p+1} = \text{const.}\, \sigma^{-2p}(2p+1)b_p.$$

(It is an easy matter to see that the odd moments of the measure $\sum_{\ell=0}^{N-1} P_\ell^2 d\mu$ are all zero.) Hence,

$$(2p+2)b_p - 4N\sigma^2(2p-1)b_{p-1} - \sigma^4(2p-1)(2p-2)(2p-3)b_{p-2} = 0.$$

In other words, for every integer $p \geq 2$,

$$b_p = 4N\sigma^2 \, \frac{2p-1}{2p+2} \, b_{p-1} + 4\sigma^4 p(p-1) \, \frac{2p-1}{2p+2} \cdot \frac{2p-3}{2p} \, b_{p-2} \tag{27}$$

($b_0 = 1$, $b_1 = \sigma^2 N^2$). This recurrence equation, reminiscent of the three step recurrence equation for orthogonal polynomials, was first put forward in an algebraic context by J. Harer and D. Zagier [H-Z]. It is also discussed in the book by M. L. Mehta [Me]. The proof above is essentially the one of [H-T].

It may be noticed that $b_p^N \to \chi_p$ as $\sigma^2 \sim \frac{1}{4N}$, $N \to \infty$, for every p, where χ_p is the $2p$-moment of the semicircle law. It may actually even be observed from (27) that when $\sigma^2 = \frac{1}{4N}$, for every fixed p and every $N \geq 1$,

$$\chi_p \leq b_p^N \leq \chi_p + \frac{C_p}{N^2}$$

where $C_p > 0$ only depends on p.

Let now $X = X^N$ be a GUE matrix with $\sigma^2 = \frac{1}{4N}$. As we have seen, the mean spectral measure $\widehat{\mu}_\sigma^N$ of X^N then converges weakly as $N \to \infty$ to the semicircle law on $(-1, +1)$. The largest eigenvalue $\lambda_{\max}^N$ of X^N is known (cf. [Bai], [H-T]) to converge almost surely to the right endpoint $+1$ of the support of the limiting semicircle law. (By symmetry, the smallest eigenvalue converges to -1.) Actually, the fact that $\limsup_{N\to\infty} \lambda_{\max}^N \leq 1$ almost surely will follow from the bounds we provide next and the Borel-Cantelli lemma. To prove the lower bound for the liminf requires a quite different set of arguments, namely the classical strengthening of the convergence of the mean spectral measure into almost sure convergence of the spectral measure $\frac{1}{N}\sum_{i=1}^N \lambda_i^N$. (This is shown in [H-T] via a measure concentration argument.) C. Tracy and H. Widom [T-W1] proved that (some multiple of) $N^{2/3}(\lambda_{\max}^N - 1)$ converges weakly to the distribution F of (6). We bound here, for fixed N, the probability that $\lambda_{\max}^N$ exceeds $1 + \varepsilon$ for small ε's in accordance with the rate $N^{2/3}$ and the asymptotics at infinity of the Airy function.

Fix thus $\sigma^2 = \frac{1}{4N}$. By (3), for every $\varepsilon > 0$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1 + \varepsilon\}\big) \leq N \widehat{\mu}_\sigma^N\big([1+\varepsilon, \infty)\big) = \int_{2\sqrt{N}(1+\varepsilon)}^{\infty} \sum_{k=0}^{N-1} P_k^2 \, d\mu \, .$$

Hence, for every $p \geq 0$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1 + \varepsilon\}\big) \leq (1+\varepsilon)^{-2p} N b_p$$

where we recall that $b_p = b_p^N$ are the $2p$-moments of the mean spectral measure $\widehat{\mu}_\sigma^N$ (that is of the trace of X^N). Now, by induction on the recurrence formula (27) for b_p, it follows that, for every $p \geq 0$,

$$b_p \leq \left(1 + \frac{p(p-1)}{4N^2}\right)^p \chi_p$$

where, by (12),

$$\chi_p = \frac{2p-1}{2p+2}\,\chi_{p-1} = \frac{(2p)!}{2^{2p}p!(p+1)!}$$

is the $2p$-moment of the semicircular law. By Stirling's formula, $\chi_p \sim p^{-3/2}$ as $p \to \infty$. Hence, for $0 < \varepsilon \le 1$ and some numerical constant $C > 0$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \ge 1+\varepsilon\}\big) \le CNp^{-3/2}\,\mathrm{e}^{-\varepsilon p + p^3/4N^2}.$$

Therefore, optimizing in $p \sim \sqrt{\varepsilon}\,N$, $0 < \varepsilon \le 1$, yields the following sharp tail inequality. By symmetry, a similar bound holds for the smallest eigenvalue.

