Invariant β-Wishart ensembles, crossover densities and asymptotic corrections to the Marčenko-Pastur law

Romain Allez[1,2], Jean-Philippe Bouchaud[2], Satya N. Majumdar[3], and Pierpaolo Vivo[3]

[1]Université Paris-Dauphine, Laboratoire CEREMADE, Place du Marechal de Lattre de Tassigny, 75775 Paris Cedex 16 - France.
[2]Capital Fund Management, 6–8 boulevard Haussmann, 75 009 Paris, France.
[3]Laboratoire de Physique Théorique et Modèles Statistiques (UMR 8626 du CNRS), Université Paris-Sud, Bâtiment 100, 91405 Orsay Cedex, France.

September 28, 2012

Abstract

We construct a diffusive matrix model for the β-Wishart (or Laguerre) ensemble for general continuous $\beta \in [0, 2]$, which preserves invariance under the orthogonal/unitary group transformation. Scaling the Dyson index β with the largest size M of the data matrix as $\beta = 2c/M$ (with c a fixed positive constant), we obtain a family of spectral densities parametrized by c. As c is varied, this density interpolates continuously between the Marčenko-Pastur ($c \to \infty$ limit) and the Gamma law ($c \to 0$ limit). Analyzing the full Stieltjes transform (resolvent) equation, we obtain as a byproduct the correction to the Marčenko-Pastur density in the bulk up to order $1/M$ for all β and up to order $1/M^2$ for the particular cases $\beta = 1, 2$.

1 Introduction

The theory of matrices with random entries, originally devised as a tool to understand and predict the spectra of heavy nuclei for which a detailed account of the interactions between particles is too complicated, has seen a spectacular resurgence of interest in recent years, with a number of unexpected and often surprising applications (see [1, 2, 3, 4, 5] for a recent overview). While Wigner and Dyson are usually regarded as the pioneers in the field, John Wishart had already introduced random matrices in 1928 in his studies of multivariate populations [6]. The Wigner-Dyson (Gaussian) and Wishart ensembles (together with a few others) lie at the core of the *classical* world of invariant matrices, characterized by the following main features:

1. The joint probability distribution (jpd) of matrix entries, collectively denoted by $P[\mathbf{X}]$, remains unaltered if one performs a similarity transformation $\mathbf{X} \to \mathbf{UXU}^{-1}$, with $\mathbf{U}$ being orthogonal (real symmetric $\mathbf{X}$), unitary (complex hermitian $\mathbf{X}$) or symplectic (quaternion self-dual $\mathbf{X}$) matrix. As a consequence, the eigenvectors of such matrices are Haar (uniform) distributed in their respective groups.

2. The joint distribution of the N real eigenvalues $P(\lambda_1, \ldots, \lambda_N)$ can be generically written in the Gibbs-Boltzmann form,

$$P(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} \exp\left(-\mathcal{H}(\lambda_1, \ldots, \lambda_N)\right) \tag{1.1}$$

with the Hamiltonian $\mathcal{H}(\lambda_1, \ldots, \lambda_N)$ given by:

$$\mathcal{H}(\lambda_1, \ldots, \lambda_N) = \sum_{i=1}^{N} V(\lambda_i) - \beta \sum_{j<k} \ln |\lambda_j - \lambda_k| \tag{1.2}$$

where $V(x)$ a confining potential that depends on the precise form of the joint distribution of matrix entries $P[\mathbf{X}]$. For example, if the entries of $\mathbf{X}$ are independent, the only allowed potential is quadratic $V(x) = \beta x^2/2$, which correspond to the Gaussian ensembles. If correlations among the entries are allowed, then different potentials (all corresponding to rotationally invariant weights) are possible. For example, in the Wishart case, $V(x) = \infty$ for $x < 0$ (so that all the eigenvalues are non-negative) and $V(x) = x/(2\sigma^2) - \alpha \log x$ for $x \geq 0$.

The normalization constant Z_N is called the partition function and is simply given by the multiple integral

$$Z_N = \int \cdots \int \prod_i d\lambda_i \exp\left(-\mathcal{H}(\lambda_1, \ldots, \lambda_N)\right). \tag{1.3}$$

From (1.1), one easily deduces that the system of N eigenvalues of a classically invariant ensemble behaves as a thermodynamic system of charged particles arranged on the real line, in equilibrium at inverse temperature β under competing interactions (the confining potential $V(x)$ and the logarithmic all-to-all repulsion term) in (1.2). In contrast with the usual canonical ensemble in statistical mechanics, however, the so-called Dyson index β is quantized and can only assume the values $\beta = 1, 2, 4$ for real symmetric, complex hermitian and quaternion self-dual matrices respectively.

Lifting the quantization of β (Dyson's *threefold way*) has been a major theoretical challenge in view of possible applications e.g. to the quantum Hall effect. Dumitriu and Edelman [7] were eventually able to construct ensembles of tridiagonal matrices with independent entries whose eigenvalues are distributed as (1.1) with general $\beta > 0$. Their ensemble is however *not* invariant under similarity transformations, and the eigenvectors are not Haar distributed in the appropriate symmetry group. After an earlier attempt in the case of 2×2 matrices [8], the explicit construction of an ensemble of $N \times N$ matrices displaying at once rotational invariance and a continuous β was put forward in [9, 10] for the Gaussian ensemble. It was further shown in [9] that only by letting the Dyson index β of that ensemble scale with the matrix size N in an appropriate way (namely $\beta = c/N$) one obtains a continuous family of deformed spectral densities parametrized by c, interpolating between Wigner's semicircle (typical for $\beta \sim \mathcal{O}(1)$ invariant ensembles) and a Gaussian law (properly describing the non-interacting limit $\beta \to 0$). This result can be established in two alternative ways:

1. Starting from the stationary joint distribution of eigenvalues (eq. (1.1)), setting $\beta = c/N$, and then finding the average density of eigenvalues $\rho(\lambda) = (1/N)\langle \sum_i \delta(\lambda - \lambda_i) \rangle$. In the limit of large N, this average density can be obtained by a saddle point analysis of the partition function Eq. (1.3) in a standard way. Usually, when $\beta \sim \mathcal{O}(1)$, only the energy term $\sim \mathcal{O}(N^2)$ dominates and the entropy term $\sim \mathcal{O}(N)$ is subleading. However, when $\beta \sim c/N$, *both* the energy and the entropy terms are of the same order ($\sim \mathcal{O}(N)$), which leads to a nontrivial modification of the density. The complicated nonlinear singular integro-differential equation for the saddle density, reduces very nicely to a Riccati equation for the Stieltjes transform of the density, which can then subsequently solved exactly. Finally one obtains the density by taking imaginary part of the Stieltjes transform (see section 3.3 for a detailed discussion).

2. Starting from the dynamical equation of motion of the eigenvalues, one first derives directly the equation of motion of the Stieltjes transform of the density via Itô's calculus, finds the stationary solution and then obtains the average density by taking the imaginary part of the Stieltjes transform (see detailed discussion in section 3.4).

   In section 3.3 and 3.4, in the context of our model, we show that both methods lead to the same solution.

The purpose of this paper is threefold. We first explicitly construct a random matrix model ι) which is invariant under similarity transformations (and thus has Haar distributed eigenvectors), and $\iota\iota$) whose jpd of eigenvalues is exactly given by the β-Wishart ensemble of random matrices with a continuous $\beta > 0$. Then, letting the Dyson index β of the ensemble scale inversely with the size of the matrix, we analytically derive the density of states for this crossover model, written in terms of the Whittaker hypergeometric function (see eq. (3.49)), and we show that it continuously interpolates between the Marčenko-Pastur law and a certain type of Gamma distribution (see subsection 3.3). Some other types of deformations of the Marčenko-Pastur distribution for Wishart-like matrix models were reported in the literature (see e.g. [11, 12]). Finally, keeping the Dyson index β unscaled (i.e. it remains of $\sim \mathcal{O}(1)$ for large matrix size N) but nevertheless continuous ($\beta > 0$), we analyze the full Stieltjes transform equation and we can compute the $1/N$ correction to the Marčenko-Pastur asymptotic density for the β-Wishart ensemble for all values of $\beta > 0$. Furthermore, using results obtained by Pastur and Lytova in [13] on the noise in the Marčenko-Pastur law, we are also able to derive the $1/N^2$ correction term in the particular cases $\beta = 1$ and 2.

The rest of the paper is outlined as follows. In section 2 we introduce the main features of the classical Wishart ensemble along with the evolution law for the eigenvalue process. In section 3, we construct a 3-parameters matrix model (and the respective evolution law for the eigenvalues) that at large times interpolates between the Wishart ensemble and so-called CIR processes whose stationary pdf is a certain Gamma distribution. The corresponding parametrical density of states is computed exactly in the two ways described above (from the saddle point route on the partition function in section 3.3 and from Itô's calculus in section 3.4) and constitutes a continuous deformation of the Marčenko-Pastur distribution (see below). In section 4 we compute the systematic $1/N$ (for all $\beta > 0$) and $1/N^2$ (for $\beta = 1, 2$) corrections to the Marčenko-Pastur law for the smoothed density in the bulk for the (scaled) β-Wishart ensemble. For the special cases $\beta = 1$, 2 and 4, the

3

$\mathcal{O}(1/N)$ correction term was computed in [14, 15]. Our result generalizes this to arbitrary β for the $\mathcal{O}(1/N)$ term and in addition, we obtain the $\mathcal{O}(1/N^2)$ correction for $\beta = 1$ and $\beta = 2$. We conclude with a summary and discussion in section 5.

2 Wishart ensembles

2.1 Real and complex Wishart ensembles

Let $\mathbf{X}$ be a real (respectively complex) Gaussian random matrix of size $M \times N$, i.e. a random matrix chosen in the space of $M \times N$ real (resp. complex) matrices according to the law:

$$P(\mathbf{X})d\mathbf{X} \propto \exp\left(-\frac{1}{2\sigma^2}\mathrm{Tr}(\mathbf{X}^\dagger\mathbf{X})\right)d\mathbf{X}, \qquad (2.1)$$

where $\mathbf{X}^\dagger$ is the Hermitian conjugate of $\mathbf{X}$. In the following, we will denote the real (resp. complex) Wishart ensemble by $\mathcal{W}^\beta$ with $\beta = 1$ in the real case (resp. $\beta = 2$ in the complex case).

The real (resp. complex) Wishart Ensemble is the ensemble of $(N \times N)$ square matrices of the product form $\mathbf{W} := \mathbf{X}^\dagger\mathbf{X}$ where $\mathbf{X}$ is a real (resp. complex) Gaussian random matrix of size $N \times M$. They have appeared in many different applications such as communication technology [16], nuclear physics [17], quantum chromodynamics [18], statistical physics of directed polymers in random media [19] and non intersecting Brownian motions [20, 21, 22], as well as Principal Component Analysis of large datasets [23, 24].