Proposition 5.2. *In the preceding setting, for every $0 < \varepsilon \le 1$ and $N \ge 1$,*

$$\mathbb{P}\big(\{\lambda_{\max}^N \ge 1+\varepsilon\}\big) \le C\,\mathrm{e}^{-N\varepsilon^{3/2}/C}$$

where $C > 0$ is a numerical constant.

A slightly weaker inequality, involving some polynomial factor, is established in [Le1] using hypercontractivity of the Hermite operator. For an argument using integral bounds, see [Au].

b) LUE random matrices. The following is the analogue of Lemma 5.1 for Laguerre polynomials. It is again a version of the Christoffel-Darboux formula. Here $\mu = \mu^\gamma$ is the Gamma distribution with parameter γ and $P_\ell = P_\ell^\gamma$, $\ell \in \mathbb{N}$, are the (normalized) orthogonal Laguerre polynomials for μ.

Lemma 5.3. *For any smooth function g on $\mathbb{R}$, and each $N \ge 1$,*

$$\int xg' \sum_{\ell=0}^{N-1} P_\ell^2\, d\mu = -\int g\sqrt{N}\,\sqrt{\gamma+N}\,P_N P_{N-1}\, d\mu.$$

Proof. Let $\mathcal{L}$ be the first order operator

$$\mathcal{L}f = xf' + (\gamma + 1 - x)f,$$

$\gamma > -1$, acting on smooth functions f on $(0, +\infty)$. The integration by parts formula indicates that for smooth functions f and g,

$$\int g(-\mathcal{L}f)\,d\mu = \int xg' f\, d\mu.$$

The recurrence relation for the (normalized) Laguerre polynomials P_N, $N \in \mathbb{N}$, is given by

$$xP_N = -\sqrt{N+1}\,\sqrt{\gamma+N+1}\,P_{N+1} + (\gamma+2N+1)P_N - \sqrt{N}\,\sqrt{\gamma+N}\,P_{N-1}$$

We also have that $xP_N' = -\sqrt{N}\,\sqrt{\gamma+N}\,P_{N-1} + NP_N$. Hence,

$$\begin{aligned}\mathcal{L}(P_N^2) &= P_N\big[2xP_N' + (\gamma+1-x)P_N\big]\\ &= \sqrt{N+1}\,\sqrt{\gamma+N+1}\,P_{N+1}P_N - \sqrt{N}\,\sqrt{\gamma+N}\,P_NP_{N-1}.\end{aligned}$$

Therefore,

$$(-\mathcal{L})\bigg(\sum_{\ell=0}^{N-1} P_\ell^2\bigg) = -\sqrt{N}\,\sqrt{\gamma+N}\,P_NP_{N-1}$$

from which the conclusion follows. □

Together with this lemma, we consider as in the Hermite case, recurrence equations for moments. Set $a_k = \int x^k P_N P_{N-1} d\mu$, $k \in \mathbb{Z}$, $k > -\gamma - 1$. By Corollary 3.2 with $\theta(x) = x^k$, $p = P_N$, $q = P_{N-1}$,

$$\begin{aligned}&- k(k-1)(k-2)(k-3)a_{k-2} - 3k(k-1)(k-2)a_{k-2}\\ &+ k(k-1)\big[a_k - 2(\gamma+2N)a_{k-1} + (\gamma^2-1)a_{k-2}\big]\\ &+ k\big[a_k - (\gamma+2N)a_{k-1}\big] - a_k = 0,\end{aligned}$$

that is

$$(k+1)(k-1)a_k - (\gamma+2N)k(2k-1)a_{k-1} + k(k-1)\big[\gamma^2 - k(k-2) - 1\big]a_{k-2} = 0$$

for each $k-2 > \gamma-1$. Note that $a_0 = 0$ and, by the recurrence relation of the Laguerre polynomials, $a_1 = -\sqrt{N}\,\sqrt{\gamma+N}$. As we have seen in Lemma 5.3,

$$k\int x^k \sum_{\ell=0}^{N-1} P_\ell^2 d\mu = -\int x^k \sqrt{N}\,\sqrt{\gamma+N}\,P_NP_{N-1}d\mu = -\sqrt{N}\,\sqrt{\gamma+N}\,a_k.$$

Set now

$$b_p = b_p^N = \int (\sigma^2 x)^p \frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu, \quad p \in \mathbb{N},$$

so that

$$a_p = \text{const.}\,\sigma^{-2p} p b_p.$$

Hence

$$(p+1)b_p - (\gamma+2N)\sigma^2(2p-1)b_{p-1} + \big[\gamma^2 - p(p-2) - 1\big]\sigma^4(p-2)b_{p-2} = 0.$$

In other words, for $p-2 > -\gamma-1$, $p \neq -1$,

$$b_p = 2(\gamma+2N)\sigma^2\,\frac{2p-1}{2p+2}\,b_{p-1} - \big[\gamma^2 - (p-1)^2\big]\sigma^4\,\frac{p-2}{p+1}\,b_{p-2} \tag{28}$$

($b_0 = 1$, $b_1 = (\gamma + N)\sigma^2$). As for (27), this recurrence relation already appeared, with the same proof, in the paper [H-T] by U. Haagerup and S. Thorbjørnsen.

Note that, for every fixed p, the moment equations (28) converge as $\sigma^2 \sim \frac{1}{4N}$, $\gamma = \gamma_N \sim c'N$, $N \to \infty$, $c' \geq 0$, to the moment equations

$$\chi_p^{c'} = \frac{c'+2}{2} \cdot \frac{2p-1}{2p+2}\, \chi_{p-1}^{c'} - \frac{c'^2}{16} \cdot \frac{p-2}{p+1}\, \chi_{p-2}^{c'}$$

($\chi_0^{c'} = 1$, $\chi_1^{c'} = \frac{c+1}{4}$). The moments $\chi_p^{c'}$ may be recognized as the moments of the Marchenko-Pastur distribution $\nu_{c'}$ of (24). It may actually even be observed from (28) that when $\sigma^2 = \frac{1}{4N}$ and $\gamma = \gamma_N = c'N$, for every fixed p and every $N \geq 1$,

$$|b_p^N - \chi_p^{c'}| \leq \frac{C_p}{N^2}$$

where $C_p > 0$ only depends on p and c' (use that $b_0^N = \chi_0^{c'}$ and $b_1^N = \chi_1^{c'}$).

We turn to the Laguerre ensemble, and consider now the Wishart matrix $Y = Y^N$ from the LUE with $\sigma^2 = \frac{1}{4N}$ and $M = M(N) = [cN]$, $c \geq 1$. As we have seen, the mean spectral measure $\widehat{\mu}_\sigma^N$ of Y^N then converges weakly as $N \to \infty$ to the Marchenko-Pastur distribution $\nu_{c'}$ of (24), $c' = c - 1 \geq 0$, with support $(v-u, v+u)$ where $u > 0$ and v are given by (23). The largest eigenvalue $\lambda_{\max}^N$ of Y^N is known to converge almost surely to the right endpoint $v+u$ of the support of ν_c while the smallest eigenvalue $\lambda_{\min}^N$ converges almost surely towards $v-u$ (cf. [Bai]). As for the GUE, the fact that $\limsup_{N\to\infty} \lambda_{\max}^N \leq 1$ almost surely will follow from the bounds we provide next and the Borel-Cantelli lemma. The lower bound on the liminf follows from the almost sure convergence of the spectral measure. K. Johansson [Joha1] and I. Johnstone [John] independently proved that (some multiple of) $N^{2/3}(\lambda_{\max}^N - (v+u))$ converges weakly to the Tracy-Widom distribution F. We bound here, for fixed N, the probability that $\lambda_{\max}^N$ exceeds $v + u + \varepsilon$ for small ε's at the appropriate rate and dependence in ε.

Fix thus $\sigma^2 = \frac{1}{4N}$ and $\gamma = [cN] - N$, $c \geq 1$. For simplicity, we assume throughout the argument below that $\gamma = c'N$, $c' = c-1 \geq 0$, the general case being easily handled along the same line. By (4), we have that

$$\mathbb{P}\big(\big\{\lambda_{\max}^N \geq (v+u)(1+\varepsilon)\big\}\big) \leq (u+v)^{-p}(1+\varepsilon)^{-p} N b_p$$

for every $\varepsilon > 0$ and $p \geq 0$, where $b_p = b_p^N$ is the p-th moment of the mean spectral measure $\widehat{\mu}_\sigma^N$ (that is of the trace of Y^N) and $u > 0$ and v are given by (23). The recurrence equation (28) for b_p indicates that

$$b_p = 2v\, \frac{2p-1}{2p+2}\, b_{p-1} - \left((v^2 - u^2) - \frac{(p-1)^2}{16N^2}\right) \frac{p-2}{p+1}\, b_{p-2}.$$

Two cases have to be distinguished. If $u = v = \frac{1}{2}$ ($c' = 0$), the argument leading to Proposition 5.2 may essentially be repeated. Assume thus in the following that