The spectral properties of the Wishart matrices have been studied extensively and it is known [25] that for $M \geqslant N$, all N positive eigenvalues of $\mathbf{W}$ are distributed via the joint probability density function (pdf)

$$P_\beta(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^N \lambda_i}\prod_{i=1}^N \lambda_i^{\frac{\beta}{2}(M-N+1)-1}\prod_{i<j}|\lambda_i - \lambda_j|^\beta \qquad (2.2)$$

where Z is a constant normalization factor and where $\beta = 1$ in the real case (resp. $\beta = 2$ in the complex case). Note that the joint distribution P_β defined in (2.2) is in fact well behaved for every $\beta > 0$.

Another classical result of Random Matrix Theory concerns the asymptotic density of states (or spectral measure) for the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ of a real Wishart matrix $\mathbf{W} \in \mathcal{W}^1$ in the limit of large matrices, i.e. when $N, M \to \infty$ with $N/M = q \in (0; 1]$ where q is a fixed parameter. Let us recall that the density of states of the matrix $\mathbf{W}$ is simply the probability measure ρ_N^β defined as

$$\rho_N^\beta = \frac{1}{N}\sum_{i=1}^N \delta(\lambda - \lambda_i) \qquad (2.3)$$

where β is introduced for later convenience ($\beta \equiv 1$ in the present case) and where $(\lambda_1, \ldots, \lambda_N)$ are the eigenvalues of $\mathbf{W}$. Setting $\lambda = M\hat{\lambda}$, the Marčenko-Pastur Theorem (see [26]) states that, in the limit $N, M \to \infty$ with $N/M = q \in (0; 1]$, the spectral

measure of a Wishart matrix $\mathbf{W} \in \mathcal{W}^1$ converges to a continuous probability density (with compact support) given by

$$\rho_\beta(\hat{\lambda}) = \frac{1}{2\pi\sigma^2\beta q} \frac{\sqrt{(b-\hat{\lambda})(\hat{\lambda}-a)}}{\hat{\lambda}}, \quad a < \hat{\lambda} < b \tag{2.4}$$

where the edges a, b of the spectrum are given by

$$a = \sigma^2\beta(1 - \sqrt{q})^2, \quad b = \sigma^2\beta(1 + \sqrt{q})^2$$

with again $\beta = 1$.

For general $\beta > 0$, the probability measure ρ_N^β is defined again as in (2.3) where this time the vector $(\lambda_1, \ldots, \lambda_N)$ is distributed according to the law P_β in (2.2). The Marčenko-Pastur theorem remains in fact valid for all $\beta > 0$ in the sense that the probability law ρ_N^β converges when $N, M \to \infty$ with $N/M = q \in (0; 1]$ to the continuous probability density ρ_β in (2.4) for every $\beta > 0$.

The probability measure ρ_N^β will sometimes be referred to as the spectral density as it corresponds to the spectral density of random matrices $\mathbf{W} \in \mathcal{W}^\beta$ at least when $\beta = 1$ or 2.

2.2   Continuous processes for real and complex Wishart ensembles

We wish to define here a diffusive matrix process depending on a fictitious time $t \geqslant 0$ that will converge to the Wishart Ensembles in the limit of large time. The idea is simply to set

$$\mathbf{W}_t := \mathbf{X}_t^\dagger \mathbf{X}_t \tag{2.5}$$

where $\mathbf{X}_t$ is a real (resp. complex) random matrix process (of size $M \times N$) following the Ornstein-Uhlenbeck law,

$$d\mathbf{X}_t = -\frac{1}{2}\mathbf{X}_t dt + \sigma\, d\mathbf{B}_t$$

where $\mathbf{B}_t$ is a real Brownian (resp. complex) random matrix, i.e. a matrix whose entries are given by independent standard Brownian motions. By a standard Brownian motion, one means a centered (zero-mean) Gaussian process with covariance function $\langle B_t B_{t'} \rangle = \min(t, t')$.

It is well known that the stationary law of a Ornstein-Uhlenbeck process is the Gaussian law and therefore, the real (resp. complex) matrix process $\mathbf{X}_t$ converges in law when $t \to \infty$ to the law of a Gaussian real (resp. complex) random matrix. Hence, we deduce that the real (resp. complex) matrix process $\mathbf{W}_t$ defines a diffusive matrix process that converges in law to $\mathbf{W} \in \mathcal{W}^\beta$ with $\beta = 1$ or $\beta = 2$.

It is also easy to check that the positive definite matrix process $\mathbf{W}_t$ verifies the following stochastic differential equation studied by Bru [27]:

$$d\mathbf{W}_t = -\mathbf{W}_t dt + \sigma\sqrt{\mathbf{W}_t}\, d\mathbf{B}_t + \sigma\, d\mathbf{B}_t^\dagger \sqrt{\mathbf{W}_t} + M\sigma^2\beta\,\mathbf{I}\, dt \tag{2.6}$$

where $\mathbf{B}_t$ is a real (resp. complex) Brownian random matrix and with $\beta = 1$ in the real (resp. $\beta = 2$ for complex) case.

The evolution of the eigenvalue process $\lambda_1(t) \leqslant \lambda_2(t) \ldots \leqslant \lambda_N(t)$ is also easy to derive [28] using perturbation theory to second order

$$d\lambda_i = -\lambda_i dt + 2\sigma \sqrt{\lambda_i} \, db_i + \sigma^2 \beta \left(M + \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) dt \qquad (2.7)$$

where b_i's are independent standard Brownian motions and with $\beta = 1$ in the real (resp. $\beta = 2$ for complex) case. The stationary distribution of the process $(\lambda_1, \ldots, \lambda_N)(t)$ is necessarily the jpdf $P_\beta(\lambda_1, \ldots, \lambda_N)$ defined in (2.2) (this is true for any $\beta > 0$ and can also be recovered using the Fokker-Planck equation for the multivariate diffusion (2.7)).

3 Crossover between Wishart and CIR processes

Following [9, 10], we aim at defining a diffusive matrix process $\mathbf{W}_t$ which converge in the limit of large time to a general β-Wishart matrix, i.e. a matrix whose eigenvalues are distributed according to P_β in (2.2) for general $\beta > 0$ and with Haar distributed eigenvectors. In this paper, we will restrict ourselves to the description of the eigenvalues process but the interested reader can find a study of the eigenvectors for a related model in [10]. To simplify notations, we will take in this section $\sigma = 1$.

3.1 Preliminary definition: CIR diffusion process

We first need to introduce a family of real diffusion processes. Let $\delta > 0$ be a fixed parameter. The CIR process (named after its creators Cox, Ingersoll, and Ross [29] and widely used in finance to model short term interest rate) is a diffusion process $x(t)$ defined by $x(0) := x_0 > 0$ and for $t \geqslant 0$ by

$$dx(t) = -x(t) \, dt + 2 \sqrt{x(t)} \, db_t + \delta \, dt \,. \qquad (3.1)$$

Using the assumption $\delta > 0$, it is easy to see that the process $x(t)$ will remain non negative for all times $t \geqslant 0$. It is also easy to verify that the stationary pdf of the Langevin equation (3.1) is the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$ defined as

$$p_\delta(x) = \frac{1}{2^{\frac{\delta}{2}} \Gamma(\frac{\delta}{2})} x^{\frac{\delta}{2}-1} e^{-\frac{x}{2}} \,. \qquad (3.2)$$

In analogy with squared Bessel processes, the parameter δ will be called the *dimension* of the process $x(t)$.

3.2 Diffusive matrix process for general β-Wishart matrices

Following [9, 10], our goal is to construct a diffusive matrix process whose eigenvalues process is asymptotically distributed according to P_β for general $\beta \in [0,1]$. This construction can be extended respectively for general $\beta \in [0,2]$ (resp. $\beta \in [0,4]$) by using complex (resp. symplectic) Brownian motions instead of real Brownian motions in the following.

We will in fact describe how to handle the value $\beta \in [0,1]$ by using real Brownian matrix. This construction can be extended for the values $\beta \in [0,2]$ using complex Brownian matrix and also $\beta \in [0,4]$ using symplectic Brownian matrix.

The idea is to slice the time interval into small chops of length $1/n$ and for each interval $[k/n; (k+1)/n]$, to choose independently Bernoulli random variables $\epsilon_k^n, k \in \mathbb{N}$ such that $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$. Then, setting $\epsilon_t^n = \epsilon_{[nt]}^n$, our diffusive matrix process evolves as:

$$d\mathbf{W}_t^n = -\mathbf{W}_t^n\, dt + d\boldsymbol{\Delta}_t^n \qquad (3.3)$$

where the increment matrix $d\boldsymbol{\Delta}_t^n$ now depends on the value of the additional random process ϵ_t^n:

- if $\epsilon_t^n = 1$, then

$$d\boldsymbol{\Delta}_t^n = \sqrt{\mathbf{W}_t^n}\, d\mathbf{B}_t + d\mathbf{B}_t^\dagger\, \sqrt{\mathbf{W}_t^n} + M\, \mathbf{I}\, dt.$$

  where $d\mathbf{B}_t$ is an $N \times N$ real [1] Brownian increment matrix whose entries have variance dt.

- if $\epsilon_t^n = 0$, then

$$d\boldsymbol{\Delta}_t^n = \sqrt{\mathbf{W}_t^n}\, d\mathbf{Y}_t + d\mathbf{Y}_t^\dagger\, \sqrt{\mathbf{W}_t^n} + \delta\, \mathbf{I}\, dt.$$

  with $\delta > 0$ and where $d\mathbf{Y}_t$ is a symmetric matrix that is co-diagonalizable with $\mathbf{W}_t^n$ (i.e. the two matrix have the same eigenvectors) but with a spectrum given by N independent real Brownian increments of variance dt.

An algorithmic description of how to build (approximatively on a discrete grid) the matrix process $\mathbf{W}_t^n$ can be found in Appendix B.