$v > u > 0$. We show by recurrence that for some constant $C > 0$ only depending on $c \geq 1$ and possibly varying from line to line below, for every $1 \leq p \leq N/C$,

$$b_p \leq (v+u)\left(1 + \frac{C}{p} + \frac{Cp^2}{N^2}\right)\frac{2p-1}{2p+2}\, b_{p-1}. \tag{29}$$

To this task, it suffices to check that

$$\begin{aligned} 2v\,\frac{2p-1}{2p+2}\, b_{p-1} - \left((v^2-u^2) - \frac{(p-1)^2}{16N^2}\right)\frac{p-2}{p+1}\, b_{p-2} & \\ \leq (v+u)\left(1 + + \frac{C}{p} + \frac{Cp^2}{N^2}\right)\frac{2p-1}{2p+2}\, b_{p-1} & \end{aligned}$$

holds as a consequence of the recurrence hypothesis (29) for $p-1$. Hence, by iteration, for every $p \leq N/C$,

$$b_p \leq (v+u)^p \left(1 + \frac{C}{p} + \frac{Cp^2}{N^2}\right)^p \chi_p.$$

We thus conclude as for the proof of Proposition 5.2 to the next sharp small deviation inequality.

Proposition 5.4. *In the preceding setting, for every $0 < \varepsilon \leq 1$ and $N \geq 1$,*

$$\mathbb{P}\big(\big\{\lambda_{\max}^N \geq (v+u)(1+\varepsilon)\big\}\big) \leq C\,\mathrm{e}^{-N\varepsilon^{3/2}/C}.$$

where $C > 0$ only depends on c.

For $c = 1$ ($c' = 0$), a slightly weaker inequality, involving some polynomial factor, is established in [Le1] using hypercontractivity of the Laguerre operator.

Interestingly enough, the same argument may be developed to establish the same exponential bound on the smallest eigenvalue $\lambda_{\min}^N$ of $Y = Y^N$ in case $v > u > 0$, that is the so-called "soft wall" in the physicist language as opposed to the "hard wall" 0 (since all the eigenvalues are non-negative). Fluctuation results for $\lambda_{\min}^N$ at the soft edge have been obtained in [B-F-P] (see also [Fo]) in a physicist language, with again the limiting Tracy-Widom law (for the hard wall with a Bessel kernel, cf. [T-W2], [Fo]). Assume thus that $v > u > 0$. For any $0 < \varepsilon < 1$,

$$\begin{aligned} \mathbb{P}\big(\big\{\lambda_{\min}^N \leq (v-u)(1-\varepsilon)\big\}\big) &\leq N\widehat{\mu}_\sigma^N\big([0,(v-u)(1-\varepsilon)]\big) \\ &= \int_0^{4N(v-u)(1-\varepsilon)} \sum_{\ell=0}^{N-1} P_\ell^2\, d\mu\,. \end{aligned}$$

Hence, for every $p \geq 0$,

$$\mathbb{P}\big(\big\{\lambda_{\min}^N \leq (v-u)(1-\varepsilon)\big\}\big) \leq (v-u)^p(1-\varepsilon)^p N b_{-p}\,.$$

Setting $\tilde{b}_p = b_{-p}$, the recurrence equations (28) (with $\gamma = c'N$, $c' = c-1 > 0$), amount to

$$\left((v^2 - u^2) - \frac{(p-1)^2}{16N^2}\right)\frac{p}{p-3}\,\tilde{b}_p = 2v\,\frac{2p-3}{2p-6}\,\tilde{b}_{p-1} - \tilde{b}_{p-2}$$

for every $(4 \leq)\ p < c'N + 1$. Arguing as for the largest eigenvalue shows that for some constant $C > 0$ only depending on c,

$$\tilde{b}_p \leq \frac{1}{v-u}\left(1 + \frac{C}{p} + \frac{Cp^2}{N^2}\right)\frac{2p-3}{2p}\,\tilde{b}_{p-1}.$$

By iteration,

$$\tilde{b}_p \leq (v-u)^{-p}\left(1 + \frac{C}{p} + \frac{Cp^2}{N^2}\right)^p \chi_p$$

for $p \leq N/C$. As for Propositions 5.2 and 5.4, we thus conclude to the following.