It is clear that the eigenvalues of the matrix $\mathbf{W}_t^n$ will cross at some points but only in intervals $[k/n; (k+1)/n]$ for which $\epsilon_k^n = 0$ (in the other intervals where they follow the SDE (2.7) with parameter $\beta = 1$, it is well known that the repulsion is too strong and thus collisions are avoided). In this case, the eigenvalues are re-numbered at time $t = (k+1)/n$ in increasing order. With this procedure, when ordered $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t)$, we can again check as in [9, 10], using perturbation theory, that the eigenvalues will remain always non-negative and will verify the Stochastic Differential System (SDS):

$$d\lambda_i^n = -\lambda_i^n dt + 2\sqrt{\lambda_i^n}\, db_i + \left(\epsilon_t^n M + (1 - \epsilon_t^n)\delta + \epsilon_t^n \sum_{k \neq i} \frac{\lambda_i^n + \lambda_k^n}{\lambda_i^n - \lambda_k^n}\right) dt \qquad (3.4)$$

where the b_i are independent standard Brownian motions, which are also independent of the process ϵ_t^n.

Note that when $\epsilon_t^n = 0$, the particles λ_i^n are evolving as independent CIR processes of dimension $\delta > 0$ as defined in paragraph 3.1. Therefore, the particles can cross in those time intervals, breaking the increasing order so that they will be re-ordered at time $([nt] + 1)/n$ but they $will$ remain non-negative as the dimension δ is strictly positive. Therefore the SDS (3.4) remains well defined at all times $t \geqslant 0$.

One can follow the proof of [10] to prove that the scaling limit (i.e. the limiting process when $n \to \infty$) of the process $(\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t))$ satisfies the following SDS

$$d\lambda_i = -\lambda_i dt + 2\sqrt{\lambda_i}\, db_i + \left(pM + (1 - p)\delta + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k}\right) dt. \qquad (3.5)$$

_____

[1] Here one can use use complex Brownian motions instead to extend the interval of β to $[0, 2]$.

One can deduce from the above equation (3.5) the Fokker-Planck equation for the joint density $P(\{\lambda_i\}, t)$, for which the stationary jpdf is readily found to be [see the derivation in appendix A]

$$P^*(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z} e^{-\frac{1}{2}\sum_{i=1}^N \lambda_i} \prod_{i=1}^N \lambda_i^{\frac{p}{2}(M-N+1-\delta)-(1-\frac{\delta}{2})} \prod_{i<j} |\lambda_i - \lambda_j|^p. \qquad (3.6)$$

The probability P_β introduced in (2.2) is recovered here by taking the values $p = \beta$ and $\delta = 0$. The corresponding large N, M-limit spectral probability density is therefore given by the Marčenko-Pastur law in the case where $p = \beta > 0$ *independent of* M. Note that with the above normalizations, the spectrum is spread over a region of $\mathbb{R}_+$ of width of order $pM = \beta M$. On the other hand, if $p = 0$, the large N, M-limit of the spectral density is the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$ (recall that it is the stationary pdf of the CIR process of dimension δ):

$$\rho_0(\lambda)d\lambda = \frac{1}{2^{\delta/2}\Gamma(\frac{\delta}{2})} \lambda^{\frac{\delta}{2}-1} e^{-\frac{\lambda}{2}} \, d\lambda. \qquad (3.7)$$

It is quite natural to ask whether a crossover regime may be found, interpolating between the Marčenko-Pastur density (p independent of M) and the Gamma distribution ($p = 0$). A good candidate for triggering such a transition is clearly a parameter p vanishing with M as $p = 2c/M$ where c is a positive fixed constant. We discuss this case in the following subsection.

In the next two subsections we compute the crossover density interpolating between the Marčenko-Pastur law and the Gamma distribution with shape parameter $\delta/2$. This family of probability densities is indexed by the three parameters c (such that $p = 2c/M$), $q = N/M$ and $\delta \geqslant 0$. More precisely, we compute the limiting density of the probability measure $\rho_N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$, when $N, M \to \infty$ with $N/M = q \in (0; 1]$ and where $(\lambda_1, \ldots, \lambda_N)$ is distributed according to the law P^* defined in (3.6) with $p = \beta = 2c/M$. As mentioned in the introduction, the crossover density can be computed via two alternative methods: (1) by a saddle point method as shown in subsection 3.3 and (2) by analyzing directly the stochastic differential systems introduced in 3.2 above following the analogous route for the Gaussian case in Ref. [9]. This is done in section 3.4. We will see that both methods yield identical result.

3.3 Crossover for the spectral density via the saddle point method

Our starting point in the joint probability law of eigenvalues in Eq. (3.6), where the normalization constant (partition function) Z is given by the N-fold integral

$$Z = \int_{[0,\infty]^N} \prod_i d\lambda_i e^{-\frac{1}{2}\sum_i \lambda_i} \prod_{i<j} |\lambda_i - \lambda_j|^p \prod_i \lambda_i^{\frac{p}{2}(M-N+1-\delta)-(1-\delta/2)}$$

$$= \int_{[0,\infty]^N} \prod_i d\lambda_i e^{-E[\{\lambda_i\}]} \qquad (3.8)$$

where the energy function $E[\{\lambda_i\}]$ is given by

$$E[\{\lambda_i\}] = \frac{1}{2}\sum_i \lambda_i - \left(\frac{p}{2}(M-N+1-\delta) - (1-\delta/2)\right)\sum_i \ln \lambda_i - \frac{p}{2}\sum_{i \neq j} \ln |\lambda_i - \lambda_j|. \quad (3.9)$$

8

Written in this form, Eq. (3.8) is the Gibbs-Boltzmann canonical weight of a system of charged particles on the positive half-line in equilibrium at inverse temperature $\beta = 1$ under the effect of competing interactions. We wish to compute the average density of states $\langle \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i) \rangle$ where $\langle O \rangle$ denote the expectation value of O with respect to the probability distribution P^* in (3.6). There are many ways to compute this average density, but the one rather convenient for large N is the saddle point route. This was originally done by Dyson [30] for the Gaussian random matrices and a physically more transparent derivation can be found in Ref. [31].

The main idea behind the saddle point method is as follows. In the large N limit, the most dominant contribution to the partition function emerges indeed from a set of configurations of λ_i's that correspond to a particular density $\rho^*(\lambda)$. Naturally then, the average computed over the ensemble of λ_i's, in this large N limit, will also be given by the saddle point density $\langle \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i) \rangle \approx \rho^*(\lambda)$. It thus suffices to analyze just the partition function Z in the large N limit and find, in particular, the saddle point density $\rho^*(\lambda)$ that maximizes the partition function Z for large N.

To analyze Z in the large N limit, one first defines a 'local' smooth density function

$$\rho(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i) \tag{3.10}$$

which is normalized to unity. The main idea then is to split the multiple integration in Eq. (3.8) in two parts: First fix the local density $\rho(\lambda)$ and sum over all microstates (i.e., configurations of λ_i's consistent with the local density defined in (3.10)) and then, sum (functionally) over all possible local density functions. Roughly speaking, the first step corresponds to a partial tracing over microstates by fixing the local density. Notationally, on can express this by

$$Z = \int \mathcal{D}[\rho] \int_{[0,\infty]^N} \prod_i d\lambda_i e^{-E[\{\lambda_i\}]} I[\rho(\lambda), \{\lambda_i\}] \tag{3.11}$$

where $\mathcal{D}[\rho]$ denotes a functional integration over the function space and $I[\rho(\lambda), \{\lambda_i\}]$ is an indicator function that is 1 if the microstate $\{\lambda_i\}$ is compatible with a given $\rho(\lambda)$, normalized to unity, as defined in (3.10) and otherwise $I = 0$. The energy function $E[\{\lambda_i\}]$ associated with a microstate can then be expressed in terms of the local density $\rho(\lambda)$ using the identity $\sum_i f(\lambda_i) = N \int d\lambda f(\lambda) \rho(\lambda)$ and one gets [31]

$$E[\rho(\lambda)] = \frac{N}{2} \int d\lambda \lambda \rho(\lambda) - \left[\frac{p}{2} \left((\frac{1}{q} - 1)N + 1 - \delta \right) - \left(1 - \frac{\delta}{2} \right) \right] N \int d\lambda \rho(\lambda) \ln \lambda$$

$$- \frac{p}{2} N^2 \int \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| + \frac{p}{2} N \int d\lambda \rho(\lambda) \ln \frac{1}{\rho(\lambda)} + C_1 \left(\int d\lambda \rho(\lambda) - 1 \right) \tag{3.12}$$

where the last term includes a Lagrange multiplier C_1 that enforces the normalization of the local density to 1. The next-to-last term accounts for the self-energy term ($\lambda \to \lambda'$) that needs to be subtracted. Note indeed that in the original discrete sum $\sum_{i \neq j} \ln |\lambda_i - \lambda_j|$, the eigenvalues do not coincide. This means that the integral over λ and λ' should exclude the region where $|\lambda - \lambda'|$ is less than the typical spacing between eigenvalues, which is

proportional to $1/N\rho(\lambda)$ [30]. The contribution of this thin sliver is the next-to-last term, up to an additional contribution that can be absorbed into C_1.

Once this is done, equation (3.11) simplifies further and one gets

$$Z \approx \int \mathcal{D}[\rho] \, e^{-E[\rho(\lambda)]} \, J[\rho(\lambda)] \tag{3.13}$$

where $J[\rho(\lambda)] = \int_{[0,\infty]^N} \prod_i d\lambda_i I[\rho(\lambda), \{\lambda_i\}]$ is an entropic factor that just counts how many microstates are compatible with a given local density function $\rho(\lambda)$. This can be estimated very simply by the following combinatorial argument involving arrangement of N particles in K boxes. Let us first divide our one dimensional line into K small boxes of equal width. We have N particles that need to be distributed into the K boxes with occupation numbers $\{n_1, n_2, \ldots, n_K\}$. The number of ways this can be done is simply

$$\frac{N!}{n_1! n_2! \cdots n_K!} \, . \tag{3.14}$$

Setting $\rho_i = n_i/N$ (the local density in box i) and using Stirling's approximation $N! \sim N^{N+1/2} e^{-N}$ (using the fact that $\sum_i n_i = N$), we have:

$$\frac{N!}{n_1! n_2! \cdots n_K!} \sim e^{-\sum_i n_i \ln n_i} \tag{3.15}$$

which in the continuum limit, becomes $\sim e^{-N \int d\lambda \rho(\lambda) \ln \rho(\lambda)}$. Thus, the entropic factor can also be expressed as a simple functional of the local density $\rho(\lambda)$. Inserting this expression in the functional integral over the density, yields:

$$Z = \int \mathcal{D}[\rho] e^{-E[\rho(\lambda)]} e^{-N \int d\lambda \rho(\lambda) \ln \rho(\lambda)} = \int \mathcal{D}[\rho] e^{-NF[\rho(\lambda)]} \tag{3.16}$$

where the free energy $F[\rho(\lambda)]$ is given by:

$$F[\rho(\lambda)] = \frac{1}{2} \int d\lambda \lambda \rho(\lambda) - \left[\frac{p}{2} \left((\frac{1}{q} - 1)N + 1 - \delta \right) - \left(1 - \frac{\delta}{2} \right) \right] \int d\lambda \rho(\lambda) \ln \lambda$$
$$- \frac{p}{2} N \int \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| + \left(1 - \frac{p}{2} \right) \int d\lambda \rho(\lambda) \ln \rho(\lambda) + C_1 \left(\int d\lambda \rho(\lambda) - 1 \right) \tag{3.17}$$

Note that for $p \sim \mathcal{O}(1/N)$ the entropy term becomes of the same order of the energy term, while in the usual case $p \sim \mathcal{O}(1)$ the entropy contribution is subdominant in the large N limit and is therefore disregarded.