Proposition 5.5. *In the preceding setting, for every $0 < \varepsilon \leq 1$ and $N \geq 1$,*

$$\mathbb{P}\big(\big\{\lambda_{\min}^N \leq (v-u)(1-\varepsilon)\big\}\big) \leq C\,\mathrm{e}^{-N\varepsilon^{3/2}/C}.$$

where $C > 0$ only depends on c.

c) JUE random matrices. The following lemma is the analogue of Lemmas 5.1 and 5.3, and is one further instance of the Christoffel-Darboux formula. Here $\mu = \mu^{\alpha,\beta}$ is the Beta distribution with parameter α, β and $P_\ell = P_\ell^{\alpha,\beta}$, $\ell \in \mathbb{N}$, are the (normalized) orthogonal polynomials for μ.

Lemma 5.6. *For any smooth function g on $\mathbb{R}$, and each $N \geq 1$,*

$$\int (1-x^2)g' \sum_{\ell=0}^{N-1} P_\ell^2\, d\mu = \int g e_N P_N P_{N-1}\, d\mu$$

where e_N is a constant.

Proof. Let $\mathcal{L}$ be the first order operator

$$\mathcal{L}f = (1-x^2)f' + \big[\beta - \alpha - (\alpha+\beta+2)x\big]f,$$

$\alpha, \beta > -1$, acting on smooth functions f on $(-1,+1)$. The integration by parts formula indicates that for smooth functions f and g,

$$\int g(-\mathcal{L}f)\,d\mu = \int (1-x^2)g' f\,d\mu.$$

The recurrence relation for the (normalized) Jacobi polynomials P_N, $N \in \mathbb{N}$, is given by

$$\begin{aligned}(\alpha+\beta+2N)(\alpha{+}\beta+2N+1)&(\alpha+\beta+2N+2)xP_N\\ &= 2(N+1)(\alpha+\beta+N+1)(\alpha+\beta+2N)\kappa_n^{-1}\kappa_{N+1}P_{N+1}\\ &\quad+(\alpha+\beta+2N+1)(\beta^2-\alpha^2)P_N\\ &\quad+2(\alpha+N)(\beta+N)(\alpha+\beta+2N+2)\kappa_N^{-1}\kappa_{N-1}P_{N-1}.\end{aligned}$$

We also have that

$$\begin{aligned}(\alpha+\beta+2N)(1-x^2)P_N' &= -N\big[\beta-\alpha+(\alpha+\beta+2N)x\big]P_N\\ &\quad+2(\alpha+N)(\beta+N)\kappa_{N-1}\kappa_N^{-1}P_{N-1}.\end{aligned}$$

Hence,

$$\begin{aligned}\mathcal{L}(P_N^2) &= P_N\Big(2(1-x^2)P_N'+\big[\beta-\alpha-(\alpha+\beta+2)x\big]P_N\Big)\\ &= P_N\bigg(\frac{4(\alpha+N)(\beta+N)}{\alpha+\beta+2N}\,\kappa_{N-1}\kappa_N^{-1}P_{N-1}\\ &\qquad-\frac{(\alpha+\beta+2N+2)(\alpha+\beta+2N)x+\alpha^2-\beta^2}{\alpha+\beta+2N}\,P_N\bigg)\\ &= 2P_N\bigg(\frac{(\alpha+N)(\beta+N)}{\alpha+\beta+2N+1}\,\kappa_{N-1}\kappa_N^{-1}P_{N-1}\\ &\qquad-\frac{(N+1)(\alpha+\beta+N+1)}{\alpha+\beta+2N+1}\,\kappa_N^{-1}\kappa_{N+1}P_{N+1}\bigg).\end{aligned}$$

By definition of κ_N, it follows that

$$\mathcal{L}(P_N^2) = e_NP_NP_{N-1}-e_{N+1}P_{N+1}P_N$$

so that

$$(-\mathcal{L})\bigg(\sum_{\ell=0}^{N-1}P_\ell^2\bigg) = e_NP_NP_{N-1}.$$

The conclusion follows. □

Turning to moment equations, set $a_k=\int x^kP_NP_{N-1}d\mu$, $k\in\mathbb{N}$. By Corollary 3.2 with $\theta(x)=x^k$, $p=P_N$, $q=P_{N-1}$,

$$\begin{aligned}&\quad k(k-1)(k-2)(k-3)(-a_k+2a_{k-2}-a_{k-4})\\ &+k(k-1)(k-2)(-6a_k+6a_{k-2})\\ &+k(k-1)\Big(\big[(\alpha+\beta)(\alpha+\beta+2)+2(\rho_N+\rho_{N-1})-6\big]a_k+2(\alpha^2-\beta^2)a_{k-1}\\ &\quad+\big[(\alpha-\beta)^2-2(\alpha+\beta)-2(\rho_N+\rho_{N-1})+2\big]a_{k-2}\\ &+k\Big(\big[(\alpha+\beta)(\alpha+\beta+2)+2(\rho_N+\rho_{N-1})\big]a_k+(\alpha^2-\beta^2)a_{k-1}\Big)\\ &-(\alpha+\beta+2N)^2a_k=0,\end{aligned}$$