Setting now $p = 2c/M = 2cq/N$, we get:

$$F[\rho(\lambda)] = \frac{1}{2} \int d\lambda \lambda \rho(\lambda) - \left[cq \left(\frac{1}{q} - 1 \right) - \left(1 - \frac{\delta}{2} \right) \right] \int d\lambda \rho(\lambda) \ln \lambda$$
$$- cq \int \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| + \left(1 - \frac{cq}{N} \right) \int d\lambda \rho(\lambda) \ln \rho(\lambda) + C_1 \left(\int d\lambda \rho(\lambda) - 1 \right) \tag{3.18}$$

We set $a = cq(1/q-1)-(1-\delta/2)$ and take $N \to \infty$ (so that the term cq/N drops out). Now, the saddle point density $\rho^*(\lambda)$ is obtained by minimizing the free energy $F[\rho(\lambda)$, i.e., by taking the functional derivative $\frac{\delta F}{\delta \rho} = 0$ and $\rho(\lambda) == \rho^*(\lambda)$. This gives the saddle point equation

$$\frac{\lambda}{2} - a \ln \lambda - 2cq \int d\lambda' \rho^*(\lambda') \ln|\lambda - \lambda'| + \ln \rho^* + C_2 = 0 \tag{3.19}$$

where $C_2 = C_1 + 1$ is just a constant. For notational simplicity, in the rest of the subsection we will denote the saddle point density $\rho^*(\lambda)$ simply by $\rho(\lambda)$.

Taking one more derivative of (3.19), we get

$$\frac{1}{2} - \frac{a}{\lambda} - 2cq \, \text{Pr} \int \frac{\rho(\lambda')}{\lambda - \lambda'} d\lambda' + \frac{\rho'(\lambda)}{\rho(\lambda)} = 0 \tag{3.20}$$

Next, we define the Stieltjes transform:

$$H(z) = \int \frac{\rho(\lambda)}{\lambda - z} d\lambda \tag{3.21}$$

for z complex and outside the support of ρ. By definition, for large $|z|$, $H(z) \to -1/z$. Multiplying eq. (3.20) by $\rho(\lambda)/(\lambda - z)$ and integrating over λ, we have:

$$\frac{1}{2} \int \frac{\rho(\lambda)}{\lambda - z} d\lambda - a \int \frac{\rho(\lambda) d\lambda}{\lambda(\lambda - z)} - 2cq \, \text{Pr} \int \frac{\rho(\lambda) d\lambda}{\lambda - z} \int \frac{\rho(\lambda')}{\lambda - \lambda'} d\lambda' + \int \frac{\rho'(\lambda) d\lambda}{\lambda - z} = 0 \tag{3.22}$$

and we analyze each of the four contributions separately.

1. $T_1 = \frac{1}{2} \int \frac{\rho(\lambda)}{\lambda - z} d\lambda = \frac{1}{2} H(z)$

2. $T_2 = -a \int \frac{\rho(\lambda) d\lambda}{\lambda(\lambda - z)}$. We rewrite this as:

$$T_2 = -a \int \rho(\lambda) d\lambda \left[\frac{1}{\lambda - z} - \frac{1}{\lambda} \right] \frac{1}{z} \tag{3.23}$$

implying:

$$T_2 = -\frac{a}{z} H(z) + \frac{b_1}{z} \tag{3.24}$$

where $b_1 = a \int d\lambda \frac{\rho(\lambda)}{\lambda}$.

3. $T_3 = -2cq \, \text{Pr} \int \frac{\rho(\lambda) d\lambda}{\lambda - z} \int \frac{\rho(\lambda')}{\lambda - \lambda'} d\lambda'$ which we rewrite as:

$$T_3 = 2cq \left\{ \text{Pr} \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \left[\frac{1}{\lambda - z} - \frac{1}{\lambda - \lambda'} \right] \frac{1}{\lambda' - z} \right\} \tag{3.25}$$

$$= 2cq H^2(z) - 2cq \, \text{Pr} \int \frac{d\lambda d\lambda' \rho(\lambda) \rho(\lambda')}{(\lambda - \lambda')(\lambda' - z)} \tag{3.26}$$

By renaming $\lambda \to \lambda'$ and $\lambda' \to \lambda$, we get:

$$T_3 = 2cq H^2(z) + 2cq \text{Pr} \int \frac{d\lambda d\lambda' \rho(\lambda) \rho(\lambda')}{(\lambda - \lambda')(\lambda - z)} \tag{3.27}$$

$$= 2cq H^2(z) - T_3 \tag{3.28}$$

Solving for T_3 we get:

$$T_3 = cq H^2(z) \tag{3.29}$$

11

4. $T_4 = \int \frac{\rho'(\lambda)d\lambda}{\lambda - z}$, which we integrate by parts, obtaining:

$$T_4 = \frac{1}{\lambda - z}\rho(\lambda)\Big|_0^\infty + \int \frac{\rho(\lambda)}{(\lambda - z)^2}d\lambda$$
$$= \frac{c_1}{z} + H'(z) \tag{3.30}$$

In the derivation above, we assumed b_1 and c_1 to be finite. This is not completely obvious, because $\rho(\lambda)$ at an edge point may diverge. However, by imposing that for large z, $H(z) \to -1/z$, it is immediate to derive that $b_1 + c_1 = 1/2$. Thus, one may regularize the density near the edge points so that b_1 and c_1 exist individually, but eventually their sum is universally $1/2$ and hence is independent of the specific regularization near the edge.

Adding up the four contributions, we get the equation:

$$\frac{dH}{dz} - \frac{a}{z}H + \frac{b_1 + c_1}{z} + cq\,H^2 + \frac{1}{2}H = 0$$

Thus we find the following differential equation for the Stieltjes transform H:

$$\boxed{\frac{dH}{dz} + \gamma\,H^2 + \frac{1}{2}\left(1 + \frac{\alpha}{z}\right)H + \frac{1}{2z} = 0} \tag{3.31}$$

where we have set

$$\alpha = (2 - \delta) - 2c(1 - q), \quad \gamma = cq\,.$$

In the next subsection, we will derive the same equation via Itô's stochastic calculus route.

The density $\rho(\lambda)$ (normalized to unity) can then be read off from

$$\rho(\lambda) = \frac{1}{\pi}\mathrm{Im}[H(z \to \lambda)] \tag{3.32}$$

where $z \to \lambda$ occurs inside the cut on the real axis.

To solve the Riccati equation (3.31), we make a standard substitution

$$H(z) = \frac{1}{\gamma}\frac{u'(z)}{u(z)} = \frac{1}{\gamma}\partial_z \ln u(z)\,. \tag{3.33}$$

This gives a second order differential equation for $u(z)$

$$u''(z) + \frac{1}{2}\left[1 + \frac{\alpha}{z}\right]u'(z) + \frac{\gamma}{2z}u(z) = 0\,. \tag{3.34}$$

It follows from Eq. (3.33) and the asymptotic behavior of $H(z)$ that

$$u(z) \xrightarrow[|z|\to\infty]{} \frac{A_1}{z^\gamma} \tag{3.35}$$

where A_1 is a constant.

To reduce Eq. (3.34) to a Schrödinger like differential equation, we make the substitution

$$u(z) = e^{-z/4} \, z^{\alpha/4} \, \psi(z) \,, \tag{3.36}$$

and we find the following equation for ψ

$$\psi''(z) + \left[-\frac{1}{16} + \frac{1}{z} \frac{4\gamma - \alpha}{8} + \frac{\alpha}{4} \left(1 - \frac{\alpha}{4} \right) \frac{1}{z^2} \right] \psi(z) = 0 \,.$$

Making further a rescaling $\psi(z) = y(z/2)$, it reduces to the standard form of the Whittaker differential equation [32]

$$y''(z) + \left[-\frac{1}{4} + \frac{\lambda}{z} + \frac{\frac{1}{4} - \mu^2}{z^2} \right] y(z) = 0 \,, \tag{3.37}$$

where

$$\zeta = \gamma - \frac{\alpha}{4} \,, \quad \mu = \frac{1}{2} |1 - \frac{\alpha}{2}| \,. \tag{3.38}$$

Note that the solution of this differential equation does not depend on the sign of μ, hence we take the absolute value. The differential equation (3.37) has two linearly independent solutions $W_{\zeta,\mu}(z)$ and $W_{-\zeta,\mu}(-z)$. The Whittaker function $W_{\zeta,\mu}(z)$ has the following asymptotic behavior [32]

$$W_{\zeta,\mu}(z) \xrightarrow[|z|\to\infty]{} z^\zeta \, e^{-z/2} \,. \tag{3.39}$$

Thus the general solution of $u(z)$, using Eq. (3.36), reads

$$u(z) = e^{-z/4} \, z^{\alpha/4} \, [\mathcal{C}_1 \, W_{\zeta,\mu}(z/2) + \mathcal{C}_2 \, W_{-\zeta,\mu}(-z/2)] \tag{3.40}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are arbitrary constants. Using the asymptotic behavior in Eq. (3.39) it is easy to check that only the second solution has the right asymptotic behavior in Eq. (3.35). Thus, finally, we have our solution

$$u(z) = \mathcal{C}_2 \, e^{-z/4} \, z^{\alpha/4} \, W_{-\zeta,\mu}(-z/2) \tag{3.41}$$

where ζ and μ are given in Eq. (3.38).

By plugging this solution (3.41) into Eq. (3.33) and using Eq. (3.32), we find the following expression

$$\rho(\lambda) = \frac{\mathcal{C}_2}{2\pi\gamma} \frac{(\mathrm{Im}(W'_{-\zeta,\mu})\mathrm{Re}(W_{-\zeta,\mu}) - \mathrm{Im}(W_{-\zeta,\mu})\mathrm{Re}(W'_{-\zeta,\mu}))(-\lambda/2)}{|W_{-\zeta,\mu}(-\lambda/2)|^2} \,. \tag{3.42}$$

where Re and Im denote respectively the real and imaginary parts. Using the linear differential equation verified by the Whittaker functions (3.37), it is easy to see that the derivative with respect to λ of the Wronskian type function $(\mathrm{Im}(W'_{-\zeta,\mu})\mathrm{Re}(W_{-\zeta,\mu}) - \mathrm{Im}(W_{-\zeta,\mu})\mathrm{Re}(W'_{-\zeta,\mu}))$ is equal to 0. Hence the Wronskian appearing in the numerator in (3.42) is simply a constant.