that is

$$
\begin{aligned}
&\big[(k^2-1)(\alpha+\beta+2N)^2 - k(k-1)(k-2)(k+3) - 6k(k-1)\big]a_k \\
&- k(2k-1)(\beta^2-\alpha^2)a_{k-1} \\
&- k(k-1)\big[4N(\alpha+\beta+N) - (\alpha-\beta)^2 - 2k(k-2) - 2\big]a_{k-2} \\
&- k(k-1)(k-2)(k-3)a_{k-4} = 0
\end{aligned}
\tag{30}
$$

for $k \geq 4$. Here $a_0 = 0$, $a_1 = \frac{e_N}{\alpha+\beta+2N}$ by the recurrence relation of the Jacobi polynomials. As we have seen in Lemma 5.6,

$$
k\int (x^{k-1} - x^{k+1}) \sum_{\ell=0}^{N-1} P_\ell^2 d\mu = \int x^k e_N P_N P_{N-1} d\mu = e_N a_k.
$$

Set

$$
b_k = b_k^N = \int x^k \frac{1}{N} \sum_{\ell=0}^{N-1} P_\ell^2 d\mu, \quad k \in \mathbb{N},
$$

with $b_0 = 1$ and

$$
b_1 = \frac{\beta^2-\alpha^2}{N} \sum_{\ell=0}^{N-1} (\alpha+\beta+2\ell+2)^{-1}(\alpha+\beta+2\ell)^{-1}.
$$

Therefore,

$$
a_k = \text{const.}\, k(b_{k-1} - b_{k+1}), \quad k \geq 1.
$$

Hence, by (30), for every $p \geq 2$,

$$
R(b_{2p} - b_{2p+2}) = S(b_{2p} - b_{2p-2}) + T(b_{2p-2} - b_{2p-4}) + U(b_{2p-1} - b_{2p+1})
$$

where

$$
\begin{aligned}
R &= (2p+2)(\alpha+\beta+2N)^2 - (2p+1)(2p-1)(2p+4) - 6(2p+1), \\
S &= (2p-1)\big[4N(\alpha+\beta+N) - (\alpha-\beta)^2 - 2(2p+1)(2p-1) - 2\big], \\
T &= (2p-1)(2p-2)(2p-3), \\
U &= (4p+1)(\beta^2-\alpha^2).
\end{aligned}
$$

In other words, provided that $R \neq 0$,

$$
b_{2p} - b_{2p+2} = \frac{S}{R}(b_{2p} - b_{2p-2}) + \frac{T}{R}(b_{2p-2} - b_{2p-4}) + \frac{U}{R}(b_{2p-1} - b_{2p+1}). \tag{31}
$$

Note that the moment equations (31) converge, as $\alpha = \alpha_N \sim a'N$, $\beta = \beta_N \sim b'N$, $N \to \infty$, $a', b' \geq 0$, to the moment equations

$$
\chi_{2p}^{a',b'} - \chi_{2p+2}^{a',b'} = (u^2 - v^2)\,\frac{2p-1}{2p+2}\,\big(\chi_{2p-2}^{a',b'} - \chi_{2p}^{a',b'}\big) + v\,\frac{4p+1}{2p+2}\,\big(\chi_{2p-1}^{a',b'} - \chi_{2p+1}^{a',b'}\big).
$$

The moments $\chi_{2p}^{a',b'}$ may be recognized to be the moments of the distribution $\nu_{a',b'}$ given in (26).

Let finally $Z = Z^N$ a random matrix from the JUE with $M_1 = M_1(N) = [aN]$, $a \geq 1$, and $M_2 = M_2(N) = [bN]$, $b \geq 1$. As we have seen, the mean spectral measure $\widehat{\mu}_\sigma^N$ of Z^N converges weakly as $N \to \infty$ to the distribution $\nu_{a',b'}$ of (26), $a' = a-1 \geq 0$, $b' = b - 1 \geq 0$, with support $(v-u, v+u)$ where $u > 0$ and v are given by (25). The behavior of the largest and smallest eigenvalues of Z^N has been settled recently in [Co] (see also [Fo], [T-W2] and the references therein). We prove here some sharp tail inequalities similar to the ones for the GUE and LUE at the rate $N^{2/3}$ at the soft edges. We actually only deal with the symmetric case $a = b$. Therefore $v = 0$ in (25). All the eigenvalues of Z^N take values between -1 and $+1$. We assume that $u^2 < 1$ and concentrate thus on a "soft wall" analysis.

As in the Laguerre case, fix thus for simplicity $\alpha = a'N$, $\beta = b'N$, $a' = b' \neq 0$, so that $u^2 < 1$. By (5),

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq u(1+\varepsilon)\}\big) \leq u^{-2p}(1+\varepsilon)^{-2p} N b_{2p}$$

for every $\varepsilon > 0$ and $p \geq 0$ where $b_{2p} = b_{2p}^N$ are the $2p$-moments of the mean spectral measure $\widehat{\mu}^N$ (that is of the trace of Z^N). Set $c_p = b_{2p} - b_{2p+2} \geq 0$. Since $\alpha = \beta$, the recurrence equations (31) for $b_{2p} = b_{2p}^N$ indicate that

$$c_p = \frac{S}{R}\, c_{p-1} + \frac{T}{R}\, c_{p-2} \tag{32}$$

where

$$\begin{aligned}
R &= (2p+2)4N^2(a'+1)^2 - (2p+1)(2p-1)(2p+4) - 6(2p+1), \\
S &= (2p-1)\big[4N^2(2a'+1) - 2(2p+1)(2p-1) - 2\big], \\
T &= (2p-1)(2p-2)(2p-3).
\end{aligned}$$

¿From (32), there exists a constant $D > 0$, depending on u, such that for every $(2 \leq)$ $p \leq N/D$,

$$c_p \leq u^2\, \frac{2p-1}{2p+2}\left(1 + \frac{Dp^2}{N^2}\right) c_{p-1} + \frac{Dp^2}{N^2}\, c_{p-2}.$$

Arguing by induction, it follows that there exists a constant $C > 0$ only depending on u, and possibly changing from line to line below, such that for every $p \leq N/C$,

$$c_p \leq u^{2p}\left(1 + \frac{Cp^2}{N^2}\right)^p \chi_p$$

where we recall that χ_p is the $2p$-moment of the semicircle law. Now, $c_p = b_{2p} - b_{2p+2}$ and $b_{2p} \to 0$ as $p \to \infty$ (for every fixed N). Therefore, for every $p \leq N/C$,

$$\begin{aligned}
b_{2p} &\leq \sum_{\ell=p}^{\infty} u^{2\ell}\left(1 + \frac{Cp^2}{N^2}\right)^\ell \chi_\ell \\
&= \sum_{\ell=p}^{\infty} \int \left(1 + \frac{Cp^2}{N^2}\right)^\ell (u^2x^2)^\ell\, d\mu_3(x).
\end{aligned}$$

As a consequence, provided that $u^2(1+\frac{Cp^2}{N^2}) \leq d < 1$,

$$b_{2p} \leq \frac{1}{1-d}\left(1+\frac{Cp^2}{N^2}\right)^p \chi_p.$$

We may then conclude, as in the Hermite and Laguerre cases, to the following result. By symmetry, a similar bound holds for the smallest eigenvalue.

Proposition 5.7. *In the preceding setting, for every $0 < \varepsilon \leq 1$ and $N \geq 1$,*

$$\mathbb{P}\big(\big\{\lambda_{\max}^N \geq u(1+\varepsilon)\big\}\big) \leq C\,\mathrm{e}^{-N\varepsilon^{3/2}/C}.$$

where $C > 0$ only depends on u.

The hypercontractive method of [Le1] may be used similarly on the Jacobi operator to prove a slightly weaker inequality, involving some polynomial factor.

As a consequence of Proposition 5.7 and the Borel-Cantelli lemma, it immediately follows that $\limsup_{N\to\infty} \lambda_{\max}^N \leq u$ almost surely. Almost sure convergence of the spectral measure is established in [Co].

The arguments developed here do not seem to yield at this point bounds at fixed N on the largest eigenvalues of the GUE, LUE or (symmetric) JUE under their limiting values (and similarly for the smallest eigenvalues). This would be of particular interest in case of the "hard walls" $v - u = 0$ for the LUE and $v \pm u = \pm 1$ for the JUE where the Airy kernel is replaced by the Bessel kernel [T-W2] at the rate $N^{1/2}$.

Acknowledgement. Thanks are due to T. Kurtz, and O. Zeitouni, for crucial comments leading to the description of equilibrium measures as mixtures of arcsine and uniform distributions. Thanks also to D. Bakry for stimulating conversations at the early stage of this work, and to Y. Doumerc for critical comments.