Collecting all the constants together, we get:

$$\rho(\lambda) = \frac{A}{|W_{-\zeta,\mu}(-\lambda/2)|^2} \,. \tag{3.43}$$

13

The overall normalization constant A has to be fixed from $\int_0^\infty \rho(\lambda)\,d\lambda = 1$. Thus we get, after rescaling $\lambda/2 \to \lambda$,

$$\frac{1}{A} = 2 \int_0^\infty \frac{d\lambda}{|W_{-\zeta,\mu}(-\lambda)|^2}\,. \tag{3.44}$$

This integral in Eq. (3.44) can be done in closed form. First, we first use the well known identity [33]

$$W_{\zeta,\mu}(z) = z^{\mu+1/2}\,e^{-z/2}\,U(\mu - \zeta + 1/2, 1 + 2\mu; z) \tag{3.45}$$

where $U(a,b;z)$ is the Tricomi hypergeometric function (or Kummer function) that behaves for large z as $U(z) \sim z^{-a}$. Using this in Eq. (3.44) gives

$$\frac{1}{A} = 2 \int_0^\infty d\lambda\, \lambda^{-2\mu-1}\, e^{-\lambda}\, |U(\mu + \zeta + 1/2, 1 + 2\mu; -\lambda)|^{-2}\,. \tag{3.46}$$

It turns out that there exists an interesting integral representation in a paper by Ismail and Kelker [34]

$$\int_0^\infty \frac{dt\, e^{-t}\, t^{-b}}{z+t}\, |U(a,b;-t)|^{-2} = \Gamma(a)\Gamma(a-b+2)\,\frac{1}{z}\frac{U(a,b-1;z)}{U(a,b;z)}; \quad \text{for } a > 0,\, 1 < b < a+1 \tag{3.47}$$

Note that in Ref. [34] they use the notation $\psi(a,b,z)$ instead of $U(a,b;z)$, but it is the same function. Our μ and ζ satisfy the condition of validity of this identity: $a > 0$ and $1 < b < a + 1$. Taking $z \to \infty$ limit on both sides and using $U(z) \sim z^{-a}$, we arrive at the following exact expression of the normalization constant

$$\frac{1}{A} = 2\,\Gamma(\mu + \zeta + 1/2)\Gamma(\zeta - \mu + 3/2)\,. \tag{3.48}$$

This leads to the following final expression for the spectral density, which is the central result of our work[2]:

$$\boxed{\rho_c(\lambda) = \frac{1}{2\Gamma(\mu + \zeta + \frac{1}{2})\Gamma(\zeta - \mu + \frac{3}{2})}\frac{1}{|W_{-\zeta,\mu}(-\frac{\lambda}{2})|^2}} \tag{3.49}$$

with the following values for the parameters

$$\alpha = (2 - \delta) - 2c(1 - q); \quad \zeta = cq - \frac{\alpha}{4}; \quad \text{and} \quad \mu = \frac{1}{4}|\alpha - 2|\,.$$

The above expression is the analogue, in the present context, of the Askey-Wimp-Kerov one-parameter family of models found in [9], that smoothly interpolates between the Gaussian distribution and Wigner's semi-circle.

Let us now consider the limiting case $c \to 0$ first. In this case, we have $\alpha = (2 - \delta)$, $\zeta = \delta/4 - 1/2$ and $\mu = \delta/4$. Thus, $W_{-\zeta,\mu}(-\lambda/2) = W_{1/2-\delta/4,\delta/4}(-\lambda/2)$. It turns out that for these special values of the indices, the Whittaker function simply reduces to $W_{1/2-\delta/4,\delta/4}(-\lambda/2) \propto \lambda^{-\delta/4+1/2}\,e^{\lambda/4}$ up to a proportionality constant [32]. Substituting this in (3.49), we then recover the CIR density in Eq. (3.2). The limit $c \to \infty$ is more

---

[2]We add the subscript c in the notation of the density ρ to recall the dependence in c; the density ρ depends also on the two parameters δ and q but we omit to subscript those.

tricky as one needs to rescale $\lambda \to \hat{\lambda}c$ and take the large c limit carefully. This can be done and one recovers the Marčenko-Pastur law. This can be rewritten (without rescaling λ) as

$$\rho_c(\lambda) \sim_{c \to \infty} \frac{1}{4\pi cq} \frac{\sqrt{(\gamma_+ - \lambda)(\lambda - \gamma_-)}}{\lambda} \mathbf{1}_{\{\gamma_- < \lambda < \gamma_+\}}, \tag{3.50}$$

where $\gamma_\pm = 2c \left(1 \pm \sqrt{q}\right)^2$.

Using standard results on Whittaker functions (see e.g. [32, 33]), it is easy to compute the asymptotic behavior of $\rho_c(\lambda)$ for $\lambda \to 0$ and $\lambda \to +\infty$. Up to multiplicative constants, we have

$$\rho_c(\lambda) \sim_{\lambda \to +\infty} \lambda^{2\zeta} e^{-\frac{\lambda}{2}},$$

and

$$\rho_c(\lambda) \sim_{\lambda \to 0_+} \lambda^{2\mu - 1}.$$

We plotted in Fig. 1 the density ρ_c for $c = 0, 1, 2, 3, 4, 5, 10$ and $q = 1/2, \delta = 1$, showing the progressive deformation of the Gamma distribution with shape parameter $\delta/2$ towards the Marčenko-Pastur distribution (3.50) with parameter $q = 1/2$. The critical value of c at which the divergence at $\lambda \to 0_+$ changes to convergence is $c^* = (2 - \delta)/(2(1 - q)) = 1$. As expected, in Fig. 1, the curve with second highest value at the origin corresponds to $c = 1$ and converges when $\lambda \to 0_+$ to $1/2$. The curve with highest value at the origin is the Gamma distribution with shape parameter $\delta/2 = 1/2$ and diverges at 0_+ to $+\infty$. The other curves corresponding to $c = 2, 3, 4, 5, 10$ converge to 0 when $\lambda \to 0_+$.

We have also verified our analytical result for the crossover density in (3.49) numerically for the sample value $c = 1$ and found very good agreement (see Fig. 2).

3.4 Crossover for the spectral density via Itô's stochastic calculus

In this subsection, we want to re-obtain the result Eq. (3.31) of the previous subsection via Itô's calculus. We therefore consider the process $(\lambda_1(t), \cdots, \lambda_N(t))$ which verifies the stochastic differential system (3.5) with the scaling relation $p = 2c/M$. The idea is to work out the evolution equation of the probability measure

$$\rho_N^t(dx) := \frac{1}{N} \sum_{i=1}^{N} \delta(x - \lambda_i(t)) \tag{3.51}$$

in the large N limit. We expect the equilibrium of this evolution equation to be the solution of (3.31).

In the following, f is a smooth function. Using Itô's formula for $\int f(x)\rho_N^t(dx)$, Eq. (3.5) and the scaling relation $p = 2c/M$, we obtain (see [35] for similar calculations)

$$d \int f(x)\rho_N^t(dx) = \int \left(-x + 2c + \left(1 - \frac{2c}{M}\right)\delta\right) f'(x)\rho_N^s(dx)dt \tag{3.52}$$

$$+ 2\left(1 - \frac{c}{M}\right) \int x f''(x)\rho_N^s(dx)dt$$

$$+ cq \int_0^t \int \int \frac{f'(x) - f'(y)}{x - y}(x + y)\rho_N^s(dx)\rho_N^s(dy)dt + dM_t^N,$$
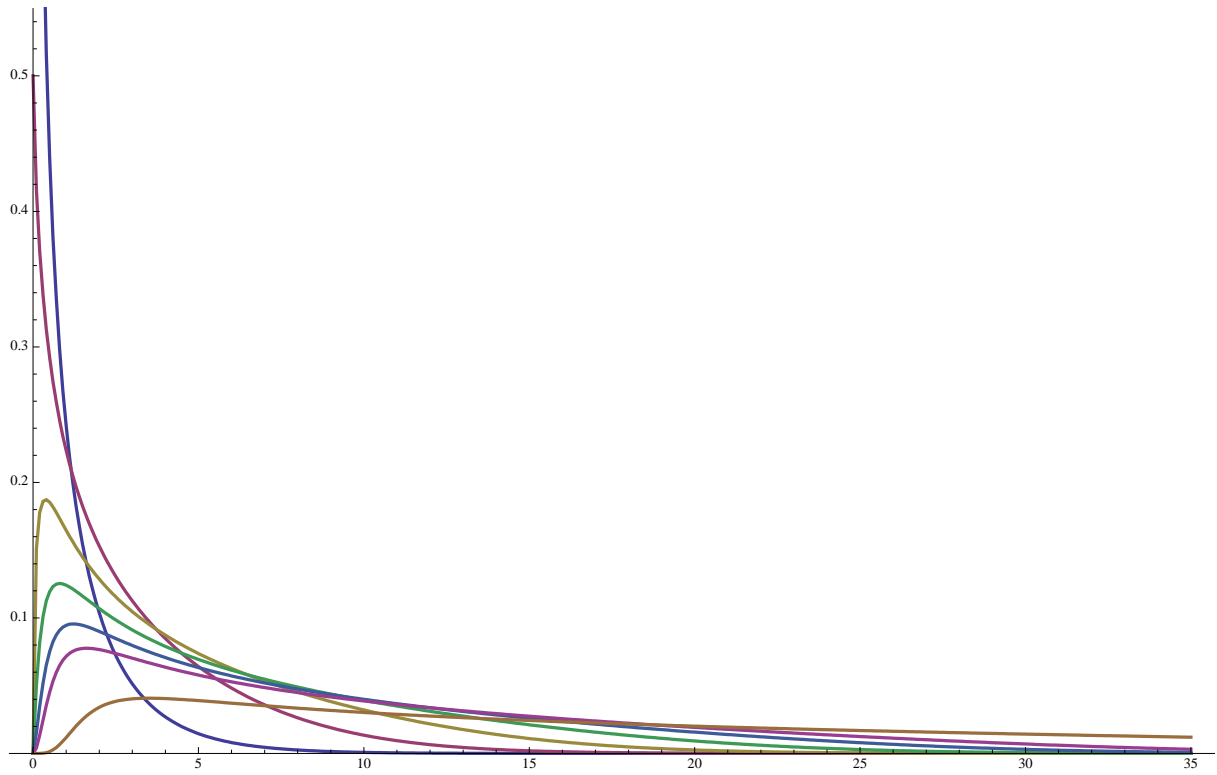
15

Figure 1: Density $\rho_c(\lambda)$ for $c = 0, 1, 2, 3, 4, 5, 10$ of Eq. (3.49) showing the progressive deformation of the Gamma distribution (3.2) with parameter $\delta = 1$ towards the Marčenko-Pastur distribution with parameter $q = 1/2$. The value $\rho_c(0)$ at the origin decreases when c increases.
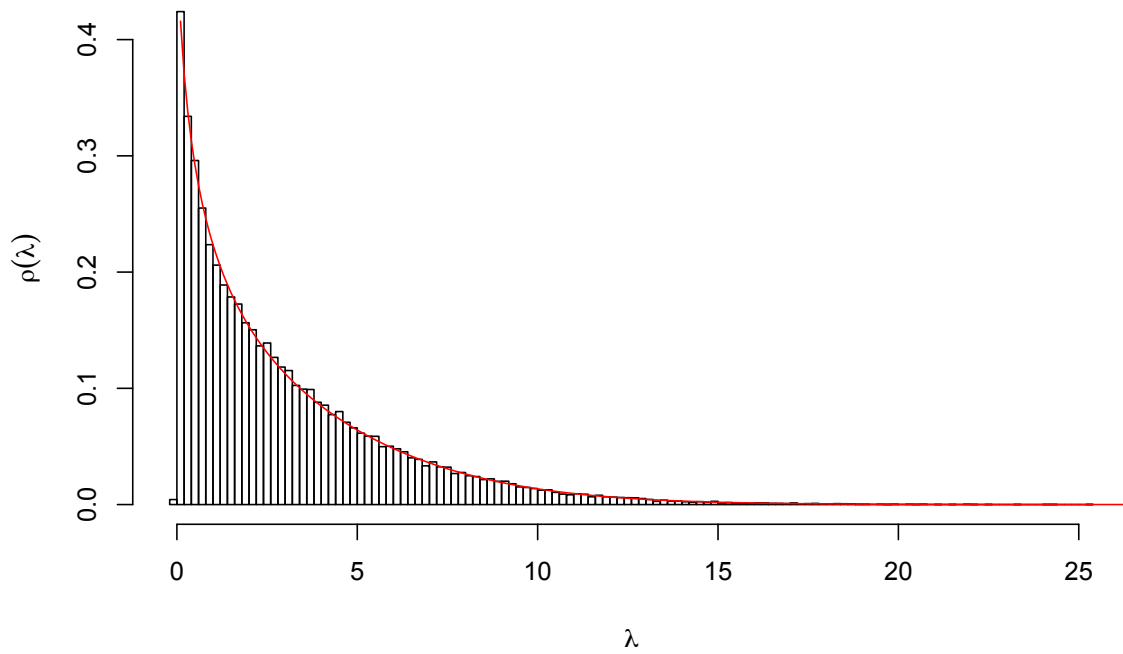
Figure 2: Numerical simulations of the state density of the random matrix $\mathbf{W}_{t=\infty}^n$ defined in Eq. (3.3) for $c = 1, p = \beta = 2c/M, M = 100, N = 50, \delta = 1, q = 1/2$.

where $dM_t^N = \frac{2}{N} \sum_{i=1}^{N} \sqrt{\lambda_i} f'(\lambda_i) db_i$ is a noise term of variance $\frac{4}{N} \int \lambda f'(\lambda)^2 \rho_N^t(d\lambda) dt$. When $N, M \to \infty$ with $N/M = q$, this noise term is of order $1/\sqrt{N}$.

In the large N, M limit, the *stationary* probability measure ρ solution of Eq. (3.52) therefore satisfies to leading order (keeping only the terms of order 1)

$$\int \left(-x + 2c + \delta\right) f'(x)\rho(dx) + 2 \int x f''(x)\rho(dx) \tag{3.53}$$

$$+cq \int \int \frac{f'(x) - f'(y)}{x - y}(x + y)\rho(dx)\rho(dy) = 0$$

Applying Eq. (3.53) to the particular function $f(x) = \frac{1}{x-z}$ for $z \in \mathbb{C} \backslash \mathbb{R}$ and denoting $H(z)$ the Stieltjes transform of the probability measure ρ, we obtain the following differential equation for H

$$\left[H(z) + zH'(z)\right] - (2c + \delta) H'(z) + 2 \left[2H'(z) + zH''(z)\right]$$
$$+ 2cq \left[H'(z) + H^2(z) + 2zH(z)H'(z)\right] = 0. \tag{3.54}$$

Eq. (3.54) can be rearranged as

$$2cqH(z)[H(z) + 2zH'(z)] + \frac{1}{2}[H(z) + 2zH'(z)] + \frac{1}{2}H(z) + [3H'(z) + 2zH''(z)]$$
$$+ [(1 - \delta) - 2c(1 - q)] H'(z) = 0. \tag{3.55}$$

Eq. (3.55) can be integrated easily by doing the change of function $G(z) = zH(z^2)$. Indeed, we just need to write (3.55) for z^2 instead of z and then multiply the corresponding equation by z to obtain the following equation

$$2cqG(z)G'(z) + \frac{1}{2}[zG'(z) + G(z)] + \frac{1}{2}G''(z) + \left[\frac{1 - \delta}{2} - c(1 - q)\right] 2zH'(z^2) = 0,$$

which can be integrated as

$$2cqG^2(z) + \left[z + \frac{(1 - \delta) - 2c(1 - q)}{z}\right] G(z) + G'(z) = -1 \tag{3.56}$$

where the integration constant is chosen so that $zG(z) \sim -1$ when $|z| \to \infty$. Note that the asymptotic behavior for H is therefore also $zH \sim -1$ when $|z| \to \infty$ as is expected for the Stieltjes transform of a probability measure. Rewriting now Equation (3.56) in term of the function H, we obtain exactly Eq. (3.31).

4 Correction to the Marčenko-Pastur law for large but finite dimension

In this section, we come back to the case of generalized Wishart matrices for which particles are distributed according to the jpdf P_β with general parameter $\beta > 0$ (*not* scaling with M). We want to compute the first correction terms to the Marčenko-Pastur density ρ_N^β (defined in Eq. (2.3)) for large but finite N, M with $N/M = q \in (0; 1]$.

We are therefore interested in computing functionals of the form $\langle \int f(\lambda)\rho_N^\beta(d\lambda)\rangle$ where f is a test function and where $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ is distributed according to the jpdf P_β defined in Eq. (2.2) for $\beta > 0$. The idea is to use the stochastic process $\lambda_1(t) \leqslant \ldots \leqslant \lambda_N(t)$ following the SDE (2.7) that converges in law when $t \to \infty$ to the vector $(\lambda_1 \leqslant \ldots \leqslant \lambda_N)$ distributed according to P_β.

Note that in order to have a well behaved limiting spectral distribution with edges that do not depend on M or on β, we will choose in this section $\sigma = 1/\sqrt{M\beta}$ (or alternatively rescale all eigenvalues by $1/M\beta$).

Using again Itô's formula for $\int f(\lambda)\rho_N^t(d\lambda)$ (where ρ_N^t is still defined by Eq. (3.51)) and Eq. (2.7), we obtain

$$d\int f(x)\rho_N^t(dx) = \int(-x+1)f'(x)\rho_N^t(dx)dt + \frac{1}{M\beta}(2-\beta)\int xf''(x)\rho_N^s(dx)dt \quad (4.1)$$

$$+ \frac{q}{2}\int\int\frac{f'(x)-f'(y)}{x-y}(x+y)\rho_N^t(dx)\rho_N^t(dy)dt + dM_t^N$$

where $dM_t^N = \frac{1}{N}\sum_{i=1}^N 2\sqrt{\frac{\lambda_i}{M\beta}}f'(\lambda_i)db_i$ is a noise term of variance $\frac{4}{q\beta M^2}\int\lambda f'(\lambda)^2\rho_N^t(d\lambda)dt$.

Let us introduce the Stieltjes transform $H_t(z)$ of the probability measure ρ_N^t defined as

$$H_t(z) = \int\frac{\rho_N^t(dx)}{x-z}\,.$$

We now apply (4.1) to the particular function $f(x) = 1/(x-z)$ and we take the expectation with respect to the b_i; Eq. (4.1) rewrites as

$$\frac{\partial\langle H_t\rangle}{\partial t} = \left[\langle H_t\rangle + z\frac{\partial\langle H_t\rangle}{\partial z}\right] - \frac{\partial\langle H_t\rangle}{\partial z} + \frac{1}{M\beta}(2-\beta)\left[2\frac{\partial\langle H_t\rangle}{\partial z} + z\frac{\partial^2\langle H_t\rangle}{\partial z^2}\right] \quad (4.2)$$

$$+ q\left[\frac{\partial\langle H_t\rangle}{\partial z} + \langle H_t\rangle^2 + 2z\langle H_t\rangle\frac{\partial\langle H_t\rangle}{\partial z}\right] + q\left[\langle H_t^2\rangle - \langle H_t\rangle^2\right] + zq\frac{\partial}{\partial z}\left[\langle H_t^2\rangle - \langle H_t\rangle^2\right]$$

where $\langle\cdot\rangle$ denotes the expectation with respect to the Brownian motions b_i. The two last terms come from the replacement of $\langle H_t^2\rangle$ by $\langle H_t\rangle^2$ in the third term of the right hand side of (4.2). By setting

$$F_t(z) = \langle H_t(z)^2\rangle - \langle H_t(z)\rangle^2\,, \quad (4.3)$$

Eq. (4.2) can be rewritten as

$$\frac{\partial\langle H_t\rangle}{\partial t} = \left[\langle H_t\rangle + z\frac{\partial\langle H_t\rangle}{\partial z}\right] - \frac{\partial\langle H_t\rangle}{\partial z} + \frac{1}{M\beta}(2-\beta)\left[2\frac{\partial\langle H_t\rangle}{\partial z} + z\frac{\partial^2\langle H_t\rangle}{\partial z^2}\right] \quad (4.4)$$

$$+ q\left[\frac{\partial\langle H_t\rangle}{\partial z} + \langle H_t\rangle^2 + 2z\langle H_t\rangle\frac{\partial\langle H_t\rangle}{\partial z}\right] + q\left[F_t + z\frac{\partial F_t}{\partial z}\right]\,.$$