References

[Au] G. Aubrun. An inequality about the largest eigenvalue of a random matrix (2003). To appear in Séminaire de Probabilités XXXVIII. Lecture Notes in Math., Springer.

[Bai] Z. D. Bai. Methodologies in the spectral analysis of large dimensional random matrices: a review. Stat. Sinica 9, 611–661 (1999).

[B-F-P] T. Baker, P. J. Forrester, P. Pearce. Random matrix ensembles with an effective external charge. J. Phys. A 31, 6087–6101 (1998).

[Bak] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. Ecole d'Eté de Probabilités de St-Flour. Lecture Notes in Math. 1581, 1–114 (1994). Springer.

[BA-D-G] G. Ben Arous, A. Dembo, A. Guionnet. Aging of spherical spin glasses. Probab. Theor. Relat. Fields 120, 1–67 (2001).

[C-C] M. Capitaine, M. Casalis. Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to Beta random matrices (2002). To appear in Indiana J. Math.

[C-I] L.-C. Chen, M. Ismail. On asymptotics of Jacobi polynomials. SIAM J. Math. Anal. 22, 1442–1449 (1991).

[Ch] T. S. Chihara. An introduction to orthogonal polynomials. Gordon & Breach (1978).

[Co] B. Collins. Product of random projections, Jacobi ensembles and universality problems arising from free probability (2003).

[De] P. A. Deift. Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. CIMS Lecture Notes 3. Courant Institute of Mathematical Sciences (1999).

[Fo] P. J. Forrester. Log-gases and random matrices (2002).

[F-O-T] M. Fukushima, Y. Oshima, M. Takeda. Dirichlet forms and symmetric Markov processes. De Gruyter (1994).

[G-S] U. Grenander, J. W. Silverstein. Spectral analysis of networks with random topologies. SIAM J. Applied Math. 32, 499–519 (1977).

[G-Z] A. Guionnet, O. Zeitouni. Concentration of the spectral measure for large matrices. Elect. Comm. Probab. 5, 119-136 (2000).

[H-T] U. Haagerup, S. Thorbjørnsen. Random matrices with complex Gaussian entries (1997).

[H-Z] J. Harer, D. Zagier. The Euler characteristic of the moduli space of curves. Invent. math. 85, 457–485 (1986).

[K-S] R. Koekoek, R. F. Swarttouw. The Askey scheme of hypergeometric orthogonal polynomials and its q-analogue (1998). http://aw.twi.tudelft.nl/~koekoek/askey/index.html.

[K-VA] A. Kuijlaars, W. Van Assche. The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients. J. Approx. Theor. 99, 167–197 (1999).

[Joha1] K. Johansson. Shape fluctuations and random matrices. Comm. Math. Phys. 209, 437–476 (2000).

[Joha2] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. Ann. Math. 259–296 (2001).

[John] I. M. Johnstone. On the distribution of the largest principal component. Ann. Statist. 29, 295–327 (2001).

[Jon] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. J. Multivariate Anal. 12, 1–38 (1982).

[Le1] M. Ledoux. A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices. Séminaire de Probabilités XXXVII. Lecture Notes in Math. 1832, 360–369 (2003). Springer.

[Le2] M. Ledoux. Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The discrete case (2003).

[M-P] V. Marchenko, L. Pastur. The distribution of eigenvalues in certain sets of random matrices. Math. Sb. 72, 507–536 (1967).

[M-N-T] A. Máté, P. Nevai, V. Totik. Strong and weak convergence of orthogonal polynomials. Amer. J. Math. 109, 239–282 (1987).

[Ma] O. Mazet. Classification des semi-groupes de diffusion sur $\mathbb{R}$ associés à une famille de polynômes orthogonaux. Séminaire de Probabilités XXXI. Lecture Notes in Math. 1655, 40–53 (1997). Springer.

[Me] M. L. Mehta. Random matrices. Academic Press (1991).

[Sz] G. Szegö. Orthogonal polynomials. Colloquium Publications XXIII. Amer. Math Soc. (1975).

[T-W1] C. Tracy, H. Widom. Level-spacing distribution and the Airy kernel. Comm. Math. Phys. 159, 151–174 (1994).

[T-W2] C. Tracy, H. Widom. Level-spacing distribution and the Bessel kernel. Comm. Math. Phys. 161, 289–309 (1994).

[Vo] D. Voiculescu. A strengthened asymptotic freeness result for random matrices with applications to free entropy. Int. Math. Res. Notices 1, 41–63 (1998).

[Wa] K. W. Wachter. The strong limits of random matrix spectra for sample matrices of independent elements. Ann. Probability 6, 1–18 (1978).

[Wi] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. 62, 548–564 (1955).