To simplify notations, we will now omit the $\langle\cdot\rangle$ and write H instead of $\langle H\rangle$. The stationary solution of (4.4) writes simply as

$$\left[H + z\frac{dH}{dz}\right] - \frac{dH}{dz} + \frac{1}{M\beta}(2-\beta)\left[2\frac{dH}{dz} + z\frac{d^2H}{dz^2}\right] \quad (4.5)$$

$$+ q\left[\frac{dH}{dz} + H^2 + 2zH\frac{dH}{dz}\right] + q\left[F + z\frac{dF}{dz}\right] = 0\,.$$

which can be rewritten as

$$qH(z)[H(z) + 2z\frac{dH}{dz}] + \frac{1}{2}[H(z) + 2z\frac{dH}{dz}] + \frac{1}{2}H(z) + \frac{1}{2M\beta}(2-\beta)[3\frac{dH}{dz} + 2z\frac{d^2H}{dz^2}] \quad (4.6)$$

$$+ \left[\frac{1}{2M\beta}(2-\beta) + (-1+q)\right]\frac{dH}{dz} + q\left[F + z\frac{dF}{dz}\right] = 0\,.$$

Eq. (4.6) can be integrated easily by doing the change of function $G(z) = zH(z^2)$. Indeed, we just need to write (4.6) for z^2 instead of z and then multiply the corresponding equation by z to obtain the following equation

$$qG\frac{dG}{dz} + \frac{1}{2}\left[z\frac{dG}{dz} + G\right] + \frac{1}{4M\beta}(2-\beta)\frac{d^2G}{dz^2} \quad (4.7)$$

$$+ \frac{1}{2}\left[\frac{1}{2M\beta}(2-\beta) + (-1+q)\right]2z\frac{dH}{dz}(z^2) + qz\left[F(z^2) + z^2\frac{dF}{dz}(z^2)\right] = 0\,.$$

Equation (4.7) can be straightforwardly integrated with respect to z as

$$qG^2 + zG + \frac{1}{2M\beta}(2-\beta)\frac{dG}{dz} + \left[\frac{1}{2M\beta}(2-\beta) + \beta(-1+q)\right]\frac{G(z)}{z} + qz^2F(z^2) = -1 \quad (4.8)$$

where the integration constant is chosen so that $zG(z) \sim -1$ when $|z| \to \infty$. Note that the asymptotic behavior for H is therefore also $zH \sim -1$ when $|z| \to \infty$ as is expected for the Stieltjes transform of a probability measure. Rewriting now Equation (4.8) in term of the function H, we obtain

$$qH^2 + H\left[1 + \frac{1}{z}\left(q - 1 + \frac{1}{M}\frac{2-\beta}{\beta}\right)\right] + \frac{1}{M}\frac{2-\beta}{\beta}\frac{dH}{dz} + \frac{1}{z} + q\,F(z) = 0\,. \quad (4.9)$$

Now, using the result about the noise in the Marčenko-Pastur law obtained in [13] by Lytova and Pastur, we know that in the limit of large M, we have for $\beta = 1$ or 2,

$$F(z) \sim_{M\to\infty} \quad (4.10)$$

$$\frac{1}{M^2}\frac{1}{q^2}\frac{1}{2\beta\pi^2}\int_{\gamma_-}^{\gamma_+}\int_{\gamma_-}^{\gamma_+}\frac{d\lambda d\mu}{(\lambda-z)^2(\mu-z)^2}\frac{4q - (\lambda-(1+q))(\mu-(1+q))}{\sqrt{4q-(\lambda-(1+q))^2}\sqrt{4q-(\mu-(1+q))^2}}$$

with $\gamma_\pm = (1 \pm \sqrt{q})^2$.

The idea to obtain the correction to the Marčenko-Pastur law is to use perturbation theory in Eq. (4.9). More precisely, we want to compute explicitly the coefficients ρ_0, ρ_1 and ρ_2 such that the eigenvalue density of a β-Wishart matrix writes under the form, in the limit of large N, M with $N/M = q$,

$$\rho(\lambda) = \rho_0(\lambda) + \frac{1}{M}\rho_1(\lambda) + \frac{1}{M^2}\rho_2(\lambda) + o\left(\frac{1}{M^2}\right)\,. \quad (4.11)$$

Note that this asymptotic expansion (4.11) is obtained by perturbation theory and therefore is valid only for the values of λ such that the correction terms $\rho_1(\lambda)/M$ and $\rho_2(\lambda)/M^2$ are negligible compared to the leading term $\rho_0(\lambda)$ in the limit of large M, i.e. for the values of λ such that $\rho_0(\lambda) \neq 0$. The expansion (4.11) is not valid outside the

Marčenko-Pastur sea, i.e., it breaks down near the edges (see below). In addition, here we are talking about smoothed density, hence it contains no oscillatory term in the finite N bulk corrections [14].

To this purpose, we first write $H(z)$ under the form

$$H(z) = H_0(z) + \frac{1}{M}H_1(z) + \frac{1}{M^2}H_2(z) + o\left(\frac{1}{M^2}\right) \tag{4.12}$$

and we plug Eq. (4.12) into Eq. (4.9). By solving the equation to leading order, we find the following expression for $H_0(z)$

$$\begin{aligned}
H_0(z) &= \frac{1}{2q}\frac{-(z+q-1)+\sqrt{(z-\gamma_-)(z-\gamma_+)}}{z} \\
&= \frac{1}{2q}\frac{-(z+q-1)+\sqrt{(z-(1+q))^2-4q}}{z} .
\end{aligned} \tag{4.13}$$

We deduce from this the famous Marčenko-Pastur result: the eigenvalue density converges in the limit of large N, M with $N/M = q$ to the Marčenko-Pastur density as expected given by

$$\rho_0(\lambda) = \frac{1}{2\pi q}\frac{\sqrt{(\lambda-\gamma_-)(\gamma_+-\lambda)}}{\lambda} .$$

For all value of $\beta > 0$, we can now compute the $1/M$ correction to the Marčenko-Pastur density by plugging Eq. (4.12) into Eq. (4.9) and solve to order $1/M$. This gives the following expression for $H_1(z)$

$$H_1(z) = -\left(\frac{1}{\beta}-\frac{1}{2}\right)\frac{1}{q}\left[\frac{1}{2}\left(\frac{1}{z-\gamma_+}+\frac{1}{z-\gamma_-}\right) - \frac{1}{\sqrt{(z-\gamma_+)(z-\gamma_-)}}\right]$$

and the corresponding $1/N$ correction to the density is then given (for all β) by

$$\rho_1(\lambda) = \left(\frac{1}{\beta}-\frac{1}{2}\right)\frac{1}{q}\left[\frac{1}{2}\left(\delta(\lambda-\gamma_+)+\delta(\lambda-\gamma_-)\right) - \frac{1}{\pi}\frac{d\lambda}{\sqrt{(\lambda-\gamma_-)(\gamma_+-\lambda)}}\right]. \tag{4.14}$$

Comparing $H_1(z)/M$ with $H_0(z)$, we see that the correction term ceases to be negligible when $|\lambda-\gamma_\pm| \sim M^{-2/3}$, as expected: this is indeed the standard edge scaling that defines the Tracy-Widom region. Note that this $\mathcal{O}(1/N)$ correction term was derived earlier [[14] and references therein] for $\beta = 1, 2$ and 4, but our result is valid for general β.

For the particular value $\beta = 1$ or 2, we can use the result of Lytova and Pastur stated above in Eq. (4.10) to compute the $1/M^2$ correction with the same method by solving the equation until order $1/M^2$.

Let us first compute an explicit expression for $F(z)$ from the integral representation

in Eq. (4.10)

$$
\begin{aligned}
q^2 M^2 F(z) &= \frac{1}{2\beta\pi^2}\left[4q\left(\int_{\gamma_-}^{\gamma_+}\frac{1}{(\lambda-z)^2}\frac{d\lambda}{\sqrt{4q-(\lambda-(1+q))^2}}\right)^2\right.\\
&\qquad\left. -\left(\int_{\gamma_-}^{\gamma_+}\frac{d\lambda}{(\lambda-z)^2}\frac{\lambda-(1+q)}{\sqrt{4q-(\lambda-(1+q))^2}}\right)^2\right]+o(1)\\
&= \frac{1}{2\beta\pi^2}\left[4q\left(-\frac{\pi}{(z-\gamma_-)(z-\gamma_+)}\frac{2z-\gamma_--\gamma_+}{\sqrt{(z-\gamma_-)(z-\gamma_+)}}\right)^2\right.\\
&\qquad\left. -\frac{\pi^2}{(z-\gamma_-)(z-\gamma_+)}\left(1-\frac{1}{2}\frac{(2z-\gamma_--\gamma_+)^2}{(z-\gamma_-)(z-\gamma_+)}\right)^2\right]+o(1)\\
&= \frac{1}{2\beta}\frac{1}{(z-\gamma_-)(z-\gamma_+)}\left[4q\frac{(2z-\gamma_--\gamma_+)^2}{(z-\gamma_-)^2(z-\gamma_+)^2}-\left(1-\frac{1}{2}\frac{(2z-\gamma_--\gamma_+)^2}{(z-\gamma_-)(z-\gamma_+)}\right)^2\right]+o(1)\\
&= \frac{1}{2\beta}\frac{1}{(z-\gamma_-)(\gamma_+-z)}+o(1)\,.
\end{aligned}
$$

Then we can turn to compute $H_2(z)$ and deduce from this computation the expression for $\rho_2(\lambda)$

$$
\begin{aligned}
\rho_2(\lambda) = -2\left(\frac{1}{2}-\frac{1}{\beta}\right)^2\frac{1}{q}\frac{1}{\sqrt{(\lambda-\gamma_-)(\gamma_+-\lambda)}}&\left[\frac{1}{2}\left(\frac{1}{\lambda-\gamma_-}+\frac{1}{\lambda-\gamma_+}\right)\right.\\
\left. -\frac{\lambda}{2}\left(\frac{1}{(\lambda-\gamma_-)^2}+\frac{1}{(\lambda-\gamma_+)^2}\right)\right]&+\frac{1}{2q\beta}\frac{\lambda}{(\lambda-\gamma_-)^{3/2}(\gamma_+-\lambda)^{3/2}}\,.
\end{aligned}
$$

Again, the comparison of this correction term with the dominant term indicates that our perturbation expansion breaks down when $|\lambda-\gamma_\pm|\sim M^{-2/3}$.

5   Conclusions

In summary, we proposed a random matrix model (invariant under similarity transformations) whose joint density of eigenvalues is given by the classical β-Wishart ensemble where the quantization of the Dyson index β is lifted. The procedure is constructive and is described in section 3.2. The resulting ensemble is by construction invariant under similarity transformations with Haar distributed eigenvectors The diffusive evolution equation for the eigenvalues involves the Dyson index of the ensemble as a free parameter. Letting it scale with the size M of the matrix, the spectral density of the ensemble becomes a one-parameter continuous family interpolating between the familiar Marčenko-Pastur distribution and a certain type of Gamma distribution. On the other hand, keeping the Dyson index unscaled but not quantized, we showed that a careful analysis of the full Stieltjes transform equation lead naturally to $1/N$ and $1/N^2$ corrections (and possibly systematically to any order) to the average spectral density (Marčenko-Pastur) for all $\beta\neq 2$. This then extends the previous work [14, 15] on the $\mathcal{O}(1/N)$ correction term for $\beta = 1$, 2 and

4. To order $O(1/N^2)$, our result is valid for $\beta = 1$ and $\beta = 2$. It would be interesting to see (or conjecture) if this formula to $O(1/N^2)$ term is valid for general β.

In this work, we have computed the crossover density as a function of the interpolating parameter c. In the limit $c \to \infty$, it reduces to the standard Marčenko-Pastur density, whereas the opposite limit $c \to 0$ corresponds to the Gamma laws associated with the CIR process. It would be interesting to extend our analysis to the distribution of the largest eigenvalue. As in the case of bulk density, we would expect a c-dependent distribution for the largest eigenvalue, properly centered and scaled, interpolating between the Tracy-Widom distribution ($c \to \infty$ limit) and Gumbel distribution (as $c \to 0$).

A   Derivation of (3.6)

The Fokker-Planck equation for the transition probability density $P(\lambda_1, \cdots, \lambda_N; t)$ of the process $(\lambda_1(t), \cdots, \lambda_N(t))$ which satisfies the stochastic differential system (3.5) reads

$$\frac{\partial P}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial \lambda_i} \left[P \left(-\lambda_i + pM + (1-\delta)p + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) \right] + 2 \sum_{i=1}^{N} \frac{\partial^2}{\partial \lambda_i^2} [\lambda_i P] . \quad \text{(A.1)}$$

The stationary solution is the solution which does not depend on time t, satisfying

$$-\sum_{i=1}^{N} \frac{\partial}{\partial \lambda_i} \left[P \left(-\lambda_i + pM + (1-\delta)p + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) \right] + 2 \sum_{i=1}^{N} \frac{\partial^2}{\partial \lambda_i^2} [\lambda_i P] = 0 . \quad \text{(A.2)}$$

It is easy to check using elementary algebra that the jpdf P^* defined in (3.6) verifies Eq. (A.1) as in fact we can verify that for all i,

$$2 \frac{\partial}{\partial \lambda_i} [\lambda_i P^*] = P^* \left(-\lambda_i + pM + (1-\delta)p + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) . \quad \text{(A.3)}$$

B   Algorithmic description of how to build the process $\mathbf{W}_t^n$ in practice

Let us describe shortly an algorithmic description of how to build the process $\mathbf{W}_t^n$ in practice, on a discrete grid. First note that this algorithmic description needs a discrete grid and that it does not reproduce exactly the process $\mathbf{W}_t^n$ but only a discretized approximation of it. Choose a large value of n and an initial symmetric matrix $\mathbf{W}_0$. The construction is iterative. Suppose that the process is constructed until time k/n and let us explain how to compute the matrix $\mathbf{W}_{(k+1)/n}^n$ at the next discrete time of the grid, $(k+1)/n$.

1. Step 1. We first need to compute the matrix $\sqrt{\mathbf{W}_{k/n}^n}$. It suffices to compute the orthogonal matrix $\mathbf{O}_{k/n}^n$ such that

$$\mathbf{W}_{k/n}^n = \mathbf{O}_{k/n}^n \mathbf{\Sigma}_{k/n}^n \mathbf{O}_{k/n}^{n\ \dagger}$$

where $\mathbf{\Sigma}_{k/n}^n$ is the diagonal matrix composed of the eigenvalues of $\mathbf{W}_{k/n}^n$ (in increasing order). The eigenvalues of the matrix $\mathbf{W}_{k/n}^n$ should be non negative as the eigenvalues process of $\mathbf{W}_t^n$ are almost surely non negative at all time t. However, due to the discretization scheme necessary for algorithmic procedure, the non-negativity can fail. To avoid this problem, we define $\sqrt{\mathbf{W}_{k/n}^n}$ as

$$\sqrt{\mathbf{W}_{k/n}^n} = \mathbf{O}_{k/n}^n \sqrt{\mathbf{\Sigma}_{k/n}^n} \mathbf{O}_{k/n}^{n\,\dagger} \tag{B.1}$$

where $\sqrt{\mathbf{\Sigma}_{k/n}^n}$ is the diagonal matrix composed of the square roots of the absolute values of the eigenvalues of $\mathbf{W}_{k/n}^n$ (again in increasing order).

2. Step 2. We sample the Bernoulli random variable ϵ_k^n with $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$.

3. Step 3. It depends on the value of ϵ_k^n:

   - if $\epsilon_k^n = 1$, we sample a $N \times N$ matrix $\mathbf{G}_n$ filled with independent Gaussian variables with mean 0 and variance $1/n$ and then we compute the matrix $\mathbf{W}_{(k+1)/n}^n$ by the formula

   $$\mathbf{W}_{(k+1)/n}^n = \left(1 - \frac{1}{n}\right)\mathbf{W}_{k/n}^n + \sqrt{\mathbf{W}_{k/n}^n}\,\mathbf{G}_n + \mathbf{G}_n^{\dagger}\sqrt{\mathbf{W}_{k/n}^n} + \frac{1}{n}M\,\mathbf{I}.$$

   - if $\epsilon_k^n = 0$, we sample N independent Gaussian variables $(z_1, \cdots, z_N)$ with mean 0 and variance $1/n$. We then compute the matrix $\mathbf{Y}_n$, which is co diagonalizable with the matrix $\mathbf{W}_{k/n}^n$, defined as the product

   $$\mathbf{Y}_n := \mathbf{O}_{k/n}^n \mathrm{Diag}\left(z_1, z_2, \ldots, z_N\right)\mathbf{O}_{k/n}^{n\,\dagger}. \tag{B.2}$$

   Finally we obtain the matrix $\mathbf{W}_{(k+1)/n}^n$ by

   $$\mathbf{W}_{(k+1)/n}^n = \left(1 - \frac{1}{n}\right)\mathbf{W}_{k/n}^n + \sqrt{\mathbf{W}_{k/n}^n}\,\mathbf{Y}_n + \mathbf{Y}_n^{\dagger}\sqrt{\mathbf{W}_{k/n}^n} + \frac{1}{n}\delta\,\mathbf{I}.$$

References

[1] G. Akemann, J. Baik and P. Di Francesco, *The Oxford Handbook of Random Matrix Theory*, (Oxford University Press 2011).

[2] G. W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices* (Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2009).

[3] Z. Bai and J. Silverstein, *Spectral Analysis of Large dimensional random matrices*, (Second Edition, Springer, 2010).

[4] P. J. Forrester, *Log gases and Random Matrices* (Princeton University Press, 2010).

[5] A. Tulino and S. Verdù, *Random Matrix Theory and Wireless Communications*, Foundations and Trends in Communication and Information Theory **1**, 1 (2004).

[6] J. Wishart, Biometrika **20**, 32 (1928).

[7] I. Dumitriu and A. Edelman, J. Math. Phys. **43**, 5830 (2002).

[8] P. Vivo and S. N. Majumdar, Physica A **387**, 4839 (2008).

[9] R. Allez, J.-P. Bouchaud and A. Guionnet, Phys. Rev. Lett. **109**, 094102 (2012).

[10] R. Allez and A. Guionnet, [arXiv:1206.1460] (2012).

[11] G. Akemann and P. Vivo, J. Stat. Mech. P09002 (2008).

[12] Z. Burda, A.T. Gorlich and B. Waclaw, Phys. Rev. E **74**, 041129, (2006).

[13] A. Lytova and L. Pastur, Ann. Probab. **37**, Number 5, 1778, (2009).

[14] P. J. Forrester, N. E. Frankel and I. M. Garoni, J. Math. Phys. **47**, 023301, (2006).

[15] P. J. Forrester, J. Phys. A.: Math. Theo. **45**, 075206 (2012).

[16] M. Sadek, A. Tarighat and A.H. Sayed, IEEE Trans. Signal Processing **55**, 1498 (2007).

[17] Y. V. Fyodorov and H.-J. Sommers, J. Math. Phys. **38**, 1918 (1997); Y. V. Fyodorov and B. A. Khoruzhenko, Phys. Rev. Lett. **83**, 66 (1999).

[18] J. J. M. Verbaarschot, Phys. Rev. Lett. **72**, 2531 (1994).

[19] K. Johansson, Comm. Math. Phys. **209**, 437 (2000).

[20] G. Schehr, S. N. Majumdar, A. Comtet and J. Randon-Furling, Phys. Rev. Lett. **101**, 150601 (2008).

[21] C. Nadal and S. N. Majumdar, Phys. Rev. E **79**, 061117 (2009).

[22] J. Rambeau and G. Schehr, Europhys. Lett. **91**, 60006 (2010); Phys. Rev. E, **83**, 061146 (2011)

[23] S. N. Majumdar and M. Vergassola, Phys. Rev. Lett. **102**, 060601 (2009).

[24] S. N. Majumdar and P. Vivo, Phys. Rev. Lett. **108**, 200601 (2012).

[25] A. T. James, Ann. Math. Statistics **35**, 475 (1964).

[26] V. A. Marčenko and L. A. Pastur, Math. USSR-Sb. **1**, 457 (1967).

[27] M. F. Bru, J. Theor. Prob. **4**, 725 (1991).

[28] M. F. Bru, J. Multivar. Anal. **29**, 127 (1989).

[29] J. C. Cox, J. E. Ingersoll and S. A. Ross, Econometrica **53**, 385 (1985).

[30] F. J. Dyson, J. Math. Phys. **3**, 140; ibid 157; ibid 166 (1962).

[31] D. S. Dean and S. N. Majumdar, Phys. Rev. E **77**, 041108 (2008).

[32] I. S. Gradshteyn and I. M. Ryzhik, (Academic Press, San Diego, 1980).

[33] M. Abramowitz and I. A. Stegun, (Dover, New York, 1965).

[34] M. E. H. Ismail and D. H. Kelker, Siam J. Math. Anal. **10**, 884 (1979).

[35] L. C. G. Rogers and Z. Shi, Probab. Th. Rel. Fields **95**, 555 (1993).